Redefining Student Success:
Learning from Nontraditional Learners

by

Jean Mallory Campbell
B.A., University of Victoria, 1972
Diploma, Education of Children with
Learning and Behaviour Disorders,
University of British Columbia, 1974
M.Ed., University of British Columbia, 1979

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the School of Child and Youth Care

0-612-36631-6

Supervisor: Jim Anglin

ABSTRACT

This study examines student success from the perspective of mature, female students in human service programs at one B.C. post-secondary institution. An exploratory, case study approach was employed, focussing on in depth open ended interviews of 36 women. Context for the educational experience of these women was explored through document review and interviews of faculty. Definitions of success and experiences in achieving success appear quite different from the traditional student success literature. These women are committed to holistic definitions of success which include not only good grades and program completion, but also personal growth and maintenance of satisfactory family, extended family, and community relationships. They are unlikely to drop out because of their intense internal drive and because of a program model which provides a credential after first and second year (as well as at the degree level), allows stopping out temporarily, and supports part-time participation. They are unlikely to access support services because of the pressures on their time. Factors which impede their progress may also support them (for example, families place demands but also provide support; negative educational experiences in the past both limit their self confidence and make them determined to do well and "prove themselves"). Poverty appears to be the greatest barrier to many, particularly single parents. Implications for theory, educational reform policy, and research are outlined.

TABLE OF CONTENTS

TABLE OF CONTENTS	iv
CHAPTER 1: INTRODUCTION AND REVIEW OF THE LITERATURE	1
Background	
The Issue of Student Retention	1
Improving Student Retention	4
Education and Retention in the Human Services: Focusing on Nontraditional Students	5
Definitions	7
Review of the Retention/Student Success Literature	
Theoretical Models	8
General Factors Influencing Student Persistence	10
Factors Influencing Nontraditional Students	12
Student Outcomes and Decisions	16
Approaches to Improving Student Retention	18
Gaps and Inconsistencies in the Literature	23
Student Retention Research and Initiatives in B.C.	27
B.C. Research Findings	29
B.C. Research Limitations	30
B.C. Research Orientation	35
Conclusions	37
Rationale and Scope of the Current Study	39
Research Focus and Questions	41
CHAPTER 2: METHODOLOGY	44
Orientation of the Case Study	45
Principles	47
Procedures	
Summary	48
Description	49
Methodology: Conclusions	55

CHAPTER 3: RESULTS	58

Findings: Interviews
- The Participants 58
- Organization of the Interview Data 61
- Student Goals 63
- Factors Affecting Achievement of Goals 77
- Considering Quitting 98
- Definitions of Success 100
- Summary: Interview Findings 103

Findings: Document Review
- Institution, Department and Program Descriptions 105
- Unpublished Reports 109
- Summary: Document Review Findings 113

Study Findings in the Context of Previous Literature and Research
- Definitions of Success 114
- Factors Predicting Success 115
- Strategies for Improving Student Success 118

CHAPTER 4: DISCUSSION	131

Unique Aspects of the Findings 131

Implications for Student Success Theory 135

Implications for Policy and Practice 142

Implications for Future Research 159

Conclusions 161

REFERENCES	167

APPENDICES	179

- Appendix A 179
- Appendix B 180
- Appendix C 181
- Appendix E 182
- Appendix F 183

CHAPTER 1: INTRODUCTION AND REVIEW OF THE LITERATURE

Background

The Issue of Student Retention

As costs of post-secondary education rise, and fiscal restraints become more severe, increasing attention is being paid to outcome measures both in terms of educational and financial returns. Student retention and completion, as well as the inverse, student attrition, are recognized as critical issues facing post-secondary education across North America. In British Columbia (B.C.), attention to outcome issues will become more acute in the next several years as the provincial government considers outcome-based funding mechanisms, and cuts to federal transfer payments lead to increasing scrutiny of post-secondary expenditures and returns.

Statistics in this area are not encouraging. Both U.S. and Canadian research indicate that 40% to 50% of post-secondary students leave the system without completing a credential (Gomme & Gilbert, 1984; Guppy & Bednarski, 1993; Noel, 1985). It appears that more students are taking longer to complete credentials; however, U.S. data indicate that even 6 years after entering degree programs, only about half of the students graduate (Dey & Hurtado, 1995).

B.C.'s attrition rates appear typical of this pattern. For example, one third of University of Victoria students drop out before completing their first year, and only 56% of students entering the University earn a credential after 7 years (University of Victoria, 1995). Some 2-year college programs graduate less than 50% of their full-time students (B.C. Council on Admissions and Transfer, 1992).

Completion rates of only 50% to 60% may be more significant for B.C. than the U.S. because overall participation rates are much lower. B.C.'s post-secondary participation rates also remain below the Canadian average (B.C. Ministry of Education, Skills and Training, 1997a). Less post-secondary spaces combined with average attrition rates means fewer credentials awarded per capita than in most other provinces (B.C. Council on Admissions and Transfer, 1992; Guppy & Bednarski, 1993). During the years 1981 to 1986, for example, B.C. produced 14 degree graduates per 1,000 population, compared to a Canadian average of 21 per 1,000 (B.C. Ministry of Advanced Education, Training and Technology, 1991). Additional post-secondary spaces have recently been created in B.C., and participation rates at both colleges and universities in the province have increased. However, the extent to which attrition mitigates the positive impact of this increase is unknown.

High attrition rates are a concern for several reasons. Costs to society include limited returns for financial investments in education and a lack of qualified graduates to enter the workforce and contribute to the economy. Costs to post-secondary institutions include the loss of tuition revenue (annual losses associated with freshman attrition at one Canadian college were estimated at $7.2 million; Dietsche, 1990) and potential loss of public confidence stemming from negative student experiences. Costs to the student are probably the most dramatic. These include the emotional strain to student and family created by a perceived failure experience, coupled with direct financial costs of the education and, more significantly, the lifetime income reduction with which non-completion is associated (Bean, 1990a; Gomme & Gilbert, 1984; Guppy & Bednarski, 1993). For example, Tinto (1993) cited evidence that, on average, male college graduates in the U.S. earn at least 23% more than men of the same age who lack a college credential.

High attrition is even more of a concern for groups that are conspicuously under-represented in the post-secondary system. For example, First Nations British Columbians who complete high school are 20% less likely than their non-First Nations peers to participate in post-secondary education; British Columbians with disabilities represent approximately 14% of the working age population, yet only 3% of the post-secondary population (B.C. Ministry of Education, Skills and Training, 1997b). Additionally, there appear to be significant variations in post-secondary participation rates across B.C.; low rates of post-secondary completion are particularly problematic in those areas of the province where fewer students even enter the post-secondary system.

Barriers to student participation and success that affect some student groups more than others (e.g., financial barriers that have particular impact on students living in poverty) have the further outcome of limiting the potential equalizing effect of a post-secondary education. The economic disadvantage typically experienced by members of all equity groups and students from low socioeconomic status backgrounds can be ameliorated, to some extent, by completion of a post-secondary credential. Attempting a post-secondary credential, but "failing," can have the exacerbating effect of increasing the debt and financial disadvantage of the student, decreasing his or her already low self-esteem, and solidifying the social stratification that adult education should assist in addressing.

Despite the many reasons that dropout and retention are important, and the prominence these issues have gained in post-secondary literature, there remains concern that some Canadian educators and administrators still do not take the matter seriously, rationalizing that attrition is to be expected.

> Indeed the senior managers of most, if not all, colleges have no idea what their dropout rate is because this group tends to subscribe to the "passive" philosophy of education; all that is required of them is to provide the necessary inputs of facilities, faculty, and programs and some students will succeed while others will not. . . . Furthermore, most administrators assume the principal factors promoting student failure/dropout are associated with the student, not the institution. (Dietsche, 1990, p. 81)

> Apparently there is an attitude pervasive in Canada that a university degree is simply not for everyone and that it is only natural to expect some students to discover their lack of interest or lack of suitability for university after a year or two of undergraduate studies. (Johnson & Buck, 1995, p. 55)

Improving Student Retention

U.S. post-secondary institutions have discovered that, in the interest of maintaining high enrollment (to maximize both tuition and alumni support), "it is easier to maintain existing students (customers) than to solicit new students" (Johnson & Buck, 1995, p. 55). In the face of declining enrollments and financial resources, American "institutions have come to view the retention of students as the only reasonable course of action left to insure their survival, and . . . a growing number have turned their energies in that direction with a renewed passion" (Tinto, 1993, p. 2).

Consistent with this trend, although with motivation less narrowly linked to institutional survival, post-secondary educators and researchers across Canada have begun to focus considerable attention on interventions designed to improve student retention. The literature contains many descriptions of student retention strategies, which will be explored in a later section of this paper. In B.C., post-secondary institutions are implementing the types of strategies generally accepted as improving student retention. Beatty-Guenter (1994) found that all but two of B.C.'s colleges increased their application of retention strategies between 1981 and 1991. Recent provincial surveys (B.C. Ministry of Education, Skills and Training, 1996c; B.C. Ministry of Skills, Training and Labour, 1995) revealed that, although the sophistication and comprehensiveness of the strategies vary between institutions, all are making increasing effort in this area.

The expanding implementation of retention strategies has not, however, been accompanied by accumulation of solid research evidence regarding effectiveness. Much of the literature on retention consists of testimonials to the value of particular approaches, rather than research: Beatty-Guenter (1994) referred to "the predominantly 'show and tell' pattern of the retention literature" (p. 19). The research that does exist has been soundly criticized, as will be discussed in a later section of this paper. The dearth of relevant research is paralleled in British Columbia. According to Springate (1991b), most intervention strategies in use in B.C. colleges have not been evaluated. Beatty-Guenter (1994) similarly concluded that most B.C. colleges do not know their retention rates and are unlikely to have studied the effectiveness of their intervention strategies.

Education and Student Retention in the Human Services: Focusing on Nontraditional Students

The human service sector comprises a variety of occupations united by their focus on provision of service to vulnerable individuals: for example, children, the elderly, people with disabilities, and victims of abuse. Availability of a qualified workforce is critical to most employment sectors, but particularly vital in human service occupations. Because the content and medium of the work is human interaction, staff skills and knowledge are the primary measure of program quality and directly influence the lives of individuals receiving care. Although it is increasingly recognized that training is key to ensuring staff competence, access to education is limited for human service providers. The human service workforce in

B.C. (including workers providing child day care, support to persons with disabilities and children and youth, and workers in community justice, immigrant and multicultural agencies, and women-serving agencies) numbers approximately 30,000 individuals with staff turnover in the range of 20% to 30% per year (Campbell, 1993), yet there are only approximately 2,000 spaces in the community college programs that provide most of the training for this sector (B.C. Ministry of Education, Skills and Training, unpublished data).

It is critical that the effectiveness of these limited training opportunities be maximized by facilitating positive outcomes, including high rates of program completion. It appears, however, that there has been minimal formal research attention paid to the outcomes of human service training programs.

One of the challenges facing human service educators is the nontraditional nature of the student population. In B.C., at least 75% of those working in the human service field are female (Erickson, 1993) and the proportion of females in human service training programs appears to be even higher. Many students work while pursuing an education, and many also balance family responsibilities.

Nontraditional students have been defined in the literature as those who are over 25, do not reside on campus, or attend part-time (Bean & Metzner, 1985). This group comprises a growing proportion of post-secondary populations (Baker & Velez, 1996; Belanger, Lynd, & Mouelhi, 1982; Devlin, 1989; Dey & Hurtado, 1995; Tinto, 1993). For example, growth in part-time post-secondary enrollment in Canada almost doubled that of full-time enrollment between 1965 and 1985 (Anisef, 1989). In B.C., overall part-time post-secondary enrollments have increased significantly more than full-time enrollments in recent years, and college part-time participation is almost double the national average (B.C. Ministry of Education, Skills and Training, 1997a).

The importance of nontraditional student enrollments has been noted across North America and internationally. Blaxter, Dodd, and Tight (1996), for example, commented that "for the last decade, almost all British universities have been paying increasing attention to the recruitment and retention of mature students" (p. 187). Despite this international trend, however, nontraditional students have been the focus of relatively little research. There are some indications that attrition rates for mature students may be higher than for traditional age students (Bean & Metzner, 1985), possibly because of the competing demands students face. "Life roles are revolving doors that move with blurring speed" for nontraditional students, according to Chartrand (1992, p. 8).

This chapter will examine the literature regarding student retention and completion, with a focus on the unique needs of nontraditional students. B.C. reports will be utilized to compare the B.C. situation to the broader literature.

Definitions

As will be discussed, one of the difficulties with the study of student retention is inconsistency in terminology and operational definitions. For the purposes of this report, *retention* and *persistence* will be used interchangeably to refer to students' continuation in their studies, either within a semester or between semesters or years. One subset of retention is *completion*: persistence to the completion of the intended credential. *Attrition* is the reverse of retention and refers to students' non-continuation in their studies, for whatever reasons. *Success* has often been used as a synonym for completion, but will be used here to refer to positive outcomes in the program of study as defined by the individual student, which may or may not include retention and completion.

Nontraditional students are defined as all students other than students entering post-secondary education directly from high school and attending on a full-time basis. Included within this group are older returning students, part-time students of any age, and many cultural minority group students and students with disabilities. This definition, and the use of the term nontraditional to describe this group, is not meant to imply that the majority of post-secondary students currently fits the traditional mold, since they may not. For example, at present even young students are taking longer to complete their post-secondary education because of work and family obligations. The 19-year-old student living at home or in residence, with no other responsibilities than those deriving from the student role, likely no longer represents the norm on our post-secondary campuses; nevertheless, this remains the stereotype of the typical post-secondary student around whom programming is generally planned and organized.

Review of the Retention/Student Success Literature

Theoretical Models

Numerous theories have evolved over the years to explain why some post-secondary students persist and others do not. The most prominent theorist has been Vincent Tinto (1975, 1982, 1985, 1988, 1993) whose original model acknowledged the importance of students' background traits, distinguished academic and social factors, and recognized the longitudinal nature of attrition/persistence. Tinto's research was primarily based on traditional students attending residential universities. Bean (1980) proposed a compatible model that introduced one additional factor, the effect of environmental variables external to the institution.

Bean and Metzner (1985) took this approach one step further in developing a model of nontraditional student attrition. The premise of this model was that nontraditional students have significantly less interaction with the college community, deriving social support instead from their external environment. Environmental variables such as family support, finances, and work demands are hypothesized to be most important for this student group. More recently, Tinto (1988, 1993) proposed and developed a model similar to his earlier version, but incorporating environmental variables and intentions, and drawing on social anthropology to identify stages ("rites of passage") through which students typically pass. A model integrating the key factors of the leading theories was proposed as most likely to explain persistence (Cabrera, Castenada, Nora, & Hengstler, 1992).

Bean (1990a) attempted to synthesize the various theories into a single model for examining persistence/attrition, as represented in diagram form in Appendix A. Elements of the proposed model include: (a) background characteristics of students; (b) organizational, academic, and social interactions between the college and student, and concurrent environmental pull; (c) outcomes including grades, attitudes, and institutional fit and commitment; (d) intent to leave or continue; and (e) eventual decisions. These categories will be used to organize a summary of existing research regarding persistence/attrition, first for all students and then with particular focus on nontraditional students.

General Factors Influencing Student Persistence

Background Characteristics

Factors that generally appear to influence whether students complete their programs or credentials encompass previous educational attainment, including secondary school grades and general academic preparedness; parental encouragement, income and educational level (directly related to both enrollment and completion); gender, with females more likely to persist; and certainty about career/educational goals (Bean, 1990a; Dietsche, 1990; Noel, 1985; Pascaralla, Duby, & Iverson, 1983; Stage & Rushin, 1993). How realistic and informed students are about the demands of post-secondary life also appears to be a factor (Holdaway & Kelloway, 1987; Noel, 1985).

Organizational Variables

Organizational variables that appear to increase students' likelihood of program completion include flexible/convenient class scheduling, availability of required courses, and availability of financial aid and academic and social support services (Aitken, 1982; Bean, 1990a; Levitz & Noel, 1985; Pappas & Loring, 1985).

Academic Integration

A key factor facilitating students' successful academic integration into the institution appears to be contact with faculty members, both in class and informally (Bean, 1990a; Gomme & Gilbert, 1984; Stewart, 1990). Additional factors include study habits and ability to handle required reading, tests, and assignments (Anderson, 1985; Bean, 1990a; Dietsche, 1990; Noel, 1985). Students enrolled in general Arts and Sciences have been reported to have the highest percentage of withdrawals, suggesting the importance of having a focus for studies (Johnson & Buck, 1995). Students enrolled in honors programs are more likely to complete (Astin, 1984), as are students in clearly defined majors such as journalism or health professions (Bean, 1990a; Stewart, 1990), which are more conducive to development of an academic identity.

Social Integration

The extent of on-campus social integration, with friends and faculty or through extracurricular involvement, was viewed as critical in earlier models but has been recognized as much more important for full-time, younger, and resident students (Bean, 1990a; Bean & Metzner, 1985; Gomme & Gilbert, 1984; Stewart, 1990). For all students, a social support network, whether on campus or through the external environment, appears important in facilitating persistence (Anderson, 1985; Astin, 1984; Bean, 1990a; Stage & Rushin, 1993).

Environmental Pull

Primary environmental factors that negatively influence persistence appear to be family responsibilities and employment, with amount of work related to likelihood to withdraw. Residency appears relevant in that on-campus residents are more likely to persist, and housing difficulties for off-campus students can be a factor leading to attrition. Other factors less frequently identified include such issues as transportation or a "significant other" living elsewhere (Anderson, 1985; Astin, 1984; Bean, 1990a; Stewart, 1990). Opportunities to transfer to other institutions also influence withdrawal decisions. For many students, the decision to leave one institution is part of a predetermined transfer plan and not a failure experience (Bean, 1990a; Bean & Metzner, 1985; Byrne, 1989; Tinto, 1982; Ungar, 1982).

Financial issues are frequently identified by students as reasons for withdrawal, but appear to be indirectly involved or, as Tinto (1982) stated, "at the margins of decision making" (p. 689). Financial concerns are likely to influence enrollment decisions, but are less likely to lead directly to withdrawal than is commonly thought (Garland, 1993; Noel, 1985; Ungar, 1980). However, the significance of the indirect effects of finances on persistence was noted by Baker and Velez (1996), who cited research underlining the impact of finances on student's academic integration, socialization processes, and resolve to persist in a post-secondary program. These same authors also noted that women are more likely than men to withdraw as a result of family responsibilities.

Factors Influencing Nontraditional Students

According to the literature, nontraditional students are qualitatively different from traditional students (Baker & Velez, 1996; Blaxter et al., 1996; Devlin & Gallagher, 1982; Hodgson, 1989; Kaplan, 1992; Pappas & Loring, 1985); therefore "the concept of persistence or retention must be thought of differently" (Pappas & Loring, 1985, p. 139). One general factor should be noted: That is, that nontraditional students are more likely to attend part-time, and part-time attendance may predispose students to higher withdrawal rates (Okin, Benin, & Brandt-Williams, 1996).

Background Characteristics

Academic preparedness and previous educational attainment are as important for mature students as for younger students. The former group not only differs in its greater clarity of occupational goals and educational expectations, but also in greater likelihood of anxiety and lack of confidence about entering or reentering (often after many years) the academic environment (Anderson, 1985; Byrne, 1989; Hodgson, 1989; Justice, 1997; Kaplan, 1992; Miller & Daloz, 1992; Pappas & Loring, 1985). The health and resilience of older students is a factor in their persistence (Justice, 1997; Scott, Burns, & Cooney, 1996; Sugrue, 1996):

"Older students bring to the educational experience an aging body and mind" (Justice, 1997, p. 31).

Organizational Variables

Several organizational variables appear critical to institutions' successful accommodation of nontraditional students. Although more likely to be self-directed, not expecting assistance, and therefore not in need of many of the supports required by younger students (Bean, 1990a; Byrne, 1989), the unique needs of this group require significant organizational accommodation. Key features appear to be flexibility in class scheduling, timing of assignments, and other program features; avoidance of excessive paperwork and bureaucracy; advance notice regarding timetabling; opportunities for timely program completion; assistance with academic upgrading and study skills; and support for dealing with external issues such as child care (Allen, 1995; Baker & Velez, 1996; Byrne, 1989; Hodgson, 1989; Shively, 1989; Smallwood, 1980; Pappas & Loring, 1985; Potter & Alderman, 1992; Thompson & Devlin, 1992; Van Stone, Nelson, & Niemann, 1994). As Baker and Velez summarized,

> Nontraditional students may need less in terms of fit and more in terms of institutional flexibility. . . . Nontraditional students may not need to make the integrative links that Tinto (1993) suggested, instead, schools may need to loosen up. (1996, p. 95)

Hodgson (1989) also identified the need to include assessment and support for previously overlooked learning disabilities, a critical issue for adults who left the school system before learning disabilities were routinely recognized and addressed.

Academic Integration

It appears that academic involvement is even more important to nontraditional than traditional students. Along with their clear career goals, mature and working students have high expectations of educational programs in terms of relevancy and quality, and may withdraw if programs do not meet these expectations (Allen, 1995; Pappas & Loring, 1985). In one study, mothers who had withdrawn from post-secondary programs cited dissatisfaction with courses and lack of academic freedom as reasons for discontinuation, but ranked these factors as less important than family, work, financial and logistical barriers (Scott et al., 1996). Many nontraditional students do not enter the post-secondary system with the intention of completing a credential, seeking instead specific skills or courses (Bean, 1990a; Tinto, 1985).

Social Integration

As indicated above, nontraditional students tend to draw their social supports from the external environment rather than from on-campus involvement, although the presence of social support appears no less important (Bean, 1990a; Tinto, 1985). Research into minority group retention highlights the importance of campus climate and attitudinal acceptance of nontraditional groups. Attrition among minority group students is associated with feelings of alienation and social estrangement. Presence of a "critical mass" of minority students, sufficient to create a social niche, appears important (Crosson, 1992; McJamerson, 1992; Suen, 1983; Tinto, 1985); the same is likely true for other nontraditional students.

Environmental Pull

Nontraditional students are probably most unique in the extent to which external environmental factors affect their persistence. The need to juggle often incompatible roles and responsibilities creates a significant challenge to such students in their pursuit of a post-secondary credential (Allen, 1995; Anisef, 1989; Bean & Metzner, 1985; Byrne, 1989; Okin et al., 1996; Pappas & Loring, 1985; Potter & Alderman, 1992; Smallwood, 1980; Sugrue, 1996). Byrne spoke of the importance of program flexibility to accommodate potential "disrupters such as sick children and overtime work assignments" (p. 10). Scott et al. (1996) commented that women continue to have greater responsibility for domestic and child rearing tasks, plus "a large percentage of mothers are also employed. . . . Women are often responsible for caring for other family members beyond the nuclear family" (p. 235), with the result that "fitting study into one's already full life is the main challenge encountered" (p. 249).

Often, environmental factors compound other issues facing this student group. For example, the effect of internal forces such as low self-confidence or academic underpreparedness can be exacerbated by limitations on study time created by work or family demands (Anderson, 1985; Ryland, Riordan, & Brack, 1994). Nontraditional students commuting more than 50 miles to class have been noted to be particularly at risk for dropping out (Allen, 1995), perhaps because of the additional time pressure this creates. Financial issues are also more likely to be acute for nontraditional students; students who delay their post-secondary entry are more likely to be from low socioeconomic status backgrounds and to have more financial stresses (Baker & Velez, 1996; Van Stone et al., 1994).

Student Outcomes and Decisions

For both traditional and nontraditional student groups, it appears that the factors outlined above interact to produce the key outcome variables that ultimately lead to the decision to continue or withdraw.

Attitudes

For both student groups, satisfaction with the institutional environment, perceived value of the education, and confidence in ability to succeed all influence the decision to stay (Bean, 1990a). As Dietsche (1990) stated, "What is of major importance, is the student's perception of his or her program as an intellectually stimulating and rewarding activity, which will be of value to his or her future occupational success" (p. 79). Garland (1993) found that students withdrawing from a distance program are likely to give environmental reasons such as time or finances initially, but reveal attitudinal issues such as dissatisfaction with course content, stress, or fear of failure when interviewed. As mentioned above, nontraditional students may be more demanding in this regard and more critical of education that does not appear to be relevant. Nontraditional students may also experience more stress, with ability to cope and specific coping strategies varying with a number of factors, including age (Thacker & Novak, 1991).

Grades

Academic performance results from the complex interaction of all the factors above and clearly influences students' decision-making, although it is a myth that academic failure is a major cause of attrition (Noel, 1985), and attrition can be significant even for students with good grades (Johnson & Buck, 1995). The impact of grades on the decision to withdraw appears to be less direct for nontraditional students, for whom environmental factors often have a greater influence (Bean & Metzner, 1985).

Institutional Fit and Loyalty

One of the most constant research findings is that students tend to leave post-secondary programs that do not fit with their needs/expectations or when they do not feel that they "fit in" (Bean, 1990a). The longitudinal interaction of the factors above produces the composite of "fit" for both traditional and nontraditional students. As suggested above, the issue of fit is likely less important for nontraditional students, for whom institutional flexibility is far more critical (Baker & Velez, 1996).

Intention and Decision

Except in the case of some mature or part-time students who enter programs with no intention of completion (Bean & Metzner, 1985), the relationship between students' intentions and eventual decisions is unclear. Bean and Metzner indicated that intent to leave is generally related to attrition (although this has not been studied with nontraditional students). Pascarella, Duby, and Iverson (1983), conversely, found that intention did not influence persistence, and neither commitment to graduating nor to the institution are significant variables. Okin et al. (1996) concluded that students' likelihood of actually acting on their intentions (to persist or withdraw) depends on complex interactions between variables including grades, prior commitment, and encouragement from others to stay; however, results remained equivocal. Clearly, the ultimate decisions made by students reflect the complex and unique situation each faces and cannot always be predicted by stated intentions.

Approaches to Improving Student Retention

Despite the limitations to student retention research, several conclusions regarding success strategies appear to be supported by the literature. These will be outlined briefly below.

Advising and Counselling

Accurate, individualized, and easily accessible academic advising and career counselling should be available to assist all students with the difficult educational and occupational decisions facing them (Allen, 1995; Bank, Slavings, & Biddle, 1990; Cooper & Bradshaw, 1984; Gardner, 1992; Gomme, Hall, & Murphy, 1993; Metzner, 1989; Smith, Lippitt, Noel, & Sprandel, 1980; White, 1983; Young, Backer, & Rogers, 1989).

Student Transitions

Collaboration with the school system and other post-secondary institutions, careful recruitment, and pre-enrollment contact need to be in place to support smooth student transitions into the institution (Byrne, 1989; Crosson, 1992; Gordon & Grites, 1984; Metzner, 1989; Pascarella, Terenzini, & Wolfe, 1986). Mature students, in particular, who are not making a direct transition from previous educational programs and are often dealing with lack of confidence and anxiety about returning to education, require support and orientation to facilitate a smooth entry into the post-secondary setting (Allen, 1995; Justice, 1997). Educational excellence, service, and support should be "front loaded" to put most emphasis and resources into new students (Gardner, 1992). Supports should also be in place for students bridging to other programs, because difficulties obtaining transfer credit can frustrate and discourage students (Monk-Turner, 1990; Tyler & Small, 1990). Lack of confidence may block some students, notably women, from persisting to upper levels of training (Fiorentine, 1987).

Student Success Courses

Many institutions have instituted orientation or student success courses for new students, popularly called "University 101" after the pioneering work done in this area at the University of South Carolina. Orientation courses have been offered on both a credit and non-credit basis in the U.S. for over 80 years; by 1928 more than 100 U.S. institutions offered such courses (Gordon & Grites, 1984). Extensive literature has been generated about student success courses, and detailed content outlines and guidelines for delivery are available (Gardner, 1978, 1981, 1992; Gordon & Grites, 1984; Toder & Hartsough, 1993). Although research regarding the effectiveness of these courses produces mixed results, they receive considerable support and have recognized utility in combination with other strategies.

Program Quality

Measures for ongoing evaluation and improvement of instructional quality and program relevance (as defined by students), including professional development activities for teaching assistants and faculty, are essential (Cooper & Bradshaw, 1984; Gomme et al., 1993; Shively, 1989; Smith et al., 1980). Quality is clearly important for all students; mature students, who have no time to spare for meaningless learning experiences (Allen, 1995; Justice, 1997) and often sacrifice to enter a post-secondary program, are particularly concerned regarding program quality.

Academic Support

Access to individualized academic support through activities such as learning skills classes, learning style analysis and intervention, tutoring, or specific remedial assistance is critical for "underprepared" students (who include many minority group students); returning students lacking confidence or out of practice at studying; and students with learning disabilities, English language limitations, or other potential learning barriers (Allen, 1995; Byrne, 1989; Crosson, 1992; Gardner, 1992; Metzner, 1989; Nelson et al., 1993; Smith, Lippitt, Noel & Sprandel, 1980; White, 1983). An early warning system, to identify students at risk academically and to make immediate assistance available to them, can be vital (Gardner, 1992; Smith et al., 1980).

Environmental Support

Support in dealing with potential financial barriers to entrance or persistence is essential, especially for students from lower socioeconomic status backgrounds (Allen, 1995; Crosson, 1992; Gilbert & Auger, 1988; Gomme et al., 1993; Smallwood, 1980; Young, 1992). Assistance in dealing with issues such as housing, child care, and transportation are particularly important for students with family responsibilities or limited financial resources (Allen, 1995; Byrne, 1989; Hodgson, 1989; Smallwood, 1980; Thacker & Novak, 1991).

Mentorship

In addition to the value of formal faculty mentorship programs, it appears critical that faculty take responsibility for more than just the cognitive development of students by making time for their students individually (Cooper & Bradshaw, 1984; Gardner, 1981, 1992). For older students in particular, faculty need to "step down from the podium", respect the skills and knowledge adult students bring, and interact as peers engaging in a common learning venture (Justice, 1997). In addition, peer support networks and peer mentoring can assist new students to cope with the potentially overwhelming demands of complex campuses and unresponsive institutional bureaucracies (Byrne, 1989; Gardner, 1992; Jewell & Lubin, 1988; Justice, 1997; Thacker & Novak, 1991).

Organizational Adjustments

Flexible course scheduling and program delivery can be essential to students, particularly mature students (Allen, 1995; Baker & Velez, 1996; Byrne, 1989; Hodgson, 1989; Justice, 1997). Examples include scheduling classes so that students with jobs or families can attend, facilitating part-time enrollment, providing opportunities for prior learning assessment so that experienced students can receive credit for previous learning, and providing a blend of on- and off-campus courses. Flexibility related to evaluation practices (e.g., with regard to deadlines for completion of course work) can also be important to students who cannot always control the external factors interfering with their studies. Clustering courses so that students are together for several classes, and content can be integrated across subjects, has been shown to support students' academic and social integration (Crampton & Holm, 1993; Metzner, 1989).

Respecting Diversity

A commitment to respecting diversity appears essential to creating a success-oriented climate for equity group students and counteracting the alienation and estrangement associated with minority group attrition (Baker & Velez, 1996; Crosson, 1992; Gardner, 1992). Effective employment equity practices are recognized as an important success component: Role models, employed on campus, can influence equity group students' perseverance in continuing their education (Crosson, 1992).

Success Strategies and Institutional Reform

It is clear from the literature that a wide variety of potential strategies exist, and that each has many advantages and advocates. It appears, however, that significant impact is only likely if a multiplicity of approaches is in place and an overall institutional commitment to student retention is demonstrated (Crosson, 1992; Pascarella, 1986; Smith et al., 1980). Creation of an overall "staying" or "holding" environment is advocated. Alfred, Peterson, and White *(1992) identified high performing colleges as having a "facilitating culture"* that empowers staff to help students achieve their goals. Attention to student interests and goals must permeate all facets of an institution in order for success to be maximized:

> In short, improved retention starts with the development of a holding environment which can be created through the careful design and the thoughtful delivery of quality academic and related student services and experiences. And it really begins when the institution enters into an internal dialogue about its mission and the quality of life it hopes to build on its campus. (Smith et al., 1980, p. 93)

The list of change indicators for increased retention provided by Smith et al. (1980) illustrates the breadth of activity required: advising procedures; curriculum; mission statement; student aid policies; procedures for assessing student opinions; and even criteria for faculty hiring, promotion and tenure are all recommended for review and revision. For nontraditional students, the requirement for broad institutional change, rather than specific "add on" programs or services, is even more intense (Allen, 1995; Baker & Velez, 1996; Justice, 1997).

The importance of creating a post-secondary environment committed to students and centred on both learners and learning has become an increasingly prominent theme in the adult education literature (Alfred, et al., 1992; Barr & Tagg, 1995; Hall & Shiffman, 1996). The language and recommendations of the student success literature are highly congruent with that of the literature on post-secondary system reform. Maximizing learner satisfaction and increasing the focus on positive outcomes of post-secondary education are currently recognized as critical to the reform of post-secondary systems around the world (Faris, 1995). Across Canada, post-secondary systems are becoming attentive to learners and learner outcomes (e.g., the Association of Colleges of Applied Arts and Technology of Ontario, 1995, 1996). This trend is echoed in British Columbia. For example, the recently released strategic plan of the college and institute system (B.C. Ministry of Education, Skills and Training, 1996a) makes movement to a learner-centred approach to post-secondary education, and a focus on outcomes including learner success, major priorities for post-secondary reform in the province.

Gaps and Inconsistencies in the Literature

It is generally agreed that despite the volume of research into persistence and attrition, results are confusing and inconclusive (Astin, 1984; Brower, 1992; Cabrera et al., 1992; Darkenwald & Gavin, 1987; Gomme & Gilbert, 1984; Tinto, 1993). At most one third of the variance in student attrition is generally accounted for by either the predictive ability of specific background factors or the impact of various interventions (Hatcher, Kryter, Prus & Fitzgerald, 1992; Metzner, 1989; Okin et al., 1996). According to Darkenwald and Gavin (1987), despite all the research, "understanding of the dropout phenomenon has progressed very little" (p. 152).

There are several reasons for this situation. Many authors cited inadequate theoretical grounding as causal to the weaknesses of the research (Attinasi, 1989; Bean, 1990a; Darkenwald & Gavin, 1987). The literature is confused by varying definitions of *dropout* and inconsistent choice of variables to measure the phenomenon (Astin, 1984; Bean, 1990a;

Darkenwald & Gavin, 1987; Tinto, 1982, 1985, 1993). Often, withdrawals related to inter-institutional transfer or *intentional withdrawal* (i.e., those students who did not intend to complete) are not distinguished in the research; nor are voluntary and involuntary withdrawals studied separately (Tinto, 1993). Studies of nontraditional students are further impacted by confusion in definitions of *part-time* (Devlin, 1989; Stewart, 1990; Thompson & Devlin, 1992) and in age parameters for *older* or *adult* students.

Methodological variations and shortcomings limit the generalizability of much of the research (Astin, 1984; Levin & Levin, 1993; McKeown, Macdonell, & Bowman, 1993; Pappas & Loring, 1985), producing sometimes conflicting results. For example, Ashar and Skenes' (1993) study refuted many standard attrition research results, but faulty procedures and definitions place their findings in question. Lewallen (1993) showed that controlling for other variables removes from the impact on attrition one of the more consistently identified persistence factors, student "decidedness." Proximity to exam period can influence students' likelihood of responding and their responses to surveys (Cooper & Bradshaw, 1984), yet few studies consider possible implications of the timing of data collection.

Some studies rely on potentially inaccurate student reports, rather than formal records, for measurement of progress, or fail to monitor the implementation of success interventions to ensure that students actually receive the intended services (Pascarella, 1986). Cross sectional and "autopsy" data (collected after students have made the decision to leave) continue to be relied upon despite evidence of their unreliability (Gass, 1990). Students' survey responses often give superficial explanations for their behaviour, providing socially acceptable reasons for their departure (such as finances, lack of time), which are not always accurate and can mask more complex, personal reasons (Garland, 1993; Gilbert & Auger, 1988).

The research also suffers from a lack of context or representation of the student perspective (Attinasi, 1987; Darkenwald & Gavin, 1987; McKeown et al., 1993; Van Stone et al., 1994). As Attinasi (1987) concluded, studies have tended to "strip away the context surrounding the student's decision to persist or not to persist in college and exclude the student's own perceptions of the process" (p. 350). In the discussion of outcomes, little attention has been paid to student goals, attitudes, or satisfaction.

Results from divergent student populations are often aggregated, despite growing recognition that persistence and attrition are highly individual phenomena that affect specific subgroups of students, and individuals within those subgroups, in unique ways (Baker & Velez, 1996; Brower, 1992; Garrison, 1985, 1987; Langenback & Korhonen, 1988; Lewis, Hearn & Zilbert, 1993; Stage, 1989; Tinto, 1982). Results for nontraditional students have not typically been distinguished in research despite the uniqueness of this student group (Baker & Velez, 1996; Bean & Metzner, 1985; Scott et al., 1996; Tinto, 1993). Thus, divergent trends (such as the association between absenteeism and attrition in traditional but not nontraditional students; Bean, 1990a) are masked and generalizability of results to either group weakened (Gilbert & Auger, 1988; Pascarella, 1986). The potential impact of variables such as gender and ethnicity, which have been shown to impact student decision-making, is obscured (Tomlinson-Clarke, 1994). As Tinto (1993) summarized, "The fact is that even with the recent surge of interest in persistence we still know relatively little about the specific attributes of attrition among females and adults" (p. 76). In addition, large scale studies are not sensitive to the impact of critical programmatic or institutional traits that affect persistence (Aitken, 1982; Lewis et al., 1993).

Research into the effectiveness of intervention techniques is frequently weakened by sampling issues. Many studies rely on volunteer samples: Students who volunteer to participate in the intervention under review are compared with those who choose not to volunteer. However, studies comparing participant and non participant traits show that volunteers tend to possess more of the variables predicting success than those who do not volunteer (Bank et al., 1990; Gomme et al., 1993; Pascarella et al., 1986; Russell & Skinkle, 1990; Steltenpohl & Shipton, 1986), thus confounding the effects of the intervention with preexisting differences. Factors other than motivation may also be involved. For example, students with heavy family or work demands may not have time to take on voluntary activities, even when these are designed to improve success. Any differences between nonparticipants' performance and that of participants would have to be viewed in the light of existing external demands, rather than attributed to nonparticipation in success interventions.

Research regarding interventions also provides insufficient analysis of indirect or conditional effects, relying primarily on unidimensional analyses that potentially overlook genuine but less obvious effects or interactions. Multivariate analysis can reveal indirect relationships that would otherwise be missed, helping to explain the large amount of attrition variance typically not accounted for by univariate designs (Bank et al., 1990; Metzner, 1989; Pascarella, 1986).

Finally, the literature is weak in several areas critical to the current paper. Canadian studies are lacking: Gomme, Hall, and Murphy (1993) speak of the "dramatic under-investigation" (p. 20) of social and academic integration and institutional fit in Canada.

Despite indications that attrition is greater in 2-year colleges, such institutions are relatively neglected in the research (Tinto, 1982; Bean, 1990a); and, again despite indications of high attrition, there are few studies of part-time students (Thompson & Devlin, 1992) or other nontraditional students (Baker & Velez, 1996; Scott et al., 1996), and none identified to date focusing specifically on human service programs.

Several directions for future research have been identified in the literature: (a) development of an improved data base regarding student completion and attrition (Dietsche, 1990; Levitz & Noel, 1985); (b) studies that specifically isolate issues affecting nontraditional students (Allen, 1995; Baker & Velez, 1996; Bean, 1990a; Bean & Metzner, 1985; Dey & Hurtado, 1995; Scott, et al., 1996; Thompson & Devlin, 1992); (c) localized, institution/situation-specific research (Baker & Velez, 1996; Bean & Metzner, 1985; Langenback & Korhonen, 1988; Lewis et al., 1993; Tinto, 1982, 1985); (d) exploration of the context for persistence/attrition, including student perceptions (Attinasi, 1989; Darkenwald & Gavin, 1987; Dey & Hurtado, 1995; McKeown et al., 1993; Noel, 1985; Stewart, 1990; Tinto, 1988; Van Stone et al., 1994); and (e) exploration of persistence that takes into account varying student motivation and intentions (Bean, 1990a; Brower, 1992; Okin et al., 1996; Stage, 1989; Stalker, 1993; Tinto, 1985, 1993).

Student Retention Research and Initiatives in B.C.

The topic of student retention is not going unnoticed in B.C. This section will review some provincial research related to student success, and will revisit the issue of research gaps and limitations in light of B.C. involvement.

In general, research and comprehensive data on student retention have been limited in the province. Most available analyses are unpublished internal documents prepared by individual institutions. Camosun College in Victoria has played a leadership role among the community colleges in developing a framework for gathering consistent retention data and in promoting utilization of student success strategies. Since 1988, Camosun College and the University College of the Cariboo (UCC) in Kamloops have gathered retention data according to a consistent format and with highly consistent results. The software used in these analyses, initially developed by Camosun College, has recently been made available throughout the post-secondary system to encourage more standardized data collection. In addition, the three established universities in B.C. maintain internal retention data. Province-wide data was formerly gathered by the Strategic Information Research Institute (SIRI), although this data focussed primarily on transfers and student flows, rather than student success. In her doctoral study, Beatty-Guenter (1994) analyzed use of retention strategies by all the province's colleges and compared these to retention rates. In addition, graduate work by other post-secondary personnel in the province (Cooper, 1993; Springate, 1993) shed further light on retention issues.

In contrast to the limited data available on student retention, B.C. has an impressive history of measuring outcomes for those students who do complete programs. Annual surveys of college and institute graduates have been conducted since 1987. Currently, telephone interviews of all students who can be contacted are conducted one year after program completion. B.C.'s universities are only beginning to collect and analyze outcome data at a provincial level; to date, analyses have been conducted 2 and 5 years after program completion. These studies provide useful information concerning the nature of graduate populations, graduates' satisfaction with their programs and their current status in terms of employment and further education. It is interesting (and consistent with the general tendency in higher education) that the province has paid far closer attention to what happens to its post-secondary graduates than to those who do not complete programs.

B.C. Research Findings

As indicated, B.C.'s post-secondary institutions are utilizing an expanding array of strategies associated with student retention (Beatty-Guenter, 1994; B.C. Ministry of Education, Skills and Training, 1996d; B.C. Ministry of Skills, Training and Labour, 1995). Beatty-Guenter (1994) found that institutions were increasingly using the strategies she identified, but particularly those characterized by their "simplicity, visibility, compatibility [with the campus culture] and pilot-ability" (p. 149), such as academic advising, co-op education programs, remedial education or learning skills classes, and financial aid. This is generally supported by Ministry surveys (1995, 1996), which highlight an impressive number of retention strategies in place in many institutions and an increase in utilization over time. Both Beatty-Guenter and Ministry studies identified low institutional involvement in peer programs, and the Ministry reviews also identified low participation in supports aimed at women, returning adults and cultural minorities other than First Nations. These findings are limited, however. Although both Beatty-Guenter and Ministry research attempted to identify the extent to which interventions are applied, for the most part the critical issue of availability/intensity was not adequately assessed. For example, studies reported high access to on-campus child care, reflecting the fact that most campuses now have on-campus facilities but obscuring the limited number of spaces this actually entailed and the fact that costs are often prohibitive for student parents.

Actual retention statistics reported in the various B.C. reports are not inconsistent with the literature, although program-specific data reveal significant variations between programs, underscoring the risk involved in aggregating data. For community college students entering in Fall, 1991, Beatty-Guenter (1994, 1995) found within-year retention rates of 76% in academic programs and 69% in career/technical programs, and that overall retention rates across the post-secondary system had risen between 1989 and 1991. Springate (1993) examined persistence rates for students in three diploma programs at Okanagan University College (OUC), finding significant differences between the three programs. Average completion rates in a longitudinal study of three OUC cohort groups per program ranged from 35% to 77%. University College of the Cariboo--UCC (1992) data indicate that "about 60% of the students at least made it through the first two semesters" (p. 3), although persistence rates (labeled "survival rates") again varied significantly by specific program and program category. For example, on average 67% of students in 2-year career/technical programs "survived" from first to second year, compared to 39% in academic programs (UCC, 1995).

B.C. Research Limitations

The interest in this critical topic is encouraging, although B.C. research echoes the approaches and limitations which predominate in the broader academic literature. These will be discussed below.

Operational Definitions

Although use in the future of common software may promote common definitions and thus comparability across institutions and studies, research to date has not been based on standardized terminology. For example, Springate (1993) limited the term *dropout* to a student who "leaves a program or course without notice and does not return to that particular program or course" (p. 16). According to this operational definition, those who withdraw formally or withdraw from one program but enter another would not be considered dropouts. *Attrition* is defined by Springate as "the decrease in student count from one period in time to another" (1993, p. 41), and by UCC (1994) as "withdrawal from the college" (p. 2). *Retention* is defined as "the rate of student count in a program or course from one semester to the subsequent semester (or from one year to the subsequent year)" (Springate, 1993, p. 17). Beatty-Guenter (1994) suggests that retention "refers to the students maintaining their enrollment through the normal length of a course of program of studies" (p. 4). Although similar, these slight variations in terminology leave room for significant differences in operational definitions of the key variables under review.

Data Limitations

The data collected to date provide only a superficial picture of the student completion situation. As indicated by UCC (1994), the inability of the data to distinguish between those who dropped out, *stopped out* (took a temporary break from studies), or transferred elsewhere is a serious limitation, leaving approximately 20% of any cohort unaccounted for. The SIRI (1994) data show significantly lower retention over a 4-year period for colleges as compared to universities, but fail to take into account stopouts, transfers, or (most significantly) students who exit from college programs because they have completed a certificate or diploma. Other factors that confound the existing B.C. data include: the large number of part-time students occupying full-time seats, making comparisons of intake to graduate numbers an inaccurate reflection of retention

(Springate, 1993); the rapid changes in institutional record-keeping systems, policies, offerings, and so on (e.g., the university colleges moving to degree granting status) that limit longitudinal comparisons (UCC, 1994); the wide divergence in student numbers between programs, making some between-program comparisons inappropriate (UCC, 1994); the failure in some data to account for programs "topping up" enrollment, to compensate for attrition, by admitting students midyear (UCC, 1991); and the lack of suitable tracking systems for continuous intake programs (UCC, 1992).

Student Perspectives

Consistent with the literature, B.C. studies are notably lacking in inclusion of student perspectives or attention to student goals and attitudes. Several of the provincial reports note this limitation. Springate (1993) commented that OUC neither required students to declare their goals on entrance nor conducted exit interviews, and suggested that "a system of measuring the achievement of student goals rather than course or program completion might be a more accurate indication of the usefulness and success of a particular course or program" (p. 112). UCC researchers (1989) suggested that the institution's withdrawal form did not gather enough information about why students drop out.

The systematic measurement of goals and attitudes of B.C.'s post-secondary graduates, by way of the outcome surveys mentioned previously, is noteworthy. Similar data collected from students who leave their programs would provide comparable perspectives from those students who do not graduate.

Nontraditional Student Issues

Consistent with the literature, B.C. research does not generally separate and analyze results for nontraditional student groups. Although some of the data is broken down by age, gender, or both (Springate, 1993), these factors are looked at separately. Results for older women, for example, are not reported. Equity group membership is otherwise not addressed in the retention research.

However, the outcomes surveys previously referenced do differentiate between outcomes for women, visible minority and First Nations students, and students with disabilities. The profile of First Nations college and institute program completers is interesting (B.C. Ministry of Education, Skills and Training, 1996b). Some of the significant differences between First Nations and non-First Nations respondents included the following: First Nations students were more likely to enroll in programs categorized as "legal/social" (the category that includes human service programs), although proportional enrollment in all other program areas was comparable to non-First Nations; the population included more females and was older; less had finished high school; more were in vocational (versus career/technical or academic) programs; and less were employed at the time of the survey.

Because of the existing paucity of data on students with disabilities, several post-secondary institutions have been funded to complete localized retention and outcome analyses for this student group, and a provincial study has been proposed. A co-op student report (Manak, 1995) identified the absence of data, including retention and outcome measures, concerning students of colour and new immigrants.

Context Issues

Although institution-specific data are broken down by program area and Springate (1993) provided analyses by demographic variables, the provincial studies (Beatty-Guenter, 1994, 1995; SIRI, 1994, 1995) generally aggregated data across large cohorts of students, thus losing the benefit of useful contextual information. Beatty-Guenter (1995) commented on the lack of individual student data in the B.C. retention research, and the need to add demographic variables:

> This study did not seek individual level explanations for differing retention rates, yet these important factors (such as economic circumstance and prior academic preparation) vary by institution, and can have profound effects on the appropriateness of certain retention programs. (p. 13)

Beatty-Guenter (1994, 1995) focussed on the relationship between implementation of retention strategies and provincial increases in retention, with limited exploration of other possible sources of these increases.

The UCC studies (1989, 1991, 1992) illustrated the value of looking at retention data in context. Institutional changes over time and their potential impact on the data were described. Program-specific results were placed in the context of essential background details that help to understand retention patterns and could assist in identifying effective intervention strategies. For example, midyear attrition in one program was explained by the fact that the "core" program, which provides the level of training necessary for entry to employment, does not require the whole year to complete. The provincial graduate outcomes surveys also highlighted both the significant variations between programs and variations within programs for specific student groups.

Evaluation of Strategies

For the most part, the B.C. research lacks evaluation of the strategies being introduced (Beatty-Guenter, 1994; Springate, 1993). Some exceptions exist. For example, OUC (Springate, 1991), UCC (1992,1993) and Camosun College (1992) have each evaluated the outcomes of student success/orientation courses. The first two found no clear evidence regarding impact; the latter ignored the sampling bias discussed above and assumed that positive outcomes for participants were attributable to the course. The UCC study (1993) concluded that quantitative analyses of results needs to be complemented by a qualitative look at outcomes: "Future study should focus on the qualitative approach and on measurement of non-cognitive achievement of students" (p. 1).

B.C. Research Orientation

The B.C. studies add to the information base on retention and completion in the province, providing the quantitative and aggregate data which are clearly needed. These studies build on a literature base which defines retention empirically and appears to assume that success can be quantified, that institutional interventions and student responses are causally related, and that student behaviour can be understood by looking at aggregate patterns. Examples of this orientation will be outlined below.

Cooper (1993), in a University of British Columbia (UBC) graduate thesis, distinguished success, "the qualitative experience of the student", from retention, "the quantitative experience of graduating from university" (p. 3), yet scarcely mentioned the former in his research analysis. Retention is accepted as an adequate proxy for success: "Student retention is a quantifiable measurement that can be used to indicate how successful both individuals and universities are in capitalizing on human potential" (p. 3).

Cooper accepted the premise that "overwhelmingly, students of today choose to attend university to gain improved employment opportunities" (p. 11), and quoted Gardner's 1989 statement that "today's students are majoring in pre-wealth." Cooper's philosophical orientation is further illustrated through his reaction to a quote from then UBC president, David Strangway, that "the prime objective of the university should be to admit only those who are the best qualified to benefit from and contribute to programs in which they are enrolled." Cooper commented, in support, that "accessibility to higher education can be accommodated through other higher education institutions which have a mission for open accessibility" (p. 29) and further concluded "Currently far too many students who attend university are unmotivated, unprepared and unfit" (p. 144). Within the context of Cooper's view of success as defined by employment outcomes, and of this traditional "selection/survival of the fittest" approach to university education, the absence of a student perspective or concern with qualitative aspects of success is not surprising.

Beatty-Guenter's (1994) dissertation focussed on institutional behaviours with regards to implementation of retention strategies. Attention remains at the systemic and generic level. An instrumental approach is demonstrated by the language of the typology of retention strategies suggested by the author, which includes: "sorting--grouping students into subsets" (p. 20) and "transforming", which focuses on strategies to teach students to perform better as students and institutions to maximize student performance, and is not about developmental transformation. Beatty-Guenter's (1995) work relied on cohort analysis, which is described as "gaining prominence in the research literature since it has become technically possible to store and match large data bases" (p. 5).

Springate's (1991, 1993) research incorporated some analyses of students by program area and demographic traits, and complemented survey and statistical data with individual interviews. An instrumental orientation is revealed, however, in a statement summarizing the benefits of improving student retention as "an enhanced image of the institution due to the increased success rate of students, an increase in student tuition fees, and pre-emption of any initiatives by the Ministry to change funding patterns [to move to outcome-based funding]" (1991, p. 24). Springate (1993) focussed on *student non-persistence*, a term that appears to place full responsibility for attrition on the student.

Conclusions

It appears that the approaches to the study of student retention that dominate the literature are also found in B.C. research. Further studies of student retention at provincial and institutional levels will undoubtedly be completed. This will provide better and more consistent data about retention and completion rates and better evaluation of interventions, thus improving the quantitative basis for analysis. However, both quantitative and qualitative analyses are needed to comprehend the complex phenomenon of student success. Without revisiting how student success is defined, questioning the equation of success with retention, and exploring new ways to research the phenomenon, results will likely continue to have limited relevance to individuals and the unique issues and contexts each experiences. Students will not be better assisted to get what they want and need from an education until success is redefined with their input, taking context into account and including a qualitative element. Better data need to be accompanied by better understanding, and this can only be derived from "connecting with" individual students in order to learn what success means to them and how they feel they can best be assisted to achieve their goals.

The UCC (1989, 1991, 1992) data provided program-specific completion figures indicating that 1-year certificate programs in the human services and the newly established Bachelor of Social Work program have high levels of student retention. If completion rates are viewed as the only relevant indicator for studying student success and this trend can be replicated province-wide, an inquiry into student success in human service programs may appear redundant. Even with high completion rates, however, questions remain.

First, success in terms of completion ignores the more important issue of success in terms of student growth and learning. Even though students complete programs, it cannot be assumed that they are successful, without exploring pertinent issues: Have students' personal goals been met? How do they define success, and do they define themselves as successful? Have their educational experiences opened doors and encouraged students to consider further learning opportunities? Have they changed as a result of the experiences? Will they be more effective in providing care to vulnerable individuals?

Second, have the unique issues affecting part-time students, single parents, returning women, First Nations students, and so on been addressed and has a climate conducive to both retention and success been created for them? Have those who did not complete stopped out (and will they return), or did they transfer or change to another program, or have they dropped out and, if so, why?

Finally, given the increased attention to student success as a measure of institutional effectiveness and one of the key variables for institutional reform, what can we learn about the approach taken in human service programs, with nontraditional students in particular (the majority of students in such programs), which can assist in maximizing success and retention for similar learners in other program areas? Can students in human service programs, who tend to complete their programs, express a high degree of satisfaction, and evidence high levels of graduate employment (B.C. Ministry of Education, Skills and Training, 1996) aid in understanding the phenomenon of success and identify what institutions can do to promote success for other nontraditional students? Can these students assist with operationalizing current rhetoric regarding post-secondary education's transformation to a more "learner-centred" focus?

Employing intervention strategies to maximize retention and collecting data to monitor student outcomes are essential elements of a high quality, cost-effective post-secondary system, but only part of what is needed. These basic standards need to be complemented by attention to other outcome measures, including such qualitative factors as whether students' goals are met and whether they perceive their educational experience as having been valuable. As Beatty-Guenter (1994) pointed out, "retention [is] a necessary condition for achieving other goals" (p. 5); however "retention per se is not the goal" (p. 6), and maximizing retention should not be presumed to be sufficient to maximize student success. More complex, subtle, and important aspects of success remain to be explored.

Rationale and Scope of the Current Study

The provincial government has become increasingly interested in the outcomes of the educational programs it funds. To date, more efforts to measure attrition have been made at the secondary than the post-secondary level. However, it appears that school system attempts may be based on flawed data: A newspaper column entitled "Numbers-challenged Bean

Counters Double the High School Drop-out Rate" (Bates, 1994) suggested that faulty dropout definitions result in reported dropout rates that represent double the number of students who actually do not finish school. Similarly, U.S. college student research has traditionally significantly underestimated the extent of eventual degree completion (Tinto, 1993, p. 25).

As the post-secondary system attempts to become more accountable in terms of student outcomes, and student success initiatives become more popular, it is critical that the phenomenon of student success be carefully defined and the issues surrounding it well understood. This is essential in order to avoid collection of misleading data and investment in inappropriate interventions. New approaches are needed if genuine progress is to be achieved in the study of this much studied phenomenon. Traditional research approaches, presenting students as members of cohorts and relying on quantitative indicators of success, have not adequately considered students' perspectives or individual and institutional/programmatic variables relevant to the success phenomenon, nor have they proven sufficient to explain the complex and individualistic issues characterizing this topic. Alternative research strategies are necessary.

The focus of this inquiry is on students considered nontraditional, with particular attention to those in 1-year certificate through to degree programs in the human services. Understanding and effectively supporting successful outcomes for nontraditional students is important due to the sacrifices many of these students make to enter the post-secondary system, the unique challenges they face in completing their education, and their capacity for exceptional contribution, both as students and eventually as employees. Recent studies affirm the potential of nontraditional students: "Nontraditional students may, in some institutions, be the most motivated, most academically talented, and the most emotionally strong on campus" (Baker & Velez, 1996, p. 96); "Studies have shown that mature age students make exceptional students who are very motivated and perform well academically" (Scott et al., 1996, p. 233).

Additionally, many nontraditional students represent groups (e.g., women, First Nations people) that are frequently disadvantaged in socioeconomic terms and for whom completion of post-secondary education is vital to altering patterns of low skill, low status, and poverty.

Successful completion of education for students entering the human services is crucial due to the significant competency and knowledge requirements of human service occupations; the shortage of adequately trained personnel in all the human service disciplines; and the limited educational programming available to respond to the demands of the field, which increases the importance of maximizing positive outcomes for every available student space. The current study will focus on human service students who are women in order to limit the scope of the study and enhance its depth and because women represent a dominant majority of nontraditional students, participants in these programs and employees in the field.

Research Focus and Questions

This study was designed to be exploratory in nature, with the goal of gaining an understanding of the phenomenon of student success from the perspective of female nontraditional students in the human service disciplines at one B.C. college. The intent of the project was to enrich the study of student success and complement traditional student success research by focusing on in-depth understanding of a single group of students, exploring students' feelings and motivations, allowing students to tell their own stories, and looking beyond the educational experience to relevant factors in students' personal lives and in the institutional and community contexts.

Written documentation from the college and program was analyzed to enhance understanding of the educational context within which the students are functioning. The primary source of information, however, was students themselves. The study attempted to answer the following questions concerning this unique group of students:

1. What brought them to college, to this field, to this program?

2. What are their goals? What motivates them?

3. What supports are important to achieving their goals, and what barriers could (or do) get in their way?

4. What does success, as a student, mean to them?

In addition, background and basic sociodemographic information was gathered concerning each participant (including program of study, community of residence, family status and background, work and educational history), not for purposes of statistical analysis but to assist with interpretation of the findings, inform identification of possible trends and patterns in the responses, and suggest potential directions for future research.

This inquiry balances the traditional quantitative and aggregate level success research with a focussed, individual, and interpretive look at a single group of students: It moves away from research approaches that tend to overlook or minimize the critical voice and perception of the student, toward an approach that emphasizes and values each student's views and conceptualizations.

The primary goal of the research is to enhance understanding of a complex and critically important phenomenon, namely student success from the point of view of female nontraditional students of human service programs. At the same time, patterns and themes identified by these participants might have relevance to other settings and student groups, and questions for future research will be suggested. The study will provide a step toward formulation of a theoretical framework to explain the concept of student success as related to this particular student group. Finally, government and post-secondary institutional policy implications of the findings will also be explored in terms of approaches to measurement of student outcomes, provision of student supports, and accommodation of nontraditional learners. The results of the study will enrich planning that is currently underway in the B.C. post-secondary system regarding movement to a more learner centred approach, adding the voice of one group of nontraditional students to discussions of how to accomplish this goal.

CHAPTER 2: METHODOLOGY

This inquiry takes the form of a case study, with the defining feature of the case being the collective nature of the student participants. The focus is on the students and not on the college or the educational programs in which they are enrolled.

The participants who are the focus of the study are female nontraditional students currently attending human service programs at one institution; these participant attributes constitute the boundaries of the case. Participants have some contextual features in common, in that they reside in a common geographical region with relatively consistent socioeconomic and labour market features, although they live in communities spread across the college region. Other variations within the identified case include the range of programs from which these students are drawn, encompassing certificate (1 year), diploma (2 year) and degree (4 year) programs in the disciplines of Child and Youth Care, Early Childhood Education, Social Services, Community Support, and Social Work. Students represent various stages of participation (i.e. first, second year, etc.), and both full- and part-time enrollment. However, these possible subunits of the case will not be studied separately; a holistic approach is taken at this point, although possibilities for further exploration of subunit differences may emerge as potential topics for future research.

Information was gathered about the programs and college in order to provide a context for exploration of the students and their experiences. Descriptive details about these environmental systems within which individual students function may assist in determining the potential generalizability of the findings, although this is not the primary goal of the inquiry. The major focus of the study is on acquiring a rich and resonant understanding of the case "in its idiosyncrasy, in its complexity" (Stake & Trumbull, 1988, p. 256), rather than pursuing comparisons and generalizations, and on exploring and articulating students' experiences and interpretations, rather than searching for universalities.

Orientation of the Case Study

The inquiry utilizes qualitative methods, with their emphasis on discovery and understanding. This represents a major shift from the empiricist approaches typical of student success research, with their focus on prediction and control of the phenomenon. Intuitive processes and interpretation guided the inquiry; the researcher's history, attitudes, and presuppositions, along with the literature, guided the interviews and the interpretation of results.

The inquiry embodies several of the key features of ethnographic case study research in that its emphasis is "on exploring the nature of particular social phenomena, rather than setting out to test hypotheses about them." It relies on relatively unstructured data that are not coded in terms of a closed set of analytic categories; it involves investigation in detail of a single case; and its data take the form primarily of verbal descriptions and explanations, with quantitative analysis playing a subordinate role (Atkinson & Hammersley, 1994, p. 248).

Another key orientation of the study is toward feminist and emancipatory traditions. Feminist research approaches are highly relevant to the current study for the following reasons:

1. The dominant majority of students and workers in the human service disciplines is female, and studies in the field have increasingly focussed on the implications of "caregiving as women's work."

2. The human service field, its workers, and the consumers of human services have long been undervalued and given only peripheral recognition in terms of wages, status, or general

societal valuing of their worth to the larger society. A key focus of feminist research is to counter this type of marginalization through reflecting and profiling women's lives, experiences, and views (Acker, Barry & Esseveld, 1983; Ball, 1992; Devault, 1990; Lather, 1986, 1991; Oleson, 1994; Reinharz, 1992). Oleson speaks of making visible "the widespread caring for children, the ill, and the elderly and exposing the taken-for-grantedness of such work, its oppressiveness, and also its value to women and their societies" (p. 159). Reinharz advocates for "making the invisible visible, bringing the margin to the centre, rendering the trivial important, putting the spotlight on women as competent actors, understanding women as subjects in their own right" (p. 248).

3. The focus of the study on understanding versus prediction also fits with feminist research tenets (Acker et al., 1983; Lather, 1991), as does the inclusion of analysis of context. Acker et al. included as a criterion for feminist research that "the underlying social relations that eventuate in the daily lives we are studying" (p. 146) be revealed in the reconstruction.

4. The intention of this inquiry, to ask women to frame their goals and personal definitions of success in their own voices and language, reflects a primary commitment of feminist research (Acker et al., 1983; Devault, 1990; Lather, 1991). As Devault (1990) suggested, "language itself reflects male experiences, and . . . its categories are often incongruent with women's lives" (p. 97). This inquiry provided women an opportunity to explore and explain the phenomenon of student success in ways that make sense and ring true for them.

5. Finally, the fact that this inquiry involved a woman interviewing women also fits with feminist research approaches, which have highlighted the unique and potentially fruitful situation of "women studying women" (Acker et al., 1983; Devault, 1990; Reinharz, 1992).

Usual conversational style between women constituted the standard for interviews, and intuitive evaluation of the progress of interviews as conversations guided the researcher, both within interviews and while analyzing transcripts to inform subsequent interviews.

The feminist research focus on emancipatory outcomes also suits this inquiry, although it creates a challenge. A commitment to consciousness-raising and furthering of a social change agenda permeates feminist research literature. Transformative and emancipatory outcomes are not the primary focus of the current inquiry, although potential opportunities for transformative experiences were not ignored. The reality that participants may not share an interest in social change, and the risk that the emancipatory intent of the researcher could lead participants or bias data analysis toward a compatible outcome, were recognized.

Principles

The principles underlying the research are summarized below. These principles guided the selection of specific research strategies and their implementation. The principles are stated as objectives that represent steps toward the overarching goal of exploring the phenomenon of student success from the perspective of female nontraditional students in human service programs. The principles/objectives are: (a) to use multiple strategies, with the goal of ensuring a deep and comprehensive understanding of the participants and elements of their context; (b) to allow students to tell their own stories in their own words; (c) to provide a diverse group of women with the opportunity to participate; (d) to pursue hints and nuances in students' stories, encouraging exploration of roles beyond that of student; (e) to rely on interactions that are not tightly scripted, following general themes rather than pursuing specific hypotheses; (f) to focus on individual situations and unique contexts rather than the unearthing of universal truths; (g) to bring the researcher's perspective into the inquiry, while being attentive (during interviews and while reviewing transcripts) to the risk of not being open to other perspectives; (h) to attend to the issue of potential power imbalances between researcher and participants; (i) to acknowledge any external factors that participants could view as related to the inquiry and ensure that these are made overt and discussed; and (j) to enhance the empowerment agenda of adult education through provision of opportunities for students' voices to be heard and through facilitation of reflection.

Procedures

Summary

The phases of the proposed research are summarized below, with details regarding each component provided in the section following. It should be noted that the proposed methodology was approved by Human Research Ethics Committees at both the college involved and the University of Victoria, and that procedures included steps designed to maximize confidentiality and ensure participants' informed consent.

Student Data

1. College program coordinators were provided with an overview of the intended research and asked to identify individuals for pilot and key informant interviews. Coordinators distributed information regarding the project to students who met defined criteria. Interested students self-identified and the researcher contacted them directly.

2. Two pilot interviews were conducted, audiotaped, and analyzed to ensure that the research approach was appropriate, effective, and consistent with the research principles.

3. All students who volunteered were contacted and interviewed. Twenty-five interviews, lasting approximately 1 hour each, were conducted. Most interviews were

audiotaped. Interview notes or transcripts were prepared immediately following the interviews.

4. An additional nine interviews were conducted following the completion and analysis of the initial interviews. Participants for this phase were solicited through open invitation.

5. Analyses of the transcripts were based on both manual and computer-based processes.

Contextual Information

1. Program coordinators were requested to suggest opportunities for participant observation throughout the research period; however, no opportunities were suggested.

2. Faculty were interviewed informally to provide background information regarding the programs and to discuss themes emerging from the student interviews.

3. Relevant documents were collected and analyzed throughout the research period.

Description

Student Data

Participant selection. Participants are selected for qualitative case studies based on their potential to shed light on the research topic from their individual perspective and to further the goals of the inquiry. As Stake (1994) recommended regarding selection of participants for case study research, "The primary criterion is the opportunity to learn" (p. 244).

Program coordinators were asked to provide invitations to participate in the study to mature female students viewed as perceptive and able to articulate their personal situation and goals as well as shed light on the broader issues affecting nontraditional students (see Appendix B). All those who subsequently volunteered (27 women) were interviewed. It was emphasized to program coordinators that the intent was to locate students able to illuminate the phenomenon of student success for nontraditional students, not to select students because they were strong academically or articulate regarding program and course content. Interview findings indicated that coordinators selected appropriate interviewees for this portion of the research; the only challenge was the number of suitable candidates identified, which was larger than expected. Representativeness of the participants was not a selection criterion, although coordinators were encouraged to promote participation of a diverse student group.

Initial interviews were followed by recruitment of additional participants. At this stage, all female nontraditional students were invited to participate through a mailing to all female students, intended to ensure that reliance on instructors to recommend participants did not eliminate other women who wanted to participate and share their stories. This generated a further nine interviews. Most women who contacted the researcher at this stage were interviewed, with the exception of one who did not qualify as nontraditional and two who made contact after interviewing was completed.

Interview techniques. It is often difficult to outline in advance the specific interview questions and techniques to be employed within a qualitative framework, where adjustments are typically made throughout the course of the inquiry (May, 1989) and a variety of approaches is typically employed (Reinharz, 1992). Guidance from research literature is limited due to the continuing focus on more structured approaches, even though unstructured approaches predominate in the qualitative paradigm (Burgess, 1984). As Oakely (1981) concluded 17 years ago "Interviewing is rather like marriage: everybody knows what it is, an awful lot of people do it, and yet behind each closed door there is a world of secrets" (p. 31).

The primary premise of interviewing in qualitative inquiries is that participants have the right to be heard and to "name" their world, and the capacity to tell their own stories in their own words (May, 1989; Mishler, 1986; Oakely, 1981; Payne, 1990; Peck, 1986; Reinharz, 1992). Efforts were made in the present inquiry to establish a comfort level with participants which encouraged them to share their personal stories, openly and in depth. Audiotaping was used wherever possible to ensure accuracy. Although participants were asked whether they minded audiotaping, none indicated any discomfort. Those few (three) interviews that were not taped occurred in restaurants or other public settings where taping would have been difficult. Detailed note-taking replaced taping in those instances. Although a form of shorthand was used and an effort made to keep the note-taking unobtrusive, the flow of these few interviews may have been affected and transcribing of the conversations was less complete.

In order to ensure that participants told full stories, reflective of the complexity and interconnectedness of the various aspects of their lives, interviews were not rigidly scripted. Participants were encouraged to include information about themselves and their lives, beyond their role as student. A list of topic areas was developed to guide the interviews (see Appendix C). Topics were general in nature. Identification of areas of interest related to each topic evolved over the course of the inquiry as issues raised in earlier interviews were pursued with later participants. Even later in the process when themes were emerging, open-ended questions were used to encourage women to express their own views in their own words (Burgess, 1984; May, 1989). Prompts were used to encourage participants to more fully explore their experiences and feelings.

Interviews represented open discourse that acknowledged the researcher's active and personal involvement. Techniques identified in the literature as effective for minimizing power imbalances between researcher and participant were employed. These include ensuring that participants are fully cognizant of the goal of the research and the destination of the information gathered (Kirby & McKenna, 1989); bringing researcher experiences into the discussion (Burgess, 1984; Kirby & McKenna, 1989); willingly answering participants' questions (Oakely, 1981); and yielding some control of the interview and research process to participants (Lather, 1991; Mishler, 1986; Oakley, 1981; Reinharz, 1992).

The interview approach was built on traditional woman-to-woman talk with its collaborative intent, responsiveness to nuances, and sensitivity to personal details of women's lives. As Devault (1990) summarized:

> Women together can more easily cooperate in understanding each other than speakers in mixed groups. . . . Women interviewing women bring to their interaction a tradition of "women talk." They help each other develop ideas and are typically better prepared than men to use interviews as "search procedures," cooperating in the project of constructing meanings together. (p. 101)

As part of this process of jointly constructing meanings, themes emerging from earlier interviews were checked out with subsequent participants to involve them in the process of interpreting and pursuing understanding of the findings.

Finally, opportunities to reduce anxiety, give outlet for verbalization of feelings and reassure participants of their normality (Oakley, 1981); confront silence (Westwood, 1992); and raise consciousness (Ball, 1992) were attended to (rather than ignored as typically occurs in the empiricist paradigm). Participants were given the opportunity to be interviewed together, to promote shared understanding of the common issues affecting them. However, women opted to be interviewed together on only two occasions (involving two women each time), largely because the complexity of their schedules made individual scheduling more realistic than trying to coordinate with others.

Contextual Information

Observation. Although observation was not intended as a significant component of this inquiry's methodology, opportunities to supplement other sources with data gleaned from observation were sought. However, no such opportunities were identified by college contacts, and this was not seen as a sufficiently significant source of information to pursue.

Instructor interviews. Although the focus of this inquiry was on students' perspectives, it was recognized that instructors could contribute toward understanding the context in which participants are studying and that faculty insights could aid in interpretation of student data. Faculty were interviewed informally throughout the student interview process. Early faculty interviews assisted with understanding the programs, expectations, and college climate; later interviews served to validate the themes that were emerging and the conclusions the researcher was drawing.

Document review. One of the array of methodological approaches commonly used in qualitative inquiries is review and interpretation of documents and records. According to Quinn (1987), reviewing documents provides information about program activities and processes, suggests important questions to pursue through other strategies, and may also identify program goals or intentions that, according to other data sources, do not appear to have been realized.

For the current inquiry, understanding the educational context was enhanced through review of documents describing program and institutional goals and priorities, including public descriptions (e.g., calendars, brochures) and internal policy documents. In addition, the perspectives of previous students were examined through review of interview notes and reports from previous relevant studies.

Data Analysis

Transcription of audiotapes and interview notes occurred as soon after interviews as possible; tapes were reviewed repeatedly to ensure accuracy and enable notation of nonverbal nuances. All transcription was completed by the researcher.

Computer-based data analysis was employed to assist with general data management and with categorizing and retrieving information from interview transcripts. The efficiencies afforded by computer applications supported the inquiry through completion of organizational tasks that would otherwise have had to be conducted manually (Fielding & Lee, 1992; Richards & Richards, 1994). The "ATLAS Text Interpretation" program was selected as the software program to be utilized with interview data. Its designer describes ATLAS as "a tool which effectively supports you, the human interpreter, particularly in handling large amounts of research material, notes, and associated theories . . . an effective and highly efficient work aid to the human interpreter (Muhr, 1993, p. 2).

Specifically, the ATLAS program assisted as follows:

1. Although decision-making regarding coding of quotations was not performed automatically (that is, the program's "automatic coding" function, based on a key word search technique, was not employed), the program automatically implemented the coding function and organized the coded quotations for storage, retrieval, and further analysis.

2. Memo writing was encouraged by the program and found useful, with the program providing automatic storage and retrieval of memos.

3. Notation of comments was encouraged throughout data analysis in the ATLAS program and ensured that interpretations emerging from the data were not forgotten; comments were stored for easy access.

4. Many forms of data manipulation were possible once quotations were coded, and were employed to assist with interpretation of results. For example, interrelationships between codes were analyzed and codes compared for frequency of use.

5. Code families were created, linking related codes and assisting with identification of themes in the data.

6. Interview texts were sorted (e.g., separating transcripts of married students from their unmarried peers; sorting transcripts by program area and year of study) and code frequency and relationships between coded quotations were compared between groups to suggest possible subunit differences.

To summarize, every available function provided by the ATLAS program, in terms of organizing, sorting, and analyzing the data, was employed to assist with interpreting the findings and to lend confidence to the conclusions reached. Quotations were assigned codes; code families identified; and memos, relating to the coded quotations, produced to point to themes emerging from the data and suggest the outline for presentation of the findings within these themes. Quantitative data analyses (such as relative frequency of occurrence of codes and code families) were used to assist in validating the researcher's interpretations. Finally, the program simplified organization and retrieval of quotations for inclusion in the text.

Analysis of documents relied on manual approaches to organizing and sorting information, because the volume of information of any one type was limited.

Methodology: Conclusions

This study aimed to develop a rich, deep understanding of student success from the perspective of female nontraditional human service students attending one B.C. college. Contextual data were reviewed and analyzed, but the primary methodology was provision of opportunities for students to describe their own goals and definitions of success, as well as the supports and barriers affecting them.

In stepping back from the large-scale studies that characterize the student success literature, this focussed inquiry attempted to illuminate the issue more personally and perceptively for a single student group. Overall student outcome patterns are certainly important; so too are individual stories and themes particular to specific sets of student. Encouraging students to talk about the issues important to them allows and elicits identification of factors that would not be captured through standard survey approaches. Looking at the student in the context of the program and institution, her community, and the many roles she plays (in addition to that of student) assists in developing a more realistic and holistic comprehension of participation and persistence patterns for these individuals than can emerge from traditional research, with its tendency to look at students in isolation from contextual factors.

The primary goal for this study was learning about and seeking meaningful patterns within the boundaries of the case under review. However, in the course of acquiring valid and detailed understanding, the learning that occurred cannot fail to suggest themes and beginning theories that may have broader application. In the summary and analysis that follows, the goal of understanding should not be preempted by a search for generalizations. Moving too quickly to generalization risks yielding neither accurate generalizations nor interpretations that assist with understanding individual students. In addition, focusing too heavily on generalization could risk implying that the views and conceptualizations of the individuals participating in the inquiry are not valid and valuable in their own right, which would be contrary to the principles of the study. Achieving the goal of developing a theoretical framework for explaining student success for this student group (which should have applicability to similar student groups and relevance to the broader development of student success theory) should not supersede the primary goal of exploring and understanding the phenomenon.

CHAPTER 3: RESULTS

Findings: Interviews

The Participants

Demographic Profile

The 36 women who participated in the study represented considerable diversity in terms of standard demographic variables (see Appendix D). They ranged in age from 21 to 51, with an average age of 36. Five of the women were First Nations; no other visible minority women participated. Four women reported having a disability; others may have lived with invisible disabilities that they did not mention.

Only 2 of the women had never been married. One was separated, 14 divorced, and 19 married or living with a partner. Six reported having no children, 13 were single parents with children at home, 14 were married and had children at home, and 3 had adult children, not living with them.

Educational histories varied significantly. One subgroup was drawn from a program requiring a bachelor's degree for admission; these 4 women all had prior degrees in human-service-related disciplines (criminology, psychology, therapeutic recreation, and education). Only one of the remaining 32 women had a degree: a 3-year pharmacy assistant degree from Hungary. Six had previous post-secondary credentials at the certificate/diploma level. A significant number of the remaining women did not graduate from high school: 2 were missing a few credits, and 12 dropped out before reaching twelfth grade (5 of these, prior to Grade 10).

The work histories of the women were equally diverse. Fourteen had experience in human service occupations prior to entering programs.Twelve had experience in a variety of jobs, generally unskilled or semiskilled and often short term. The work histories of the remaining women were more focussed, primarily in unskilled or semiskilled fields (such as clerical, housecleaning, hairdressing, waitressing) although 2 participants had achieved senior office management positions.

These women were paying for their education and supporting themselves (and their families in many cases) through a variety of means. Typically, multiple approaches to financial survival were in place. Sixteen women had student loans, which in some cases were likely to total in the $20,000 to $30,000 range. Of these, 8 were formerly income assistance recipients who were affected by recent changes in income assistance provisions requiring students in post-secondary programs to go off "welfare" and rely wholly on student loans. Eighteen women mentioned part-time jobs; 5 studied part-time while working full-time. Six mentioned financial supports from spouses, and 3 identified supports from parents. In addition, individual women were receiving funding support from other sources, including Employment Insurance, Vocational Rehabilitation Services (for people with disabilities), a band or tribal council, an insurance settlement, and maintenance from a divorce settlement. Two of the women interviewed, students in a special Ministry for Children and Families (MC&F) training program, were fully sponsored with all costs covered and receiving full salary.

Subgroup Comparisons

The participating women could be viewed as representing three distinct groups: those from the ongoing human service programs, encouraged to participate by instructors; those encouraged by instructors but drawn from the special MC&F program identified above; and finally, those from the ongoing programs who volunteered in response to the mailing sent to all female students. Twenty-three students were in the first group, 4 in the second, and 9 in the third.

The 4 women attending the special MC&F program differed from the other participants in several ways. The key differences included educational background, because a previous degree was a requirement for entry; work history, because only senior Ministry staff and experienced human service practitioners were accepted into the program; and financial status, since half (2) of the women interviewed were sponsored and on full salary. Other differences, which may have been incidental, included age, with average age at 34 slightly younger than the other groups, and family status, with a much lower proportion having dependent children (1 of 4) than in either other group.

The comparison between the remaining two groups was also interesting. Age was fairly similar. Five First Nations students were part of the first group; none responded to the mailing. The group encouraged by instructors included students from all four ongoing programs; students from only two programs volunteered in response to the mail-out. This could be attributable to timing: At the time of the mailing, students from the two programs from which there were no responses were just entering a block practicum.

The majority of students in the instructor-encouraged group was divorced (14 of 23), while none of the self-identified group was divorced (1 was single, 1 separated and 7 married). Over half of the first group were single parents, while only 1 in 9 of the other group was a single parent. These striking differences in marital and parental status may reflect instructor assumptions that the single parent student would have a particularly relevant story to tell, single parents' reluctance to volunteer for anything that further impinges upon their already stretched schedules, or a combination of factors. Four of 23 students (17%) in the first group were attending part-time while 8 of 9 students (89%) in the self-identified group were attending part-time. This may have reflected the difficulty that full-time students experience in finding time for extra activities like interviews, or the fact that married students (the majority of the second group) appear more likely to attend part-time. The reliance on student loans by students in the first group (15 of 23) was also not matched in the second group of whom none had loans. This related to the largely part-time status of the self-identified group and could also be related to marital status and lower reliance on loans by women able to count on a second income. Finally, the large number of students in the first group (12 of 23) that had dropped out of high school was not matched in the second group, wherein 7 of 9 had graduated from high school and the remaining 2 had completed twelfth grade with their age peers, missing only a few credits.

Organization of the Interview Data

The following sections of the paper will outline findings in three broad content areas, paralleling the interview topics: students' goals; supports and barriers to attaining those goals; and finally, what (for them) student success is and is not. The presentation of findings will outline general trends and apparent patterns in the views expressed, but will also highlight unique views. Quotations from the participants will be used generously. Names of both participants and the family members they mentioned have been changed, and alternate names assigned using a random procedure. Each quotation in any grouping is from a different participant. Because the focus is on this group of women as the case under study, and not on individual or subunit patterns, quotations are not assigned to specific participants.

Three primary themes that emerged in each content area will be traced throughout the presentation of findings, reflecting their predominance in the data: age and history, family, and finances. These themes are briefly introduced at this point, then explored further in later sections of this chapter.

Age and History

Because the defining feature of this student group is their age, and age is directly related to history (the more years lived, the more history accumulated), it is not surprising that age and history were among the most frequently mentioned topics. What may be surprising is that age is as often viewed positively (because of the resultant clarity of goals, maturity, experience and knowledge) as negatively (e.g., the feeling of not belonging, health and energy issues, pressure to complete quickly and get into the work force). Similarly, a participant's history was as likely to be a positive as a negative influence. The self-esteem issues associated, for some women, with negative personal or educational histories were balanced by the determination and drive associated with these same negative past events. For example, for some women experiences of working in impersonal, "no brain" jobs assisted in crystallizing alternative education and career goals and inculcating a resolution to succeed in accomplishing these goals.

Family

Family, in whatever configuration happened to be in place for each individual, significantly impacted these students. References to children predominated, although references to spouses/partners and parents/extended families were also frequent. Families influenced goals (e.g., interest in a discipline because of family experiences, or wanting to be a role model for children), acted as both support and barrier to meeting these goals, and figured in individual definitions of success.

Finances

As indicated above, far more of the instructor-encouraged sample than of subsequent groups were reliant on student loans, and finances figured most prominently for this initial group. Factors decreasing the impact of finances on the two other groups included the fact that the MC&F program provided full sponsorship for Ministry employees, and the possibility that individuals experiencing financial stress were less likely to respond to the mailed invitation to participate (most of those who did respond were reducing financial stress through attending part-time).

Nevertheless, participant comments indicated that finances were sufficiently important to be considered a major theme. The intensity of concern expressed by those experiencing financial pressures balanced the fact that not all students were in this situation.

Student Goals

At a superficial level, the goals of mature female human service students are not dissimilar to those of other post-secondary students: that is, to complete a credential and gain employment in the intended field of study. These overt goals will be explored below, followed by a discussion of the subtle characteristics of these goals, and their apparent origins. The focus of this section of the paper is on the goals that initially brought students into human service programs; however, these initial goals cannot be separated from the evolving goals of women once engaged in programs.

Overt Goals

Students entering human service training, like most students, are intent upon employment in their field of choice. The women interviewed had entered their programs of study clear on their career goals, with already-developed commitment to working with people.

> I finally decided to go for it because everything kept leading me back [to this type of work].
>
> I really want to work with youth. There are so many out there that need the support and the guidance along the way. So I know that's where I want to be.
>
> I want to work with young offenders and I knew that when I came into the program. The only thing maybe that's changed is that it's more fine tuned.
>
> I would like to work with kids who are having problems, and parents--family work. There's a terrific need for it.
>
> I love kids. I love them, they're just so interesting. . . . In ECE (Early Childhood Education) the wages aren't that good, but it's very rewarding so that's O.K. for me.
>
> I wanted to work with kids because I always thought that kids were honest and they told the truth, like it was.
>
> I think, if anybody had ever said to me, two years ago, what about working with people with disabilities, it had never crossed my mind. And until I actually tried it, then I thought, yeah, this would be O.K. People actually get paid for doing this?

These students entered programs with enough work, volunteer, or personal experience to be sure of the type of work they wanted to do, although the specific details of their ideal job within the human service field were shaped over the course of class and practicum experiences. Their personal career planning derived its basis from firsthand, experiential knowledge rather than from the theoretical conjecture that forms the basis of much of the formal or informal career exploration available to young students. Each chose a specific career/education path because they had experienced working with people, enjoyed it and were convinced it was what they wanted, not because it sounded interesting and they thought they might like it. Younger students might also enter human services with firm, experience-based goals, perhaps derived from family circumstances, part-time jobs, or work placements during high school. However, it is logical that the age of the students interviewed is a factor in explaining their consistent experiential base and solid commitment to fields of study, and likely that a cohort of recent high school graduates would be less uniform in this regard.

A second overt goal of the women interviewed was completion of their program of studies and receipt of the related credential. Women who are motivated to participate in post-secondary education by a variety of other factors are also interested in the goal of obtaining a credential, because of the additional credibility and job opportunities they perceive this will provide. Although completion of the credential was important for all the women interviewed, only one indicated that the credential was "all she was after" (and even for her, comments suggested otherwise). As will be discussed below, the credential represented a significant goal for these women, but was only one of many indicators of success.

> Every job I ever did, I excelled at, you know. But I never had the education to put me to the next level up. In the last 10 years, I've noticed that the education had to be there. It didn't matter if you had 15 years experience. You get to the point where you have to have the degree or diploma.

> With education, once you start, and if you want to, you know, hope to use this to gain employment, half a diploma, half a degree, that means nothing. I mean it's not meaningless 'cause you do have the training out of the half of the courses that you did, but, on a resume, it means nothing.

> I wanted to get the degree. I need it for credibility, I need the education.

The design of the programs in which these women were participating is significant. The program model provides students with certification after the first and second year (at certificate and diploma levels, respectively) as well as at the degree level. This clearly makes completion of a credential a more realistic goal for all students, no matter what their personal circumstances, and provides more immediate reinforcement for completing steps along the path to a degree for those who choose to continue. An interesting phenomenon with the laddered approach taken by this institution is that students often enter hoping to achieve the initial credential (the certificate), then shift their goals upward to diploma or degree. This goal elevation appears to come about because students experience unexpected success and enjoy their education, their passion for their field of study intensifies, and they recognize the hierarchies that exist within the field and the importance attached to degrees.

> I want the degree because, it's not that I have this ideal that I'm going to go out and change the system. But you can't go out and implement your own theories unless you've got at least a degree. I cannot go out with just a certificate and say, "Look, I've got a good idea for a program for people with mental illness," and implement that, because nobody's going to listen.
>
> I've always liked learning, and I thought I may as well keep going. It's only another 2 years. I'll have my degree and it will be like a dream fulfilled, I guess, that I've wanted to do for some time.
>
> I planned on the diploma. I just wanted to stick the diploma with my psych. nursing and I thought that would be good. But once I got into it, I mean, I'd like to get my masters!
>
> Now I want to be you, sitting there asking people questions for my doctorate!

For those women with limited confidence in their own academic ability or uncertainty regarding their ability to cope (financially or in any other way) with post-secondary education, the laddered program delivery model was sometimes a key factor in encouraging initial enrollment as well as continuation from one credential to the next.

> So the certificate was just a convenient way to get started, that little chunk.
>
> And it was easier for me to take those bites, a year at a time, and the diploma was just another year on top of the social service worker certificate. So I knew that if I had to stop, I wouldn't lose anything.

Nature of the Goals

Although, on the surface, the reasons given by mature female human service students for entering training programs may not appear to differ from those of most traditional aged students, the subtle characteristics of these goals are distinctive.

The students interviewed made it clear that their motivation for entering a post-secondary program, although influenced by family members, was predominantly intrinsic. These women made independent decisions to enter human service training programs and continued making independent decisions to remain in these programs. Some identified this as a factor unique to mature students, which distinguished them from their younger classmates.

> Some of the younger students are just "not there."

> Mature students typically make better grades than young students because they are focussed, this is what they want to do. They're not here because of mom and dad, and it's true, it's their money, they've probably taken a student loan out, they're going to make it. They're going to do everything in their power to pass, right?

Recognizing the intrinsic nature of mature students' motivation to enter programs is critical in understanding their determination to complete these programs and in contemplating interventions to support them. Although students can be influenced by programmatic variables and availability of supports (as will be discussed below), it is their internal drive that motivated them to enter and acts as the primary force that keeps them going.

> This is what I wanted to do; this is mine.

> This is what I've always wanted to do. So for myself, I feel like I'm doing something for myself at the same time as I'm achieving more security for myself and my family.

> What's kept me going is my will to be here. I'm my own biggest fan. . . . It just boils down to something I wanted to do. You've got to have a lot of discipline--well, you know that! You've got to want it, or else you don't do it.

> You really have to know in your heart that it's what you want to do. I think that's the biggest thing, to become emotionally tied to a goal of some sort.

> Now I'm not here for teachers, or my family, or the system; now I'm totally here for me . . . just seeing it's really me who's reaping the benefits of being in school, and it's me who lives this life, it's me who's going to achieve goals.

A second notable characteristic of the goals driving these women is their intensity. The terms some women used to describe their goals in pursuing human service education suggested determination and sometimes passion. They were where they wanted to be and were firmly committed to the paths they had chosen.

> It's like getting paid to go and play and do something you love. I absolutely love it, it's so fun and it's so rewarding. It's so neat, it's a great field.

> Nothing will discourage me. I will continue to go on, and it may be really difficult at times, and I might need more support than I normally would. But I will never be discouraged, in the end.

> I think that, um, yeah, I'm on a mission. It's like being on a mission--exactly.

Origin of the Goals

What is emerging is a picture of women keenly committed to working in the human service field and to completing their educational program. The following discussion will attempt to shed light on the basis for this level of commitment, examining the issue according to the three themes identified earlier: age and history, family, and finances. As the student comments will illustrate, however, this breakdown is somewhat arbitrary and the actual sources of motivation tend to be multifaceted and interrelated.

Age and history. Some of the women interviewed attributed their clarity of purpose to their age, maturity, and experiences, and this explanation rings true in general for the participants. Several had aspired to post-secondary education for many years or been working toward that goal gradually over a long period (through upgrading basic skills, getting personal issues under control, and so forth). This long-standing interest, once acted upon, leaves these women particularly determined to realize their vision.

> Just as I got older and older, and I kept thinking, "No I can't go back now." That's why when I finally said, "I'm going to do it," I did it.
>
> So this was sort of always something in the back of my mind, "One day I will go and I will do it."
>
> For myself it was really going back and sticking to a plan I had. You know, wishing that I had done it, but I hadn't.
>
> I've worked years to get here, literally years.

The combination of age and experience promotes this determination: age because of the pressure women feel to "get on with" reaching their goal, once it becomes clear to them, and the determination to make the most of what could be their last educational opportunity; experience because the sum total of their history has defined their intended path and made deviation personally unacceptable.

> I think it was a turning 40 kind of thing. I think that things fall into place.
>
> I'll be entering the work force when my kids go to school, and I'll be 40 years old! I've got to have something behind me.
>
> I thought, this is my one and only chance. If I don't do it now, I probably won't do it.
>
> Living life and having my kids, I really knew what I wanted to do.
>
> Setting goals comes from that same self-awareness, recognizing who you are, where you're going. I attribute all my experiences, through a very challenging time, as making it that much easier to follow through. The more challenges, the more experiences, be it negative or positive, the better off we are.

Along with the accumulation of years, women have accumulated history that contributes to formulating and solidifying their goals. For those women who had negative prior educational experiences, receiving a credential has become an essential part of "proving themselves" and demonstrating that they are not educational failures.

> I was just treated as if I was "slower than", and I just basically went through life thinking that that was true.
>
> I don't know if I'm proving it to myself, or what, but I always felt that if I'm really proud of something then I know I've done the best I could. So I think that was in my mind to come back here. Actually I was really surprised they accepted me because I didn't think that I was good enough to come back to school.

For the women who had worked in dissatisfying jobs for many years, there is unbending resolution to leave these experiences behind and move on to more rewarding work in the human services. For some, the passionate commitment to human services is matched by a passionate aversion, developed over years, to a previous career. A large number of quotes are included here to illustrate the diversity of scenarios, but the consistency of pattern.

> I knew that I was really unhappy, and that I couldn't do psychiatry. . . . I felt so disillusioned from where I'd come from. . . . I couldn't articulate it until I got into CYC [Child and Youth Care], but I found it [psychiatry] too pathologically oriented.
>
> I learned that it didn't matter, monetary worth is not the end all, be all. That was not what I wanted out of life. Yes, I could get through and keep surviving. But I also wanted some inner rewards which I wasn't getting from that.
>
> Before that I was actually working in the banking business, and at one time I enjoyed it because of the personal contact with people. And as it became more and more impersonal and data-based, I really lost my interest. The connection wasn't there. I wanted something more, where I thought I could get connected with people.
>
> Hairdressing is no brain, it really is. I mean, the people part of it I loved, because they come in, sit down, tell all their secrets. I loved that part of it, loved listening to people's life stories. But it wasn't challenging, or stimulating, or fulfilling. It was just, stand there and have a conversation.

> I was the personnel manager and I didn't really agree with the policies. And it was either go along with it and not be true to myself or get out and find what I wanted to do. . . . I thought, well, one of these days I'm just going to blow up and make a fool of myself. It's time to start looking around. . . . If I can find a job working with youth, I can be really happy. I wasn't before; it was very stressful, very long hours. I was there 6 or 7 days a week. . . . I'm just so happy, I'm so relaxed compared to what I used to be like. Like I feel I'm going to live past 50, except at exam time!

> I got hired as an FAW [Financial Assistance Worker], my dream job! Interesting! It wasn't exactly what I wanted it to be; there wasn't enough counselling or enough one-on-one for me to really enjoy it.

> Long-term care work was pretty depressing, and shift work. And talk about lazy people, where I worked, handling the residents and getting them to bed at 7:00 so they could sit and smoke till 9:00. Just awful. I couldn't handle it.

> I was tired of the business world. Like, I thought, it's such a dog-eat-dog world, everyone trying to climb up the corporate ladder, and it just seemed to me a futile experience. . . . I had got up there and pretty well had a nice position but it wasn't fulfilling any more.

> I'd quit my job as a medical secretary because I knew I didn't want to do that, too much pressure, and I quit my job at the biological centre because I'd gone as far as I could with ground fish. I can't even eat it any more!

> I worked as a secretary for 16 years, and I just hated it. . . . Going to work was so humdrum. It eats away at you when you don't like what you're doing, and so I would change secretarial jobs thinking that would change, would be more exciting, but it never was.

There were also unique aspects of students' histories that helped to strengthen their resolve to pursue human service training and employment. Three women entered a First Nations focussed program because of its cultural components; a number of women linked the decision to enter the social services field to personal experiences dealing with the income assistance system; some knew that the programs involve an element of self-reflection and focus on personal growth that suited their individual circumstances.

> I never really knew anything of my own culture. And, I don't know why, when I got here, it seemed to be like a calling or something, and I wanted to learn it even though it was very different, probably, from my grandmother's culture.

> One other goal was getting in touch with my culture, through the generative curriculum. I really took advantage of that.
>
> So what brought me to this program specifically was that I found that the social assistance, income assistance, system needs a lot more people who are educated in a more holistic approach.
>
> I thought, somebody's got to be in there [the income assistance system] who's treating people right.
>
> A lot of dysfunction in my history. And probably I'm here because I believe that all of that has to have some value to it. It wasn't just good for nothing.
>
> I set personal goals that weren't related to education and career. They were more things around doing some personal work, around my divorce. Personal things around separation and divorce, and who I am.

Family. For many women, and in many varied ways, the goal of entering human services education and employment relates to their families. Getting an education in order to improve life for both self and family is a common theme, particularly for single parent students. Knowing that they are undertaking an education on behalf of their children balances, to some extent, the guilt that student parents experience related to the time and money pressures accompanying the student role.

> Also part of it was the fact that I had become a single parent with two children, ages three and four. So I knew that in order to be able to give them the care that I wanted while I was working, I had to reassess where I wanted to be and what kind of work I wanted to be in. I wanted to do something that would benefit me as well as my family.
>
> Jeff goes, "Kyle's mom is home all the time." "Yeah, well, they're supported by the government, Jeffrey. I don't want to be supported by the government. I want us to have more. I want us to be able to spend time together as a family, and go on vacation, maybe buy a house. I want those things and we can't have those things if we're supported by the government."
>
> It's still, for me, the guilt that I'm not providing for them. But I also remember that that's why I'm doing this.

> But what I always have to reassure myself, is that even if it doesn't look like I'm doing something for the two of us, as a family, I am. It's for the better of us.
>
> At our graduation we're only allowed to bring two people so I invited my son and my mom. And I'm in the middle, sort of thing. And I think it's important that he comes because he needs to see that there is a completion, that there's an end, and that what I've been doing has been worth it.

Acting as a role model for their children is an important goal for women students. They see their efforts to take control of their lives, embark on an education and career, and do well as students as providing a positive and important example for their children, whether young or adult.

> And I guess that's the other thing that motivated me was to provide a role model for my children, to show them that it doesn't matter at what age you decide to make changes in our life. You can do it. . . . And for them to see that, right now, they might not care if they go to college . . . but somewhere down the line they're going to remember the hard work and commitment that I put in and be able to follow that example.
>
> I think they're proud. They look at my report card, like I used to look at theirs, and I think they're pretty proud of what I'm doing. And I hope it's an example to them because they didn't go on after graduating. . . . And I just figured, well, I'm not going to push them because they should go on when they're ready, like I did. And I think they will some day.
>
> I'm hoping that what they're going to see is a role model. We took our time, and we crashed, and now we're putting our life back together. And this is what I want in my life whether Dad's here or not. . . . I wanted to use the time now to model for the girls, so that when they leave and my mothering days are over, I would still have a role, that I'm not just a mother and a wife.
>
> It's neat that it sets a pace for the kids, like yes, you can, there's no excuse to go and jerk around. Go and do what you have to do and do the best you can.

Equally important for some students is providing a model for extended family members and (particularly for those women from families with limited education) achieving an educational milestone on behalf of the whole family.

> I'm the first one to go this far. I tease [my parents] every once in a while when I get stressed out: "I've taken the genes as far as I can take them; I can't take them any more!"

> I think about my family. That's my most important reason to continue because I'm a role model. I'm the only one in my family to go to college--that keeps me going. So I gotta finish. . . . It isn't about letting myself down, it's about letting my family down. We were a family of five kids and not one of us graduated from high school.

Finally, for some women the specific career goal relates to a family member: for example, a desire to work with people with developmental disabilities stemming from having a child with Down's syndrome, or wanting to advocate for income assistance clients because of familial experience.

> My passion lies with mental illness because it's so, so close. . . . I have a brother who has schizophrenia. Watching him struggle and seeing the holes again; they're huge out there, in mental illness.

Finances. Although it is clearly not lucrative salaries that attract women to human service fields, financial considerations do enter into the thinking and planning of potential students, especially single parents who must consider the financial well-being of their children. Many of the students interviewed were relying on student loans and consider this a valid investment. For students who have had to rely on income assistance, becoming self-sufficient is a powerful motivator.

> Older students have a pretty good sense of the reality of the jobs they'll be going into, and the money they'll earn.

> I look at it this way: I thought, O.K., I'm going to owe all this money. But I can sit on welfare for the next ten years and not have anything. O.K. so I'm going to owe, I think it's $16,000 or $17,000, but some people pay that for a car. I'm paying $16,000 to $17,000 to get a job--to get a job and then be able to pay for a car. That's the way I look at it; I don't look at it like a debt. I look at it as an opportunity.

> My kids were now teenagers and I needed money to survive. I knew I could keep walking into the welfare office, but I didn't want to be on that system too long.

> I was on social assistance at the time. Actually lots of my life; second generation. So, poverty was a big motivator. . . . I could not go back on welfare. I would never do that to myself again. I was on it so long and I hate it so much.

Individual Stories

Fragmenting and isolating the details of women's motivation for entering human service training does not do justice to their individual stories or reflect the interconnectedness of the factors that influenced each person to take that major life step. Following are brief descriptions of the backgrounds of some of the women participating in the study and what drove them to enter human service programs. Many other women interviewed have similarly complex and fascinating stories; these stories were selected as a representation of the array of diverse situations from which these participants come.

Jana. Jana was born and educated in Hungary. Her formal training was in pharmacy and, after coming to Canada, she ended up working in a pharmacy in downtown Vancouver, or as she puts it, "pushing pills on Denman Street". She traveled extensively as a young single person. On one of her trips she met a Syrian man whom she married. Jana lived in Syria with her husband, quite happily through the births of her first two children. Her third child was born with Down's syndrome, and Jana soon realized that attitudes toward disability in Syria would condemn her third child, Karl, to institutionalization and an extremely limited future. Not willing to accept that, she moved back to B.C. with Karl when he was 3. After a number of unsuccessful efforts to secure meaningful employment, including starting a small business that went bankrupt and securing a student loan to take private training as a medical office assistant, Jana decided that she wanted to work in the disability field. Her goal is to operate a group home in which her son can live. She is on income assistance and attending school part-time, but is committed to working and becoming independent. One reason for this determination is that she would like to sponsor her mother to move to Canada, but cannot as long as she is on income assistance. Her husband remains in Syria, with her two older children, and continues to block her involvement with the family so long as Karl is a part of her life. This is one of her primary stresses, but has not weakened her resolve. At 51, Jana is completing a 1-year certificate.

Cecile. Not long after graduating from high school, Cecile was a successful hair stylist and instructor of other hairdressers. She was earning a good income and, although not completely satisfied with her work, would likely have continued had she not become pregnant. As a single parent, she began to question where she wanted to be in life. A short time later, she married, had a second child, and lived through the collapse of a brief and painful marriage. Fleeing an abusive situation, she decided to enter a career more suited to her and enrolled in a human service certificate program. After completing the certificate, and realizing how successful she could be and how relevant the content was to her both personally and professionally, Cecile committed to completing first a diploma, then a degree. She recognizes that she needs a job with adequate income to support her family, but also wants a job consistent with her values. Her negative experiences dealing with the income assistance system have firmed her resolve to complete an education and become independent. Cecile is 29 and close to completing a degree program.

Christa. Christa depicts her background, both in her birth family and in subsequent relationships, as involving "a lot of dysfunction." She came from an alcoholic and abusive family, did not complete high school, and spent her early adult years in a string of low-skill jobs and unsatisfactory relationships (all short term). She experienced ongoing substance abuse problems and, like her birth family, often relied on income assistance for her survival and that of her three children. Once she "turned herself around" and dealt with the substance abuse issues, she determined to make something of her years of dependency on the social service system. She felt that all she had learned through that experience had to pay off in some fashion and that the system needed more people working in it who could relate compassionately to clients. Her initial goal of working for the Ministry of Human Resources has shifted, but she is still determined to work in a social service agency. Christa is 37 and completing a 1-year certificate program.

Lindsay. Lindsay is 28 and has a bachelor's of education degree from another province. She moved to B.C. to look for a teaching job and searched widely but in vain. In order to find employment in a teaching-related area, she completed the requirements for licensure as an Early Childhood Educator and has been working for several years in a nonprofit agency child care program. Although she enjoys this work, she has been feeling the need for a change. She enrolled in a Bachelor of Social Work program, looking for training that, along with her previous education, will prepare her for work with the Ministry of Children and Families or somewhere in the child welfare field. She would like a position that recognizes (in salary and societal value) her considerable education and experience and sees this degree opportunity as potentially opening doors to professional employment in a way that her teaching degree does not.

Factors Affecting Achievement of Goals

As mentioned earlier, these women lead complex lives and the factors that lend support to them in achieving their educational goals are often also the stressors that get in the way of success. The predominant themes of age and history, family, and finances will again be used to organize the data.

Age and History

Mature students have lived longer, bring more "baggage," have more commitments and responsibilities than typical young students. Many of the impacts of age are indirect. Family responsibilities are more likely an issue than for younger students, as will be discussed below. History shapes development of personal traits, with both positive and negative outcomes. Influences are interrelated and present an overall pattern likely different from that for younger classmates.

One factor clearly attributable to age, and with more consistently negative effects than for other factors, is energy level. Many women mention that they do not have the energy they had when they were younger.

> I find that if I don't have enough sleep, I just get overwhelmed. I felt like that more in the first year, and I think it was probably due to getting back into the routine. I didn't realize at the time what a big load I'd taken on.
>
> I think I've worked my own body clock to winding down in the evening. So especially Thursdays, I'm not there from 4:00 to 7:30--I'm barely in class. . . . I'm in bed by 10:00, 10:30 now.
>
> At this age, I don't have the physical energy. At that time, too, when I went back to school, unbeknownst to me, I was in early menopause, so I had short-term memory stuff, emotional stuff.
>
> I come home, I can't get off the couch, and Travis goes "Mom, Mom" and I just don't have the energy to listen.
>
> I'd be dead tired by the time I got home, like I could hardly move. Unreal.
>
> I don't mind [evening classes] as long as it's not every evening. . . . I find that the older I get, though, the less appealing that is.
>
> I find with young people they have so much energy and so much drive, they're here and they're there and they're doing this and they want to get it done, and I'm like "Hold on!" . . . I need to evaluate before I can do something, and their minds go 100 miles an hour, which is cool. I love their energy and how they do stuff. But when I'm

> involved, they want me to speed up and I want them to slow down, and we need to find a balance somewhere.

Age and energy also influence learning, although surprisingly few negative comments were made regarding effects of age on learning. What did come through for some women--possibly associated with age and the need to "get on with it," plus competing time pressures--was some impatience with spending time on activities not seen as generating worthwhile learning rewards. Time is so scarce for these students that any unnecessary encroachment is resented.

> I mean, sometimes it probably takes a lot longer [than for younger students] doing homework.

> I don't have the photographic memory that I used to have. . . . Sometimes now I'll study for the exam and think I really know this, and two weeks later, "What was that?" So that's a little different.

> I was really annoyed at this age. I thought, I have spent hours and hours and I want to learn this stuff. And this exam thing is totally useless. You don't remember anything from an exam. . . . Again, it was all part of that: Yeah, if you've got the time and you don't have kids and a life and a family, yeah, you can cram it all into this concentrated period.

A primary challenge indirectly related to age is the likelihood that students are juggling myriad other responsibilities with the demands of their studies. Many of these demands relate to family and will be discussed below. However, another category of "juggling" warrants discussion as it may be unique to mature students. Many of the women interviewed mentioned continuing their volunteer commitments despite their busy schedules as students: Anna coaches a girls' softball team and has signed up to volunteer with the John Howard Society; Gwen and her husband provide volunteer recreational leadership for First Nations teens in their neighbourhood; Frances volunteers at Victims' Services every weekend; Thelma spends one morning a week with an adult with intellectual disabilities, whom she met on an earlier practicum; and last spring Alma volunteered at the school where she had done a practicum, to maintain continuity after the practicum ended. For some women, these commitments predated becoming students; for others the specific commitments are recent but the pattern of volunteering is long-standing. Contributing to society through volunteer work appears to be an integral and ingrained part of who these women are, important enough to maintain even when educational pressures make time precious.

Another challenge associated with age that was mentioned by many women is the difficult reentry adjustment associated with returning to school after an often long absence. Some women worried about being out of place due to their age, and initially felt awkward and intimidated in the college environment.

> I think it's easier for younger students because they've come right out of high school. They're coming here, they've got friends here, they've got family. You know what I mean? But for me, it was really hard.
>
> I think that's why I was so scared when I came back. I realized that I'd been gone for so long.
>
> That first day, I think I was late for every class. I felt really intimidated, so intimidated. For one thing, I was over-dressed. I had no idea, and I was so used to getting dressed up for work. So I came in all dressed up, and I think someone commented, when I walked into one class, they thought I was the teacher.
>
> I'm not one to drum up a conversation. I tried it a couple of times and they looked at me like, "The old lady is hitting on me!"
>
> When I started up here, I didn't know where I was going or what I was doing. Like, I don't know the ropes. I guess a lot of the young ones have been up here off and on. There's lots still I don't know.
>
> At first [being older] upset me, kind of bothered me. Because I thought they'll be thinking, "What's she doing here?"

Many women enter programs with additional stress due to their lack of confidence.

> For me, like I wasn't independent. I was very dependent on my ex-husband. I felt like I was flailing in the water and trying to keep above ground. . . . Like, I know, I had low self-esteem, I was really not sure I was going to be able to do this. . . . So when I came here, I was really lost. I had been in this abusive marriage and I had absolutely no confidence. I just thought, God, if I could just do this. I had this hope, but I didn't have the courage or the confidence to think I could do this. So when I came here, I was really lost.

> I'm very nervous before I do something and I have a lot of fear around it.

> I'm almost finished the year and I still feel like a fraud. I'm getting good grades, but I wonder if I'll retain anything?

> I always feel that I don't really know what I'm doing and they're going to find out. But it's scary; for me, I'm in my last year and still having these feelings, like, "Oh gee, do you ever feel O.K. about what you're doing?"

Even those students with prior degrees, in the MC&F sponsored program, experienced confidence issues on entering: Would they be able to do this after being out of school so long? Were they too old? Did this signal that they were "bad social workers," unable to perform competently?

For many of the others, a more direct link between the lack of confidence and school history is suggested, particularly for those who had not completed high school. Anyone going back to school after a considerable gap experiences doubts about their ability to cope and succeed; this is intensified for women whose previous educational history was unsuccessful.

> I never had positive experiences in school, so I dropped out in Grade 8. . . . Went back, did my upgrading, filled with all these tapes, you know: "You can't do this, you're a fake," and all this, and came up with straight *A+'s* and that was the turning point. I thought, hmm! This wasn't true!

> I wanted to push myself to get good grades--I was never academic in high school.

> Grades have been important because I dropped out in Grade 8 and I've always had this belief that I was stupid and all those kind of things.

In the long run, as illustrated by the quotations above, this self-doubt pushes women to perform, deliver, succeed. Once they overcome the initial confidence crisis, their determination to prove themselves almost ensures that they will complete programs and do well academically. Their age and histories create personal commitments to take full advantage of this "last chance," manage both academically and personally, and get good grades. Again, this is a source of stress that many younger students may not feel (at least with the same consistency or intensity), but also predisposes older students to meet academic expectations and complete programs.

> I'm not going to fail, I'm not going to. I'm going to give it my best shot.
>
> I can remember first semester getting a quiz back; it was out of 10 and I got 9.5 and the first thing I looked for was where did I lose the point. And as soon as we saw that (a friend was with me), we looked at each other because we'd just been talking about how we focus on what's not O.K. And that had been part of it--I did it again!
>
> I'm so hard on myself--what am I trying to do? What am I trying to prove? I should just be happy to get what I get, but I want an *A*!
>
> I also have this compulsive personality and I think that most people who go back at this age do. Because this is your last shot, last "kick at the can." So you tend to try to do the very best that you possibly can. It's not like when you're a younger person, and it doesn't matter, you've got lots of time. So you put a lot into it.

Many women spoke of the need to "let go" of some of these high expectations in response to the reality of their personal lives and to come to terms with not getting the top grades they would like. A common pattern appears to be entering college with little confidence, realizing that high grades are possible and striving for that, then gradually recognizing that high academic achievement cannot override all other aspects of life. Women appear to become more forgiving of themselves as they move through their programs, although high expectations remain both a motivator and a stressor throughout women's educational careers.

> I was raised that nothing was ever good enough. So I actually beat myself up the whole first year: Do the best, do the best, do the best.
>
> One personal goal is to not get so stressed out about it. And fourth year, I'm probably not getting as stressed out. But I think it's because of my background, I worry that I might miss something.
>
> Letting go is a very difficult thing for me, lowering my high expectations. And then, seeing that I have achieved everything that I wanted to achieve because I've still managed to do the best that I can, be the best that I can given the circumstances.
>
> I've learned quite a bit. [Grades] were the be all and end all when I first started. I was totally destroyed if I didn't get what I expected. Now I'm prioritizing, now I'm beginning to realize that it's not going to kill me if I don't get that *A* and I get a *B+* instead. Or, if I get a *B*, because, depending on what's going on for me as a whole, at work, at home, that definitely influences it.
>
> When I came back, I realized that I'd been gone for so long, maybe I couldn't get straight *A's* any more. Now, for me, I'm happy with myself because I'm doing the best that I can but not trying to be the best. . . . It's too hard, 'cause life is out there still! . . . It's for me to remember that I can't, I'm doing the best I can with the life that I'm living. That's just, getting through school, and maintaining pride in my work, and remembering that I'm not superhuman.

For some mature female students, age and high expectations also translate into a pressure to complete quickly and move into the workforce. Although this likely contributes to high completion rates, it also adds to the stress: Going part-time or at a slower pace is not seen as an option for some, because of age.

> If I could have gone at a slower pace, it would have taken away some of the stress and that would've been quite nice to be able to really focus. . . . I've pushed myself; I'm on a time frame because of my age. I feel that this is really pushing me to go as fast as I can.
>
> I mean, I'm 37, and that's also what weighs in taking the course part-time. It's like, when I'm done I'll be ready to retire!

Related to expectations is the issue of self-reliance. Although many women indicate that they have no difficulty seeking help if they need it ("There have been a few students who have consistently taken advantage of [the help that's available], and I've been one"), for many others, asking for help is difficult, either because of confidence issues or because of a desire to demonstrate an independent capability to solve their own problems. Some describe the dawning awareness, experienced over the course of their educational programs, that acknowledging a need for assistance is acceptable.

> And I guess, for me, because I've always been so independent and have basically done everything on my own without any help in any way, it's been really difficult for me to recognize that I can stop and go ask for help. I put all the weight on my shoulders of having to do things on my own and of being strong and a person capable of doing that.
>
> I would just struggle till I got it. I don't know why, I don't know why, I should go for help. I'm usually pretty shy around people when I first meet them.
>
> So I'm beginning to risk letting everybody know I'm not perfect, wonderful, et cetera. I'm like everybody else and I need help.
>
> I made a point of going to all the instructors and telling them about my experiences and what I wanted out of the program. I couldn't have done that before; I had the ability but I wasn't able to do that. . . . What I liked was that I recognized about myself after third year . . . how I do things and that when I'm feeling stressed and overwhelmed I isolate myself and don't get help. It's a catch-22 situation. So now I recognize that and see what's happening, and I'll try to phone people.
>
> Because I've been on my own since 16, I trust that I can take care of myself, you know, and get done what I need to get done, but I don't trust other people to do that for me or with me. So that's been a 4-year lesson!

Women view the experience gained with age as a positive factor. Entering human service fields with a background of parenting, for example, facilitates learning about children and about family relationships. Experience has familiarized mature women with the content of many courses, making mastery easier than it is for younger students, for whom the courses may be more abstract and theoretical.

> I do feel that as a mature student, I can relate a lot to what they're saying because, in my own mind, I just think of my own kids.
>
> My ECE marks haven't gone down because I don't find that hard. I think it's maturity and having kids and having worked with kids. . . . So I'd had a lot of experience in areas where some kids straight out of school were struggling with the concepts. Making it real when you haven't been through it is hard; it's very abstract.

Family

Family considerations were among the most frequently mentioned factors affecting women's achievement of their post-secondary goals. Although (as has been discussed) family factors constitute one of the major motivators for entering an educational program and families provide a significant source of support along the way, family issues are also a primary stressor for women.

Family-related factors cannot be viewed separately from other issues because they interrelate with all (positive and negative) factors impacting women students. One of the primary effects of family relationships on students' capacity to succeed is in terms of the time demands such relationships entail, and it is here that the interrelatedness of family with other factors becomes clear. The predominant struggle these women face is in juggling their various commitments, including academic requirements. Families generally top the list of demands and therefore are major contributors to the stress and pressure women encounter.

> I didn't know if I'd make it through the first semester. My grades are fine, but it was very, very stressful. 'Cause I'm the one, I have to remember all those appointments, play therapy, psychiatrists, pediatricians. Like, the Mom does that. . . . So I pretty well do two jobs. I do laundry before I get going in the morning, and I come home and do housework and dishes, and lots of time on Saturdays, instead of studying, I do housework.

> I feel like I'm a part-time everything: part-time student, part-time worker, part-time Mom, and I wonder just who's getting shortchanged.
>
> I think that the system is set up for 18-, 19-, 20-, 21-year olds who are living at home or with roommates who don't care if the house is a mess, who have no other responsibilities. I remember going to university at that age and being able to keep up and head off to the bar on Friday nights, and that sort of thing. And being able to have a life. . . . It asks way too much for me, way too much when you're a mature student. Like, you have a life, you have demands, you have children at home, you have a partner that you must attend to. Like, that's your life and there's no such thing as balance.
>
> When you have a family, your life just takes you. You can set aside two hours to do something and don't necessarily get it.
>
> The problem is the juggling act and the reality that we still have more than one hat to wear. You'll always be Mom, first and foremost, so that's been a struggle.
>
> I think the most difficult part is balancing study time and family time and commitments in the community and then still being able to have that time for yourself and that's the part that goes down the tubes as you become more and more overloaded, right? So I found that was the most difficult. . . . Being able to juggle and balance and take care of everything; it's a totally different perspective. It's difficult because of the high expectations you set as a student and a Mom.

One interesting admission made by only a few women was that school is sometimes the least stressful component of their juggling act. In fact, school is sometimes the positive part of the equation and the element that compensates for some of the other stresses.

> It's home and school; school's not usually too bad. It's mostly, like, living at home with teenagers. That's the biggest part of it [the stress].
>
> I haven't thought about dropping out of school--just moving out of home!
>
> This has really been my escape I think.

One of the issues related to family stress that became apparent through the interviews is that, while becoming a student creates additional challenges for families, the typical stressors do not go away. Dealing with divorce, difficult teenagers, or family health crises, for example, are pressures mature women commonly encounter. These typical family stresses also emerged as significant for the women interviewed, compounding the pressure of academic demands and the financial stresses associated with student life.

> What's happening within the family and the whole system around you ultimately impacts what decisions I might make. Because I felt that I had put my finances in order, knowing that I'd be able to go back and make it through without being totally stressed. But at Christmas time my boys were in a serious car accident and my car was totalled.
>
> My son has Attention Deficit Hyperactivity Disorder, and my younger son has just been diagnosed too. So we have a really hard time in the mornings, and I don't get a lot of study time.
>
> This year I'm realizing that I have teenagers too. There are times when I just have to let things go and say, I can't do it all.
>
> There's a lot of stress, at work, my own kids, my daughter being in her adolescent years. She just turned 14 and it's been difficult.
>
> It was really hard to reorganize my whole life because I wasn't in a relationship with anyone, had two children, and was dealing with a terrible divorce. . . . [My son] had two more open heart surgeries, he had three altogether. So during the time I was in school, he was in Vancouver part of the time.

Some of the family stresses typically faced by women involve extended family, and again these continue for mature students and may be more intense for human service students because of the personal traits they seem to share.

> I tend to have a lot of those helper skills. When I'm around other people and they need something, then I'll put their needs before mine.

References to extended family stresses came most frequently from First Nations women, which makes sense given the focus on extended family within that culture. Although family is viewed as a strong supporter, for some First Nations women the expectation of extensive connection to extended family can create a barrier to academic achievement, given other constraints on their time.

> So you know, sometimes our culture gets in the way of school, but it's important to us. . . . I've had more of my family move here and my husband's family moved here. They can't figure out sometimes when we have to say, "I'm sorry, school is a priority." We can't just up and leave to do these family things.

Another key and common issue regarding families is the guilt students feel in terms of their inability to spend sufficient time with family members. This is a generic issue, as one student pointed out: "The guilt that mothers feel! The kids are saying, 'Let's go and play', and the mother's saying, 'No we have to do this.' And I'm sure all parents deal with this whether it's work or school or volunteer positions. " However, the difference for these student mothers is that they are often fulfilling all of these roles: employee, student, and volunteer, as well as parent. Their capacity to spend time with family is multiply eroded, leaving them with a huge potential for guilt.

> And [my son], there's part of him that would just like me to work and be a regular Mom. . . . He doesn't go to bed, usually until I go to bed. And so I'm not there. I'm at the computer, I'm working and I'm reading and I'm studying. So he'd like me to just not do homework. . . . He'll say it: "Stop doing homework!"

> [My kids] are 12 and 13, so there are limits. They still know what their needs are so, although they can intellectually understand that Mom needs to study, their needs are still, "C'mon Mom, let's go play basketball."

> I believe in attachment, in a really strong bond with your child, and when you're trying to do school, even if you're only doing two or three courses in a semester, that still gets in your way of being a parent. Timelines, deadlines, extra stress. This year he's into his friends, he's three and a half so he's almost at the age of wanting to be in preschool and in groups, but the guilt!

> I also get overwhelmed sometimes from the guilt. I'll realize that it's been a length of time since I read to the little one, or I'm not jumping up to say "bath time" so our mornings are rushed because you were too tired to do it the night before. And sending them off, you know, that does it, the guilt and remorse.

A further potential for guilt exists for students in human service programs who are learning about appropriate child care practice at the same time as they are facing the real demands of parenting under trying circumstances and often alone.

> I struggle with, here I am, the CYC worker, and here I am, the parent, and do the beliefs I have around CYC, do they fully transfer over to me as the parent? And the answer is, no they don't. Because I don't practice what I preach 100% of the time, so there's a conflict there for me. Because I want to be the CYC worker with my children, but I can't be that person all the time. Because I get tired, because old, not so much ways of being as ways of seeing things done in my own family come back, and because there's a lot more invested in my own children than the children I work with. So I constantly struggle and kind of beat myself up about it.

Lack of support from family was referenced by only a few women. The extreme difficulty that active nonsupport can create likely explains the infrequency with which it was raised: It appears that few women coping with a nonsupportive family made it into this sample, either because they are not enrolled or because they did not have the time to participate in an interview. A few women mentioned lack of family support either as a past irritant or something with which they were dealing, but not as a major barrier. These women's intrinsic resolve is generally too firm to be shaken by minor family resistance. However, the only participant for whom an actively nonsupportive spouse was a major concern had, in fact, temporarily withdrawn from her studies.

> I don't know about my brothers, I don't think they've got their heads around it. I think they see it more as me being frivolous, to be a student and single parent. I should be out there working and slogging away, at 8 dollars an hour.

> I have a lot of cousins that are my age. We were raised together so we were friends and family; we didn't have extended friends almost. And it was O.K. when I went through the first 2 years, but when I hit my 3rd year they pulled back.

> Every time I talked about going to college or university, there was a great uproar. It was, "Well, who do you think you are? Putting a bunch of letters behind your name won't change who you are." From my husband it was more, why would I want to do that kind of work. Why would I want to waste my time with other people's problems.

> Well, sometimes [my husband] comes out with a snarky retort like "I didn't tell you to go to school."

> And it's quite damaging to your relationship, too, if your partner doesn't have his feet on the ground. . . . My husband has not been really supportive . . . so I often thought he was a lot of the stress. I couldn't come home and relax. I couldn't depend on him . . . and I had a lot of resentment about it. I resented the hell out of him and it did damage, and I know that's happened with a couple of other people.

The issue of child care must also be addressed. Access to child care was not raised frequently by the participants overall; however, this does not indicate that access to child care is not an issue for nontraditional students. Few of the women interviewed have young children; those who do invariably raised access to quality child care as an issue. In addition, child care was raised by some of the participants as an issue affecting their classmates, if not themselves.

Finding and keeping affordable, quality care for children is a challenge to any parent working or studying outside the home. The specific difficulties faced by student parents include their low income status and their complex and changing schedules (including evening classes some semesters and schedule changes for practicum). For example, two part-time students with young children described the challenge of scheduling classes around their husbands' availability in order to minimize the cost of child care. The guilt that many parents feel at leaving young children in care is even greater for these women, who often already feel guilty that their study schedule limits the time they have for their families. Parents are always concerned about the quality of nonparental care their children receive; this may be intensified for individuals in the human services given their education in principles of quality care.

> I have a girlfriend in town who takes him on Friday, someone on the island one day and my Mom one day. It's funny, because I have it all set up but every week is different. It's like, this can't be, when is it going to get normal?

> They've never gone to day care and they've developed in their own minds that it's awful, from kids who go and don't enjoy it. . . . And I went through all these old things, you know, not my babies! And it was me; the kids weren't saying they didn't want a babysitter!

> So if I'm going out of the home, for school and practicum, and hiring a baby sitter, then there was always that part of me that said I'm not going to go out for myself now because they've already had a babysitter once or twice this week.

> For me, it's my sanity, if I know my kids are well cared for. If I hadn't found a place where they would be happy, I would have quit school. For sure.

As mentioned previously, although families place pressure on women's time, finances, and emotional resources, they also constitute the most frequently referenced source of support. References to family support include parents, extended family, spouses, and children, and although the relationships differ, they are all generally characterized by both ongoing encouragement and more tangible practical assistance.

Parents are mentioned frequently, often with gratitude that they can be counted on when things get really tough. Parents provide emotional support and encouragement, help with child care, and provide direct and indirect financial assistance. The intense appreciation these women express for their parents raises the question of how women in comparable situations, but without comparable support, manage to cope.

> My parents are supportive in that they've been there if there's an emergency, and I've always known that if I'm not available the boys can call them. So that's supportive. Otherwise I'm sure [the boys] would have felt quite abandoned at times, with the busy schedule that I have.

> Finances have been tight all along, and when crises happen, I've quite often had to turn to my parents to get through the crisis. I was lucky. If my parents weren't there, it would have been a lot more difficult.

> My parents actually moved here from England at the same time and that worked out real nice; we actually stayed with them for a few months. My parents bought the

> trailer so I just pay the lot rent. So I think if I'd been paying--I know other girls who are paying $600 a month rent for a place--that might be tough.
>
> So I actually have a lot of support from my parents with regards to my son. When I was doing a lot of work, school work, they would take him a week out of the month. So I had, every three weeks, I had a week break. And they'll pick him up from school, and if I can't find a babysitter, they take him. And they just phone, connect: "Hi, how are you doing, keep going, whatever".
>
> My parents listened to me when I've said, "Hey, maybe I might not go back. I'm real tired of this, this is real hard." Or if I miss birthdays, or if I'm not there for Thanksgiving dinner. There's a lot of support around that and that was something I was worried about. . . . Actually what they do [when I'm stressed], they don't talk me through it. They've learned that when I phone up and say, "I'm coming home, I'm quitting," they don't say "Oh no, honey, you're so close. You just stay there." Because that's not what I want to hear. I want to hear, "O.K., you can come home." So we've turned this 20 minute wailing conversation into 5 minutes. I'll phone my Mom, "I'm coming home and I'm quitting," . . . and I told her, I need you to say, "It's O.K., honey, come home." So now she goes, "O.K., honey, come home."

Although support from both parents is important to these women, there is no question that for many the relationship and support from their mothers is most critical. One participant mentioned her father as particularly supportive; many mentioned their mothers. One woman identified her mother as a role model. More typically, the involvement of mothers is in being available to listen, in providing practical support such as assistance with child care, and in "rooting for" their daughters.

> My mother went back to school as a mature student; after a divorce she went to university. She was a little younger than me when she started to go. And, she's also helped me financially to a certain extent so it's not only, you know, moral support, but physical support as well.
>
> My Mom especially, like today, she's one of my biggest supporters to be in school. She comes over once a month to help out, and her boyfriend as well is really supportive. That helps, and when things go wrong or get really stressful, she's there.
>
> I'll cry "Mom, I can't handle this". [She'll say] "Keep going, it's all right, you're O.K."

> I can say, "I'm tired, I need to phone Mom." And she can come and help out if she takes the girls.

Women with a spouse or partner identified the support of that individual as critical. As mentioned previously, only one woman with an actively nonsupportive spouse was in the sample. As with parents, spouses' support is both emotional and pragmatic: providing moral support to get through the pressures and frustrations of the student role, and easing the path through assuming a greater share of domestic responsibilities.

> My husband actually takes over a lot of the time because he's worked his schedule out. He doesn't have night classes, and he's worked it out so most of his classes are before the kids get home, and that helps.

> My husband has taken over the Mr. Mom role. He's retired from the military and now he works out of the home.

> He's really supportive and he brings me coffee. He makes supper and cleans house and vacuums and stuff, so that really helps. And I think he's proud of me too. . . . So now he sees the back of my head, but at least I'm happy!

> If I need to vent, I can usually talk to my husband and family.

> I had a partner that I lived with and he was supportive. . . . He knew how hard it was for me to come up and go back to school. He knew about my confidence issues, so he would spend a lot of time listening to me cry and wail.

> I'm fortunate because my husband is very supportive. He offers a lot of encouragement. There were a few times when I got frustrated, and I'd say "I'm going to quit." He'd say, "Now, you don't really want to do that."

Finally, but often most importantly, women draw support from their children, whether young or adult. As mentioned above, many women see themselves as role models for their children and this motivates them to continue their training after bringing them into training in the first place. They relish the pride their children take in them and value their support.

> From what I understand, my kids really like the idea of me being in school too. We talk about it and it helps them that I understand what they're going through because I am too.

> I have really high standards for myself and when I don't live up to them, I feel real disappointed. So my goal is to follow my own things I say to my kids: I say, "It's O.K., you did your best." I say that to them when they're disappointed in how they're accomplishing things. So they turn it around on me and say, "Mom, you've got to remember, you're doing your best." They say that to me quite often, 'cause I'll say, "Man, I put so much work into this!" They'll say, "Mom, you tried!"

> My children are extremely supportive. They're very happy that I'm doing what I want to do.

> Because I've been on my own with [my son] for so long, we've always talked about things, and if I have issues, he's a good little counsellor!

> So any time I think of letting go, [my sons] say "No we can all stick together." So that's been really supportive.

Parents of younger children sometimes have to develop a structure to ensure that they receive the support they need.

> I made a game plan with the children. I told them that for every *A* I got, they got pizza; for every *B* they got a sleepover. We had a board to keep track, so they were very supportive. When I said I had to study, there weren't any qualms, there wasn't any arguing, it was very quiet in the house.

Several mothers of older children mention the healthy competition regarding grades that has developed between them and their children; two women interviewed together laughed that their teenage children think they're "nerds" for getting such good grades, although they are actually proud of their mothers.

> Yeah, the kids think it's great. My daughter and I were going to university at the same time and we got into a competition because she was getting *B's* and I was getting *A's*, which made her get *A's*!

An interesting downside to all the support women receive from partners and children was noted and can be traced to the internal push to be "super woman" with which many participants struggle: that is, the stress women encounter related to giving up roles they still feel they should fulfill. Some women are uncomfortable sharing control of domestic chores, for example, and express this through criticism of their partner's domestic competence. One woman with teenage sons, who were helping financially, expressed her distress at this situation because in her mind she (as parent) should be fully responsible for the family's financial well-being.

> So that was difficult for me, to give up [the domestic] role . . . but it's been a great learning tool, because I did control the home. Especially with him being away a lot, with his job. It was very difficult for me to realize what was wrong the first term. I couldn't figure out why I was so angry. Then I realized that it's all right if he screws up, and he doesn't do things just the way I'd do them. So those kind of things I've had to let go. And now I just, you know, clean the bathroom myself!

> My sons say "You can do it. We'll do it. If you're stuck for part of the rent, we'll help pay it" or whatever. It's just been this year, and for me it's a matter of pride and a matter of knowing that I'm the supporter of the family, and it's been very difficult to, to acknowledge that the children are helping me financially. . . . They're 17 and 18 and they want to help. It's just for me, the guilt that they have to do that, that I'm not providing for them.

Finances

Unlike the other factors influencing women's achievement of their educational goals, which generally have both positive and negative effects, financial issues have a consistently negative impact. The intensity of the stress varies, but in no circumstances are financial issues a positive factor, except in their contribution to students' motivation to seek training and complete the credential.

Only a few women described themselves as free of financial pressures: the two women in the MC&F program who had all expenses covered and received full salary; one woman with no children and a supportive, well-paid husband; one single parent with a generous maintenance settlement from her divorce; and one single parent who received parental support plus a grant adequate to cover expenses for her 1-year program. For all others, financial considerations were a major influence.

Probably the most significant and stressful financial hardship is experienced by single parents: Eleven of the 13 single mothers interviewed were struggling to get by on student loans. Over half of this group (six women) were also managing part-time jobs, adding further time pressures to the stress of raising a family in poverty. Evelyn dashes off campus every noon hour to work as playground supervisor at an elementary school; Cecile does hairdressing in the evenings at home so she'll be with her young children; Jean struggles to fit time with her 7-year old daughter into a schedule that includes full-time school and two part-time jobs; and Jennie earns a little extra money helping her mother with her family child care business after class every day.

> I'm working part-time at the school district as well, and on a student loan, so it's been tight, tight, tight. And so my greatest stress is finances, and at times that gets to be overwhelming. . . . I think, fine, I've only got so much left, can I make it through? Can I manage to pay my rent? Or is this it: has the crunch come?

> The first year was really hard for my children, not having, not having, not having. . . . It's a struggle every day to make ends meet.

> My costs of living are not high because of frivolity; they're high because I have two children. And to have a home that's adequate, that's just the way it works.

> You know, I'm from a middle-class background, and it's hard to live in poverty when you're used to certain standards. I don't mean high standards, but not poverty where I have to worry about what food I have to put in my body.

Two particular financial challenges emerged from single parent participants. One relates to the changes in B.C.'s income assistance provisions in 1996, requiring students in post-secondary programs to go off income assistance and rely fully on student loans. Although students affected by this change were glad to leave the stigma and devaluation of the welfare system behind, they were concerned about loss of medical, dental, and other benefits. The second financial challenge impacting single parents is the difficulty of stretching student loan funds to cover the full ten months of a program: Two women with teenage dependents, interviewed in March, were unsure how they would cope financially until their programs ended in June.

> I have to pay prescriptions now. . . . So it means putting money away every month so that I can afford my prescription.

> [Going off income assistance] was bittersweet: bitter because I knew I'd have a higher debt load and struggle financially, and sweet because I knew I didn't have to answer to anybody any more.

> I mean, I've lived in poverty my whole life. . . . It was rough enough to get your money all at once for the month, never mind all at once for a year.

> I'm scared to death that I'm not going to have enough money for May and June, and I have my practicum so I don't know what's going to happen. . . . I have an 18-year old son who isn't working. So to me, it's like I'm feeding a man, you know, he has needs.

> [The student loan] isn't enough. It's not lasting. It just all goes--rent and bills and food--when you have a teenager. . . . And it's not like I've wasted it or anything, it's just not enough.

Even single women without dependents struggle with loans and part-time jobs. Ruth stays with her mother in another city every weekend, so she can work two full days at the job she had before entering college; Lindsay has cashed in her retirement savings bonds and works in a child care centre before class every day, plus after class when she can; and Rita lives with her parents but works full-time hours in a restaurant in order to pay the expenses for her full-time course load. Married women also cite financial pressures, less intense but still stressful.

> It's very difficult to go from two wages to three-quarters of one. . . . I tried to do it without [a student loan] and finally, last week, I was kicking myself and saying, "Why are you doing this? You can't make this stretch into that. It doesn't work." Thinking it was better not to take a loan, but instead all I've got is a huge credit card bill.

> I was sponsored through UI [Unemployment Insurance] then that was cancelled, so my husband had to pay for my courses because of course I didn't have any money. And of course he's working in a job that's shut down part of each year, and this is a strike year. . . . I had to pay last fall, and it did cause hardship.
>
> We're just working poor. I have a student loan. We've been married about 8 years and we've always been working poor.
>
> I've done loans, and scholarships, and bursaries, and summer employment, and part-time work, which is again stressful.

One concern raised by some of the married women is that financial assistance appears to be less available to them: "The prejudice of that 'tics me off', that they think the women who are single parents are more deserving than the ones who are struggling with a partner."

Twelve of the women interviewed are attending part-time, generally in order to manage financially without a student loan. Two of this group are single parents; the remainder are married, 8 of 10 with children at home. For some parents with young children, studying part-time is essential in order to minimize child care costs; class schedules are arranged around the schedules of family members who can look after the children.

Although studying part-time allows these women to adhere to the "pay as you go" philosophy espoused by some, this financial solution is not without it's inherent limitations, including the scheduling difficulty of juggling a job, child care, or both (issues also faced by many full-time students), and, more significantly, the number of years required for program completion. Mary has been studying part-time for 8 years to complete her degree; Marilyn worries that she'll be ready to retire by the time she graduates.

<u>Considering Quitting</u>

Despite dedication to their chosen fields and determination to complete their programs, many of the participants have considered taking time out from their studies. Often, a combination of pressures precipitated contemplation of stopping out or quitting; generally this thinking appears to have been a short-lived reaction to a seemingly overwhelming situation involving financial, family, and academic stressors.

> I wasn't willing to accept that I couldn't be super everything, but last year and this year, to realize that if I don't take care of myself, if I don't lighten up here and let go of some of the high expectations, I'm not going to make it through.
>
> It wasn't just school. I wanted to quit my family, I wanted to quit everything. . . . I wanted to have my kids go live with their dad. I wanted to get out of my house, put everything in storage, live like a hermit.
>
> Sometimes I just get fed up and I want to go home. I just get tired and I want to "bag it" and I don't want to do this any more.
>
> Quitting? Yeah, what happens is usually when I'm feeling really overwhelmed by it all and I'm thinking, I can't even deal with it any more, it's too much.
>
> There was a time period when I ended up having what I call "the college blues": "Man, I've been doing this forever and ever; I just want it to be done." But that was short lived.
>
> I had to stop school, it was wearing me right out. Physically I collapsed. And I thought, I have to focus on my family and get this thing back together and that will take priority.

Some specific themes emerged regarding women's consideration of quitting. For some women, financial stresses precipitated thoughts of withdrawal. Specific difficulties dealing with income assistance (in the past, when full-time students were still eligible for income assistance) had provoked thoughts of quitting for 2 women.

> A lot of it [considering quitting] was to do with, I'm going to get a job, I'm going to quit, because I cannot stand [dealing with income assistance], I cannot stand being devalued this way.
>
> When I get to that point [of considering quitting], it's always financial stress, all of it financial. And that's what's so crazy because with everything else going on, it's never been anything else but financial.

> I came in here crying a few months ago, bawling my eyes out, because I was running out of money then. Thinking I had to leave the program and that tore me apart because I've worked too hard to get to where I am, for money to delay me again.

Academic pressures contributed to thoughts of quitting for many women, usually in combination with other factors and representing brief periods of panic.

> When things like, at the end of the semester, I think how much longer can I do this. The energy level, it's hard to keep it up. Especially when I have a couple of exams back to back and it's study and study, and it gets pretty hairy.

> I still have those days when I just want to quit. . . . Like Monday, we have three main activities due plus we have a midterm on Wednesday. It's just, like, you get a little panicked inside.

> It's like here I am, no self-confidence, can't do it, ready to quit, then I push and I push and I push and then all of a sudden I've done it, I've handed in my paper.

Additional, individual reasons for considering quitting also emerged. Anna came closest to quitting when she was diagnosed with cancer for the second time: "When I got sick I couldn't see how I was ever going to finish. . . . The only time [I considered quitting] was when I got the second round of cancer and that was not because of school, it was because I was pretty much fed up with all the stress in my life." Christa considered quitting when a particularly intense conflict with classmates threatened to ruin one course and relationships within the program. Kathy, a First Nations student, considered quitting only once, when "fear of success got involved, because I'm the only one from my family [to complete a post-secondary credential] and it makes me different. That's quite common actually in the Indian community, fearing success."

Definitions of Success

One of the primary goals of this research was to examine how this particular group of students views the concept of success: What makes the educational experience successful in their eyes? What will stand out for them in terms of the success of their own educational accomplishments? In summary, these mature female human service students view educational success in transformative terms, as evidenced by major and long-lasting personal changes. Interviews were held close to the end of an academic year, and students (whether finishing a degree or a 1-year program) were generally extremely positive about their programs and the success of the experience in terms of personal growth and transformation.

> There's so much I've gotten out of it. It's like a dream come true.
>
> We'd love to just go to Europe and relish all this growth and knowledge. We're the same people, basically, but we're different than we were four years ago. And it's a neat difference, it's a really neat difference. So it's like, I don't know, a butterfly thing. I don't know how to describe it. I wouldn't have missed it for the world.
>
> Even if I never work in the field, what I've gotten here, I'll take it wherever I go.
>
> What stands out is the knowledge, the learning, the opening up of all my senses, the self awareness part of the program is excellent. Nobody seems to come out of that course the same: You go in and then you come out the other end.

These women identify a holistic set of occurrences and outcomes, including learning, personal growth, and relationships, as indicators of a successful educational experience.

> It will be all combined, the growth, the self awareness, the connections, the growth within the family too. There's a real special bond there when sacrifices are made jointly by everybody, to contribute to the goal. It gives such a different feeling.
>
> What makes it a success? It's a success when you're in an environment where you feel supported and respected and valued. And when you can learn something that you can actually use.

Many women identified the importance to them of all components being in place; for example, emphasizing that grades were relevant but not in isolation. The focus on self and personal growth within these programs was one of the most frequently mentioned determinants of success; another was the relevance and applicability of the learning.

> Now, I realized that--and a lot of it's through the holistic approach of the First Nations people--that the best isn't really the best. You don't have to have the *A*+ if the *A*+ costs you your sanity for a month, or your marriage, or your harmony at home. There's more to it.
>
> It wasn't just about academics and jumping hoops, it was about self and about how self influences practice. And I would say it's been a healing journey too.
>
> Grades have been important, but more my understanding. Like, if I get a good grade and I don't feel I understand the stuff enough, I'll go to the instructor and say "You know, I still feel like I'm not getting this." . . . It doesn't matter what your grade is. For me it matters how much you understand it, how well you can do the practical.
>
> The self-work, challenging core beliefs . . . that's the most important part. I mean the grades have to be there if I want to go on, but for this kind of work the most important part is self-work.
>
> It really helps you to think in a different way. It teaches you how to pull something apart and look at things differently.
>
> It's exposed me to different ways of looking at things, new ideas. Maybe I'm more capable of expressing myself, more comfortable in a wider range of situations. [What will stand out] is the chance to go and learn new things, to have my thinking challenged, and to be exposed to other perspectives, the chance to talk to other people that want to work in the same field, hear their viewpoints. I'll miss that a lot.
>
> I'm using tons of stuff out of the courses; I've learned scads and droves of stuff that all the time, every day, I use.

For some women, accomplishing a personal goal is a hallmark of success: completing a program, improving their education, achieving what they set out to achieve are key features of success to these women, in addition to the learning and growth.

> Success is really important, but I think what it's about for me is all the support I got along the way. Family, friends, coworkers; the whole idea of pulling together and also my own self-awareness. Knowing that, yes, I can do it. Knowing that there's no obstacle too big.
>
> I want to be recognized for the work that I put into what I do. . . . I want to know that I can meet these challenges and I can do it.
>
> I wanted the personal satisfaction of completing what I started.

> [Success for me] is that I raised myself; I raised myself through the system.

The success of the educational experience was measured in unique ways for some women. For example, the First Nations focus of one of the programs was the trait contributing to its perceived success for some participants: "What stands out is the wonderful way in which I have been taught the native culture and spirituality and approaches to child care and harmony that I don't think are taught in other programs." For some, success included the program's reaffirmation or validation of existing knowledge, values, and beliefs.

> I needed to know for myself a lot of the theory and work my way up; to see what my beliefs were and what theories they really fit into and where I'm coming from. . . . The program, it's helped me understand a lot of where I'm coming from. It just firmed up my beliefs, a lot of them that I wasn't too sure about.

As mentioned earlier, many of the women interviewed had financed their education through student loans and were facing the prospect of significant debt upon graduation. A very pragmatic measure of success for this group is their perception of the value of this financial investment. For all who commented, the cost of their education was assessed as a worthwhile investment despite concerns regarding their accumulated debt.

> It's been worth it but it's a lot of money to pay back. . . . It's an experience that I don't think I could have gotten anywhere else.

> We just feel it's a wonderful investment and I just try to think of it that way. I've got to better myself. I want to do this. We'll get it paid off.

> People say, do you regret taking $30,000 out for a student loan to go to school? I say "No, I've got 30 years left to work." . . . And I raised myself, and the skills I've learned, while my son was still young.

Summary: Interview Findings

These women followed a variety of paths to arrive at human service education programs, but were alike in their clarity of goals and high level of commitment to human service work.

They set high standards for themselves, in terms of their performance as students and in terms of maintaining balance in their family, work, and community lives. Most entered college with considerable trepidation, because they had not been strong students in the past or because they had been out of school for so long, or both. Once in the programs, they discovered a joy in the experience and a capacity to exceed their own initial academic expectations. Their lives are complex; factors that add stress (such as family relationships) are often also powerful motivators to keep them going.

They ascribe to a holistic definition of a successful education, which encompasses personal growth, accomplishing individual goals, the value of the content learned, and the relationships built along the way. Completing a credential and getting good grades were definitely important to these women, but not sufficient in themselves as indicators of success. Furthermore, achieving educational goals at the expense of family life would not be viewed as success.

> Like, I love college, but at the same time, I'm getting a Ph.D. in life, and the balancing comes back into it for me. Sure all those practical and theoretical things are great but there has to be a balance.

One final general comment is warranted. That is that I found these women not only strong and determined, but also astute, humorous, and an inspiration to me as a woman and a human services professional. They represented an affirmation of the current and future strength of the human service field, but also an embodiment of the potential of adult education. I feel privileged that they willingly gave me some of their scarce time, openly shared their stories, views, and anxieties, and demonstrated keen interest in my research topic and the questions I posed.

Findings: Document Review

Although the focus of this inquiry is on female students and their personal stories, explored through interviews, relevant documents were also reviewed to gain some understanding of the context for participants' educational experiences. Two types of documents were reviewed: current internal and public descriptions of the college, department and programs, and previous unpublished reports related to the college and its programs. These documents will not be named or cited specifically, as this would identify the institution and program areas.

Institution, Department, and Program Descriptions

The college calendar, departmental handbooks for students and new faculty, departmental statement of Core Values and Beliefs, program brochures, and a sample of course outlines were reviewed for comments or implications regarding student goals, supports to achieving those goals, and indicators of success. Interestingly, there appeared to be more congruence with student views and experiences in these areas at the macro level (such as, college mission and goals as stated in the calendar, the introductory statement in the student handbook from the departmental Dean, departmental statement of Core Values and Beliefs) than in the program-specific literature.

Goals

At the institutional level, the college supports the importance of students achieving their individual goals and recognizes goals related to personal growth, as reflected in the following statement from the 1996/97 college calendar:

> [The College Board] adopted a statement of mission and goals which is based on a commitment to individual growth. . . . [The Mission is to] develop the full potential of

> students in every college program; promote the development of analytical and creative thinking skills in all students [et cetera].

This validation of goals related to personal growth and individual aspirations is also found at the departmental level, in a 1996 "Core Values and Beliefs" statement, "Our curriculum is based on a humanistic framework in which people are encouraged and supported in working towards becoming aware individuals and achieving healthier societies."

The importance of career/employment goals is emphasized at both institutional and departmental levels. The President's message in the 1996/97 calendar states, "[The college] has been a leader in developing innovative programs and courses . . . that are addressing both local and national career and employment needs." The Human Services Department description in the same calendar states "Graduates of our programs have been successful in finding employment working with people in a variety of settings. . . . Our programs meet emerging community and professional needs, and provide flexible opportunities for students to pursue career and educational goals."

In program-specific literature, the balance between career and personal goals shifts to an emphasis on career. Programs are described in the calendar as designed to provide students with the knowledge and practice skills necessary: "to be skilled working with individuals, their families and small groups across a variety of practice settings"; "to work with young children and their families in child care settings"; "for employment in the field of social services to perform many para-professional jobs in the community"; or "[for] working with individuals who are challenged."

Goals in terms of laddering to further credentials are emphasized at both departmental and program level. The departmental Student Handbook makes the commitment that "We strive to provide transferability from entry-level certificate programs through career ladders to diploma and degree programs."

Supports to Students in Achieving Their Goals

At both institutional and departmental levels there is a stated commitment to accommodating and supporting a diverse student population. The calendar commits that "[The college] will . . . promote fair access to programs for students with varied backgrounds and needs through effective support services, flexible learning formats and locally-based instruction." The departmental Core Values and Beliefs document further states:

> We believe there should be access for mature students and part-time students. We value diversity in our students and believe that the classroom should reflect a mix of ages and cultural backgrounds, including members of both sexes and equitable representation of First Nations students, visible minorities, gay/lesbian and bisexual students, and students with disabilities.

With the exception of the First Nations-focussed program, this commitment to accommodating diversity is not echoed in program-specific literature, except through provision of part-time programming options in all but one of the programs.

The stated commitment to a learner-centred, flexible approach does not appear to translate into formal policies that ensure flexibility. For example, the first "policy and procedure" issue addressed in the 1996/97 Student Handbook is that of attendance: "[The college] has the right to cancel your registration in any course or program where you fail to attend and [the college] deems attendance to be important." Some of the course outlines reviewed indicate 20% of course grades assigned to "attendance and participation" without definition of grading criteria for this component or indication of flexibility in interpretation. Policies outlining penalties for late submission of assignments underscore the caveat that instructors "may" be flexible under "exceptional circumstances." Although a few course outlines indicate an admirable attempt at student involvement in course evaluation (that is, by allowing students to determine the percentage of final grade attached to each evaluation component), most reflect no student input to this area.

This institution is clearly committed to supporting students, as evidenced in the array of support services provided and described in the calendar. The departmental Student Handbook outlines all of these supports while program-specific literature references only financial and educational advising.

Definitions of Success

Few post-secondary students pay attention to institutions' stated mission and goals. Those students in the current sample who happened to do so would find congruence between their own holistic definitions of success and those of the institution. At the departmental level as well, print materials suggest a commitment to a holistic definition of successful student outcomes as the following quote from the Student Handbook illustrates;

> Our students generally do very well, and graduates of our programs have secured work in a range of people-serving agencies around the province. . . . Students report a high degree of satisfaction with the work, and the large majority of our graduates plan to continue to work in the field and would recommend the field to others. As well, students have noted that their studies helped them to become more confident, increased their self-esteem, and taught them valuable communication and community-building skills.

Other departmental depictions of success tend to focus more on gains in terms of performance on the job than on personal growth, although the two domains are clearly not separate. The Core Values and Beliefs document establishes the following goals:

> Graduates will establish relationships which are caring, purposeful, respectful, and collaborative. . . . Graduates must have high ethical standards, values and skills. Graduates will use flexible, critical and creative thinking that will support them in their practice and in their own lifelong learning.

Program-specific literature tends to be more narrowly focussed on positive outcomes related to employment. High employment rates are suggested as evidence of program and graduate success. Although there is some reference at the departmental level to the challenge and satisfaction associated with human service work, this potential marker of student success is not highlighted at the program level.

Unpublished Reports

A number of unpublished reports related to the institution and its programs were made available to the researcher, and reviewed in order to shed further light on the context in which the participants are studying and to permit preliminary comparison of participants' expressed views and feelings with those of students sharing aspects of their experiences. These reports included: (a) institutional data on student outcomes by program area, based on province-wide follow-up surveys of students who have completed programs, approximately one year after leaving (detailed 1995 data and a summary combining 1995 and 1996 data were made available.); (b) a 1996 report on feedback from focus groups made up of degree students from across the institution, on the topic of student retention; (c) a 1995 report of a survey of graduates of 3 human service programs, focussed on employment status; (d) a 1994 report of focus group, interview and survey data for female, part-time human service students, examining supports and barriers to educational participation; and (e) a 1992 evaluation of a special offering of one of the programs.

Goals

Few references to student-identified goals for entering and continuing in programs were located in the documents reviewed. One comment from a focus group participant indicated that although her original goal had merely been "the piece of paper," she was pleased by how much she was actually learning. The rationale given for conducting one of the departmental graduate employment surveys was the presumption that students' primary ambition is employment.

The province-wide graduate survey asks students whether their main reason for enrolling was to "complete a diploma, certificate or degree; complete requirements for another credential; qualify to enter another program; improve existing job skills or learn new job skills; decide on a career or change careers; personal interest; or other." (B.C. Ministry of Education, Skills and Training, 1996b). Responses for 1993/94 human service students, interviewed in 1995, indicated that personal interest predominated as the goal for graduates of two programs, whereas improving job skills or learning new skills was most important for students in another program.

Supports and Barriers to Achieving Goals

This topic was a major emphasis of several reports, although the focus was generally on factors potentially affecting retention rather than factors influencing actual learning or personal growth. Several consistent themes emerged, common to the institution-wide data as well as that gathered specifically from the department or its individual programs.

Students of this college generally felt supported by the overall atmosphere of the institution and liked its physical setting and small classes. The quality of instruction was viewed positively, and knowledgeable, supportive faculty were perceived to be in place across the institution. With the exception of one campus, library staff were described as helpful. Within the human services programs, the laddered program model was identified as supportive.

Students from across the institution identified some of the same barriers as students in the human services. Lack of accurate, up-to-date information was a commonly cited problem, with examples including inadequate information regarding financial aid, course prerequisites, and other program requirements. A related issue was class scheduling and access to course options. Some of the difficulties in this category appear to result from information gaps (e.g., course schedules not being published far enough in advance), although actual limits on access to the courses students need, at times that suit them, was also a frequently identified barrier. The need to travel between campuses in order to access required courses was identified as a barrier for some students.

Although the supports offered in the library were generally viewed as a strength, actual library resources were viewed as inadequate. Third and fourth year students used the partner university's library, but this was not viewed as a satisfactory solution. Similarly, access to computer terminals was seen as problematic. Students, institution-wide and within the human services, pointed to text costs as a barrier, and concerns were raised regarding the "value for money" of student society fees. Mature students in particular begrudged these fees, complaining that their unique needs were not addressed by the student society: "Mandatory student activity fees are a scandal! The activities can't be used by adult working students. The Student Society should look at the profile of the student body and provide age-appropriate options--like Day Care."

At all levels, students complained about inadequate opportunities for their input: Examples include frustration that instructor evaluations were used inconsistently, and lack of confidence that student feedback regarding instructors had any impact. Students also shared a complaint regarding the lack of institution-wide cohesion and community spirit; this appeared particularly true of part-time students who described themselves as feeling "alienated."

In addition to these generic barriers, human service students (who are predominantly mature women) identified difficulties with child care and other external pressures (such as family or finances), excessive demands on their time, and frustration with the stress of course evaluation in general and examinations in particular. Time was cited as the major barrier for mature students, who identified the juggling of home, work and school responsibilities as their primary challenge. A previous study of human service students summarized "Comments indicate that anything requiring time or energy that doesn't contribute directly to improving in-class performance would be considered an indulgence."

Definitions of Success

There is some recognition in the various outcome studies that student definitions of success are not limited to employment outcomes. For example, 60% of the respondents in one human service program follow-up study indicated that the self-development aspect of their program contributed a great deal to their professional competence. One of the human service follow-up studies concluded:

> In addition many students reflected on the personal change that they experienced through their post-secondary experience. They cited enhanced personal growth, greater confidence, better communication skills, greater empathy, better abilities in understanding and working within group structures, and an overall increase in self-esteem. Many students felt that their program of studies helped them to become better citizens because of the emphasis on community building and understanding change dynamics. Even students who went on to work in different employment fields cited their post-secondary education as an important personal growth event and one that contributed to their employment and personal success.

Survey respondents leaving human service programs in 1994 and 1995 indicated a high level of satisfaction (defined as having their goals met, plus finding studies worthwhile);

responses for human service programs on this measure were third highest of any program cluster in the institution. The proportion of respondents indicating that they were working or studying in a related field and found their studies to be very useful was also high, again ranking third highest of all program clusters in the institution.

Summary: Document Review Findings

Documents reviewed included program, department and institutional literature, as well as relevant reports (both published and unpublished) provided by the institution. The analysis of documents achieved the following:

1. The educational context within which study participants functioned was explored, and students' general impressions of an institution, department and programs committed to them and their success was supported.

2. Documents illustrated that, despite indications of this overall commitment, vestiges of traditional, instruction- versus learner-focussed approaches remained. Examples included departmental attendance policies that did not reflect the flexibility actually allowed, and course outlines that did not appear to allow any student input to evaluation practices. It appeared that a commitment in principle to a learner-centred approach had been made, but had yet to be fully operationalized.

3. The importance of personal (as well as employment) goals and the high level of satisfaction with their program of studies expressed by participants in the current study were shown to be fairly typical of students in these programs.

4. Many of the concerns raised by the women interviewed in the current study were shared by like participants in previous studies, providing some confidence that cautious

application of the findings to other mature female human service student populations may be warranted.

Study Findings in the Context of Previous Literature and Research

Thus far, the findings have been organized according to the research questions and the themes that emerged in the interviews. Another way to organize the data is under the categories identified in traditional student success literature. A brief look at the findings organized in this fashion follows.

Definitions of Success

For the most part, the post-secondary student success literature assumes that *success* is equivalent to *retention*, and opposite to *attrition*. Student success initiatives are predominantly measures designed to prevent students from dropping out.

According to this definition of success, the programs included in the current study appear to be highly successful. The laddered program model, along with the option to attend full- or part-time and provision for stopping out, promote completion of a credential. Coupled with these program variables was the women's fierce determination to complete, which rendered dropping out an unacceptable course of action for them personally.

However, an institutional definition of success that equates success to program completion is not adequate for the women interviewed. These students would not see merely staying in school, passing courses, and even completing a credential as success. As discussed previously, success for them involves much more, including mastery of usable knowledge and skills, personal growth, and development of relationships. Given this difference in definition, it is difficult to compare the experience of these women, struggling to achieve their holistic and transformative goals, to the traditional research with its narrow focus on student persistence from semester to semester. Continuing in school, for these women, is almost a "given," although continuing without interruption is sometimes not possible, nor are temporary disruptions in studies viewed as indicative of failure.

Some of the women participants, studying part-time, are taking 8 years or more to complete degrees. According to some quantitative definitions, which define success as completion of credentials within prescribed time periods, these women would be labeled as unsuccessful, despite the fact that they are meeting their goals in the face of tremendous odds and will eventually complete credentials.

Factors Predicting Success

Given the dissonance between typical research definitions of success and the experience of these women, attempting to apply identified predictive variables would have limited value. Factors identified in the literature as predicting dropout cannot be used similarly with this population, for whom dropping out is such a remote possibility. As will be discussed below, it is not that the identified factors are unimportant to this population, rather that their relevance is not reflected in terms of dropout prediction.

Background

Background is identified in the literature as an important predictor of retention and, as has been seen, background does in fact strongly influence women's selection of human service education and their determination to complete educational programs. Whereas limited previous educational attainment is generally found to be a risk factor in terms of retention, for some of these women a negative educational history intensified their determination to "prove themselves" through high achievement.

Organizational Variables

Organizational variables within the institution and program are identified in the literature as related to retention and such issues as class scheduling, flexible student evaluation, access to support services, and so on, were definitely important to the women interviewed. However, although these institutional features may influence the ease with which women complete their studies, they do not appear to be as critical as intrinsic motivation in influencing persistence.

> I think you have to have an inner drive, an inner motivation, or it wouldn't matter how many support systems were in place. If that inner energy level, that core stuff, if that wasn't there, it wouldn't matter.

Academic Integration

Academic involvement or integration is identified in the literature as a predictor of retention for all students and particularly mature students, who tend to be critical consumers of educational programming. For the women interviewed, dedication to their intended field of employment was so strong that concerns regarding the academic experience would not likely lead to withdrawal, although such concerns could discourage continuation to subsequent levels of study. Some participants mentioned that they viewed occasional frustration with classes and instructors as just one of the challenges they expected to face. The students interviewed were generally highly involved and satisfied with their academic experience, and often cited their enthusiasm for what they were learning as a prime motivator for continuing. This particular student group was fortunate to be studying meaningful material in a context of strong learner support. The predictive impact of academic nonintegration, cited in the literature, did not appear to be a factor in this department, where full academic involvement prevailed.

Social Integration

Social involvement within a program or institution is identified in the literature as a predictor of retention, moreso for young than mature students (the latter tending to draw their social support from external sources). Positive relationships with classmates were identified as important to many participants in the current study, and relationships with classmates and faculty were highlighted as one of the indicators of the success of the educational experience. However, participants did not suggest that the absence of such relationships would have led them to drop out. Many women, particularly those studying part-time, accepted the reality that there simply is not enough time for relationships outside classes. Even those few who did not feel supported by classmates or faculty did not indicate that this weakened their resolve to continue their studies.

Environmental Pull

For nontraditional students in particular, environmental pull or external environmental factors are identified in the literature as influencing persistence. External factors such as illness, family issues, and financial stresses were among the potential disruptions most frequently identified by the women interviewed. However, given the program model and participants' intrinsic determination, these factors were not predictive of attrition so much as potentially predictive of level of achievement or length of time for program completion. That is, women appear unlikely to drop out because of environmental pull, but were influenced to accept lower grades than they might have liked or to consider stopping out, attending part-time, or stopping at a certificate or diploma rather than continuing to a degree.

Strategies for Improving Student Success

Most of the research attention to success strategies focuses on approaches to minimizing dropout, rather than approaches to maximizing learning. Whereas these strategies might make the difference between persistence and attrition for other student groups, for the women interviewed they more likely made a qualitative difference related to achievement of personal goals.

These women were not prepared to let stressors push them into dropping out; however, they were willing to admit that stressors interfered with their ability to get the most possible out of their education. Given this, appropriate success strategies could reduce stress for these students, make the educational experience more positive, assist in enhancing learning, and thereby support them to achieve success as they define it. Notable in participants' discussion of factors supporting or hindering their success was the high degree of individuality in their needs and preferences. Some consistent themes emerged. However, for many issues, although the importance of the issue is consistent across participants, the preferred solutions were not.

Advising and Counselling

Advising and counselling services were identified in the literature as key determinants of success for most students, in order to ensure a fit between student and program. The participants in this study had already determined the career path they wished to follow; however, access to advising regarding prerequisites, program requirements, laddering opportunities, financial assistance, and so on, was definitely still viewed as essential.

> I had done an extra psychology course that I didn't need to do. I had paid for another course that I couldn't finish and I'd paid for an extension. O.K., so we're talking, like, $800! And my time. So if I'd been aware of that, if someone had looked a little closer at the transcript originally.

Most students found adequate advising available within the department (rather than from generic institutional services), but some felt access to information could be improved. Several of the students took advantage of the personal counselling services available through the college and spoke highly of the assistance provided; others appreciated that counselling was available but had no time (or need, in some cases) to use it.

> I have used the counselling. I did go and see them at the beginning of the term. They were excellent and I would use them again . . . but I just don't get up to the college except for classes.

> I was going to make an appointment [for counselling] last semester, after the car accident, but I just never had the time. I'd missed school already and I didn't want to miss more time to go there. I still haven't made an appointment but I'd like to. I hear they're quite helpful, talking to you, you know.

<u>Student Transitions</u>

Support to students in making the transition into a post-secondary environment has been identified in the literature as critical to new students, and the importance of this factor was supported by comments from the participants in this study. Many were reentering school after long absences and were dealing with significant challenges related to self-confidence. The need for better support for reentering students at the beginning of their studies was raised by many of the women interviewed.

> The only thing that I'd have to say the college lacks, and this is from a single parent perspective, is that I wish that there had been someone to take me by the hand when I came here, the first year, and lead me through things. I mean, where to start.

Specific examples of the types of transition support required included an overview of the services available; physical orientation to the campus; advice regarding financial assistance, study skill and time management strategies; and orientation to the library and use of computers for library research.

> Like, I didn't know that we're eligible for so many hours a week to learn computers, and to go to the computer lab and stuff. There's all kinds of stuff they don't tell you. And there's still a lot I'm not aware of, which is frustrating. Like the financial stuff. I didn't have a clue what I was doing when I went there.

> Information to do with scholarships and bursaries and stuff--it's like, it's clouded. And I think it should be out there, the route to go, how to get it, what you need, all that's available.

> You just go in [the library] and sit at the computer and look at everyone, and think "Well, they all know what they're doing!" . . . There should be more instruction to people. It's busy at those terminals, people are waiting to get on. All these people looking at you like, "Are you almost finished?"

Although women recognized that some of these services were available if sought out, they still felt a need for readier access to information and more proactive support at the beginning of their studies. Print information provided to students may need to be reviewed to ensure that student transition issues are adequately addressed and highlighted. One woman in the intensive MC&F program referenced the belief of many classmates, before entering, that this would be "a breeze" compared to their jobs, their subsequent shock at the work-load, and the wish that they had been advised in advance to "put their affairs in order" for the year so that they could fully concentrate on studies.

<u>Student Success Courses</u>

One of the key retention strategies profiled in the literature is provision of student success courses. Some of the students in this study took part in workshops that included components of typical student success courses and found them helpful. Others appeared not to have had access to such courses, or knew workshops were available and potentially beneficial but were unable to participate due to the other pressures on their time.

Program Quality

A key retention strategy for all students is ensuring program quality. Any student is more likely to remain in a program that is viewed as high quality, and mature students in particular have little tolerance for programming that is not perceived as valuable and a good use of their limited time. Most of the women interviewed were highly appreciative of the quality of the programs included in the current study and placed high value on course and program contents.

One aspect of program quality that appeared important to these mature learners is the experiential nature of the courses, including opportunities for active involvement within classes so that student knowledge and experience can be brought into discussions, plus ample opportunity to apply skills in practica.

> We were presented with the theories and it was, "These are the theories; how do they fit for you? Do they work for you? Is this a tool for you?"
>
> I think we're so fortunate to be able to share experiences. There are individual assignments that we do on our own, but within the class we get the opportunity to discuss things, brainstorm, address our concerns and issues that we might have.
>
> Mature students can challenge what's presented, where inexperienced students can't. It's more of a mutual experience, if it's allowed to be.

These programs used typical evaluative measures, including end-of-course student feedback and annual graduate surveys (conducted one year post completion) plus occasional external program evaluations, to monitor and improve quality. Even so, some students felt a need for further opportunities for input, including midpoint course evaluations so that concerns regarding specific courses could be raised and addressed before these courses came to an end.

Academic Support

Academic support is seen in the literature as a key institutional provision, particularly for retention of "underprepared" students, a category that would include the many students in this study who had not graduated from high school. Indeed, many students spoke positively of the out-of-class supports they had received through this college, including a writing centre used by many students early in their studies. Many also, however, spoke of their difficulty finding time to utilize these supports even though such time was recognized as well spent. Confidence issues also came up; some students who had struggled to feel comfortable within their department and discipline were still nervous about taking part in support activities delivered outside the department.

> The writing centre, like, it's been great for me. It got me out of a *C-* in English; I finished with a *B+* because of the writing centre.
>
> I'd like to come to some of the study skill classes, but I just can't do it. I guess if I gave up sports with my kids I could do it, but I don't want to do that.
>
> It's really scary for me when it's part of the rest of the college. If there were a writing centre down here I'd do that. If you go there you're just nobody, right, just a student at the college. . . . I'd wonder what do I have to do, say, where do I go, what does it look like, do I bring all my books and say, "Here's my project, help me", or what?

The importance of specialized academic supports for adult students with disabilities is recognized in the literature and was reiterated by the participants in this study. Two women dealing with learning disabilities and one with the aftermath of a serious head injury expressed frustration that more assistance was not available regarding these unique needs. As one student commented, what she needed was not sympathy so much as someone with whom to talk and strategize: "I didn't want 'awww,' I wanted 'O.K., let's get on with it!'"

Academic support provided directly by program faculty is highlighted in the literature as key to promoting student retention. Participants in the current study emphasized the importance they attach to support from program instructors in assisting them to be successful academically. Most felt strongly supported, although some commented that students need to be proactive in order to access instructor support.

> The instructors have been very supportive. They're supportive when we're doing personal work, around making sure that extra support is there and offered. With assignments, taking time, if students want to go, taking the time to go over it before it's handed in to make sure the student is getting it or getting what they want from it.

> The support from instructors, that's been first and foremost for me. It's been my experience that they've always been available, and they put that out to us.

> I think all the instructors are actually quite supportive, but you know, it's becoming aware that they're there and making the choice to reach out and ask for that help and support.

However, a few women commented that they did not feel supported by instructors, with the perceived gap including emotional as well as academic support issues. This negativeness, in the context of overall positive views regarding the support received from faculty, highlights the importance of responding to each mature student individually.

> I think the one thing that I really struggled with, especially in my fourth year, is the vacuum of support. I know that they're trying to push us away in order that we get ready to go out there, but I'm still a student, and I still need that support. I want to be able to say, "I'm struggling and I have no idea what I'm doing!"

> Well, the instructors have been good, but it would be nice to get a little bit more support from them. I think you have to go to them, then they'll give it to you if you need it.

> A lot of us feel that they're trying to weed us out. There's so much work, they're trying to crush us.

Environmental Support

Provision of support to students in dealing with nonacademic issues such as finances and child care is recognized in the literature as important for nontraditional students, and support regarding finances was certainly raised by the participants in this inquiry. As mentioned above, several complained about the lack of information available to them regarding financial assistance issues, for example, loan remission, scholarships, and bursaries. Students' perception is that information about this critical aspect of their educational experience must generally be gleaned from fellow students, rather than being presented to them in a proactive manner.

Nonacademic support issues are given relatively little prominence in the print literature from this department and its programs. However, support regarding nonacademic issues other than finances was not generally identified as problematic by these students, possibly due to the overall commitment of the department to maintaining a supportive environment: "Here, they're family-centred, they're child-centred, they're student-centred." Those who had to make child care arrangements did not seem to feel the institution should play more of a role in this, although one student did mention that improved access to on-campus care would be helpful.

Social Integration

Social support from classmates and faculty is identified in the literature as critical to student retention, and the importance of such supports was generally reinforced by participants in the current study.

> I liked that sense of people pulling together to help me become the best that I can.

> It's very caring, very supportive; we learn by being with each other, by being able to express our feelings. The whole team thing, it's always there.
>
> Our program, we're more like a family. Come April, we're saying good-bye to family, which I think is going to be difficult.
>
> The CSW [Community Support Worker] classes--it's like going into a different world of people. They seem to be more laid back, social people. You know, more happy, always showing that they're happy, more people-orientated. I find being in the academic courses very serious: Don't say a lot, don't smile a lot.

Comments regarding social integration issues were highly individualized, however. Many students, particularly those with families, commented that their classmates were supportive in class but that they did not feel the need to develop friendships with fellow students nor have time to socialize out of class. One student commented that personal relationships with instructors were not important: "At this age, it doesn't matter to me. If I have an instructor that I don't really like, that's just one more challenge."

One pattern suggested by the interview data is that part-time students appear to have greater difficulty connecting with faculty or classmates and tend to attach less importance to these relationships. Another suggested pattern is that students attending closed enrollment programs, especially on small campuses, attach particular value to the role of relationships in their educational experience.

Participants also differed concerning the impact of age on their relationships with their classmates. The wide range of responses included occasional hostility toward students perceived to have less maturity or responsibility, affirmations that age makes no difference in relationships with fellow students, and enthusiasm regarding interaction with younger classmates.

> There's contention at times. We have people that come in right from high school, and I'm 32, and there's people in the class who are grandmas and grandpas, so we have

this whole range. So, some of us can support each other nonjudgmentally, and others just can't.

And when they give extensions in school, that really bothers me, because I've tried so hard to stay on top of everything. I almost want to shout, "My God, you know, you don't have any kids, you go home and sleep after school! Why can't you do your work?"

For some reason, whatever reason, I don't know, you don't tend to start looking at stuff till you're in your thirties. . . . I could feel myself react like, "Jesus Christ, we compromised on this already!" You know, it's, like, "You little brats, smarten up or go to your rooms!"

I felt I was being treated like the old lady. . . . Like, excuse me! I kind of get my back up.

It's funny, there's three of us really hooked up and one's 19, one's 31, and I'm 42 and it's really funny, age doesn't matter. . . . We talk on the phone every day.

We're all [ages] at the same level, you know. We're in there to learn the same things and I've never felt anything from the other students like, "What's an old woman like you doing here?" No, they're great.

Some of the younger ones have gravitated to me for nurturing. . . . Maybe it's there and I don't see it, but [age] is just not an issue with me. My friends have always been all over the board in ages, and so it's not an issue. So [my classmates] have been excellent support.

I've always been interested in young people, and I find coming back to school, that sitting in a class with all these young people, I really admire them. They work so hard, you know, and they're just really nice. That's where I want to be, I feel young when I'm around young people, and that's the way I want to be.

The kids are really neat. They really enjoy someone their parents' age that they can sit and shoot the breeze with. And it was fun to be with them. They were sometimes in awe of the mature students, because mature students are such keeners. So I was very aware of that, but then I work with kids. I'd always say, when I got a chance to stand up in class, how neat it was to be with the younger people and how many insights I'd gotten, and how I'd forgotten the fresh insights of youth, and how, when you get older, you get jaded and miss things. You know, because your eyes aren't wide open.

Organizational Issues

Organizational adjustments at the program and institutional level are identified in the literature as potentially supportive to retention of mature students. Such practices as flexible class scheduling, flexibility in evaluation processes, and provision for part-time enrollment options and assessment of prior learning are highlighted as valuable in facilitating student persistence. The programs involved in the current review appear generally to model the flexibility that mature female students require, whereas the print materials reflect more traditional and inflexible models of operation.

The students interviewed here supported the critical need for flexibility in order to allow students with family and other responsibilities to meet program demands. Students did not generally suggest that program demands be reduced for them, but many mentioned how essential it is for them to have flexibility in terms of how and when they meet requirements. Several mentioned, appreciatively, that extensions on assignments are routinely available. Reflective of their high standards, however, some added that they do not like to use such extensions because they might encourage falling farther behind.

> I've rarely missed classes, but I quite often have to leave early to pick up [the children], 15 minutes or so if I have a class till 4:00, and I've never been made to feel like I'm expecting special treatment or that I shouldn't be doing it. It's accepted as just a part of my life, and an important part, and I need to do that, and it's O.K.
>
> I really like the way it's set up. I get a lot of half days, so I have time to do other stuff. . . . Having the time is important, and [the instructor] is really understanding, saying "O.K., we don't have to meet every Friday. You have these things to do."
>
> The instructors were really good about letting us bring our kids and sometimes there'd be a ton of kids in the class.
>
> When I got sick the faculty rearranged my exams, they rearranged my schedule for me so I could go to Victoria for my treatments.

Students also highlighted the importance of consultation and communication regarding organizational changes in programming, given their many other commitments. Some part-time students did not feel fully apprised of program modifications, and students with children felt a need for earlier notice regarding scheduling changes.

> I've found that there is a lot of last minute rearranging of times and schedules, and it's for good reasons, but I need a month to tell my children, arrange the babysitter, prepare my mind, and do all the things I need to do to get ready.

As indicated previously, organizational details impact individual women differently, and specific logistical arrangements suit some but not others. For example, although many complained of the difficulty created by the need to travel between campuses, some commented on the relief that this travel time provided to them (not eliminating the difficulties also identified, but providing some compensation).

> The drive is good. I put my tapes in, tapes that nobody else wants to hear.
> I play them really loud in the car and sing, and think. And next thing you know, I'm home again.

> For half an hour, I don't have anywhere to be but in my car. The music's there and I take a coffee and relax. So it's a point of view--once a week you have to do it so I can either relax or feel lousy that I have to do it.

Class scheduling is another issue of great importance to mature students, but an issue where preferences again diverge. Scheduling classes in the evening is probably the most controversial issue: For some students, this scheduling is what enables them to participate, but for others it creates resentment.

> I really appreciate the fact that there are evening classes to accommodate me when I couldn't do the daytime ones.

> Evening classes have worked out all right for me 'cause it's allowed me to work during the days.

> I don't like evening classes. They're 3 hours long, and even for the students who work full-time and then come to class, they find it a real stretch. . . . I don't like it at all and I don't think they're productive.
>
> No more night classes! . . . Those of us who had kids at home were very upset, because it took away our family time.
>
> I didn't like the fact that it was evening and I had to hire a babysitter again, after they'd been in day care all day.

Some students had participated in prior learning assessment to gain credit for previously acquired knowledge and skills, and advocated for increased use of this approach. Others spoke generally of the need to do more to validate the knowledge brought by returning students: "I still feel that mature students have a lot of wisdom, they come with a lot of experience. I just don't think that mature students get the respect that they need."

How they are evaluated is an important issue for these high-achieving women, but they do not agree regarding evaluation preferences. For example, some interviews reflected passionate aversion to exams ("The only thing I have a big beef with is final exams. What are you measuring, my endurance?"), whereas others indicated a definite preference for exams over papers. Although evaluation issues appear to be more an annoyance than a potential major influence on persistence, the importance of flexibility in this aspect of course and program organization is again highlighted. The course outlines examined suggest that in most courses students are not given the opportunity for input to methods of evaluation.

Respecting Diversity

Key to facilitating success for students from diverse backgrounds is program and institutional respect for diversity. The programs included in the present inquiry have a strong commitment to diversity, in particular with respect to First Nations learners. Students appreciated the respect for diverse family and economic situations reflected in the flexible program policies and the overall student-centred approach mentioned above.

CHAPTER 4: DISCUSSION

Unique Aspects of the Findings

This study took a focussed look at a single group of students (women in human service programs within one institution), rather than examining aggregate data for a diverse cohort as is characteristic of most student success research. Given this approach, the findings must be viewed with some caution in that they cannot be assumed to suggest patterns applying to other student groups.

Student Traits

The generalizability of the findings to broader post-secondary populations is limited by the dissimilarities between the participants and either B.C. norms or norms for the institution studied. Some of the differences that make this student group unique are summarized below. Some of the figures cited below for the post-secondary educational system and this institution are taken from college and institute student outcomes data, representing only students who have completed most or all of their programs. Some variations may exist between this group and the total population of enrollees.

1. The participants in the current inquiry are all female. Approximately 61% of part-time and 52% of full-time post-secondary students in B.C. were female in 1995/96.

2. All but one of the participants are over age 24, with an average age of 36. In 1995/96, 40% of B.C. university students and 45% of college and institute students were over 24.

3. Provincial (college and institute) and institutional averages for equity group participation in 1994/95 were: (a) 3% both provincially and at this institution for First Nations students, compared with 14% for this study; (b) 5% provincially and 6% at this institution for students with disabilities, compared with 11% for this study; and (c) 18%

provincially and 8% at this institution for visible minority students, with no participants in this study.

4. Ninety-five percent of the participants are or have been married. Although comparison figures are not available for the general student population in the province or at this institution, marriage rates are unlikely to be comparably high due to the relative youth of the general student population.

5. Eighty-three percent of the participants have children and, of this group, 90% have their children living with them. Again, comparison figures are not available but this likely exceeds typical averages.

6. Thirteen percent of the participants are single parents with children living at home. Overall comparison figures are not available; however, 8.5% of B.C. student loan recipients in 1995/96 were single parent females, and it is likely that single parents are "over represented" among loan recipients, compared with overall post-secondary populations.

7. Thirty-nine percent of the participants did not graduate from high school with their age peers. For a post-secondary population this is likely considerably higher than average. For the general population ages 20 to 24, only 21% indicate that they have not graduated from high school (Statistics Canada, 1991, cited in B.C. Ministry of Education, Skills and Training, 1997b).

8. Approximately 30% of students in B.C.'s post-secondary system rely on student loans, compared with 44% of study participants. Removing the part-time students who would not be eligible for student loans, 57% of potentially eligible students in the current sample have student loans.

Program Choice

As well as the sociodemographic differences from the norm that these women bring to their education, their chosen program of studies also distinguishes them. Unlike students entering many other disciplines, it is clearly not the prospect of lucrative salaries or high status jobs that motivates them. Human service positions are noted for low salaries, and the general devaluing of the field and its workers (as expressed through these low wages) is commonly recognized. Further, possession of a credential is not required to gain employment in most human service employment areas. Despite this, the motivation of these participants is firm and sometimes fervent.

Motivation to enter a specific field is often shaped by family and background forces, and the same is true in the human services, although in a decidedly different manner. For many adults, career choices result in part from knowledge about the work world and about specific occupations gleaned from observing and hearing about the careers of family members. In human services, probably more than in many fields, motivation is generally shaped by personal events unrelated to individual or family employment history. Students experience helping relationships through the course of their own or family members' resolution of personal issues or through volunteer work; they discover their joy in working with people similarly, through family involvement, work with sports teams, and so on. None of the participants in the current study mentioned employment of family members in the human service field as having influenced their own career choice.

Further, although many students are attracted to a field strictly as a job, these students are drawn to human services in a more holistic fashion, reflecting an interest in human relationships across their work, family, and community lives. They generally appreciate that effective work with vulnerable individuals requires a certain amount of personal growth and level of personal competence. They have recognized (some before entering, most once in programs) that what they are learning is not abstract or separate from themselves, nor are the competencies acquired useful only on the job. They value the fact that what they are learning has bearing on their personal lives and will benefit them even if they do not work in the field.

Students expressed appreciation for program content, both in terms of preparation for employment and in terms of its relevance for them personally, with reference to several specific elements of the programs. These include the following:

1. Students are enthusiastic about the actual human service knowledge they are acquiring. They are passionate advocates for vulnerable people (such as young children, at-risk adolescents) and love learning about these individuals and the issues affecting them, both because they see the relevance of this content to their professional practice and because it applies to them and their personal lives.

2. One of the program components receiving strongest student support is the focus on personal growth and understanding of "self." Students identify this as one of the most challenging aspects of the programs, but also one of the most valuable.

3. Students are positive about the programs' instruction regarding, and provision of opportunities to practice, group and teamwork skills.

The uniqueness of the programs included in the current review also warrants mention. Although it is likely that human service programs are generally more "student friendly" than many others, due to the orientation of the field and the faculty working within it, it cannot be assumed that the program features so important to this group of students are replicated in all human service programs. Key features of this program include its climate of flexibility and learner-centredness; the laddered model that allows students to successfully complete a credential whether they finish the equivalent of 1, 2 or 4 years of full-time study; the support for part-time studies; the accommodation of students who stop out; and the validation of mature students through provision for assessment of prior learning as well as through a focus on participative and experiential learning. Generalization of findings to other mature, female human service students would be inappropriate unless these program features were also in place.

Summary

This study explores the goals and feelings of a unique group of students studying in a unique set of programs. Understanding the phenomenon of student success for this group was the primary goal of the study, and this has been accomplished. Implications of the findings in terms of a theoretical framework for this group will be explored below. A secondary goal relates to exploration of possibilities for applicability of the results beyond this group. A discussion of implications of the findings in terms of educational policy and practice comprises a later section of this chapter.

Implications for Student Success Theory

Limitations of Existing Theoretical Formulations

The retention literature recognizes that nontraditional students are qualitatively different from young, full-time students and that their completion patterns must be studied separately. However, the basis for studying success for this student group remains firmly rooted in traditional models. Traditional student success theories are not reflective of the participants in the current study for two primary reasons. First, success continues to be defined as *persistence within a program or completion of a credential*, a definition far too limited for these women. Second, clusters of factors are viewed as relating in a linear, causal manner (see Appendix A). The complex, dynamic, and multidirectional interrelationships between these clusters as they apply to mature women are not adequately considered. These theoretical weaknesses are discussed in more detail below.

Adult learning embodies both mastery of skills and knowledge and the transformation of beliefs and perspectives. Learning that encompasses both components is most likely to appeal to mature students: that is, learning that is relevant, makes a difference to them both personally and professionally, and adequately recompenses their significant efforts to participate.

Mezirow (1994) defines transformative learning as "the social process of construing and appropriating a new or revised interpretation of one's experience as a guide to action" (pp. 222-223). Participants in the current study soundly supported a definition of success emphasizing transformative outcomes. A commitment to transformative learning is particularly congruent with the human service disciplines, where much of the curriculum focuses on recognizing and transforming one's own and society's views about devalued individuals. Transformative learning models the transformations that human service students are encouraged to support in others.

For these students, success must be defined in terms of both relevant learning and profound personal change, consistent with both ecological and transformative theory: "Development is defined in this work as a lasting change in the way in which a person perceives and deals with his environment" (Bronfenbrenner, 1979, p. 9). A successful educational experience is one that "helps learners learn what they want to learn and at the same time acquire more developmentally advanced meaning perspectives" (Mezirow, 1991,

p. 168). Given the subjective nature of these definitions, it is clear that learner progress and outcomes must be defined and measured on an individual basis. Merriam and Clark (1991) suggested that transformation, to occur, must be subjectively valued by the person.

Theoretical frameworks traditionally focus on retention as a proxy for success, perhaps because of the intangible and individual nature of success when viewed more qualitatively. Therefore, at least for the participants in the current study, such frameworks are of limited value. Theories that address issues of keeping students enrolled, without addressing issues related to promotion of transformative learning, add little to our understanding of success as defined by this student group. Theories that focus on students in isolation from the important ecological contexts of their lives are also inadequate. Beatty-Guenter (1994) suggested that retention is a necessary condition for accomplishing other goals. For mature students, even this could be debated. However, accepting this premise would mean that retention models only help us understand one precondition for success, not success itself.

Traditional retention models express the relationship between relevant factors in linear, causal terms. Although attractive in its simplicity, this perspective does not adequately address the complex interrelationships that actually exist between personal, environmental, and other variables. Particularly for mature women, motivation to enter and continue in adult education has been shown to be multidimensional, changeable, and interconnected with all other aspects of the learner's life (Merriam & Clark, 1991; Woodley, Wagner, Slowey, Hamilton & Fulton, 1987). As Caffarella and Olson (1993) pointed out, "A single linear pattern of psychosocial development appears to be almost the antithesis of what might be termed the 'norm' for women. Rather women's development is characterized by multiple patterns, role discontinuities, and a need to maintain a 'fluid' sense of self" (p. 143).

Further, a causal, linear model does not reflect the fact that mature learners are not only impacted by the educational system, but also exercise influence upon it: "A more complete view is one in which the relationship between students and the college environment is seen as both reciprocal and dynamic. . . . The interplay between students and institutions is both subtle and complex, where direct cause-and-effect relationships are difficult to detect" (Dey & Hurtado, 1995, pp. 208-209).

Tinto's (1993) most recent theoretical model reflects the interaction between students, institutions, and communities that produces persistence, and recognizes the likelihood that students' progress through post-secondary education will not be continuous. The model focuses on "rites of passage": "Our model of student institutional departure sees the process of persistence as being marked over time by different stages in the passage of students from past forms of association to new forms of membership in the social and intellectual communities of the college" (Tinto, 1993, p. 135). This linear emphasis does not fit for the women interviewed, as will be discussed below.

Toward a Theoretical Formulation That Fits

For the participants in this study, simple linear models of causation and prediction need to be replaced with a dynamic, recursive perspective. A holistic approach is required, a "panoramic view" (Merriam & Clark, 1991, p. 224) of the interconnected spheres of women's lives.

It must again be emphasized that generalizations about the women who participated in this study cannot be assumed to suggest theories of student success for other learner groups. Further, discussing generalizations is not intended to minimize the uniqueness of individual participants' situations or views. However, some general statements can be made with confidence, based on consistent themes in participants' forthright disclosures. No attempt will be made to enshrine these general principles in a diagram: The interrelationships between relevant factors and variables are too fluid and changeable to be captured in any static representation.

Ambivalent effects. First, at each stage of women's pursuit of human service education, the same factors simultaneously impede progress and provide support. For example, lack of success in previous educational experiences seems to both disadvantage women, by predisposing difficulties in academic preparedness and confidence, and, at the same time, facilitate success by pushing women to "prove themselves." Family responsibilities are both a barrier to success, through the time demands and stresses families present, and a motivator of success, in terms of determination to succeed on behalf of their families and families' provision of emotional and practical support. This deviates from the "black or white" approach of traditional theories, which attempt to categorize factors as singularly positive or negative in terms of prediction of, and support for, student success.

Multiple roles. Second, the mature women in the present study cannot be viewed strictly as students. For them, their multiple roles (student, mother, partner, employee, volunteer, support to extended family and friends) merge into the complex reality of who they are. They are not students who incidentally have other responsibilities; they are mothers, daughters, volunteers, and so on, pursuing an educational goal relevant to them in all their roles. This contrasts with typical theories that suggest that nontraditional students, as students, are influenced by external factors such as family. Family, career, and community commitments are not external factors pulling women away from their studies (as the concept of "environmental pull" suggests), they are integral to these women's identities and to their interaction with the educational program. Further, women do not leave past relationships behind as they integrate into post-secondary institutions, as Tinto's (1988, 1993) rites of passage model suggests. Merriam and Clark (1991) illustrated that women tend to perceive themselves to be more actively engaged in learning during periods when things are going well in their work and emotional lives, reflecting the interrelationship between learning and other spheres of women's lives.

Learner supports. Third, because the motivation of these women to attend a post-secondary program is so intense and strongly internal, the interventions required are different from those for students whose motivation is weaker and less intrinsic. Students straight out of high school, attending post-secondary institutions either because they are not sure what else to do or because of parental pressure or encouragement, require different supports than mature students who are self-directed and determined. As Tinto (1993) summarized, "Unlike the typical youthful high school graduate who goes to college *instead of doing something else*, the typical adult student goes to college *in addition to doing other things*" (p. 76). To achieve success this young student group requires a combination of marketing (the values of post-secondary education), motivating (to stay in school), and nurturing. For mature women, marketing is unnecessary, motivation is not an issue, and nurturing is provided by family and other supports. What is needed isn't provision of services so much as elimination of barriers, and flexibility in response to individual needs. The types of supports that could be helpful are also different (e.g., child care versus social or sports activities), as is the overriding requirement that supports not impinge significantly on learners' time.

Success indicators. Fourth, for these women, there is no question that success involves more than passing courses or completing a credential. Success for them is too comprehensive and profound to be arbitrarily delimited and quantified into number of courses passed or length of time to complete a program. Models of success that address only factors related to persisting or dropping out are not adequate. For many of the women in this study, dropping out is not even an option: There is little doubt that virtually all of them will complete a credential, even if it takes them far longer than anticipated or they do not complete the level of credential to which they might have aspired. To be relevant to these students, a theory of student success must include factors promoting learning and transformation, not simply factors impacting continuation, and focus on successfully accommodating education, jobs and personal demands in a balanced and satisfying way, rather than subordinating other areas of life to academic success.

Summary. In summary, a new model of student success, applicable to this particular student group, would look very different from the models traditionally employed to explain students' decisions to complete or withdraw from educational programs. Rather than linking factors in a linear fashion, a diagrammatic representation would comprise a web of interconnected factors. The complexity of mature female students' lives and the interrelationship between the components of these lives would be illustrated, without implying prediction or causation.

Causal models suggest that the environment can be manipulated in ways that influence students to stay in educational programs. For these mature women, with their strong internal drive, external manipulation is likely to make only marginal difference in terms of persistence patterns. A model that makes more sense for these learners assumes that they *will* be successful in some fashion, with success defined individually according to their goals and ecological reality. This model focuses on process instead of prediction. It assumes that a process embodying adequate connection with learners, support to learners in balancing needs and priorities, and flexibility in responding to learners and facilitating their progress, will result in institutions successfully ensuring that students, in turn, are successful by their own definitions.

Appendix E provides a model developed by Alfred, et al. (1992) that outlines assessment practices in a reformed post-secondary educational system, but which also could be viewed as a model of institutional organization designed to promote student success. This model avoids identifying factors predictive of success or specific interventions designed to influence student behaviours. Instead, it depicts processes that need to be in place to ensure that each student's needs are assessed and met, through interventions designed with individual students and their goals in mind. A model suitable to participants in the current study would be based on the assumption that they can and will define and achieve success in individual terms, so long as institutions assess and adjust to their needs and minimize obstacles in their path.

Implications for Policy and Practice

Implications for post-secondary policy and practice have been implicit throughout this paper and will be articulated and consolidated in this section. Issues related to student finances will be discussed first, followed by a discussion of educational reform initiatives relevant to student success.

Financial Implications

Finances are addressed first because of the prominence of their impact on participants, and the significance of this finding related to previous student retention literature. As discussed earlier, the literature typically downplays the importance of financial issues to students' persistence patterns. Although Tinto's (1982) statement that finances are "at the margins of decision making" (p. 689) may hold true for traditional students, it does not appear accurate for the population studied. The greater significance of finances to this older student group--most with dependents and many solely responsible for supporting their families--is not surprising.

Despite recent tuition fee freezes, student loan awards have not kept pace with either educational costs or costs of living (Beauchesne, 1997; B.C. Ministry of Education, Skills and Training, 1997b). Single parents in particular are disadvantaged by the student loan system (B.C. Ministry of Education, Skills and Training, 1997b), and this group is most likely of all students to default on their loans (B.C. Ministry of Education, Skills and Training, 1996, unpublished data). Previous to 1996, single parent students could receive income assistance while attending post-secondary programs, relying on loans only for tuition and books. This option is no longer available: Single parent students attending school full-time must now rely wholly on loans, potentially resulting in increased debt load, and no longer have access to medical and dental coverage or to grants for emergency situations. Financial support from family members in times of crisis is often not available to women who have been on income assistance, for the same array of reasons that precipitated their dependence on government support in the first place. For some women, the sudden switch from monthly cheques to "lump sum payments" of student loan funds for a full semester (or full year) creates a challenge in terms of money management.

Finances are also an issue for participants not relying on loans, but in a more indirect fashion. For this group, financial pressures often require students to maintain part- or full-time jobs while studying. This pattern, coupled with the other responsibilities carried by mature women, not only interferes with their integration into the program and institution and potentially limits their capacity to focus on learning, it also results in the frustration of elongated time periods for completion of credentials.

Despite this, Tinto's (1982) assertion is not totally incorrect: Finances are at the margins of decision-making related to dropping out, because dropping out itself is rarely considered by these women no matter how acute their financial crises. Finances are, however, central to decision-making related to significant choices such as whether to work while studying, attend part-time, stop out, or stop at completion of a lower level of credential than might otherwise be attained. Further, financial challenges can significantly influence the ability of these determined women to achieve the high level of learning that, for them, is equated with success.

Questions regarding the adequacy of B.C.'s student financial aid system generally elicit the response that the system is among the best in Canada, and that students are, in fact, better off than had they been on income assistance. It is not within the purview of this paper to explore this, but it is essential that notice be taken of the very real trauma that financial challenges are creating for some women students, particularly those who are single parents. The women in this study are hard-working, determined and have accepted the barriers facing them; however, the stresses they are enduring and the sacrifices they are making appear excessive. Circumstances are forcing some to raise their children in poverty and exhaust themselves juggling jobs, along with school and other commitments, in order to achieve an education and enter their chosen field of work.

Female students remain financially disadvantaged. Women's wages continue to be lower than men's, so women earn less in the part-time and summer jobs they take to supplement student loans. Single mothers are limited in the amount they can work while studying and looking after their families, and single parents are more likely to have child care and medical or dental expenses for which they must also borrow (if possible, within available loan limits). Women's lower incomes upon graduation make loan repayment more difficult, and both married and single mothers are constrained in their ability to relocate or accept positions involving shift work or other scheduling complexities. Further, student loan repayment remains a particular challenge for graduates entering human service occupations due to the low wages they can anticipate. In 1995/96, the average debt load for a 4-year degree in B.C., after loan remission, was $17,200 (B.C. Ministry of Education, Skills and Training, 1997b), a daunting amount for any low-wage worker. Single parents can borrow more than this average, resulting in an even more overwhelming total debt.

Government needs to take note that women and their families are suffering in the pursuit of valid educational goals. Post-secondary institutions and personnel also need to take note. Financial assistance policies are clearly not the responsibility of human service faculty, however it is incumbent on all personnel involved with nontraditional students to recognize the likelihood that significant financial challenges are being faced routinely and are impacting students' ability to succeed. Institutions can demonstrate empathy for the financial plight of their students through small actions such as implementation of consistent policies to minimize costs for text books, or establishing emergency grant funding for students in dire financial need (ideally with contributions from government, faculty, staff and student associations), as has occurred in some institutions.

In the programs included in this inquiry, faculty have generally moved beyond a narrow course content focus to respond to students and their needs in a holistic fashion. However, there appears to be a hole in this holistic approach, and that is recognition of students' financial challenges and the significant direct and indirect effects these challenges can have on all aspects of students' experience and learning.

Educational Reform Implications

As mentioned previously, there is considerable congruence between the educational reform literature circulating in B.C. and the findings of this inquiry. Educators and policy makers world-wide are exploring issues and strategies related to post-secondary educational reform.

Need for reform. The reasons for this are many. First, there are financial factors forcing this analysis. Governments in North America and around the world are wrestling with the challenge of meeting increasing demands for tertiary education and training in times of fiscal restraint (Association of Applied Arts and Technology of Ontario, 1995, 1996; B.C. Ministry of Education, Skills and Training, 1997a; Faris, 1997). The Government of Canada has eliminated transfer payments to the provinces for post-secondary education, with the result that several provinces have cut spending in this area. B.C. has chosen to increase spending on post-secondary education despite federal cuts. However, increases have not kept pace with inflation or institutional costs, resulting in significant financial pressures on the institutions (B.C. Ministry of Education, Skills and Training, 1997c). Although enrollment remains strong at many B.C. post-secondary institutions, others struggle with reduced enrollment and resultant low utilization of funded spaces. This produces a loss of tuition revenue and also the risk that government funding will be reduced in future. Financial pressures push institutions to explore more cost effective ways of delivering education, but also ways to attract and keep tuition-generating "customers."

Second, institutional change is prompted by the dramatic alteration in the nature of post-secondary student populations. The traditional student, entering directly from high school, attending full-time, and supported by his or her parents, currently represents only about half of the post-secondary student population in B.C. (and less than half of part-time students). This shift in the nature of the student population necessitates a parallel shift in post-secondary educational programming (Association of Colleges of Applied Arts and Technology of Ontario, 1995, 1996; Barr & Tagg, 1995; B.C. Ministry of Education, Skills and Training, 1996a; Simosko, 1995). The strategic plan for the B.C. college and institute system, "Charting a New Course," summarizes the issue:

> B.C.'s colleges and institutes were originally based on traditional post-secondary structures, policies, programs and services. They were designed around campuses using classroom and lecture formats that served young, full-time students and that scheduled courses and class times Monday to Friday, primarily during daytime hours. . . . The overall make-up of the student body has, however, undergone significant change over the years. The so-called "traditional" learner is no longer predominant in the system. Learners come from a variety of ethnic and cultural backgrounds, more are taking courses part-time and a larger proportion are older and consequently, have been out of high school for several years. . . . The diversity of the learner population requires that the system respond to the needs of all learners--male and female, young and old, and those from varied cultures. The system must continue its attempts to break down barriers, provide new pathways to non-traditional learners, and eliminate gender and cultural stereotyping. . . . The diversity of the population requires diversity in the approaches taken to learning. . . . Opening doors to women, visible minorities and other equity groups and creating a supportive internal environment once those doors are open, are necessary parts of "redefining the classroom." (B.C. Ministry of Education, Skills and Training, 1996a, pp. 9-19)

Susan Simosko (1995), in a paper commissioned by the Ministry of Education, Skills and Training, sheds further light on the student changes that are driving reform of post-secondary programming in the province:

> Most programs and course offerings in the post-secondary system are built on certain premises that reflect a particular paradigm of education and training that was established to meet the mass market needs of growing numbers of homogeneous

> groups of young people twenty or thirty years ago. Scheduling classes for the day, monitoring class attendance, limiting the amount of control individuals have over how and what they learn--these strategies may have been effective in the past, but in today's world they no longer hold up so well for many individuals. (p. 27)

The no longer atypical mature student population is impacting upon the environment of colleges and institutes both through its nature and needs and through its assertiveness and expectation for input to all aspects of the education being purchased (Dey & Hurtado, 1995).

Third, educational reform is necessary because of increasing public and government concern regarding accountability for the returns on public funding. Post-secondary institutions are not completely comfortable with being held accountable for the outcomes of their programs. As the B.C. Auditor General commented in a 1993 report on post-secondary education, "The natural tendency of institutions is to report on effort--because they know and are proud of how much effort they put into the job--rather than accomplishment, which can take a long time to achieve, and which can be affected by forces outside their control" (quoted in Association of Colleges of Applied Arts and Technology of Ontario, 1996). Along with public concern regarding the value per taxpayer dollar of post-secondary education, students themselves are increasingly insistent that their personal investment of time and money be worthwhile. It is no longer enough that post-secondary institutions manage the "inputs," ensuring that students enroll and teaching occurs. Students want their learning goals to be met; governments and taxpayers want to see evidence of "outputs" such as high rates of completion and outcomes in terms of graduate competency: "Our colleges rise or fall in the eye of the public based on what students do, not what is done to them" (Alfred et al., 1992, p. 24).

As focus shifts to the outcomes of education, some shortfalls of the current system become apparent. For example, it is clear that the effectiveness of the dominant didactic model and traditional institutional practices is in question:

> An abundance of research shows that alternatives to traditional semester-length, classroom-based lecture methods produce more learning. . . . We waste our students' time with registration lines, bookstore lines, lockstep class scheduling, and redundant courses and requirements. We do not teach them to learn efficiently and effectively. (Barr & Tagg, 1995, p. 23).

B.C.'s performance in terms of overall post-secondary participation and success has received its share of criticism:

> Douglas College, among others, often silently and usually blindly witnesses enormous failure rates--almost 1/3 of all students repeat one or more courses, 16% become placed on academic probation at some point, 475 officially withdrew from at least one course, and only 21% earn any college credential from us. Few people appreciate the carnage taking place in our classrooms. (James, 1995, p. 17)

> Traditionally, British Columbia's overall participation rate in post-secondary education and training has been lower than the national average. This is increasingly problematic given the requirement for higher levels of education for successful participation in the labour market. . . . Students with disabilities, visible minorities and women have, each in their own way, faced physical, attitudinal and cultural barriers. (B.C. Ministry of Education, Skills and Training, 1996a, p. 25)

It is recognized that future funding will not keep pace with demand and that current models are not delivering adequate returns for the billions of dollars already invested. Students, employers, taxpayers, and government each have concerns regarding the system's ability to deliver the level of education and training required for the complex world of today and tomorrow. A new paradigm is needed.

Nature of reform. The new approaches to post-secondary education currently being proposed embody several key themes, woven around a central commitment to learning and learners, that are also central to promotion of student success. Successful institutions are identified as ones that view students holistically and make student well-being and goal-achievement their primary focus. Their commitment to students underpins and is reflected in everything they do. In a successful institution, under the new paradigm, "Student outcomes are supported by the organizational culture. . . . Colleges become involved with students as 'whole persons' holding multiple roles and responsibilities. . . . Virtually every institutional policy and practice--from class schedules, faculty office hours, student orientation, parking, child care, and access to support services--is carefully considered in relationship to student needs." (Alfred, et al., 1992, pp. 4, 26). Similarly, Tinto (1993) emphasized, "In the interactive system of a college, almost any institutional action, whether in admissions, counseling, advising, academic programs and classrooms, or student life, will continually affect student persistence and will do so in often unintended and quite unexpected ways" (p. 205). Continuous evaluation of needs and services produces a model finely tuned to assist learners achieve their goals: "High performing institutions consistently and aggressively reach out to students by assessing needs, adjusting programs and services, and maintaining a powerful client orientation. . . . Student success is the ultimate criterion for institutional effectiveness" (Alfred, et al., 1992, pp. 16, 24). As Barr and Tagg (1995) summarized, "The Learning Paradigm envisions the institution itself as a learner--over time, it continuously learns how to produce more learning with each graduate class, each entering student" (p. 14). B.C.'s strategic plan for its colleges and institutes picks up on the themes from the educational reform literature, making a firm commitment to an outcome-oriented, learner-centred system (B.C. Ministry of Education, Skills and Training, 1996a). Some of the components of a reformed system, identified in the literature, are discussed below in more detail and related to findings of the current study.

Characteristics of a reformed education system.

1. In a learner-centred system, the ultimate agenda is ensuring that students' individual goals are identified and addressed. Success looks somewhat different for each learner, depending on what she or he wants to achieve: "Defining outcomes too narrowly or rigidly misses the point. . . . And assessing college performance without information about student goals is generally misleading. . . . Student success is in the eye of the beholder--it all depends on what is important" (Alfred et al., 1992, pp. 5, 7). Success and achievement of individual goals may or may not include completing a credential: "Institutions should not define dropout in ways which contradict the students' own understanding of their leaving. If the leaver does not define his/her own behavior as representing a form of failure, neither should the institution" (Tinto, 1993, p. 141). This holds true for the women in the current study, for whom personal growth and transformation are most critical, although these students are also solidly committed to the goal of completing their programs and attaining credentials.

2. Active outreach to students ensures that their needs are addressed. Efforts are made to connect with students proactively instead of waiting until they cry out for help or drop out. Supports are integrated into the educational whole, not compartmentalized or marginalized. Students are viewed holistically and their personal, as well as academic, needs are considered in every academic and nonacademic interaction with them. This approach is essential for mature women such as those interviewed, whose complex personal lives are fully interconnected with their lives as students and who do not have time to seek out supports that are not integrated into their educational program.

3. Students are "recognized for what they already know" (Hall & Shiffman, 1996, p. 1), and "prior learning assessment is not a separate activity oriented to the subject credit; it is

woven into the student's learning experience" (Association of Colleges of Applied Arts and Technology of Ontario, 1996, p. 6). "Assessment and learning are part of a continuous process that enables learners to monitor their own progress and set new targets" (Simosko, 1995, p. 3). Few of the women in the current study had been provided with formal opportunities to receive credit for prior knowledge, but those few were positive about the experience and felt there should be more such opportunities. Otherwise, however, evaluation processes were perceived as a major stressor and intrusion on learning for many women, rather than part of a continuous, self-directed cycle of learning and assessment.

4. A key and related component of the new paradigm is improved transferability of courses and credentials between institutions so that learners are not faced with the costly and time-consuming need to repeat courses unnecessarily. For example, two major reports have recently been distributed on this topic in B.C., one addressing issues of block transfer between institutions (B.C. Council on Admissions and Transfers, 1997), the other addressing transfer between private and public post-secondary institutions (B.C. Ministry of Education, Skills and Training, 1997c). A third recent report also highlights the importance of improved articulation of educational opportunities in the province: "Training providers need to facilitate rather than complicate processes by which British Columbians can move through education and training systems" (Gallagher, Sweet & Rollins, 1997, p. 10). For the women in the current study, transfer between institutions did not emerge as a major concern, likely because personal circumstances would make relocating difficult and because their educational laddering aspirations can be satisfied within the college they are attending. Even within this one institution, however, some students did express frustration regarding

determining and meeting prerequisite requirements for moving from one level to the next, supporting the importance of this issue overall.

5. Faculty shift roles under the new paradigm, from lecturer to mentor, guide, and coach. Further, everyone in the environment assumes responsibility for student learning and for responding to students as customers (Alfred et al., 1992; Association of Colleges of Applied Arts and Technology of Ontario, 1995; Barr & Tagg, 1995; Simosko, 1995). As Barr and Tagg summarized, under traditional instruction-oriented paradigms "Faculty classify and sort students. . . . Under the learning paradigm, faculty--and everybody else in the institution--are unambiguously committed to each student's success . . . faculty find ways to develop every student's vast talents and clear the way for every student's success" (p. 23). It seems that most of the faculty in the current study have "stepped down from the podium"; a few have not. Overall, women perceive the environment within the department as demonstrating commitment to them and their learning. As has been discussed, however, even within that supportive environment some students feel that their needs are not being met, a finding that highlights the importance of sensitive and individualized outreach to each learner.

6. The key criterion for program curriculum and delivery under the new paradigm is flexibility. Curriculum is experiential and allows room for learner contribution to both content and delivery of courses. Courses are modularized to prevent learners from having to "sit through" content with which they are already familiar: "Curriculum will be adaptive rather than dictated in form and fashion. . . . Curricula will be arranged in segments or modules, and organized around a predetermined set of competencies" (Hall & Shiffman, 1996, p. 3). Self-directed learning is a valued part of every educational program and every course: "The subject is as long as needed, not as long as the semester. . . . There are fewer

scheduled class hours; students use the college's learning centres at any time of the day and any time of the week. . . . Materials and resources to assist the learner in acquiring learning outcomes are available when the learner requires them through a variety of media" (Association of Colleges of Applied Arts and Technology of Ontario, 1995, p. 6). The B.C. college and institute strategic plan defines flexibility as "providing customized curriculum and learning opportunities anytime and anywhere, including the workplace, the community, and the home" (B.C. Ministry of Education, Skills and Training, 1996a, p. 2). Educators and educational institutions in B.C. are only beginning to explore more flexible, learner-centred delivery approaches; the programs under study are similarly still dominated by "old paradigm" approaches. There is no doubt that for mature women (like those interviewed) who are juggling myriad responsibilities, successful post-secondary participation could be facilitated through access to modularized, self-paced, and self-directed learning, allowing them to segment and schedule learning conveniently according to the other demands of their lives.

7. Critical to the new paradigm is constant evaluation and adjustment of every aspect of programming, to ensure that all stakeholders (students, employers, and so on) are satisfied and that desired outcomes are attained. Both quantitative and qualitative measures must be included, and a commitment to making changes based on results is essential (Alfred et al., 1992; Association of Colleges of Applied Arts and Technology of Ontario, 1996; Barr & Tagg, 1995; Tinto, 1993). Barr and Tagg summarized what is needed:

> The key structure for changing the rest of the system is an institution-wide assessment and information system--an essential structure in the Learning Paradigm and a key means for getting there. It would provide constant, useful feedback on institutional performance. It would track transfer, graduation, and other completion rates. It would track the flow of students through learning stages . . . and the development of in-depth

> knowledge in a discipline. It would measure the knowledge and skills of program completers and graduates. It would assess learning along many dimensions and in many places and stages in each student's college experience. (Barr & Tagg, 1995, p. 20)

The Province of British Columbia, and the institution and department involved in the current study, make commendable efforts toward evaluating their activities, as discussed above. From the perspective of some mature students, however, there remains a need for even more attention to students' input.

Reform-related recommendations. The Ministry's college and institute strategic plan poses the following questions:

> How can the system effectively respond to the learning needs of an increasingly heterogeneous population and remove barriers which have inhibited access for many other potential learners, especially those from equity-seeking groups? . . . How can learners receive the appropriate preparatory, developmental and other learning supports to ensure retention and completion of their learning goals? (B.C. Ministry of Education, Skills and Training, 1996a, pp. 20, 21)

Responses are suggested in the literature cited above. Further, the participants in the current study have contributed their experiences and ideas to the discussion and pointed toward potential solutions (at least for them and other women in like situations). Specific reform-related recommendations, linked to the characteristics identified above and derived from the input of the women interviewed, follow. It must be emphasized that these recommendations are intended to address the needs of the learner group which was the focus of this inquiry, and may not be applicable to other learners.

To move toward a learner-centred, success-oriented educational environment for learners like these, institutions and their personnel need to:

1. Take time early on to identify each individual students' unique goals and priorities. Ask students what the indicators of success will be for them and demonstrate commitment to

individuals by assisting them to achieve success as they define it, even if that does not match with institutional, program, or faculty expectations.

2. Remember the importance of women's personal lives and that women are being true to human service values when they place family commitments before studies. Offer personal and academic supports proactively, recognizing that many students lack the confidence or simply do not have time to ask for help, or may not realize that support is available.

3. Provide opportunities in each course for students to indicate what elements of the content they know already, what elements are priorities, and what needs to be added. Take every opportunity to recognize the skills and knowledge mature students bring with them, through formal and informal assessment of prior learning and offering opportunities for students to contribute to the learning of their classmates. Maximize the experiential aspect of each class and tailor each to students' preexisting competencies as well as their needs and interests.

4. Recognize the importance for busy and financially strapped students of ensuring seamless movement from one credential to the next, or between programs or institutions, in order to avoid wastage of time or repetition of coursework. Ensure that students are getting sound advice regarding course prerequisites and transfer issues and make transfer as painless as possible for incoming students. Develop policies and procedures that welcome the return to education of students who have stopped out (whether from the same or another institution or program).

5. Make a conscious effort to mentor and support students, involve everyone in the learning environment in ensuring that each student is successful, and shift the focus from instruction to the facilitation of learning. Look for ways to maximize student-directed

learning, rather than assuming class attendance will be the primary learning vehicle. Recognize that changing the role of faculty will require effort, peer and administrative support, and access to appropriate professional development opportunities.

6. Under the guidance of students, carefully examine every program and course rule, policy, and procedure to ensure that flexibility for students is enhanced. Consider the unique needs of students from a variety of family, economic and cultural backgrounds throughout this exercise. Given the contradictory needs and preferences of individual students, seek approaches that provide options and maximize flexibility.

7. Provide regular, ongoing opportunities for student input to all aspects of programming, using a variety of approaches (such as class discussions, focus groups, surveys) to ensure all learners are able to contribute in a manner comfortable for them. For example, allow student input to how they will be evaluated and ask for feedback on each evaluative exercise to measure whether it encouraged and accurately assessed their learning and was worth the time consumed.

Although these recommendations are directed at institutions and their personnel, there are parallel recommendations that can be directed at government. Clearly, for each of the recommendations identified above, government needs to fulfill its role in provision of leadership, coordination, and funding in order to promote the necessary changes. For example, steps that government can take in support of educational reform include: modelling inclusion of the student voice; emphasizing qualitative as well as quantitative outcome measures, taking students' goals into consideration; putting greater emphasis on provision of supports for nontraditional students; facilitating movement to a learner-centred system by supporting professional development opportunities, funding innovative pilots, cataloguing best practices, and so forth. Although the changes needed will actually occur at the level of individual programs and personnel, government can play a positive role in promoting a climate for such change to take place.

Further, government has a definite role in recognizing and addressing the financial barriers that emerged as pivotal for these students and others like them. B.C. is taking the lead nationally in promoting improvement to the Canada student loan system, but this is only one of many steps that need to be taken. For example, access to financial aid for married women needs to be reviewed and the situation of former income assistance clients on student loans needs to be revisited, to ensure that students who experience costly medical crises or who run out of funds before the end of the year are not at risk of dropping out, reverting to income assistance, and potentially defaulting on their loans.

The women who participated in this inquiry are fortunate to be studying in a department that is beginning to apply the tenets of educational reform, prides itself on being learner-centred and is committed to continual evaluation and improvement. These students generally feel supported, recognize just how fortunate they are, and are strongly committed to achieving their educational goals. Nevertheless, they are frequently struggling on many fronts, and their success--as they define it, in transformative and holistic terms--is sometimes diminished by factors outside their control. They are articulate and extremely self-aware. They can in many ways be viewed as key informants in an educational reform pilot. As such, and as able spokespersons for other women in circumstances similar to their own, their insights and impressions are invaluable.

Implications for Future Research

The narrow scope of the current inquiry raises many possibilities for future research. A few possibilities, derived from questions and issues that have intrigued the researcher, will be highlighted.

First, the impact of the students' field of study on the results of this inquiry is intriguing and warrants investigation. Human service program content is unique in its relevance to personal and family life: Child development, ethical behaviour, relationships, and so on, are all content areas germane to every adult. Therefore, it is reasonable to question whether the close interconnection between school and personal lives, observed here, also occurs (and with the same intensity) for mature students in other disciplines. Is an integrated approach to students as necessary, or a holistic definition of success as prevalent, in disciplines where the course content is not so relevant to personal life? Further, will students in programs that place less emphasis on examining "self" be as forthright in disclosing their thoughts and concerns in an open interview format such as was used here, or would a different methodology be required?

Second, some tentative patterns emerged from the data and warrant further investigation. Trends that were suggested but not confirmed were that part-time students feel less attached to their peers or faculty, and place less emphasis on social integration, than full-time students; that students in closed enrollment programs on small campuses place greater emphasis on social integration than others; and that students' commitment to transformative learning and ability to articulate their goals and perceptions grow during their years in the human service programs studied. Further research would be needed to assess these hypotheses. Further, some patterns that might have been expected did not emerge from the data and warrant investigation in future studies. For example, response patterns in this study did not suggest variation by program area or intended level of credential (certificate, diploma, or degree). Additional investigation would be required to explore the relevance to the findings of participants' specific discipline or intended credential.

Third, a similar research methodology applied to human service students in other institutions could highlight institutional influences on mature students. To what extent were the findings related to the institutional context in which this particular student group was studying?

Fourth, although group interviews were envisaged as part of this inquiry, they did not occur. This approach would be particularly beneficial in research focusing on the changes needed to better promote success. Exploring this issue with students in groups would not only generate productive discussions and solutions, it would also provide an opportunity for growth and learning for the participants, as a result of the research process.

Fifth, a longitudinal study of these students or others like them would be valuable, to determine whether their long-term commitments to the human service field and their positive assessment of the value of their education (despite the debt load associated with it for some) endure over time. Although attrition from human service education programs appears to be low, turnover in the field is high: How do retention rates (in specific jobs and in the field overall) for workers as committed and well-prepared as these compare with the average for all workers?

Finally, the province should parallel its annual survey of program completers with a survey of those who do not complete. This would balance the qualitative institution and program-specific data recommended above with quantitative, provincial data, prompting an exploration on a provincial and program-specific basis of such critical issues as why students do not complete programs, whether their goals were met despite non-completion, and how the system and individual institutions can better address concerns that may have contributed to their early departure.

Conclusions

My own thinking has been influenced through becoming acquainted with these women and reflecting on their forthright disclosures. Although I have not taught in the college system for several years, I do have considerable experience working closely (as instructor and program head) with human service students similar to those interviewed. Despite this, I underestimated the participants and was only superficially aware of the issues facing them. The revelations I experienced are outlined below.

First, I launched the research with some trepidation regarding the motivation and perspectives on success that would be revealed by participants. I hoped (because of my own commitment to transformative education) that they would be pursuing transformative goals and definitions of success, but did not have complete confidence that this would consistently be the case. My doubts were clearly unjustified.

Second, the interviews affirmed the commonalties between these women and myself, which I expected, but also highlighted and gave me a clearer understanding of our differences. Aspects of my own background made it easy to relate to the participants in a number of key areas: My history as a single parent with young children; more recent experience as parent of challenging teenagers; experience with juggling work, parenting, and school; and exposure to the impacts of age on educational participation and performance, all provided common ground with these women. However, more striking were the areas where I could empathize, but not relate from my own personal experience. Many of these women came from family backgrounds where educational achievement was uncommon and post-secondary participation rare or wholly unknown. This is the opposite to my family situation, where university participation was an expectation. More significantly, my own struggles with single parenthood and the trials of balancing school with other responsibilities were never endured in poverty. I have extensive indirect experience with so-called "disadvantaged groups", but was not prepared for the extreme poverty-related challenges facing many of these women. Further, I was struck by the matter-of-fact courage with which they cope.

Third, I began this research with a conviction that traditional student success models, and even models developed with particular focus on nontraditional students, were not adequate or accurate for mature women in the human services. However, I was not prepared for the profound nature of the differences. These have been highlighted throughout this paper. Primary was the fact that dropping out was not even a real option for these women. I was also not prepared for the recognition that in many cases the most effective support that could be provided a mature female student would be financial: Alleviating financial pressures (and the need to work several jobs, inability to buy texts, stress related to family financial management, and so on, which poverty entails) would eliminate many of the barriers limiting women's ability to achieve success as they define it.

Fourth, and related to the above point, was the realization that I have been viewing mature students as students first and foremost, encumbered with external demands potentially impeding their progress. I believe that as an instructor I was sensitive to the pressures of women's family and other commitments, though unaware of the intensity of these pressures. However, I viewed mature learners as experiencing more challenges than other students, but otherwise qualitatively the same. I did not realize that their goals and definitions of success included satisfaction and balance in home, job and community as well as academic spheres. I did not recognize these students as representing their family and social unit as they met with me or sat in my classes, persisting in order to improve the lot of their children, achieving on behalf of family members who had never had the opportunity to go to college, pursuing knowledge and skills to benefit them and their families personally (as well as for application in the work-place). I did not recognize these distinguishing traits as qualitatively different from many traditional students.

Finally, I have realized that as democratic as I would like to be, I have never as instructor, administrator, or policy maker genuinely "stepped down from the podium" to give mature learners control of their learning. I have unthinkingly gone along with the traditional beliefs that government, institutions, and faculty know best and that learners' voices, although they should be heard, can often be discounted. These women and their comments are not equally articulate or sophisticated, but their wisdom and perceptiveness are consistent and surpassed my own prior insight into their needs, as well as my expectations.

I feel that I have come full circle in my career as an educator. I began, over 20 years ago in Special Education, with a particular interest in teaching those who have the most difficulty learning, those students labeled "mentally handicapped". I had the good fortune to participate in an intense post-baccalaureate program enrolling only eight students, with an instructor who provided the kind of holistic learner-centred teaching and mentoring that emerges in current literature as if newly discovered. The philosophy of the instructor, the program, and favourite theorists in the field at that time foreshadowed current human service approaches, with an emphasis on relationship between educators and individual learners, respect for diversity, and a belief in the potential and value of every human being, no matter how challenged by circumstances or personal limitations.

At that time, a medical approach to disabilities was being replaced by a commitment to supporting individuals to overcome their disabilities and live full, rich lives. The challenge to educators and other care providers was to find ways to help learners reach their potential. The basic premise was that everyone can learn. If students are not progressing it is because we have not yet designed the necessary accommodations or figured out how to facilitate their learning, not because of innate limitations: "A lack of learning in any particular situation should first be interpreted as an inappropriate or insufficient use of teaching strategy, rather than an inability on the part of the learner" (Gold, 1980, p. 3). Directive and mechanistic behavioural approaches certainly continued to be popular, but at the same time there was a rejection of operant techniques in favour of approaches that respected and empowered even the most challenged of learners.

I have found throughout my career that the educational field overall could take many lessons from Special Education. This makes some sense in that special educators have had to refine and perfect planning, teaching, and evaluation techniques in order to experience success. Unlike most learners, students with cognitive disabilities, in particular, are generally unlikely to learn in the absence of good teaching. Many of the thrusts of current educational reform have been recognized as best practice for years in working with students with cognitive or multiple disabilities, such as recognition of the need for constant evaluation of student outcomes as a guide for educational program design.

The results of this inquiry promote the philosophy that everyone can learn. At least for the students in this inquiry, every one can be successful in meeting her goals and it is up to the post-secondary system and all its component parts to facilitate (or at the very least, avoid encumbering) that process.

The difference from Special Education is that the mature learners interviewed here can do much more to assist in designing the educational environments they need. However, the high degree of perceptiveness and responsiveness required to work effectively with students who have limited communication skills also has a parallel here. Many mature women, even though they clearly know what they need and how they could be assisted, are lacking in self-confidence and are intimidated by the post-secondary bureaucracy. Discovering how to help them be successful requires attentiveness, sensitivity, and proactivity in order to elicit their views, respond to the cues they provide, and support them in taking control of their educational experiences.

Success for these women is both transformative and ecologically based: It encompasses true learning and achievement of personal goals, assessed in the context of their personal lives. Satisfaction with the learning experience will not occur if satisfactory work, family and community lives are not maintained. But these women want it all--all of this plus good grades. To achieve such demanding goals, they do not need to be sorted into programs--they know where they want to go. They do not need to be told what supports they need--they generally know that too, but cannot participate in supports that are not tailored to them. They do not need help in becoming connected--they are already overburdened with connections and the commitments resulting from these connections. What they do need from us as post-secondary educators, administrators and policy makers is that we examine every requirement we put upon them, asking if it is necessary and valuable for their learning, worth their time, and sensitive to the realities of their lives. They need us to evaluate our performance as reflected in their realization of personal goals, and over all, to shift our attention from teaching to learning and from systemic policies to person-centred flexibility. Finally, they need us to remember that being person- or learner-centred means, for these women, being family-centred. Student success for this population, so enmeshed with their families, goes beyond even transformative learning to encompass positive family outcomes.

Success for this student group truly requires redefinition. However what is required is not that educators struggle with the task of redefinition, but that learners be supported to define their own goals, needs, and indicators of success and that educators and policy makers listen, learn, and respond.

REFERENCES

Acker, J., Barry, K., & Esseveld, J. (1983). Objectivity and truth: Problems in doing feminist research. *Women's Studies International Forum, 6(4),* 423-435.

Aitken, N. (1982). College student performance, satisfaction and retention: Specification and estimation of a structural model. *Journal of Higher Education, 53(1),* 32-50.

Alfred, R., Peterson, R., & White, T. (1992). *Making community colleges more effective: Leading through student success* (Community College Consortium). Ann Arbor, MI: University of Michigan.

Allen, B. (1995). The student in higher education: Nontraditional student retention. *Catalyst [On-line serial], 23(3).*

Anderson, E. (1985). Forces influencing student persistence and achievement. In L. Noel, R. Levitz, D. Saluri, & Associates (Eds.), *Increasing student retention* (pp. 44-61). San Francisco: Jossey-Bass.

Anisef, P. (1989). Studying part-time in Canada's universities: A social change perspective. *Canadian Journal of Higher Education, 12(1),* 11-26.

Ashar, H. & Skenes, R. (1993). Can Tinto's student departure model be applied to nontraditional students? *Adult Education Quarterly, 43(2),* 90-100.

Association of Colleges of Applied Arts and Technology of Ontario. (1995). *Learning centred education in Ontario's colleges: A white paper prepared for the Council of Presidents.* Toronto, ON: Author.

Association of Colleges of Applied Arts and Technology of Ontario. (1996). *Accountability in a learning centred environment: Using performance indicators in reporting, advocacy and planning.* Toronto, ON: Author.

Astin, A. (1984). Student involvement: A developmental theory of higher education. *Journal of College Student Personnel, 25,* 297-308.

Atkinson, P. and Hammersley, M. (1994). Ethnography and participant observation. In N. Denzin & Y. Lincoln (Eds.), *Handbook of qualitative research* (pp. 248-260). Thousand Oaks, CA: Sage.

Attinasi, L. (1989). Getting in: Mexican Americans' perceptions of university attendance and the implications for freshman year persistence. *Journal of Higher Education, 60(3),* 247-277.

Baker, T., & Velez, W. (1996). Access to and opportunity in postsecondary education in the United States: A review. *Sociology of Education,* 82-101.

Ball, W. (1992). Critical Social Research, adult education and anti-racist feminist praxis. *Studies in the Education of Adults, 24(1),* 1-25.

Bank, B., Biddle, B., & Slavings, R. (1992). What do students want? Expectations and undergraduate persistence. *The Sociological Quarterly, 33(3),* 321-335.

Bank, B., Slavings, R., & Biddle, B. (1990). Effects of peer, faculty, and parental influences on students' persistence. *Sociology of Education, 63,* 208-225.

Barr, R., & Tagg, J. (1995, November). From teaching to learning: A new paradigm for undergraduate education. *Change,* 13-5.

Bates, L. (1994, October 22). Numbers-challenged bean counters double the high school drop-out rate. *Vancouver Sun*, p. B1.

Bean, J. (1980). Dropouts and turnover: The synthesis and test of a causal model of student attrition. *Research in Higher Education, 12(2),* 155-187.

Bean, J. (1990a). Using retention research in enrollment management. In D. Hossler, J. Bean, & Associates (Eds.), *The strategic management of college enrollments* (pp. 170-185). San Francisco: Jossey-Bass.

Bean, J. (1990b). Why students leave: Insights from research. In D. Hossler, J. Bean, & Associates (Eds.), *The strategic management of college enrollments* (pp. 147-169). San Francisco: Jossey-Bass

Bean, J., & Metzner, B. (1985). A conceptual model of nontraditional undergraduate student attrition. *Review of Educational Research, 55(4),* 485-540.

Beatty-Guenter, P. (1994). *Retention strategies at community colleges: Diffusion, legitimacy and effectiveness.* Unpublished doctoral dissertation, University of California, Berkeley.

Beatty-Guenter, P. (1995). *Student retention at B.C. community colleges.* Victoria, BC: B.C. Council on Admissions and Transfer.

Beauchesne, E. (1997, September 26). Cuts force students to pay more. *Vancouver Sun,* p. A4.

Belanger, R., Lynd, D., & Mouelhi, M. (1982). *Part-time degree students: Tomorrow's majority.* Ottawa: Statistics Canada.

Blaxter, L., Dodd, K., & Tight, M. (1996). Mature student markets: An institutional case study. *Higher Education, 31,* 187-203.

British Columbia Council on Admissions and Transfer. (1992). *Student success/completion: A discussion paper.* Victoria, BC: Author.

British Columbia Council on Admissions and Transfer. (1997). *Block transfer: Issues and options.* Vancouver, BC: Author

British Columbia Ministry of Advanced Education, Training and Technology. (1991). *Forces of change influencing education and training.* Victoria, BC: Author.

British Columbia Ministry of Education, Skills and Training. (1996a). *Charting a new course: A strategic plan for the future of British Columbia's college, institute and agency system.* Victoria, BC: Author.

British Columbia Ministry of Education, Skills and Training. (1996b). *1995 B.C. college and institute student outcomes.* Victoria, BC: Author.

British Columbia Ministry of Education, Skills and Training. (1996c). *1995 B.C. student outcomes: Survey results of former college and institute aboriginal students.* Victoria, BC: Author.

British Columbia Ministry of Education, Skills and Training. (1996d). *Student retention in British Columbia.* Unpublished manuscript.

British Columbia Ministry of Education, Skills and Training. (1997a). *Annual report 1995/96.* Victoria, BC: Author.

British Columbia Ministry of Education, Skills and Training. (1997b). *Moving on: Secondary to post-secondary transition. A report to the Minister of Education, Skills and Training.* Victoria, BC: Author.

British Columbia Ministry of Education, Skills and Training. (1997c). *Report of the working committee on public-private articulation agreements.* Victoria, BC: Centre for Curriculum, Transfer and Technology.

British Columbia Ministry of Skills, Training and Labour. (1995). *Student retention in British Columbia.* Unpublished manuscript.

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design.* Boston: Harvard University Press.

Brower, A. (1992). The "Second Half" of student integration: The effects of life task predominance on student persistence. *Journal of Higher Education, 63(4),* 441-462.

Burgess, R. (1984). *In the field: An introduction to field research.* London: Unwin Hyman.

Byrne, A. (1989). The process of adult socialization to higher education. *Equity and Excellence, 24(3),* 9-10.

Cabrera, A., Castaneda, M., Nora, A., & Hengstler, D. (1992). The convergence between two theories of college persistence. *Journal of Higher Education, 63(2),* 143-164.

Caffarella, R., & Olson, S. (1993). Psychosocial development of women: A critical review of the literature. *Adult Education Quarterly, 43(4),* 227-242.

Camosun College (1993). *Learning skills centre: August 1992 and Fall 1992 follow-up.* Unpublished manuscript.

Campbell, J. (1993). Community social services project: Training report. In J. Korbin, *Commission of inquiry into the public service and public sector: Background paper--contracted community social services* (pp. A1-A19). Victoria, BC: Crown Publications.

Chartrand, J. (1992, December). Nontraditional students: Dropping the prefix, redefining "traditional". *The Teaching Professor,* 8.

Cooper, C. (1993). *Factors affecting first year student retention at university.* Unpublished manuscript.

Cooper, C., & Bradshaw, R. (1984). How green is your academic climate? Check it out with MOSS: A monitor of student satisfaction. *College and University, 59(3),* 251-260.

Crampton, B., & Holm, K. (1993). A learning-community program for transfer and returning students. *Journal of College Student Development, 34,* 439-440.

Crosson, P. (1992). Environmental influences on minority degree attainment. *Equity and Excellence, 25(2-4),* 5-15.

Darkenwald, B., & Gavin, W. (1987). Dropout as a function of discrepencies between expectations and actual experiences of the classroom social environment. *Adult Education Quarterly, 37(3),* 152-163.

Devault, M. (1990). Talking and listening from women's standpoint: Feminist strategies for interviewing and analysis. *Social Problems, 37(1),* 96-116.

Devlin, L. (1989). Part-time students: A closer look. *Canadian Journal of University Continuing Education, 15(1),* 29-38.

Devlin, L., & Gallagher, P. (1982). Age-related needs of community college students. *Canadian Journal of Higher Education, 12(2),* 33-42.

Dey, E., & Hurtado, S. (1995). College impact, student impact: A reconsideration of the role of students within American higher education. *Higher Education, 30,* 207-223.

Dietsche, P. (1990). Freshman attrition in a College of Applied Arts and Technology of Ontario. *Canadian Journal of Higher Education, 20(3),* 65-84.

Erickson, D. (1993). Contracted community social services in B.C.: A technical report. In J. Korbin, *Commission of Inquiry into the Public Sector: Background paper--contracted community social services* (pp. B1-B35). Victoria, BC: Crown Publications.

Faris, R. (1995). *Major reforms in training systems in three countries* (submitted to the B.C. Ministry of Education, Skills and Training). Unpublished manuscript.

Fielding, N., & Lee, R. (1991). *Using computers in qualitative research.* Newbury, CA: Sage.

Fiorentine, R. (1987). Men, women, and the pre-med persistence gap: A normative alternatives approach. *American Journal of Sociology, 92(5),* 1118-1139.

Gallagher, P., Sweet, R., & Rollins, R. (1997). *Intermediate skill development in British Columbia: New policy and research directions.* Vancouver, BC: Archetype Print Limited.

Gardner, J. (1978, Summer). How to make students more effective consumers of their education: University 101 at the University of South Carolina. *Southern College Personnel Association Journal,* 61-68.

Gardner, J. (1981). Developing faculty as facilitators and mentors. *New directions for student services, 14,* 67-80.

Gardner, J. (1992). *Model elements of campus approaches to enhancing the first year experience.* Paper presented at the First Year Experience Canadian-American Conference, Victoria, BC.

Garland, M. (1993). Ethnography penetrates the "I didn't have time" rationale to elucidate higher order reasons for distance education withdrawal. *Research in Distance Education,* 6-10.

Garrison, D. (1985). Predicting dropout in Adult Basic Education using interaction effects among school and nonschool variables. *Adult Education Quarterly, 36(1),* 25-38.

Garrison, D. (1987). Dropout prediction within a broad psychosocial context: An analysis of Boshier's congruence model. *Adult Education Quarterly, 7(4),* 212-222.

Gass, M. (1990). The longitudinal effects of an adventure orientation program on the retention of students. *Journal of College Student Development, 31,* 33-38.

Gilbert, S., & Auger, M. (1988). *Student finances and university attrition (PCS-6-00173).* Ottawa, ON: Department of the Secretary of State of Canada.

Gold, M. (1980). *Try another way training manual.* Champaign, IL: Research Press.

Gomme, I., & Gilbert, S. (1984). Paying the cost: Some observations on the problem of postsecondary student attrition. *Canadian Journal of Higher Education, 14(3),* 95-100.

Gomme, I., Hall, M., & Murphy, T. (1993). In the shadow of the tower: The view of the undergraduate experience. *Canadian Journal of Higher Education, 23(3),* 18-35.

Gordon, V., & Grites, T. (1984). The freshman seminar course: Helping students succeed. *Journal of College Student Personnel, 25,* 315-320.

Guppy, N.,& Bednarski, V. (1993). *Enhancing student retention in higher education: A literature review.* Victoria, BC: B.C. Council on Admissions and Transfer.

Hall, J., & Shiffman, P. (1996). Embracing change: Evolving to the campus of the future. *New Horizons in Adult Education, 10(1),* 1-4.

Hatcher, L., Kryter, K., Prus, J., & Fitzgerald, V. (1992). Predicting college student satisfaction from investment model constructs. *Journal of Applied Social Psychology, 22(16),* 1273-1296.

Hodgson, T. (1989). Support services for adult learners in higher education. *Equity and Excellence, 24(3),* 20-24.

Holdoway, E. & Kellowy, K. (1987). First year at university: Perceptions and experiences of students. *Canadian Journal of Higher Education*, 17(1), 47-63.

James, T. (1995). Quo Vadis? Douglas College and the paradoxes of change. *Discourse 1(1)*, New Westminster, BC: Douglas College.

Jewell, L., & Lupin, B. (1988). Effectiveness of peers who re-entered college in a program to enhance retention. *Psychology Reports, 63,* 921-922.

Johnson, G., & Buck, G. (1995). Students' personal and academic attributions of university withdrawal. *Canadian Journal of Higher Education, 25(2),* 53-77.

Justice, D. (1997, Winter). Facilitating adult learning in a liberal education context. *Liberal Education,* 28-33.

Kaplan, S. (1992, June/July). Providing academic options for nontraditional students. *American Association for Colleges and Junior Colleges (AACJC) Journal,* 6-7.

Kirby, S., & McKenna, K. (1989). *Experience, research, social change: Methods from the margins.* Toronto, ON: Garamond Press.

Langenback, M., & Korhonen, L. (1988). Persisters and non-persisters in a graduate level, nontraditional, liberal education program. *Adult Education Quarterly, 38(3),* 136-148.

Lather, P. (1986). Research as praxis. *Harvard Educational Review, 56(3),* 257-277.

Lather, P. (1991). *Getting smart: Feminist research and pedagogy with/in the postmodern.* New York: Routledge.

Lee, V., & Frank, K. (1990). Students' characteristics that facilitate the transfer from two-year to four-year colleges. *Sociology of Education, 63,* 178-193.

Levin, J., & Levin, M. (1993). Methodological problems in research on academic retention programs for at-risk minority college students. *Journal of College Student Personnel, 34,* 119-124.

Levitz, R., & Noel, L. (1985). Using a systematic approach to assessing retention needs. In L. Noel, R. Levitz, D. Saluri, & Associates (Eds.), *Increasing student retention* (pp. 345-361). San Francisco: Jossey-Bass.

Lewallen, W. (1993). The impact of being "undecided" on college student persistence. *Journal of College Student Development, 34,* 103-12.

Lewis, D., Hearn, J., & Zilbert, E. (1993). Efficiency and equity effects of vocationally focused postsecondary education. *Sociology of Education, 66,* 188-205.

Manak, S. (1995). *Racism and other issues affecting students of colour and new immigrants in post-secondary education* (submitted to the B.C. Ministry of Skills, Training and Labour). Unpublished manuscript.

May, K. (1989). Interview techniques in qualitative research: Concerns and challenges. In J. Morse (Ed.), *Qualitative nursing research: A contemporary dialogue* (pp. 171-182). Chicago: Aspen.

McJamerson, E. (1992). Undergraduate academic major and minority student persistence: Individual choices, national consequences. *Equity and Excellence, 25(2-4),* 35-48.

McKeown, B., Macdonell, A., & Bowman, C. (1993). The point of view of the student in attrition research. *Canadian Journal of Higher Education, 23(2),* 65-85.

Merriam, S., & Clark, M. (1991). *Patterns of work, love and learning in adulthood.* San Francisco: Jossey-Bass.

Metzner, B. (1989). Perceived quality of academic advising: The effect on freshman attrition. *American Educational Research Journal, 26(3),* 422-442.

Mezirow, J. (1991). *Transformative dimensions of adult learning.* San Francisco: Jossey-Bass.

Miller, M., & Daloz, L. (1992). Assessment of prior learning: Good practices assure congruity between work and education. *Equity and Excellence, 25(2-4),* 30-34.

Mishler, E. (1986). *Research interviewing: Context and narrative.* Cambridge, MA: Harvard University Press.

Monk-Turner, E. (1990). The occupational achievements of community and four-year college entrants. *American Sociological Review, 55,* 719-725.

Muhr, T. (1993). *ATLAS computer-aided text interpretation and theory building: User's manual.* Berlin, Germany: Thomas Muhr.

Nelson, B., Dunn, R., Griggs, S., Primavera, L., Fitzpatrick, M., Bacilious, Z., & Miller, R. (1993). Effects of learning style intervention on college students' retention and achievement. *Journal of College Student Development, 34,* 364-369.

Noel, L. (1985). Increasing student retention: New challenges and potential. In L. Noel, R. Levitz, D. Saluri, & Associates (Eds.), *Increasing student retention* (pp. 1- 27). San Francisco: Jossey-Bass.

Oakely, A. (1981). Interviewing women: A contradiction in terms. In H. Roberts (Ed.), *Doing feminist research* (pp. 30-61). Boston: Routledge & Kegan Paul.

Okin, M., Benin, M., & Brandt-Williams, A. (1996). Staying in college: Moderators of the relation between intention and institutional departure. *Journal of Higher Education, 67(5),* 577-596.

Oleson, V. (1994). Feminisms and models of qualitative research. In N. Denzin & Y. Lincoln (Eds.), *Handbook of qualitative research* (pp. 138-161). Thousand Oaks, CA: Sage.

Pappas, J., & Loring, R. (1985). Returning learners. In L. Noel, R. Levitz, D. Saluri, & Associates (Eds.), *Increasing student retention* (pp. 138-161). San Francisco: Jossey-Bass.

Pascaralla, E. (1982). *Studying student attrition.* San Francisco: Jossey-Bass.

Pascarella, E. (1986). A program for research and policy development on student persistence at the institutional level. *Journal of College Student Personnel, 27,* 100-107.

Pascarella, E., Duby, P., & Iverson, B. (1983). A test and reconceptualization of a theoretical model of college withdrawal in a commuter institution setting. *Sociology of Education, 56,* 88-100.

Pascarella, E., Terenzini, P., & Wolfe, L. (1986). Orientation to college and freshman year persistence/withdrawal decisions. *Journal of Higher Education, 57(2),* 155-174.

Payne, J. (1990). Doing adult education research. *Studies in the Education of Adults, 22(1),* 77-93.

Peck, T. (1986). Women's self-definition in adulthood: From a different perspective. *Psychology of Women Quarterly, 10,* 274-284.

Potter, J., & Alderman, T. (1992). *A profile of adult learners at the University of New Brunswick: Balancing responsibilities.* Fredericton, NB: University of New Brunswick.

Quinn, M. (1987). *Qualitative evaluation methods.* Thousand Oaks, CA: Sage.

Reinharz, S. (1992). *Feminist methods in social research.* New York: Oxford University Press.

Richards, T. & Richards, L. (1994). Using computers in qualitative research. In N. Denzin & Y. Lincoln (Eds.), *Handbook of qualitative research* (pp. 445-462). Thousand Oaks, CA:Sage.

Russell, J., & Shinkle, R. (1990). Evaluation of peer-adviser effectiveness. *Journal of College Student Development, 31,* 388-393.

Ryland, E., Riordan, R., & Brack, G. (1994). Selected characteristics of high-risk students and their enrollment persistence. *Journal of College Student Development, 35,* 54-58.

Scott, C., Burns, A., & Cooney, G. (1996). Reasons for discontinuing study: The case of mature age female students with children. *Higher Education, 31,* 233-253.

Shively, H. (1989). Opening the college to non-traditional adult learners. *Equity and Excellence, 24(3),* 11-12.

Simosko, S. (1995). Prior learning assessment and educational reform: A vision for now. Victoria, BC: Centre for Curriculum and Professional Development.

Smallwood, K. (1980). What do adult women college students really need? *Journal of College Student Personnel, 21,* 65-73.

Smith, L., Lippitt, R., Noel, L., & Sprandel, D. (1980). Mobilizing the campus for retention: An innovative quality of life model. In J. Bean & L. Noel (Eds.), *What works in student retention* (pp. 74-93). Iowa City, IA: The American College Testing Program and the National Center for Higher Education Management Systems.

Springate, G. (1991). *Identification of successful retention techniques used by community colleges in B.C.* Unpublished manuscript.

Springate, G. (1993). *Identification of reasons for student non-persistence and development of a comprehensive strategy for diploma programs at a community college in B.C.* Unpublished doctoral dissertation, Nova University, Florida.

Stage, F. (1989). Motivation, academic and social integration, and the early dropout. *American Educational Research Journal, 26,* 385-402.

Stage, F., & Rushin, P. (1993). A combined model of student predisposition to college and persistence in college. *Journal of College Student Development, 34,* 276-281.

Stake, R. (1994). Case studies. In N. Denzin & Y. Lincoln (Eds.), *Handbook of qualitative research* (pp. 236-247). Thousand Oaks, CA: Sage.

Stake, R., & Trumbull, D. (1988). Case study methods in educational research: Seeking sweet water. In R. Jaeger (Ed.), *Complementary methods for research in education* (pp. 253-265). Washington, DC: American Educational Research Association.

Stalker, J. (1993). Voluntary participation: Deconstructing the myth. *Adult Education Quarterly, 43(2),* 63-75.

Steltenpohl, E., & Shipton, J. (1986). Facilitating a successful transition to college for adults. *Journal of Higher Education, 57(6),* 637-657.

Stewart, M. (1990). Attrition from health professional schools in a Canadian university. *Canadian Journal of Higher Education, 20(3),* 43-63.

Strategic Information Research Institute--SIRI. (1994). *Student flows into and within the B.C. post-secondary system.* Victoria, B.C.: B.C. Council on Admissions and Transfer.

Suen, H. (1983). Alienation and attrition of black college students on a predominantly white campus. *Journal of College Student Personnel, 24,* 117-121.

Sugrue, C. (1996, September). Life as a mature student. *Focus on Women,* 41-47.

Thacker, C., & Novak, M. (1991). Student role supports for younger and older middle-aged women: Application of a life event model. *Canadian Journal of Higher Education, 21(1),* 13-36.

The University Presidents' Council of B.C. (1996). *Two year graduate follow-up survey of 1993 baccalaureate graduates of B.C.'s public universities: Major findings.* Victoria, BC: Author.

Thompson, G., & Devlin, L. (1992). Access by part-time students: A question of openness in Canadian universities. *Canadian Journal of Higher Education, 22(3),* 57-75.

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research, 45(1),* 89-125.

Tinto, V. (1982). Limits of theory and practice in student attrition. *Journal of Higher Education, 53(6),* 687-700.

Tinto, V. (1985). Dropping out and other forms of withdrawal from college. In L. Noel, R. Levitz, D. Saluri, & Associates (Eds.), *Increasing student retention* (pp. 28-43). San Francisco: Jossey-Bass.

Tinto, V. (1988). Stages of student departure: Reflections on the longitudinal character of student leaving. *Journal of Higher Education, 59(4),* 438-455.

Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition.* Chicago: University of Chicago Press.

Toder, F., & Hartsough, C. (1993). A collaborative approach to retaining first year undeclared students. *Journal of College Student Development, 34,* 159-160.

Tomlinson-Clarke, S. (1994). A longitudinal study of the relationship between academic comfort, occupational orientation and persistence among African American, Hispanic and White college students. *Journal of College Student Development 35,* 25-28.

Tyler, D., & Small, J. (1990). Persistence of college transfer students. *Alberta Journal of Educational Research, 36(2),* 181-188.

Ungar, S. (1980). The retention problem: An analysis of enrollment attrition at a Canadian college. *Canadian Journal of Higher Education, 10(1),* 57-74.

Unger, R. (1983). Through the looking glass: No wonderland yet! (the reciprocal relationship between methodology and models of reality). *Psychology of Women Quarterly, 8(1),* 9-32.

University College of the Cariboo--UCC. (1989). *Cariboo College attrition study* (Office of Institutional Research). Kamloops, BC: Author.

University College of the Cariboo--UCC. (1991). *Cariboo College student retention: A follow up study* (Office of Institutional Research and Evaluation). Kamloops, BC: Author.

University College of the Cariboo--UCC. (1992a). *A study on student enrollment retention and attrition, Fall 1990 to Fall 1991* (Office of Institutional Research and Evaluation). Kamloops, BC: Author.

University College of the Cariboo--UCC.(1992b). *Student retention and participation: A preliminary study of the effect of student success seminar attendance at the University College of the Cariboo* (Office of Institutional Research and Planning). Kamloops, BC: Author.

University College of the Cariboo--UCC. (1993). *A follow-up report on the impact of student success seminars* (Office of Institutional Research and Planning). Kamloops, BC: Author.

University College of the Cariboo--UCC. (1994). *A report on UCC student enrollment attrition, Fall 1991 to Winter 1993* (Office of Institutional Research.) Kamloops, BC: Author.

University of Victoria. (1995). *UVic Trends* (Institutional Analysis Office). Victoria, BC: Author.

Van Stone, N., Nelson, J., & Niemann, J. (1994). Poor single-mother college students' views on the effect of some primary sociological and psychological belief factors on their academic success. *Journal of Higher Education, 65(5),* 571-584.

Westwood, S. (1992). Power/knowledge: The politics of transformative research. *Studies in the Education of Adults, 24(2),* 191- 98.

White, W. (1983). Increase of college retention by an information systems approach to instruction. *Psychological Reports, 52,* 306.

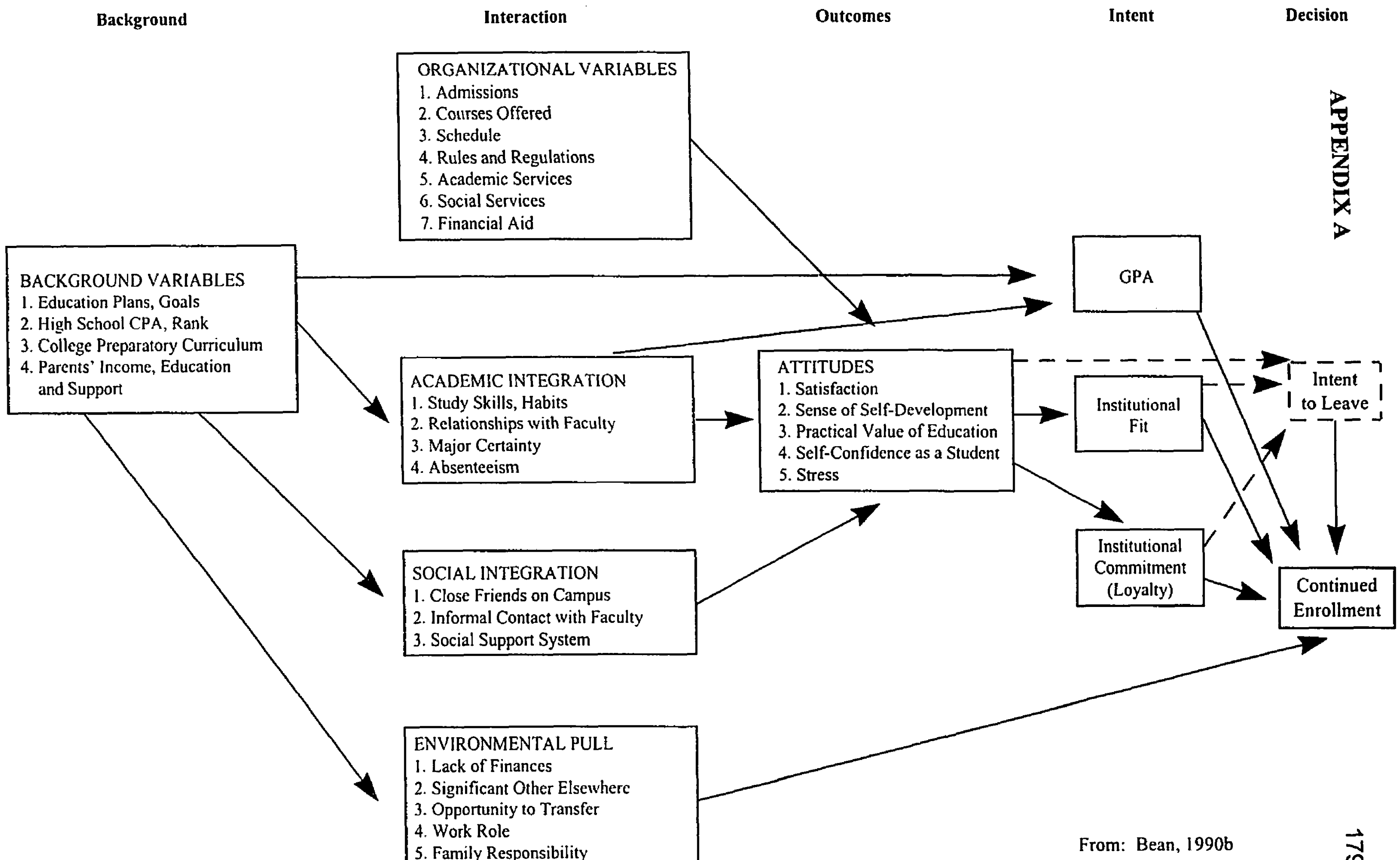
Woodley, A., Wagner, L., Slowey, M., Hamilton, M., & Fulton, O. (1987). *Choosing to learn: Adults in education.* Milton Keynes, England: Open University Press.

Young, C. (1992). Issues of access and retention for Afro-Americans in higher education. *Phylon, 44(1-2),* 55-60.

Young, R., Backer, R., & Rogers, G. (1989). The impact of early advising and scheduling on freshman success. *Journal of College Student Development, 30,* 309-312.

APPENDIX A

A Longitudinal Model of the
Type of Factors That Affect Retention Decisions



From: Bean, 1990b

APPENDIX B

Dear Human Service student;

My name is Jean Campbell, and I am a doctoral student in the School of Child and Youth Care at the University of Victoria. I need your assistance to complete my doctoral research.

My topic is post secondary student success, and my goal is to explore the concept of success from the perspective of 'nontraditional' female students in Human Service programs. Most of the research on student success has focused on traditional students coming straight out of high school, and tended to look at all students (both genders, all programs, etc.) together. Studies have relied on quantifiable measures of success (numbers of students continuing from one semester to the next, for example) and information gathered from formal institutional records or student surveys.

I want to look at success, and the factors which contribute to it, quite differently: from your perspective. Rather than gathering superficial data from a huge sample, I want to spend time talking with a small number of students so that I can understand their goals, supports, and barriers. My background and field of study is the Human Services, so I'm most interested in students in that field. In order to keep the study manageable, I've chosen to focus even further: on women (who make up the majority of students and employees in this field) and, in particular, on women defined as 'nontraditional' (in my definition, anyone *other than* a student entering college directly from high school and attending full time).

To begin, I have asked program coordinators to assist me in identifying a few students who fit this criterion and who will be able to help me understand why they are here, what they are striving for, and what's helping them or getting in their way. I want to spend about an hour with students, individually, and if they are agreeable, will tape the interview. Based on what they tell me, I'll then interview more students, individually or in small groups, to ensure that I fully understand the issues and the variations particular to individuals or subgroups within the Human Service student population. I'm requesting that you participate in this initial stage to help me make sure I'm asking the right questions, in the right way, when I approach your classmates.

I think that, by sharing your thoughts and stories, you will be able to add new and valuable information to the study of student success, and assist post secondary institutions to better meet the needs of all students. I know that your goals and the issues affecting you are different than for young students just out of high school, and I think it's important that your perspective is heard.

If you're willing to give me an hour of your time, please fill in the attached sheet and return to your instructor. Any information you provide me will be held in total confidentiality, and neither shared nor discussed with anyone at the University College. If you wish to review the transcript of the interview, you will be most welcome. I would also be pleased to have your input to the summary which I eventually pull together, after all the information gathering is complete. I will ensure that all opinions expressed are reported anonymously in surmmary documents.

Thank you for your assistance in exploring a critically important topic.

Sincerely;

Jean Campbell

APPENDIX C

STUDENT
INTERVIEW PROTOCOL

Interviews will be semi-structured; general topics will be raised for discussion, and open dialogue encouraged. An attempt will be made to include the following topics in the discussion, while keeping direct questions to a minimum:

1. **Background/History** ***(what brought you here - to college, this field, this program?)***

   - Tell me about yourself and your recent (personal, professional, educational) history

2. **Goals and Definitions of Success** ***(what are you wanting out of this and how will you know you've got it?)***

   - Tell me about your (personal, professional, educational) goals: when you started, currently, for the future

   - What does success, as a student, mean to you?

3. **Supports and Barriers** ***(what is helping and what is getting in your way?)***

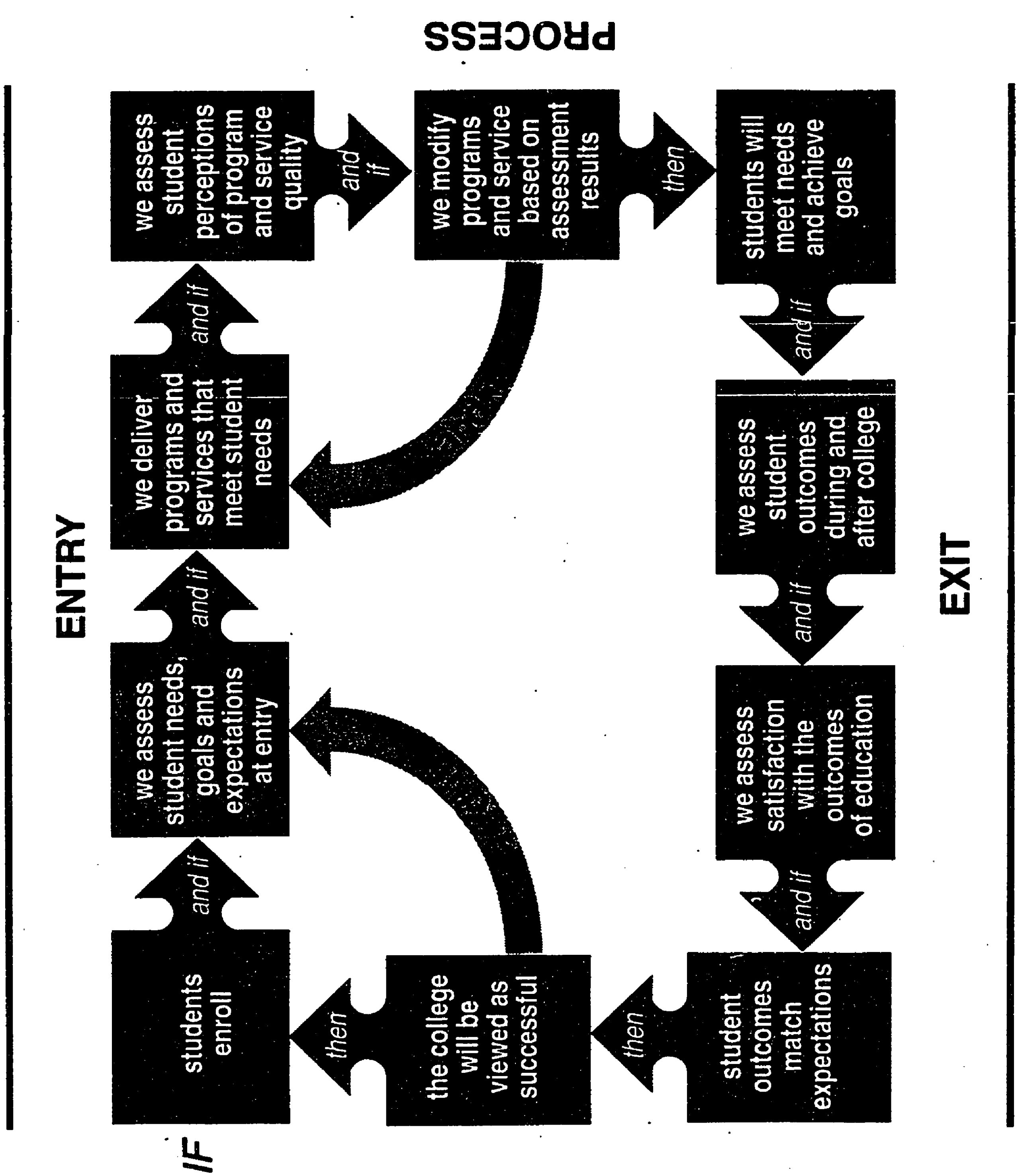
   - Tell me about the supports that are important to you as a student

   - Tell me about any barriers that you have encountered/might encounter

   - Tell me what would make this experience more successful for you

APPENDIX D
INTERVIEWEE PROFILE

		Initial	BSW	Mailout	Total
	N =	23	4	9	36
AGE	Average	39	34	37	36
	Range	29 - 49	28 - 42	21 - 51	21 - 51
FIRST NATIONS ANCESTRY		5	0	0	5
MARITAL STATUS					
	Single	0	1	1	2
	Separated	0	0	1	1
	Divorced	14	0	0	14
	Married	9	3	7	19
PARENTAL STATUS					
	no kids	1	3	2	6
	single parent, kids at home	12	0	1	13
	married, kids at home	7	1	6	14
	grown kids/not at home	3	0	0	3
EDUCATIONAL HISTORY					
	did not graduate	12	0	2	14
	missing credits	0	0	2	2
	dropped out	12	0	0	12
	[< gr. 10	5	0	0	5]
	previous post secondary credential				
	degree	0	4*	1**	5
	(* criminology, psychology, education, therapeutic recreation)				
	(**pharmacy assistant - 3 yr. degree from Hungary)				
	other	4*	0	2**	6
	(*psychiatric nursing, long term care, practical nursing, business administration)				
	(**recreation admin. diploma, U.K. equivalent of ECE plus special needs)				
WORK HISTORY					
	health/human service	4	4	6	14
	various	10	0	1	11
	clerical, bookkeeping	3	0	0	3
	banking	3	0	0	3
	housecleaning	2	0	0	2
	hairdressing	1	0	0	1
	waitressing	0	0	1	1
	pharmacy assistant	0	0	1	1
FINANCIAL STATUS (categories not exclusive)					
	SFA	15	1	0	16
	IA	0	0	1	1
	former IA	8	0	0	8
	p.t. work	13	1	4	18
	f.t. work	2	0	3	5
	EI (former or current)	3	0	0	3
	ICBC	0	0	1	1
	VRS	1	0	0	1
	band	2	0	0	2
	parents	2	0	1	3
	spouse	4	1	1	6
	maintenance	1	0	0	1
	fully sponsored	0	2	0	2



Alfred, Peterson & White, 1992, p.24)

VITA

Surname: Campbell Given Names: Jean Mallory

Place of Birth: Edmonton, Alberta, Canada

Educational Institutions Attended:

University of Victoria	1968 to 1972
University of British Columbia	1973 to 1974
	1977 to 1979

Degrees Awarded:

B.A.	University of Victoria	1972
Diploma	University of British Columbia	1974
M.Ed.	University of British Columbia	1979

Honours and Awards:

B.C. Psychology Association Gold Medal	1972

PARTIAL COPYRIGHT LICENSE

Title of Dissertation:

Redefining Student Success - Learning from Nontraditional Learners

Author: Jean Campbell

Jean Mallory Campbell
May 29, 1998

IMAGE EVALUATION
TEST TARGET (QA-3)

1.0 1.1 1.25 1.4 1.6 1.8 2.0 2.2 2.5 2.8 3.2 3.6 4.0 4.5 5.0 5.6

150mm

6"

APPLIED IMAGE . Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989

Global Optimization Using Interval Constraints

by

Huaimo Chen
B.Sc., Shenyang Institute of Aeronautical Technology, 1982
M.Sc., Changsha Institute of Technology, 1987

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Computer Science

We accept this dissertation as conforming
to the required standard

M. H. van Emden

Dr. M.H. van Emden, Supervisor (Department of Computer Science)

Mantis Cheng

Dr. M.H.M. Cheng, Department Member (Department of Computer Science)

Dr. M.R. Levy, Department Member (Department of Computer Science)

Dr. Z. Dong, Outside Member (Department of Mechanical Engineering)

Dr. T. Hickey, External Examiner (Department of Computer Science, Brandeis University)

0-612-36632-4

Supervisor: Dr. M.H. van Emden

Abstract

Global optimization methods can be classified into two non-overlapping classes with respect to accuracy: those with guaranteed accuracy and those without. The former are called *bounding methods*, the latter *point methods*. Bounding methods compute lower and upper bounds of function over a box and give a lower bound and an upper bound for the minimum. Point methods compute function values at points and output as the minimum the function value at a point.

R. E. Moore was the first to propose the bounding method using interval arithmetic for unconstrained global optimization. The first bounding method using interval arithmetic for constrained global optimization was due to E. R. Hansen and S. Sengupta. These methods are the well known bounding methods. Since these methods use interval arithmetic, we call them interval arithmetic methods. This dissertation studies the new bounding methods that use interval constraints, which is called interval constraint methods.

We prove that interval constraints is a generalization of interval arithmetic, computing an interval function in interval constraints gives the same result as in interval arithmetic. We propose a hypernarrowing algorithm using interval constraints. This algorithm produces a smaller interval result for the range of function f over a given domain than interval arithmetic. We present a generic Branch-and-Bound algorithm for unconstrained global optimization, prove the properties of the algorithm, and propose improvements on the algorithm. From this algorithm, we can obtain its interval arithmetic version and interval constraint version. We investigate the role of interval

constraints in global optimization and discuss the performance and characteristics of interval arithmetic methods and interval constraint ones.

Based on the Branch-and-Bound algorithm for unconstrained global optimization, we present a generic Branch-and-Bound algorithm for constrained global optimization, study the effect of Fritz-John conditions as redundant constraints and compare the interval arithmetic method for constrained optimization with the interval constraint one.

Examiners:

M.H. van Emden

Dr. M.H. van Emden, Supervisor (Department of Computer Science)

Mantis Cheng

Dr. M.H.M. Cheng, Department Member (Department of Computer Science)

Dr. M.R. Levy, Department Member (Department of Computer Science)

Dr. Z. Dong, Outside Member (Department of Mechanical Engineering)

Dr. T. Hickey, External Examiner (Department of Computer Science, Brandeis University)

Contents

List of Symbols ix

List of Tables xii

List of Figures xiv

Acknowledgements xvi

1 Introduction 1

1.1 Motivation 1

1.2 Global Optimization 2

1.3 Methods for Solving the Global Optimization Problem 4

1.3.1 Point Methods 4

1.3.2 Bounding Methods 5

1.4 Related Work 8

1.5 An Overview of the Dissertation 10

2 Interval Arithmetic and Interval Constraints **12**

2.1 Basics of Interval Arithmetic 13

2.1.1 Interval Arithmetic Operations 14

2.1.2 Inclusion Functions 18

2.1.3 How to Get Inclusion Functions 18

2.2 Solving Interval Constraints 19

2.2.1 Interval Constraint Systems 20

2.2.2 Consistency Operators 23

2.2.3 Consistency Algorithms 28

2.2.4 Interval Constraint Programming Languages 32

2.3 Computing Interval Functions in Interval Constraints 34

3 Hypernarrowing **41**

3.1 Where the Idea of Hypernarrowing Comes From 41

3.2 A Hypernarrowing Algorithm 44

3.3 Waltz, Solve and Hypernarrowing 46

3.3.1 Comparing Waltz and Hypernarrowing 48

3.3.2 Comparing "Solve" with Hypernarrowing 49

3.4 Applications of the Hypernarrowing 51

4 Unconstrained Global Optimization **54**

4.1 Overview of Interval Arithmetic Methods 55

4.2 Branch-and-Bound for Unconstrained Optimization 62

4.2.1 The Major Components of the Algorithm 62

4.2.2 The Data Structures Used in the Algorithm 63

4.2.3 The Description of the Algorithm 64

4.3 Interval Arithmetic vs Interval Constraints 69

4.3.1 Translating Conditions into an Interval Constraint System . . 70

4.3.2 Comparing Interval Arithmetic with Interval Constraints on Small Examples 71

4.3.3 Comparing Interval Arithmetic with Interval Constraints in General 74

4.3.4 Overview of the Implementation Variations 75

4.3.5 Implementations of the Algorithms 76

4.3.6 Computational Results 79

4.4 Time vs Memory Trade-off 82

4.4.1 Reducing Memory Use 82

4.4.2 Implementations of the Algorithms with Improvements on Memory Use 84

4.4.3 IAU_0 and IAU_1 vs IAU_0' and IAU_1' 87

4.4.4 ICU_0 and ICU_1 vs ICU_0' and ICU_1' 89

5 Constrained Global Optimization **97**

5.1 Unconstrained vs Constrained Global Optimization 98

5.1.1 Existence of the Global Minimum 99

5.1.2 Necessary Conditions 99

5.2 Overview of Interval Arithmetic Methods 103

5.3 Branch-and-Bound for Constrained Optimization 108

5.3.1 Methods for Finding An Upper Bound 109

5.3.2 Conditions for Rejecting Boxes 109

5.3.3 Conditions for the Answer List 111

5.3.4 Input and Output of the Algorithm 113

5.4 An Interval Constraint Method 114

5.4.1 Finding An Upper Bound in Interval Constraints 115

5.4.2 Using Conditions as Constraints 117

5.5 An Implementation of the Interval Constraint Method 118

5.6 Test Results 121

5.7 Discussion 123

6 Concluding Remarks **126**

6.1 Summary and Contributions 126

6.2 Suggestions For Future Work 128

List of Symbols

$f(x)$	objective function.
$p_i(x)$	inequality constraint functions, $i = 1, \ldots, m$.
$q_j(x)$	equality constraint functions, $j = 1, \ldots, r$.
f^*	the global minimum of the function f subject to $p_i(x) \leq 0$ $(i = 1, \ldots, m)$ and $q_j(x) = 0$ $(j = 1, \ldots, r)$, i.e., $f^* = \min \{f(x) \mid p_i(x) \leq 0 \ (i = 1, \ldots, m), q_j(x) = 0 \ (j = 1, \ldots, r)\}$.
x^*	a global minimizer, $f(x^*) = f^*$.
f^*_{ub}	the lowest upper bound of the global minimum f^* obtained so far.
R	the set of real numbers.
$\mathcal{F}$	a finite subset of R. Typically, $\mathcal{F}$ is the set of floating-point numbers of a computer.
$I(\mathcal{F})$	the set of floating-point intervals, i.e., the set of all $[a, b]$, where $a, b \in \mathcal{F}$.
$\mathcal{F}_n$	$\mathcal{F}_n = \{I_1 \times \cdots \times I_n \mid I_i \in I(\mathcal{F}) \text{ for } i = 1, \ldots, n\}$.
B	floating-point box $B \in \mathcal{F}_n$.
X	floating-point box $X \in \mathcal{F}_n$.
$\Box f(X)$	$\Box f(X)$ is the smallest interval containing the range of f over X.
lb	lower bound function $lb([a, b]) = a$ for any $[a, b] \in I(\mathcal{F})$.

ub upper bound function $ub([a,b]) = b$ for any $[a,b] \in I(\mathcal{F})$.

$|B|$ the width of box B. For any $B = [a,b]$, $|B| = b - a$;
for any $B = X_1 \times \cdots \times X_n$, $|B| = \max\{|X_i| \mid i = 1, \ldots, n\}$.

$\mathrm{mid}(B)$ the midpoint of box B. For any $B = X_1 \times \cdots \times X_n$,
$\mathrm{mid}(B) = (c_1, \ldots, c_n)$, where c_i is the midpoint of the interval X_i.

$\pi_i(r)$ the i-th projection of relation r. For any n-ary relation $r \subset R^n$,
the i-th projection of r
$\pi_i(r) = \{x_i \in R \mid \exists x_1 \cdots \exists x_{i-1} \exists x_{i+1} \cdots \exists x_n \text{ such that } (x_1, \ldots, x_n) \in r\}$.

$g_i(x)$ $g_i(x) = \partial f(x)/\partial x_i (i = 1, \ldots, n)$, i.e., the i-th projection of gradient of f.

$\nabla f(x)$ $\nabla f(x) = (g_1(x), \ldots, g_n(x))^T$, i.e., the gradient of f.

$h_{ij}(x)$ $h_{ij}(x) = \partial^2 f(x)/\partial x_i \partial x_j \quad (i,j = 1, \ldots, n)$, i.e., the element of Hessian.

$J_{ij}(x,X)$ $J_{ij}(x,X) = h_{ij}(X_1, \ldots, X_j, x_{j+1}, \ldots, x_n)\ (i,j = 1, \ldots, n)$,
i.e., the element of Jacobian.

$J(x,X)$ the Jacobian of f, i.e., $J(x,X) = [J_{ij}(x,X)]$, $i,j = 1, \ldots, n$.

BBUGO the Branch-and-Bound algorithm for Unconstrained Global Optimization.

BBCGO the Branch-and-Bound algorithm for Constrained Global Optimization.

IAU the Interval Arithmetic version of the Branch-and-Bound algorithm
for Unconstrained Global Optimization.

ICU the Interval Constraint version of the Branch-and-Bound algorithm
for Unconstrained Global Optimization.

$f_b(B)$ $$f_b(B) = \begin{cases} ub(f(c)) & \text{if no equalities and} \\ & c \in B, ub(p_i(c)) \leq 0 \quad \text{for all } i = 1, \ldots, m \\ ub(f(B)) & \text{if equalities exist and B contains a feasible point} \\ +\infty & \text{otherwise.} \end{cases}$$

IAC the Interval Arithmetic version of the Branch-and-Bound algorithm for Constrained Global Optimization.

ICC the Interval Constraint version of the Branch-and-Bound algorithm for Constrained Global Optimization.

List of Tables

2.1 Primitive constraints 21

3.1 Running results of Program A and Program B 53

4.1 Names of different algorithms 76

4.2 Test problems for unconstrained global optimization 80

4.3 Running results of IAU_0 and ICU_0 90

4.4 Running results of IAU_1 and ICU_1 91

4.5 Comparisons between IAU_0 and ICU_0 92

4.6 Comparisons between IAU_1 and ICU_1 92

4.7 Running results of IAU_0' 93

4.8 Running results of IAU_1' 93

4.9 Comparisons between IAU_0 and IAU_0' 94

4.10 Comparisons between IAU_1 and IAU_1' 94

4.11 Running results of ICU_0' 95

4.12 Running results of ICU_1' 95

4.13 Comparisons between ICU_0 and ICU_0' 96

4.14 Comparisons between ICU_1 and ICU_1' 96

5.1 Test problems for constrained global optimization 122

5.2 Running results of IAC and ICC 124

List of Figures

2.1 Consistency algorithm 31

3.1 Intervals for $f(x) = 12x_1^2 - 6.3x_1^4 + x_1^6 + 6x_2(x_2 - x_1)$ over $[-2,4] \times [-2,4]$ 43

3.2 A hypernarrowing algorithm 45

3.3 Algorithms for hypernarrowing the lower and upper part of an interval 47

3.4 A rough map of the poset $\mathcal{F}_n$ 50

3.5 The behavior of "solve" 51

3.6 The behavior of hypernarrowing 52

4.1 Newton algorithm in Interval Arithmetic 61

4.2 Branch-and-Bound algorithm for unconstrained optimization 66

4.3 Code of Branch-and-Bound algorithm for unconstrained optimization 78

4.4 Procedure "reduce" 85

4.5 Procedure "reduce" considering box shrunk 85

4.6 Branch-and-Bound algorithm with improvements on memory use . . . 86

4.7 Code for the procedure "reduce" 88

4.8 Code for the procedure "reduce" considering box shrunk 88

5.1 Branch-and-Bound algorithm for constrained optimization 110

5.2 Code of interval constraint method for constrained optimization . . . 120

Acknowledgements

First and foremost, I would like to express my deep appreciation to my supervisor, Professor Maarten van Emden, who guided me with extreme patience during the course of my study. He introduced me to the beautiful world of constraint logic programming and global optimization. His insights and careful guidance has made this dissertation possible. Last but not least, Maarten provided me with adequate financial support when I needed money most.

I would like to thank Professors T. Hickey, Z. Dong, M.H.M. Cheng and M.R. Levy for their conscientious reading of the dissertation, their comments and suggestions.

I would also like to thank all members of the Department of Computer Science at University of Victoria for the friendly and pleasant working environment, especially members of logical and functional programming group in the department.

Chapter 1

Introduction

1.1 Motivation

Global optimization is concerned with the determination of the global optimum (maximum or minimum) of a function. Many practical engineering applications can be formulated as global optimization problems, which are difficult to solve. For quite a long time, it has been held that no numerical method could guarantee a global optimum to a general nonlinear global optimization problem [81].

Interval arithmetic provides an upper bound and a lower bound for the range of the values of a function over a given domain [64, 61, 74, 36]. The global optimization methods using interval arithmetic can produce an upper bound and a lower bound for the global optimum of a given function [61, 78, 34, 35, 36], which are the well known methods that produce results with guaranteed accuracy [81]. These methods are called interval arithmetic methods. However, they are not efficient enough for practical uses.

This dissertation will study a new global optimization method which uses interval constraints. This new method inherits the advantages of the interval arithmetic methods. More importantly, it is more declarative and faster than interval arithmetic methods.

1.2 Global Optimization

Definition 1.1:

The global optimization problem is defined as finding

$$f^* = \min \{ \ f(x) \ |$$

$$p_i(x) \le 0 \ (i = 1, \ldots, m) \ \text{ and } \tag{1.1}$$

$$q_j(x) = 0 \ (j = 1, \ldots, r) \tag{1.2}$$

$$\}$$

where f, p_i $(i = 1, \ldots, m)$ and q_j $(j = 1, \ldots, r)$ are scalar functions of a vector x of n components and f^* is the global minimum.

Without loss of generality, we restrict our attention to minimization. Since minimizing $f(x)$ is equivalent to maximizing $-f(x)$, we can find the maximum for $f(x)$ by finding the minimum of $-f(x)$. If $m = 0$ and $r = 0$, we get an unconstrained global optimization problem; otherwise, we get a constrained one.

The global optimization is to find the global minimum f^* as well as global minimizers or locations at which this minimum value occurs. Formally, we say that a point x^* is a global minimizer if

$$f(x^*) = f^*$$

and x^* satisfying the constraints in 1.1 and 1.2.

Contrary to the global optimization, local optimization is to find local minima and local minimizers. A point x^* is a local minimizer if there exists a number $\epsilon > 0$ such that

$$f(x^*) \leq f(x)$$

for all x satisfying

$$||x - x^*|| < \epsilon$$

and the constraints in 1.1 and 1.2; where $x - x^*$ is the result of vector subtraction, the *norm* $||v||$ of a vector v with n components is a measure of the size of v [28, 81]. One common norm is defined as

$$||v|| = \max_{1 \leq i \leq n} |v_i|.$$

Rinnooy Kan and Timmer [45] claim that the global optimization problem as stated above is unsolvable in a finite number of steps. However, for a given tolerance $\epsilon > 0$, one can, in a finite amount of time, find a point x' such that the difference between $f(x')$ and the global optimum is within ϵ [62]. Such a point x' is called an ϵ-global minimizer. Formally, an ϵ-global minimizer is defined as a point, x', such that

$$| f(x') - f^* | \leq \epsilon$$

and

$$p_i(x') \leq \epsilon \ (i = 1, \ldots, m) \tag{1.3}$$

$$|q_j(x')| \leq \epsilon \ (j = 1, \ldots, r) \tag{1.4}$$

where $f(x')$ is called an ϵ-global minimum.

In the following, the terms *minimum* and *minimizer* will mean *global minimum* and *global minimizer* respectively if we do not specify otherwise.

1.3 Methods for Solving the Global Optimization Problem

Among many kinds of classifications of methods for solving the global optimization problem [81, 88], we are interested in dividing the methods into the following two categories:

- point methods, and
- bounding methods.

Point methods compute function values at points and output as the minimum the function value at a point. Bounding methods compute lower and upper bounds of function over a box and give a lower bound and an upper bound for the minimum. Point methods are incapable of reliably solving the global optimization problem. Bounding methods produce correct global optimization solutions even in the presence of round-off errors. The strength of the point methods is their efficiency.

1.3.1 Point Methods

Point methods are iterative. From a starting point x_i (initially $i = 0$), a point method tries to converge to a minimum by finding another point x_{i+1} such that $f(x_{i+1}) < f(x_i)$

and using x_{i+1} as a new starting point.

Point methods may converge to a local minimum. Through starting the iteration at a number of randomly selected points or at the points of a grid, we increase the probability of finding all of the local minima. Thus it is possible that the smallest local minimum among all the local minima found is the global minimum.

A point method gives an upper bound for the global minimum without an indication of how close this bound is. Without additional information about the objective function, point methods can not even guarantee that an ϵ-global minimizer has been found [74].

1.3.2 Bounding Methods

Starting with a given box, a Cartesian product of intervals, bounding methods produce: (1) a lower bound and an upper bound for the global minimum of the function f, and (2) a list of small boxes. The union of these boxes will contain all global minimizers.

All bounding methods, such as Lipschitzian methods [81] and interval arithmetic methods [81], consist of the following three steps:

(1) **partitioning** the initial box into smaller boxes,

(2) **bounding** the function (and possibly its derivatives) over the boxes and the global minimum of the function, and

(3) **rejecting** (by using the bounds calculated in step 2) those boxes which do not contain a global minimizer.

The above three steps will be repeatedly executed until a certain termination condition is satisfied. The union of the remaining boxes will contain all global minimizers.

The bounding algorithm searches for the global minimum by exhaustively partitioning and "pruning" all of the feasible space, which do not contain a global minimizer.

Step 1 (i.e., the partitioning step) usually splits the initial domain, a given box, into smaller boxes. Choosing boxes as the geometric shape to be split has the following advantages:

- boxes are easily partitioned,
- boxes cover the feasible region without overlap,
- lower and upper bounds of the function f over boxes are easily computed.

If the feasible space itself is not a box, then an initial box containing it is used as the initial search domain. The equalities and inequalities which are used to characterize the true feasible space are then additionally used in the rejection phase of the algorithm to eliminate those subboxes lying in the initial box, but outside of the feasible region.

Step 2 (i.e., the bounding step) computes the lower bounds of the function f over boxes and finds an upper bound of the global minimum. Initially, the upper bound of the minimum is set to $+\infty$. At each iteration, it may be improved by the value of the function f at a point in the feasible region or a local minimum. For a given box B, the lower bound of f can be computed in one of the following ways:

- **Lipschitzian Approach.** It is based on the assumption that a Lipschitz constant, L, exists such that for any $u, v \in B$, we have

$$|f(u) - f(v)| \leq L||u - v||.$$

  If the value of function f is known at point u, then a lower bound on the function value for all x between u and v can be determined by the following formula

$$f(u) - L||u - v||.$$

- **Interval Arithmetic.** Lower and upper bounds of the function f over the given box can be calculated by using interval arithmetic. An "*interval extension*" (to be defined later) F for f can be obtained easily. Such an interval function operates over a given box and returns an interval result bounding the range of the function over the given box.

- **Interval Constraints.** For the function f and box B, the interval constraints approach produces the same interval result as interval arithmetic if the initial interval for y is $(-\infty, +\infty)$, where y is the variable for the value of the function f. More importantly, when various additional conditions are used as constraints, not only can it produce a smaller interval result than interval arithmetic, but also it can make the given box B become smaller. This will be explained in more detail later.

Among the lower bounds of the function f over all the unrejected boxes, the lowest lower bound is a lower bound of the global minimum.

Step 3 (i.e., the rejecting step) rejects boxes by using various conditions. The following are some of commonly used conditions.

- rejecting boxes over which the lower bound of the function f is greater than an upper bound on the global minimum known so far;
- rejecting boxes which do not intersect with the feasible space (i.e., in which any point does not satisfy the conditions 1.1 and 1.2);
- rejecting boxes B not on the border of the initial box for which $0 \notin \{g(x) \mid x \in B\}$, where g is the gradient of the objective function, f;
- rejecting boxes for which the function is not concave anywhere within the boxes.

1.4 Related Work

Applications of BNR-Prolog. BNR-Prolog [10, 67, 7] is an interval constraint logic programming language. W. J. Older explored in BNR-Prolog the applications of interval constraints in several areas [73, 69, 65, 72, 70]. In global optimization, Older solved a global optimization problem from [29] by using a branch and bound algorithm and the Kuhn-Tucker conditions [72]. But he did not study the properties of the branch and bound algorithm for solving the global optimization problem.

UniCalc Solver. The UniCalc solver [3] was designed to solve systems of nonlinear equalities and inequalities. A technique analogous to interval constraints, called *subdefinite computations method* [76], is used in the UniCalc solver. It gives bounds on solutions. But it does not contain special methods for solving global optimization problems. To find the global minimum of a second order differentiable function without constraints, one has to follow the following three steps:

(1) create a system of equalities and inequalities from the known necessary conditions for a point x to be a local minimizer: the first order partial derivative of the objective function must be zero, and the diagonal elements of the Hessian matrix must be greater than or equal to zero.

(2) apply the UniCalc solver to the system, and thus find the intervals for the variables in the system.

(3) separate minimizers by using the root locating tool.

Numerica. Recently, P. Van Hentenryck and L. Michel reported their Numerica system for global optimization in [37, 38]. They use interval constraints, therefore obtaining the same advantages over interval arithmetic as reported in [16, 84, 17]. Numerica is built on Newton [5], an interval constraint logic programming language. An iterative interval Newton method was embodied in Newton. And the interval Newton method was combined with an internal splitting operation on intervals. It reported in [5] that Newton achieved one to two order magnitudes in speed over BNR-Prolog. They demonstrated the performance of their system. It seems that their system outperforms existing ones. Comparison are hard to make from their papers, since they only give timings and number of splits. The number of splits has the advantage of being machine independent. However, they did not report the "internal splits" performed by means of the iterative interval Newton method. Since there are hierarchical redundant conditions that can be used as constraints for solving the global optimization problem, it is not clear what role of each constraint is played.

1.5 An Overview of the Dissertation

In chapter 2, we first review the basic concepts of interval arithmetic, followed by interval constraints. We then demonstrate how the interval functions can be computed by using interval constraints and prove that computing an interval function in interval constraint gives the same result as in interval arithmetic.

In chapter 3, we give a hypernarrowing algorithm which can produce a smaller interval result for the range of function f over a given box than the corresponding interval function in interval arithmetic. And then we compare the semantics of hypernarrowing with that of a constraint solver, which is followed by applications.

In chapter 4, we first review interval arithmetic methods for unconstrained global optimization, describe a generic Branch-and-Bound algorithm for unconstrained global optimization, prove the properties of the algorithm, and propose improvements on the algorithm. We investigate the role of interval constraints in global optimization and show how to obtain an interval arithmetic version and interval constraint version of the algorithm, which is followed by the implementation of a variety of versions of the Branch-and-Bound algorithm in BNR-Prolog. We then compare the computational results produced by the interval arithmetic versions with those produced by the interval constraint ones.

In chapter 5, we discuss the differences between unconstrained and constrained optimization, which is followed by the review of interval arithmetic methods for constrained global optimization. We then present the transition from the Branch-and-Bound algorithm for the unconstrained global optimization to the variety for the constrained global optimization. We also study the effect of redundant conditions as interval constraints and compare the interval arithmetic methods with the interval constraint ones.

In chapter 6, we summarize our results and contributions, and indicate directions of future researches.

Chapter 2

Interval Arithmetic and Interval Constraints

Interval arithmetic is an arithmetic defined on intervals, rather than on real numbers. A form of interval arithmetic perhaps first appeared in 1924 in [12, 47]. Modern development of interval arithmetic began with R. E. Moore's dissertation [59, 47]. The key idea of interval arithmetic is to bound by an interval the effect of errors from all sources, including approximation errors and errors in data. Since then, applications [2, 74, 36, 8] of interval arithmetic have been developed. In most applications, interval functions [64, 2, 74, 36] are used to bound the ranges of given functions. In general, the interval yielded by an interval function of a given function f is much larger than the range of f. It is therefore important to find ways to bound the range of f as closely as possible.

Using interval constraints, we can bound the range of the function f better than using interval arithmetic. The concept of constraints was formed gradually. Sutherland [80] is one of pioneers of constraints. A basic consistency technique for interval constraints, the Waltz consistency algorithm, was given by Waltz [89]. Interval constraints has been embedded in several programming languages [18, 44, 23, 20, 5, 57]. It provides a general approach in computer problem solving where one expects the problem to be solved by merely entering the constraints among the variables of the problem without need for any algorithm in addition to a general-purpose constraint solver that comes with the system. Of course, in interesting problems the general-purpose constraint solver is neither efficient nor convenient to use. This is also the case in global optimization problems.

In this chapter, we prove that for any interval function F of a given function f, if we leave the value Y for F unconstrained initially and compute it in interval constraints, we get the same result for Y as we compute it in interval arithmetic. This indicates that interval arithmetic is a special case of interval constraints.

2.1 Basics of Interval Arithmetic

Before the use of interval arithmetic, bounds on the range of a function were sometimes obtained with Lipschitz constants. Judicious use of interval arithmetic allows such range bounds to be computed without extensive analysis. With *outward rounding* (to be explained later), interval arithmetic provide correct results from floating-point operations on computers in the presence of rounding errors.

We use R to denote the set of real numbers and $\mathcal{F}$ a finite subset of R. Typically, $\mathcal{F}$ is the set of floating-point numbers of a computer.

Definition 2.1: [40]

For every $a, b \in R$, the real interval (or R-interval for short)

$$[a, b] \text{ represents } \{x \in R \mid a \leq x \leq b\},$$
$$[a, +\infty] \text{ represents } \{x \in R \mid a \leq x\},$$
$$[-\infty, b] \text{ represents } \{x \in R \mid x \leq b\}, \text{ and}$$
$$[-\infty, +\infty] \text{ represents } R.$$

Note $+\infty \notin R$ and $-\infty \notin R$. $[a, +\infty]$, $[-\infty, b]$ and $[-\infty, +\infty]$ are just more convenient notations for $\{x \in R \mid a \leq x\}$, $\{x \in R \mid x \leq b\}$, and R respectively.

If we replace "$a, b \in R$" by "$a, b \in \mathcal{F}$" in the definition 2.1, then we obtain the definition of floating-point interval (or $\mathcal{F}$-interval for short). We denote the set of all $\mathcal{F}$-intervals by $I(\mathcal{F})$. For any $[l, u] \in I(\mathcal{F})$, we call l the lower bound and u the upper bound of $[l, u]$. For convenience, we will use lower bound function $lb : I(\mathcal{F}) \to \mathcal{F}$ and upper bound function $ub : I(\mathcal{F}) \to \mathcal{F}$, which are defined as follows:

$$lb([l, u]) = l, \ \ ub([l, u]) = u \quad \text{for any } [l, u] \in I(\mathcal{F}).$$

Note that real and floating-point intervals differ only in the bounds but every interval, real or floating-point, denotes a set of real numbers.

2.1.1 Interval Arithmetic Operations

Definition 2.2: [74, 36]

If $\circ \in \{+, -, *, /\}$ and $A, B \in I(\mathcal{F})$, then $A \circ B$, the result of the interval operation $\circ$, is the smallest interval in $I(\mathcal{F})$ containing

$$\{x \circ y \mid x \in A \text{ and } y \in B\}.$$

Other operations over $I(\mathcal{F})$ like sin, cos, log, etc. can be defined in the same way.

The image of each of the four basic interval operations is the range of the corresponding real operation except for A/B which is complicated if $0 \in B$. Although the above definition characterizes these operations mathematically, the practical use of interval arithmetic is due to the following theorems.

Theorem 2.1: [83]
For any $\circ \in \{+, -, *, /\}$ and $[a, b], [c, d] \in I(\mathcal{F})$

$$[a, b] \circ [c, d] = [\min(a \circ c, a \circ d, b \circ c, b \circ d), \max(a \circ c, a \circ d, b \circ c, b \circ d)]$$

except for $[a, b]/[c, d]$, which is more complicated if $0 \in [c, d]$.

Theorem 2.2: [74, 36]
Let $[a, b], [c, d] \in I(\mathcal{F})$, then

$$\begin{aligned}
[a, b] + [c, d] &= [a + c, b + d] \\
[a, b] - [c, d] &= [a - d, b - c] \\
[a, b] * [c, d] &= [\min(ac, ad, bc, bd), \max(ac, ad, bc, bd)] \\
[a, b]/[c, d] &= [a, b] * [\tfrac{1}{d}, \tfrac{1}{c}], \quad \text{if } 0 \notin [c, d]
\end{aligned}$$

Theorem 2.3: [83]

Let $0 \in [c, d]$, then

$$[a,b]/[c,d] = \begin{cases} \emptyset & \text{if } 0 \notin [a,b] \text{ and } c = 0 \text{ and } d = 0 \\ [-\infty, +\infty] & \text{if } a < 0 \text{ and } 0 < b \\ [-\infty, +\infty] & \text{if } c < 0 \text{ and } 0 < d \\ [\frac{a}{d}, +\infty] & \text{if } c = 0 \text{ and } 0 < d \text{ and } a \geq 0 \\ [-\infty, \frac{b}{d}] & \text{if } c = 0 \text{ and } 0 < d \text{ and } b \leq 0 \\ [-\infty, \frac{a}{c}] & \text{if } c < 0 \text{ and } d = 0 \text{ and } a \geq 0 \\ [\frac{b}{c}, +\infty] & \text{if } c < 0 \text{ and } d = 0 \text{ and } b \leq 0 \end{cases}$$

It is obvious that both interval operations $+$ and $*$ are commutative and associative. However, the distributive laws do not hold. For any intervals $A, B, C, D \in I(\mathcal{F})$, we have [1]

$$A * (B + C) \subset A * B + A * C.$$

This property of the interval operations is called *subdistributive*. Moreover, although $[0,0]$ is an identity for addition and $[1,1]$ is an identity for multiplication, that is that for any interval $X \in I(\mathcal{F})$, we have

$$[0,0] + X = X$$

and

$$[1,1] \times X = X,$$

but we do not have

$$[0,0] = X - X$$

or

$$[1,1] = X/X.$$

[1]As we regard Bourbaki [11] and Halmos [31] as authorities on set-theoretic notation, we prefer $\subset$ for the subset relation to $\subseteq$. Similarly, $\supset$ rather than $\supseteq$.

This is due to the fact that the multiple occurrences of one variable X vary independently from the definition 2.2. This phenomenon is called *interval dependency*. Because of the interval dependency, the interval result of evaluating a function f over a given interval with interval arithmetic usually is much larger than the range of the function f. For example, if

$$f(x) = x^2 - x,$$

then evaluating f over $[0, 1]$ with interval arithmetic gives

$$[0,1]^2 + [0,1] = [0,1] - [0,1] = [-1,1],$$

but the range of f over $[0, 1]$ is $[-\frac{1}{4}, 0]$.

In floating-point arithmetic, real numbers are approximated by floating-point numbers using rounding. This rounding introduces roundoff errors. During the process of a computation, the rounding errors are accumulated. This may lead to the result of the computation to be wrong. This is not the case in floating-point interval arithmetic. In an interval operation, the lower bound of the result interval is rounded down to the largest floating-point number less than the exact lower bound, and the upper bound of the result interval is rounded up to the smallest floating-point number larger than the exact upper bound. This rounding process is called *outward rounding*. The outward rounding does not introduce roundoff errors. It guarantees that no answer, if there exists any, escapes from the rounded interval. This contributes to the guaranteed accuracy property of interval arithmetic. Therefore, a sound implementation of the outward rounding is essential in an interval arithmetic system.

The IEEE binary floating point standard [21, 30] prescribes three rounding modes: *nearest* (round to the nearest floating-point number), *round down* (round toward $-\infty$), and *round up* (round toward $+\infty$). The *nearest* mode is the default rounding mode. For the implementation of outward rounding, we use the *round down* mode when computing lower bounds and the *round up* mode when computing upper bounds.

2.1.2 Inclusion Functions

In order to introduce the inclusion function of the real function $f : R^n \to R$, we define floating-point box. Let

$$\mathcal{F}_n = \{I_1 \times \cdots \times I_n \mid I_i \in I(\mathcal{F}) \text{ for } i = 1, \ldots, n\} \tag{2.1}$$

the set of Cartesian products of n floating-point intervals. We call $I_1 \times \cdots \times I_n \in \mathcal{F}_n$ floating-point box or $\mathcal{F}$-box for short, and $I \in \mathcal{F}_1$ (i.e., $I(\mathcal{F})$) floating-point interval or $\mathcal{F}$-interval for short.

Definition 2.3: [74]
Let $f : R^n \to R$. Let furthermore $\Box f(X)$ be the smallest interval containing the range of f over X, where X denotes $X_1 \times \cdots \times X_n$, and $X_i \subset R$. A function $F : \mathcal{F}_n \to \mathcal{F}_1$ is called an inclusion function for f if $\Box f(X) \subset F(X)$ for any $X \in \mathcal{F}_n$.

2.1.3 How to Get Inclusion Functions

Definition 2.4: [74]
Let $f : R^n \to R$, E be an expression for f, and $X \in \mathcal{F}_n$. The *natural interval extension* for E of f to X is the function $F : \mathcal{F}_n \to \mathcal{F}_1$ defined by the expression E' that is obtained from E by replacing each occurrence of the variable x by X, each arithmetic operation by the corresponding interval arithmetic operation and each predefined function by the corresponding inclusion function. We also call F an interval

function of f.

Lemma 2.1: [74]
For any $f, g, h : R \to R$, if $f(x) = g(h(x))$, G and H are inclusion functions for g and h, and $F(X) = G(H(X))$, then F is an inclusion function for f.

Corollary 2.1: [74]
For any real number functions $f, h_i (i = 1, \ldots, n) : R^n \to R$, $g : R^m \to R$, if $f(x) = g(h_1(x), \ldots, h_m(x))$, G and H_i $(i = 1, \ldots, n)$ are inclusion functions for g and h_i, and $F(X) = G(H_1(X), \ldots, H_m(X))$, then F is an inclusion function for f.

Theorem 2.4: [74]
The natural interval extension for any expression for f as defined above is an inclusion function for f.

The mean-value form and Taylor form [74, 64] of f are two other kinds of inclusion functions for f.

In general, the value of an inclusion function for a given function f over $X \in \mathcal{F}_n$ is much larger than $\Box f(X)$, the smallest interval containing the range of the function f over X. Therefore, it is important and challenging to find ways to approximate $\Box f(X)$ as well as possible. It is challenging as $\Box f(X)$ itself requires solving global minimization and maximization problems where f may not be a convex function.

2.2 Solving Interval Constraints

Interval constraints, built on interval arithmetic, is a generalization of interval arithmetic. The fundamental algorithm for solving interval constraint systems was proposed by Davis [22]. Following the work presented in [18, 52, 68], F. Benhamou and W. Older [6] introduced the notions of approximation and narrowing, and applied them to constraints over real numbers, integers and Booleans. M. H. van Emden [82] generalized these notions to Herbrand universes and finite domains.

2.2.1 Interval Constraint Systems

Definition 2.5: [84]

An interval constraint system is an entity consisting of

(1) A **constraint conjunction**, $A_1 \wedge \cdots \wedge A_m$, where A_i $(i = 1, \ldots, m)$ are atomic formulas of first-order predicate logic. These formulas are called *primitive constraints* or *primitive relations*, each of which has one of the forms in Table 2.1, where the variables are interpreted as reals.

(2) A **state**, $I_1 \times \cdots \times I_n$, which is an $\mathcal{F}$-box. Each component (i.e., an interval) of the box is associated with a variable occurring in the constraint conjunction. It may happen that one or more of the intervals are empty. In this case the state denotes an empty set of tuples of values for the variables. Such a state is called a *failure state* or an *inconsistent state*.

For example,

$$exp(x, 2, y) \;\wedge\; exp(y, 2, z) \;\wedge\; sum(y, z, 1)$$

is a constraint conjunction, which is interpreted as

$$x^2 = y \;\wedge\; y^2 = z \;\wedge\; y + z = 1.$$

formula	interpretation
sum(x, y, z)	$x + y = z$
times(x, y, z)	$x * y = z$
exp(x, n, y)	$x^n = y$ for integer $n \geq 2$
eq(x, y)	$x = y$
lt(x, y)	$x < y$
le(x, y)	$x \leq y$
gt(x, y)	$x > y$
ge(x, y)	$x \geq y$
sin(x, y)	$\sin(x) = y$
cos(x, y)	$\cos(x) = y$
tan(x, y)	$\tan(x) = y$
asin(x, y)	$\arcsin(x) = y$
acos(x, y)	$\arccos(x) = y$
atan(x, y)	$\arctan(x) = y$
abs(x, y)	$\mathrm{abs}(x) = y$
ln(x, y)	$\ln(x) = y$
min(x, y, z)	if $x < y$ then $z = x$ else $z = y$
max(x, y, z)	if $x > y$ then $z = x$ else $z = y$

Table 2.1: Primitive constraints

Suppose that we also have the following $\mathcal{F}$-box

$$[-5.5, 6.5] \times [-1.5, 1.5] \times [-1.6, 1.6].$$

If we associate the interval $[-5.5, 6.5]$, $[-1.5, 1.5]$ and $[-1.6, 1.6]$ with the variables x, y and z in the above constraint conjunction respectively, then we get a state. This state and the above conjunction constitute an interval constraint system.

When writing a primitive relation, we prefer to write the interpretation as shown in the right-hand column of the table. This has the advantage of improved readability at the expense of a risk of confusion. For example, when writing "$x + y = z$" as a primitive relation, we should keep in mind that it is a ternary relation written in a sort of distributed infix notation (i.e., it is the set $\{< x, y, z >| x + y = z\}$.) and that it is not an instance of the binary equality relation involving the result of an addition.

Usually, we denote a constraint system by $C = < S, D >$, where $S = A_1 \wedge \cdots \wedge A_m$ is the constraint conjunction, $D = I_1 \times \cdots \times I_n$ is the state, in which intervals $I_1, \ldots, I_n$ are associated with the variables $x_1, \ldots, x_n$ in the constraint conjunction S respectively.

Definition 2.6:
A *solution* of an interval constraint system is an n-tuple of values for the variables that makes the constraint conjunction $A_1 \wedge \cdots \wedge A_m$ become true if each variable is replaced by the corresponding value.

Thus the set of solutions is a set of n-tuples, hence an n-ary relation, say r, where $r \subset I_1 \times \cdots \times I_n$.

Even though the interval constraint system defined here allows only a limited number of primitive constraints, it is general enough to represent equalities or inequalities between polynomials of any degree in any number of variables. Of course such polynomials have to be translated to the primitive constraints by introducing auxiliary variables. Such a translation is similar to the one performed by a Fortran or C compiler, where the target code's arithmetic instructions play the role of the primitive constraints here.

For example,

$$y = x_1^2 + x_2,\ x_1 \in [-10, 10],\ x_2 \in [-100, 50],\ y \in (-\infty, +\infty)$$

is translated to an interval constraint system where the constraint conjunction is

$$x_1^2 = z \ \wedge \ z + x_2 = y$$

and the state is

$$[-10, 10] \times [-100, 50] \times [-\infty, +\infty] \times [-\infty, +\infty]$$

which is associated with the variables x_1, x_2, y and z, where z is the auxiliary variable introduced.

2.2.2 Consistency Operators

Definition 2.7:

Let $p(x_1, \ldots, x_n)$ be a primitive constraint, and $X = X_1 \times \cdots \times X_n$ be the state, in which $X_1, \ldots, X_n$ are associated with the variables $x_1, \ldots, x_n$ in the constraint. A value $v \in X_i$ for the variable x_i is said to be consistent for p and X if we can find $x_j \in X_j$ ($j \in \{1, \ldots, n\}, j \neq i$) such that

$$p(x_1, \ldots, x_{i-1}, v, x_{i+1}, \ldots, x_n)$$

holds.

The functionality of the consistency operator associated with a primitive constraint p and a given box is to remove inconsistent values from each interval of the box and find the smallest box containing all the consistent points in the given box.

For example, if $p(x, y, z)$ is $x+y = z$ (the primitive constraint *sum* from Table 2.1) and the intervals associated with x, y and z are $[0, 2], [0, 2]$ and $[3, 5]$ respectively, then all three intervals contain inconsistent values. Now $y \leq 2$ (from $y \in [0, 2]$) and $z \geq 3$ (from $z \in [3, 5]$) imply that $x = z - y \geq 1$. Hence the values in $[0, 1)$ for x are *inconsistent*. Thus we get the interval $[1, 2]$ for x from $[0, 2]$. Similar considerations rule out values in $[0, 1)$ for y and values in $(4, 5]$ for z. Removing all inconsistent values from the given intervals leaves the intervals $[1, 2]$ for x and y and $[3, 4]$ for z.

This is an example of the consistency operator associated with the primitive constraint *sum* acting on intervals associated with variables related by *sum*. In general there is a consistency operator associated with each primitive constraint relation p that acts on intervals associated with argument places of p by first removing all inconsistent values. As the resulting sets may not be intervals, the consistency operator includes a second step, which is to replace these sets by the least intervals containing them.

Before giving the definition of consistency operator formally, we introduce *projection* and *approximation*.

Definition 2.8: [6]

For every n-ary relation $r \subset R^n$, the projection of r, denoted by $\pi_i(r)$, is defined as follows:

$$\pi_i(r) = \{x_i \in R \mid \exists x_1 \cdots \exists x_{i-1} \exists x_{i+1} \cdots \exists x_n \text{ such that } (x_1, \ldots, x_n) \in r\}.$$

Definition 2.9: [6, 85]

The approximation of a relation r, denoted by $ap(r)$, is the least (w.r.t. inclusion relation) $\mathcal{F}$-box containing r.

Definition 2.10: [6]

Let $r \subset R^n$. The consistency operator of r is the function C_r: $\mathcal{F}_n \rightarrow \mathcal{F}_n$, such that for any $u \in \mathcal{F}_n$,

$$C_r(u) = ap(u \cap r).$$

If r is one of the primitive constraints in Table 2.1, the corresponding consistency operator can be simply computed. The formulas used for computing consistency operators can be found in [18, 71, 85, 6, 51, 83]. We list a few of important ones in the following lemmas.

Lemma 2.2: [18, 71, 85, 6, 51]

If sum is the ternary relation

$$sum = \{< x, y, z > \mid x + y = z, \quad x, y, z \in R\}$$

and C_{sum} is the consistency operator of sum, then for any $[a, b] \times [c, d] \times [e, f] \in \mathcal{F}_3$ we have

$$\begin{aligned} C_{sum}([a,b] \times [c,d] \times [e,f]) \quad = \quad & [a,b] \cap ([e,f] - [c,d]) \times \\ & [c,d] \cap ([e,f] - [a,b]) \times \\ & [e,f] \cap ([a,b] + [c,d]). \end{aligned}$$

Thus, the consistency operator C_{sum} can be easily computed in terms of the interval arithmetic operations + and −. However, the formulas for computing C_{times} in [18, 71, 51] are complicated and incomplete, because the result of interval division $[a,b]/[c,d]$ is not defined if $0 \in [c,d]$. Even in extended interval arithmetic [33], $[a,b]/[0,0]$ is not defined. Because of the discontinuity at 0, W. Older broke the computation of C_{times} into twenty seven cases. By using symmetry, he reduced these twenty seven cases to three essentially different ones. The discontinuity also led to the consistency operator C_{times} incompletely implemented in BNR-Prolog for a number of years [71].

After deriving the actually executable computation rules for interval division $[a,b]/[c,d]$ especially when the divisor is near 0, M. H. van Emden [83] has presented the simple formula for the consistency operator of *times*.

Lemma 2.3: [83]
If *times* is the ternary relation

$$times = \{< x, y, z > \mid x * y = z, \quad x, y, z \in R\}$$

and C_{times} is the consistency operator of *times*, then for any $[a,b] \times [c,d] \times [e,f] \in \mathcal{F}_3$ we have

$$\begin{aligned} C_{times}([a,b] \times [c,d] \times [e,f]) \quad = \quad & [a,b] \cap ([e,f]/[c,d]) \times \\ & [c,d] \cap ([e,f]/[a,b]) \times \\ & [e,f] \cap ([a,b] * [c,d]). \end{aligned}$$

Lemma 2.4: [71]
If eq is the binary relation

$$eq = \{< x, x > \mid x \in R\}$$

and C_{eq} is the consistency operator of eq, then for any $[a,b] \times [c,d] \in \mathcal{F}_2$ we have

$$\begin{aligned} C_{eq}([a,b] \times [c,d]) &= [a,b] \cap [c,d] \times \\ &\quad [a,b] \cap [c,d]. \end{aligned}$$

Theoretically speaking, the relation r defined by the constraint system C can be approximated by the consistency operator C_r. However, there is no algorithm for C_r. The following prepares our mathematical model for the practical consistency operator of the relation r.

Suppose that we have a constraint system $C = < S, D >$, where $S = A_1 \wedge \cdots \wedge A_m$, $D = I_1 \times \cdots \times I_n$. For all $i = 1, \ldots, m$, if $x_{j_1}, \ldots, x_{j_{n_i}}$ are the variables occurring in A_i, then A_i denotes the n_i-ary relation $r_i \subset I_{j_1} \times \cdots \times I_{j_{n_i}}$ as the set of tuples that, when substituted for $x_{j_1}, \ldots, x_{j_{n_i}}$, make A_i true.

Lemma 2.5: [85]
$r = r_1 \bowtie \cdots \bowtie r_m$, where $\bowtie$ denotes the natural join.

The specification of r by means of $\exists x_1 \ldots \exists x_n(A_1 \wedge \cdots \wedge A_m)$ may suggest that $r = r_1 \cap \cdots \cap r_m$, but this is only the case if each of $A_1, \ldots, A_m$ contains all the variables in the constraint conjunction. Typically, however, every constraint contains only a small subset of all the variables.

Definition 2.11: [85]
For any $u \in \mathcal{F}_n$ and $r = r_1 \bowtie \cdots \bowtie r_m$,

$$T(u) = C_{r_1}(u^1) \bowtie \cdots \bowtie C_{r_m}(u^m)$$

where for all $i = 1, \ldots, m$, u^i is the projection of u on the subset of the variables that occur in A_i.

Proposition 2.1: [85]

(i) T is monotonic; (ii) For any $u \in \mathcal{F}_n$, $u \supset T(u) \supset C_r(u)$; (iii) $T(C_r(u)) = C_r(u)$.

Theorem 2.5: [85]

For any $u \in \mathcal{F}_n$, there exists a finite M such that $T(T^M(u))) = T^M(u)$. Moreover, $u \supset T^M(u) \supset C_r(u)$.

This suggests a consistency operator for approximating $C_r(u)$ and hence r.

Definition 2.12: [85]

The consistency operator of a constraint conjunction S is the function $\Psi_S : \mathcal{F}_n \to \mathcal{F}_n$, such that for any $u \in \mathcal{F}_n$,

$$\Psi_S(u) = T^M(u).$$

Lemma 2.6: [85]

(i)	Ψ_S is contracting:	$\Psi_S(u) \subset u$ for all $u \in \mathcal{F}_n$.
(ii)	Ψ_S is monotonic:	$u_1 \subset u_2 \Rightarrow \Psi_S(u_1) \subset \Psi_S(u_2)$ for all $u_1, u_2 \in \mathcal{F}_n$.
(iii)	Ψ_S is idempotent:	$\Psi_S(u) = \Psi_S(\Psi_S(u))$ for all $u \in \mathcal{F}_n$. In other words, Ψ_S maps $\mathcal{F}_n$ to the fixpoints of Ψ_S.

2.2.3 Consistency Algorithms

Definition 2.13:

A state is a consistent state for an interval constraint system if this state is unchanged under the consistency operator of any of the primitive constraints in the constraint conjunction.

A consistency algorithm takes as input an interval constraint system, which has a state. The algorithm reduces the intervals in this initial state to one of the interval constraint system's "consistent states". It performs this reduction in such a way that any solution contained in the initial state is also contained in the consistent state obtained from it. That is, it does not eliminate any solution. It may happen that the consistent state has intervals that have a width in the same order of magnitude as the precision of the machine arithmetic. This suggests that a unique solution is contained in the consistent state.

It may also happen that the consistency algorithm reduces the state to the failure state. In that case it has been shown that the original state contains no solutions. Finally, it may happen that the intervals of the consistent state that has been reached provide too little information about solutions to be useful. In this case, the state can be split into two or more sub-states. By applying the consistency algorithm to each of these sub-states, we can obtain more information about solutions.

One can get a consistent state by starting from an initial state and repeatedly execute a cycle in which all consistency operators are applied to the current state. As intervals never increase in size and as there are only finitely many machine numbers, such an iteration reaches a consistent state.

It is usually unnecessary to apply all consistency operators in each cycle. For a primitive constraint in the constraint system, its consistency operator needs to be applied only if the intervals associated with some of the variables occurring in the constraint are changed. This optimization is achieved by Waltz/Davis algorithm [22]. The process is often referred to as "constraint propagation" or "filtering". Thus the Waltz/Davis algorithm maps an initial state of an interval constraint system to a consistent state contained in it. In the process we obtain information about solutions possibly contained in the initial state to the extent by which the intervals have contracted.

The algorithm in Figure 2.1, which is based on the algorithms in [22, 6], is the pseudocode of an efficient procedure. In the algorithm, $I(x_i)$ represents the interval associated with the variable x_i in the constraint system. The input of the algorithm is a constraint system $C = < S, D >$. The output is a failure state or a consistent state D'.

For example, if the constraint conjunction S is

$$y = x^2 \wedge z = y^2 \wedge y + z = 1$$

and the initial state D is

$$[0, 1] \times [0, 1] \times [0, 1],$$

then no information is gained as this initial state is a consistent state. Through splitting the interval $[0, 1]$ for x into two sub-intervals $[0, 0.5]$ and $[0.5, 1]$, we get two states:

$$D1 = [0, 0.5] \times [0, 1] \times [0, 1]$$

and

$$D2 = [0, 0.5] \times [0, 1] \times [0, 1].$$

Thus we have two interval constraint systems $C1 =< S, D1 >$ and $C2 =< S, D2 >$. The consistency algorithm will map the state $D1$ of the constraint system $C1$ to empty. That is, we will obtain a failure state from $D1$. This indicates that $D1$ does not contain any solution. Applying the consistency algorithm to $D2$, we obtain as corresponding consistent state

$$0.78615137775742[3, 4]^2 \times 0.61803398874989[4, 5] \times 0.38196601125010[5, 6].$$

[2] This is a notation invented by M. H. van Emden [84]. We use this notation instead of the

input: a constraint system $C = < S, D >$
output: inconsistency or a consistent state D'

initialize Q to the queue of all the constraints in S
while Q is not empty **do**
    remove a constraint $r(x_1, \ldots, x_m)$ from Q
    $X := I(x_1) \times \cdots \times I(x_m)$
    apply C_r to X to obtain X', i.e., $X' = C_r(X)$
    if $X' = \emptyset$ **then** exit with inconsistency
    foreach x_i in $\{x_1, \ldots, x_m\}$ **do**
        if $X_i' \neq I(x_i)$ **then**
            $I(x_i) := X_i'$
            foreach $r' \neq r$ and $r' \notin Q$ in which x_i appears **do**
                put r' into Q
        end-if
    end-foreach
end-while
output: $D' = I(x_1) \times \cdots \times I(x_n)$

Figure 2.1: Consistency algorithm

This suggests that a unique solution for x, y and z is contained in this state.

The above states are obtained from BNR-Prolog [10, 6], which includes an implementation of the consistency algorithm for interval constraint systems.

2.2.4 Interval Constraint Programming Languages

So far we have discussed interval constraints as mathematical entities. It is the basis of an interval constraint programming language (or constraint language for short). Among several paradigms for combining the current programming languages with interval constraints, the Constraint Logic Programming (CLP) is the most natural one [58].

Constraint languages are declarative [53, 39, 87]. In a constraint language, it is easy to specify as a constraint system constraints (i.e., equalities and inequalities) of arbitrary complexity and a domain for the variables in the constraints. We call this type of constraint system a *high-level constraint system*. To distinguish it from the constraint system defined in Definition 2.5, we call the latter a *low-level constraint system*. A high-level constraint system can not be solved directly. To solve it, a constraint language compiler or interpretor translates it into a low-level version by

conventional

$$[0.7861513777574 23, 0.7861513777574 24],$$

which is not only cumbersome, but requires close attention to determine the most significant digit at which the two numerals differ. We regard

$$0.78615137775742[3, 4]$$

as a *scaled interval* notation. The numeral before the brackets modifies what is inside by shifting and scaling.

transforming the constraints to a conjunction of primitive constraints and the domain to a state.

For example, the high-level constraint system

$$y = x_1 x_2, \ x_1 + x_2^2 \leq 0, \quad x_1, x_2 \in [0, +\infty]$$

will be translated into a low-level one, of which the constraint conjunction is

$$y = x_1 * x_2 \ \wedge \ y_1 = x_2^2 \ \wedge \ y_2 = x_1 + y_1 \ \wedge \ y_2 \leq 0,$$

and the state is

$$I(x_1) \times I(x_2) \times I(y) \times I(y_1) \times I(y_2)$$

where $I(x_1) = I(x_2) = [0, +\infty]$ and $I(y) = I(y_1) = I(y_2) = [-\infty, +\infty]$. Note that the domains for the auxiliary variables y_1 and y_2 are $[-\infty, +\infty]$. In the high-level constraint system, if the domain of a variable is not specified, then the domain for that variable in the corresponding low-level constraint system will be also $[-\infty, +\infty]$.

Constraint languages provide operations on a given high-level constraint system. The operations include:

- adding an equality or inequality to the constraint system.
- splitting the interval for a variable in the constraint system into two or more sub-intervals. Using (1) the constraints of the constraint system, (2) one of these sub-intervals for this variable, and (3) the intervals for the other variables in the constraint system, one can specify a new constraint system.

Constraint languages also allow one to explicitly or implicitly invoke the consistency algorithm to provide information about possibly existing solutions of a high-level constraint system. The amount of information ranges from nil to near the maximum allowed by the machine numbers.

To simplify notation, we will use the term *constraint system* for either a high-level constraint system or a low-level constraint system. Which one it represents should be determined from the context.

2.3 Computing Interval Functions in Interval Constraints

The result of an interval function of f over a box X calculated in interval arithmetic contains the range of f over X. The value of the interval function can also be computed in interval constraints. In this way we can compare interval arithmetic with interval constraints within the same software system. This is a fundamental requirement for the research reported in this dissertation.

Example. Suppose that we have function

$$f(x_1, x_2) = x_1 x_2 + x_2.$$

Using interval arithmetic,

$$\begin{aligned} f([-5,5],[-10,10]) &= [-5,5] * [-10,10] + [-10,10] \\ &= [-50,50] + [-10,10] \\ &= [-60,60]. \end{aligned}$$

The same result can be obtained in interval constraints.

At first, the function f can be equivalently translated to the following conjunction of primitive constraints

$$S = \text{times}(x_1, x_2, y_1) \wedge \text{sum}(y_1, x_2, y)$$

i.e., S is

$$y_1 = x_1 * x_2 \;\wedge\; y = y_1 + x_2$$

where y_1 and y are auxiliary variables introduced; y is the variable for the function; constraints *times* and *sum* correspond to the interval operations $*$ and $+$ respectively.

And then, associate each variable in the constraint conjunction with an interval. Associate x_1 with $X_1 = [-5, 5]$, x_2 with $X_2 = [-10, 10]$, the auxiliary variables y, y_1 with $Y, Y_1 = [-\infty, +\infty]$ respectively.

Thus we get the interval constraint system $C =< S, X_1 \times X_2 \times Y \times Y_1 >$.

The execution of the consistency algorithm with input C yields the consistent state $D' = [-5, 5] \times [-10, 10] \times [-60, 60] \times [-50, 50]$, where the interval for y is $[-60, 60]$, which is the same as that of $f([-5, 5], [-10, 10])$ computed in interval arithmetic. Initially, Q in the algorithm contains two primitive constraints "times(x_1, x_2, y_1)" and "sum(y_1, x_2, y)". From the initial state

$$X_1 \times X_2 \times Y \times Y_1 = [-5, 5] \times [-10, 10] \times [-\infty, +\infty] \times [-\infty, +\infty]$$

the consistent state D' can be reached in the following two iterations.

1. remove "times(x_1, x_2, y_1)" from Q and apply C_{times} to $X_1 \times X_2 \times Y_1$

$$\begin{aligned} C_{times}(X_1 \times X_2 \times Y_1) &= C_{times}([-5, 5] \times [-10, 10] \times [-\infty, +\infty]) \\ &= [-5, 5] \times [-10, 10] \times [-50, 50] \end{aligned}$$

The state of the constraint system becomes

$$X_1 \times X_2 \times Y \times Y_1 = [-5, 5] \times [-10, 10] \times [-\infty, +\infty] \times [-50, 50].$$

2. remove "sum(y_1, x_2, y)" from Q and apply C_{sum} to $Y_1 \times X_2 \times Y$

$$\begin{aligned} C_{sum}(Y_1 \times X_2 \times Y) &= C_{sum}([-50, 50] \times [-10, 10] \times [-\infty, +\infty]) \\ &= [-50, 50] \times [-10, 10] \times [-60, 60] \end{aligned}$$

The state of the constraint system becomes

$$X_1 \times X_2 \times Y \times Y_1 = [-5, 5] \times [-10, 10] \times [-60, 60] \times [-50, 50].$$

Definition 2.14:

Let E be the expression of f used for the interval function F. The constraint conjunction S and the variable y translated from E are defined as follows:

1. If E is a variable v, then it is only translated to v.

2. If E is a constant, then it is translated to the variable y.

3. If (1) E is (E_1 op E_2), (2) E_1 is translated to the constraint conjunction S_1 and variable y_1, (3) E_2 to S_2 and y_2, and (4) y_1 is not in S_2 and y_2 is not in S_1, then E is translated to the constraint conjunction $S = A_{op}(y_1, y_2, y) \wedge S_1 \wedge S_2$ and the variable y, where A_{op} is the primitive constraint corresponding to the interval operation op, y is different from y_1, y_2 and any other variable in $S_1 \wedge S_2$.

4. If (1) E is op(E_1) and (2) E_1 is translated to the constraint conjunction S_1 and variable y_1, then E is translated to the constraint conjunction $S = A_{op}(y_1, y) \wedge S_1$ and the variable y, where A_{op} is the primitive constraint corresponding to the interval operation op, y is different from y_1 and any other variable in S_1.

Definition 2.15:

Let

1. E be the expression of f used for the interval function F,
2. S be the constraint conjunction translated from E as defined in Definition 2.14,
3. x_i $(i = 1, \ldots, n)$ be the variables in E,
4. x_{n+j} $(j = 1, \ldots, m)$ be the variables translated from constants in E, and
5. x_{n+m+k} $(k = 1, \ldots, t)$ be the auxiliary variables introduced in S.

The constraint system C translated from E over $X = X_1 \times \cdots \times X_n \in \mathcal{F}_n$ is defined as $< S, I_1 \times \ldots \times I_{n+m+t} >$, where I_i $(i \in \{1, \ldots, n\})$ is X_i, I_{n+j} $(j \in \{1, \ldots, m\})$ is the constant from which the variable x_{n+j} is translated to, and I_{n+m+k} $(k \in \{1, \ldots, t\})$ is $[-\infty, +\infty]$.

Lemma 2.7:
If $op \in \{+, -, *, /\}$, then for any $A, B \in \mathcal{F}_1$,

1. A op B, the result of the interval operation op, can be obtained by applying the corresponding consistency operator to the box consisting of A, B and $[-\infty, +\infty]$;
2. applying the operator to the box does not change A or B.

Proof: Let x, y and z be the variables occurring in the primitive constraint corresponding to op. Let furthermore A and B be the intervals for x and y respectively, $[-\infty, +\infty]$ the interval for z.

For $A + B$, suppose that "$sum(x, y, z)$" is the corresponding primitive constraint and C_{sum} is the consistency operator of sum. From Lemma 2.2, we have the following by applying C_{sum} to $A \times B \times [-\infty, +\infty]$.

$$
\begin{aligned}
& C_{sum}(A \times B \times [-\infty, +\infty]) \\
= \; & A \cap ([-\infty, +\infty] - B) \times B \cap ([-\infty, +\infty] - A) \times [-\infty, +\infty] \cap (A + B) \\
= \; & (A \cap [-\infty, +\infty]) \times (B \cap [-\infty, +\infty]) \times (A + B) \\
= \; & A \times B \times (A + B).
\end{aligned}
$$

Thus, the interval for z is the result of $A + B$ and applying C_{sum} to the box $A \times B \times [-\infty, +\infty]$ does not change A or B.

Similarly, we can prove that (1) $A - B$, $A * B$ and A/B can be obtained by applying the corresponding consistency operator to the box consisting of A, B and $[-\infty, +\infty]$ and (2) applying the consistency operator to the box does not change A or B. ∎

Theorem 2.6:
For any $X \in \mathcal{F}_n$, if E is the expression of f used for the interval function F, y is the variable translated from E, and C is the constraint system translated from E over X as defined in Definition 2.15, then the output of the consistency algorithm with input C will be a consistent state D' and the interval in D' for y equals the value of $F(X)$ computed in interval arithmetic.

Proof: The proof is by induction on the number m of interval operations in E.

If $m = 0$, then E is a variable or a constant. There is no primitive constraint in C. If E is the variable x and the interval for x is X, the consistency algorithm will output the consistent state $D' = X$. The interval for y (note that y is the same as x)

in D' is X, which equals the value of $F(X)$ computed in interval arithmetic. If E is the constant $[a,b]$, the algorithm will output $D' = [a,b]$. The interval for y in D' is $[a,b]$, which is the same as $F(X) = [a,b]$.

Suppose that for any $m \leq k$, the theorem is true.

If $m = k+1$, i.e., the number of interval operations in E is $k+1$, then E has the form $(E_1 \text{ op } E_2)$ or $\text{op}(E_1)$. For the former, let y_i and $C_i =< S_i, I_{1_i} \times \cdots \times I_{s_i} >$ be translated from E_i over X^i $(i = 1,2)$, where X^i is the projection of X on the subset of the variables that occur in E_i, and $S_i = A_{1_i} \wedge \cdots \wedge A_{m_i}$. Suppose that A_{op} is the primitive constraint corresponding to the interval operation op and y is the variable translated from E, then $C =< S, I_1 \times \cdots \times I_s >$, where $S = S_1 \wedge S_2 \wedge A_{op}(y_1, y_2, y)$, is the constraint system translated from E over X.

It has been shown in [66] that the result of executing the consistency algorithm with a constraint system C as its input does not depend on the order in which primitive constraints are chosen from the constraint queue Q in Algorithm 2.1. Thus we can select constraints in the following order:

1. as long as there exists a primitive constraint of C_1 in Q, choose that constraint;
2. if there is a primitive constraint of C_2 in Q and no constraint of C_1 in Q, then choose the constraint of C_2;
3. choose A_{op} if there is no constraint of C_1 or C_2 in Q.

From Lemma 2.7, applying $C_{A_{op}}$ does not change the intervals for y_1 and y_2, so it does not lead to any constraints of C_1 or C_2 to be put back in Q. For each constraint of C_1 (C_2), applying the corresponding consistency operator does not affect any constraint of C_2 (C_1) since there is not an auxiliary variable shared by C_1 and C_2

according to Definition 2.14. Thus, the execution of the algorithm with input C until there is not any constraint of C_1 in Q is equivalent to the execution of the algorithm with the input C_1. And then the execution of the algorithm with input C until there is not any constraint of C_2 in Q is equivalent to the execution of the algorithm with the input C_2.

Suppose that E_i $(i = 1, 2)$ is the expression used for the interval function F_i. Since the number of interval operations in E_i is less than or equal to k, according to induction hypothesis, the execution of the consistency algorithm with the input C_i $(i = 1, 2)$ will output the consistent state D'^i and the interval for y_i in D'^i equals the value of $F_i(X^i)$ computed in interval arithmetic.

Thus from the interval for y is $[-\infty, +\infty]$ and Lemma 2.7, after applying the consistency operator of A_{op}, the consistency algorithm will output the consistent state D', in which the interval Y for y is:

$$Y = Y_1 \text{ op } Y_2,$$

where Y_i is the interval for y_i in D'^i. Therefore

$$Y = F_1(X^1) \text{ op } F_2(X^2) = F(X),$$

i.e., the interval for y in D' equals $F(X)$ computed in interval arithmetic.

Similarly, we can prove that the theorem is true if E has the form $\text{op}(E_1)$. ∎

Chapter 3

Hypernarrowing

Hypernarrowing is built on a consistency algorithm. It can be used to narrow the given intervals in a constraint system. The interval Y for F, hypernarrowed by the hypernarrowing, bounds the range of f closer than the corresponding interval function in interval arithmetic. We will present its origin and a hypernarrowing algorithm. And then we compare the semantics of hypernarrowing with that of a constraint solver, which is followed by applications.

3.1 Where the Idea of Hypernarrowing Comes From

Given a function f, an interval function F for f and an element $X \in \mathcal{F}_n$, the value Y of $F(X)$ computed in interval arithmetic is generally much larger than $\Box f(X)$, the smallest interval containing the range of f over X. Since $F(X)$ can be translated to a constraint system C, and after a consistency algorithm is applied to C, the interval for Y in C is the same as the value of $F(X)$, why not probe the large interval for Y obtained from applying the consistency algorithm to C by means of constraints on Y? We show by example how to obtain a smaller interval for Y by adding constraints to the constraint system C.

Suppose we have the following function

$$f(x) = 12x_1^2 - 6.3x_1^4 + x_1^6 + 6x_2(x_2 - x_1)$$

Here we choose the natural interval extension of f (see Definition 2.4) as the inclusion function for f (see Definition 2.3). The BNR-Prolog system [10] for interval constraints allows one to enter the definition of the relation `f` defined as

```
f(X1,X2,Y) :- Y is 12*X1**2 - 6.3*X1**4 + X1**6 + 6*X2*(X2 -X1).
```

Read this as: "The relation `f` holds between `X1`, `X2` and `Y` if `Y` equals the polynomial shown." The definition of the relation `f` is translated to an equivalent conjunction of primitive constraints, typically introducing auxiliary variables.

For $[-2, 4] \times [-2, 4] \in \mathcal{F}_2$, we can enter the query

```
?- X1:real(-2, 4), X2:real(-2, 4), f(X1,X2,Y).
```

Because we leave `Y` initially unconstrained, the consistency algorithm yields an interval for `Y` equal to the result of the interval function computed in interval arithmetic. In this example the interval for `Y` is $[-1757, 4432]$.

When we add the constraint "$Y < -137$" to the constraint system by entering the query

```
?- X1:real(-2, 4), X2:real(-2, 4), f(X1,X2,Y), Y < -137.
```

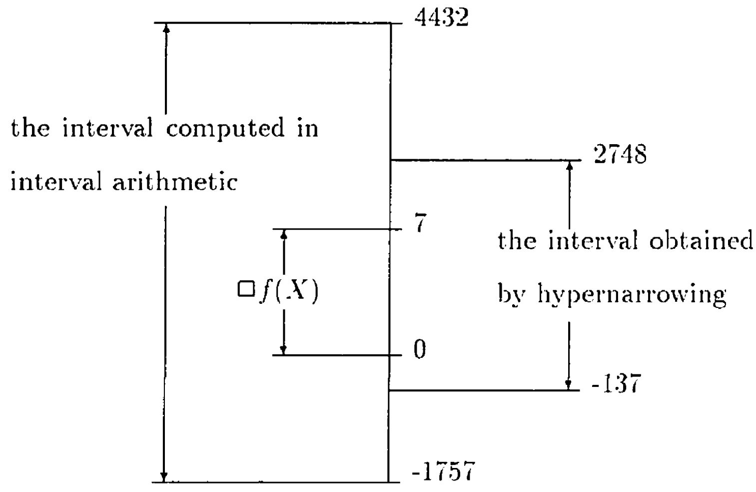


Figure 3.1: Intervals for $f(x) = 12x_1^2 - 6.3x_1^4 + x_1^6 + 6x_2(x_2 - x_1)$ over $[-2,4] \times [-2,4]$

The BNR-Prolog reports that the constraint system has no solution. This implies that there are no solutions for "$f(x_1, x_2) < -137$" within the given intervals for x_1 and x_2. Therefore, -137 is a lower bound of f over the given domain $[-2,4] \times [-2,4]$. When we add the constraint "$Y < -136$", the BNR-Prolog leaves open the possibility of solutions. Similarly, 2748 is a upper bound of f over the given domain. Thus with the interval constraint system, we can improve the interval $[-1757, 4432]$, obtained from the interval function used, to $[-137, 2748]$. W. J. Older discovered this idea and implemented it in "absolve" [67]. The hypernarrowing is a generalization of this idea, which will be described later.

By applying hypernarrowing to the interval for Y in the constraint system translated from an interval function, we can obtain an interval for Y that contains the range of f that is smaller than the one obtained by computing the interval function in interval arithmetic.

Figure 3.1 illustrates three kinds of intervals for the given function f over domain $X = [-2, 4] \times [-2, 4] \in \mathcal{F}_2$. The smallest interval is $\Box f(X)$. The largest interval is the one obtained from the interval function F of f in interval arithmetic. The one in between is the interval obtained by hypernarrowing the interval Y for F in the constraint system translated from the interval function. This middle one gives a better approximation of $\Box f(X)$.

3.2 A Hypernarrowing Algorithm

The hypernarrowing algorithm is based on a consistency algorithm. Suppose that we have a consistency algorithm, **Narrowing**, which takes the constraint system $C = < S, I_1 \times \cdots \times I_n >$ as its input, and $I_1' \times \ldots \times I_n'$ as its output, where $I_1, \ldots, I_n \in \mathcal{F}_1$. $I_1' \times \cdots \times I_n'$ is the narrowed version of $I_1 \times \cdots \times I_n$. The hypernarrowing algorithm takes as its input (1) the constraint system C, (2) a nonempty subset S_I of $\{I_1, \ldots, I_n\}$, in which we want the hypernarrowing algorithm to hypernarrow all the intervals, and (3) a predetermined tolerance ϵ. Its output is the hypernarrowed version of $I_1 \times \cdots \times I_n$.

The hypernarrowing algorithm is described in Figure 3.2. In the algorithm, the variable S_a is used for storing the intervals which the hypernarrowing algorithm will hypernarrow, S_d for storing intervals temporarily. Initially, $S_a = S_I$, $S_d = \emptyset$.

At first, the algorithm chooses an interval I_c from S_a, moves I_c from S_a to S_d, and hypernarrows the interval I_c. And then, for each narrowed interval I_i in the constraint system, it moves all the intervals in S_d which are related to I_i from S_d to S_a. This procedure is repeated until S_a is empty, i.e., all the intervals, which we want the hypernarrowing algorithm to hypernarrow, can not be narrowed any more.

```
Hypernarrowing(< S, I_1 × ··· × I_n >, S_I, ε, output)
begin
        S_a := S_I; S_d := ∅;
        while  S_a ≠ ∅  do
                choose an interval I_c from S_a;
                move I_c from S_a to S_d;
                force(I_c, < S, I_1 × ··· × I_n >, ε, output);
                foreach  narrowed I_i ∈ output  do
                        move all I_j related to I_i from S_d to S_a.
        end-while
end
```

```
force(I_c, < S, I_1 × ··· × I_n >, ε, output)
begin
        forceL(I_c, < S, I_1 × ··· × I_n >, ε, output1);
        forceU(I_c, < S, output1 >, ε, output);
end
```

Figure 3.2: A hypernarrowing algorithm

For a given constraint system C and an interval I_c in C, to hypernarrow the interval I_c is to hypernarrow the lower part and upper part of I_c. Two procedures named *forceL* and *forceU* hypernarrow the lower part and upper part of I_c respectively. (see Figure 3.3). Since these two procedures are symmetric, we only explain the procedure forceL.

For a given interval I_c and the constraint system

$$< S, I_1 \times \cdots \times I_{c-1} \times I_c \times I_{c+1} \times \cdots \times I_n >$$

at first, *forceL* tries to find a lower part I_{c_1} of the interval I_c such that the constraint system

$$< S, I_1 \times \cdots \times I_{c-1} \times I_{c_1} \times I_{c+1} \times \cdots \times I_n >$$

is not consistent. And then, it applies the consistency algorithm to the following constraint system :

$$< S, I_1 \times \cdots \times I_{c-1} \times I_{c_2} \times I_{c+1} \times \cdots \times I_n >$$

where the interval $I_{c_2} = \{x \mid x \in I_c, x \notin I_{c_1}\}$. Thus the interval I_c can be at least hypernarrowed to I_{c_2} and the other intervals in the constraint system may also be narrowed if a nonempty interval I_{c_1} is found.

3.3 Waltz, Solve and Hypernarrowing

Waltz algorithm is a basic consistency algorithm. Both solve and hypernarrowing are built on the Waltz algorithm.

forceL($I_c, < S, I_1 \times \cdots \times I_n >, \epsilon$, output)

begin

  $I_{c_1} := I_c$; $I_{c_2} := \emptyset$; output1 $:= I_1 \times \cdots \times I_n$;

  while output1 $\neq$ empty and width(I_{c_1})$> \epsilon$ **do**

    bisect I_{c_1} into two intervals $I_{c_{11}}$ and $I_{c_{12}}$;

    $I_{c_1} := I_{c_{11}}$; $I_{c_2} := I_{c_{12}} \cup I_{c_2}$;

    Narrowing($< S, I_1 \times \cdots \times I_{c-1} \times I_{c_1} \times I_{c+1} \times \cdots \times I_n >$, output1)

  end-while

  if width(I_{c_1}) $> \epsilon$ **then**

   Narrowing($< S, I_1 \times \cdots \times I_{c-1} \times I_{c_2} \times I_{c+1} \times \cdots \times I_n >$, output)

  else output $:= \langle I_1, \ldots, I_n \rangle$

end

forceU($I_c, < S, I_1 \times \cdots \times I_n >, \epsilon$, output)

begin

/***** symmetric to forceL *****/

end

Figure 3.3: Algorithms for hypernarrowing the lower and upper part of an interval

3.3.1 Comparing Waltz and Hypernarrowing

Although hypernarrowing is built on the Waltz algorithm, the iteration procedures in these two algorithms are the same. The differences between them exist in each step of the iterations.

1. Waltz maintains a queue Q of constraints, hypernarrowing a queue Q_I of intervals.

2. Waltz removes a constraint A_i from Q and applies the function similar to C_{A_i} to the domains of variables in A_i, hypernarrowing removes an interval I_k (the domain of a variable in some constraint) from Q_I and applies the function *force* to I_k.

3. For each narrowed interval I_j of the variable X_j in A_i, Waltz adds every constraint $A_s \neq A_i$ that contains X_j and is not in Q into Q; for each narrowed interval I_j, hypernarrowing adds every interval $I_s \notin Q_I$ that is related to I_j through a constraint into Q_I.

The hypernarrowing algorithm makes an improvement on Waltz. For example, suppose that we want to find the roots of the following function over [-1.0e+15,1.0e+15]$\times$ [-1.0e+15,1.0e+15].

$$f(x_1, x_2) = x_1^2 + x_1 x_2 + x_2^2$$

The solution of $f(x_1, x_2) = 0$ given by BNR-Prolog with the consistency algorithm similar to Waltz is the interval [-1.0e+15,1.0e+15] for x_1 and x_2. Here is the BNR-Prolog query for finding the solution.

```
?- [X1,X2]:real(-1.0e+15, 1.0e+15), X1**2 + X1*X2 + X2**2 == 0.
```

However, the solution given by our hypernarrowing algorithm with $\epsilon = 0.1$ is the interval $[-0.04, 0.04]$ for x_1 and $[-0.04, 0.04]$ for x_2. Here follows the query with hypernarrowing.

```
?- [X1,X2]:real(-1.0e+15, 1.0e+15), X1**2 + X1*X2 + X2**2 == 0,
     hypernarrowing([X1,X2], 0.1).
```

3.3.2 Comparing "Solve" with Hypernarrowing

Lemma 3.1:

The set $\mathcal{F}_n$ together with the set inclusion as the binary relation is a poset.

Figure 3.4 is a rough representation of the poset $\mathcal{F}_n$. In the poset $\mathcal{F}_n$, R^n is the largest element and $\emptyset$ is the least one.

Certain n-ary relations on R are singletons. They contain only one tuple. Let s be one such. In general, the least element $\breve{s} \in \mathcal{F}_n$ such that $\breve{s} \supset s$ is not a singleton. We call it the *approximation singleton*, being the image of a singleton under approximation. In the poset $\mathcal{F}_n$ there is nothing between $\emptyset$ and the approximation singletons.

Suppose that an n-ary relation $r \subset R^n$ is defined by a conjunction S_r of primitive constraints. For any $u \in \mathcal{F}_n$, $\Psi_{S_r}(u)$ is a fixpoint under u. But this fixpoint often gives too little information about the tuples in $r \cap u$.

In order to get more information about the tuples in $r \cap u$ for any $u \in \mathcal{F}_n$, **"solve"** splits u into $u_1, \ldots, u_k$ ($u_1 \cup \cdots \cup u_k = u$ and $u_1 \cap \cdots \cap u_k = \emptyset$). In general, $u_1, \ldots, u_k$ are not fixpoints. Hence $\Psi_{S_r}(u_1), \ldots, \Psi_{S_r}(u_k)$ can be usefully lower than $u_1, \ldots, u_k$ respectively. Because $u_1, \ldots, u_k$ are disjoint, so are $\Psi_{S_r}(u_1), \ldots, \Psi_{S_r}(u_k)$. If $\Psi_{S_r}(u_i)$

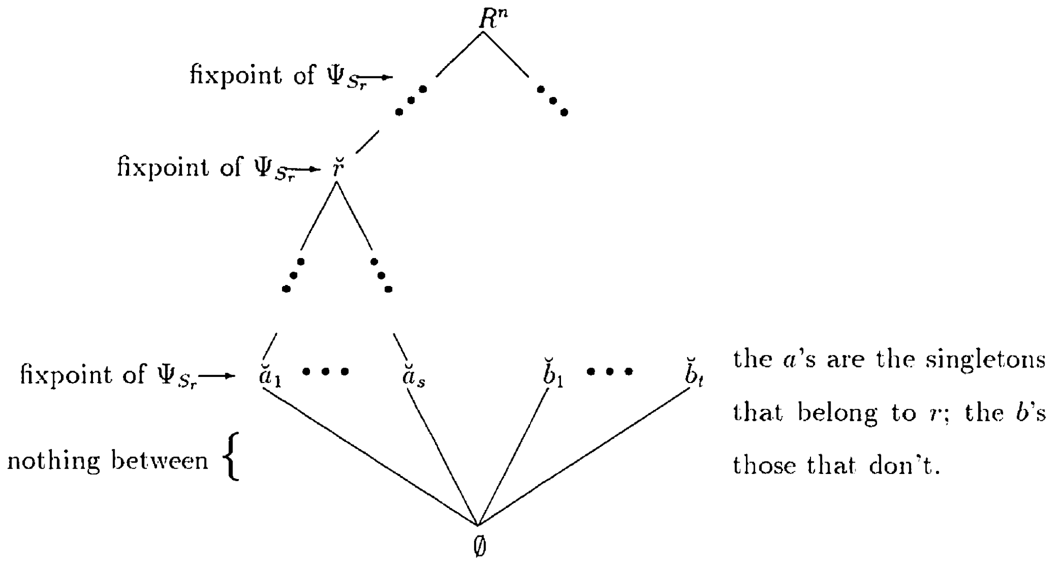


Figure 3.4: A rough map of the poset $\mathcal{F}_n$

($i \in \{1, \ldots, k\}$) is not low enough, **"solve"** will split u_i into sub-boxes. This process is performed on each sub-box u_s of u until $\Psi_{S_r}(u_s)$ is empty or low enough. If we make the partition fine enough, then $\Psi_{S_r}(u_s)$ will be either $\emptyset$ or an approximation singleton that may contain one or more tuples in $r \cap u$. That is what **"solve"** attempts to do (see Figure 3.5).

Hypernarrowing is different (see Figure 3.6). It searches for a partition that splits u into two subsets u_1 and u_2 such that $\Psi_{S_r}(u_2) = \emptyset$. One way to obtain u_2 consists of (1) selecting an interval I_i of u, (2) using binary search to find the largest possible low (or upper) part I_{i_1} of the interval I_i such that $\Psi_{S_r}(u_2) = \emptyset$, where

$$u_2 = I_1 \times \cdots I_{i-1} \times I_{i_1} \times I_{i+1} \cdots \times I_n.$$

u_1 is in general not a fixpoint, so $\Psi_{S_r}(u_1) \neq u_1$. Hypernarrowing is to repeat this

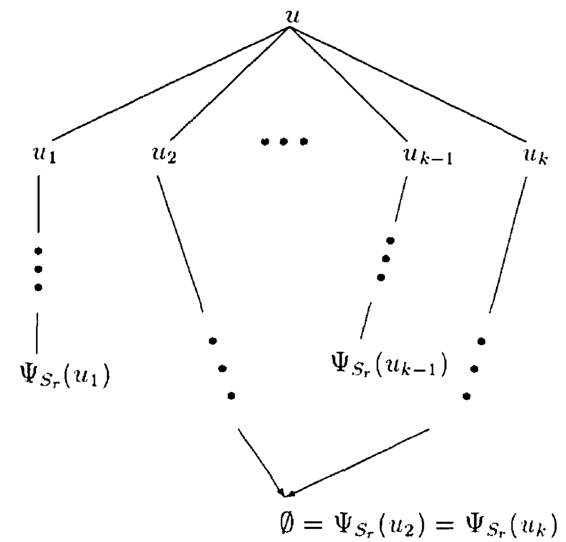


Figure 3.5: The behavior of "solve"

starting at $\Psi_{S_r}(u_1)$ until we can not find I_{i_1} whose width is greater than a given tolerance for any interval of u or sub-box of u. The best result of hypernarrowing is the least $\mathcal{F}$-box containing $r \cap u$.

3.4 Applications of the Hypernarrowing

Interval arithmetic has been used to solve many problems. However, most applications of interval arithmetic use the largest interval mentioned in Figure 3.1. Why not try the middle one which is obtained by hypernarrowing?

Global optimization using interval arithmetic is a typical application. In order to compare the differences between using the largest interval and the hypernarrowed

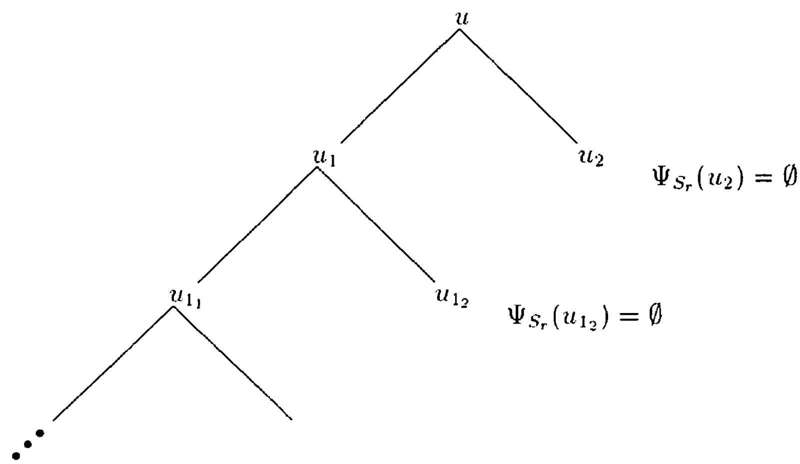


Figure 3.6: The behavior of hypernarrowing

one, at first, we implemented in BNR-Prolog a simple version of Hansen's algorithm [36] for unconstrained global optimization, which uses the largest interval. The simple version does not use the monotonicity, convexity tests, or the one step interval Newton method. And then we wrote a BNR-Prolog program to implement an improved version, which uses the middle interval in Figure 3.1. The only difference between these two programs is that the former uses the largest interval obtained by interval arithmetic and the latter uses the one obtained by hypernarrowing. The two programs are tested on the test problems used for testing the Hansen's algorithm. In order to avoid ambiguity, we denote the former by *Program A*, the latter by *Program B*. The computational results (see table 3.1) show that the Program B achieves a factor of two to eighteen in inclusion function evaluations. It is 1.8 to 9.7 times as fast as the Program A. However, compared to the program using interval arithmetic, the

		Problem 1	Problem 2	Problem 28	Problem 29
Input	W	6	8	3.78	2.4
	ϵ	10^{-1}	10^{-2}	10^{-2}	10^{-2}
Program A	N	2067	319	161	61
	t	505.9	49.5	27.7	5.4
Program B	N'	161	17	45	27
	t'	52.2	5.8	8.8	3
Comparison	N/N'	12.8	18.7	3.5	2.2
	t/t'	9.7	8.5	3.1	1.8

Table 3.1: Running results of Program A and Program B

program using interval constraints but not hypernarrowing only achieves a factor of two to five in inclusion function evaluations and is 1.8 to 4.9 times faster [84, 16]. The more interesting thing is that the execution of the Program B needs much less memory space than the Program A. All this suggests that we can improve most of the applications of interval arithmetic by using the middle interval which is obtained by hypernarrowing.

Here is an explanation of the symbols used in the table 3.1.

W Width of initial box.

ϵ Intended absolute accuracy.

N Number of all inclusion function evaluations till termination.

t Running time.

Chapter 4

Unconstrained Global Optimization

Given a function $f : R^n \to R$ and a domain $X \in \mathcal{F}_n$ ($\mathcal{F}_n$ is defined in equation 2.1), the unconstrained global optimization problem is to find

$$f^* = \min \{f(x) \mid x \in X\} \tag{4.1}$$

where f^* is the global minimum. We study interval constraint methods for seeking the minimum value f^* of f and the locations x^* at which this minimum value occurs. The interval constraint methods are based on interval arithmetic methods. Like the interval arithmetic methods, they can give a solution with guaranteed accuracy.

After reviewing the interval arithmetic methods and describing a Branch-and-Bound algorithm for solving 4.1, we prove the properties of the algorithm. We analyze the role of interval constraints in global optimization. Next we describe the implementations of both the interval arithmetic version and the interval constraint version of the algorithm. The implementations are done in the constraint logic programming language BNR-Prolog [10, 67, 7], and are followed by computational results. Finally, we introduce improvements on the memory use of the Branch-and-Bound algorithm.

4.1 Overview of Interval Arithmetic Methods

Given a function $f : R^n \rightarrow R$ and a domain $X \in \mathcal{F}_n$, R. E. Moore [64] discovered that the combination of an inclusion function of f with a certain method of partitioning X could be used to determine the range and thus the global minimum and maximum values of the function f. S. Skelboe [78] was able to reduce the number of inclusion function evaluations used in Moore's method by combining Moore's method with the *Branch-and-Bound* principle. H. Ratschek and J. Rokne [74] call this method Moore-Skelboe algorithm.

The Moore-Skelboe algorithm first partitions the initial box X into smaller sub-boxes. The search for the global minimum f^* is performed iteratively by (1) selecting those sub-boxes, for which the lower bound on the function value is the least, and (2) partitioning these sub-boxes further. It is more likely that these boxes contain a global minimizer x^*.

For some sub-boxes, a simple test can show that they do not contain any global minimizer. But the Moore-Skelboe algorithm does not eliminate any. An improvement to the Moore-Skelboe algorithm was made by K. Ichida and Y. Fujii [43]. Their method combined Branch-and-Bound with a test that allows one to reject a sub-box. In the following, we discuss this test and some other tests proposed by E. R. Hansen [34, 35, 36].

Midpoint Test. Suppose that f^*_{ub} is the lowest upper bound of the global minimum obtained by evaluating the function f at the *midpoint* of each sub-box and choosing the smallest value. For a given box B, if we can determine that

$$f(x) \leq f^*_{ub} \quad \text{for some } \; x \in B \tag{4.2}$$

does not hold, then no minimizer can be in B. Thus B can be removed from further consideration.

We call the inequality 4.2 *midpoint condition*. One way to determine that the midpoint condition does not hold is to see whether the following inequality holds:

$$lb(f(B)) \leq f^*_{ub} \tag{4.3}$$

where $f(B)$ is the interval value of the natural interval extension of the function f over B. If the inequality 4.3 does not hold, then the midpoint condition does not hold and we can reject B without sacrifice of correctness. This way of rejecting boxes is called *midpoint test*.

Stationarity Test. The midpoint test only uses the information about the function f. It can and should always be used in Branch-and-Bound to reject sub-boxes. The stationarity test uses the gradient of f to determine whether a given box B can be rejected. The use of this test in Branch-and-Bound is subject to the existence of the gradient of f.

Definition 4.1: [29]

Suppose that

$$g_i(x) = \frac{\partial f(x)}{\partial x_i} \quad (i = 1, \ldots, n).$$

The vector

$$\nabla f(x) = (g_1(x), \ldots, g_n(x))^T$$

is the *gradient* of the function f at x, where ∇ is the *gradient operator*.

For a given box B, if we can determine that

$$g_i(x) = 0 \quad \text{for } i = 1, \ldots, n \quad \text{for some } x \in B \tag{4.4}$$

does not hold, then the gradient of f is not zero in B. Thus the global minimum can not occur in B and B can be rejected.

The equality 4.4 is called *stationarity condition*. One way of determining that the stationarity condition does not hold is to check whether the following inequality holds:

$$lb(g_i(B)) \leq 0 \leq ub(g_i(B)) \quad \text{for } i = 1, \ldots, n \tag{4.5}$$

where $g_i(B)$ is the interval value of the natural interval extension of the function g_i over B. If the inequality 4.5 does not hold, then the stationarity condition does not hold and we can reject B for further consideration. This way of discarding boxes is called *stationarity test* or *monotonicity test*.

Convexity Test. The convexity test uses the second-order derivatives of the function f to determine whether a given box B can be rejected. It can be used in Branch-and-Bound if f is twice differentiable.

Definition 4.2: [29]

Suppose that

$$h_{ij}(x) = \frac{\partial^2 f(x)}{\partial x_i \partial x_j} \quad (i, j = 1, \ldots, n).$$

The matrix

$$[h_{ij}(x)], \quad i,j = 1,\ldots,n$$

is the *Hessian matrix* of the function f at x.

Since the Hessian matrix of f at x must be positive semi-definite at a minimizer x^*, we can reject a box B if we can guarantee that the Hessian matrix is *not* positive semi-definite anywhere in B. One necessary condition for the Hessian matrix to be positive semi-definite is that its diagonal elements $h_{ii}(x)$ $(i = 1,\ldots,n)$ are non-negative. If we can determine that the following inequality does not hold,

$$h_{ii}(x) \geq 0 \quad \text{for } \; i = 1,\ldots,n \quad \text{for some } \; x \in B \tag{4.6}$$

then B can be rejected.

We call the inequality 4.6 *convexity condition*. One way to determine whether the convexity condition holds is to see if the following inequality holds:

$$lb(h_{ii}(B)) \geq 0 \quad \text{for } \; i = 1,\ldots,n \tag{4.7}$$

where $h_{ii}(B)$ is the interval value of the natural interval extension of the function h_{ii} over B. If the inequality 4.7 does not hold, then the convexity condition 4.6 does not hold and B can be rejected. This way of rejecting boxes is called *convexity test*.

In addition to these two tests, Hansen used *the linear method*, *the quadratic method* and *the one-pass interval Newton method* to reject a sub-box B' of B. These three methods are based on Taylor's theorem and interval analysis, and will be discussed in the following.

Linear Method. Consider the one-dimensional case for the linear method. From Taylor's theorem, expanding $f(x)$ about a point x_0, we have

$$f(x) = f(x_0) + (x - x_0)f'(\xi) \quad \text{for some } \xi \text{ between } x_0 \text{ and } x. \tag{4.8}$$

For a given box B, a point $x_0 \in B$ and an upper bound f^*_{ub} of f^*, if we can determine that

$$f(x_0) + (x - x_0)f'(\xi) \leq f^*_{ub} \quad \text{for some } x \in B \text{ and some } \xi \text{ between } x_0 \text{ and } x \tag{4.9}$$

does not hold, then B can be rejected. The condition 4.9 is called *linear condition.*

The linear method tries to find a sub-box B' of B such that the linear inequality

$$U + (x - x_0)V \leq 0 \quad \text{for some } \ x \in B' \tag{4.10}$$

does not hold (i.e., there is not any point $x \in B'$ at which the value of f is less than or equal to f^*_{ub}), where x_0 is a point of B, $U = f(x_0) - f^*_{ub}$ and $V = f'(B)$ are constant intervals. After B' is found, we can get the box B'_c complementary to B'. Thus the sub-box B' of B can be rejected through replacing B by $B \cap B'_c$. If $B \cap B'_c = \emptyset$, then B is totally rejected.

Quadratic Method. Still consider the one-dimensional case. Suppose that f has the derivative of the second order. From Taylor's theorem, expanding $f(x)$ about a point x_0, we have

$$f(x) = f(x_0) + (x - x_0)f'(x_0) + \frac{1}{2}(x - x_0)^2 f''(\xi) \quad \text{for some } \xi \text{ between } x_0 \text{ and } x.$$

Given a box B, a point $x_0 \in B$ and an upper bound f^*_{ub} of f^*, if we can determine that

$$f(x_0)+(x-x_0)f'(x_0)+\frac{1}{2}(x-x_0)^2f''(\xi) \le f^*_{ub} \text{ for some } x \in B \text{ and } \xi \text{ between } x_0 \text{ and } x \tag{4.11}$$

does not hold, then B can be rejected. We call the inequality 4.11 *quadratic condition.*

The quadratic method tries to find a sub-box B' of B such that the quadratic inequality

$$f(x_0) - f^*_{ub} + (x - x_0)f'(B) + \frac{1}{2}(x - x_0)^2 f''(B) \le 0 \quad \text{for some } x \in B' \tag{4.12}$$

does not hold. After B' is found, we can get the box B'_c complementary to B'. Thus the sub-box B' of B can be rejected through replacing B by $B \cap B'_c$. If $B \cap B'_c = \emptyset$, then B is totally rejected.

Newton Method. For the Newton method, consider the one-dimensional case. Suppose that g is the gradient of the function f and B is a box. From Taylor's theorem, we have

$$g(x) = g(x^*) + (x - x^*)g'(\xi) \quad \text{for some } \xi \text{ between } x^* \text{ and } x. \tag{4.13}$$

If x^* is the minimizer, then $g(x^*) = 0$. Thus we have

$$x^* = x - \frac{g(x)}{g'(\xi)} \quad \text{for some } \xi \text{ between } x^* \text{ and } x. \tag{4.14}$$

We call this equation *Newton condition.* One way of determining whether there exists $x^*, x \in B$ such that the condition 4.14 holds is to use interval arithmetic to try to find an $x^* \in B$.

Let B contain both x^* and x. Since ξ is between x and x^*, it follows that $\xi \in B$. Therefore, $g'(\xi) \in g'(B)$. Thus $x^* \in x - g(x)/g'(B)$. Based on this fact, the algorithm described in Figure 4.1 can be used to find x^*. The algorithm produces two possible outputs. One is $X = \emptyset$, the other is $X \neq \emptyset$ and X is small enough. In the former case, it is determined that B does not contain any minimizer. Thus B can be rejected. For the latter case, B is replaced by X.

```
input:  g, g', B;
X := B;
while  X is not small enough do
        x := midpoint(X);
        X' := x - g(x)/g'(X);
        X := X ∩ X';
end-while
output:  X = ∅ or X ≠ ∅;
```

Figure 4.1: Newton algorithm in Interval Arithmetic

It is wasteful to use the interval Newton method to iterate to convergence. The reason is that it may be converging to a local minimizer which is not a (global) minimizer. Because of this, Hansen uses one execution of the body of the loop of the Newton method, which is called one-pass interval Newton method [36].

4.2 Branch-and-Bound for Unconstrained Optimization

Branch-and-Bound [56, 77, 81] is an algorithm for finding the minimum of a function over a given set of points. This set may be countable, as in the discrete optimization problems, or the set may be a box in Euclidean n-space, as here. We will consider from the start the continuous instance of the Branch-and-Bound algorithm.

4.2.1 The Major Components of the Algorithm

The essential constituents of the algorithm to determine the global minimum are: (1) partitioning a box into sub-boxes, (2) finding an upper bound f^*_{ub} for f^*, and (3) using conditions to reject sub-boxes.

Partitioning a box into sub-boxes. Let

$$|[a,b]| = b - a$$

denote the width of the interval $[a,b]$. The width of a box $B = X_1 \times \cdots \times X_n$ is defined as the following

$$|B| = \max\{|X_i| \mid i = 1,\ldots,n\}.$$

Given a box $B = X_1 \times \cdots \times X_n$, if $X_k = [a_k, b_k]$ has the largest width among those for $X_1,\ldots,X_n$, and X_k is split into $X_{k_1} = [a_k, c_k]$ and $X_{k_2} = [c_k, b_k]$, where $c_k = a_k + (b_k - a_k)/2$ is the midpoint of the interval X_k, then the box B is partitioned into two sub-boxes

$$B_1 = X_1 \times \cdots \times X_{k-1} \times X_{k_1} \times X_{k+1} \times \cdots \times X_n$$

and

$$B_2 = X_1 \times \cdots \times X_{k-1} \times X_{k_2} \times X_{k+1} \times \cdots \times X_n.$$

Finding an upper bound f_{ub}^* of f^*. Suppose that f_{ub}^* is the lowest upper bound of f^* currently known. Initially $f_{ub}^* = ub(f(\text{mid}(X)))$, where X is the given box over which we want to find the global minimum. For a given sub-box B of X, we may be able to improve f_{ub}^* by arbitrarily selecting a point c in B and obtaining a new f_{ub}^* by

$$f_{ub}^* := \min(f_{ub}^*, ub(f(c))).$$

Using conditions to reject boxes. Given a box B and the lowest upper bound f_{ub}^* of f^* found so far, we can use the midpoint condition 4.2 in any circumstances. If the midpoint condition does not hold, then the box B can be rejected. One way of determining whether the midpoint condition holds is the midpoint test, which is the implementation of the midpoint condition in interval arithmetic. In the following section, we will give another way.

If the function f is differentiable, we can use the stationarity condition 4.4 and the linear condition 4.9 to reject the box under consideration.

When f is twice differentiable, we can use the convexity condition 4.6, the quadratic condition 4.11 and the Newton condition 4.14 to reject the box under consideration.

4.2.2 The Data Structures Used in the Algorithm

The data structures and variables manipulated by the algorithm are:

(1) Floating-point number f^*_{ub}, which is initialized as $ub(f(\mathrm{mid}(X)))$.

(2) A list A and a priority queue L of tuples $< B, l, u >$, where B is a box in which a minimizer may occur, l is a lower bound of f over B, u is an upper bound for the minimum of f over B. The priority queue L is ordered according to increasing l.

A non-empty priority queue allows one to perform a "remove" operation. The result of this operation is the removal of a tuple $< B, l, u >$ with the lowest value of l. The priority queue also allows one to perform an "add" operation, which adds an arbitrary tuple to the priority queue. A priority queue is an "abstract data type" in the sense of [55]. Its significance is based on the fact that certain data structures allow *add* and *remove* to be performed in time $O(\log n)$ where n is the number of elements in the priority queue.

The list A is initially empty and contains on termination the tuples whose boxes are small enough to qualify as part of the final result. Priority queue L contains at any time tuples with boxes that may intersect with the minimizer. Initially, it only contains $< X, lb(f(X)), ub(f(\mathrm{mid}(X))) >$, where X is the box of the given optimization problem. The function f always yields an interval, even when its argument is a point.

4.2.3 The Description of the Algorithm

After the data structures and initializations of the Branch-and-Bound algorithm thus given, we specify the iteration of the algorithm as in Figure 4.2.

The input data for the algorithm are :

- a function f and a given box X,
- a tolerance ϵ_X on box size and a tolerance ϵ_F on function width.

The output data of the algorithm are :

- a lower bound f^*_{lb} and an upper bound f^*_{ub} for f^*,
- a list
$$A = [< B_1, l_1, u_1 >, \ldots, < B_s, l_s, u_s >]$$
where for $i = 1, \ldots, s$, $l_i = lb(f(B_i))$, $u_i = ub(f(mid(B_i)))$, $|B_i| \leq \epsilon_X$ and $(u_i - l_i) \leq \epsilon_F$.

In the algorithm, the key part is "conditions($< B, l, u >, f^*_{ub}$)", which is underlined. The failure of any single one of the conditions implies that B does not contain any minimizer. Hence neither it, nor any subset of it should appear on the answer list A. Thus it can be omitted from L without sacrificing the correctness of the algorithm.

In the beginning, the algorithm initializes f^*_{ub}, L and A. In each iteration, it removes a tuple $< B, l, u >$ from L. If the tuple satisfies the following criterion for the answer list A:

$$|B| \leq \epsilon_X \wedge (u - l) \leq \epsilon_F \tag{4.15}$$

then it will be added into A. Otherwise, an attempt is made to reject the tuple by subjecting it to the conditions. If it is not rejected, then (1) the box B in the tuple will be partitioned into two sub-boxes B_1 and B_2, (2) lower bound l_i ($i = 1, 2$) and upper bound u_i ($i = 1, 2$) of f^* over B_i will be computed, (3) f^*_{ub} will be replaced by $\min(f^*_{ub}, u_1, u_2)$, and (4) two new tuples $< B_1, l_1, u_1 >$ and $< B_2, l_2, u_2 >$ will be added to L.

input: $f, X, \epsilon_X, \epsilon_F$;
$f^*_{ub} := ub(f(\text{mid}(X)))$;
$L := [< X, lb(f(X)), f^*_{ub} >]$;
$A := \emptyset$;
while $L \neq \emptyset$ **do**
    remove $< B, l, u >$ from L;
    if $|B| \leq \epsilon_X \wedge (u - l) \leq \epsilon_F$ **then**
        add $< B, l, u >$ to A
    else if $\underline{\text{conditions}(< B, l, u >, f^*_{ub})}$ **then**
        partition B into B_1 and B_2 with midpoints m_1 and m_2;
        $l_1 := lb(f(B_1))$; $l_2 := lb(f(B_2))$;
        $u_1 := ub(f(m_1))$; $u_2 := ub(f(m_2))$;
        $f^*_{ub} := min(f^*_{ub}, u_1, u_2)$;
        add $< B_1, l_1, u_1 >$ and $< B_2, l_2, u_2 >$ to L;
    end-if
end-while
$f^*_{lb} = \min\{l_i \mid < B_i, l_i, u_i > \in A\}$;
output: f^*_{lb}, f^*_{ub}, A;

Figure 4.2: Branch-and-Bound algorithm for unconstrained optimization

In the algorithm, "partition" is assumed to have the property of that both B_1 and B_2 are non-empty and smaller than B, otherwise the algorithm may not terminate. The algorithm does not handle the case that B can not be partitioned. This case arises only for functions that are pathological in the sense that when the size of box B is less than or equal to the distance between two adjacent floating point numbers, the value of $(u - l)$ is still greater than ϵ_F.

Theorem 4.1:

When the algorithm terminates, it is the case that

$$f^*_{lb} \leq f^* \leq f^*_{ub}, \tag{4.16}$$

$$f^*_{ub} - f^*_{lb} \leq \epsilon_F, \tag{4.17}$$

$$x^* \in B_1 \cup \cdots \cup B_s \tag{4.18}$$

and

$$u_i - f^*_{lb} \leq 2\epsilon_F \quad \text{for any } i \in \{1, \ldots, s\}. \tag{4.19}$$

Proof: Since f^*_{lb} is a lower bound of the function f over the original given domain X and f^*_{ub} is the value of f at the midpoint of a sub-box of X, it is obvious that the inequality 4.16 holds.

Suppose that f^*_{lb} equals to l_j of the tuple $< B_j, l_j, u_j >$ in the answer list A, where $j \in \{1, \ldots, s\}$. From the criterion 4.15 for the answer list A, we have

$$u_j - l_j \leq \epsilon_F$$

i.e.,

$$u_j - f^*_{lb} \leq \epsilon_F.$$

Since f_{ub}^* is the lowest one among all u_i $(i = 1, \ldots, s)$, we have

$$f_{ub}^* - f_{lb}^* \leq \epsilon_F.$$

That is that the inequality 4.17 holds.

The Branch-and-Bound algorithm only rejects the sub-boxes of the original given box X that do not satisfy the conditions for the global minimizers. Thus all the minimizers that occur in X are still in the union of the sub-boxes of X that are not rejected. Therefore we have

$$x^* \in B_1 \cup \cdots \cup B_s.$$

That is that the formula 4.18 holds.

Suppose that the inequality 4.19 does not hold. Thus we have

$$u_k - f_{lb}^* > 2\epsilon_F \quad \text{for some } k \in \{1, \ldots, s\}, \tag{4.20}$$

i.e.,

$$u_k - \epsilon_F > f_{lb}^* + \epsilon_F \quad \text{for some } k \in \{1, \ldots, s\}. \tag{4.21}$$

From the criterion 4.15 for the answer list A, we have

$$u_k - l_k \leq \epsilon_F, \tag{4.22}$$

i.e.,

$$l_k \geq u_k - \epsilon_F. \tag{4.23}$$

From the inequalities 4.23 and 4.21, we have

$$l_k > f_{lb}^* + \epsilon_F. \tag{4.24}$$

From the inequality 4.17, we have

$$f_{lb}^* + \epsilon_F \geq f_{ub}^*. \tag{4.25}$$

From the inequalities 4.24 and 4.25, we have

$$l_k > f^*_{ub}. \tag{4.26}$$

This contradicts with "conditions($< B_k, l_k, u_k >, f^*_{ub}$)". If the inequality 4.26 holds, then the midpoint condition implemented either in interval arithmetic or in interval constraints can not be satisfied and the tuple $< B_k, l_k, u_k >$ will be rejected. Thus $< B_k, l_k, u_k >$ can not be in the answer list A. Therefore, the inequality 4.19 holds. ∎

This is the Branch-and-Bound algorithm for unconstrained global optimization. Its performance varies with the effectiveness of the implementation of "conditions($< B, l, u >, f^*_{ub}$)". If we use the midpoint test, the stationarity test, the convexity test, the linear method, the quadratic method and the one-pass interval Newton method, we obtain Hansen's Branch-and-Bound algorithm for unconstrained global optimization. Since all the tests and methods use interval arithmetic, we call Hansen's algorithm an interval arithmetic version of Branch-and-Bound. In the next section, we translate all the conditions into a pure constraint processing task in Interval Constraints.

4.3 Interval Arithmetic vs Interval Constraints

"conditions($< B, l, u >, f^*_{ub}$)" in the Branch-and-Bound algorithm can be implemented in two ways. One way is in interval arithmetic. The other is in interval constraints. Since BNR-Prolog can simulate interval arithmetic, we can implement these two ways in this same language.

Suppose that Y is the variable for the value of the function f over a box B and F_{ub}

is the lowest upper bound of f^* found so far. The following one line of BNR-Prolog code implements the midpoint condition in interval arithmetic:

```
less(Y, Fub) :- range(Y, [Ylb, Yub]), Ylb =< Fub.
```

This line of code means that the formula "less(Y, Fub)" holds if $lb(Y) \leq F_{ub}$, where $Y = f(B)$. In BNR-Prolog, the implementation of the midpoint condition in interval constraints is as follows:

```
less(Y, Fub) :- Y =< Fub.
```

This line of code means that the formula "less(Y, Fub)" holds if adding the interval constraint "$y \leq F_{ub}$" to the existing constraint system "$y = f(x), x \in B$" does not produce a failure state.

Although these two lines of code look similar, there is a huge difference between the behaviors of the program with one line and that with the other. This difference brings great effect on the performance of the algorithm.

4.3.1 Translating Conditions into an Interval Constraint System

For a given box B and an upper bound f^*_{ub} of f^*, we can create the following interval constraint system from the midpoint condition 4.2:

$$x \in B, \quad f(x) \leq f^*_{ub}. \tag{4.27}$$

For any other applicable conditions, we can add them as interval constraints into the constraint system 4.27 and obtain a new constraint system. For example, if the function f has the first-order derivative, we can add the stationarity condition 4.4 and the linear condition 4.9 as interval constraints into the constraint system 4.27 and obtain the following interval constraint system:

$$x \in B, \quad f(x) \leq f^*_{ub}, \quad f'(x) = 0, \quad f(x_0) + (x - x_0)f'(\xi) \leq f^*_{ub} \tag{4.28}$$

where ξ is between x_0 and x.

Whenever we have an interval constraint system, we can invoke the consistency algorithm described in Figure 2.1 with this constraint system as an input. In BNR-Prolog, when we add an interval constraint into the existing constraint system and obtain a new constraint system, the consistency algorithm is automatically called with this new constraint system as its input. The output of the consistency algorithm will be inconsistency (i.e., failure) or a reduced non-empty box B'. The inconsistency means that some of the conditions translated into the interval constraints in the constraint system can not be satisfied for the given B and f^*_{ub}. That is that we can not find some $x \in B$ such that all the conditions translated into the interval constraints in the constraint system are satisfied. Thus the box B can be rejected. The inconsistency also implies that the box B is reduced to an empty box $\emptyset$. When the box B is reduced to an non-empty box B', we have that some of the conditions can not hold for $B - B'$.

4.3.2 Comparing Interval Arithmetic with Interval Constraints on Small Examples

Let us call *feasible* any box that contains at least one point that satisfies all the applicable conditions. The interval constraints translated from the midpoint condition 4.2, the stationarity condition 4.4 or the convexity condition 4.6 can be used to *reduce* the size of the box B and also, in certain cases, to tell whether the box B in its entirety is infeasible.

The midpoint test, the stationarity test and the convexity test are the interval arithmetic implementations of the midpoint condition, the stationarity condition and the convexity condition respectively. These three tests do not change the size of the box B. They can only be used to determine whether the box B in its entirety is infeasible. If it is infeasible, we can eliminate it. However, if these tests can not tell us the box B in its entirety is infeasible, then the box B is not changed. In this case, these tests do nothing.

By using the linear method, the quadratic method or the Newton method, we may delete the entire box B or some of it. It seems that these methods have the same functionality as the interval constraints corresponding to them. In fact, the interval constraint versions work differently and have more power. To illustrate this, we present some examples; and then explain why interval constraints are more powerful than their corresponding interval arithmetic tests.

Example 1: Interval constraints more powerful for midpoint condition. Suppose that we wish to find the global minimum of

$$f(x_1, x_2) = x_1^2(4 + x_1^2(-2.1 + (1/3)x_1^2)) + 4x_2^2(x_2^2 - 1) + x_1x_2.$$

For the given box $B = [-1.9, -1.8] \times [0.6, 0.8]$ and $f_{ub}^* = -0.8$, the lower bound of $f(B)$ is -2.13, which is below f_{ub}^*. So we can not eliminate B by using the midpoint test, which is the implementation of the midpoint condition in interval arithmetic. Even though the lower bound of $f(B)$ is below f_{ub}^*, the box B is deleted just by the interval constraint $f(x) \leq f^*_{ub}$, which is the implementation of the midpoint condition in interval constraints. The reason is that this constraint reduces the box B to an empty result.

For the given box $B = [0, 1.3] \times [0, 1.5]$ and $f^*_{ub} = -0.2$, the lower bound of $f(B)$ is -9.0, which is below f^*_{ub}. The midpoint test does nothing for this box. By the interval constraint $f(x) \leq f^*_{ub}$, the box $B = [0, 1.3] \times [0, 1.5]$ will be shrunk to $B = [0, 1.3] \times [0.2, 1.0]$, and the interval for $f(x)$ is shrunk to $[-3.6, -0.2]$ from $[-9.0, 20.0]$.

Example 2: Interval constraints more powerful for linear condition. Suppose that we wish to find the global minimum of

$$f(x) = x^6 - 15x^4 + 27x^2 + 250.$$

For the given box $B = [-3.0, -2.0]$ and $f^*_{ub} = 7.00000000000245$, using the linear method, we can not eliminate any part of the box B. But by using the interval constraints corresponding to the linear method, we get box B shrunk to $[-3.0, -2.2]$.

For the given box $B = [-3.147, -3.06067]$ and $f^*_{ub} = 7.00000000000245$, the linear method narrows the box to $[-3.0704, -3.06067]$. The interval constraints corresponding to the linear method produces an empty box. This means that the entire box is deleted.

4.3.3 Comparing Interval Arithmetic with Interval Constraints in General

The examples show that the linear method *can* be less effective than the same idea expressed as interval constraints. Below we show that this is *always* the case.

For a given box B and an upper bound f^*_{ub} of f^*, the linear method tries to delete infeasible points from B, i.e., tries to find a sub-box B' of B such that the linear inequality

$$U + (x - x_0)V \leq 0 \quad \text{for some } \ x \in B'$$

does not hold and removes B' from B, where x_0 is a point of B, $U = f(x_0) - f^*_{ub}$ and $V = f'(B)$ are constant intervals. Since both x and x_0 are contained in B, ξ is in B and $f'(\xi) \in f'(B)$.

If we use interval constraints for unconstrained global optimization, we just use the formula

$$f(x_0) - f^*_{ub} + (x - x_0)f'(\xi) \leq 0 \tag{4.29}$$

as an interval constraint, where ξ is between x_0 and x. It is not necessary to write a function or procedure to solve the linear inequality.

From the point of view of interval constraints, the linear method only considers the following linear constraint indirectly and independently

$$U + (x - x_0)V \leq 0, \quad x \in B$$

where x_0 is a point of B, U and V are constant intervals. $U = f(x_0) - f^*_{ub}$, $V = f'(B)$. Thus only the interval for x may be shrunk. U and V are not changed since they are constants. If we use interval constraints, the nonlinear constraint 4.29 is considered.

Not only may the interval for x be shrunk, but also the interval for V. In the place of V, is $f'(\xi)$, which may get shrunk. This shrunk interval may make the interval for x become smaller further.

In addition, the interval constraints in a constraint system interact with each other. For example, suppose that we have a box V and a constraint $A_1(x)$ in the constraint system, where $x \in V$. If adding constraint $A_2(x)$ (where $x \in V$) makes V shrink, this shrunk V will propagate to A_1, which may make the box V shrink further. This procedure repeats until the box V can not be shrunk or one of two constraints A_1 and A_2 can not be satisfied (i.e., an empty result is produced).

Thus we see that interval constraint versions of the linear condition are more effective than the linear method. Similar considerations will show that the interval constraint versions of the quadratic condition and the Newton condition are also more effective.

4.3.4 Overview of the Implementation Variations

There are six conditions (i.e., the midpoint, the stationarity, the convexity, the linear, the quadratic and the Newton condition). The Branch-and-Bound algorithm described in Figure 4.2 may use any non-empty sub-set of these six conditions. Each of these six conditions can be implemented in two different ways, in interval arithmetic and in interval constraints. There are many combinations.

We will focus on two sets of conditions. One set only contains the midpoint condition. The Branch-and-Bound algorithm using this set of conditions is called *order zero algorithm*, since it only uses the information about the objective function f. The other set includes the midpoint condition and stationarity condition. The Branch-

IAU_0	Interval Arithmetic version of order 0 algorithm
ICU_0	Interval Constraints version of order 0 algorithm
IAU_1	Interval Arithmetic version of order 1 algorithm
ICU_1	Interval Constraints version of order 1 algorithm

Table 4.1: Names of different algorithms

and-Bound algorithm using this set of conditions is called *order one algorithm*, since it uses the information about the objective function f and the first-order derivative of f.

Thus there are four combinations. We have four names for these combinations (see Table 4.1) if we use strings IAU and ICU to represent the Interval Arithmetic and Interval Constraint version of algorithm for Unconstrained global optimization, and subscript 0 and 1 to represent order 0 and 1 algorithm respectively.

4.3.5 Implementations of the Algorithms

We first present a generic implementation of the algorithm described in Figure 4.2. The implementations of IAU_0, ICU_0, IAU_1 and ICU_1 can be easily derived from this generic implementation. Figure 4.3 shows the main part of the BNR-Prolog program for this generic implementation. We omit the definitions of *partition*, *small*, *fb* and *min*. The user should provide the definitions of f, g_i and h_{ii} $(i = 1, \ldots, n)$.

For example, suppose that we want to find the minimum value of the following function

$$f(x_1, x_2) = 4(x_1 - 5)^2 + (x_2 - 6)^2.$$

We can define relation f in BNR-Prolog as follows:

```
f([X1,X2],Y) :- Y is 4*(X1 - 5)**2 + (X2 - 6)**2.
```

Similarly we can define g_i and h_{ii} $(i = 1, \ldots, n)$.

Here we briefly explain the relations defined in the program.

The formula "$partition(B, B_1, B_2)$" holds if the sub-boxes B_1 and B_2 are the result of partitioning the box B along the dimension in which the interval is the largest.

The formula "$small(B, Y, U)$" holds if $|B| \leq \epsilon_X$ and $U - lb(Y) \leq \epsilon_F$.

The formula "$fb(B, U)$" holds if U equals $ub(f(m))$, where m is the middle point of the box B.

The formula "$min([X_1, ..., X_n], M)$" holds if M is the minimal one of $X_1, ..., X_n$.

The formula "$insert([B, Y, U], L, L_a)$" holds if $L_a = [[B_1, Y_1, U_1], \ldots, [B_n, Y_n, U_n]]$ is the result of adding $[B, Y, U]$ into the list L such that $lb(Y_1) \leq \cdots \leq lb(Y_n)$.

The formula "$g\_min(L, F_{ub}, A0, A)$" holds if

(1) $L = [[B_1, Y_1, U_1], \ldots, [B_s, Y_s, U_s]]$, where B_i $(i = 1, \ldots, s)$ is a sub-box of the initial domain X, Y_i is the interval value of the function f over B_i, and U_i is the value of f at the middle point of B_i.

(2) F_{ub} is the lowest upper bound of f^* found so far.

(3) $A0$ has the same form as A, which is described just below.

(4) $A = [[B_1', Y_1', U_1'], \ldots, [B_n', Y_n', U_n']]$ contains all elements which satisfy that $|B_i'| \leq \epsilon_X$ and $U_i' - lb(Y_i') \leq \epsilon_F$ $(i = 1, \ldots, n)$.

```
g_min([],Fub,A,A) :- !.           % L = [], terminate, Answers are in A
g_min([[B,Y,U]|L0],Fub,A0,A) :- % remove [B,Y,U] from L=[[B,Y,U]|L0]
   small(B,Y,U), !,               %  width(B)=<Ex and U-lb(Y)=<Ey
   insert([B,Y,U],A0,A1),         %  insert [B,Y,U] into A0 and get A1
   g_min(L0,Fub,A1,A).
g_min([[B,Y,U]|L0],Fub,A0,A) :-
   conditions([B,Y,U],Fub), !,   % conditions may be satisfied
   partition(B,B1,B2),            % partition box B into B1 and B2
   f(B1,Y1), fb(B1,U1),           % Y1=f(B1),U1=ub(f(m1)),m1=midpoint(B1)
   f(B2,Y2), fb(B2,U2),           % Y2=f(B2),U2=ub(f(m2)),m2=midpoint(B2)
   min([Fub,U1,U2],Fub1),         % Fub1=min(Fub,U1,U2)
   insert([B1,Y1,U1],L0,L1),      % insert [B1,Y1,U1] and [B2,Y2,U2]
   insert([B2,Y2,U2],L1,L2),      % into L0 and get L2
   g_min(L2,Fub1,A0,A).
g_min([[B,Y,U]|L0],Fub,A0,A) :- % conditions can not be satisfied
   g_min(L0,Fub,A0,A).            % remove [B,Y,U] from [[B,Y,U]|L0]

conditions([B,Y,U],Fub) :-        %    conditions:
   less(Y, Fub),                                 % f(x) =< Fub
   g1(B,G1),eq(G1,0),...,gn(B,Gn),eq(Gn,0),    % gi(x) = 0, i=1,...,n
   h11(B,H1),ge(H1,0),...,hnn(B,Hn),ge(Hn,0), % hii(x) >= 0, i=1,...,n
   t1(B,Y,Fub),  % f(x)=f(x0)+(x-x0)f'(E)=<Fub, E is between x0 and x
   t2(B,Y,Fub),  % f(x)=f(x0)+(x-x0)f'(x0)+0.5(x-x0)f"(E)(x-x0)=<Fub
   newton(B,Y).  % x = x0 + f'(x0)/f"(E)

insert([B,Y,U],[],[[B,Y,U]]) :- !.
insert([B,Y,U],[[B1,Y1,U1]|T],[[B,Y,U],[B1,Y1,U1]|T]) :-
   range(Y, [L,_]), range(Y1, [L1,_]), L =< L1, !.
insert([B,Y,U],[[B1,Y1,U1]|T],[[B1,Y1,U1]|R]) :- insert([B,Y,U],T,R).
```

Figure 4.3: Code of Branch-and-Bound algorithm for unconstrained optimization

The Implementations of IAU_0, ICU_0, IAU_1 and ICU_1. The implementation given in Figure 4.3 becomes the implementation of an order zero algorithm if we just keep "*less(Y, Fub)*" in the definition of the relation *conditions*. If we have both "*less(Y, Fub)*" and "*g1(B,G1),eq(G1,0),...,gn(B,Gn),eq(Gn,0)*", we obtain the implementation of an order one algorithm.

We obtain the Interval Constraint version of the Branch-and-Bound algorithm by providing the definitions for the relations *less* and/or *eq* in Interval Constraint as follows:

```
less(Y, Fub) :- Y =< Fub.
```

and/or

```
eq  (Y, 0)   :- Y == 0.
```

If we define the relations *less* and/or *eq* in Interval Arithmetic in the following, we have the Interval Arithmetic version of the algorithm.

```
less(Y, Fub) :- range(Y, [Ylb, Yub]), Yub =< Fub.
```

and/or

```
eq  (Y, 0)   :- range(Y, [Ylb, Yub]), Ylb =< 0, 0 =< Yub.
```

4.3.6 Computational Results

Two groups of computational results are given in this section. The first group is the results from running IAU_0 and ICU_0. The second group is from running IAU_1 and ICU_1.

All the results are from running the algorithms on the test problems selected from [36, 42, 75]. We choose the test problems of polynomials with a few of variables. These test problems are listed in Table 4.2.

Problem	Objective Function
1	$f(x) = 4(x_1 - 5)^2 + (x_2 - 6)^2$
2	$f(x) = 0.26(x_1^2 + x_2^2) - 0.48x_1x_2$
3	$f(x) = (1.5 - x_1 + x_1x_2)^2 + (2.25 - x_1 + x_1x_2^2)^2 + (2.625 - x_1 + x_1x_2^3)^2$
4	$f(x) = 100(x_2 - x_1^2)^2 + (x_1 - 1)^2$
5	$f(x) = 12x_1^2 - 6.3x_1^4 + x_1^6 + 6x_2(x_2 - x_1)$
6	$f(x) = x^6 - 15x^4 + 27x^7 + 250$
7	$f(x) = \sum_{i=1}^{3} [(x_1 - x_i^2)^2 + (x_i - 1)^2]$
8	$f(x) = \sum_{i=1}^{3} [(Ax_1 - Bx_i^2)^2 + (Cx_i - D)^2], A = B = C = D = [0.999, 1.001]$
9	$f(x) = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4$
10	$f(x) = \sum_{i=2}^{3} [(x_1 - x_i^2)^2 + (1 - x_i)^2]$

Table 4.2: Test problems for unconstrained global optimization

Before giving the computational results, we explain the symbols used for showing and comparing the results as follows:

- IA Interval Arithmetic version of an algorithm.
- IC Interval Constraint version of an algorithm.
- W_i Width of initial box.
- ϵ Intended tolerance for box size and function width.
- W_{x^*} Width of final box for x^*.

W_{f^*} Width of bound on f^*.

N_{max} Maximum Number of elements in L during the execution.

t Running time.

NA The result is not available because of memory overflow.

In Table 4.5, 4.6, 4.9, 4.10, 4.13 and 4.14, each of these symbols with superscript A represents that the corresponding results are from running the interval arithmetic version of an algorithm, that with superscript C represents that the corresponding results are from running the interval constraint version of the algorithm, and that with M represents that the corresponding results are from running a version of the algorithm with the improvements on Memory use.

Running Results of IAU_0 and ICU_0. Table 4.3 shows the computational results of IAU_0 and ICU_0. Table 4.5 gives some comparisons between the computational results of these two implementations.

From Table 4.5, we can see that ICU_0 achieves a factor of 2.7 to 27.5 in speedup on the unconstrained optimization test problems over IAU_0. Moreover, most of the output results produced by ICU_0 are better than those produced by IAU_0. the final boxes for f^* and x^* produced by ICU_0 are usually smaller. Finally, running IAU_0 needs more memory space, the maximum numbers of elements (tuples) in L during the execution of IAU_0 on the test problems are about 2.7 to 27 times as large as those of elements in L during the execution of ICU_0.

Running Results of IAU_1 and ICU_1. Table 4.4 shows the computational results of IAU_1 and ICU_1. Table 4.6 gives some comparisons between the computational results of IAU_1 and ICU_1.

From Table 4.6, we can see that ICU_1 achieves a factor of 5.8 to 123.5 in speedup on the unconstrained optimization test problems over IAU_1. The maximum numbers of elements in L during the execution of IAU_1 on the test problems are about 5.9 to 126 times as large as those of elements in L during the execution of ICU_1.

4.4 Time vs Memory Trade-off

We present improvements on memory use of the Branch-and-Bound algorithm, which are applicable both to the interval arithmetic version and to the interval constraint version of the algorithm. This is followed by test results.

In general, the improvements on memory use imply additional computation time. The interesting thing is that these improvements speed up the execution of the Branch-and-Bound algorithm for solving some unconstrained global optimization problems.

4.4.1 Reducing Memory Use

During the iteration of the Branch-and-Bound algorithm, the upper bound f^*_{ub} may be updated to a new value, which is less than the previous f^*_{ub}. Whenever this happens, some tuples in L may be discarded by checking the conditions with the new f^*_{ub}. This is because for a given tuple, the conditions that succeed with the old f^*_{ub} may fail with the new one. Thus we may reduce the size of L by checking each tuple in L with the new f^*_{ub}.

This can be achieved by changing the line that updates f^*_{ub} in the algorithm and defining a procedure called "reduce" that discards tuples from a given priority queue that do not satisfy the conditions.

The line in the algorithm

$f^*_{ub} := min(f^*_{ub}, u_1, u_2)$

is replaced by the following segment

$f^*_{ub}{}' := min(f^*_{ub}, u_1, u_2)$
if $f^*_{ub}{}' < f^*_{ub}$ **then**
  reduce(L, $f^*_{ub}{}'$);
end-if
$f^*_{ub} := f^*_{ub}{}'$;

This code segment checks whether the new upper bound $f^*_{ub}{}'$ is less than the old upper bound f^*_{ub}. If $f^*_{ub}{}'$ is less than f^*_{ub}, it will call the procedure "reduce" to discard all the tuples in L that do not satisfy the conditions with $f^*_{ub}{}'$. Finally, it updates f^*_{ub} by $f^*_{ub}{}'$.

It is possible that the change of f^*_{ub} is quite small. In this case, calling "reduce" is not very useful. To avoid this kind of situation, we can replace

$$f^*_{ub}{}' < f^*_{ub}$$

in the **if-then** line of the segment by

$$f^*_{ub}{}' * c < f^*_{ub}$$

where $c > 1$ is a constant. The choice of c should suggest that $f^*_{ub}{}' * c < f^*_{ub}$ imply that $f^*_{ub}{}'$ is significantly less than f^*_{ub} . We choose $c = 1.5$ in our implementation.

The procedure "reduce" is described in Figure 4.4. For a given priority queue L of tuples and an upper bound $f_{ub}^{*}{}'$ of f^{*}, it checks every tuple in L with $f_{ub}^{*}{}'$. It removes all the tuples in L that do not satisfy the conditions.

In the procedure "reduce", "conditions($< B, l, u >$, $f_{ub}^{*}{}'$)" may shrink the box B. If B is shrunk to, say, B', $ub(f(m))$ may be less than u, where m is the midpoint of B'. Thus it is better to update u to $\min(u, ub(f(m)))$ and $f_{ub}^{*}{}'$ with $\min(f_{ub}^{*}{}', u)$ whenever B shrinks.

This can be achieved through replacing the **if-then** statement in the procedure "reduce" by one assignment statement and one complicated **if-then** statement. The improved procedure "reduce" is shown in Figure 4.5. It considers that the box B may shrink. In the improved "reduce", m is the midpoint of the shrunk box B. "shrunk(B, B_{old})" checks whether the box B is shrunk after the conditions are executed.

The Branch-and-Bound algorithm with the improvements is given in Figure 4.6. From this improved version, both an interval arithmetic and an interval constraint version can be obtained in the same way as from the algorithm in Figure 4.2.

4.4.2 Implementations of the Algorithms with Improvements on Memory Use

The first improvement on the Branch-and-Bound algorithm can be implemented through changing the last line of the third clause that defines *g_min* and defining in BNR-Prolog the relation *reduce* for the procedure "reduce" given in Figure 4.4. The last line of the third clause

```
g_min(L2,Fub1,A0,A).
```

reduce(L, ${f^*_{ub}}'$)
begin
    $L' := \emptyset$;
    while $L \neq \emptyset$ **do**
        remove $< B, l, u >$ from L;
        if $\underline{\text{conditions}(< B, l, u >, {f^*_{ub}}')}$ **then**
            add $< B, l, u >$ to L';
        end-if
    end-while
    $L := L'$;
end

Figure 4.4: Procedure "reduce"

reduce(L, ${f^*_{ub}}'$)
begin
    $L' := \emptyset$;
    while $L \neq \emptyset$ **do**
        remove $< B, l, u >$ from L;
        $B_{old} := B$;
        if $\underline{\text{conditions}(< B, l, u >, {f^*_{ub}}')}$ **then**
            if shrunk(B, B_{old}) **then**
                $l := lb(f(B))$;
                $u := min(u, ub(f(m)))$;
                ${f^*_{ub}}' := min({f^*_{ub}}', u)$;
            end-if ;
            add $< B, l, u >$ to L';
        end-if
    end-while
    $L := L'$;
end

Figure 4.5: Procedure "reduce" considering box shrunk

```
input:  f, X, ε_X, ε_F;
f*_ub := ub(f(mid(X)));
L := [< X, lb(f(X)), f*_ub >];
A := ∅;
while  L ≠ ∅  do
        remove < B, l, u > from L;
        if  |B| ≤ ε_X ∧ (u − l) ≤ ε_F  then
            add < B, l, u > to A
        else  if  conditions(< B, l, u >, f*_ub)  then
            partition B into B_1 and B_2 with midpoints m_1 and m_2:
            l_1 := lb(f(B_1)); l_2 := lb(f(B_2));
            u_1 := ub(f(m_1)); u_2 := ub(f(m_2));
            f*_ub' := min(f*_ub, u_1, u_2)
            if  f*_ub' * c < f*_ub  then
                reduce(L, f*_ub');
            end-if
            f*_ub := f*_ub';
            add < B_1, l_1, u_1 > and < B_2, l_2, u_2 > to L;
        end-if
end-while
reduce(A, f*_ub);
f*_lb = min{l_i |< B_i, l_i, u_i >∈ A};
output:  f*_lb, f*_ub, A;
```

Figure 4.6: Branch-and-Bound algorithm with improvements on memory use

can be changed by the following two lines

```
may_reduce(L2, Fub1, Fub, Lr),
g_min(Lr, Fub1, A0, A).
```

where the definition of *may_reduce* is described in Figure 4.7. The definition of the relation *reduce* is also given in BNR-Prolog in Figure 4.7.

The other improvement on the Branch-and-Bound algorithm can be implemented by adding a new parameter for a possibly lower f_{ub}^* in the relations *may_reduce* and making some changes for the definitions of *reduce* accordingly. The definitions of *may_reduce* and *reduce* for both of the improvements is given in Figure 4.8.

Using the same way as implementing IAU_0, ICU_0, IAU_1 and ICU_1, we obtain the implementations of IAU_0', ICU_0', IAU_1' and ICU_1' from the code given in Figure 4.3 for the implementation of the algorithm with the improvements. Here the $'$ represents an algorithm with the improvements.

4.4.3 IAU_0 and IAU_1 vs IAU_0' and IAU_1'

The results of running IAU_0 and IAU_1 are contained in Table 4.3 and 4.4 respectively. Here we just give the results of running IAU_0' and IAU_1' in Table 4.7 and 4.8 respectively.

Table 4.9 (4.10) gives some comparisons between the results of running IAU_0 and IAU_0' (IAU_1 and IAU_1').

From Table 4.9, we can see that IAU_0' is faster than IAU_0 on most of the test problems. Moreover, IAU_0' uses less memory space than IAU_0, the maximum numbers

```
may_reduce(L, Fub1, Fub, Lr) :-
    Fub1*1.5 < Fub, !,    % if Fub1 is significantly less than Fub
    reduce(L, Fub1, Lr). % then  reduce  L to Lr
may_reduce(L, Fub1, Fub, L). % otherwise, do nothing

reduce([], Fub1, []) :- !.
reduce([[B,Y,U]|T], Fub1, [[B,Y,U]|R]) :- % keep [B,Y,U] if
    conditions([B,Y,U],Fub1),!, % conditions may be held for Fub1
    reduce(T, Fub1, R).         % reduce T, the Tail of the List
reduce([_|T], Fub1, R) :- % conditions can not be held for Fub1
    reduce(T, Fub1, R).   % discard [B,Y,U] and reduce T
```

Figure 4.7: Code for the procedure "reduce"

```
may_reduce(L,Fub1,Fub,Lr,Fub2) :-
    Fub1*1.5 < Fub, !,       % if Fub1 is significantly less than Fub
    reduce(L,Fub1,Lr,Fub2). % then reduce L to Lr and get a new Fub2
may_reduce(L,Fub1,Fub,L, Fub2). % otherwise, do nothing

reduce([], Fub1, [], Fub1) :- !.
reduce([[B,Y,U]|T], Fub1, [[B,Y,U]|R], Fub2) :- % keep [B,Y,U] if
   copy(B, Bold),                    % Bold := B
   conditions([B,Y,U],Fub1), !,      % conditions may be held for Fub1
   may_shrink(B,Bold,Fub1,Fub11), % get new bound Fub11 if B is shrunk
   reduce(T, Fub11, R, Fub2).        % reduce T with the new bound Fub11
reduce([_|T], Fub1, R, Fub2) :- % conditions can not be held for Fub1
    reduce(T, Fub1, R, Fub2).   % discard [B,Y,U] and reduce T
```

Figure 4.8: Code for the procedure "reduce" considering box shrunk

of elements in L during the execution of IAU_0 are upto 1.9 times as large as those of elements in L during the execution of IAU_0'.

The results in Table 4.10 are similar to those in Table 4.9.

4.4.4 ICU_0 and ICU_1 vs ICU_0' and ICU_1'

The results of running ICU_0 and ICU_1 are contained in Table 4.3 and 4.4 respectively. Here we give the results of running ICU_0' and ICU_1' in Table 4.11 and 4.12 respectively.

Table 4.13 (4.14) gives some comparisons between the results of running ICU_0 and ICU_0' (ICU_1 and ICU_1').

From Table 4.13, we can see that ICU_0' is faster than ICU_0 on almost half of the test problems. It uses less memory space than ICU_0, the maximum numbers of elements in L during the execution of ICU_0 are upto 4.5 times as large as those of elements in L during the execution of ICU_0'.

The results in Table 4.14 are similar to those in Table 4.13.

Problem	Method	W_i	ϵ	W_{x^*}	W_{f^*}	N_{max}	t
1	IA	$2 \cdot 10^6$	10^{-1}	$6.1 \cdot 10^{-2}$	$2.9 \cdot 10^{-2}$	378	8.232
	IC	$2 \cdot 10^6$	10^{-1}	10^{-1}	$2.0 \cdot 10^{-2}$	14	0.299
2	IA	6	10^{-1}	$9.5 \cdot 10^{-2}$	10^{-1}	2267	305.2
	IC	6	10^{-1}	$9.7 \cdot 10^{-2}$	$9.6 \cdot 10^{-2}$	358	23.645
3	IA	9	10^{-5}	$8.6 \cdot 10^{-6}$	$2.5 \cdot 10^{-9}$	769	44.619
	IC	9	10^{-5}	$7.4 \cdot 10^{-6}$	$9.7 \cdot 10^{-11}$	95	4.767
4	IA	$2 \cdot 10^3$	10^{-5}	$7.5 \cdot 10^{-6}$	$2.5 \cdot 10^{-8}$	568	12.908
	IC	$2 \cdot 10^3$	10^{-5}	$8.8 \cdot 10^{-6}$	$2.7 \cdot 10^{-9}$	67	1.512
5	IA	6	10^{-1}	NA	NA	NA	NA
	IC	6	10^{-1}	$9.5 \cdot 10^{-2}$	$9.0 \cdot 10^{-2}$	114	8.262
6	IA	9	10^{-1}	NA	NA	NA	NA
	IC	9	10^{-1}	$6.2 \cdot 10^{-5}$	10^{-1}	434	23.635
7	IA	$2 \cdot 10^6$	10^{-5}	$7.3 \cdot 10^{-6}$	$3.1 \cdot 10^{-10}$	232	12.623
	IC	$2 \cdot 10^6$	10^{-5}	$7.4 \cdot 10^{-6}$	$5.4 \cdot 10^{-11}$	86	4.731
8	IA	$2 \cdot 10^6$	10^{-1}	$6.0 \cdot 10^{-2}$	$2.7 \cdot 10^{-2}$	154	13.593
	IC	$2 \cdot 10^6$	10^{-1}	$7.1 \cdot 10^{-2}$	$5.0 \cdot 10^{-3}$	23	2.178
9	IA	6	10^{-1}	$9.5 \cdot 10^{-2}$	$9.4 \cdot 10^{-2}$	900	56.392
	IC	6	10^{-1}	$9.1 \cdot 10^{-2}$	$1.2 \cdot 10^{-2}$	41	2.571
10	IA	3.78	10^{-5}	$7.2 \cdot 10^{-6}$	$4.7 \cdot 10^{-10}$	354	13.853
	IC	3.78	10^{-5}	$6.3 \cdot 10^{-6}$	$1.1 \cdot 10^{-10}$	28	1.444

Table 4.3: Running results of IAU_0 and ICU_0

Problem	Method	W_i	ϵ	W_{x^*}	W_{f^*}	N_{max}	t
1	IA	$2 \cdot 10^6$	10^{-1}	$6.1 \cdot 10^{-2}$	$2.0 \cdot 10^{-2}$	378	10.742
	IC	$2 \cdot 10^6$	10^{-1}	$1.6 \cdot 10^{-15}$	10^{-31}	3	0.087
2	IA	6	10^{-1}	$9.5 \cdot 10^{-2}$	10^{-1}	163	9.752
	IC	6	10^{-1}	$3.9 \cdot 10^{-3}$	$2.3 \cdot 10^{-6}$	3	0.093
3	IA	9	10^{-5}	$8.6 \cdot 10^{-6}$	$2.5 \cdot 10^{-9}$	769	78.645
	IC	9	10^{-5}	$7.4 \cdot 10^{-6}$	$9.7 \cdot 10^{-11}$	95	9.211
4	IA	$2 \cdot 10^3$	10^{-5}	$7.5 \cdot 10^{-6}$	$2.5 \cdot 10^{-8}$	568	16.964
	IC	$2 \cdot 10^3$	10^{-5}	$8.3 \cdot 10^{-6}$	$1.8 \cdot 10^{-9}$	34	1.484
5	IA	6	10^{-1}	$9.5 \cdot 10^{-2}$	$5.3 \cdot 10^{-2}$	31	2.421
	IC	6	10^{-1}	$7.7 \cdot 10^{-29}$	$4.6 \cdot 10^{-56}$	3	0.142
6	IA	9	10^{-1}	$3.4 \cdot 10^{-5}$	$5.6 \cdot 10^{-2}$	135	7.822
	IC	9	10^{-1}	$1.9 \cdot 10^{-13}$	$3.1 \cdot 10^{-10}$	7	0.236
7	IA	$2 \cdot 10^6$	10^{-5}	$7.3 \cdot 10^{-6}$	$3.1 \cdot 10^{-10}$	232	17.434
	IC	$2 \cdot 10^6$	10^{-5}	$7.1 \cdot 10^{-6}$	$8.3 \cdot 10^{-11}$	39	3.021
8	IA	$2 \cdot 10^6$	10^{-1}	$6.0 \cdot 10^{-2}$	$2.7 \cdot 10^{-2}$	154	22.455
	IC	$2 \cdot 10^6$	10^{-1}	$6.9 \cdot 10^{-2}$	$1.4 \cdot 10^{-2}$	13	1.959
9	IA	6	10^{-1}	$9.5 \cdot 10^{-2}$	$9.4 \cdot 10^{-2}$	770	62.673
	IC	6	10^{-1}	$7.5 \cdot 10^{-2}$	$2.6 \cdot 10^{-3}$	33	3.271
10	IA	3.78	10^{-5}	$7.2 \cdot 10^{-6}$	$4.7 \cdot 10^{-10}$	354	19.044
	IC	3.78	10^{-5}	$6.2 \cdot 10^{-6}$	$1.4 \cdot 10^{-11}$	39	2.895

Table 4.4: Running results of IAU_1 and ICU_1

Problem	$W_{x^*}^A/W_{x^*}^C$	$W_{f^*}^A/W_{f^*}^C$	N_{max}^A/N_{max}^C	t^A/t^C
1	610	1.45	27	27.5
2	0.98	1.04	6.3	12.9
3	1.16	25.8	8.1	9.4
4	0.85	9.3	8.5	8.5
5	NA	NA	NA	NA
6	NA	NA	NA	NA
7	0.99	5.7	2.7	2.7
8	0.85	5.4	6.7	6.2
9	1.04	7.8	22.0	21.9
10	1.43	4.3	12.6	9.6

Table 4.5: Comparisons between IAU_0 and ICU_0

Problem	$W_{x^*}^A/W_{x^*}^C$	$W_{f^*}^A/W_{f^*}^C$	N_{max}^A/N_{max}^C	t^A/t^C
1	$3.8 \cdot 10^{13}$	$2.0 \cdot 10^{29}$	126	123.5
2	24.4	$4.3 \cdot 10^{4}$	54.3	104.9
3	1.16	25.8	8.1	8.5
4	0.90	13.9	16.7	11.4
5	$1.2 \cdot 10^{27}$	$1.2 \cdot 10^{54}$	10.3	17.1
6	$1.8 \cdot 10^{8}$	$1.8 \cdot 10^{8}$	19.3	33.1
7	1.03	3.7	5.9	5.8
8	0.87	1.9	11.9	11.5
9	1.27	36.2	23.3	19.2
10	1.16	33.6	9.1	6.6

Table 4.6: Comparisons between IAU_1 and ICU_1

Problem	Method	W_i	ϵ	W_{x^*}	W_{f^*}	N_{max}	t
1	IA	$2 \cdot 10^6$	10^{-1}	$6.1 \cdot 10^{-2}$	$2.9 \cdot 10^{-2}$	269	7.902
2	IA	6	10^{-1}	$9.5 \cdot 10^{-2}$	10^{-1}	1950	262.75
3	IA	9	10^{-5}	$8.6 \cdot 10^{-6}$	$2.5 \cdot 10^{-9}$	599	40.418
4	IA	$2 \cdot 10^3$	10^{-5}	$7.5 \cdot 10^{-6}$	$2.5 \cdot 10^{-8}$	378	14.273
5	IA	6	10^{-1}	NA	NA	NA	NA
6	IA	9	10^{-1}	NA	NA	NA	NA
7	IA	$2 \cdot 10^6$	10^{-5}	$7.3 \cdot 10^{-6}$	$3.1 \cdot 10^{-10}$	121	12.393
8	IA	$2 \cdot 10^6$	10^{-1}	$6.0 \cdot 10^{-2}$	$2.7 \cdot 10^{-2}$	121	13.543
9	IA	6	10^{-1}	$9.5 \cdot 10^{-2}$	$9.4 \cdot 10^{-2}$	698	51.071
10	IA	3.78	10^{-5}	$7.2 \cdot 10^{-6}$	$1.5 \cdot 10^{-10}$	251	12.603

Table 4.7: Running results of IAU_0'

Problem	Method	W_i	ϵ	W_{x^*}	W_{f^*}	N_{max}	t
1	IA	$2 \cdot 10^6$	10^{-1}	$6.1 \cdot 10^{-2}$	$2.0 \cdot 10^{-2}$	269	10.472
2	IA	6	10^{-1}	$9.5 \cdot 10^{-2}$	10^{-1}	163	9.792
3	IA	9	10^{-5}	$8.6 \cdot 10^{-6}$	$2.5 \cdot 10^{-9}$	599	74.515
4	IA	$2 \cdot 10^3$	10^{-5}	$7.5 \cdot 10^{-6}$	$2.5 \cdot 10^{-8}$	568	16.964
5	IA	6	10^{-1}	$9.5 \cdot 10^{-2}$	$5.3 \cdot 10^{-2}$	31	2.50
6	IA	9	10^{-1}	$3.4 \cdot 10^{-5}$	$5.6 \cdot 10^{-2}$	134	7.812
7	IA	$2 \cdot 10^6$	10^{-5}	$7.3 \cdot 10^{-6}$	$3.1 \cdot 10^{-10}$	121	17.444
8	IA	$2 \cdot 10^6$	10^{-1}	$6.0 \cdot 10^{-2}$	$2.7 \cdot 10^{-2}$	121	22.485
9	IA	6	10^{-1}	$9.5 \cdot 10^{-2}$	$9.4 \cdot 10^{-2}$	568	58.032
10	IA	3.78	10^{-5}	$7.2 \cdot 10^{-6}$	$4.7 \cdot 10^{-10}$	251	17.914

Table 4.8: Running results of IAU_1'

Problem	$W_{x^*}^A/W_{x^*}^{AM}$	$W_{f^*}^A/W_{f^*}^{AM}$	N_{max}^A/N_{max}^{AM}	t^A/t^{AM}
1	1	1	1.41	1.04
2	1	1	1.16	1.16
3	1	1	1.28	1.1
4	1	1	1.5	0.9
5	NA	NA	NA	NA
6	NA	NA	NA	NA
7	1	1	1.92	1.02
8	1	1	1.27	1.0
9	1	1	1.29	1.1
10	1	3.13	1.41	1.1

Table 4.9: Comparisons between IAU_0 and IAU_0'

Problem	$W_{x^*}^A/W_{x^*}^{AM}$	$W_{f^*}^A/W_{f^*}^{AM}$	N_{max}^A/N_{max}^{AM}	t^A/t^{AM}
1	1	1	1.41	1.03
2	1	1	1	1.0
3	1	1	1.28	1.06
4	1	1	1	1
5	1	1	1	0.97
6	1	1	1.01	1.0
7	1	1	1.92	1.0
8	1	1	1.27	1.0
9	1	1	1.36	1.08
10	1	1	1.41	1.06

Table 4.10: Comparisons between IAU_1 and IAU_1'

Problem	Method	W_i	ϵ	W_{x^*}	W_{f^*}	N_{max}	t
1	IC	$2 \cdot 10^6$	10^{-1}	$1.8 \cdot 10^{-2}$	$1.9 \cdot 10^{-2}$	7	0.230
2	IC	6	10^{-1}	$9.7 \cdot 10^{-2}$	$9.6 \cdot 10^{-2}$	358	24.10
3	IC	9	10^{-5}	$7.7 \cdot 10^{-6}$	$3.5 \cdot 10^{-8}$	21	2.412
4	IC	$2 \cdot 10^3$	10^{-5}	$6.8 \cdot 10^{-6}$	$1.4 \cdot 10^{-9}$	36	1.298
5	IC	6	10^{-1}	$9.5 \cdot 10^{-2}$	$7.7 \cdot 10^{-2}$	114	8.032
6	IC	9	10^{-1}	$6.1 \cdot 10^{-5}$	10^{-1}	434	24.255
7	IC	$2 \cdot 10^6$	10^{-5}	$2.3 \cdot 10^{-6}$	$1.3 \cdot 10^{-12}$	33	4.761
8	IC	$2 \cdot 10^6$	10^{-1}	$9.6 \cdot 10^{-2}$	$9.8 \cdot 10^{-2}$	20	3.026
9	IC	6	10^{-1}	$7.1 \cdot 10^{-2}$	$1.6 \cdot 10^{-2}$	24	2.625
10	IC	3.78	10^{-5}	$2.7 \cdot 10^{-9}$	$3.7 \cdot 10^{-18}$	18	1.540

Table 4.11: Running results of ICU_0'

Problem	Method	W_i	ϵ	W_{x^*}	W_{f^*}	N_{max}	t
1	IC	$2 \cdot 10^6$	10^{-1}	$1.6 \cdot 10^{-15}$	10^{-31}	3	0.103
2	IC	6	10^{-1}	$3.9 \cdot 10^{-3}$	$2.3 \cdot 10^{-6}$	3	0.097
3	IC	9	10^{-5}	$7.7 \cdot 10^{-6}$	$3.5 \cdot 10^{-8}$	21	3.88
4	IC	$2 \cdot 10^3$	10^{-5}	$9.8 \cdot 10^{-6}$	$2.4 \cdot 10^{-9}$	19	1.095
5	IC	6	10^{-1}	$7.7 \cdot 10^{-29}$	$4.1 \cdot 10^{-56}$	3	0.189
6	IC	9	10^{-1}	$5.1 \cdot 10^{-14}$	$8.7 \cdot 10^{-11}$	7	0.286
7	IC	$2 \cdot 10^6$	10^{-5}	$9.7 \cdot 10^{-6}$	$3.8 \cdot 10^{-11}$	20	3.491
8	IC	$2 \cdot 10^6$	10^{-1}	$5.6 \cdot 10^{-2}$	$7.5 \cdot 10^{-3}$	8	2.001
9	IC	6	10^{-1}	$8.8 \cdot 10^{-2}$	$1.1 \cdot 10^{-1}$	20	3.051
10	IC	3.78	10^{-5}	$6.7 \cdot 10^{-6}$	$2.3 \cdot 10^{-11}$	17	2.121

Table 4.12: Running results of ICU_1'

Problem	$W^C_{x^*}/W^{CM}_{x^*}$	$W^C_{f^*}/W^{CM}_{f^*}$	N^C_{max}/N^{CM}_{max}	t^C/t^{CM}
1	5.6	1.05	2	1.3
2	1	1	1	0.98
3	0.96	2.77	4.5	1.98
4	1.29	1.93	1.86	1.16
5	1	1.17	1	1.03
6	1.02	1	1	0.97
7	3.22	41.5	2.6	0.99
8	0.74	5.1	1.15	0.72
9	1.28	0.75	1.71	0.98
10	$2.3 \cdot 10^3$	$3.0 \cdot 10^7$	1.56	0.94

Table 4.13: Comparisons between ICU_0 and ICU'_0

Problem	$W^C_{x^*}/W^{CM}_{x^*}$	$W^C_{f^*}/W^{CM}_{f^*}$	N^C_{max}/N^{CM}_{max}	t^C/t^{CM}
1	1	1	1	0.84
2	1	1	1	0.96
3	0.96	$2.8 \cdot 10^{-3}$	4.5	2.37
4	0.85	0.75	1.8	1.36
5	1	1.12	1	0.75
6	3.73	3.56	1	0.83
7	0.73	2.18	1.95	0.87
8	1.23	1.87	1.63	0.98
9	0.85	$2.4 \cdot 10^{-2}$	1.65	1.07
10	0.93	0.61	2.29	1.36

Table 4.14: Comparisons between ICU_1 and ICU'_1

Chapter 5

Constrained Global Optimization

The constrained global optimization problem is much harder to solve than the unconstrained variety. This holds in the interval arithmetic approach as well as in other approaches. Using interval constraints, we found that the constrained optimization problem is not as much harder to solve than the unconstrained version as it is in the other approaches. W. J. Older explored in BNR-Prolog the application of interval constraints in several areas [73, 69, 65, 72, 70]. In constrained optimization, Older used the Kuhn-Tucker condition [49, 28] and demonstrated it on one problem from [29]. Recently, P. Van Hentenryck and L. Michel reported their Numerica system for global optimization in [37, 38]. They use interval constraints, therefore obtaining the same advantages over interval arithmetic as reported in [16, 84, 17]. They presented a Branch-and-Bound algorithm and discussed the performance of their system. Here, we investigate the transition from solving the unconstrained global optimization problem to the constrained one, present the characteristics of the interval constraint method, study the effect of the Fritz-John condition [36, 46] as redundant constraints, and compare the interval arithmetic approach with the interval constraint method.

To avoid ambiguity between the "constraints" in the sense of constrained global optimization and interval constraints, we call the former "conditions", the latter "constraints".

5.1 Unconstrained vs Constrained Global Optimization

From the definition 1.1 of a global optimization problem, we obtain a *constrained global optimization problem* if the conditions 1.1 or 1.2 are present; otherwise, we get an unconstrained one. For the unconstrained optimization problem, x is still constrained to lie in the initial box. But usually the initial box is chosen to be so large as to contain all the minimizers and such that no occurrence of the global minimum is on the boundary of the box.

When we solve the constrained global optimization problem using interval constraints, the problem's defining conditions 1.1 and 1.2 are translated to constraints. From the point of view of interval constraints, constrained global optimization is the same type of problem as unconstrained one. But this does not mean that one can use the algorithm in Figure 4.2 for solving constrained global optimization problems. In this chapter we investigate the necessary modifications.

5.1.1 Existence of the Global Minimum

For an unconstrained global optimization problem, if the objective function is bounded in a given domain, then the global minimizer(s) exists in the domain. This is implied by Bolzano-Weierstrass theorem [41, 24]. In constrained global optimization, there is the additional complication of a feasible region: a minimizer must be a feasible point. If there exists at least one feasible point and the objective function is bounded, then it follows that a global minimizer exists; otherwise, the problem has no solution [29].

A constrained global optimization problem is a mathematical model of a physical problem. Given a physical problem, we usually know that it has solutions and optimal solution(s). For example, there exist designs and best design(s) for a chemical reactor design problem.

However, it is difficult to know whether there are feasible points (i.e., solutions) for a constrained global optimization problem of mathematical form. For a given point x, even though we can check whether it satisfies the inequality condition 1.1 using interval arithmetic or interval constraints, it is generally impossible to assure that x satisfies the equality condition 1.2. If we make a single rounding error in computing $q_j(x)$ for some $j = 1, \ldots, r$, we do not know whether $q_j(x)$ is zero.

5.1.2 Necessary Conditions

The unconstrained global optimization problem has the properties that at any occurrence of the global minimum:

- the gradient of the objective function (see definition 4.1) has to be zero, and

- the Hessian matrix of the objective function (see definition 4.2) has to be positive semi-definite.

These two properties are referred to as *first order necessary condition* and *second order necessary condition* respectively [28]. In the previous chapter we explored the idea of using these conditions as constraints. As shown there, the idea resulted in a dramatic improvement in solving unconstrained global optimization problems in interval constraints.

However, we can not use these two conditions for solving constrained global optimization problems. In a constrained global optimization problem, it is quite common that the only occurrences of the global minimum are on the boundary of the feasible region. At such occurrences the gradient of the objective function is typically not zero, and the Hessian matrix is not positive semi-definite. Fortunately, some first order necessary conditions for constrained global optimization were introduced [49, 28, 36, 46].

The most general necessary condition for the constrained optimization problem described in the definition 1.1 is the Fritz-John condition [36, 46]. The Kuhn-Tucker condition is also well known. It requires a type of condition called *regularity condition*. The regularity condition is rather troublesome to verify in practice.

Fritz-John Condition. Assume that the objective function f, constraint functions p_i $(i = 1, \dots, m)$ and q_j $(j = 1, \dots, r)$ in the problem given by the definition 1.1 are differentiable. If a point x is a local minimizer of the problem, then there exist $u_i \geq 0, u_i \in R$ $(i = 0, 1, \dots, m)$ and $v_j \in R$ $(j = 1, \dots, r)$ such that

$$F(t) = (F_1(t), F_2(t), \dots, F_{m+r+2}(t))^T = 0 \tag{5.1}$$

where

$$F_1(t) = u_0 \nabla f(x) + \sum_{i=1}^{m} u_i \nabla p_i(x) + \sum_{j=1}^{r} v_j \nabla q_j(x), \quad (5.2)$$

$$F_2(t) = \sum_{i=0}^{m} u_i + \sum_{j=1}^{r} v_j - 1, \quad (5.3)$$

$$F_{i+2}(t) = u_i p_i(x) \ (i = 1, \ldots, m), \quad (5.4)$$

$$F_{m+j+2}(t) = v_j q_j(x) \ (j = 1, \ldots, r), \quad (5.5)$$

$t = (x_1, \ldots, x_n, u_0, u_1, \ldots, u_m, v_1, \ldots, v_r)^T$ and ∇ is the gradient operator.

The equation $F_2(t) = 0$, i.e.,

$$\sum_{i=0}^{m} u_i + \sum_{j=1}^{r} v_j - 1 = 0 \quad (5.6)$$

is called the *normalization condition*. The components u_i $(i = 0, \ldots, m)$, v_j $(j = 1, \ldots, r)$ are called the *Lagrangian multipliers*. The function whose gradient with respect to x occurs in the equation 5.2

$$\Psi(x, u, v) = u_0 f(x) + \sum_{i=1}^{m} u_i p_i(x) + \sum_{j=1}^{r} v_j q_j(x) \quad (5.7)$$

is called the *generalized Lagrangian function* of the problem given by the definition 1.1.

For an unconstrained optimization problem, the Fritz-John condition reduces to

$$\nabla f(x) = 0,$$

which is the first-order necessary condition for the unconstrained optimization problem.

Kuhn-Tucker Condition. Before describing the Kuhn-Tucker condition, we present the definition of the regularity condition.

Definition 5.1: [74]

Let x be a feasible point of the problem given by the definition 1.1. Let $A(x) = \{i \mid p_i(x) = 0, i \in \{1, \ldots, m\}\}$ be the so-called *active index set*. The point x is said to satisfy the *regularity condition* if the set of gradients $p_i'(x)$, $q_j'(x)$ with $i \in A(x)$, $j = 1, \ldots, r$ is linearly independent.

Assume that the objective function f, constraint functions p_i $(i = 1, \ldots, m)$ and q_j $(j = 1, \ldots, r)$ in the problem given by the definition 1.1 are differentiable. The Kuhn-Tucker condition says that if a point x is a local minimizer of the problem and satisfies the regularity condition, then there exist $u_i \geq 0, u_i \in R$ $(i = 1, \ldots, m)$ and $v_j \in R$ $(j = 1, \ldots, r)$ such that

$$K(t) = (K_1(t), K_2(t), \ldots, K_{m+r+1}(t))^T = 0 \tag{5.8}$$

where

$$K_1(t) = \nabla f(x) + \sum_{i=1}^{m} u_i \nabla p_i(x) + \sum_{j=1}^{r} v_j \nabla q_j(x), \tag{5.9}$$

$$K_{i+1}(t) = u_i p_i(x) \; (i = 1, \ldots, m), \tag{5.10}$$

$$K_{m+j+1}(t) = v_j q_j(x) \; (j = 1, \ldots, r), \tag{5.11}$$

$t = (x_1, \ldots, x_n, u_1, \ldots, u_m, v_1, \ldots, v_r)^T$ and ∇ is the gradient operator.

The function whose gradient with respect to x occurs in the equation 5.9

$$L(x, u, v) = f(x) + \sum_{i=1}^{m} u_i p_i(x) + \sum_{j=1}^{r} v_j q_j(x) \tag{5.12}$$

is called the *Lagrangian function* of the problem given by the definition 1.1.

To solve the equality system corresponding to the Fritz-John condition or the Kuhn-Tucker condition is very hard, since the equalities in the system are nonlinear and quite complex.

E. R. Hansen explored the Fritz-John condition in his interval arithmetic method for constrained global optimization through using the interval Newton method to solve the equality system [36]. But computational results are lacking [36, 46, 47].

In an interval constraint method for the constrained optimization, we do not solve the Fritz-John or Kuhn-Tucker conditions, we just add them in the form of their definitions into an interval constraint system. This will be discussed in detail in the following.

5.2 Overview of Interval Arithmetic Methods

Hansen and Sengupta [36] were the first to use interval arithmetic to solve the constrained global optimization problems with inequality conditions. Their algorithm is the combination of the Branch-and-Bound with several tests such as the midpoint test. The value of the objective function f at a feasible point in the original domain is an upper bound of f^*. In the midpoint test, Hansen used the lowest upper bound of f^* found so far.

For the constrained global optimization problem with equality conditions, there are difficulties in determining whether a given point is feasible, thus it is hard to find an upper bound of f^*. Because of rounding errors, it is generally impossible to verify that a point c satisfies an equality condition. To overcome the difficulties, Hansen and Walster [74, 32] suggested that Moore's existence theorem [60] should be applied to the constrained global optimization with the equality condition to check whether a given box contains a feasible point. Bao and Rokne [74, 4] generalized Moore's existence theorem to include the equality condition.

The algorithm for constrained global optimization is a modification of the one in Figure 4.2 for unconstrained global optimization. One important point is that $ub(f(mid(X)))$ in line 2, $ub(f(m_1))$ and $ub(f(m_2))$ in line 12 can no longer be counted on to provide an upper bound for f^*, as $mid(X)$, $m_1 = mid(B_1)$ or $m_2 = mid(B_2)$ may not be a feasible point in constrained global optimization. Instead we define $f_b(B)$ and use it in the place of $ub(f(mid(B)))$ to find an upper bound for f^*.

Another important observation is that the necessary condition for unconstrained global optimization is different from that for constrained one. This suggests that some of tests used for unconstrained global optimization can not be used in the constrained case.

In the following, we define $f_b(B)$ to find an upper bound of f^*, which is followed by the explanation of several tests that can be used for constrained global optimization.

Finding an Upper Bound. For the constrained global optimization problem only with inequality conditions, the value of the objective function f at a point c in the original domain is an upper bound u of the global minimum f^* if the point c satisfies

$$ub(p_i(c)) \leq 0 \quad \text{for all } i = 1, \ldots, m. \tag{5.13}$$

Formally, for a given box B, we can obtain u by using the following formula:

$$u = f_b(B) = \begin{cases} ub(f(c)) & \text{if } c \in B \text{ and } ub(p_i(c)) \leq 0 \quad \text{for } i = 1, \ldots, m \\ +\infty & \text{otherwise.} \end{cases} \tag{5.14}$$

Hansen calls the point that holds for the condition 5.13 a *certainly feasible* point. Hansen set an upper bound f^*_{ub} of f^* equal to the smallest upper bound found in this way and used it in the midpoint test.

However, it generally impossible to assure that the point c satisfies

$$q_j(c) = 0 \qquad \text{for all } j = 1, \ldots, r \tag{5.15}$$

i.e.,

$$lb(q_j(c)) = ub(q_j(c)) = 0 \qquad \text{for all } j = 1, \ldots, r \tag{5.16}$$

because of rounding errors. If we make a single rounding error in computing $q_j(c)$ for some $j = 1, \ldots, r$, we do not know whether $q_j(c)$ is zero or not. Thus the value of f at the point c can not be used as an upper bound of f^* for the constrained global optimization problem with the equality condition.

One way to obtain an upper bound of f^* is due to Hansen-Walster[32] and Bao-Rokne[4]. Suppose that we have proved that a feasible point exists in a box B. We compute $f(B)$, getting the result $[lb(f(B)), ub(f(B))]$. Then $ub(f(B))$ is an upper bound u of f^*. Formally, for a given box B, u is obtained according to the following formula:

$$u = f_b(B) = \begin{cases} ub(f(B)) & \text{if B contains a feasible point} \\ +\infty & \text{otherwise.} \end{cases} \tag{5.17}$$

In order to prove that a given box B contains a feasible point, we proceed as follows:

(1) Without restriction of generality, let

$$\begin{aligned} & lb(p_i(B)) \leq 0 < ub(p_i(B)) && \text{for } i = 1, \ldots, m'(\leq m) \\ & 0 \in q_j(B) \wedge q_j(B) \neq 0 && \text{for } j = 1, \ldots, r'(\leq r) \\ & s := m' + r'. \end{aligned} \tag{5.18}$$

where m' is the number of inequalities we can not determine whether they hold over the given box B, r' is the number of equalities we can not determine if they hold over B. The remaining conditions should not be violated:

$$\begin{aligned} &ub(p_i(B)) \le 0 && \text{for } i = m'+1,\ldots,m \\ &lb(q_j(B)) = ub(q_j(B)) = 0 && \text{for } j = r'+1,\ldots,r. \end{aligned} \tag{5.19}$$

(2) Select s components of B (for simplicity the first s) and build a box

$$Z = X_1 \times \cdots \times X_s.$$

Let $c = (c_1,\ldots,c_n)$ be the midpoint of B. Let

$$\begin{aligned} &h_i(z) = p_i(z, c_{s+1},\ldots,c_n) && i = 1,\ldots,m' \\ &h_{m'+j}(z) = q_j(z, c_{s+1},\ldots,c_n) && j = 1,\ldots,r', \end{aligned} \tag{5.20}$$

where $z = (x_1,\ldots,x_s)^T$. Then $h = (h_1,\ldots,h_s)^T$ is an s-dimensional vector-value function.

(3) Apply one step of the interval Newton method to solve $h(z) = 0$. Suppose that Z' is the result of solving $h(z) = 0$ by using one step of the interval Newton method. The property of the interval Newton method guarantees that there is a unique solution to $h(z) = 0$ within Z if Z' is contained in the interior of Z and not empty. So if $Z' \subset Z$ and $Z' \ne \emptyset$, there exists a feasible point $x' = (x'_1,\ldots,x'_s, c_{s+1},\ldots,c_n)^T$ in B.

Here we integrate formulas 5.14 and 5.17 into one formula for finding an upper bound u of f^* as follows:

$$u = f_b(B) = \begin{cases} ub(f(c)) & \text{if no equalities and } ub(p_i(c)) \le 0 \text{ for all } i = 1,\ldots,m \\ ub(f(B)) & \text{if equalities exist and applying one step of the interval} \\ & \quad \text{Newton method to solve } h(z) = 0 \text{ for } z \in Z \text{ gives the} \\ & \quad \text{result } Z' \subset Z \text{ and } Z' \ne \emptyset \\ +\infty & \text{otherwise} \end{cases} \tag{5.21}$$

where $c = (c_1, \ldots, c_n)^T$ is the midpoint of $B = X_1 \times \cdots \times X_n$, $Z = X_1 \times \cdots \times X_s$, h is an s-dimensional vector-value function whose components are defined in formula 5.20.

It is possible that for every sub-box B split in a Branch-and-Bound algorithm we can not conclude that there is a feasible point in B. In this case, we do not obtain an upper bound of f^*. This makes the constrained global optimization problem hard in interval arithmetic.

Inequality and Equality Test. Given a box B, if the problem's defining conditions (i.e., the inequality condition and the equality condition) can not be satisfied anywhere in B, then B is rejected. In interval arithmetic, $p_i(B)$ $(i = 1, \ldots, m)$ and $q_j(B)$ $(j = 1, \ldots, r)$ are computed and the results are denoted by

$$p_i(B) = [lb(p_i(B)), ub(p_i(B))]$$

and

$$q_j(B) = [lb(q_j(B)), ub(q_j(B))].$$

If we can determine that

$$lb(p_i(B)) \leq 0 \quad \text{for all } i = 1, \ldots, m \tag{5.22}$$

does not hold, then $p_i(x) > 0$ for all $x \in B$ for some $i \in \{1, \ldots, m\}$ (i.e., there is no point $x \in B$ such that $p_i(x) \leq 0$) and B is rejected. We call the inequality 5.22 the *inequality test.*

If we can determine that

$$0 \in [lb(q_j(B)), ub(q_j(B))] \quad \text{for all } j = 1, \ldots, r, \tag{5.23}$$

i.e.,

$$lb(q_j(B)) \leq 0 \text{ and } 0 \leq ub(q_j(B)) \quad \text{for all } j = 1, \ldots, r, \tag{5.24}$$

does not hold, then there is no point $x \in B$ such that $q_j(x) = 0$ for some $j \in \{1, \ldots, r\}$ and B is rejected. We call the inequality 5.24 the *equality test*.

Note: Hansen treated each equality condition $q_j(x) = 0$ $(j = 1, \ldots, r)$ as two inequality conditions

$$q_j(x) \leq 0 \text{ and } q_j(x) \geq 0.$$

If $lb(q_j(B)) \leq 0$ and $ub(q_j(B)) \geq 0$ does not hold (i.e., $lb(q_j(B)) > 0$ or $ub(q_j(B)) < 0$), then the inequality conditions do not hold and B is rejected. This is equivalent to treating $q_j(x) = 0$ as an equality condition.

Fritz-John Test. Hansen solves the Fritz-John condition $F(t) = 0$ using the interval Newton method. Suppose that we seek the solution of $F(t) = 0$ in a given box B and the interval Newton method produces the result B'. If

$$B' \cap B \neq \emptyset \tag{5.25}$$

does not hold, then there is no solution in B and B is rejected. The inequality 5.25 is called the *Fritz-John test*.

5.3 Branch-and-Bound for Constrained Optimization

The Branch-and-Bound algorithm for Constrained Global Optimization, which is denoted by *BBCGO*, is described in Figure 5.1. The algorithm is applicable to both interval arithmetic and interval constraints. We obtain this algorithm from the Branch-and-Bound algorithm shown in Figure 4.2 for Unconstrained Global Optimization (denoted by *BBUGO* in the following) by making a few changes. The underlined parts in Figure 5.1 are the main changes.

5.3.1 Methods for Finding An Upper Bound

In *BBUGO*, we use

$$f^*_{ub} := ub(f(mid(X))), \quad u_1 := ub(f(mid(B_1))), \quad u_2 := ub(f(mid(B_2)))$$

to obtain an upper bound of f^* over boxes X, B_1 and B_2 respectively. For an unconstrained global optimization problem, the value of the objective function f at an arbitrary point of a box is an upper bound of f^* over the box. We choose the midpoint there.

In *BBCGO*, we use

$$f^*_{ub} := f_b(X), \quad u_1 := f_b(B_1), \quad u_2 := f_b(B_2)$$

to obtain an upper bound of f^* over boxes X, B_1 and B_2 respectively, where one definition of f_b is given in Formula 5.21, another definition will be discussed in the following section. For a constrained global optimization problem, the value of the objective function f at an arbitrary point of a box is only usable if that point is feasible.

5.3.2 Conditions for Rejecting Boxes

In *BBUGO*, we can always use condition 4.2 (i.e., the midpoint condition) for rejecting boxes. If the function f is differentiable, conditions 4.4 and 4.9 (i.e., the stationarity condition and the linear condition) can be used. If f is twice differentiable, we can use

---

input: $f, p_i(i = 1, \ldots, m), q_j(j = 1, \ldots, r), X, \epsilon_X, \epsilon_F, \epsilon_P, \epsilon_Q$;
$f^*_{ub} := f_b(X)$;
$L := [< X, lb(f(X)), f^*_{ub} >]$;
$A := \emptyset$;
while $L \neq \emptyset$ **do**
    remove $< B, l, u >$ from L;
    if $|B| \leq \epsilon_X \wedge |f(B)| \leq \epsilon_F \wedge ub(p_i(B)) \leq \epsilon_P \wedge |q_j(B)| \leq \epsilon_Q \wedge 0 \in q_j(B)$ **then**
        add $< B, l, u >$ to A
    else **if** conditions($< B, l, u >, f^*_{ub}$) **then**
        partition B into B_1 and B_2 with midpoints m_1 and m_2;
        $l_1 := lb(f(B_1))$; $l_2 := lb(f(B_2))$;
        $u_1 := f_b(B_1)$; $u_2 := f_b(B_2)$;
        $f^*_{ub} := min(f^*_{ub}, u_1, u_2)$;
        add $< B_1, l_1, u_1 >$ and $< B_2, l_2, u_2 >$ to L;
    end-if
end-while
$f^*_{lb} := \min\{l_i \mid < B_i, l_i, u_i > \in A\}$;
if $f^*_{ub} = +\infty$ **then**
    $f^*_{ub} := \max\{ub(f(B_i)) \mid < B_i, l_i, u_i > \in A\}$;
    output: f^* is in $[f^*_{lb}, f^*_{ub}]$ if there exists a feasible point in X, A;
else
    output: There is a feasible point in X, f^* is in $[f^*_{lb}, f^*_{ub}]$, A;
end-if

Figure 5.1: Branch-and-Bound algorithm for constrained optimization

---

conditions 4.6, 4.11 and 4.14 (i.e., the convexity condition, the quadratic condition and the Newton condition).

In *BBCGO*, the problem's defining conditions 1.1 and 1.2 can be used for rejecting boxes. Condition 4.2 (i.e., the midpoint condition) is always applicable. We can also use conditions 4.9, 4.11 (i.e., the linear condition, the quadratic condition) and the Newton condition if the function f is twice differentiable. Although we can not use conditions 4.4 and/or 4.6 (i.e., the stationarity condition and/or the convexity condition) in *BBCGO*, we can use condition 5.1 (i.e., the Fritz-John condition).

5.3.3 Conditions for the Answer List

In *BBUGO* shown in Figure 4.2, a triple $< B, l, u >$ is put into the answer list A if it satisfies the following condition:

$$|B| \leq \epsilon_X \wedge (u - l) \leq \epsilon_F. \tag{5.26}$$

This condition guarantees that the box B of every triple $< B, l, u >$ in the answer list A is small enough and the value of the function f at the midpoint of B is close enough to the real minimum f^*. However, this condition is not suitable for constrained global optimization. It does not consider the problem's defining conditions. Some conditions should be added for considering the problem's defining conditions. Furthermore, it is hard for the triple $< B, l, u >$ to satisfy the condition

$$(u - l) \leq \epsilon_F, \tag{5.27}$$

since u is $+\infty$ in general. It is better to change this condition.

Adding More Conditions for the Answer List. Suppose that $\epsilon_P > 0$ and $\epsilon_Q > 0$ are the tolerances for the inequalities and equalities of the problem's defining conditions respectively. We usually want to assure that any point x in the box B of every triple $< B, l, u >$ in the answer list satisfies the conditions

$$p_i(x) \leq \epsilon_P \quad \text{for } i = 1, \ldots, m \tag{5.28}$$

and

$$-\epsilon_Q \leq q_j(x) \leq \epsilon_Q \quad \text{for } j = 1, \ldots, r. \tag{5.29}$$

We can achieve this by adding the conditions

$$ub(p_i(B)) \leq \epsilon_P \quad \text{for } i = 1, \ldots, m \tag{5.30}$$

and

$$|q_j(B)| \leq \epsilon_Q \wedge 0 \in q_j(B) \quad \text{for } j = 1, \ldots, r \tag{5.31}$$

into condition 5.26, i.e., by using conditions 5.26, 5.30 and 5.31 instead of just using condition 5.26.

Condition 5.30 guarantees that any point x in B satisfies condition 5.28. This indicates that $p_i(x) \leq 0$ holds within the given tolerance ϵ_P for $i = 1, \ldots, m$. Condition 5.31 makes sure that any point x in B satisfies condition 5.29. This means that $q_j(x) = 0$ holds within the given tolerance ϵ_Q for $j = 1, \ldots, r$.

Replacing the Condition $(u - l) \leq \epsilon_F$. For a triple $< B, l, u >$ in *BBUGO*, an upper bound u of the minimum f^* over the box B is always defined. The value of the function f at an arbitrary point of B is defined and is an upper bound of f^*. We choose the midpoint of B. However, this is not the case for constrained global optimization. An arbitrary point of B is not necessarily a feasible point, neither is the midpoint of B. Thus u is not defined in general. Therefore, condition 5.27 may not be satisfied. One suitable replacement of condition 5.27 is

$$|f(B)| \leq \epsilon_F. \tag{5.32}$$

After the replacement and the addition of conditions 5.30 and 5.31, we have the following condition in *BBCGO* for adding a triple $< B, l, u >$ into the answer list A:

$$|B| \leq \epsilon_X \wedge |f(B)| \leq \epsilon_F \wedge ub(p_i(B)) \leq \epsilon_P \wedge |q_j(B)| \leq \epsilon_Q \wedge 0 \in q_j(B) \tag{5.33}$$

where $i = 1, \ldots, m$ and $j = 1, \ldots, r$.

5.3.4 Input and Output of the Algorithm

The input and output of the algorithm for constrained global optimization contain more information than those of the algorithm for unconstrained one.

Input. In addition to f, p_i $(i = 1, \ldots, m)$ and q_j $(j = 1, \ldots, r)$, the input of *BBCGO* includes the tolerances ϵ_P and ϵ_Q. ϵ_P and ϵ_Q are for the inequalities and equalities of the problem's defining conditions respectively.

Output. The **output** statement in *BBUGO* is replaced by the following **if-then-else** statement:

if $f^*_{ub} = +\infty$ **then**

$f^*_{ub} :=$ $\max\{ub(f(B_i)) \,|< B_i, l_i, u_i >\in A\}$;

output: f^* is in $[f^*_{lb}, f^*_{ub}]$ if there exists a feasible point in X, A;

else

output: There is a feasible point in X, f^* is in $[f^*_{lb}, f^*_{ub}]$, A;

end-if .

It is possible that we can not assure that there exists a feasible point during any iteration of the algorithm. In this case, f^*_{ub} is initialized as $+\infty$ and keeps being $+\infty$ until the end of the algorithm. We can only make sure that f^* is between

$$f^*_{lb} = \min\{l_i \mid < B_i, l_i, u_i > \in A\}$$

and

$$f^*_{ub} = \max\{ub(f(B_i)) \mid < B_i, l_i, u_i > \in A\}$$

if there exists a feasible point in the original domain X. The problem we solve using the algorithm is the mathematical model of a physical problem. In general, we know that the physical problem has solutions. So we can make sure that f^* is between f^*_{lb} and f^*_{ub}.

In the case that at least one feasible point exists, we can guarantee that f^* is between f^*_{lb} and f^*_{ub}.

5.4 An Interval Constraint Method

In the previous section, we described the Branch-and-Bound algorithm for constrained global optimization. The performance of the algorithm depends on the effectiveness of the implementation of "conditions($< B, l, u >, f^*_{ub}$)" and "$f_b(B)$". If we use the mid-point test, inequality test, equality test and the Fritz-John test, we obtain Hansen's interval arithmetic method for constrained global optimization. Since all the tests use interval arithmetic, we call Hansen's algorithm an interval arithmetic version of the Branch-and-Bound algorithm for constrained global optimization, denoted by IAC.

In this section, we demonstrate how to translate all the interval arithmetic computations in "conditions($< B, l, u >, f^*_{ub}$)" and "$f_b(B)$" into an interval constraint processing task. Thus we can obtain an interval constraint version of the Branch-and-Bound algorithm for constrained global optimization, which is denoted by ICC in the following.

5.4.1 Finding An Upper Bound in Interval Constraints

In Branch-and-Bound, we need an upper bound of f^* over a given box. In unconstrained global optimization, the upper bound of the interval obtained by evaluating the objective function over the box is an upper bound for the minimum over that box. Another upper bound is the value of the objective function at an arbitrary point in the box. The latter is usually much better.

In constrained global optimization with Branch-and-Bound, an upper bound is harder to obtain, and this makes constrained optimization harder also in interval constraints. An upper bound of the interval for the objective function evaluated over the box in the presence of the constraints is easily computed in interval constraints. However, it is only valid as an upper bound in Branch-and-Bound if there exists at least one feasible point in the box. This is not known. Similarly, an upper bound obtained from evaluating the objective function at an arbitrary point is only usable if that point is feasible.

In section 5.2, we described the interval arithmetic methods for finding an upper bound u of f^* for constrained global optimization. Through translating interval arithmetic computations in formula 5.21 into an interval constraint processing task, we obtain an interval constraint method for finding u. This method is formulated as:

$$u = f_b(B) = \begin{cases} ub(f(c)) & \text{if no equalities and } ub(p_i(c)) \leq 0 \text{ for all } i = 1, \ldots, m \\ ub(f(B)) & \text{if equalities exist and applying the consistency} \\ & \quad \text{algorithm to the constraint system} \\ & \quad \text{``} z \in Z, h(c') + J(z, Z)(z - c') = 0\text{'' shrinks } Z \text{ but} \\ & \quad \text{does not produce any failure state.} \\ +\infty & \text{otherwise} \end{cases} \tag{5.34}$$

where c, B, Z and h are the same as in formula 5.21, $c' = (c_1, \ldots, c_s)^T$ is the midpoint of Z, and $J(z, Z)$ is the Jacobian matrix of h over Z.

We have three different cases for computing u:

(1) There is a feasible point c in B. In this case, we compute $ub(f(c))$ as the value of u.

(2) The box B contains a feasible point. We compute $ub(f(B))$ as the value of u.

(3) No feasible point in B. We set u to $+\infty$.

Since interval constraints is not less effective than interval arithmetic, there is not less chance to find a real upper bound u of f^* using the interval constraint method formulated in 5.34 than using interval arithmetic method formulated in 5.21. The computational results in Table 5.2 suggest that interval constraints is more effective.

5.4.2 Using Conditions as Constraints

In constrained global optimization, interval constraints are especially interesting. All the problem's defining conditions, both inequality 1.1 and equality 1.2, can be directly used as constraints. In addition, the midpoint condition 4.2 can be used as constraints. Furthermore, we can also use the Fritz-John condition or the Kuhn-Tucker condition as constraints if the function f, p_i $(i = 1, \ldots, m)$ and q_j $(j = 1, \ldots, r)$ are differentiable.

For a given box B, an upper bound f^*_{ub} of f^* and the conditions used as constraints, we have the following interval constraint system (note that we use the Fritz-John condition here):

$$\begin{array}{ll}
x \in B & \\
y = f(x), y \leq f^*_{ub} & \\
p_i(x) \leq 0 & (i = 1, \ldots, m) \\
q_j(x) = 0 & (j = 1, \ldots, r) \\
u_i \in [0, +\infty] & (i = 0, 1, \ldots, m) \\
v_j \in [-\infty, +\infty] & (j = 1, \ldots, r) \\
u_0 \nabla f(x) + \sum_{i=1}^{m} u_i \nabla p_i(x) + \sum_{j=1}^{r} v_j \nabla q_j(x) = 0 & \\
u_i p_i(x) = 0 & (i = 1, \ldots, m) \\
v_j q_j(x) = 0 & (j = 1, \ldots, r) \\
\sum_{i=0}^{m} u_i + \sum_{j=1}^{r} v_j - 1 = 0 &
\end{array} \tag{5.35}$$

where the second line is the midpoint condition, the third and fourth line are the problem's defining conditions (i.e., the inequality condition and equality condition), and the rest is the Fritz-John condition. If applying the consistency algorithm to the constraint system 5.35 results in a failure state (i.e., produces an empty box), then the box B can be rejected. It is also possible that the output of the consistency algorithm is a shrunk non-empty box B'. In this case, some of the conditions used can not hold for $B - B'$, so that we conclude that no point in $B - B'$ satisfies all the conditions used as constraints. Note that in this case we can not conclude that every point $x \in B'$ satisfies all the conditions.

In interval arithmetic the problem's defining conditions 1.1 and 1.2 can not be used directly. The Fritz-John condition or the Kuhn-Tucker condition is essential, because using either of them is the only way in which the problem's defining conditions can be indirectly taken into account.

A novel feature of interval constraint approach, and perhaps first demonstrated in [16, 84, 17], is that it allows one to select from a hierarchy of redundant conditions. The conditions on the minimization are completely determined by the problem's defining conditions 1.1 and 1.2. In principle it is possible in interval constraints to only include conditions 1.1 and 1.2.

5.5 An Implementation of the Interval Constraint Method

We have implemented in BNR-Prolog the interval constraint method for constrained global optimization. Figure 5.2 shows the main part of the source code. We omit the definitions of *small*, *fb*, *partition* and *min*. The user should provide in BNR-Prolog the definitions of f, p_i $(i = 1, \ldots, m)$ and q_j $(j = 1, \ldots, r)$. The user should also give the definitions of f', p_i' $(i = 1, \ldots, m)$ and q_j' $(j = 1, \ldots, r)$ if f, p_i $(i = 1, \ldots m)$ and

q_j $(j = 1, \ldots, r)$ are differentiable and she/he wants to use the Fritz-John condition as constraints.

For example, suppose that we want to find the minimum value of the following function

$$f(x_1, x_2) = (x_1 - 2)^2 + (x_2 - 1)^2 \qquad \text{subject to}$$
$$p_1(x_1, x_2) = x_1^2 - x_2 \tag{5.36}$$
$$p_2(x_1, x_2) = x_1 + x_2 - 2.$$

We should define the relations f, p_1 and p_2 in BNR-Prolog as follows:

```
f([X1,X2],Y) :- Y is (X1 - 2)**2 + (X2 - 1)**2.
p1([X1,X2],P1) :- P1 is X1**2 - X2.
p2([X1,X2],P2) :- P2 is X1 + X2 - 2.
```

Similarly we can define f', p_i' $(i = 1, \ldots, m)$ and q_j' $(j = 1, \ldots, r)$ in BNR-Prolog.

Here we explain the relations defined in the program for the *ICC* that are different from those defined for *ICU*.

The formula "$small(B, Y)$" holds if $|B| \leq \epsilon_X$, $|Y| \leq \epsilon_F$, $ub(p_i(B)) \leq \epsilon_P$ (for all $i = 1, \ldots, m$), $|q_j(B)| \leq \epsilon_Q$ and $0 \in q_j(B)$ (for all $j = 1, \ldots, r$).

The formula "$fb(B, U)$" holds if the formula 5.34 holds.

The formula "$conditions([B, Y], Fub)$" holds if applying the consistency algorithm to the constraint system 5.35 does not produce any failure state.

```
g_min([],Fub,A,A) :- !.            % L = [], terminate, Answers are in A
g_min([[B,Y,U]|L0],Fub,A0,A) :-% remove [B,Y,U] from L=[[B,Y,U]|L0]
   small(B,Y), !,                  % |B|=<Ex, |Y|=<Ey, ub(pi(B))=<Ep,
                                   % |qj(B)|=<Eq, and  0 in qj(B)
   insert([B,Y,U],A0,A1),          % insert [B,Y,U] into A0 and get A1
   g_min(L0,Fub,A1,A).
g_min([[B,Y,U]|L0],Fub,A0,A) :-
   conditions([B,Y],Fub), !,       % conditions may be satisfied
   partition(B,B1,B2),             % partition box B into B1 and B2
   f(B1,Y1), fb(B1,U1),            % Y1=f(B1),U1=ub(fb(m1)),m1=midpoint(B1)
   f(B2,Y2), fb(B2,U2),            % Y2=f(B2),U2=ub(fb(m2)),m2=midpoint(B2)
   min([Fub,U1,U2],Fub1),          % Fub1=min(Fub,U1,U2)
   insert([B1,Y1,U1],L0,L1),       % insert [B1,Y1,U1] and [B2,Y2,U2]
   insert([B2,Y2,U2],L1,L2),       % into L0 and get L2
   g_min(L2,Fub1,A0,A).
g_min([[B,Y,U]|L0],Fub,A0,A) :-% conditions can not be satisfied
   g_min(L0,Fub,A0,A).             % reject [B,Y,U] from [[B,Y,U]|L0]

conditions([B,Y],Fub) :-           %   constraints for:
   Y =< Fub,                                 % condition f(x) <= Fub
   p1(B,P1),P1=<0, ..., pm(B,Pm),Pm=<0, % inequalities
   q1(B,Q1),Q1==0, ..., qr(B,Qr),Qr==0, % equalities
   fritz_john_condition(B).                  % Fritz-John condition

insert([B,Y,U],[],[[B,Y,U]]) :- !.
insert([B,Y,U],[[B1,Y1,U1]|T],[[B,Y,U],[B1,Y1,U1]|T]) :-
   range(Y, [L,_]), range(Y1, [L1,_]), L =< L1, !.
insert([B,Y,U],[[B1,Y1,U1]|T],[[B1,Y1,U1]|R]) :- insert([B,Y,U],T,R).
```

Figure 5.2: Code of interval constraint method for constrained optimization

5.6 Test Results

In addition to conditions 1.1 and 1.2 (i.e., the problem's defining conditions), there are several conditions that are redundant, but are extremely effective in reducing the required amount of computation. Condition 4.2 (i.e., the midpoint condition) and condition 5.1 (i.e., the Fritz-John condition) are independent of each other. In our first series of tests (ICC_1 in Table 5.2) we only use conditions 1.1, 1.2 and 4.2. The novelty here is that we take the problem's defining conditions into account without using the Fritz-John condition.

In ICC_2 of Table 5.2, we improve on ICC_1 by *adding* the Fritz-John condition. As one can see, this gives a considerable improvement in performance in two of the three test problems.

Since our results give guaranteed inclusions, it only makes sense to compare performance with the methods that give the same guarantees. As far as we know, these can only be found in [74, 36, 37, 38]. Although E. R. Hansen, in [36], gives a copious variety of test results on unconstrained optimization, he gives the test result for only one problem on constrained optimization, which we include in Table 5.1. Even for these few problems, the performance measures given in [74, 36] are spotty; see the open spaces in Table 5.2.

Although Van Hentenryck and Michel reported the performance of their Numerica system in [37, 38], they only give timings and numbers of splits. Comparisons are hard to make from [37, 38]. The number of splits have the advantage of being machine-independent. However, they do not report the splits performed in narrowing by means of Newton's method and an inequality ("internal splits").

Problem	Objective Function and Conditions	Domain
1	$f(x) = 0.1(x_1^2 + x_2^2)$ $p_1(x) = 2\sin(2\pi x_2) - \sin(4\pi x_1)$	$[-1,1] \times [-1,1]$
2	$f(x) = (x_1 - 2)^2 + (x_2 - 1)^2$ $p_1(x) = x_1^2 - x_2$ $p_2(x) = x_1 + x_2 - 2$	$R \times R$
3	$f(x) = 12x_1^2 - 6.3x_1^4 + x_1^6 + 6x_1x_2 + 6x_2^2$ $p_1(x) = 1 - 16x_1^2 - 25x_2^2$ $p_2(x) = 13x_1^3 - 145x_1 + 85x_2 - 400$ $p_3(x) = x_1x_2 - 4$	$[-2,4] \times [-2,4]$

Table 5.1: Test problems for constrained global optimization

We also successfully ran our algorithm on many of the problems with three or fewer variables in [29, 42, 75].

Table 5.2 shows the performance figures from [74, 36], as far as given, (in the part labeled IAC, which is short for Interval Arithmetic method for Constrained global optimization) and compares with two versions of our algorithm (in the parts labeled ICC_1 and ICC_2; see above for their differences).

Here is an explanation of the symbols used in the tables.

W_i Width of initial box.

ϵ Intended tolerance for box size, function width, inequalities and equalities.

W_{x^*} Width of final box for an occurrence x^*.

W_{f^*} Width of bound on the global minimum f^*.

N_i Number of iterations of an algorithm till termination.

N_f Total number of function evaluations.

Each of these symbols with superscript A represents that the corresponding results are from running the interval Arithmetic version of an algorithm, that with superscript C represents that the corresponding results are from running the interval Constraint version of the algorithm.

From Table 5.2, one can see that the interval constraint algorithm using conditions 1.1, 1.2, 4.2 and 5.1 as constraints achieves a factor of 4 to 58 in the number of iterations on the constrained global optimization test problems and a factor of 23 to 76 in the number of function evaluations over the interval arithmetic methods. Moreover, the final boxes for x^* and f^* produced by our interval constraint algorithm are smaller.

Comparing the running results of ICC_1 and ICC_2, we find that using the Fritz-John conditions as redundant constraints gives a considerable improvement in performance in two of the three test problems. For these two problems, ICC_2 achieves a factor of 10 to 18 in the number of iterations and a factor of 9 to 34 in the number of function evaluations over ICC_1.

5.7 Discussion

We have compared our algorithm with the one of Hansen. He uses Interval Newton method to solve the Fritz-John condition. Conventionally, some form of Newton method has to be used to solve this condition. We only use the condition in its definitional form to prune the Branch-and-Bound search. In our approach, the use of Inter-

		Problem 1	Problem 2	Problem 3
Input	W_i	2	$+\infty$	6
	ϵ	10^{-4}	10^{-6}	10^{-2}
IAC	N_i^A	175		72
	N_f^A	686	252	
	$W_{x^*}^A$	10^{-4}	10^{-6}	10^{-2}
	$W_{f^*}^A$	10^{-10}	10^{-6}	10^{-2}
ICC_1	$N_i^{C_1}$	3	51	335
	$N_f^{C_1}$	9	101	995
	$W_{x^*}^C$	0	10^{-7}	10^{-3}
	$W_{f^*}^C$	0	10^{-7}	10^{-3}
ICC_2	$N_i^{C_2}$	3	5	19
	$N_f^{C_2}$	9	11	29
	$W_{x^*}^C$	0	10^{-16}	10^{-5}
	$W_{f^*}^C$	0	10^{-16}	10^{-5}
Comparison	$\frac{N_i^A}{N_i^{C_2}}$	58		4
	$\frac{N_f^A}{N_f^{C_2}}$	76	23	
	$\frac{N_i^{C_1}}{N_i^{C_2}}$	1	10	18
	$\frac{N_f^{C_1}}{N_f^{C_2}}$	1	9	34

Table 5.2: Running results of IAC and ICC

val Newton amounts to the addition of constraints that are redundant with respect to the definition of the Fritz-John condition, which is itself redundant to conditions 1.1 and 1.2. So far experience shows that adding additional redundant conditions as constraints generally improves performance. Without this additional possibility we already achieve better performance than the results published by Hansen.

Chapter 6

Concluding Remarks

6.1 Summary and Contributions

Interval Arithmetic and Interval Constraints. We have specified the essential components of interval arithmetic and interval constraints, which include interval functions, interval constraint systems and consistency algorithms. Interval constraints is based on interval arithmetic, but is a generalization of interval arithmetic. An interval function F of a real function f over a given domain X can be computed in interval constraints. $F(X)$ can be translated into an interval constraint system C, in which there is a variable y for the value of F. We have proved that applying a consistency algorithm to C gives the same interval result for y as computing $F(X)$ in interval arithmetic.

The interval value of $F(X)$ is an approximation of $\Box f(X)$, the smallest interval containing the range of function f over X. In general it is much larger than $\Box f(X)$.

Hypernarrowing. We have presented an algorithm called hypernarrowing algorithm, which is based on a consistency algorithm. Through applying the hypernarrowing algorithm to the constraint system C translated from $F(X)$ and the interval for y, we usually obtain a much smaller interval for y than applying the consistency algorithm to C. Combining the hypernarrowing with the simple version of Hansen's algorithm for unconstrained global optimization that uses interval arithmetic and the midpoint test, we have achieved a factor of 1.8 to 9.7 in speedup over the simple version. We have also compared the semantics of the hypernarrowing algorithm with that of a constraint solver.

Unconstrained Global Optimization. After reviewing interval arithmetic methods and describing a generic Branch-and-Bound algorithm for solving unconstrained global optimization problems, we have proved the properties of the algorithm. We have investigated the role of interval constraints in global optimization, explained why interval constraints is more powerful than interval arithmetic and demonstrated how to obtain an interval arithmetic version and an interval constraint version of the Branch-and-Bound algorithm. The implementations of a variety of the interval arithmetic and interval constraint versions of the Branch-and-Bound algorithm have been given in BNR-Prolog. Computational results have showed that the interval constraint version of the order zero algorithm achieves a factor of 2.7 to 27.5 in speedup on the unconstrained global optimization test problems over the corresponding interval arithmetic version. The interval constraint version of the order one algorithm achieves a factor of 5.8 to 123.5 in speedup over the corresponding interval arithmetic version. In addition to the higher speed, the interval constraint versions are more declarative and use less memory space. We have also proposed two improvements on the memory use of the Branch-and-Bound algorithm. The interesting thing is that these improvements also speed up the execution of the algorithm on about half of the unconstrained global optimization test problems.

Constrained Global Optimization. We have discussed the differences between solving the unconstrained global optimization problem and the constrained one. After reviewing the interval arithmetic methods for constrained global optimization, we have presented the transition from the Branch-and-Bound algorithm for unconstrained global optimization to the one for constrained global optimization. We have demonstrated how to obtain an interval constraint version of the Branch-and-Bound through using the applicable conditions as constraints. The interval constraint version has been implemented in BNR-Prolog and run on a few of test problems. The test results indicate that the interval constraint method using the problem's defining conditions, the midpoint condition and the Fritz-John condition as interval constraints in their definitional form achieves a factor of 4 to 58 in the number of iterations on the constrained optimization test problems over the interval arithmetic method. We have also investigated the effect of using the redundant Fritz-John condition as constraints. Computational results show that it gives considerable improvement in performance in most of cases.

6.2 Suggestions For Future Work

Combining a Point Method with an Interval Constraint One. The merit of point methods is their high efficiency. A point method can provide a real-valued approximate minimum f_m and the minimizer. The minimum f_m can be used as an upper bound of f^* in an interval constraint method. In the interval constraint method, an upper bound u of f^* is obtained by computing the function value at a sampled point or the upper bound of function f over box B. Since f_m is generally much lower than u, using f_m as the upper bound of f^* may greatly speed up the interval constraint method. In addition, we can use the minimizer to guide the splitting of a box into sub-boxes. This leads to a considerable improvement of interval arithmetic methods [14]. Therefore, it is worth studying the combination of a point method and an interval constraint method.

Among point methods, a local optimization method is much faster than a global one [92]. So the higher priority should give to the study of the combination of a local optimization method and an interval constraint method.

Interval Constraint Compiler A high-level interval constraint system is translated into a low-level one, which consists of a conjunction of primitive constraints and an initial state. The consistency algorithm transforms the initial state into a consistent state or a failure state through "interpreting" the primitive constraints (i.e., applying the corresponding consistency operators). The consistency algorithm is an interpretor and the primitive constraints are the intermediate instructions to be interpreted. It is possible to design and implement an interval constraint compiler that translates a low-level (or high-level) interval constraint system into a sequence of C/C++ or assembly instructions [15]. This is analogous to designing a Prolog compiler [1, 91, 90, 54, 19] based on the techniques used in a Prolog interpretor [86, 13, 79]. In the place of a Prolog interpretor is the consistency algorithm. In general, the object code generated by a compiler from a source program achieves a factor of one to two order magnitudes in speed over the interpreted program [27].

Parallel Interval Constraint Methods. Interval constraint methods for solving the global optimization problem are based on an exhaustive search in a given box. The box is split into sub-boxes and the search continues in a selected sub-box. Since there are a number of sub-boxes during the search, we can use N processors and execute the interval constraint algorithm on each processor for searching a selected sub-box in parallel.

Much research work has been done about parallel interval arithmetic algorithms for global optimizations [48, 9, 14, 25, 26, 50, 63]. It is a good starting point for studying parallel interval constraint algorithms.

Improving the Consistency Algorithm. There exists potential parallelism in computing each consistency operator and the consistency algorithm. Exploiting this parallelism, we can improve the efficiency of solving interval constraints systems. There are two levels of parallelism that can be exploited to speed up the consistency algorithm. The computation of the consistency operator for a primitive constraint can be parallelized. For example, from the formula for computing C_{sum}, the consistency operator for the primitive constraint *sum*,

$$\begin{aligned} C_{sum}([a,b] \times [c,d] \times [e,f]) &= [a,b] \cap ([e,f] - [c,d]) \times \\ &\quad [c,d] \cap ([e,f] - [a,b]) \times \\ &\quad [e,f] \cap ([a,b] + [c,d]) \end{aligned}$$

we can see that three interval operations and intersections can be parallelized. In the consistency algorithm, we only choose one primitive constraint at each iteration and apply the consistency operator of the constraint. It is possible to select several primitive constraints at an iteration and apply the consistency operators corresponding to these constraints in parallel.

Bibliography

[1] Hassan A¨-Kaci. *Warren's Abstract Machine.* MIT Press, 1991.

[2] Götz Alefeld and Jürgen Herzberger. *Introduction to Interval Computations.* Academic Press, 1983.

[3] A. B. Babichev, O. B. Kadyrova, T. P.Kashevarova, A. S. Leshchenko, and A. L. Semenov. UniCalc, a novel approach to solving systems of algebraic equations. *Interval Computations*, 1993(2):29–47, 1993.

[4] P.G. Bao and J.G. Rokne. Existence of a unique zero of nonlinear systems. 1987.

[5] F. Benhamou, D. McAllester, and P. Van Hentenryck. CLP(Intervals) revisited. In *Logic Programming: Proc. 1994 International Symposium*, pages 124–138, 1994.

[6] Frédéric Benhamou and William J. Older. Applying interval arithmetic to real, integer, and Boolean constraints. *Journal of Logic Programming.* To appear.

[7] Frédéric Benhamou and William J. Older. Programming with CLP(BNR): Examples on finite domains. To be published Journal of Logic Programming. 1993.

[8] A. Bernat and V. Kreinovich. Data processing beyond traditional statistics: applications of interval computations. a brief introduction. In V. Kreinovich, editor, *Extended Abstracts of International Workshop on Applications of Interval Computations*, pages 13–21, El Paso, TX, 1995.

[9] S. Berner. A parallel method for verified global optimization. In G. Alefeld, A. Frommer, and B. Lang, editors, *Scientific Computing and Validated Numerics*, Mathematical Research, volume 90, pages 200–206, Berlin, 1996. Akademie Verlag.

[10] BNR. BNR Prolog user guide and reference manual. 1988.

[11] Nicolas Bourbaki. *Elements of Mathematics: Theory of Sets.* Addison-Wesley Publishing Company, 1968.

[12] J. C. Burkill. Functions of Intervals. In *Proceedings of the Lodon Mathematical Society*, pages 22:375–446, 1924.

[13] J.A. Campbell. *Implementations of Prolog.* Ellis Horwood, 1984.

[14] O. Caprani, B. Godthaab, and K. Madsen. Use of a real-valued local minimum in parallel interval global optimization. *Interval Computations*, 1993(2):71–82, 1993.

[15] H. M. Chen and M. H. van Emden. Compiling interval constraint systems. In preparation.

[16] H. M. Chen and M. H. van Emden. Adding interval constraints to the Moore-Skelboe global optimization algorithm. In V. Kreinovich, editor, *International Journal of Reliable Computing (Formerly Interval Computations) Supplement*, pages 54–57, 1995.

[17] H. M. Chen and M. H. van Emden. An interval constraint method for constrained global optimization. In *Workshop on Constraint Programming Applications: An Inventory and Taxonomy*, Cambridge, Massachusetts, USA, August, 1996.

[18] J. G. Cleary. Logical arithmetic. *Future Computing Systems*, 2:125–149, 1987.

[19] Philippe Codognet and Daniel Diaz. wamcc: compiling Prolog to C. In Leon Sterling, editor, *Proceedings of the Twelfth International Conference on Logic Programming*, pages 317–331. MIT Press, 1995.

[20] Alain Colmerauer. An introduction to Prolog III. *Communications of the ACM*, 33(7):69–90, 1990.

[21] D. Stevenson, chairman, Floating-Point Working Group, Microprocessor Standards Subcommittee. IEEE standard for binary floating point arithmetic (IEEE/ANSI 754-1985). Technical report, IEEE, 1985.

[22] E. Davis. Constraint propagation with labels. *Artificial Intelligence*, 32:281–331, 1987.

[23] M. Dincbas, P. Van Hentenryck, H. Simonis, A. Aggoun, T. Graf, and F. Berthier. The constraint programming language CHIP. In *Proc. Int. Conf. on Fifth Generation Computer Systems*, 1988.

[24] L. C. W. Dixon and G. P. Szego. *Towards Global Optimization 2*. North-Holland, Amsterdam, 1978.

[25] J. Eriksson. *Parallel Global Optimization using Interval Analysis*. PhD thesis, University of Umeå, Institute of Information Processing, 1991.

[26] J Eriksson and P. Lindstrœm. A Parallel Interval Method Implementation for Global Optimization Using Dynamic Load Balancing. *Reliable Computing*, 1(1):77–92, 1995.

[27] Charles N. Fischer and Jr. Richard J. LeBlanc. *Crafting a Compiler with C*. The Bejamin/Cummings Publishing Company, Inc., 1991.

[28] R. Fletcher. *Practical Methods of Optimization*. John Wiley and Sons Ltd., 1990.

[29] C. A. Floudas and P. M. Pardalos. *A Collection of Test Problems for Constrained Global Optimization Algorithms*. Springer-Verlag, 1990. Lecture Notes in Computer Science 455.

[30] David Goldberg. What every computer scientist should know about floating-point computation. *Computing Surveys*, 1991.

[31] Paul R. Halmos. *Naive Set Theory*. D. Van Nostrand, 1960.

[32] E. R. Hansen and G. W. Walster. Nonlinear equation and optimization. 1987.

[33] Eldon Hansen. On solving equations using interval arithmetic. *Math. Comput.*, 22:374–384, 1968.

[34] Eldon Hansen. Global optimization using interval analysis – the one-dimensional case. *Jour. Optimiz. Theory Applic.*, 29:331–344, 1979.

[35] Eldon Hansen. Global optimization using interval analysis – the multi-dimensional case. *Numer. Math.*, 34:247–270, 1980.

[36] Eldon Hansen. *Global Optimization Using Interval Analysis*. Marcel Dekker, 1992.

[37] P. Van Hentenryck and L. Michel. Helios: a modeling language for global optimization. In *Proceedings of the Second International Conference on the Practical Application of Constraint Technology*, pages 317–335. The Practical Applications Company, 1996.

[38] Pascal Van Hentenryck, Laurent Michel, and Yves Deville. *Numerica: A Modeling Language for Global Optimization.* MIT Press, 1997.

[39] T. Hickey. CLP(F) and constrained ODEs,. In Yap Editors Jourdan, Lim, editor, *Proceedings of the Workshop on Constraint Languages and their use in Problem Modelling*, pages 69–79, Nov. 1994.

[40] T. Hickey, M.H. van Emden, and H. Wu. A unified framework for interval constraints and interval arithmetic. Technical report, University of Victoria, Canada, 1998.

[41] Einar Hille. *Analysis.* Blaisdell Publishing Company, 1964.

[42] Willi Hock and Klaus Schittkowski. *Test Examples for Nonlinear Programming Codes.* Springer-Verlag, 1981. Lecture Notes in Economics and Mathematicl Systems 187.

[43] K. Ichida and Y. Fujii. An interval arithmetic method for global optimization. *Computing*, 23(1):85–97, 1979.

[44] J. Jaffar and J.-L. Lassez. Constraint logic programming. In *ACM Conference on Principles of Programming Languages*, Munich, 1987.

[45] A. H. G. Rinnooy Kan and G. T. Timmer. New methods in optimization and their industrial uses. In *Argument for the Unsolvability of Global Optimization Problems*, pages 133–155. Birkhauser Verlag, 1989. Basel.

[46] R. B. Kearfott. An interval branch and bound algorithm for bound constrained optimization problems. *Journal of Global Optimization*, 2:259–280, 1992.

[47] R. Baker Kearfott. *Rigorous Global Search: Continuous Problems.* Kluwer Academic Publishers, 1996. Nonconvex Optimization and Its Applications.

[48] Vladik Kreinovich and Andrew Bernat. Parallel algorithms for interval computations: An introduction. *Interval Computations*, (3):6–62, 1994.

[49] H. W. Kuhn and A. W. Tucker. Nonlinear programming. In *The Second Berkeley Symposium on Mathematical Statistics and Probability*. University of California Press, 1951. Ed. J. Neyman.

[50] A. Leclerc. Parallel interval global optimization in C++. *Interval Computations*, 1993(3):148–163, 1993.

[51] J. H. M. Lee. *Numerical Computation As Deduction In Constraint Logic Programming*. PhD thesis, University of Victoria, Victoria, BC, Canada, 1992.

[52] J. H. M. Lee and M. H. van Emden. Adapting CLP(R) to floating-point arithmetic. In *Proceedings of the International Conference on Fifth Generation Computer Systems 1992*, pages 996–1003, 1992.

[53] Wm Leler. *Constraint Programming Languages: Their Specification and Generation*. Addison-Wesley Publishing Compnay, Inc., 1988.

[54] M. R. Levy and R. N. Horspool. Translator-based multi-paradigm programming. *Journal of Software and Systems*, 1993.

[55] H. R. Lewis and L. Denenberg. *Data Structures and Their Algorithms*. Harper Collins Publishers, 1991.

[56] J. D. C. Little, K. G.M urty, and D. W. Karel. An algorithm for the travelling salesman problem. *Operations Research*, 11, 1963.

[57] Gus Lopez, Bjorn Freeman-Benson, and Alan Borning. Kaleidoscope: A constraint imperative programming language. In Brian Mayoh, editor, *Constraint Programming, NATO ASI Series*, pages 313–329. Springer-Verlag, 1994.

[58] Brian Mayoh, Ann Tyugu, and Tarmo Uustalu. Constraint satisfaction and constrain programming: A brief lead-in. In *Constraint Programming, NATO ASI Series, Series F: Computer and System Science Vol. 131*. Springer-Verlag, 1994. Ed. Brian Mayoh.

[59] R. E. Moore. *Interval Arithmetic and Automatic Error Analysis in Digital Computing*. PhD thesis, Stanford University, October 1962.

[60] R. E. Moore. A test for existence of solutions to nonlinear systems. *SIAM J. Numer. Anal.*, 14(4):611–615, September 1977.

[61] R. E. Moore. Method and applications of interval analysis. *SIAM Studies in Applied Mathematics. SIAM, Philadelphia*, 1979.

[62] R. E. Moore. Global optimization to prescribed accuracy. *Computers Math. Applic.*, 21:25–39, 1991.

[63] R. E. Moore, E. Hansen, and A. Leclerc. Rigorous methods for parallel global optimization. In A. Floudas and P. Pardalos, editors, *Recent Advances in Global Optimization*, pages 321–342, Princeton, N.J., 1992. Princeton Univ. Press.

[64] Ramon E. Moore. *Interval Analysis.* Prentice-Hall, 1966.

[65] W. J. Older. CLP(BNR) algorithm for travelling salesman. Technical report, Bell-Northern Research Computing Research Laboratory, August, 1993.

[66] W. J. Older and A. Vellino. Constraint arithmetic on real intervals. In *Constraint Logic Programming: Selected Research.* MIT Press, 1993. Eds. F. Benhamou and A. Colmerauer.

[67] William Older. Constraints in BNR Prolog. Bell-Northern Research Technical Report, 1993.

[68] William Older and André Vellino. Constraint arithmetic on real intervals. In Frédéric Benhamou and Alain Colmerauer, editors, *Constraint Logic Programming*, pages 175–195. MIT Press, 1993.

[69] William J. Older. Application of relational interval arithmetic to ordinary differential equations. In Jean Jourdain, Pierre Lim, and Roland H.C. Yap, editors, *Workshop on Constraint Languages and their Use in Problem Modelling*, pages 60–69, Ithaca, New York, November 1994.

[70] W.J. Older. The application of relational arithmetic to X-ray diffraction crystallography. Technical report, Bell-Northern Research Computing Research Laboratory, 1989.

[71] W.J. Older. Interval arithmetic specification. Technical report, Bell-Northern Research Computing Research Laboratory, 1989.

[72] W.J. Older. Using interval arithmetic for non-linear constrained optimization. Technical report, Bell-Northern Research Computing Research Laboratory, 1993.

[73] W.J. Older. Note on computing square-well spectra using interval arithmetic. Technical report, Bell-Northern Research Computing Research Laboratory, 1994.

[74] H. Ratschek and J. Rokne. *New Computer Methods for Global Optimization.* Ellis Horwood/John Wiley, 1988.

[75] Klaus Schittkowski. *More Test Examples for Nonlinear Programming Codes.* Springer-Verlag, 1987. Lecture Notes in Economics and Mathematicl Systems 282.

[76] Alexander L. Semenov. Solving optimization problems with help of the UniCalc solver. In *Applications of Interval Computations.* Kluwer Academic Publishers. 1996. Eds. R. Baker Kearfott and Vladik Kreinovch.

[77] M. Simonnard. *Linear Programming.* Prentice-Hall, Inc., Englewood Cliffs, N.J. 1966. translated by W.S.Jewell.

[78] S. Skelboe. Computation of rational interval functions. *Nordisk Tidskrift for Informationsbehandling (BIT)*, 14:87–95, 1974.

[79] Leon Sterling and Ehud Shapiro. *The Art of Prolog.* MIT Press, Second edition, 1994.

[80] I. Sutherland. *Sketchpad: a Man-Machine Graphical Communication System.* PhD thesis, Dept. of Electrical Engineering, MIT, 1963.

[81] A. Torn and A. Zilinskas. *Global Optimization.* Springer-Verlag, Berlin Heidelberg New York, 1989. Lecture Notes in Computer Science.

[82] M. H. van Emden. The compatibility operator for real intervals, herbrand universes and finite domains. Technical report, Department of Computer Science, University of Victoria, 1994.

[83] M. H. van Emden. Canonical extensions as common basis for interval constraints and interval arithmetic. In *Proceedings of the Sixth French Conference on Logic and Constraint Programming*, Orléans, France, 1997.

[84] M. H. van Emden and H. M. Chen. Interval constraints for unconstrained global optimization. *International Journal of Reliable Computing (formly Interval Computations)*, 1995. Submitted.

[85] M.H. van Emden. The compatibility operator in numerical computation. In preparation.

[86] M.H. van Emden. An algorithm for interpreting Prolog programs. In *Implementations of Prolog*, pages 93–110. Ellis Horwood, 1984.

[87] M.H. van Emden. Algorithmic power from declarative use of redundant constraints. Technical report, University of Victoria, Canada, 1998.

[88] G. R. Walsh. *Methods of Optimization*. John Wiley and Sons Ltd., 1975.

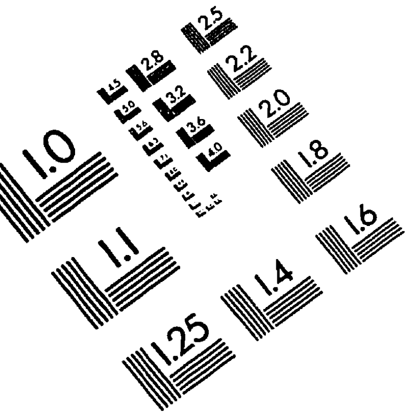
[89] D. Waltz. Understanding line drawings in scenes with shadows. In Patrick Henry Winston, editor, *The Psychology of Computer Vision*, pages 19–91. McGraw-Hill, 1975.

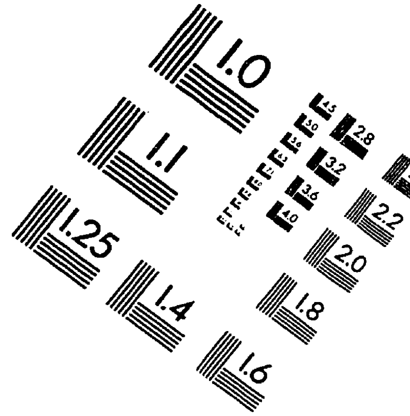
[90] David H. D. Warren, Luis M. Pereira, and Fernando Pereira. Prolog - the language and its implementation compared to lisp. In *Proc. Symp. on AI and Programming Languages*, 1977. SIGPLAN Notices vol. 12.

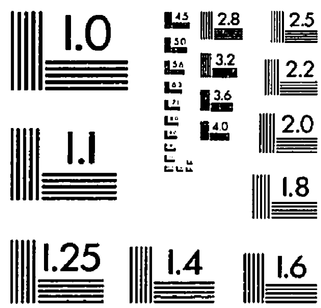
[91] D.H.D. Warren. Implementing prolog — compiling predicate logic programs. D.A.I. Research Report 39–40, Department of Artificial Intelligence, University of Edinburgh, 1977.

[92] D. Xue and Z. Dong. Improvement and testing of local and global optimization programs for NGFT project. Technical report, Department of Mechanical Engineering, University of Victoria, 1995.

IMAGE EVALUATION TEST TARGET (QA-3)

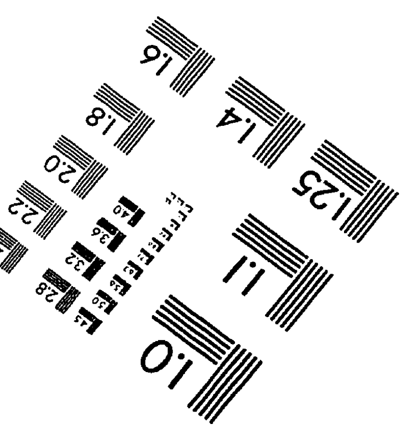


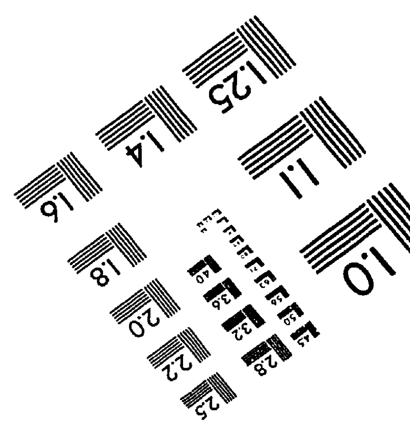




150mm

6″





APPLIED IMAGE, Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989