THE UNIVERSITY OF CALGARY

Developing a Knowledge Base to Incorporate Traditional Knowledge with Geographic Information Systems

by

Deena Clayton

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF SCIENCE

DEPARTMENT OF GEOGRAPHY

CALGARY, ALBERTA

MAY, 1998

0-612-34951-9

ABSTRACT

This thesis presents research using a case based reasoner (CBR) to develop a traditional knowledge base for the Gwich'in of the Northwest Territories. A prototype was developed, on a stand-alone machine and on the Internet, to assist Gwich'in resource managers with decision-making by incorporating the knowledge from interviews with Gwich'in elders into a CBR. Resource managers are able to search the prototype by entering a query, answering a series of interactive questions and receiving an appropriate response through the CBR. On a stand-alone machine, the CBR responses include scripts which launch images created in a geographic information system (GIS). This allows the user access to descriptive data from the interviews, as well as spatial data in the GIS images. Access on a stand-alone machine is limited, but integration of the knowledge base with GIS on the Internet offers potential solutions which should be further explored in the future.

ACKNOWLEDGEMENTS

This thesis could not have been written without the assistance and generosity of many individuals and institutions. I would like to thank the Northern Scientific Training Program, the Province of Alberta, the Margaret P. Hess Graduate Scholarship, the Industrial Research Assistance Program and the Faculty of Graduate Studies at the University of Calgary for their financial assistance.

I would also like to thank the Gwich'in Nation in the Northwest Territories for their support and encouragement throughout the research. Specifically, I would like to thank the Gwich'in Renewable Resource Board and their staff for their unfailing support. The Gwich'in and the Board provided the context and the substance for this research and it will hopefully further their interests in the future. I would also like to thank the staff of the Gwich'in Land and Water Board and the Gwich'in Interim Land Use Planning Board for their assistance while I was in Inuvik.

Appreciation must also be extended to my thesis committee members, Dr. Dianne Draper (University of Calgary) and Mr. Michael Robinson (Arctic Institute of North America), both of whom were in this from beginning to end. An even greater appreciation must be extended to my supervisor, Dr. Nigel Waters (University of Calgary) for his ability to launch headlong into little known territory.

On a personal note, I would like to thank my mother, father and family for supporting me every step of the way. Finally, I would like to thank my partner, Dave Watt, for proof-reading every sentence, listening to every presentation and following me to the icy climes of the Canadian North.

TABLE OF CONTENTS

APPROVAL PAGE . . . . -ii-

ABSTRACT . . . . -iii-

ACKNOWLEDGEMENTS . . . . -iv-

TABLE OF CONTENTS . . . . -v-

LIST OF FIGURES . . . . -viii-

LIST OF ABBREVIATIONS . . . . -ix-

CHAPTER 1: INTRODUCTION . . . . -1-

1.1: Project Objectives . . . . -3-

1.2: Thesis Organization . . . . -5-

CHAPTER 2: BACKGROUND TO RESOURCE MANAGEMENT IN THE GWICH'IN SETTLEMENT AREA . . . . -7-

2.1: The Gwich'in Renewable Resource Board's Traditional Knowledge Project . . . . -9-

2.2: Goals for Management Plans within the GSA . . . . -10-

CHAPTER 3: TEK: WHAT IT IS, WHAT IT MEANS AND WHY IT IS DIFFERENT FROM WESTERN KNOWLEDGE . . . . -11-

3.1: Practical Benefits from the Complement of Indigenous and Western Knowledge . . . . -13-

3.2: Differences Between Traditional and Western Knowledge . . . . -15-

3.2.1: Epistemological Differences . . . . -15-

3.3: Differences in Knowledge Organization . . . . -17-

3.3.1: Science .... -17-
3.3.2: Spiritual, Social and Economic .... -20-

CHAPTER 4: AN OVERVIEW OF CASE-BASED REASONING TOOLS .... -24-
4.1: Generic Issues .... -24-
4.1.1: What is a Case? .... -25-
4.1.2: Indexing and Case Retrieval .... -26-
4.1.2.1: Nearest neighbour .... -27-
4.1.2.2: Induction .... -27-
4.1.2.3: Template Retrieval .... -28-
4.1.3: Case Adaptation .... -28-
4.1.4: Retaining Cases .... -29-
4.2: Inference's CBR Software .... -30-
4.2.1: Knowledge Acquisition and Management .... -31-
4.2.1.1: CBR Express .... -31-
4.2.1.2: Generator .... -33-
4.2.2: Knowledge Searching .... -33-
4.2.2.1: CasePoint .... -33-
4.2.2.2: CasePoint WebServer .... -33-

CHAPTER 5: MODELLING TEK WITH CBR .... -35-
5.1: Manual Case Authoring with CBR Express .... -35-
5.1.1: Developing a Traditional Knowledge Model .... -36-
5.1.2: Case Base Design and Design Issues .... -39-
5.1.2.1: Identifying the Case .... -39-
5.1.3: Case Base Components .... -40-
5.1.3.1: Designing the Case Title .... -40-
5.1.3.2: Designing the Case Description .... -41-
5.1.4: Designing Case Questions .... -42-

5.1.4.1: Question Definition . . . . -42-

5.1.4.2: Context Questions and Confirmation Questions . . . . -43-

5.1.4.3: Question Structure . . . . -47-

5.1.4.4: Question Scoring . . . . -48-

5.1.5: Designing Actions . . . . -49-

5.2: An Evaluation of the Model Designed Using CBR Express . . . . -49-

5.3: Automatic Case Authoring: CBR Generator . . . . -52-

CHAPTER 6: USING CBR SOFTWARE FOR SEARCHING AND RULE GENERATION . . -58-

6.1: How a User Searches and Retrieves Information . . . . -58-

6.2: CasePoint: The User's Search Interface . . . . -58-

6.2.1: How Does a User Interact with CasePoint? . . . . -58-

6.3: Writing Rules to Automatically Answer Questions . . . . -61-

6.4: Actions, or Answers in CasePoint . . . . -63-

6.5: Using CasePoint WebServer to Develop Internet Access . . . . -64-

CHAPTER 7: PATHWAYS TO KNOWLEDGE INTEGRATION . . . . -71-

7.1: GIS Integration on a Stand-Alone Machine . . . . -71-

7.1.1: Operating a GIS as a Client Application . . . . -72-

7.1.2: Attaching GIS Files as 'Browsable' Files . . . . -74-

7.2: Internet Access to the CBR and the GIS . . . . -79-

7.3: Difficulties of Integrating GIS on the Internet and on a Stand-Alone Machine . . . . -82-

CHAPTER 8: CONCLUSIONS AND FUTURE IMPLICATIONS . . . . -84-

REFERENCES . . . . -89-

LIST OF FIGURES

FIGURE 1. Location of Gwich'in Settlement Area . . . . -8-

FIGURE 2. CBR Software Used For The Research . . . . -30-

FIGURE 3. Process of Case Base Development . . . . -32-

FIGURE 4. Decision Tree Highlighting a Partial Branch . . . . -37-

FIGURE 5. Examples of Case Descriptions . . . . -43-

FIGURE 6. Relationship of Questions and Cases . . . . -44-

FIGURE 7. Decision Tree Highlighting Context and Confirmation Questions . . . - 46-

FIGURE 8. Automatically Generated Case Example . . . . -54-

FIGURE 9. Manually Generated Case Example . . . . -55-

FIGURE 10. CasePoint Search Screen . . . . -60-

FIGURE 11. Examples of Rules . . . . -63-

FIGURE 12. CasePoint WebServer Solution Screen . . . . -67-

FIGURE 13. Sample CasePoint WebServer Interface . . . . -69-

FIGURE 14. Example of Answers as Text or Image . . . . -77-

FIGURE 15. GIS Image Attached to CBR Action . . . . -78-

LIST OF ABBREVIATIONS

Case Based Reasoner	CBR
Dynamic Data Exchange	DDE
Geographic Information System	GIS
Gwich'in Comprehensive Land Claim Agreement	GCLCA
Gwich'in Environmental Knowledge Project	GEKP
Gwich'in Renewable Resource Board	GRRB
Gwich'in Settlement Area	GSA
Integrated Resource Management	IRM
Northwest Territories	NWT
Traditional Environmental Knowledge	TEK

CHAPTER 1: INTRODUCTION

> [Indigenous knowledge is] knowledge that derives from, or is rooted in the traditional way of life of Aboriginal people. [Indigenous] knowledge is the accumulated knowledge and understanding of the human place in relation to the universe. This encompasses spiritual relationships, relationships with the natural environment and the use of natural resources, relationships between people, and is reflected in language, social organization, values, institutions and laws (cited in Roberts 1994:23).

The above definition of indigenous knowledge was put forth by a traditional knowledge working group in the Northwest Territories (NWT) and attempts to incorporate all the most salient features which define traditional knowledge. The recent Royal Commission on Aboriginal Peoples adopted another definition of traditional knowledge, coined by Berkes (1993) and interpreted: "as a cumulative body of knowledge and beliefs, handed down through generations by cultural transmission, about the relationship of living beings (including humans) with one another and their environment" (Berkes and Henley 1997:30). Although each of these definitions is generally similar, specific differences in semantics can lead to varied reading and understanding. No universally accepted definition of traditional knowledge exists, but this is by no means a reflection of disorder or disarray on the part of traditional knowledge experts. Indeed, no universal definition of 'western science' exists either, but western scientists are quite confident they have a vast and definitive knowledge of the world around us. Similarly, traditional knowledge experts have a vast and detailed knowledge of local ecology, even if that knowledge cannot be easily categorized or delimited.

As no universally accepted definition for traditional knowledge exists, no universally accepted label exists either. Traditional knowledge is variably named as some combination of the following: traditional, indigenous or aboriginal, environmental or ecological and knowledge or science. For instance the term science rather than knowledge is considered inherently more respectable by some, while the term traditional connotes an immutable and unchanging reflection of the past which indigenous does not. While excellent arguments are put forth for the use of one term over another, in a single work, one label must be chosen for consistency and coherency in the work. The phrase traditional environmental knowledge (TEK) is used throughout this research. Without entering into debates for and against the use of this term, TEK appears to be the most common phrase utilized by the Gwich'in in the NWT. As this work was conducted in conjunction with the Gwich'in, the phrase was adopted.

The widespread interest, for non-Aboriginals, in TEK arguably is a reflection of the growing recognition of the deplorable state of the world's environment. Western scientists may be coming to terms with the fact that their particular model of knowledge is fallible and as such, are beginning to accept the use of other paradigms. Thus interest is growing in inter-disciplinary studies and in 'new' sciences, like ecology.

Key papers and reports in the non-Aboriginal world have noted the strengths to be found in traditional knowledge and have pushed for its recognition as an alternative science. The report of the World Commission on Environment and Development, *Our Common Future*, was tabled in 1987 and brought widespread recognition to the state of the environment, as well as identifying the key roles indigenous people have played in resource management around the world (World Commission on Environment and Development 1987:12).

In the NWT and in Canada, the Commission's work was prefaced by the seminal *Report of the Mackenzie Valley Pipeline Inquiry*, presented by Justice Thomas Berger in 1977. The report underlined the moral, ethical and environmental imperative in including Aboriginal people in the NWT in decisions facing their land and lifestyle. The map of the NWT has changed significantly since 1977, as has the policy of the government. Within the last six years TEK has been adopted by the NWT government as official policy and it must be incorporated into government decision and actions where appropriate, including everything from the setting of hunting quotas to environmental impact assessments (Laghi 1997). Comprehensive land claim agreements have also opened up the possibility for the joint aboriginal—non-aboriginal management of resources.

For Aboriginals the knowledge of their skill and understanding of processes of resource management have not simply appeared with recognition in federal and international reports. Aboriginal people have been, for the most part, calmly and patiently stating their case since Europeans first arrived in North America. As early as the nineteenth century, Chief Seattle predicted the destruction which would be wrought by western resource managers, or mis-managers. The call for the inclusion of indigenous systems of knowledge to resource management practices in North America is not new, but non-aboriginals have only recently taken heed.

Despite the theoretical recognition of the value of TEK in resource management, the inclusion of TEK, in practice, has met with formidable barriers. The barriers are partly a reflection of differential power structures and partly a comment on the difficulty of understanding the world view of another culture and the intransigence of western scientists to even attempt to understand. Numerous suggestions have been put forth and actions taken— some successful, others not—to break down the barriers between western science and traditional knowledge. In a report prepared by the Inuit Circumpolar Conference on the inclusion of TEK in the Arctic Environmental Protection Strategy (1993), the need for indigenous people to have the ability to establish their own capacity to create and utilize information systems is identified as a priority.

That priority has become the challenge of this research. The goal of the current research is to develop a traditional knowledge decision support system using case-based reasoning (CBR) tools to allow Gwich'in resource managers direct access to the vast body of recorded, Gwich'in TEK. The knowledge system should be simple to create, utilize and maintain, and should operate in conjunction with tools currently used by Gwich'in resource managers, namely GIS and Internet applications.

1.1 PROJECT OBJECTIVES

More specifically, the goal of the project is to develop a knowledge support system, based on indigenous knowledge, to assist in the management of resources and the maintenance of TEK. Although TEK is a necessary component in many northern environmental studies, few models exist to show resource managers how to integrate TEK into their resource management practices and decisions (personal communication A. Veitch

1997). This is perhaps best exemplified in a political policy paper written by Howard and Widdowson (1996) in which the authors are confused and disgruntled by requirements to integrate TEK into their work because they view TEK as 'unscientific'. This research seeks to address this confusion by focussing on access, rather than philosophy.

In developing a knowledge support system, the goal in this project is not to integrate the two systems of knowledge conceptually, but simply to make them equally available to a resource manager using the current management tools. The first step toward using TEK is to record it, the second is to make it accessible and the third and final step is dependent upon the managers and decision makers to incorporate it. The goal of this project is *accessibility*.

The objectives are to make a computer-based system that is easy to develop, use and maintain. Case-based reasoning (CBR) software was chosen as the tool to achieve these objectives and to manage the TEK base. The system is also required to maintain the context and philosophy of TEK as far as is possible. The simple act of recording TEK in interviews already removes it from its context. Ideally, TEK must be learned and understood on the land and in the local language. In reality however, recording traditional knowledge through interviews has become a necessary practice to attempt to maintain TEK (Seminar on the Documentation and Application of Indigenous Knowledge 1996). With that in mind, the goal of this research is to maintain the context of TEK as much as is possible with interview material previously recorded in English or transcribed into English. The aim is to allow an end-user to query the knowledge support system using free form text and to return to the end-user a response that will be modelled after the words of the traditional knowledge expert. A representative subset of information from the GEKP was used to develop a prototype knowledge support system.

The final objective is to make the system accessible within a framework of computer tools currently used by resource managers. In the case of the Gwich'in Boards and Renewable Resource Councils, their current goal is to make all of their data accessible across the organization via the Internet. Although the details of an integrated data structure have yet to be established, GIS is used to represent a significant portion of their data and will thus be an integral part of the organization-wide database. The objectives of this research are to make the TEK base functional first over the Internet, and second, in conjunction with GIS databases.

1.2: THESIS ORGANIZATION

Following the introduction and project objectives, which have just been outlined, the thesis provides a brief overview of the Gwich'in in the NWT and the efforts of the Gwich'in Renewable Resource Board to record TEK and to include TEK in their overall management strategy.

Chapter 3 presents a comparison of the general similarities and differences of TEK and western science and points to the important contribution both can make to resource management. A general overview is presented in this section because of the need to recognize the theoretical and practical barriers to, and implications of, integrating two systems of knowledge for a common purpose.

This is succeeded by an overview of CBR tools in Chapter 4. The overview explains what CBR tools are and why they were chosen for this research. Inference Corporation provided the software for the research and the suite of Inference tools used is outlined. The tools are broadly divided between knowledge acquisition, or authoring, tools and knowledge management, or searching, tools. The following sections are thus divided between the authoring of the knowledge system and the searching of the knowledge system.

Chapter 5 describes the process of modelling TEK with CBR. The authoring process began with the manual development of a case base and the method of knowledge modelling and system development is explained. Further, the potential limits of the system are identified. This is followed by a description of the use of an automatic case base developer and the problems inherent in automatic case generation.

After a brief introduction to searching tools at the beginning of Chapter 6, the subsequent sections describe the use of those tools. First there is an evaluation of searching tools on a stand-alone machine together with a description of the development of rules to assist that search. This is then followed by an evaluation of Internet-based searching tools and their potential benefits.

Chapter 7, entitled "Pathways to Knowledge Integration" addresses the possibilities for the integration of GIS and the case base. The limited success with GIS integration on a stand-alone machine is described, followed by a look forward to the potential benefits the Internet may provide for integration. Finally, a brief look at how the limits in the current data set limited GIS integration.

The research is concluded in Chapter 8 and areas of future research potential are addressed.

CHAPTER 2: BACKGROUND TO RESOURCE MANAGEMENT IN THE GWICH'IN SETTLEMENT AREA

In 1992, the Gwich'in of the NWT signed the Gwich'in Comprehensive Land Claim Agreement (GCLCA) with the Canadian government, establishing the Gwich'in Settlement Area (GSA) (Gwich'in Tribal Council 1992). Figure 1 shows the extent of the GSA and its location within the NWT. The claim includes the following goals:

> 1.1.3 To recognize and encourage the Gwich'in way of life which is based on the cultural and economic relationship between the Gwich'in and the land;
> 1.1.6 To provide the Gwich'in with wildlife harvesting rights and the right to participate in decision making concerning wildlife harvesting and management;
> 1.1.7 To provide the Gwich'in the right to participate in decision making concerning the use, management and conservation of land, water and resources (Gwich'in Tribal Council 1992).

These goals indicate that traditional and local knowledge must be part of the resource management framework.

Under the auspices of the agreement, and the currently tabled Mackenzie Valley Resource Management Act, resource planning and management is divided between essentially three different organizations: co-management boards, Gwich'in organizations and government agencies (Gwich'in Land and Water Board 1996:3). The co-management boards and their responsibilities are as follows:

- the Gwich'in Renewable Resource Board (GRRB) is the main instrument of renewable resource management in the settlement area;
- the Gwich'in Land and Water Board (GLWB) regulates land and water use throughout the settlement area;
- the Gwich'in Land Use Planning Board (GLUPB) is charged with developing a land use plan for the settlement area;
- the Mackenzie Valley Environmental Impact Review Board (MVEIRB) reviews all proposed development in the Mackenzie Valley;
- the Surface Rights Board (SRB) maintains jurisdiction over surface entry, and;
- the Mackenzie Valley Land and Water Board regulates land and water use throughout the Mackenzie Valley (Gwich'in Land and Water Board 1996:4).

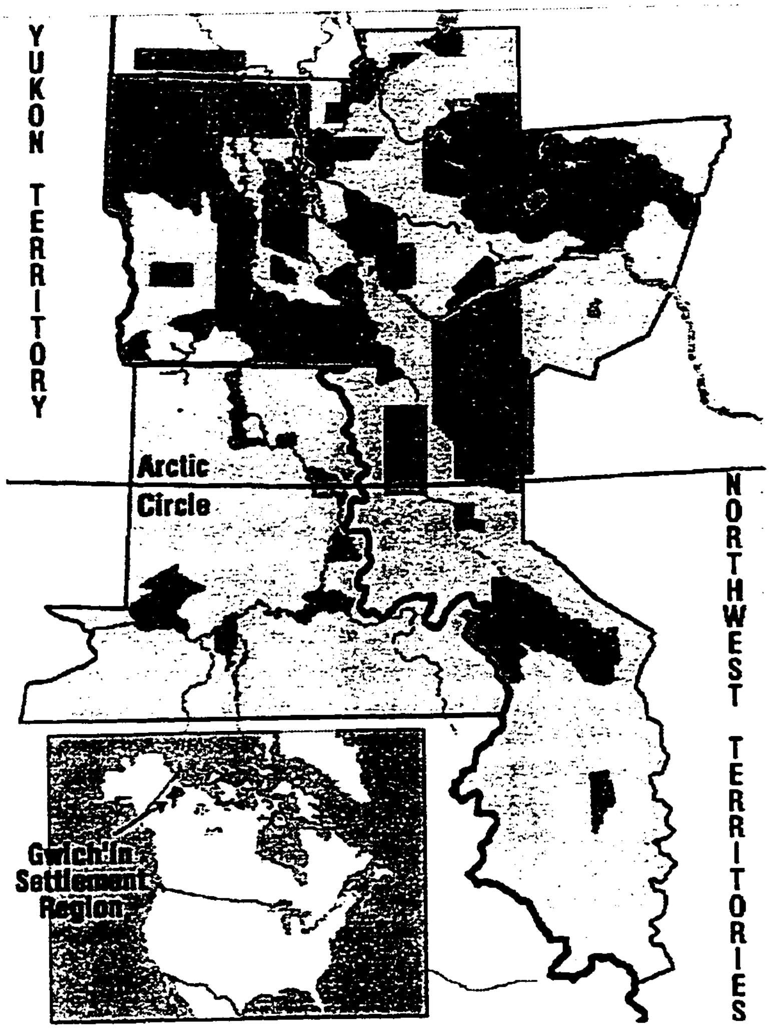

Figure 1. The Gwich'in Settlement Area and its location within North America (Gwich'in Graphics Ltd., http://209.68.1.63/inuviktv/atgraphics/gsrmap.html)

The boards are directed jointly by Gwich'in and local representatives and territorial and federal government appointees. Government appointees represent agencies which continue to operate and conduct research within the GSA and the Mackenzie Valley. At the local level, public resource and development responsibilities are undertaken by Renewable Resource Councils while private Gwich'in lands are administered by the Gwich'in Tribal Council, Gwich'in designated organizations and the Gwich'in Land Administration. The Gwich'in Social and Cultural Institute is mandated to document, promote and preserve the practice of Gwich'in culture, language, traditional knowledge and values (Gwich'in Land and Water Board 1996). Together these organizations have collected a significant amount of resource and management related data for the region and have directed efforts toward meeting claim objectives in their management strategies.

2.1: THE GWICH'IN RENEWABLE RESOURCE BOARD'S TRADITIONAL KNOWLEDGE PROJECT

The GRRB is striving to achieve the claim objectives of respecting the spiritual values of the land, training younger generations in traditional ways and passing knowledge on to future generations by, among other things, making TEK a key feature in their strategic plan (Gwich'in Renewable Resource Board 1998). In June 1995, the GRRB initiated the Gwich'in Environmental Knowledge Project (GEKP). While the overall goal of the project was to meet the objectives of the land claim agreement, the immediate goal was to document the traditional Gwich'in environmental knowledge on twenty wildlife species (Gwich'in Elders 1997). To meet this goal, assistants from each Gwich'in community in the settlement area (Fort McPherson, Tsiigehtchic, Aklavik and Inuvik) conducted interviews with local elders. A total of 198 interviews were conducted with 46 elders, fishers, hunters and trappers. Each interview focussed on a single species and interviewees, therefore, often contributed a number of individual interviews to the collection. The GEKP has also included Gwich'in knowledge which was recorded in the past by the Hudson Bay Company, the Committee on Aboriginal People's Entitlement and various anthropologists and researchers. All pertinent materials have been transcribed and entered into a database and a book, based on the interviews, has been published (see Gwich'in Elders 1997).

2.2: Goals for Management Plans within the GSA

One of the immediate goals within the GSA is to develop an integrated resource management (IRM) system so the boards, resource councils, Gwich'in organizations and government agencies can share their information and knowledge to make decisions which are consistent across the GSA and the Mackenzie Valley (Gwich'in Tribal Council, 1992; Gwich'in Land and Water Board 1996). Purely from a data management perspective, integration is not a simple task - data has been collected and catalogued by a variety of people in a variety of organizations for a variety of purposes. How can a resource manager at one of the boards access relevant information collected by other researchers, sitting on other boards and agencies, to respond to their question quickly and efficiently?

This research sought to integrate data from the GEKP into a knowledge base which would complement current Gwich'in efforts to develop a shared database. The proposed research was approved by the University of Calgary and the GRRB in March 1997. Prototype development took place between June and December 1997 in Inuvik. Software for prototype development was provided by Inference Corporation (http://www.inference.com) in Novato, California. By the end of December 1997, a stand-alone and Internet capable prototype version, modelling two of the twenty GEKP species, was complete. The system has been evaluated and recommendations based on its strengths and weaknesses are summarized in this thesis, along with a complete description of prototype development.

CHAPTER 3: TEK: WHAT IT IS, WHAT IT MEANS AND WHY IT IS DIFFERENT FROM WESTERN KNOWLEDGE

Both traditional and western knowledge, while quite different from each other, have a valuable contribution to make in our overall understanding of the northern environment. Each approach to knowing the environment has limitations, but together the two systems of knowledge can complement each other, and create a better, more informed framework within which to make resource management decisions. Although this project does not seek to integrate the two systems of knowledge, it does seek to place recorded TEK alongside western knowledge. This section highlights some of the differences between traditional and western knowledge and underlines the importance of including traditional knowledge in northern resource management.

First, the terminology used throughout this research must be defined and a disclaimer must be made. Indigenous people are the native inhabitants of a land. Although the Scottish are indigenous to Scotland, and the Egyptians are indigenous to Egypt, the term indigenous is rarely applied to these people. Here, the term 'indigenous' applies to people who are native to lands which were colonized, primarily by Europeans, and who now live as the minority population within those lands. Traditional, indigenous environmental knowledge refers to knowledge of local ecosystems that is accumulated through a subsistence lifestyle. Knowledge is transmitted orally and includes relationships within the environment, religion, society and economy. For this research, indigenous systems of knowing will be described primarily in a North American context with some contributions from Australia.

Similar to the term 'indigenous', the term 'western' does not simply describe people that live west of Greenwich, but people of a shared culture. Western culture is predominantly Judaeo-Christian, and has a value system rooted in European history and thought. The scientific method defines a western 'way of knowing' the environment.

A disclaimer must also be made because both terms, western and indigenous, refer loosely to two distinct, culturally constructed epistemologies. The terms only loosely apply because no single statement can be used to identify characteristics that apply to each and every member of a cultural community. The paradigms used to understand the world are very complex, and may be highly differentiated even within a culture. The definitions used to describe the two frameworks of knowledge in this paper are, admittedly, broad generalizations. One cannot assume that there is no diversity of thought within western and indigenous frameworks, nor can one assume that individuals subscribe completely and unfailingly to one of these frameworks. Nevertheless, members of each culture do share a common understanding. A generalized understanding of each culture provides a framework for understanding different approaches to resource management in the north.

In order to begin discussing traditional environmental knowledge, a myth must be dispelled. A common myth, within western society, is that traditional knowledge has disappeared. Native people, according to the myth, have been assimilated into the western mainstream culture and their subsistence lifestyle is an artifact of the distant past. Admittedly, traditional life has been disturbed by western influences in the north. Many native people do rely on the monetary economy, but a reliance on subsistence based activities continues. Wenzel and Usher estimate that: "...smaller native communities depend on wild foods for upward of 50 per cent of caloric needs and often as much as 80 per cent of their protein, whereas larger communities (more than 1000 people) obtain at least 30 per cent of their caloric intake and half their protein from wild food" (cited in Osherenko 1988:46). Obviously the reliance on the subsistence harvest is still very strong. In order to be successful hunters and trappers, natives must maintain a rigorous knowledge of the environment. Anthropologists confirm that native traditions are adapting to change and continuing, rather than being assimilated or disappearing (Osherenko 1988:5). Intimacy with the local environment remains an integral feature of northern native communities today.

3.1: Practical Benefits from the Complement of Indigenous and Western Knowledge

The complement of two systems of knowledge will help to overcome the practical limits within each system. Non-native management has often been ineffective in the north because decisions are based on incomplete data sets (Riewe and Gamble 1988; Sallenave 1994; Freeman 1985; Feit 1988). The basic data required to make informed decisions about wildlife resources, according to western scientific models, is absent:

> "Scientists believe they can control wildlife population behaviour and attain management objectives if they can gather the correct data, such as information on species distribution, habitat requirements, population size, structure and discreteness, reproductive potential, and natality and mortality rates including those occasioned by domestic and commercial harvesting. Unfortunately, this comprehensive data base is rarely attained in the south, and has never been obtained for any northern species (Davis et al 1980; Dickinson and Herman 1979; Freeman 1985; cited in Riewe and Gamble 1988).

Baseline data is limited because research in the north is exceedingly expensive, and is constrained seasonally and historically. Historically, western observations have been recorded for, at most, one hundred and fifty years in the Canadian north. Those observations are often limited to short intense surveys of single species in the summer months. In the absence of adequate data, managers must extrapolate from old data gathered in other areas. In his review of environmental impact assessments of the Hudson Bay oil industry, Nakashima found:

> "From the Canadian Occidental Petroleum Ltd.'s environmental overview (EAG 1984) and Canterra's "Oil Spill Sensitivity Assessment of Hudson Bay and James Bay Shorelines" (EAG 1985), the poverty of environmental baseline data for the region is starkly apparent. Neither study includes original research...the industry's environmental statements serve more to highlight the absence of critical baseline data than to provide an adequate basis for effective EIA" (1990:3).

In the absence of baseline knowledge, how can managers continue to make decisions?

Traditional knowledge of the environment provides an alternative form of baseline data. Hunters and trappers are continually making observations, then sharing their observations with the community. Gunn, et al note (1988:25): "Hunters make observations in the course of their travelling and hunting and the intensity of their vigilance can be awe-inspiring. New observations are also acquired through discussion with other hunters...The organization of the observations is akin to a mental encyclopaedia used for successful hunting and travelling". Traditional knowledge is passed through generations orally. Oral histories are often embellished with every telling, but highly technical and historically accurate information remains a crucial part of the telling (Cruikshank 1981:72). Data is collected without the exorbitant costs of western research and is collected throughout the year, both winter and summer. Studies have illustrated the depth of native knowledge and its ability to complement scientific knowledge. Often, these studies also show that in specific areas of northern research, native knowledge is much more thorough than western.

The historical perspective that is absent from western knowledge of the northern environment is an important part of native knowledge. Western knowledge in the north dates back, as has been noted, perhaps one hundred and fifty years, whereas native knowledge of the north dates back thousands of years. Traditional knowledge of significant environmental events can confirm western scientific hypotheses about natural history events. Cruikshank (1981) studied oral histories in the Yukon Territory and found that they supported glaciological and dendroclimatological research. Scientific research confirms oral traditions which date from 1400 AD. Oral traditions describe Yakutat and Icy Bay on the Alaskan coast as filled by ice, while the coastline to the east and west of these bays was free of ice. "This is confirmed by archaeological and geological evidence, radiocarbon dates and tree-ring dates. Lobes of the Malaspina Glacier once filled Icy Bay and Yakutat Bay, then began to retreat by 1400AD to positions near or behind present limits" (Cruikshank 1981:78). The investigation and integration of oral histories can provide western science with a much greater historical understanding of natural history processes.

In certain cases native knowledge of animal behaviour, and the concomitant harvest decisions, are more informed than that of wildlife managers. The Peary caribou herd on Ellesmere Island was flourishing in 1953, when an Inuit village was established there by the

Department of Indian and Northern Affairs (Freeman 1985: 266). Non-Inuit authorities restricted the harvest to include only large male caribou, and restricted the harvest to a few animals from each herd. The Inuit protested the restrictions because they felt that such a management strategy would endanger the herd. Observations of caribou behaviour had made clear to Inuit hunters that each small group in the caribou herd represents a social group. Older/larger males provide a leadership role for the group that is critical for the group's survival (Freeman 1985: 267). The Inuit predictions were correct. Under the state management strategy, Riewe notes that even the minuscule harvest had deleterious effects on the herd. By 1968, the Peary herd was virtually eliminated in that region (cited in Freeman 1985: 267). Similar cases are presented for Bowhead whales, arctic char, musk oxen (Freeman 1985), beaver (Feit 1988), eider ducks (Nakashima 1990) and many other species. Thus, drawing on both forms of knowledge when decisions are made presents a better possibility for successful resource management.

Traditional data sets can also be supplemented by western knowledge. First, traditional knowledge is limited geographically. Although traditional people have an excellent knowledge of the immediate environment in which they hunt and trap, regional knowledge is limited. Hunters may have intimate knowledge of where species are, and at which times of the year, within traditional territories, but little knowledge of the complete migration pattern of the species. Second, western science also has an understanding of the microbiological and cellular structures of organisms. This type of knowledge is absent from traditional knowledge (personal communication O. Kawagley).

3.2: Differences Between Traditional and Western Knowledge

3.2.1: Epistemological Differences

Psychologists suggest that human cognition is marked by two distinct ways of thinking; analytical and intuitive (Murdoch, 1988). Every individual has the ability to operate in both spheres, but within cultures one sphere tends to be more highly valued than the other. Western culture primarily operates in, and overwhelmingly values, analytical thought:

> "People of Western culture, Bastick contends, become increasingly immersed through socialization and formal schooling grounded in literacy and numeracy, in an analytical mode of thinking which is conscious, cerebral, literate, and linear...Indeed, those socialized and trained in analytical thinking tend to rationalize the products of intuitive thought, such as hunches and gut reactions. Insights, ideas or decisions arrived at intuitively are verbalized with explanations" (in Wolfe 1992:8).

Logical, rational explanations take precedence over unverbalized feelings. Hunches are not considered thoughts until they are verbalized and rationalized. True 'thinking' on the other hand, is understood by western culture as a conscious, verbal and rational process, that relies on words to analyse and define thoughts.

Conversely, indigenous culture organizes understanding with an intuitive, rather than an analytical, approach. 'Thinking' does not use symbols, or words, to define the understanding. Brown (1973:59) notes that: "...in the Indians' cognitive orientations, meanings generally are intuitively sensed and not secondarily interpreted through analysis; there tends to be a unity between form and idea or content. Here the 'symbol' is, in a sense, that to which it refers". The reliance of indigenous culture on intuitive thinking is confirmed by Strachan (1988), in her study of Inuit environmental perception. Strachan concludes that intuition defines the predominant cognitive style among the Inuit. She also notes that verbal explanations or rationalizations are not stressed and are often virtually impossible to elicit. Although Murdoch (1986) uses the term visual spatial, rather than intuitive, to describe Cree cognition, his conclusions are similar. Thus, in indigenous culture, 'gut reactions' and hunches are directly acted upon, without being rationalized, verbalized or analysed.

Western culture, in its valuation of the two cognitive modes, subordinates those cultures that rely on intuitive expression. Western thought places cultural evolution on a hierarchy from intuitive to analytic, or from less advanced to more advanced. At the bottom of the ladder is the hunter-gatherer, and at the top is the modern-industrial society. Hunter-gatherer societies are understood as wholly reliant on intuition; as societies that do not actively shape their environment, but merely respond to environmental changes. An analogy is often made between hunter-gatherers and children. They are children that have not yet developed the ability to analyse critically, nor to think about, a situation. Modern-industrial societies define the pinnacle of the hierarchy. Rather than respond to environmental change, the modern-industrial individual can analyse the components of the environment and orchestrate the change. By western standards, this is the top of the evolutionary societal ladder.

The two modes of cognition must be placed on a continuum rather than a hierarchy. To describe hunter-gatherer societies as being like mere children, reacting to the needs of their environment, is grossly simplistic. Likewise, assuming that modern-industrial societies have developed an infallible system for shaping the environment is simply incorrect. The two modes of cognition should be considered complementary, rather than exclusionary. The greatest scientists and philosophers in western culture are those that can identify with both the analytical and the intuitive. Music and mathematics go hand in hand. Fritjof Capra (1982:38) writes: "The rational and the intuitive are complementary modes of functioning of the human mind. Rational thinking is linear, focussed, and analytic...Intuitive knowledge, on the other hand, is based on a direct, nonintellectual experience of reality arising in an expanded state of awareness". As the complement between mathematics and music creates an exceptional individual, so too the complement of intuition and analysis creates a broader society and better strategies for resource management.

3.3: DIFFERENCES IN KNOWLEDGE ORGANIZATION

3.3.1: Science

Differences in dominant cognitive modes between western and indigenous cultures have also created or, been created by, different systems of knowledge organization.

The organizational framework that dominates western thought stems from the scientific revolution of the 17th century. This revolution, spearheaded by philosophers and scientists such as Descartes, Bacon and Newton, introduced a mechanistic, reductionist approach to science. The world, they said, operates like a machine. Each component in the machine can be isolated, understood and reassembled. By understanding how each component operates, one can understand how the collection of components operate (Capra 1982). The middle-Christian concept of the earth as a nurturing mother, was discarded and replaced with the western scientific concept of the earth as a machine (Capra 1982, Watson and Chambers 1989). The shift of the scientific paradigm, from an organic to a mechanistic view, quickly expanded western understanding of the world and continues to dominate the western scientific method.

Western knowledge is structured according to the rigours of the scientific method. Laws are objectively formulated to reflect a scientific truth. All scientific truths are understood to have a validity and objectivity which is separate from the influence of the researcher, and from humans in general. Good scientific research must be supported by data that is quantifiably sound and mathematically proven. In reference to the management of wildlife resources, Johnson (1992:3) states: "Western scientists gather quantitative data to build mathematical models of population dynamics. The models are used to calculate sustainable yields of a resource. The yields are then recommended for implementation to decision-makers as wildlife harvest regulations". The reliance of wildlife managers on quantitative data is noted throughout the literature (Freeman 1985; Gunn, et al 1988; Feit 1988). Wolfe, et al (1992) note that western scientific knowledge accumulates data quickly and selectively. Natural phenomena are isolated and studied intensively for a short period of time (Usher 1993:117). Data are tested for their validity and if they are accepted general laws and theories are formulated to describe the phenomena.

Indigenous culture organizes knowledge using a holistic rather than reductionist approach, and a systemic rather than a mechanistic approach. (Kawagley 1995, Wolfe et al 1992; Bird Rose 1988, Hughes 1983). Brown (1973:59) notes that: "Unlike the conceptual categories of the Western man, American Indian traditions generally do not fragmentize experience into mutually exclusive kinds of dichotomies, but tend rather to stress modes of interrelatedness across categories of meaning, never losing sight of an ultimate wholeness". Rather than understand the world as a machine with individual parts, indigenous systems of knowledge organization focus on the relationship between the parts. The interaction of the parts, not the parts themselves, is crucial in understanding the whole.

Rather than viewing the world as a machine, indigenous cultures view the world as a system. Freeman (1992:9) notes that traditional ecological knowledge is: "...directed toward gaining a useful understanding of how ecological systems generally work, to how many of the key components of the total ecosystem interrelate, and how predictive outcomes in respect to matters of practical concern can best be effected". Holistic, system-based approaches define the native view of 'seeing' the world.

Indigenous knowledge of the environment is based on accumulated experience of the entire environment, not on individual, intensively studied features. This experience is very subjective and always relates the individual and the community to the environment. Trends and patterns are observed, and decisions are made using qualitative data. "The Dene harvester is more concerned with conditions in general...or in trends...than he is in precise numbers and averages" (Johnson 1992:3). Knowledge is accumulated slowly and inclusively into a natural history framework (Usher 1993:117). The framework is shared by the community of decision-makers.

The scientific method is so pivotal in western culture that it defines the only valued approach to western science. However, the reductionist method does have weaknesses, as well as strengths. Descartes believed that mathematical certainty reflects scientific certainty. All science is certain, evident knowledge that can be explained mathematically (Capra 1982). The certainty that is applied to the scientific method hides its limitations. "The Cartesian approach has been very successful, especially in biology, but it has also limited the directions of scientific research. The problem is that scientists, encouraged by their success in treating living organisms as machines, tend to believe that they are nothing but machines" (Capra 1982:62).

The limits of a reductionist approach are becoming apparent in many disciplines. Models of environmental processes, no matter how mathematically sound, are often unable to predict change accurately. The discovery of quantum physics in the early part of this century forced physicists to recognize the limits to the truth of their science. This discovery shook the foundations of conventional reductionist science. Capra states:

> "Twentieth century physics has shown us very forcefully that there is no absolute truth in science, that all our concepts and theories are limited and approximate. The Cartesian belief in scientific truth is still widespread today and is reflected in the scientism that has become typical of our Western culture. Many people in our society, scientists as well as non-scientists, are convinced that the scientific method is the only valid way of understanding the universe...The acceptance of the Cartesian view as absolute truth and of Descartes' method as the only valid way to knowledge has played an important role in bringing about our current cultural imbalance" (1982:58).

3.3.2: Spiritual, Social and Economic

The differences between the reductionist and holistic approaches are apparent not only in organizing knowledge about the natural environment, but also in discerning relationships between the natural world and the spiritual, economic and social worlds. From a western perspective, neither the spiritual nor the social, although perhaps the economic, fit into a discussion of environmental knowledge. However, spiritual, economic and social relations are essential components of traditional ecological knowledge and must be included in any consideration of the environment.

In indigenous culture all the universe is a living, balanced, spiritual creature. Everything in the universe is imbued with spirit: "there are primary elemental spirits such as Thunder, Wind, Earth and Sun, as well as a spirit for every species of animal, bird, and insect on the earth. Every living creature is a tangible expression of the spirit for that species" (Young, Ingram and Swartz, 1989:25). Navajo artist Carl Gorman writes:

> "It has been said by some researchers into Navajo religion, that we have no Supreme God, because He is not named. This is not so. The Supreme Being is not named because He is unknowable. He is simply the Unknown Power. We worship him through His creation...We believe that this great unknown power is everything in His creation. The various forms of creation have some of this spirit within them...As every form has some of the intelligent spirit of the Creator, we cannot but reverence all parts of the creation" (in Brown, 1973:58).

God did not simply create the earth as in western religious traditions, God is the earth itself.

In this light, natural phenomena are explained by, and understood in terms of, spiritual phenomena. Freeman (1985:25) states: "hunters observations are loosely organized in an informal and flexible system which may equally include a spiritual or mythical interpretation". The goal of indigenous resource management therefore, is to maintain a balance because all things on the earth are to be treated equally as manifestations of the Creator (Deloria 1973). The reference to the spiritual is an essential part of the framework of indigenous environmental knowledge.

In western thought, the relationship between the spiritual and the natural is very tenuous and is usually removed from explanations of environmental phenomena. God is not everywhere in everything. 'God' is not of this earth but is considered to be above and beyond the mechanics of nature, supernatural. Humans are the only creatures made in the image of God. The only spirit that is widely recognized in western culture, resides in humans. Even this concept is reduced in western thought. The spirit is associated with the mind, not with the mechanical body. Rather than view all things as equal manifestations of the Creator, western religions subordinate nature. A spiritual hierarchy places humans on a divine plane, closer to an unknown world above, than to the earth. Humans dominate nature.

The spiritual relationship with the earth is reflected in each culture's relationship of society to nature. Western culture separates society from nature, whereas native culture links society as an integral part of nature. The divergence of thought is highlighted with an example of the different concepts used by each culture to describe the natural environment: wild and quiet.

From an indigenous perspective wild lands are those that have been degraded by humans. Wild lands are those that have not been managed sustainably and have had their internal balance upset. Quiet or tame lands on the other hand, are those that include humans living sustainably (Bird Rose 1988). Wild lands are places where humans have had too great an impact, tame lands are those where humans have had minimal impact. In western thought the terms are reversed. Wild lands are those on which humans have had minimal impact, tame lands are those that are controlled by humans. From a western perspective, true wilderness does not include humans. This conceptual opposition is an excellent reflection of how western culture views itself outside of nature and native culture views itself as an integral part of nature.

The economics of resource management strategies also reflect a different relationship for each culture. The driving force of western-based resource management is neo-classical economics. Resources are managed according to the capitalistic ideals of individual profit maximization. In terms of the management of wildlife resources, quotas are set at the maximum allowable harvest that will maintain population stability. Surplus accumulation is encouraged. Freeman (1986:86) notes the preference for the capitalistic ideal in resource management strategies: "Those who maximize their return - and can continue to do so - are held to be behaving rationally, and those who don't should, in the interest of the system, be encouraged to change their strategies to conform to the preferred ideal".

Native management is very different from the capitalist perspective. Resources are managed not for individual profit maximization, but for the reproduction of the social unit (Freeman 1986). Surplus, should it occur, is re-distributed within the society: "...the individual so favoured (with surplus) converts their economic capital to social capital by distributing the surplus production. Indeed widespread sharing and community feasting is a characteristic feature of all hunting and fishing societies" (Freeman 1986:88). The notion that resources will always be used for individual gain must be dispelled from the establishment of resource management quotas in the north.

In order to approach resource management from an ecosystem perspective, management institutions must include socio-economic processes. Management should not be focussed toward manipulating the physical and biological components of the environment only. Management should be focussed toward understanding the human components of the environment. The exclusion of socio-economic factors from management results in ineffective resource policies. Walters (cited in McDonald 1988:66) states:

> "...problems in the resource sciences result from the interaction between ecological and socio-economic systems, and the tendency of scientists and managers to reduce complex ecological-economic relationships into simplified associations...By concentrating on biological and technical harvesting issues and ignoring the socio-economic implications of management decisions, human resource-use activities are never completely controlled in the conventional, scientific management process".

Human use must be considered an integral part of any management strategy. The acceptance and incorporation of traditional knowledge will help to create more inclusive management frameworks.

The review of the strengths and weaknesses of TEK and western science inevitably lead one to appreciate the benefits of integrating both in a complementary management framework. In this research, the first step toward achieving the more philosophical goals of integration is to have access to each on an equal footing. At this point, the introduction of CBR as a tool for achieving equal access is necessary.

CHAPTER 4: AN OVERVIEW OF CASE-BASED REASONING TOOLS

4.1: GENERIC ISSUES

CBR tools are a type of expert system that operate from the notion that new problems are solved based on previous experience. CBR tools utilize the specific knowledge of previously experienced, concrete problem situations to solve new problems based on similar past cases (Aamodt and Plaza 1994). The CBR functions to retrieve cases that are similar to the current problem and attempts to reuse the case in a new solution. When a new solution is found the manager of the case-base has the opportunity to revise and to retain the new solution. Case bases are never considered to be complete. They can operate with only a partial case base, and are always expected to add new cases. Cases can be added according to questions asked and solutions given, or with new information. Attached to the case solution can be full text descriptions, images, mapped images and/or audio and video links. This section reviews case identification, indexing and retrieval methods, adaptation and reuse.

First, it is important to note how CBR differs from other database and information retrieval systems in its adaptation, information retrieval and learning processes. The ability to adapt old cases to fit new situations is dependent upon reasoning processes found in the CBR, though not in conventional database and information retrieval systems. Even in CBR without the adaptation facility, retrieving information is quite different from other systems. CBR is more inter-active:

> "Database systems and IR [information retrieval] systems leave the problem of how to formulate the right query largely to the user. In CBR systems, the system itself is often designed to start from an input description using features that are quite different from those included in the case memory, and to determine appropriate retrieval cues" (Leake 1996:14).

The user can interactively arrive at an answer without having to correctly join all keywords with Boolean operators. This separates CBR from standard search engines in which it is incumbent upon the user to develop their search criteria. Finally, CBR has the ability to store and learn from past consultations which is absent in most database and information retrieval systems.

4.1.1: What is a Case?

Each case in a case base represents a single piece of knowledge or a series of events which lead to a particular outcome (Watson 1995). For example, a doctor will record a patient's symptoms, make a diagnosis, prescribe a course of action and evaluate the treatment's success. This is a case. This example describes three typical features in case identification:

- ▸the problem that describes the state of the world when the case occurred, or the symptoms;
- ▸the solution which states the derived solution to that problem, or the diagnosis and prescribed action; and/or
- ▸the outcome which describes the state of the world after the case occurred, or the final evaluation of the patient's health (Watson and Marir 1994).

In describing a case, the term 'problem solving' has a very broad application. In a strictly mechanical sense a case is a data object with associated indices (Inference Corporation 1995a). In a practical sense, a case is anything you want a user to be able to find. The only true criterion for identifying a case is its functionality: if a case is worth keeping, it must simply have a lesson to teach to the user (Kolodner 1995).

CBR is considered a great step forward in the artificial intelligence community because of its ability to avoid the 'knowledge elicitation bottleneck'. In conventional rule based expert systems, an inordinate amount of time was required to elicit and formalize an expert's knowledge into a series of rules, thus creating a 'bottleneck' (Waters 1989; Shaw and Gaines 1986). CBR avoids the difficulty of knowledge elicitation by dispensing with the need to formalize and structure knowledge into rules.

However, defining case boundaries is one of the most difficult parts of CBR and this may be considered the 'case identification bottleneck'. Determining appropriate case features is still a necessary and very difficult part of designing a CBR system. The process mirrors the difficult knowledge engineering phase in a rule base system (Aha and Breslow 1997; Barletta 1991). In order to make information into 'CBR-ready' cases, it must be thoroughly analysed and reformulated. Cases, in effect, must be engineered. In evaluating the success of CBR, researchers note:

> "In summary, we have consistently found that creating a CBR system in a new domain requires a full-scale knowledge acquisition effort of the usual sort: designing an appropriate organization, representation, and set of reasoning methods; putting the knowledge into the designed framework; and debugging and refining the knowledge base in close interaction with domain experts." (Mark, Simoudis, and Hinkle 1996:292).

Despite the difficulty of case representation, case base systems can be designed and developed more quickly than rule base systems, are better in explaining their reasoning and are easier to maintain (Watson 1995).

4.1.2: Indexing and Case Retrieval

Case retrieval is facilitated through indexing significant case features (Watson and Marir 1994). Indexes, like keywords, point to the purpose of a case and should be chosen according to the circumstances under which the case is likely to be used. Indices should meet four criteria. They should:

1. be predictive;
2. address the purposes the case will be used for;
3. be abstract enough to allow for widening the future use of the case base; and
4. be concrete enough to be recognized in the future (Watson and Marir 1994:331).

Both automated and manual indexing methods are available. Automatic case indexing may be seen to reduce the time noted above for case design, but it has not been successful across varying domains. Kolodner (1993) believes that people, be they case base authors or domain experts, tend to do better at choosing indices than algorithms do, and therefore for practical applications indices should be chosen by hand. Similar to case identification, indexing poses a challenge for the development of case bases.

Case retrieval is guided by one or a combination of strategies including nearest neighbour approaches, induction and template retrieval (Aamodt and Plaza 1994; Barletta 1991).

4.1.2.1: Nearest neighbour

Nearest neighbour is an approach to case retrieval which is guided by matching features of the user's query with weighted indices in the case base memory. In the medical case example described above, symptoms would have weights attached according to their usefulness in determining a diagnosis. The features of a user's query which match those indices are weighted accordingly. The case displaying the greatest weighted sum would be the most likely candidate for retrieval. For example, if a patient entered a doctor's office complaining of severe stomach cramps, vomiting and diarrhea, giving equal weight to each of the symptoms would probably produce quite a long list of equally possible ailments. If, upon further investigation, the doctor discovered the patient had eaten a tuna sandwich that had been unrefrigerated for a week, the doctor would give greater weight to this new piece of information and the diagnosis would undoubtedly be food poisoning. The case base can be developed in the same way to allow some pieces of information to have equal weight, while others can be weighted to confirm a case absolutely.

Nearest neighbour approaches are excellent if the retrieval goal is not well defined or if few cases are available (Barletta 1991). The difficulty in designing a system based on nearest neighbour retrieval is the determination of feature weights (Watson and Marir 1994). In the simplest approach to nearest neighbour matching all features would have equal weight.

4.1.2.2: Induction

Inductive algorithms may be used to determine which features do the best job in discriminating between cases. Which symptoms, for example, are most useful in determining whether a patient is suffering from illness A or illness B. Inductive approaches require a large number of cases to generate accurately the discriminating features between cases, making it a poor approach in certain domains (Watson and Marir 1994). Inductive approaches also require that the case outcome be well-defined, for example where the outcome is the identification of a single illness (Barletta 1991). Again, this makes induction a poor approach in some domains.

4.1.2.3: Template Retrieval

Template retrieval returns all cases that fit within certain parameters. For example, a medical case base may be divided into groups which separate viral illnesses from traumatic injuries. If a patient's symptoms include vomiting and a high fever, only cases that fit those parameters should be included and those cases are likely to be found in the viral illness group. Cases that definitely do not fit those parameters, a broken bone for instance, should be immediately excluded from the search. This technique is often used before other techniques to limit the search space to a relevant section of the case base.

4.1.3: Case Adaptation

Adaptation is the 'reasoning' part of CBR and it is also the weakest link in most systems. The goal of case-retrieval is to return the most similar past case to the user. The goal of case adaptation is to adapt the most similar past case to meet all the needs of the new situation (Barletta 1991). Case adaptation is based on a set of domain rules or on a domain model. Consider for example, a situation where a user wanted to cook a spicy dish with shrimp and coconut milk and the case base contained a recipe for spicy chicken with coconut milk. The system would retrieve the chicken recipe, and based on certain formulae or rules integral to the system, would suggest substituting shrimp for chicken to meet the user's needs, thus adapting a stored case to meet new requirements.

In practice however, successful case adaptations are difficult to achieve. Watson notes:

> This cycle (retrieve, reuse, revise, retain) currently rarely occurs without human intervention. For example, many CBR tools act primarily as case retrieval and reuse systems. Case revision often being undertaken by managers of the case base (Watson 1995:2).

Case adaptation is domain specific. Adaptive rules are difficult to generalize across domains and as experiences with rule-based expert systems have shown, codifying knowledge into rules and debugging them is difficult and time consuming. It must be accepted that at present CBR systems retrieve-evaluate-adapt cycle has been replaced by retrieve and propose with the adaptation and evaluation being done by the user (Watson 1995, Kolodner 1995, Leake 1996, Mark, Simoudis and Hinkle 1996).

Although adaptation is a weak link in CBR theory, it does not mean that it is a weak link in CBR applications. One of CBR's strengths is that it does include human collaboration in case adaptation and decision support. Riesbeck summarizes: "Adaptation techniques are hard to generalize, hard to implement, and quick to break. Furthermore, adaptation is often unnecessary. The originally retrieved case is often as useful to a human as any half-baked adaptation of it" (Riesbeck 1996:388). CBR does not create an expert; it merely tries to mimic one. The human decision maker will always be the best judge of a case's merits and its current applicability.

4.1.4: Retaining Cases

Retention, or the storing of new cases, is the learning part of CBR. Each time a user interacts with the CBR, the process of interaction or the search consultation, can be saved. If the search is saved as a successful solution, the CBR now 'knows' given the same input features to retrieve the successful solution. "When the CBR process is successful, the resulting solution is stored for future reuse, avoiding the need to rederive it from scratch" (Leake 1996:9) . If the search is unsuccessful, this too can be recorded and provide a signal for the system manager that this is an unknown area or that the case base is missing important information. Failed searches may be stored for future analysis when new information becomes available, they may provide a warning about possible future failures that should be avoided or they may prompt the revision of indexing criteria to retrieve a better case in the future (Leake 1996).

4.2: INFERENCE'S CBR SOFTWARE

Inference Corporation has launched an extremely successful set of CBR software packages (Watson 1995). Although designed specifically with the support-desk market in mind, Inference has been a leader in the CBR field from its beginnings. The suite of Inference products is broadly divided between two functions: knowledge acquisition and management and knowledge searching, as shown in Figure 2. This means that case authoring and problem solving, or case searching, are functionally separate.

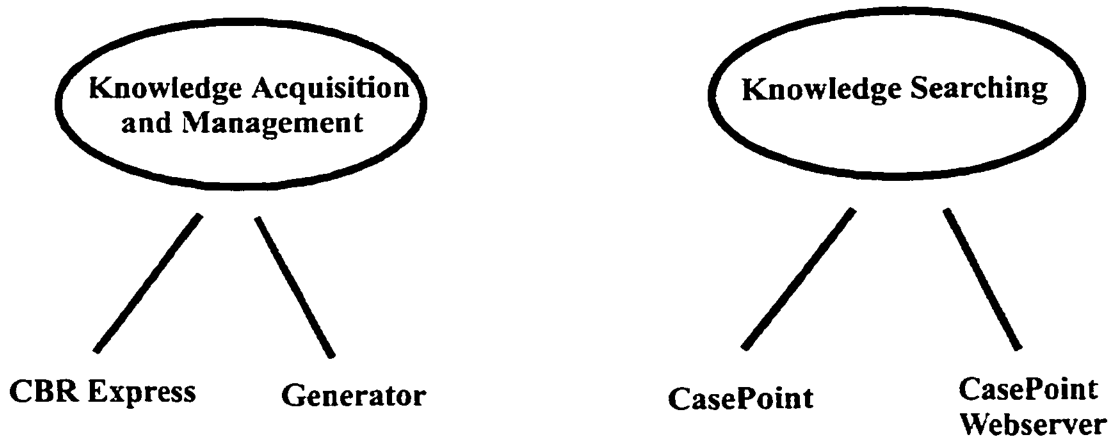

Figure 2. The products used for this research are broadly divided between knowledge acquisition tools and knowledge searching tools.

This research involved two knowledge management and acquisition tools; CBR Express and Generator, and two knowledge searching tools; CasePoint and CasePoint WebServer. After a brief introduction to each of the tools, subsequent sections describe the use of these tools specifically for this research, first by examining case authoring techniques, then by examining case searching abilities. An overview of the process is provided graphically in Figure 3.

4.2.1: Knowledge Acquisition and Management

4.2.1.1: CBR Express

CBR Express is a manual case authoring tool which uses the nearest neighbour approach to retrieve cases. Case development, creation and editing is done through a graphical user interface (GUI). All programming elements operate behind the interface which allows people without programming experience to develop case bases quickly in a syntax free environment (Watson and Adbullah cited in Watson and Marir 1994). Case development involves assigning a title, description, questions and action to each case. An end user can enter a free form text query, rather than being led through a decision tree style question and answer session, to match initially against case titles and descriptions (Watson and Marir 1994). Solutions are ranked, on a percent-matched basis, and the user is offered both the solutions and a set of questions. Questions are associated with each case by the case author to help narrow the search and provide a more accurate solution. Mark, Simoudis and Hinkle describe the process:

> "Associated with every attribute in CBR Express is a question that is used during the problem-solving session to obtain the attribute's value during problem solving. The retrieval algorithm of CBR Express, a variant of nearest neighbour, always asks the first unanswered question of the case that is considered most relevant at each iteration, given the available information. Since the nearest neighbour algorithm is executed after each question is answered, a question from a different case may be asked at each iteration." (1996:289).

If the search is not successful an entire transcript of the consultation can be saved as an unresolved case.

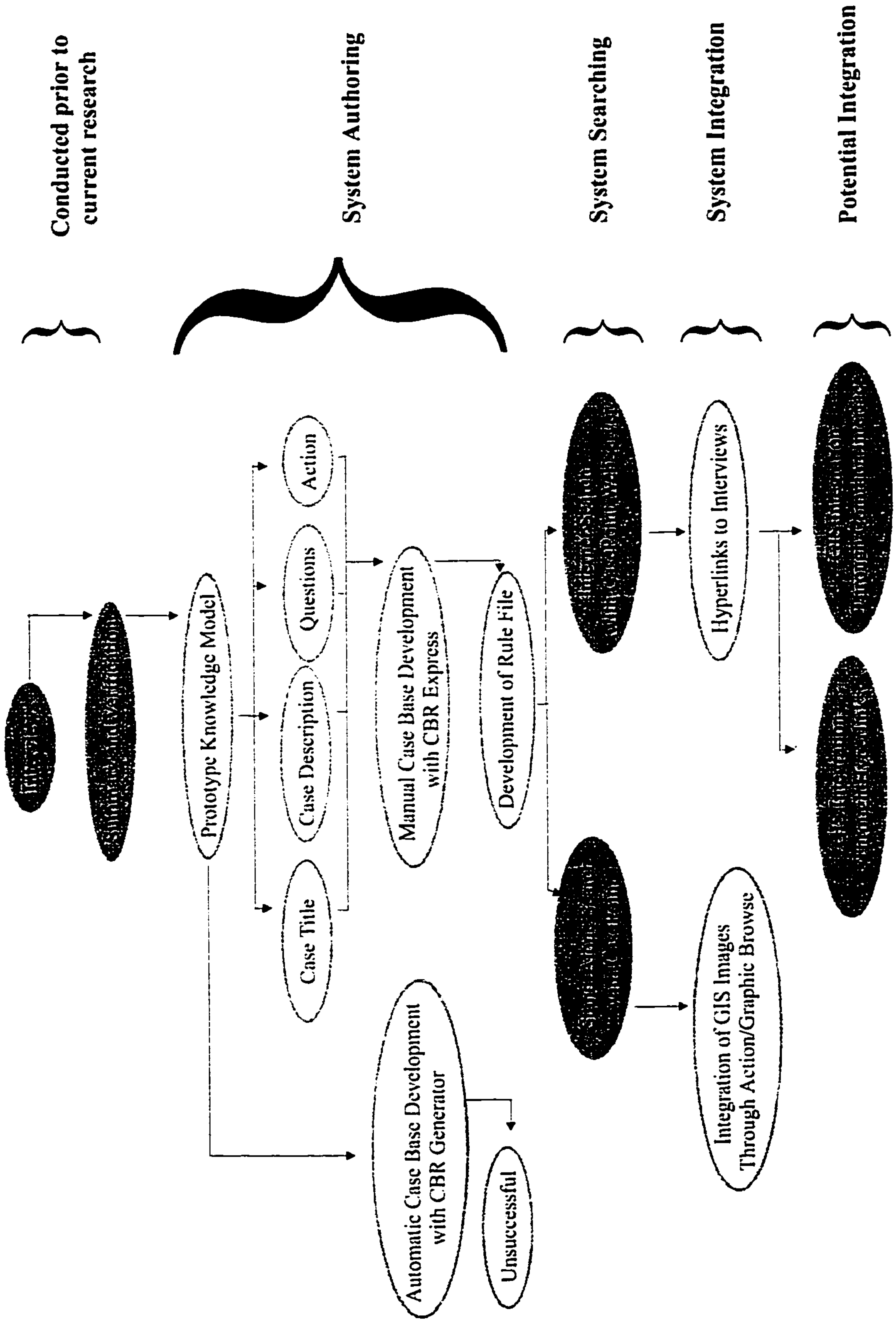

Figure 3. Process of Case Base Development, Presentation and Integration

4.2.1.2: Generator

Inference's Generator software provides an automatic indexing facility for case base development, circumventing the need for lengthy time investments in case base development. From a directory of text files, Generator is designed to create a case base index automatically (Inference Corporation 1995b). The indexing facilities are designed to operate in any domain thus providing a case base author with 'universal expertise' (ibid).

4.2.2: Knowledge Searching

4.2.2.1: CasePoint

CasePoint is the run time version of CBR Express. CasePoint provides a case base searching environment for users but has no authoring capability and thus is purely a knowledge search tool. CasePoint may be embedded in other software applications through its use of Windows Dynamic Data Exchange (DDE) protocols (Watson 1995). DDE protocols allow companies to embed CasePoint in applications, databases or GIS for example, that are already employed by their company. CasePoint also supports the use of a rule file which identifies keywords in a user's search and automatically answers questions based on heuristics. Automatic answers may be overridden by the user.

4.2.2.2: CasePoint WebServer

CasePoint WebServer is similar to CasePoint but rather than operate on a stand-alone machine, as does CasePoint, CasePoint WebServer operates on networked machines: WebServer is Internet ready. WebServer is an interactive Web application that makes search and retrieval facilities available using any World Wide Web (WWW) browser. The WebServer interface may be fully customized through a text based resource file. This allows WebServer to share a common 'look and feel' with a company's other web pages and applications. The application may be secured to limit access or prevent unauthorized access. Because the WebServer, like CasePoint, provides knowledge searching tools only, content is centralized and managed elsewhere through knowledge acquisition and management tools. Only the case base managers can alter the case base. As an Internet application, WebServer offers a range of possibilities for data integration in this application, as described in Chapter 6.5.

CHAPTER 5: MODELLING TEK WITH CBR

5.1: MANUAL CASE AUTHORING WITH CBR EXPRESS

The authoring process began with a thorough review of the GEKP final report. The report, rather than the interviews, provided the basis for the study. The report was chosen because the document has been peer-reviewed, verified and approved by the elders who contributed to it. After the original interviews were completed, a summary was made and presented to the elders over the course of a two day workshop. The workshop provided a forum to verify, clarify and expand upon the information from the original interviews. To work directly from the interviews for this research would have negated the verification process. The final summary report incorporated new information and clarified and edited existing information and therefore formed the basis of this research.

Since the GEKP was conducted on a species by species basis, the development of a prototype case base system followed a similar approach. The case base is loosely structured on a species by species basis. The research focussed on two species for the development of the prototype: grizzly bear and caribou. These species were chosen because each has a vastly different amount of information associated with it. Caribou is a major source of food and in the past was a major source of clothes and materials. Owing to its significance within the Gwich'in culture, information about caribou is vast and detailed both for its use today and in the past. The grizzly bear on the other hand has never been a major source of food, nor skins for the Gwich'in. The animal features prominently in their traditions and current hunting and trapping lifestyles, but it was never sought out as the caribou was and is. Relative to the text describing caribou, the information about the grizzly is much shorter. In the final book twenty-two pages are devoted to a description of caribou, while seven pages are devoted to the grizzly. This is not necessarily a direct reflection on the knowledge of the elders, but simply reflects what has been recorded.

Therefore, the development of a prototype, to represent both grizzly and caribou would test the ability of the system to search and retrieve from domains with a vastly different number of cases. A user should still be able to find exactly what they are looking for whether the associated number of cases is large or small.

5.1.1: Developing a Traditional Knowledge Model

The first task was to develop a case-base to operate on a stand-alone machine. In order to represent fully the knowledge collected for the GEKP, its breadth and depth needed to be considered even though not all the knowledge would be included in the prototype. The breadth of knowledge in this case was the number of species which the GEKP covers: twenty species in three broad categories — fish, birds and animals. This roughly divides the breadth of the case-base, potentially, into three equal sections covering five fish species, seven bird species and nine animal species. The depth of knowledge in the GEKP is the movement between the most general information to the most specific. For example, the mere identification of the animal is the most general level of information and when grizzly cubs are born is the most specific information.

Developing a clear picture of the extent of knowledge embodied in the GEKP is crucial in developing a case base which seeks to represent that knowledge. In the end, the case base must adequately represent the entire body of knowledge: it must successfully lead a user to the information they require. Second, the case base must be flexible enough to include new and different information in order to be a valuable tool for the future. For example, will the case base have the flexibility to include traditional knowledge not focussed on a species by species basis but focussed perhaps on geographical areas or ecozones. These questions are further developed and evaluated at the end of this section.

The model which was developed draws on the natural conceptual divisions, across the breadth of the knowledge, for fish, birds and animals. This provides the first step in a hierarchical model, as represented in Figure 4. Paths, much like the branches in a decision

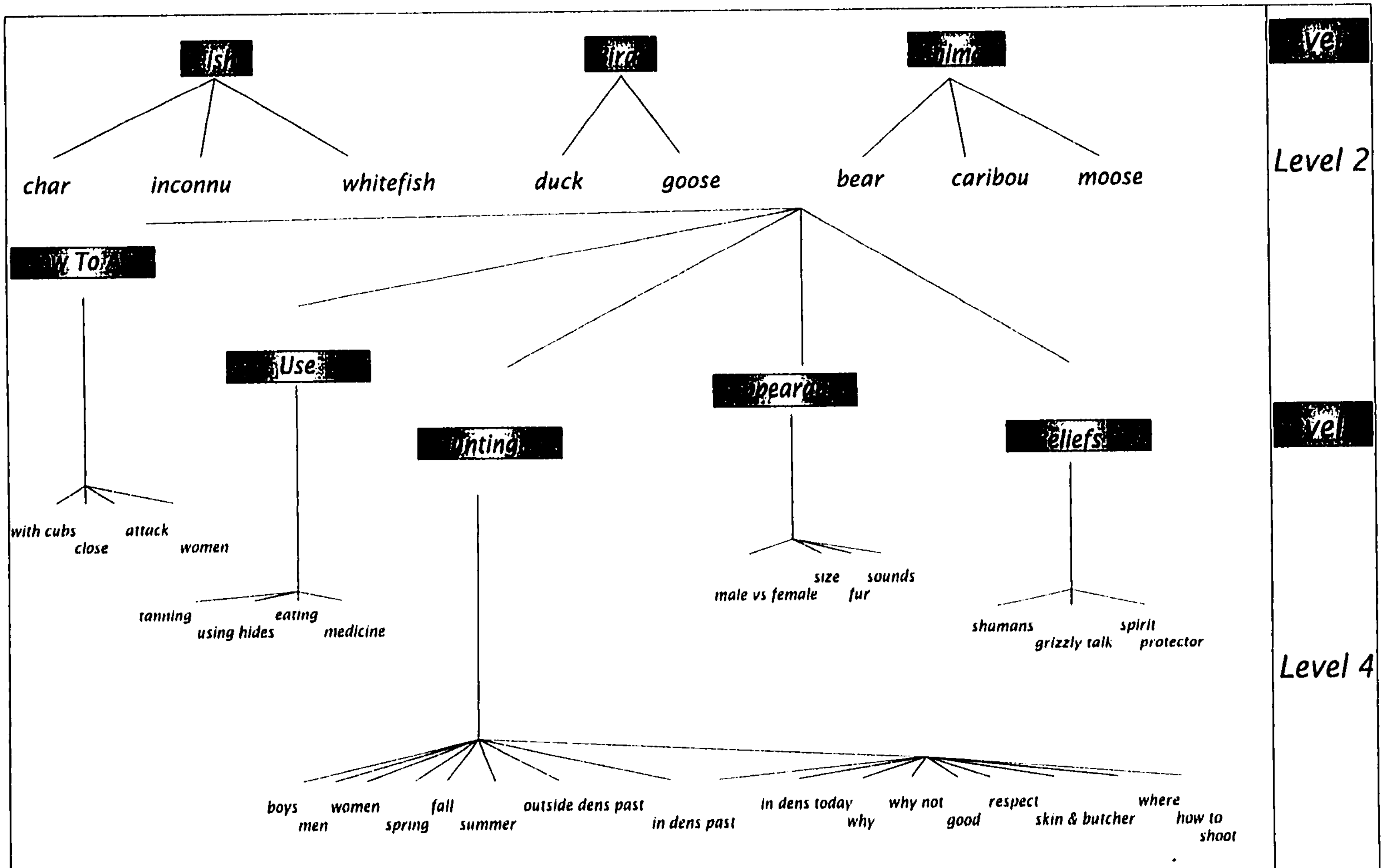

Figure 4. Decision Tree Highlighting a Partial Branch for Grizzly Bear

tree, were then created to move a user from that very general starting point, of birds, fish and animals, to the very specific answers that a user would be searching for in the end. One section of the tree is illustrated in Figure 4.

For the model to operate successfully as a case base, a certain balance had to be achieved. The paths leading from each of the most general information – fish, birds, animals – to the most specific information at the bottom of the path must be close to the same length. Each path requires a similar number of steps, or nodes, to get to the leaf. The section of the tree illustrated in Figure 4, is parallelled across the number of species represented. If information is weighted equally and one path leads a user to the end in five steps and one leads a user to the end in twenty steps, a user will rarely get past the first five steps. For example, if a patient describes five symptoms to a doctor and each of those symptoms are matching along a five step path, then the case's match value is 100. If the same five symptoms are simply the first matching steps along a path that has twenty steps, then based only on those five symptoms, the case will only score a match value of 25. Based solely on those five symptoms, however, each case could be equally correct and the first one will give a false match. The user will not get beyond the first five symptoms. Therefore each path in the prototype contains the same number of steps.

A model with, for the most part, four steps, or levels, in the path from the most general information (animals, birds, fish) to the most specific was developed. This is illustrated in Figure 4. Each level in the hierarchy, and each node within each level, to the extent possible, is paralleled across the twenty species. This gives the model consistency, coherence and elegance. For example, at the same level in their respective hierarchies is a node which asks about Gwich'in hunting of caribou and grizzly, the habitat of caribou and grizzly and the feeding habits of caribou and grizzly. This type of hierarchy was found to be successful for both species in the prototype despite the differing amounts of information associated with each. In the final level of the hierarchy, just above the leaf, the branches are quite different across the species. The leaf is the outcome, the final, most specific piece of information, and is therefore very specific to each domain or species. Although caribou and grizzly share an equal number of steps in the hierarchy, there are many more leaves on level four at the end of the caribou branches. The expediency of this model for developing a case base will be discussed in the next section which describes how the case base was designed, why it was designed that way and how it works.

5.1.2: Case Base Design and Design Issues

5.1.2.1: Identifying the Case

In many case base systems, case identification is a relatively simple matter. In a medical case base for example, a case is one patient's file. In a legal case base, one case is based on a single judgement. However, in a traditional knowledge case base, case identification is a more difficult issue. There are no cases which define a single event in history or an associated chain of events which describe a problem and a solution as there are with a patient's illness, diagnosis and recovery. What boundaries can be used to define unique and meaningful cases from a traditional knowledge text?

There are natural conceptual divisions which help to divide descriptive textual knowledge into cases. The division is similar to the structure in a book. A book is divided into chapters, which are divided into subheadings and so forth down to individual sentences. Sentences on their own convey only part of an idea. Out of context a sentence is often meaningless. Paragraphs usually describe a complete thought or idea. Their goal is to illustrate fully one piece of information. Similarly, in a semi-structured interview, each question often opens for the elucidation of a more or less singular, fully illustrated idea or story. This is how a case was defined, as a fully illustrative piece of knowledge. Each case should provide a unique and meaningful piece of information within its own context.

The separation of the text into cases distills the information into loosely connected parts. This may be doing an injustice to the knowledge because TEK represents, after all, a holistic way of thinking. Certainly the best way to understand fully the knowledge is to be out on the land experiencing it with the elders. Writing answers to English questions, or transcribing answers into English, as was done in the GEKP, already removes the knowledge from its context and therefore does somewhat of an injustice to that knowledge. The division of knowledge into cases does not dilute the knowledge any further. The knowledge is broken into cases, but the cases are presented within a context of similar cases and are linked to the original interviews, to the words of the elder. The presentation and links are described in subsequent sections. The design of the case base does not shatter the framework of the knowledge and it offers significantly better ways to have the knowledge accessed and used.

5.1.3: Case Base Components

Each node in the model is a question in the case base. Each case is designed with four parts:

- a title which is simply designed to identify the case to the user;
- a description which is comprised of keywords in the case;
- questions which are designed to guide the user to a specific answer; and,
- an action or answer which provides the solution to the user's query.

External to the case design, but critical to case searching, is a series of rules which help to guide the search efficiently. Rules are discussed in the section on information searching.

5.1.3.1: Designing the Case Title

Associated with each case in the case base is a case title. The case title is limited to a single line of text and must communicate to the user the features which make the case important (Inference Corporation 1995a). Each title, as well as each case, must be unique. In the title the user should be able to identify their problem and recognize a possible solution. Because the title field is limited it must describe the nature of the case very succinctly.

Sample titles from the prototype case base include:

- The Gwich'in name for the grizzly bear...
- Hunting grizzlies in summer...

- Preferred summer habitat for caribou....
- Preferred winter habitat for caribou...

The 'problem' described in these case titles uses 'problem' in a very broad sense, as is acceptable in case base design (Aamodt and Plaza 1994). The only 'problem' is that the user does not know the answer to a question. The user is searching for information. The title gives a quick glimpse at how the case may answer their question or find a 'solution' to their 'problem'.

5.1.3.2: Designing the Case Description

The case description is a critical paragraph of text because it, along with the case title, provides the basis for the first search. The description, therefore, must fully describe the characteristics of the case including all keywords and pertinent information.

A matching algorithm scores the user's query based on how much of the search description is contained within the case description and the title. Because the algorithm matches both the title and the description there is no need for redundant information. If a keyword is in the title, it is useless to include it in the description as well. The algorithm ignores articles, prepositions and some verbs and it removes most punctuation and suffixes (Inference Corporation 1995b). With the remaining words, the algorithm performs a trigram match (Watson 1995). For example, the word caribou would be scored based on matches with CAR, ARI, RIB, etc. This makes the matching process extremely tolerant of subtle differences in words including spelling mistakes.

The case description must be written clearly and concisely. If the case description is employing obscure and little used terms, it is unlikely a user will launch a query that can match those terms (Inference Corporation 1995b). Descriptions must be concise because superfluous words in the description can lead to falsely high scoring matches: the more words in a description, the more likely a match. For this same reason, descriptions must be of approximately the same length in every case. If each word in the description is given equal weight, a case with a very long description will have a much greater chance of scoring highly than one with a very short description. The case description relies heavily on the original identification of the case. Each case identified must be describing a singular and unique idea or event for the description to be successful.

Descriptions were designed for the traditional knowledge prototype with these tools in mind. Rather than use jargon and formal language, the words in the case description are the words the elders themselves used in interviews. Each description is a complete paragraph of text that can be understood regardless of context. The description includes all of the necessary words and information which help to guide a user from the most general information, at the top of the tree, to the most specific information, contained in the leaves of the tree (Figure 4). Some concepts left out of the descriptions are addressed by a series of rules which are described in a later section. Figure 5 shows examples of case descriptions.

5.1.4: Designing Case Questions

5.1.4.1: Question Definition

Questions serve two functions in a case base: they confirm characteristics already mentioned in the description and they guide the user into a dialogue to elicit more exact information. Because questions are essential guides in a user's consultation with the system, questions should separate clearly and define the entire population of cases. Questions can be thought of as coordinate axes which define the domain space (Inference Corporation 1995a). Cases are separated along the axes into various groups. This relationship is illustrated in Figure 6. Each group of cases shares a similar theme which separates it from other groups in the domain. Within each group are smaller clusters of cases. Each cluster shares a common theme until the individual cases can be identified. As one moves down the decision tree (Figure 4) each node signifies a common theme shared by all cases below that node. The final node, or grouping of cases, leads one to the leaf, or the individual case.

TITLE
The Gwich'in name for the grizzly bear.
DESCRIPTION
The name for the grizzly bear in the Gwich'in language, both male and female. The name, or what they call the grizzly bear, is Shih.

TITLE
Hunting grizzlies in spring or late winter
DESCRIPTION
Gwich'in hunt grizzly bears at this time of the year if the bear is close to camp and is a threat. The meat does not taste good.

TITLE
Number of Porcupine caribou in the area, past and present.
DESCRIPTION
More caribou today than in previous years. At the beginning of the century, herds stayed on the Alaskan side of the Richardson Mountains.

Figure 5. Examples of Case Descriptions

5.1.4.2: Context Questions and Confirmation Questions

Ideally the first question presented to the user will evenly divide the case base. Each question answered should definitely separate one group of cases from another. If a question does not help to distinguish at least a few of the cases from the general population, it is useless. To design a question space that distinguishes groups of related cases, as well as individual cases, questions must first identify the context, then confirm the case. Context questions distinguish groups of related cases from one another while confirmation questions separate individual cases from a small group of cases (Inference Corporation 1995a).

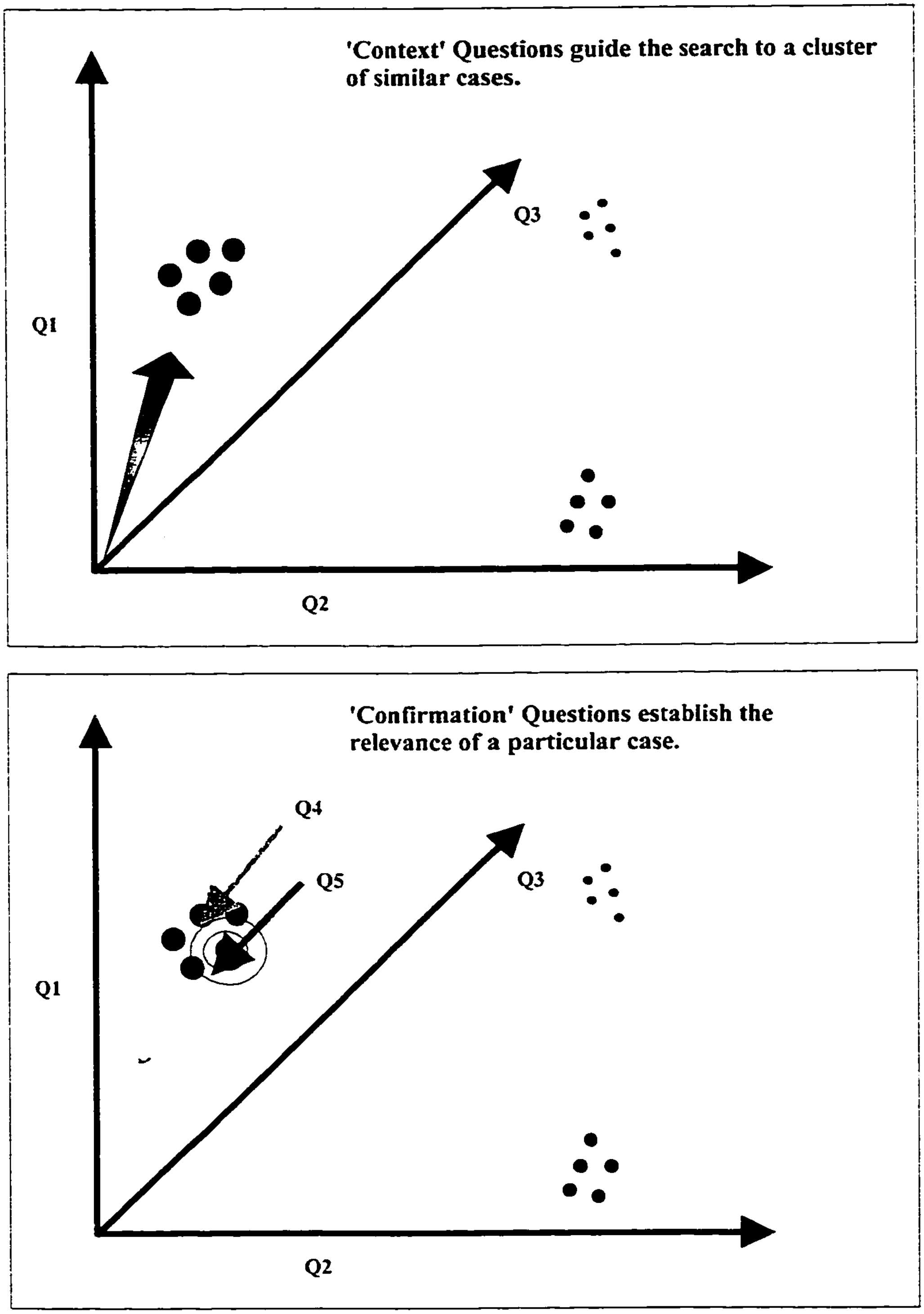

Figure 6. Relationship of 5 distinct questions and cases in the domain space (adapted from Inference Corporation 1995a).

Context questions define the framework of the inquiry. The case description may help to identify the framework for the search, but explicit context questions sharpen the focus of the search. A context question usually appears in many of the cases in the case base (Inference Corporation 1995a). In the prototype case base, of the four questions associated with each case, the first two are always context questions. The first two questions are represented as the first two nodes on the decision tree, as illustrated in Figure 7. The first question effectively divides the entire domain space into thirds by focussing the search on one of the three general topics, birds, animals or fish. The second question further establishes the context by identifying the species in question. Both context questions may be automatically answered by a series of IF...THEN rules which are triggered by the initial search. The use of rules is discussed in a subsequent section.

More precise confirmation questions are used to distinguish one neighbouring case from another. Confirmation questions shift the focus of the search from differentiating between broad groups of cases to differentiating among similar cases. The final two questions in each of the cases in the prototype are confirmation questions. On the decision tree, the final node and leaf are defined by confirmation questions, as illustrated in Figure 7. This ensures correct case identification.

The order in which case questions are written is extremely important. In the prototype, context questions are always written before confirmation questions. After the initial search, the case base displays questions drawn from the top-scoring cases. If each of these cases have context questions listed before confirmation questions, then the question list on the search panel will display context questions before confirmation questions. This ensures the correct context is established first. The top-scoring cases share a common theme which is in concert with the search. Confirmation questions can then be answered by the user to highlight the individual case or cases within the cluster that satisfy their search. By establishing the context first, not only does it help to direct the search, it also helps the user

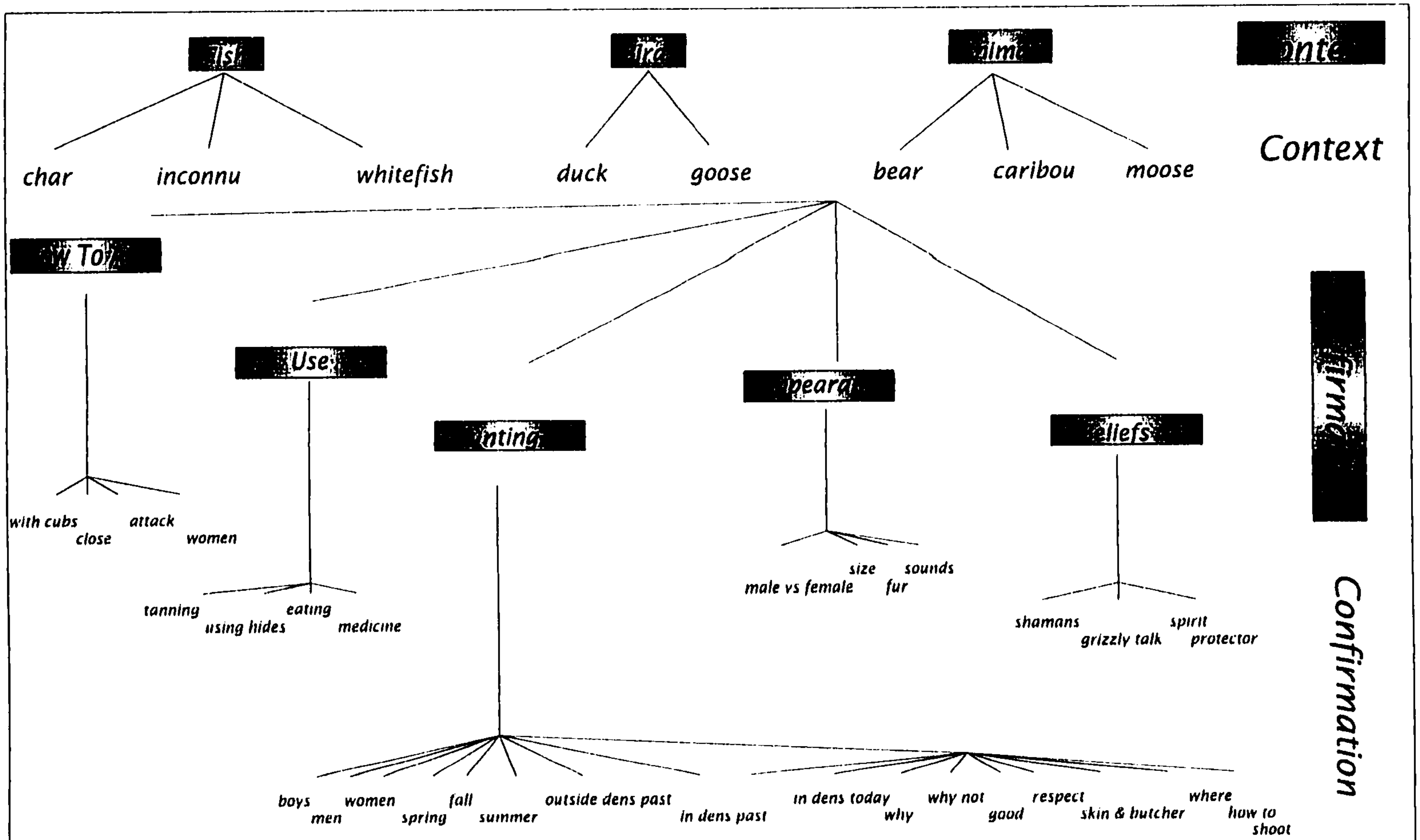

Figure 7. Decision Tree Highlighting a Partial Branch for Grizzly Bear and Displaying Hierarchy of Context and Confirmation Questions.

to explore cases which are perhaps only similar to the one for which they are searching, but nevertheless provide them with interesting information. Sometimes being able to look at a whole group of closely related cases is better than looking at a single case.

5.1.4.3: Question Structure

The CBR system allows four basic types of questions; yes/no, text, numeric and list. Yes/no and list questions are scored based on a simple string match of the answer against the cases in the case base. Text answers are scored based on the number of words in the search answer that appear in the case answer and numeric answers are scored based on similarity and range. For example, numeric ranges account for the fact that on a scale from 1-10, the number 8 is quite dissimilar from the number 10. On a scale from 1-100 however, the number 8 is more similar to the number 10. This is how numeric answers are scored in terms of similarity and range. Each type of question is scored on a percent-match basis.

In the prototype case base, yes/no questions were chosen predominantly. In attempting to design an optimal case base in which leading context questions divide the entire domain space evenly, yes/no questions provide an assurance for maintaining the structure. The user is asked for a single piece of information to which they can respond yes, no or not answered. Regardless of the answer, the system will narrow the search in a predictable direction. At this point, with the traditional knowledge system it is not known who will be using the system and for what purpose. The nature of the searches is not predictable and it is therefore important to impose a structure on the case base so it can be successfully negotiated. Yes/no questions are concise, easy to understand and easy to answer.

List questions were used sparingly and experimentally. Unless a question can be answered by a short and exhaustive list, users may find list questions to be frustrating. If the list of possible answers to the question is too long, then users will not want to scroll through the entire list to find the answer. Further, if the list does not exhaust the possibilities, then the user may simply be unable to answer the question. List questions are better served by more structured domains where the range of answers is predictable and easily identified.

Numeric questions are not applicable to this domain. Specific numbers and estimates are rarely presented in the elders interviews. Text questions, which can be answered with a user's free form text, were avoided for two reasons. First, because the domain is so unstructured, the use of text questions can lead to unpredictable answers. The answers to the questions will not necessarily clearly divide the range of cases and may therefore hinder a search. Secondly, the goal of case questions is to elicit more exact information from the user and to focus the search. In such a broad domain, it can be difficult to elucidate an exact text answer. The user may be unable to answer or may give multiple answers.

All questions, regardless of the type, must ask for a single unambiguous piece of information. Questions must make sense regardless of the context. If a case author designs one question to lead into another and those questions are not presented sequentially to the user, then there could be some confusion. In this context, each node in the decision tree is represented by a single question, often with a yes/no answer. The user is asked if they are looking for information on a certain topic relating to a certain species — yes or no, no ambiguity. Case base designers suggest reusing existing questions wherever possible. In the prototype each leaf on the tree with a shared node has a shared question. Cases can share as many as three of the four questions. This saves redundant and repetitious questions in the case base and maintains a concise and consistent structure.

5.1.4.4: Question Scoring

Each time a question is answered, the system revises its search based on the new information. Questions can be weighted according to their importance for determining a case. If questions are weighted equally then the number of questions in each case must be standardized. If some cases have very few questions and they are competing with cases that have a long series of questions, then the latter will be at a disadvantage whenever a shared question is answered. The contribution of the answer in the case with a few questions will be relatively larger than the contribution to the case with a long series of questions. The first case's score will rise substantially, but inappropriately.

In the prototype, questions were designed with equal weight for two reasons:

▸because of the importance of context in this case base; and,
▸because of the unknown uses to which it will be put.

Any similarity between the search criteria and a stored case, in the context of this research, is important. Answers which are similar to, but not exactly what the user is searching for, are important in this case base to transmit some of the important contextual issues surrounding TEK. For example, the user may be searching for information on the spring migration route of the Porcupine Caribou Herd. Although the exact answer can be found in the case base, it may be just as enlightening for the user to view a similar case on historical changes of the spring migration route. Elimination and confirmation scoring would eliminate these important similar cases, as would uneven question weighting.

5.1.5: Designing Actions

The action, or the case solution is, in this case, pre-determined. By identifying the case, the action has necessarily been identified as well. The action is really the starting point in the process because it is the single, unique piece of information for which the user is searching. How the user will find that action, through the case title, description and questions, is determined after the action is identified. What a user will find in the case actions is more fully developed in the section describing knowledge searching using CasePoint.

5.2: An Evaluation of the Model Designed Using CBR Express

After having developed the case base, it is important to evaluate both the strengths and weaknesses of the underlying model of knowledge. The best evaluation for any model is trial and error. Ideally the prototype would be made accessible to resource managers for a trial period and the resulting feedback would serve to better hone the model. Even without this process it is still possible to examine critically the model's ability to:

1. cover the breadth of knowledge currently in the GEKP and to incorporate new knowledge (i.e. flexibility);
2. represent successfully the knowledge with an essentially hierarchical model;
3. represent objectively the knowledge and be replicable.

First, as regards the model's flexibility and coverage, the prototype suggests the model is extensive though it does raise questions about its future adaptability. The current model was developed from the current GEKP material and is successful, as a prototype, in fully representing knowledge of grizzly bear and caribou, and will be successful in fully representing all of the GEKP material. The model contains all the knowledge about caribou and grizzly in a structure that is consistent and allows equal access to both data sets. The model can easily be extended to include all the other wildlife species in the database. However, the flexibility of the model to include new and perhaps very different information in the future is a point of concern. For example, if knowledge were collected on an ecosystem or spatial basis, rather than a species basis, can the existing model be adapted to include new information? As a resource management tool, it is critical the system be flexible enough to incorporate this new information.

The CBR tool certainly provides the flexibility to include new and varying data by allowing contradictory cases, shared actions or solutions and shared questions. Essentially, what one would expect from information collected on a spatial basis is much the same information collected on a species basis, simply from a different perspective. The CBR does have the flexibility to get to the same place in different ways and the model therefore should be expandable. The best way to examine the flexibility is to incorporate new and different data.

Second, the decision tree model developed for this case base is hierarchical and the question arises as to whether this is an acceptable representation for TEK knowledge. This research supports the view that the model is acceptable while appreciating the need for improvement. At first glance a hierarchical model would seem antithetical to TEK, but upon further investigation the case base model is more than a simple, flat hierarchy. The ability to view a number of top-scoring answers provides context and flexibility that would not be found in a conventional decision tree hierarchy. The user is not simply guided to a single answer, but to a range of related answers which provides the context that is more in keeping with TEK representation. As is discussed above concerning the use of CasePoint WebServer, links from answers or actions directly to interviews again provide the required and appropriate context.

Further, the hierarchy is appropriate because of the speed and the ease of conducting searches. When a user reaches an answer at the bottom of the decision tree and seeks more and selected information, a click of a button will launch a new search immediately and direct them down any other path in the hierarchy. For example, if they are reading about how caribou try to avoid predation from wolves, a search can quickly and easily be launched to take them to wolf habitat. So although the leaves on the decision tree are not directly linked, the ease of conducting new searches can quickly take a user to related leaves. In an ideal model, the leaves could perhaps be better linked.

The final evaluation is directed toward the objectivity of model development and its replicability: if someone else did it, would they do the same thing? This is a particularly important question when the differences between western and indigenous systems of knowledge organization are taken into account. The model was developed by a researcher trained in a western-based education system, and is therefore necessarily imbued with a western logic. However, the fact remains the knowledge of the GEKP is fully represented across its breadth and depth, and context is provided both with multiple answers and with the use of the words of the elders. Although another researcher from a different background may have developed the model differently, the end result would be the same representation of the knowledge. The success of the current model can only be evaluated after a thorough testing by researchers in the Gwich'in organization, both non-native and Gwich'in. This is an evolutionary process. Changes will take place and the system's merits and failings will be more clearly understood.

5.3: *AUTOMATIC CASE AUTHORING: CBR GENERATOR*

After the initial case base was designed and launched on a stand-alone machine, as described above, experimentation began with an automatic indexing tool developed by Inference Corporation called CBR Generator (http://www.inference.com). The goal of Generator, and of indexing tools in general, is to create case bases automatically, thus facilitating efficient case base design and ensuring impartial case indexing; two significant advantages over manual case base development.

However, reviews of automatic indexing facilities in CBR have been largely negative. People choose indices manually better than algorithms do automatically (Kolodner 1993 cited in Watson and Marir 1994). The experimentation with Generator in this research supports an extremely cautious view of the utility of automatic indexing tools. This section gives an overview of Generator's functions and capabilities, then describe why it was rejected as a tool to develop further a TEK case base.

The purpose of the CBR Generator module from Inference is to automate case base creation by automatically indexing large collections of existing on-line documents. Generator reads document files, then develops a case base to serve as an index to the documents. Using a lexicon file, Generator identifies and categorizes statistically significant words in the documents. The lexicon, or text file, in Generator maintains a starting list of 50 000 mainly English words and identifies the words on the list by part of speech and by the number of meanings the word may hold (Inference Corporation 1995c). The words are, in turn, used as indicators to help distinguish one document from another. Although Inference will not reveal the algorithms Generator employs, significance is drawn from analysing the patterns and the frequency of word use in the document set. The result is a case base complete with titles, descriptions and questions. As its answer, each case displays the original file that created the case.

Unfortunately, Generator does not know what the words actually mean. A word may be statistically significant yet be totally irrelevant to a case. Not only are statistically significant words sometimes irrelevant, they are also often illogical and non-intuitive when strung together. Generator uses the significant words to develop the case title and description, but because Generator cannot interpret the meaning of the words, it merely strings the words together without regard to structure or form. Similarly, key words are inserted into questions without regard for form or meaning, often making the questions unintelligible. Further the supposed key words inserted into a question rarely help to highlight the important features of the case. They are not the key words a person would usually identify.

The initial development of a case base using Generator is very simple and direct. Within the Generator application, the author simply identifies the path or file names of the documents to be indexed and Generator develops the case base using its default parameters. Figure 8 displays a generated case. Note the difference in intelligibility between this and the manually created case shown in Figure 9. This is where the difficulty in using CBR Generator lies.

Each of these cases is drawn from the same material regarding preferred summer habitat for caribou, but the indices chosen to identify the material are very different. Chosen indices are in the 'Title' and 'Description' boxes in Figures 8 and 9. Even to a non-author, the title and description in the manually created case are clear, concise and easy to understand. The automatically generated title and description on the other hand, contain many of the same words but the relationship between words is difficult to comprehend and therefore the nature of the case itself is difficult to comprehend. Some of those words are very poor indicators for the case: words such as 'snowmobile', 'disappear', 'become' and 'come'. None of these words really tell you much about caribou summer habitat.

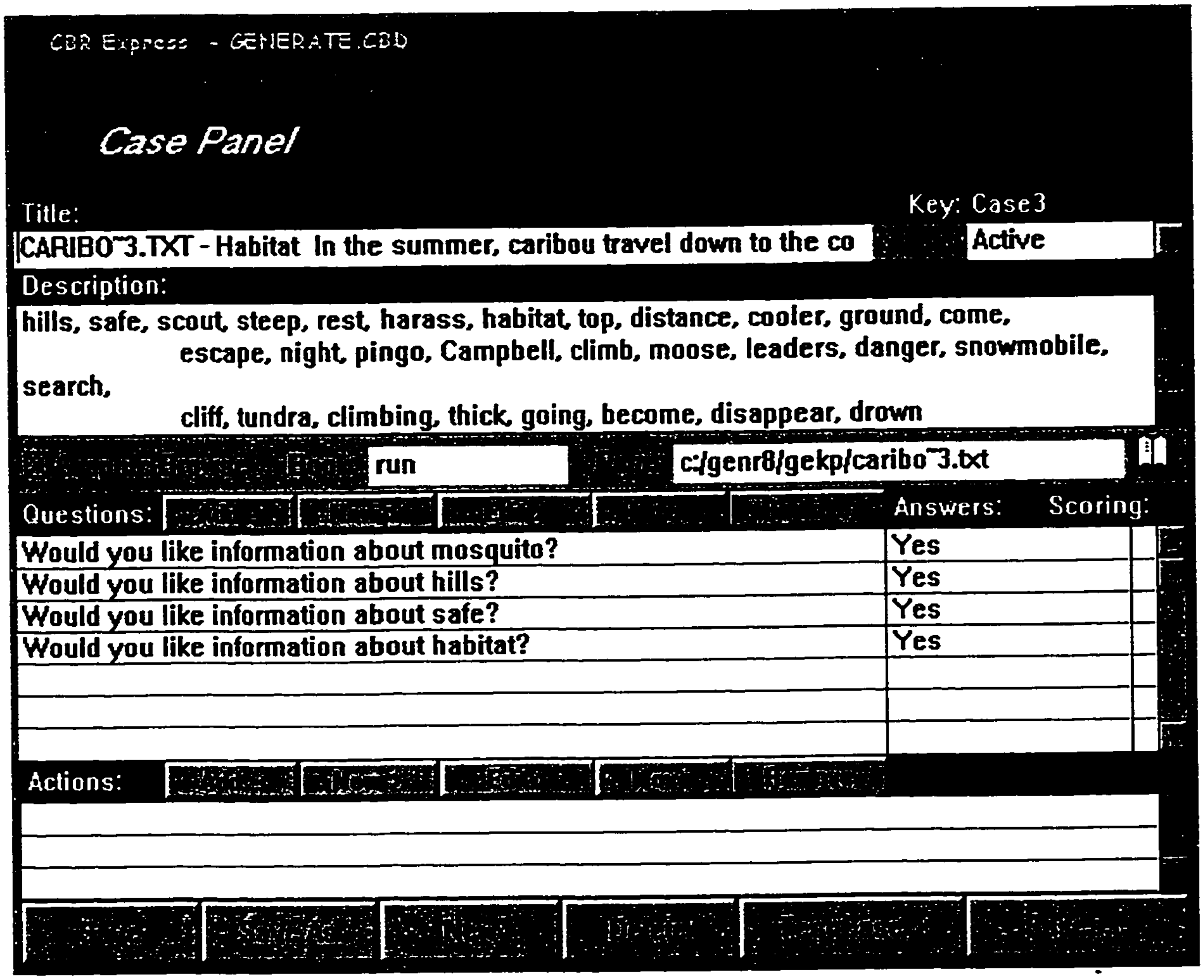

Figure 8. An example of an automatically generated case displaying information about preferred summer habitat for caribou.

Similarly the questioning frame in the generated case uses keywords which will not successfully guide a user to the case action. In wondering about caribou summer habitat, very few users would answer 'yes' to questions about mosquitos, hills, or safety. If they answer 'yes' to these questions, they probably already know the nature of caribou habitat and that caribou like hills, don't like mosquitos and like to feel safe. However, the manually created questioning frame clearly leads a user, regardless of their level of prior knowledge, toward the appropriate action or answer.

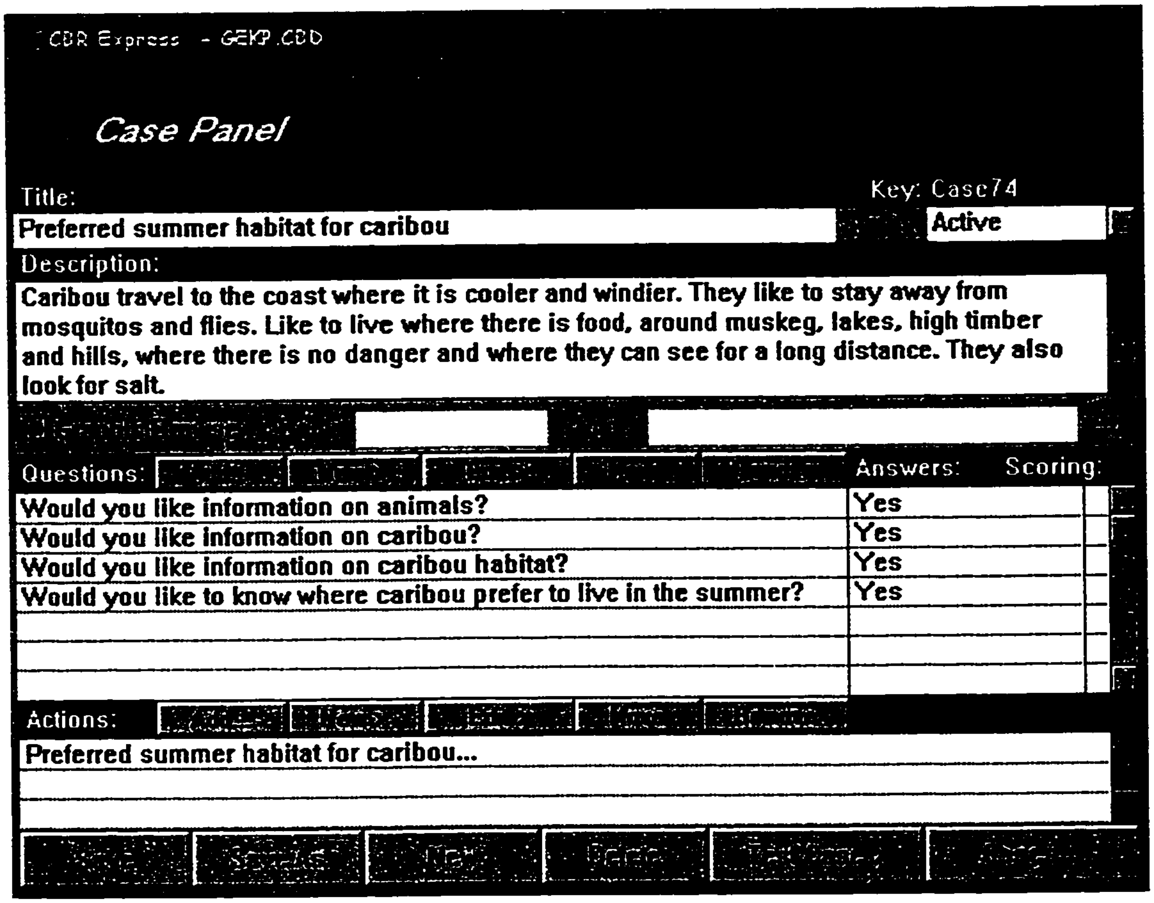

Figure 9. An example of a manually generated case displaying information about preferred summer habitat for caribou.

The question then becomes, if the generated results are not acceptable from the outset, can the parameters subsequently be modified to produce satisfactory results and thus still allow Generator to serve as an efficient 'first step' in case base development. Key word choices, and the parameters used to make those choices, can be edited in Generator, but for the purposes of this research it was determined that the degree of editing required to produce a satisfactory case base is too great. Parameter modification is discussed and evaluated in the following paragraphs.

The first parameter to be altered, in an attempt to improve upon generated results, was document length. The logic in changing document length stems from the notion that if you have long documents covering a variety of topics, the variation in each document will be difficult to index. Alternatively, if short documents covering specific topics are indexed, then the indexing should be more successful in identifying the variation between documents. Documents were indexed first as original interviews, then as summary chapters and finally as topical sections of summary chapters. Neither length, nor style of document, could be varied to produce a readily acceptable case base.

Second, to make the generated case base more acceptable, parameters which guide the indexing algorithm were also changed. The number of questions which Generator automatically associates with each case can be set from a minimum of zero to a maximum of six (Inference Corporation 1995c). The optimum number of questions for a case depends on the length of each document and the number of significant words it contains. Further, the threshold which makes a word significant enough to include in a question can be altered. Neither the threshold, nor the number of questions selected, developed a questioning structure and order that were acceptable. The difficulty in defining an acceptable threshold or number of questions stems from the fact that the automatic generation of questions follows a model which is opaque to the author. Without being able to view or comprehend the underlying model, it is very difficult to understand what the questions are asking and where they are leading. There are no context or confirmation questions and therefore no discernible directions. Generator allows the author to select new keywords to make the questions more comprehensible and more readable for a user, but the time required to understand the logic and the goals of the questioning structure would be better spent developing the case manually.

Along with altering the parameters surrounding question development, significant words chosen for the title and description can be manually altered and/or weighted. Words can be biassed, or weighted, to make them more significant and phrases may be indexed rather than individual words. For example, in a document describing the Porcupine Caribou Herd migration, the word 'migration' should be disproportionately weighted as a key word. Similarly, 'Porcupine Caribou Herd' is a phrase that should be weighted. After the initial case generation, any identified word may be manually removed and replaced with a more appropriate keyword. The lexicon file which Generator uses to identify and categorize words can also be edited to reflect better the documents being indexed. Again, with no view of the underlying case base structure, the time required to re-index and bias words and phrases to make the case base more acceptable, would be better spent developing the case manually.

Despite the flexibility of Generator, the problem of Generator stringing together statistically significant words without form or structure remains. The flexibility allows you to make changes but the scope and the nature of the changes is so great, automatic case generation is no longer attractive. The time spent 'decoding' and replacing the automatically indexed words and cases is greater than that required to develop the case base manually from the outset. Manual case base development gives the author more control and a better understanding of the case base, and results in a more intuitive and intelligible case base for a user to search.

CHAPTER 6: USING CBR SOFTWARE FOR SEARCHING AND RULE GENERATION

6.1: HOW A USER SEARCHES AND RETRIEVES INFORMATION

After examining Inference's tools for knowledge management and acquisition, the research moved toward examining tools for knowledge searching. The search tools are tied intimately with the knowledge acquisition process because the user searches the case base which resulted from the authoring, or knowledge acquisition process. From the authoring process it was shown that manual case base authoring with CBR Express was more effective than automatic case authoring with Generator, therefore the manual case base developed using CBR Express was incorporated into two knowledge searching tools, CasePoint and CasePoint WebServer. The following sections describe the searching process.

6.2: CASEPOINT: THE USER'S SEARCH INTERFACE

CasePoint is Inference Corporation's knowledge searching tool and is functionally separate from their knowledge acquisition and management tools. With an associated case base, CasePoint can provide a number of users with the ability to search and retrieve information from the case base, though it does not give them the ability to edit a case base. CasePoint has a straightforward interface that solicits information from a user to perform an interactive search, then presents the user with potential answers to their search. This section describes CasePoint searches from a user's perspective (i.e. what the user sees), how rules are invoked to guide the search more efficiently and the range of possible answers presented to a user.

6.2.1: How Does a User Interact with CasePoint?

The initial screen presented to the user is divided into three sections: description, questions and actions. At the description prompt the user types, in natural language, a few words or phrases describing their search. For example, Figure 10 shows the CasePoint search screen after the initial search based only on the user's description, 'spring migration of the Porcupine Caribou Herd'. CasePoint looks for closely matching titles and descriptions in the case base by performing an initial trigram match (recall Section 5.1.3.2 for a description of the matching algorithm) on the description alone and returns, typically, the five closest cases and lists their actions (Inference Corporation 1995d). Five cases are the default returned by CasePoint, though this number may be changed by the user. Each action, or answer, is displayed with a matching score, between 0 and 100, to highlight for the user the 'best fitting' cases. In the example in Figure 10, the five top-scoring cases range from a score of 80-76 based on the initial search.

CasePoint also retrieves a set of questions for each of the matching cases and displays the combined set of questions on the panel. The user now has the option of answering any or all of the questions, in any order, to help sharpen the focus of their search or, if they feel they have found suitably matching cases, they may view the matching actions. If the user chooses to answer a question(s), after each answer, the case base is re-searched based on the new information provided in that answer. If the user in Figure 10 answered the first questions 'Would you like to know the Porcupine Caribou Herd spring migration route?' with either a 'Yes' or 'No', CasePoint would immediately re-search the case base with the new information. The answers to the questions will change the scores of matching cases and may eliminate some cases entirely, while presenting new, possibly matching cases. Each time a new case is presented as a possible match, the associated questions will also be presented to the user. Thus, with every answered question, a new search is launched, matching scores are changed and new questions and actions are presented to continue the iterative search process with the user. In Figure 10, answering the first question affirmatively will boost the top-scoring answer 'The Porcupine Caribou Herd spring migration route and gathering areas...' from a match score of 80 to 100, therefore indicating to the user that this is their best solution. The user can choose to view an action at any time.

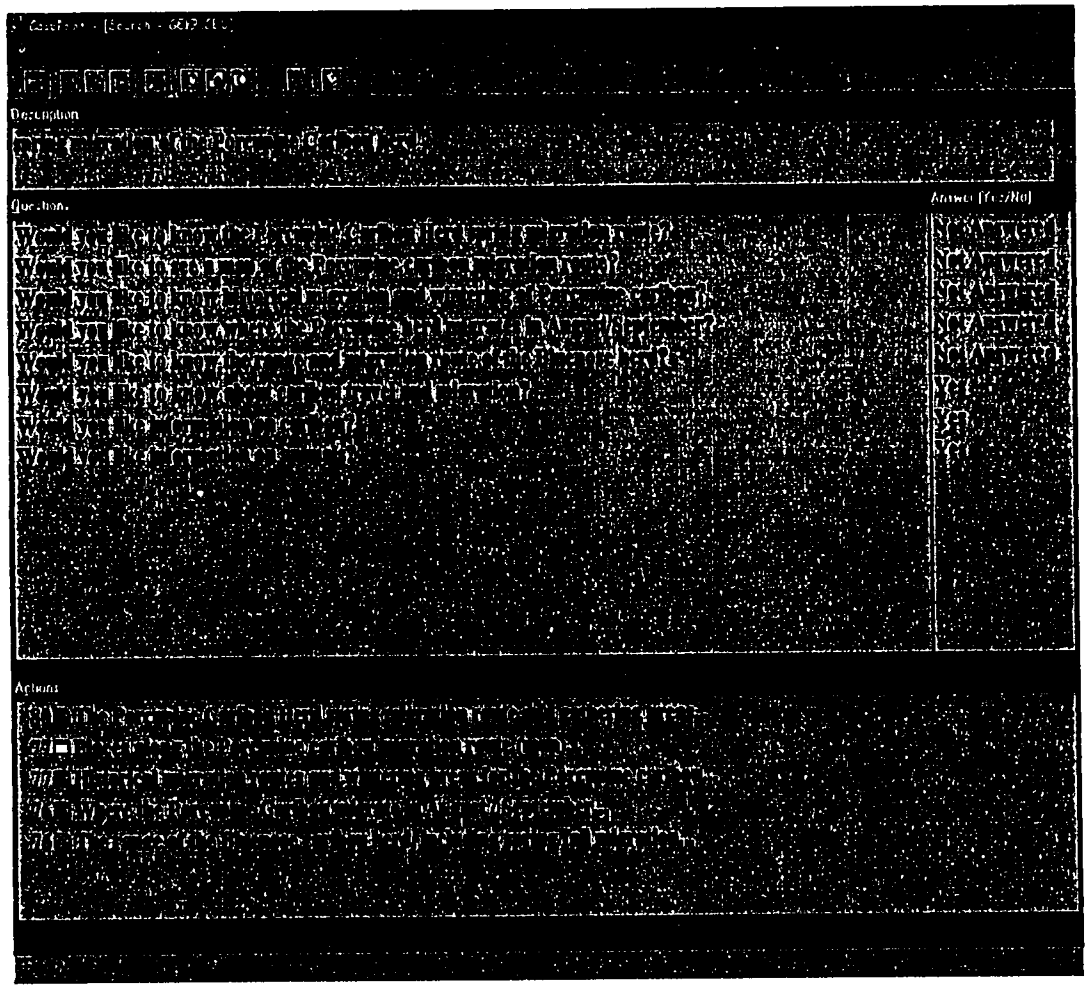

Figure 10. The CasePoint search screen after an initial search based only on the user's description.

Depending on the icon associated with the answer, an experienced user will have a preview of what the action contains. For example, the two top-scoring actions in Figure 10 have different icons associated with them. The top-scoring answer is text, while the second answer indicates a file to be launched and viewed. This is not critical information for new users but can be helpful to experienced users. If they are not satisfied, they can return to modify the search or conduct a new search.

If the user feels they have exhausted their search with no satisfactory results, they can save the entire search consultation in CasePoint's memory at the click of a button. The unresolved searches become the responsibility of the case base author. It is incumbent upon the author to find the information if it is not in the case base or to revise the case base if the information is there, but was not indexed properly.

6.3: Writing Rules to Automatically Answer Questions

Based on the initial search in Figure 10 some questions have already been answered. The three final questions in the Figure 10 question frame are answered 'Yes'. These questions are answered according to a rule file that is invoked whenever the CasePoint search engine is started. The rule file contains a series of 'If...Then' rules which answer automatically questions based on the user's initial search description. If the words in a search description match the condition in a rule, the result is an answer to a specific question or to a series of questions. The rules 'fire' whenever their conditions are met, regardless of sequence. This is referred to as opportunistic execution (Inference Corporation 1995e). CasePoint thus appears to be operating 'intelligently' by recognizing what the user is searching for.

In the prototype case base rules were developed to answer context questions. Each of the three questions with a 'Yes' answer in Figure 10 is a context question. Since the search description contained the word 'caribou' for example, a rule fired automatically to answer 'Yes' to two of the context questions in the case base:

1. 'Would you like information on animals?'

2. 'Would you like information on caribou?'.

Similarly, the word 'migration' triggered a rule.

Figure 11 illustrates the required rule syntax and provides examples from the prototype. The rules are hidden during the search: the user does not see the rules. The rules are developed in a separate text file by the case base author and can be accessed only by the author. The rule examples in Figure 11 are only a few of many rules developed for this case base.

Unfortunately the rule file is not very tolerant of spelling mistakes or differences. Unlike the trigram matching used in the initial search, for a rule to fire, the match must be exact. A near match is not enough to trigger a rule. However, the rule file can be written to tolerate subtle differences between words, like suffixes and prefixes, as in the use of 'migrat' in Figure 11. The rule will fire when 'migrat' is succeeded by any combination of letters: migration, migratory and migrate. Because of the exactness of rule requirements, rules sometimes do not fire when they would be expected to, but thoughtful rule composition can avoid mismatches in most instances.

If these questions were not answered automatically, they would be redundant to the user. Why should the user answer 'Yes' to wanting information about caribou when they have already asked for information about caribou. By not answering questions automatically, CasePoint would waste the user's time unnecessarily and would make the system appear 'stupid'. In the event a user does not agree, the answers which the rule file provides may be overridden. However, in most cases the rules should help to focus quickly the user's search and move it toward a rapid conclusion.

```
//Whenever the description contains "caribou" then the answer
to "Would you like information on the caribou?" should be "Yes".

0 "caribou"          if the condition is met
=>
6 "Yes"              then action
;

//Whenever the description contains "caribou" then the answer
to "Would you like information on animals?" should be "Yes".

0 "caribou"
=>
6 "Yes"
;
//Whenever the description contains "caribou" and "travel" or "migration, then the answer to
"Would you like information about caribou travel and migration?" should be "Yes".

0 "caribou"
0 "travel"
=>
105 "Yes"
;

0 "caribou"
0 "migrat "
=>
105 "Yes"
;
```

Figure 11. Examples of Rules and Their Required Syntax

6.4: ACTIONS, OR ANSWERS IN CASEPOINT

As shown in Figure 10, the user has a list of actions from which to choose, based on the degree of match with the user's description and answers. Actions can be simply a textual solution to the user's search, comprising a few words or a series of paragraphs, or actions can launch other windows-based programs and display specified files. This is the 'graphic browse' feature of CasèPoint. By issuing a run command in the authoring portion and identifying the complete path name for a file, CasePoint will launch the program that created the file. If, for example, a question in the case base could be better illustrated with a diagram than with text, then the graphic browse feature would be invoked. This feature has been found to be very valuable in technical troubleshooting case bases. For example, a wiring diagram may be the most appropriate solution in an electronics case base, whereas a textual solution would simply not provide the user with the information they require.

In this instance, some actions in the case base are better illustrated by media other than text. Diagrams showing how caribou corrals are built or photographs showing preferred browsing species for caribou, for example, would greatly enhance the case base. Audio and/or video files with case explanations given by the elders would provide the context and, in some sense, be more representative of the holistic outlook engendered by TEK. The ability to expand the case base to include a multitude of data sources has significant appeal for a TEK case base.

This research focussed on the incorporation of locational information displayed in mapped images or in GIS files. The possibilities for this type of incorporation are explained in a subsequent section.

6.5: Using CasePoint WebServer to Develop Internet Access

CasePoint WebServer is an interactive web application designed to provide CBR search and retrieval capabilities via any standard World Wide Web (WWW) browser. Superficially, WebServer operates like CasePoint. Users can search and retrieve information by simply typing a natural language description and interactively answering questions. Fundamentally however, the capabilities of, and therefore the possibilities for using WebServer are very different from CasePoint. WebServer provides screen customization opportunities, through hyper-text markup language (HTML) commands, centralized content management and a greater ability to incorporate data from other sources. This section discusses each of these features, how they have been applied and why WebServer shows great potential for the TEK case base.

First, WebServer allows centralized content management, as well as de-centralized content access. Any user with an Internet connection and web browsing software can access the case base and therefore the case base itself is open to all WWW users or, if preferred, to only a select few. Because WebServer is a search and retrieval tool only, the case base author(s) can edit the case base. This ensures that the integrity of the case base will not be compromised by unauthorized user editing. Issues related to case base security and content are incumbent upon the case base author — either a single individual or a small group. Centralized content management ensures that the case base remains consistent and maintains its integrity. Authors also have the power to invoke case base security measures and therefore to limit access to a select group. What centralized content management and de-centralized content access means to the end user is equal access. All users have access to the same version of the same case base at the same time.

Thus far, there has been a single author of the TEK case base, the researcher. This established the researcher as the case base 'gatekeeper', deciding what went in to the case base and what did not. This approach worked well for prototype development as it allowed for a singular consistent process of trial and error. If the case base is to be fully developed, a small group of people providing gatekeeping services is ideal for maintaining some breadth in decision making, while maintaining consistency and integrity in presentation. If the number of people making changes is too great, the model could quickly become unrecognizable to the authors themselves. Having gatekeepers and releasing the case base on the web means everyone across the settlement area has access to identical data. As soon as a change is made and approved, everyone can access it immediately.

A second key feature in WebServer is the screen customization option. The user interface can be changed to satisfy the target audience and to complement other Gwich'in web pages and applications. WebServer provides a text-based resource file which defines the user interface. The resource file gives the author quick and easy access to change the HTML code and modify the WebServer installation settings, appearance and behaviour. Customization options include:

▸changing preferences such as the maximum number of cases displayed, threshold percentages, and the position of questions and actions;
▸modifying the placement of descriptions, questions and actions so they appear on separate pages and therefore simplify the interface for inexperienced users;
▸developing links to actions which may be Web pages on the same server, URL addresses or attachments on the client, as in CasePoint (Inference Corporation 1995f).

The options available for displaying actions in WebServer are far greater than with CasePoint and these options help to solve many of the problems encountered with CasePoint. WebServer allows greater opportunity for the integration of other data sets. In the TEK WebServer case base, each action is linked to one, or to a number of web pages. Figure 12 displays a sample action with links to two interviews, that of Tony Andre and Elizabeth Greenland. The full text of each of these interviews was developed as a web page. When the user reaches a solution with the CBR, they can explore the hyper link to the original interview that contributed to that solution. In the case of Figure 12 both Tony Andre and Elizabeth Greenland refer to grizzly bears not hibernating in the winter. The link takes the user to the web pages displaying the interviews and places the exact text which contributed to the answer at the top of the user's screen. Thus the user has the option of quickly finding the 'short answer' to their question, which is summarized in the 'action', exploring further those parts of interviews which contributed to the answer, or reading the entire species-based interview.

Integration of the original interviews provides an important measure of context to the solutions that was missing from CasePoint. The user has the option to explore interviews with many different elders and can read the subtle differences in each person's description

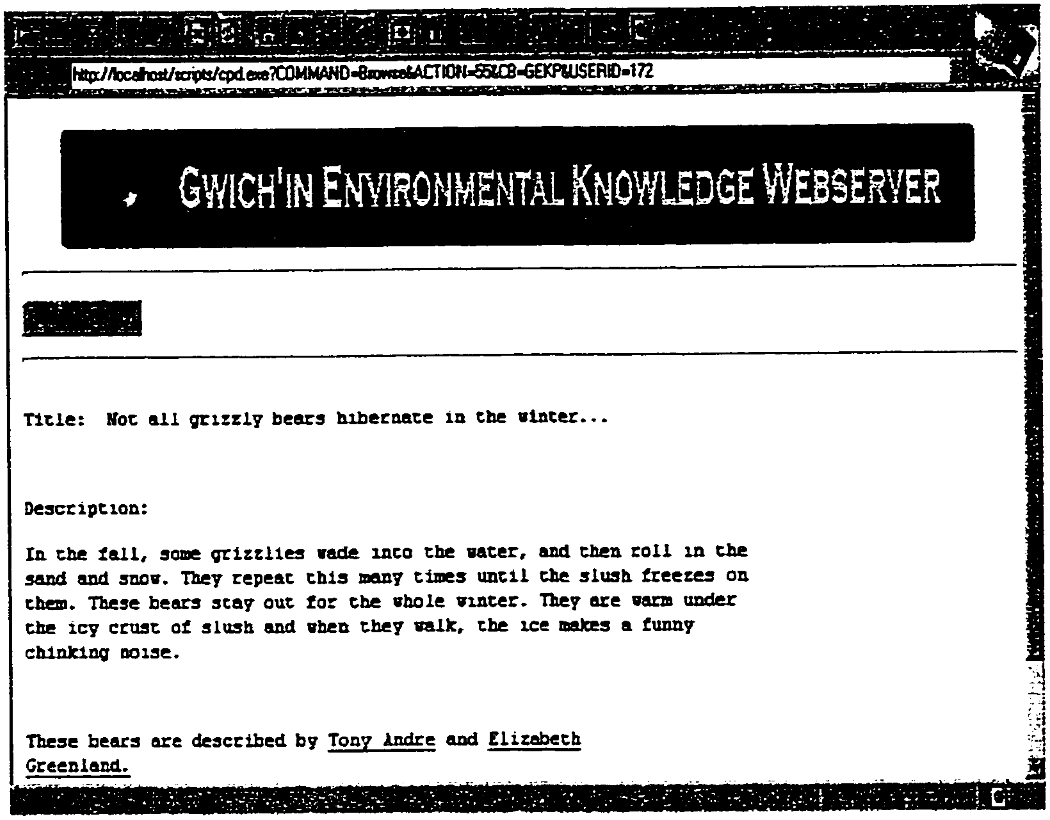

Figure 12. The solution to a WebServer search of the TEK base. Hyper links to interviews supply the user with more information and provide context for their search.

which were previously amalgamated into single solutions in CasePoint and were glossed over in text summaries. The solution is now truly in the words of the elder. As with CasePoint a number of top-scoring answers are suggested by WebServer for viewing. This allows a user to access similar solutions with links to different interviews to provide closely related, and sometimes even contradictory information. The ability to link actions and web pages in WebServer comes closer to the ideal of representing TEK in an appropriate context: in a context which is inclusive rather than exclusive and in a framework that is more applicable for holistic, contextual information.

Rudimentary changes to the resource file were made to modify the screen appearance and behaviour of the Gwich'in TEK base. The interface was modified to represent better the project and to include links to other Gwich'in web pages (see Figure 13). For each application interface, icons were developed to display the Gwich'in logo and the project name, and user guides and instructions were modified to fit the TEK base.

Further experimentation was made to change the preferences and the interface, but developing too far in this direction, at the moment, is premature. An integrated resource management database, although discussed, has yet to be implemented by the Boards. Ideally the TEK case base should be integrated and developed with this strategy and should therefore have a similar appearance. Until the new database strategy takes greater shape, any significant investment in WebServer interface development is likely to be premature.

Future developments for the TEK case base and other database strategies that develop on the Web will require that a considerable emphasis be placed on interface development. To allow people to access web-based information and to encourage people to use web-based information to its full extent, web applications and pages have to be exceedingly well designed. With the amount of information available to people, particularly on the web, the information must be well formatted and be easy to use and navigate, particularly for inexperienced users. If the user does not get satisfactory results and/or the process is painful and awkward, they will not return to seek new information.

Simply making information available over the web does not ensure its successful use: "We believe that computers, and the WWW in particular, have in many ways failed to meet their potential to enable users to work smarter and more productively" (Grose, Forsythe and Ratner 1998:122). The goal of the integrated database management strategy and the TEK base is to make data equally accessible over the Internet. Unless careful attention is given to interface development, web based implementations may have the opposite effect: they may serve to further the distance between the 'technology haves' and the 'technology have-nots', as has been the case in other studies (see Ratner 1998). In order for a web medium to be

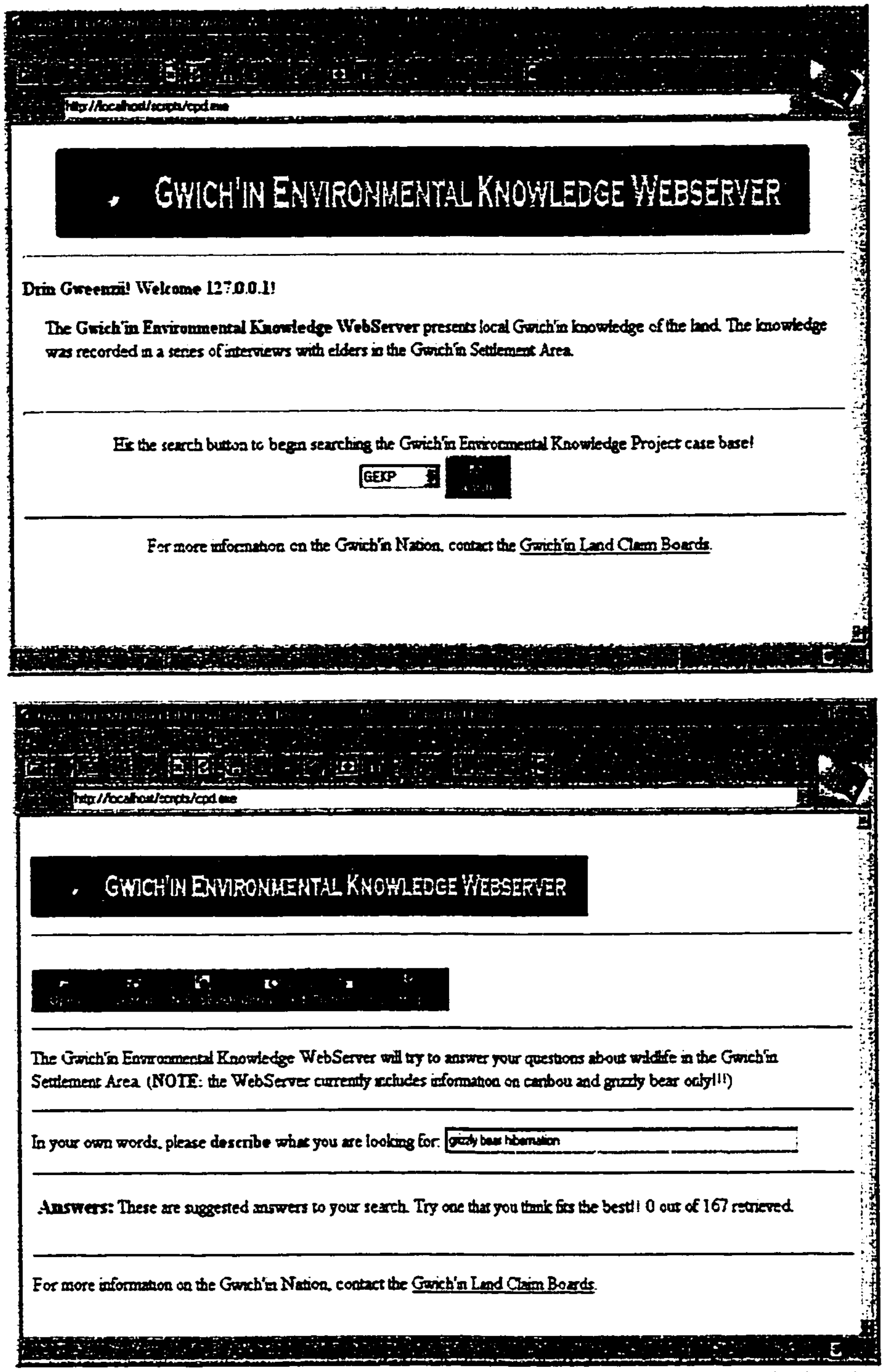

Figure 13. Sample interface to introduce a user to the knowledge base and guide them to initiate a search.

successful in the Gwich'in organization, not only are aesthetics and computational efficiency of importance, but so are overall design features that enhance the experience for new users. Human interaction and web interface development is a relatively new area of study and has an important future in this area.

CHAPTER 7: PATHWAYS TO KNOWLEDGE INTEGRATION

7.1: GIS INTEGRATION ON A STAND-ALONE MACHINE

After discussing the authoring of case bases and searching methods, it is important now to turn toward the incorporation of other data, such as GIS, with the case base. GIS is a key data management and analysis tool used in most resource and land management agencies. The importance of GIS holds true within the Gwich'in organization as well. In the Gwich'in resource management strategy, a GIS platform will provide the backbone for the integration and presentation of data across the organization. Given its importance in an overall resource management strategy, this research sought to develop active links, or 'graphic browse' links, between the CBR and the GIS, rather than exploring possibilities such as audio or video links. Since the goal is to use and access TEK on an equal footing with Western science data for use in resource management decision-making, the integration with GIS will be critical.

This integration will be beneficial to the data stored on both platforms, the GIS and the CBR, by increasing the depth and form of available information. The GIS would benefit by having, essentially, extended attribute files which could include lengthy descriptive text segments rather than just limited attributes. Similarly, the CBR would benefit by having detailed spatial information added to cases along with the flexibility to analyse spatial information in the light of descriptive information.

For the integration of a CBR and a GIS, there are essentially three options:

1. have the GIS files attached as 'browsable' files at the end of CBR actions or answers,
2. have the GIS operate as a client program for CBR, sending requests and initiating the CBR when required, or,
3. have both the CBR and the GIS operate through a common interface on a network.

None of these options is easy to implement with consistent and satisfactory results. The choice of option used in this research is defended, and the benefits and drawbacks of each of the three approaches are explained.

7.1.1: Operating a GIS as a Client Application

The first and third options were chosen for exploration in this study: attaching GIS files to the CBR and launching both the CBR and the GIS from the Internet. The second option, having the GIS act as a client program to the CBR, was discounted in this research though it may prove valuable in future research with TEK and/or resource management. As regards the second option, using the Dynamic Data Exchange (DDE) protocol in a Microsoft Windows operating environment, it is possible to have a client Windows application drive CBR Express (Inference Corp. 1995g). The client application need only support DDE programming. The approach to integration is relatively simple and straightforward from a programming perspective. The client application executes a set of DDE commands, allowing CBR Express to be initialized remotely from within the client application. The client sends data to the CBR where the search is conducted and the CBR responds with a solution to be displayed in the client application window. This data exchange can be configured in one of two ways: with CBR Express running entirely in the background, invisible to the user or, by moving the user's focus from the client application directly to CBR Express and making it entirely visible to the user (Inference Corporation 1995g). In the first configuration the search process operates remotely with no participation from the user. In the second configuration the user completes the search by operating CBR Express directly, then moving back to the client application.

With direct interaction between two different data sets in two different programs, a GIS and a CBR, there is great potential for data integration. While using the GIS a user can have direct access to TEK or, to other resource management related data located within the CBR, which is operating in the background. For example, an attribute field in the database table of a GIS layer, can be commanded to initialize the CBR. By initializing the CBR, data stored in a record in the attribute field would be used to launch the CBR search. The record may include, for example, an attribute called 'caribou corral', this information would then be taken to the CBR to search for any information the CBR has on caribou corrals (see

Gwich'in Elders 1997 for a description of caribou corrals). This is where the options for initialization differ. The user could go directly to the CBR to guide their own search on caribou corrals, or the CBR can be commanded to complete the search and display the solution in the GIS. With either option, the user is dynamically exchanging and integrating data between two data sets and programs.

However, the drawbacks to this approach, in the context of this research, outweigh the benefits. Having the GIS operate as a client application to the CBR is not ideal for this project as one of the goals of the integrated resource management plan in the settlement area is to make data readily accessible to people throughout the resource boards and councils; people who may or may not be familiar with GIS software. To have the GIS operate as a client program for the CBR requires that the user be familiar with the GIS and this would not satisfy the goals of the management plan. This would limit the number of users. Therefore, the cost of investing time in establishing links between the two programs would outweigh the limited benefits that could be gained.

Similarly, it was noted that developing DDE links between programs is relatively straightforward from a programmer's perspective. For people without programming experience however, the requirements for developing links would be onerous. Again, to comply with the goals of the Gwich'in management strategy, all aspects of data development and deployment should be accessible across the organization. Further, if the tools are to be successful in a resource management framework, they need to be dynamic. Data and information changes every day, and the information in the tools must therefore change along with it. If developing links requires a contracted programmer, then the cost of development is amplified. If the programmer must be hired each time changes are required, then the utility of the development is reduced. With a limited number of experienced GIS users in the Gwich'in organization, coupled with significant development costs, the deployment of a GIS as a client application to a CBR is not practical.

With the appropriate skill level, this remains an attractive opportunity for other resource management agencies, and may still hold promise for the Gwich'in in the near future. Use of GIS is increasing and therefore the number of experienced users within Gwich'in organizations is increasing. Training and maintaining an 'in-house' programmer could benefit a number of applications, and in this instance would reduce development costs and ensure the utility of the program by making sure it is updated. Although operating a GIS as a client application to the CBR is not a viable option for the Gwich'in at the present, it may very well be in the near future. If the option can be made viable, then the potential is considerable.

7.1.2: Attaching GIS Files as 'Browsable' Files

Attaching GIS files to CBR actions and/or developing a common network interface for the two databases are the best options for this project because they do not require the user to be an expert in GIS. Attaching GIS files to CBR actions is simple and direct, but again the benefits are limited. The CBR allows any multi-media file to be attached to the answer of a case. In CBR development, a file may be attached to each case for browsing by specifying the path to the file when authoring the case's answer. When the user searches the case base and chooses that answer the CBR opens the file by launching the program that created it. If the user is searching for information on the migration of the Porcupine Caribou Herd, for example, this information exists in two locations. Descriptive information on where important migratory areas are located is in the case base, while the geographic location of the areas is in the GIS. The CBR allows a user to find descriptive information as one solution to the question of caribou migration. As an alternative solution the CBR should allow a user to launch and view the GIS file that contains the geographic locations.

However, the difficulty is in launching the GIS. GIS file structures are quite unique and therefore it is not as straightforward to launch a GIS program as it is to launch many other Windows-based programs. With most Windows-based applications, be they text, graphic, audio or video, the text document, the graphic image or the audio or video clip is stored as a single file. When the CBR points to the file and requests it be opened, the operating system recognizes the program that created the file, then opens the file, much like 'double-clicking' on a file on the desktop. However, with each layer in most vector-based GIS packages there are at least two associated files. One file contains the coordinate/spatial information for the layer and the second file contains the attribute information for the layer. In the example of the Porcupine Caribou Herd migration noted above, the coordinate information locating migration points is in one file, while attributes describing those points – attributes like place names – is in another file. Each file containing information on a single layer is stored separately in a different format. When opening the layer in a GIS both files are accessed and used to present a single layer. In a CBR graphic browse only one file can be specified to open. Pointing to either of the files is not enough to launch the GIS and it is not possible for the CBR to first execute the GIS then open a file. This requires two steps, but only one step is allowed with the graphic browse option.

While the CBR cannot launch the GIS, it can launch a stored GIS image. The CBR and/or GIS developer can use a GIS to create images with a number of layers, then store those images as graphics files. Tables with associated attribute information can be appended to those images. Rather than launching a GIS, the CBR can read and launch, essentially, a digital map. Launching a digital map does meet the objectives of the research by integrating data sources, but it does not fully meet the objectives, as is discussed in the following paragraphs. First, the creation, development and launch of digital maps is described, followed by an explanation as to why this approach is not ideal for resource managers.

As part of the GEKP, all locations and place names referred to in interviews were located on paper maps, but were not yet digitized when this research was conducted. As part of this research, all the locations were digitized and an attribute file was attached which includes the Gwich'in and English name, along with the activity and animal associated with the place and the season of importance. This created a TEK knowledge layer in a GIS database with just a few short words in the attribute file to describe the significance of those locations.

The attribute file provides little information about the knowledge associated with each location, however the interviews give far more information. This is the juncture at which the CBR and GIS images can be integrated to complement greatly both data sets. For example, a search for information on the Porcupine caribou migration yields, with an equally high score, textual descriptions drawn from interviews as well as a mapped image of places mentioned in the interviews (see Figure 14). Selecting 'Places along the Porcupine Caribou Herd migration route...', shows the user the mapped image displayed in Figure 15. The image can be annotated with the same place names that are described together with their significance, in the accompanying text action. Any number of layers of potential interest could have been added to this image including, for example, trail data, cabin data and sacred sites. Therefore, a certain level of integration has been achieved. A CBR user can view descriptive information as one potential solution to their query and can view spatial information in an accompanying solution.

This has seemingly achieved some integration between the CBR and the GIS, but the integration is superficial and limiting. One of the major problems encountered was with the original data itself. Specific locational and spatial information was not recorded consistently in the interviews. Wildlife information is often generally described and is only interspersed with specific place names. Relying solely on a place name as an indicator of the spatial coverage of data simply is not acceptable. If the elders say in an interview that the Porcupine River is a particularly good area for hunting caribou in the fall, does that mean one specific point on the river is good for hunting? Does it mean the hunting is good along the length of the river or does it mean the entire Porcupine River watershed is good for hunting?

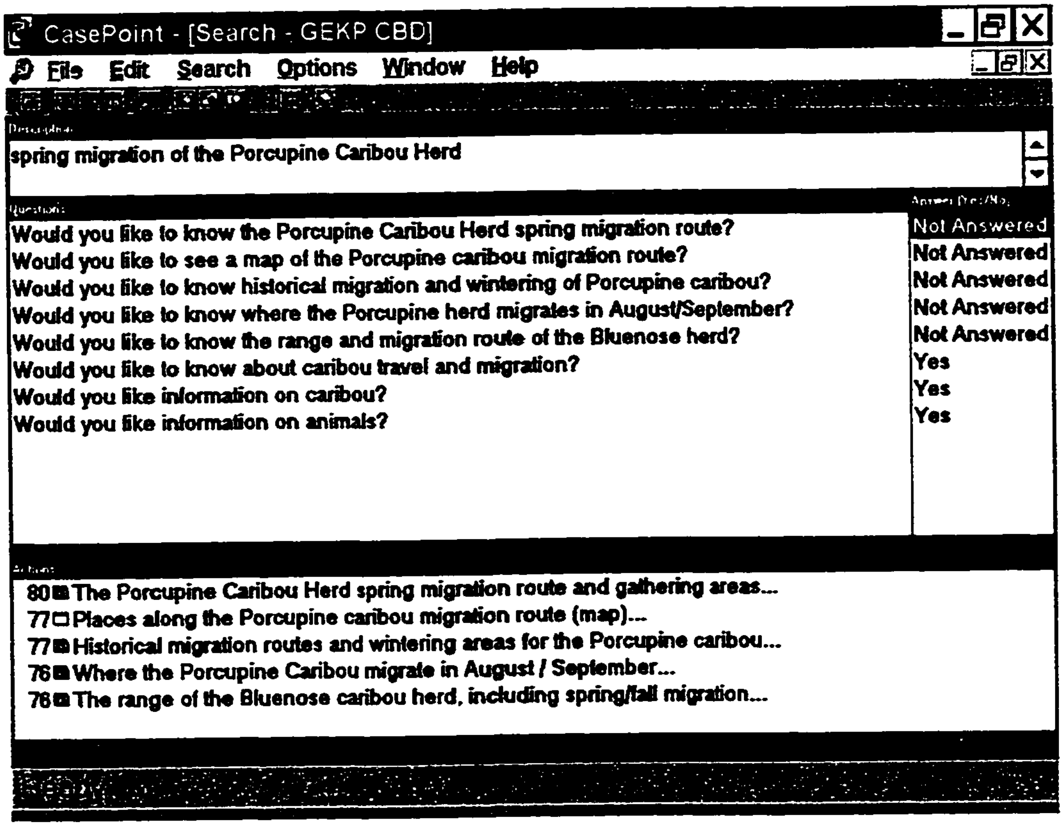

Figure 14. A user is presented with two high-scoring answers to their initial search; one a text description and one a mapped image (the first two actions shown with scores of 80 and 77 respectively).

These types of questions do not have to be answered in the CBR; the more general textual description will suffice. However, in the context of mapped images, GIS and resource management, these are critical questions to address. When compiling interview data it is very simple to move from names mentioned in the interviews, to labelling points on a map. To infer the extent of the area to which these points relate is, however, impossible. The Porcupine River is easy to find on a map, but the extent of good hunting areas can only be discerned by the elders. Using names mentioned in interviews to label points on a map presents a false image that does not fully represent the depth of the elder's knowledge.

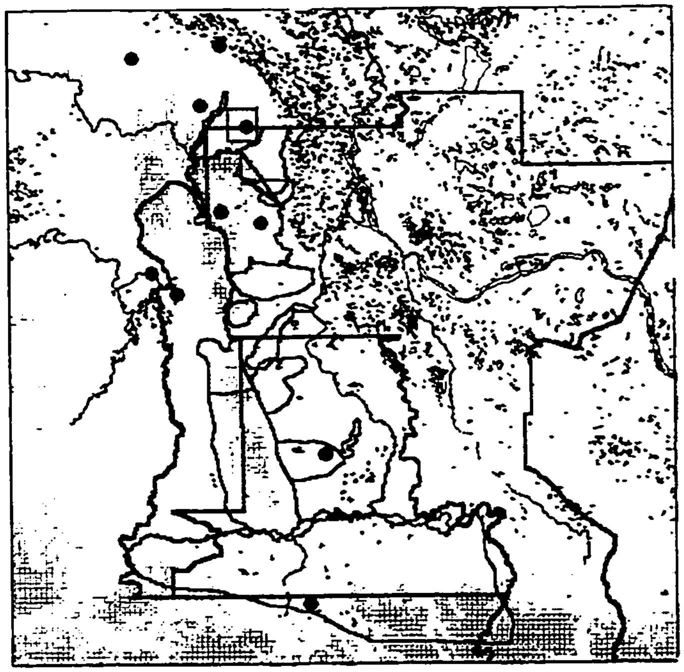

Figure 15. An image created with a GIS. then attached to a CBR action.

Resource managers require more precise spatial information. If a development is proposed in a certain area, a manager needs to know TEK related to the area, not just about a point which may or may not lie within that area. Because of problems with the spatial information in the original data, no scales or legends were added to the images produced as this may lend a false sense of accuracy to the data, as in Figure 15. In this research, integrating the CBR and the GIS image allows the development of a methodology and provides a framework to examine the tools, but it does not provide solid spatial information.

The second difficulty encountered with attaching 'browsable' mapped images relates to access. A user's access to the images is limited because those images are somewhat hidden. The mapped image is always the end result of the search: it is the solution to the user's question and therefore the last item the user sees. If the user starts the search looking for descriptive traditional knowledge about caribou migration, the spatial images are viewed as a solution. However, if the user starts the search by identifying the spatial distribution of wildlife in a GIS or on a mapped image, how then do they find any descriptive information that may relate to the area? The search can lead from descriptive information to a mapped image, but not from an image to descriptive information. The Internet may provide a solution to the problem by providing a platform to begin a full-scale GIS query and end with a case description of an area.

Finally, the utility of attaching 'browsable' images is limited because the images are simply flat records or digital maps created by an author. Each image must be created and stored individually. In image development an author cannot anticipate the needs of the user and ensure that all the layers of information a user requires is in that image. The author has a choice to include any variety of layers, but it is impossible to prejudge the needs of different users. If an image meets the needs of the user, it can only be by chance. The users themselves will be frustrated. For a truly successful integration between the CBR and GIS data, the user must have access to a fully functional GIS to have the freedom to choose the information they require. Again, the Internet may provide a solution by providing a platform to begin a full-scale GIS query.

7.2: Internet Access to the CBR and the GIS

The Internet, rather than a stand-alone machine, offers a better platform to link GIS data with Gwich'in knowledge in the case base. These links however, will not be developed within the context of this research. The Internet-based integrated resource management database within the Gwich'in organization is still in the planning stage. Although the impetus and the desire to share data across the organization by way of the Internet is a very definite move the Gwich'in will make in the near future, the direction has yet to be taken. As such, any links developed for this research would be premature, as was identified in Section 6.5, 'Using CasePoint WebServer to Develop Internet Access'. A web based GIS application will certainly be a central feature of the integrated resource management database, but issues of access, cost, content and usability continue to be debated within the settlement area. Consequently, this research can look toward future linkages between CasePoint WebServer and a web based GIS application, but those developments cannot currently be explored.

In looking forward to what the Internet offers, there are many potential linkages between a web based CBR and GIS, each with advantages and disadvantages. Ideally, a resource manager should be able to query any area in a GIS and have traditional knowledge pertinent to the area immediately accessible. Conversely, the manager should also be able to query the traditional knowledge base and have related GIS coverages and analysis tools available. If, for example, a manager is responsible for the preliminary screening of a development proposal concerning the Rat River, the manager should have a number of avenues open to him/her. First, s/he should be able to view the area of concern on a GIS layer, then query the system to display any related traditional knowledge of the area. Alternatively, s/he should be able to search the traditional knowledge base for information relating to Arctic char, then view the identified char-related areas in a GIS. This research is striving to achieve these goals and the Internet may offer the best possibilities to achieve them.

Two complementary developments are possible. The first is to link attribute files in the GIS to solutions in the CBR. The second is to link solutions in the CBR to attributes in a GIS image. In the first suggestion, each GIS layer has an associated attribute table which conventionally contains a limited number of characters in each record, in a virtually unlimited number of fields. One of the attribute fields can contain records which offers hyper-links to the solution or solutions in the TEK base. This solution can further guide the user to search the traditional knowledge base for more and relevant TEK. For example, associated with specific locations on the Rat River may be traditional information about char spawning. By querying the Rat River location in the GIS, the user can go to the CBR solution where the elders have referred specifically to char spawning and can follow hyper-links to the original interviews. Whilst searching the CBR, it is now incumbent upon the user to investigate any other links of interest. For instance, in this example it may be pertinent to search information on the importance of char as a food source for the Gwich'in.

A complementary development could link solutions in the CBR to attributes in a GIS layer or file. This would be similar to the process just described, but working in the opposite direction. Rather than initiating their query in the GIS, the user could first search the traditional knowledge base for information related to char spawning on the Rat River, and go from this solution to a GIS image which shows the spawning areas. Once launched into the GIS application, the user now has the power to use GIS analysis tools and the flexibility to include and overlay other data. Between these two links the user should have complete access to both databases as well as both applications.

However, difficulties may arise with the development of the links. Technically, the flexibility of hyper-links and the language standardization between applications and pages on the Internet will offer more potential and better compatibility than would be possible on a stand-alone machine. Nevertheless, all of the links, whether they link CBR solutions to the GIS or vice versa, have to be developed manually and continually upgraded to accommodate new and changing data. This will certainly require an initial time investment, particularly to link the vast amount of traditional knowledge data already collected. The benefits, in the end, may outweigh the cost but both must be carefully considered.

Difficulties may also arise in moving the user between unfamiliar applications. The Internet does offer the benefit of quickly guiding a user between very different programs and databases, while always maintaining a familiar environment for the user. Regardless of the application, the user is always viewing the information through a web browser which has consistent, universally recognized options. The user is never faced with an entirely new screen in an entirely new program. If it is a well designed application, the experience for the user will be just be like 'surfing the net'. Nonetheless, the user will, at the click of a button, be faced with very different information in a very different application when they move between CasePoint WebServer and a GIS. In the absence of good design features, the user may simply become frustrated and 'lost' in the new environment.

The problems associated with developing links and using unfamiliar applications may be answered by the creation of a 'front-end' interface. The user would conduct all their queries on a single web page. Behind that web page both CasePoint WebServer and a GIS application would be accessible to the 'front-end' interface. Any query made on the 'front-end' would be sent to either, or to both applications and would then return the results to the single 'front-end' interface. The user would have the benefit of having the data from both applications and the benefit of only interacting with one interface. The developers would benefit by not having to establish each and every potential link manually. The cost, of course, comes with the development. Developing a front-end interface to exchange data between both applications would require a person well versed in Internet application design. Again, the benefits accrued may be worth the cost of design and would have to be considered carefully.

7.3: Difficulties of Integrating GIS on the Internet and on a Stand-Alone Machine

In any of the above mentioned possibilities for integrating information from the traditional knowledge base with GIS, whether on a single machine or on the Internet, two difficulties consistently arise: first, with the original data itself and second, with the levels of uncertainty. The knowledge originally recorded for the GEKP did not always include spatial information and is therefore difficult to integrate consistently with GIS data. In the interviews, wildlife knowledge is only interspersed with specific place names. Because interviews were not conducted with the aid of maps, the completed interviews simply provide a list of place names related to each species. For example, the specific points and place names related to caribou migration can be mapped, but these points do not show the actual corridors and routes used by the caribou. The elders certainly have an intimate knowledge of these routes, but the spatial illustrations gleaned from the interviews do not successfully display that knowledge (see previous section).

How then do you query a specific area in the GIS and maintain consistent links to descriptions? Such a query is not possible if the areas are only spatially described by points. The series of points and place names currently in place need to be replaced with more exhaustive spatial data that display information such as caribou migration routes. Until this more exhaustive information is recorded, the links between the traditional knowledge base and the GIS will be sporadic and inconsistent.

The second difficulty in incorporating GIS with a traditional knowledge base revolves around the question of uncertainty, or how close is close enough? If caribou migration corridors based on the elder's knowledge were recorded and a development was proposed just beyond the boundary of that corridor, is the knowledge base information regarding migration pertinent? In real terms, when the resource manager reviews the proposed development and queries the GIS for information regarding that area, should the knowledge about migration corridors be displayed only if the development is within one kilometre of the boundary, or only if the development is within ten or fifteen kilometres of the boundary? The answer obviously depends on the sensitivity of the area and the nature of the proposal, but the difficulty in designing the system is to determine how much freedom is given to the manager and how much emphasis should the system designer place on the establishment of buffer zones. The notion of unclear boundaries is addressed in the GIS literature by fuzzy modelling (Borrough and McDonnell, 1998). Although the question cannot be answered definitively here, it is an important one to raise in the development of an integrated resource management system.

CHAPTER 8: CONCLUSIONS AND FUTURE IMPLICATIONS

Undoubtedly, First Nations, including the Gwich'in, need to continue to develop new and innovative ways to incorporate TEK into their management strategies. The goal of placing TEK on an equal footing with western science is often elusive as the many barriers to integration attest. Western science is certainly adapting quickly to the new 'digital' age and computer tools to aid resource management decision making abound. The representation of TEK is 'hot on the heels' of this new trend, as GIS-based land use mapping gains strength and organizations, like the GRRB, strive to develop information retrieval systems to increase access to recorded TEK. This research sought to advance those goals by:

1. developing a computer-based system, to make TEK more available, that is easy to author, use and upgrade, while maintaining the words and context of TEK experts, as much as possible;
2. making it accessible within a framework of tools already used by the Gwich'in boards, primarily through GIS integration and Internet accessibility.

The goal was not to 'integrate' TEK with western science, but simply to provide practical and equitable access. While the objectives of the research were successfully met, limitations and costs were also identified. Although the prototype developed for this research highlights the potential of a CBR system in TEK representation, future investment to improve and refine the system is required.

CBR tools were chosen for the research for three reasons: their basic design philosophy, user-friendly environment and advantages over other information retrieval systems. The goal of a CBR is to allow non experts to solve new problems based on the previous experience of experts. Here, the elders are the experts and the resource managers are seeking the knowledge of the elders. CBR operate in a windows-based environment and can be authored and searched with no knowledge of programming, and thus are user-friendly. Finally, CBR have an advantage over other information retrieval and database software systems because of their interaction with the user and ability to remember and to learn. The only drawbacks to CBR are their cost and uniqueness in this domain. CBR have not been tested nor proven in their ability to represent TEK and are therefore a more potentially costly undertaking at this point because no generally agreed upon methodologies have been established. At this point because of the risk involved in CBR development, it perhaps requires more extensive university-based research in TEK representation rather than research at the Gwich'in boards.

To summarize this research, the CBR authoring process focussed on the GEKP final report. The report is a peer-reviewed document so it became the primary information source supported by original interviews. As the GEKP was conducted on a species by species basis, so the case base was developed in the same manner. Manual case authoring, rather than automatic authoring, was found to be the most valuable.

The prototype developed for this research covers the breadth and depth of the knowledge included in the GEKP. To assist in the process of organizing the knowledge and the interviews, a hierarchical decision tree model was developed as a precursor to case base authoring. A hierarchical model is a successful tool here because of a user's ability to view similar top-scoring answers and the ease of launching new searches. As such, a hierarchical model does not preclude the provision of contextual knowledge in the case base. While this structure does raise questions about the adaptability of the model to include new and different information in the future, the flexibility of CBR software tools indicate these problems can be overcome.

The second issue raised by model development relates to the replicability of the model. While the model was developed by a western-based researcher it can be expected that another researcher, western-based or not, would develop the model quite differently. Regardless of the model structure however, if the end result is to guide a user to the information for which they are searching, then the model is successful in meeting its objectives.

All case titles, descriptions, questions and actions are designed using common language to ensure the case base is not exclusive. Cases are difficult to identify because first, there are no explicit problems and solutions with TEK and second, the separation of the knowledge into cases runs the risk of de-contextualizing and compartmentalizing the knowledge. This risk is successfully avoided however, with the presentation of closely related cases for the user to view and/or the presentation of the original interviews for the user to browse.

Balance and simplicity are achieved in the model by sharing questions as much as possible, by having the same number of questions associated with each case and by trying to establish a parallel questioning structure across the knowledge base. Questions are evenly divided between context and confirmation questions, with each case being equally comprised. The questioning structure provides an exceptional tool for the user to navigate efficiently the knowledge domain, and the use of 'Yes-No' questions provides the author with an effective tool for predicting potential search interactions. Additionally, a rule file was developed to answer context questions where possible, thereby further streamlining searches for the user.

With the prototype case base, searches can be conducted on a stand alone machine using CasePoint, or on the Internet using CasePoint WebServer. A case base accessible via the Internet has significant advantages over a stand-alone case base. First, information and data exchange via the Internet is an increasingly attractive option within the Gwich'in organization and an integrated management system, which encompasses Internet data accessibility, has been proposed. Second, the incorporation of the original interview material has greater potential on the Internet and therefore offers greater potential for an acceptable representation of TEK. CasePoint WebServer allows multiple hyper links to be associated with each action to guide a user back to the section of the original interviews which contributed to that answer.

Perhaps the most significant advantage of WebServer over CasePoint will be fully realized in the future with further integration of GIS. A search using CasePoint yields links to mapped images created in a GIS. However, it does not offer the user full GIS analysis capabilities and is therefore limited as a resource management tool. Links from the CasePoint WebServer application to a web-based GIS application may offer greater potential for developing a truly integrated resource management tool that incorporates a CBR and a GIS. Potential linkages have not been fully explored at this time primarily because the Gwich'in Internet-based GIS and data management tools are only beginning to take shape. The goal of the CBR is, of course, to operate in concert with these tools and it is therefore imperative that each be developed first, then integrated.

The exploration of GIS integration is also constrained because of limited information. While the GEKP interviews were conducted on a species by species basis, very little explicit and definitive locational information was recorded. This step must be taken for any extensive integration of TEK with GIS. The elders must record on maps the spatial information that complements their vast knowledge on species and species interactions. Exciting avenues, such as fuzzy modelling and object-oriented tools, are available which offer excellent representation possibilities for TEK within GIS. These tools may provide context and holism for TEK representation that is absent from conventional GIS representations, and may provide a more accurate reflection of a TEK viewpoint.

Further, CBR are a promising tool to be used in other areas of resource management, particularly those with easily identifiable problems and solutions. It offers the burgeoning field of resource management the ability to collect and provide access to 'corporate' knowledge and to provide direct solutions to similar problems previously encountered by other resource managers. With the increasing reliance on GIS as a tool in resource management, CBR also offer a potential to expand the explanatory capabilities of GIS by allowing textual information to be juxtaposed with spatial information. This is a new area that deserves future research.

The use of CBR to represent TEK is unique and offers great potential as a resource management tool. Owing to that same uniqueness, however, are the costs involved with venturing into new and unexplored territory. Potential costs may be too great to incur at the level of the Gwich'in boards, but do point to the need for continued research in this area in order that risks may be reduced.

Regardless of where future research is conducted into TEK representation in information systems, the position of this type of research must be viewed in context. Information systems are one potential stepping stone along the path to full integration of TEK with resource management. Continued research and effort made in the direction of TEK education and documentation, coupled with efforts to allow western scientists and traditional knowledge experts to discuss issues on a personal and equitable level are necessary for the maintenance of TEK and its inclusion in resource management.

REFERENCES

Aha, David and Breslow, Leonard. 1997. "Refining Conversational Case Libraries", in David Leake, and Enric Plaza eds., *Case-Based Reasoning Research and Development: Second International Conference on Case-Based Reasoning, ICCBR-97, Providence, RI, USA, July 25-27, Proceedings.* Lecture Notes in Artificial Intelligence, No. 1266. New York: Springer-Verlag. pp. 267-278.

Aamodt, Agnar. and Plaza, Enric. 1994. "Case Based Reasoning: Foundational Issues, Methodological Variations and System Approaches". URL: http://www.mines.unancy.fr/~guenniffe/CoursEMN/131/AICom/AICom-ToC.html

Arctic Environmental Protection Strategy (AEPS). 1993. *The Participation of Indigenous Peoples and the Application of their Environmental Knowledge in the Arctic Environmental Protection Strategy: Volume 1: A Report on Findings.* Prepared by the Inuit Circumpolar Conference for Indian and Northern Affairs, Canada. August 1993.

Barletta, Ralph. 1991. "An Introduction to Case-Based Reasoning". *AI Expert*, vol.6 (8): 42-49.

Berger, Thomas. 1977. *Northern Frontier, Northern Homeland: The Report of the Mackenzie Valley Pipeline Inquiry.* Toronto: Douglas & McIntyre.

Berkes, Fikret. 1993. "Traditional Ecological Knowledge in Perspective." In *Traditional Ecological Knowledge: Concepts and Cases*, ed. J.T. Inglis. Ottawa: International Program on Traditional Ecological Knowledge and International Development Research Center: 1-9.

Berkes, Fikret and Henley, Thomas. 1997. "Co-Management and Traditional Knowledge: Threat or Opportunity?" *Policy Options*, (March): 29-31.

Bird Rose, Deborah. 1988. "Exploring an Aboriginal Land Ethic." *Meanjin* 41: 378-387.

Brown, John Epes. 1973. "Modes of Contemplation Through Actions: North American Indians." *Main Currents in Modern Thought* 30: 192-197.

Burrough, Peter and McDonnell, Rachel. 1998. *Principles of Geographical Information Systems.* Oxford: Oxford University Press.

Campbell, Tracey. 1996. *Aboriginal Co-management of Non-renewable Resources on Treaty or Traditional Territory.* Master of Arts Thesis, Committee on Resources and the Environment, University of Calgary.

Capra, Fritjof. 1982. *The Turning Point: Science, Society and the Rising Culture.* New York: Bantam Books.

Clayton, D. 1994. *An Assessment of the Porcupine Caribou Herd Harvest Reporting System in Alaska.* Unpublished Paper. Fairbanks, Alaska: Alaska Department of Fish and Game, Division of Subsistence.

Cronon, William. 1983. *Changes in the Land: Indians, Colonists, and the Ecology of New England.* New York: Hill and Wang.

Cruikshank, Julie. 1981. "Legend and Landscape: Convergence of Oral and Scientific Traditions in the Yukon Territory." *Arctic Anthropology* XVIII, No. 2: 67-93.

Deloria, Vine. 1973. *God is Red.* New York: Grosset and Dunlap.

Everitt, Robert. 1986. "Native Roles in Monitoring of Energy Developments." In *Native People and Renewable Resource Management, Symposium of the Alberta Society of Professional Biologists, 29 April - 1 May.* Edmonton: Alberta Society of Professional Biologists: 44-48.

Feit, Harvey. 1988. "Self Management and State Management Forms of Knowing and Managing Northern Wildlife." In *Traditional Knowledge and Renewable Resource Management*, eds. M. Freeman and L. Carbyn. Edmonton: Boreal Institute for Northern Studies: 72-91.

Freeman, Milton. 1985. "Appeal to Tradition: Different Perspectives on Arctic Wildlife Management." In Native Power: *The Quest for Autonomy and Nationhood of Indigenous Peoples*, eds. J. Brosted, et al. Norway: Universitetsforlaget:265-282.

— 1986. "Renewable Resources, Economics and Native Communities." In *Native People and Renewable Resource Management, Symposium of the Alberta Society of Professional Biologists, 29 April - 1 May.* Edmonton: Alberta Society of Professional Biologists: 85-93.

— 1992 "The Nature and Utility of Traditional Ecological Knowledge". *Northern Perspectives* 20, No. 1 (Summer): 9-12.

Grose, E., Forsythe, C., and Ratner, J. 1998. "Using Web and Traditional Style Guides to Design Web Interfaces." In *Human Factors and Web Development*, eds. Forsythe, Grose and Ratner. Lawrence Erlbaum Associates: New Jersey, 121-131.

Gunn, A., Arlooktoo, G., and Kaomayok, D. 1988. "The Contribution of the Ecological Knowledge of Inuit to Wildlife Management in the Northwest Territories." In *Traditional Knowledge and Renewable Resource Management*, eds. M. Freeman and L. Carbyn. Edmonton: Boreal Institute for Northern Studies: 22-30.

Gwich'in Elders. 1997. *Nanh' Kak Geenjit Gwich'in Ginjik: Gwich'in Words About the Land.* Inuvik: Gwich'in Renewable Resource Board.

Gwich'in Land and Water Board. 1996. *Development of an Integrated Resource Management System/Framework for the Gwich'in Settlement Area: The Land Claim Agreement and Management Structures. Discussion Paper 1.* Inuvik.

Gwich'in Renewable Resource Board. 1998. *Gwich'in Renewable Resource Board Strategic Planning Report.* Inuvik.

Gwich'in Tribal Council. 1992. *Gwich'in Comprehensive Land Claim Agreement, Volume 1.* Ottawa: Gwich'in Nation and Indian and Northern Affairs Canada.

Howard, Albert and Widdowson, Frances. 1996. "Traditional Knowledge Threatens Environmental Assessment." *Policy Options*, November: 34-36.

Hughes, J. Donald.1983. *American Indian Ecology.* The University of Texas at El Paso: Texas Western Press.

Inference Corporation. 1995a. *CBR Express for Windows: Designing Case Bases.* El Segundo, California: Inference Corporation.

— 1995b. *CBR Express for Windows: User's Guide.* El Segundo, California: Inference Corporation.

— 1995c. *CBR Generator for Windows: User's Guide.* El Segundo, California: Inference Corporation.

— 1995d. *CasePoint for Windows: User's Guide.* El Segundo, California: Inference Corporation.

— 1995e. *CasePoint for Windows: Adding Rules to CasePoint*. El Segundo, California: Inference Corporation.

— 1995f. CasePoint WebServer Product Summary. URL: http://www.inference.com/products/webserverps.html

— 1995g. *CBR Express for Windows: Application Programming Interface User's Guide*. El Segundo, California: Inference Corporation.

Johnson, Martha. 1992. "Dene Traditional Knowledge" *Northern Perspectives* 20, No.1 (Summer): 3-5.

Johnson, Martha and Ruttan, R. 1993. *Dene Environmental Knowledge A Pilot Project Conducted in Fort Good Hope and Coville N.W.T., 1989-1993.* Report prepared for the Dene Cultural Institute. Hay River: Dene Cultural Institute.

Kawagley, Oscar. Personal Communication. Faculty of the University of Alaska, School of Education, Anchorage, Alaska.

Kawagley, Oscar. 1995. *A Yupiaq Worldview: A Pathway to Ecology and Spirit*. Illinois: Waveland Press Inc.

Keith, Robert. 1994. "The Ecosystem Approach, Implications for the North." *Northern Perspectives* 22, No.1 (Spring): 3-6.

Kolodner, Janet. 1995. "A Case-Based View of Case-Based Reasoning: What Have We Wrought?". College of Computing, Georgia Institute of Technology. URL: http://www.cc.gatech.edu/cogsci/edutech/cases&problems/what_have_we_wrought.html

Laghi, Brian. 1997. "Tempest in a Teepee: Getting Into the Spirit of Things". In *The Globe and Mail*. Saturday, August 9: D1.

Leake, David. 1996. "CBR in Context: The Present and Future." In *Case-Based Reasoning: Experiences, Lessons and Future Directions,* ed. David Leake. Massachusetts: American Association for Artificial Intelligence and MIT Press: 3-30.

Lloyd, Kevin. 1986. "Co-operative Management of Polar Bears on Northeast Baffin Island". In *Native People and Renewable Resource Management, Symposium of the Alberta Society of Professional Biologists, 29 April - 1 May*. Edmonton: Alberta Society of Professional Biologists: 108-116.

Mark, William, Simoudis, Evangelos and Hinkle, David. 1996. "Case-Based Reasoning: Expectations and Results." In *Case-Based Reasoning: Experiences, Lessons and Future Directions,* ed. David Leake. Massachusetts: American Association for Artificial Intelligence and MIT Press: 269-294.

McDonald, Miriam. 1988 "An Overview of Adaptive Management of Renewable Resources." In *Traditional Knowledge and Renewable Resource Management*, eds. M. Freeman and L. Carbyn. Edmonton: Boreal Institute for Northern Studies: 65-71.

Murdoch, John. 1988. "Cree Cognition in Natural and Educational Contexts." In *Indigenous Cognition: Functioning in Cultural Context*, eds., J.W. Berry, S.H. Irvine, and E.B. Hunt. Dordrecht: Martinus Nijhoff Publishers.

Nakashima, Douglas. 1990. *Application of Native Knowledge In EIA: Inuit, Eiders and Hudson Bay Oil.* Canadian Arctic Environmental Assessment Research Council.

Osherenko, Gail. 1988. *Sharing Power With Native Users: Co-management Regimes for Arctic Wildlife.* Canadian Arctic Resources Committee Policy Paper, No.3. Ottawa: Canadian Arctic Resources Committee Publishing Program.

Ratner, Julie. 1998. "Easing the Learning Curve for Novice Web Users", in Forsythe, Grose and Ratner, eds., *Human Factors and Web Development.* Lawrence Erlbaum Associates: New Jersey: 111-118.

Riesbeck, Christopher. 1996. "What Next? The Future of Case-Based Reasoning in Post-Modern AI." In *Case-Based Reasoning: Experiences, Lessons and Future Directions,* ed. David Leake. Massachusetts: American Association for Artificial Intelligence and MIT Press: 371-388.

Riewe, R. and Gamble, L. 1988. "The Inuit and Wildlife Management Today." In *Traditional Knowledge and Renewable Resource Management*, eds. M. Freeman and L. Carbyn. Edmonton: Boreal Institute for Northern Studies: 31-37.

Roberts, Karen. 1994. *Co-Management: Learning from the Experience of the Wildlife Management Advisory Council for the Northwest Territories.* Master of Environmental Design (Planning) Project, University of Calgary.

Sallenave, John. 1994. "Giving Ecological Knowledge Its Rightful Place in Environmental Impact Assessment." *Northern Perspectives* 22, No. 1 (Spring):16-19.

Schaeffer, et al. 1986. *A Review of the Game Regulations Affecting Northwest Alaska.* Kotzebue: Kotzebue Fish and Game Advisory Committee.

Seminar on the Documentation and Application of Indigenous Knowledge. 1996. *Recommendations on the Integration of Two Ways of Knowing: Traditional Indigenous Knowledge and Scientific Knowledge*, November 15-17. Inuvik, Northwest Territories.

Shaw, Mildred and Gaines, Brian. 1986. "Interactive Elicitation of Knowledge from Experts." *Future Computing Systems* Vol. 1, No. 2: 151-190.

Strachan, A.M. 1988. *The Truth is Changing All the Time: An Exploration into the Nature of Inuit Logic*. Unpublished Major Paper. University of Guelph: University School of Rural Planning and Development.

Tobias, Terry. 1993. "Stereotyped Village Economies and the Pinehouse Harvest Research". In *Traditional Ecological Knowledge: Concepts and Cases*, ed. J.T. Inglis. Ottawa: International Program on Traditional Ecological Knowledge and International Development Research Center: 87-97.

Usher, P.J. 1986. *The Devolution of Wildlife Management and the Prospects for Wildlife Conservation in the Northwest Territories*. Canadian Arctic Resources Committee Policy Paper, No.3. Ottawa: Canadian Arctic Resources Committee Publishing Program.

— 1993. "The Beverly-Kaminuriak Caribou Management Board:: An Experience in Co-Management". In *Traditional Ecological Knowledge: Concepts and Cases*, ed. J.T. Inglis. Ottawa: International Program on Traditional Ecological Knowledge and International Development Research Center: 111-118.

Veitch, Alasdair. 1997. Personal Communication. Supervisor Wildlife Management, Sahtu Region. Department of Resources, Wildlife and Economic Development, Norman Wells, NWT.

Waters, Nigel. 1989. "Expert Systems within a GIS: Knowledge Acquisition for Spatial Decision Support Systems". *Proceedings of the National Conference for Geographic Information Systems*. Ottawa: The Canadian Institute for Surveying and Mapping: 740-759.

Watson and Chambers. 1989. *Singing the Land, Signing the Land.* Victoria, Australia: Deakin University Press.

Watson, Ian. 1995. "Case-Based Reasoning Development Tools: A Review." URL:http://www.salford.ac.uk/survey/staff/IWatson/cbrtools.html

— 1995. "The Case for Case-Based Reasoning" URL:http:/146.87.176.38/ai-cbr/Papers/ita01.htm

Watson, Ian and Marir, Farhi. 1994. "Case-Based Reasoning: A Review". *The Knowledge Engineering Review*, Vol 9(4): 327-354.

Wolfe, J., Bechard, C., Cizek, P., and Cole, D. 1992. *Indigenous and Western Knowledge and Resources Management System.* First Nations Series. University of Guelph: University School of Rural Planning and Development.

World Commission of Environment and Development. 1987. *Our Common Future*. New York: Oxford University Press.

Young, D., Ingram, G., and Swartz, L. 1989. *Cry of the Eagle: Encounters with A Cree Healer*. Toronto: University of Toronto Press.

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