# FOCES: An Experimental Expert System to Select Appropriate Foster Care Homes for Children

by

Sheila Ginsberg Winett

Thesis submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

**Master of Science**

in

Computer Science

**Approved:**

Edward A. Fox, Chairman

Ezra A. Brown

John W. Roach

February, 1987

Blacksburg, Virginia

# FOCES: An Experimental Expert System to Select Appropriate Foster Care Homes for Children

by

Sheila Ginsberg Winett
Committee Chairman: Edward A. Fox
Computer Science

## ABSTRACT

The FOster Care Expert System (FOCES) was developed to provide advice to social workers of the Roanoke City Department of Social Services who must select foster care homes for children who cannot remain with their own families. It was implemented using the General pUrpose Expert Shell System (GUESS) and Horn Clause Prolog. The system's design was greatly influenced by unique features of the problem domain. Among the key concerns were: unresolved questions within the social work profession about foster home selection and evaluation, serious methodological and philosophical difficulties associated with defining a good "person-environment fit", and the volatile, free-form narrative nature of the information maintained by social services agencies about children and homes. "Traditional" approaches to knowledge acquisition and representation adopted by developers of expert systems were of limited use. Adaptation of extended "p-norm" Boolean queries previously used in information retrieval work simplified the knowledge representation and matching tasks for this human services application. Evaluation of FOCES' performance, using a small database of children and homes, has shown that the system can select appropriate foster care placements at least as well as some experienced social workers.

# Acknowledgements

The work described in this thesis could not have been done without the assistance of many people. My personal thanks to the following members of the Roanoke City Department of Social Services staff: Mr. James Ritchey, Ms. Doris Kinsey, Ms. Carolyn Doyle, Ms. Judy Brown, Ms. Joyce Minnick, and Ms. Layne Canada. Without the continuing interest and assistance of these individuals, this project could never have been done. Week after week they gave of their time so that work on this project could proceed.

The contributions of Dr. Edward A. Fox, my committee chairman and advisor, and Michael Furey, B.A. in Computer Science, Virginia Tech, Class of 1986, deserve special thanks. Michael's work to implement the matching algorithms used by FOCES to select foster care homes for children was most helpful. The continuing encouragement and patience shown by Dr. Fox, in the face of innumerable delays and problems, at times was one of the few barriers to giving up and starting over with a new topic!

My thanks also to my committee members, Dr. Roach and Dr. Brown, who have shown continuing interest in my progress, and to Dr. Linda Shapiro, whose early belief in my ability, helped me over some rough spots.

Finally, I would like to thank my family, who has supported me from day one of this endeavor. Without them, none of this would mean anything.

# Table of Contents

# List of Figures

# List of Tables

# 1. INTRODUCTION

An expert system is a computer program that performs a narrowly defined and complex task which is ordinarily performed by a human expert [RICHa]. This thesis describes the research done to develop and implement one such system, the FOster Care Expert System (hereafter referred to as FOCES).

FOCES was developed over a two year period in cooperation with the Roanoke City Department of Social Services. Its task is to provide advice to social workers who must either:

- select an initial foster care placement for a child who is no longer able to remain with his/her biological parents or legal guardians, or
- select a new foster care placement for a child whose current placement has broken.

To accomplish this task, FOCES must suggest an appropriate match between a child and foster care home.

Foster care services are provided by designated social services agencies to a child and his/her family to prevent removal of the child from the family home. However, when parents are unwilling or unable to use these services to change conditions in the family home, then placement of the child in a foster home may be sought by the agency empowered to address child welfare cases in the locale in which the child lives at the time the situation arises [VDSS].

When placement in a foster care home is necessary, the social worker assigned to investigate the situation – the individual directly responsible for protecting the child's health and welfare – selects the home. That choice is made in consultation with other social work professionals and appropriate supervisory personnel. The decision is constrained by state and federal laws and administrative regulations, heuristics of 'good social work practice', and local exigencies.

Few, if any, expert systems have been developed to address social services problems [SCHOECH]. As a consequence of how agencies that provide social services to children and families work and the nature of the problems that those agencies address, immense and unusual knowledge acquisition and data representation problems occur. For example, working with data in the form that they are typically maintained by those agencies about the entities of interest is especially difficult. Because of these circumstances it seemed appropriate to make use of information retrieval methods as an integral part of FOCES.

In the past, efforts have been made to apply artificial intelligence techniques to the design of information retrieval systems [YIP[MCCUNE]. However, the converse -- adapting information retrieval techniques to solve an artificial intelligence problem -- has not usually been thought of. Chapter 2 of this thesis will discuss the relationship between this project and previous work in artificial intelligence and information retrieval and will address the domain-specific problems that were encountered in pursuing this research. Chapter 3, System Implementation, examines how those domain-related issues were dealt with and describes the implementation methodology used. The results of the project and suggestions for future research are discussed in Chapters 4 and 5, respectively.

The central hypothesis addressed by this research concerns the applicability of extended "p-norm" Boolean queries as a method of selecting a foster home which can be expected to meet the needs of a child requiring placement. P-norm queries had previously been used to express user requests in information retrieval systems. This research sought to test their usefulness in an altogether different domain. The developers of FOCES hypothesized that a system using p-norm queries to select homes for children would perform at least as well as social workers performing the same task. To test this hypothesis, an experimental expert system was developed and used to match children with potentially available homes. The results obtained were then evaluated. Using a small test database of children and foster care homes, the recommendations developed by FOCES were compared to the recommendations made by two social workers. The results of these tests, described fully in Chapter 4, indicate that FOCES performed as well as the developers had hypothesized it would.

# 2 PROJECT BACKGROUND

## 2.1 Relationship to Other Research

The development of expert systems has been an active and visible subspecialty of artificial intelligence research since the mid-1960s [HAYES-ROTH]. Table 1 highlights some of the domains to which this technology has been applied.

None of the systems listed has been developed for the social services domain. Nonetheless, some of the systems perform diagnostic tasks that are functionally related to the match problem addressed by FOCES.

FOCES' primary goal is to find an appropriate match between a child and a foster care home. For each child with particular characteristics and needs, there is a formal representation, C. Similarly, each potential foster home, $H_i$, has a detailed description. FOCES aims to select all homes such that MATCH(C,$H_i$ ) is acceptable and to rank those homes on a number of dimensions so that the most appropriate homes for a child can be considered.

### 2.1.1 Medical Diagnosis

One can emphasize the conceptual relationship between the task performed by FOCES and the medical diagnosis task performed by some other expert systems. By describing medical diagnosis in language that is similar to that used to describe FOCES' task, the relationship between the two is highlighted. For example, the medical diagnosis task might be formalized as follows:

> Assign to each patient with a particular set of symptoms a formal representation, P. Assign to each potential diagnosis a formal representation, $D_i$. A medical diagnosis expert system aims to select a

# Table 1. Selected Domains for Expert Systems (adapted from HELLY)

| Function | Domain | System Name |
|---|---|---|
| Diagnosis | Medicine | CASNET |
| | Medicine | INTERNIST |
| | Medicine | MYCIN |
| | Medicine | PUFF |
| | Engineering | SACON |
| | Geology | PROSPECTOR |
| Search | Chemistry | DENDRAL |
| | Chemistry | SYNCHEM |
| Problem Solving and Planning | Mechanics | MECHO |
| | Programming | PECOS |
| | Configuring Computers | R1 |
| | | DEVISER |
| | Procedures | REF-ARF |
| Measurement Interpretation | Medicine | VM |
| Computer-aided Instruction | Electronics | SOPHIE |
| | Medicine | GUIDON |
| | Medicine | FLUIDMOD |
| Knowledge Acquisition | Diagnosis | TEIRESIAS |
| | Diagnosis | EMYCIN |
| | Diagnosis | EXPERT |
| | Diagnosis | SEEK |
| System Building | | ROSIE |
| | | AGE |
| | | HEARSAY III |
| | | XPLAIN |
| Temporal References | General Robots | CHRONOS |

diagnosis such that MATCH(P,$D_i$ ) is acceptable according to sound medical practice.

When the medical diagnosis task is formulated in these terms, the parallels between it and the task addressed by FOCES are obvious. INTERNIST, an expert system designed to make a diagnosis in the domain of internal medicine, originally used a related approach. It built up disease models dynamically and partitioned these groupings according to data about the patient's symptoms. Each grouping, i.e. constellation of idiosyncratic symptoms of a particular disease, was represented in a tree data structure. The match task performed by FOCES is also conceptually similar to the task performed by an information retrieval system.

### 2.1.2 Information Retrieval

Information retrieval systems are generally concerned with the representation, storage, organization, and accessing of information items [SALTON]. One specialized type of information retrieval system, referred to as a textual information retrieval or document retrieval system, selects appropriate documents from a large collection in response to a user query [FOX]. There are notable similarities between some of the tasks performed by a document retrieval system and the task addressed by FOCES.

Salton and McGill state [SALTON] that:

> "Every information retrieval system can be described as consisting of a set of information items (DOCS), a set of requests (REQS), and some mechanism (SIMILAR), for determining which, if any, of the information items meets the requirements of the requests."

One of the key problems in such a system is to find 'good' matches between a surrogate form of a person's query and items in the document collection. This task is very similar to the problem of matching a description of a child in need of foster care with descriptions of potentially available foster homes. The description of the child in need of placement may be conceptualized as the query, while the descriptions of potential foster homes play the role taken by the document collection in information retrieval. The key difference is that document retrieval can be based on simple similarity measures while selection of homes requires complex matching strategies.

While commercially available information retrieval systems typically employ rather unsophisticated selection methods, a great deal of expertise has been accumulated by the information retrieval research community. The use of extended Boolean logic queries to represent information requests has been well researched and implemented in at least two experimental information retrieval systems, SMART and SIRE [SALTON] [FOX]. One expert system has made extensive use of Boolean logic as a means of describing rules [HELLY].

FOCES makes use of p-norm queries to help formally represent children and homes. The early impetus for development of p-norm queries grew out of efforts to extend the power of Boolean queries used in document retrieval by incorporating the ideas of fuzzy set theory and permitting assignment of relative weights to query terms and clauses [FOX]. The formalism of the p-norm query has been adapted by FOCES to construct **prototypes**, that is models for classes of children and foster care homes.

### 2.1.3 Prototypes and Expert Systems

The use of prototypes, that is models or archtypes, has been suggested by others. However, p-norm queries have not previously been used as a representation for prototypical knowledge in another expert system. Some systems which have used prototypes are described in this section.

In the CENTAUR system twenty–four (24) prototypes are used to represent knowledge [AIKINS]. Twenty–one (21) represent stereotypical disease patterns in the pulmonary function domain. The remainder represent metalevel tasks such as consultation and explanation.

In the context of CENTAUR, a prototype is a type of frame [MINSKY]. Rules are one type of value for slots in a prototype. Other slots may contain a variety of data values. By explicitly providing slots in the prototypes for rules, CENTAUR blurs the distinction between 'data' and 'control' that is ordinarily part of an expert system's architecture [HAYES-ROTH]. Other systems, such as PIP [PAUKER] and INTERNIST [POPLE], also use frames to represent some kinds of prototypical disease models. However, unlike CENTAUR, those systems maintain a clear distinction between 'data' and 'control'.

Prototypes have also been used to represent hypothetical individuals who may be members of a class [BOBROW]. Rich suggests that the development of user models or prototypes may be an appropriate mechanism for tailoring the human-computer interface to meet the needs of individual users [RICHb].

Other research work of interest highlights the similarity between much of the work in information retrieval and a number of artificial intelligence tasks. For example, there are systems concerned with using artificial intelligence based techniques for natural language processing to analyze texts and automatically produce abstracts or summaries of input documents [LEHNERT] [DYER] [HAHN]. Other research has focused on applying artificial intelligence techniques to the problem of understanding and fulfilling a user's information request, a task that is more closely aligned with the concerns addressed by FOCES.

RUBRIC is an expert system that uses production rules to express a query used to retrieve documents [TONG]. Another rule-based system, EXPERT-1, models the intermediary role usually played by the human search expert to guide inexperienced users through the process of developing Boolean queries [YIP]. Several systems aim to rank documents with respect to their relevance to a user's initial information request [SALTON]. Recent work at Virginia Tech on the CODER system suggests a unified artificial intelligence-based approach that addresses all phases of the information retrieval process, from document analysis to user interface management [FRANCE].

## 2.2 DESCRIPTION OF THE DOMAIN

### 2.2.1 Foster Care in the United States

There are approximately 500,000 children in some form of foster care throughout the United States [SHYNE]. Further, foster care has an historic place in our human services delivery system, having developed out of the indenture system codified in the Elizabethan Poor Laws. It took its modern form under the leadership of Charles Brace in the mid-1850's [KADUSHIN].

Our society has a commitment to the provision of foster care as a short term alternative to permanent placement of a child, either back with his/her biological or legal parents or with a suitable adoptive family. The laws and administrative regulations that govern foster care programs customarily include language that is intended to guide those programs toward achieving permanency for a child. The following excerpt from the policy manual of the Virginia Department of Social Services is illustrative:

> The basic philosophy of the ... program is to maintain family unity and keep children in their own homes.... When parents are unwilling or unable to use these services to change conditions in the home, placement ... may be necessary. The objective of the program then becomes the provision of services to return the child home, or if this is not possible, to achieve another permanent home....

Despite the pervasiveness and historical significance of foster care, some key domain issues remain unresolved:

Is the foster home to serve a therapeutic role for the child? What characteristics must be present/absent from a particular home for it to be judged as 'good' for a particular child? How should a child's placement be evaluated?

As a result of these unresolved questions the development of FOCES has been difficult. Its developers have been required to rely on unconventional means to accomplish the knowledge acquisition tasks that conventional wisdom considers crucial to success, such as defining a benchmark for success of the system.

### 2.2.2 The Goal of FOCES

In most domains there are at least two levels at which one can expect some measure of consensus about goals -- philosophical and operational. The philosophical level refers to broad statements about the goals of professional activities. At the philosophical level, goals do not easily translate to action statements.

The philosophical level goal for foster care, i.e., to achieve permanency for a child, is quite clear and accepted by social services professionals. However, like other philosophical level goals, it lacks an operational component and, therefore, was deemed inappropriate as a metalevel goal for FOCES. While permanency may be the ultimate goal in this domain, its achievement or non-achievement is, in most cases, wholly independent of what a child may experience while he/she is in a foster care home.

The developers of FOCES had to construct their own definition of a successful placement: a successful placement is one that continues without any **unplanned** interruption caused by a foster parent requesting the child's removal. This definition was adopted for two reasons.

- The decision to place a child in a foster care home is, prima facie, a disorienting and disruptive experience. In addition to separating children from their parents, placement often separates children from their siblings, friends and classmates. FOCES assumes, therefore, that there is an intrinsic good attached to finding a foster care placement that carries with it a high probability of continuity.

- The foster care literature reveals that achievement of this goal is difficult. Most children in care for over one year will experience at least one unplanned move following initial placement. As a child stays in care longer, his/her chances of 'bouncing' from placement to placement increase dramatically [KADUSHIN] [OLSEN].

### 2.2.3 Information to Achieve the Goal

Achieving the goal established for FOCES by its developers required that the system be capable of discovering those patterns of child-foster home interactions that consistently result in a stable placement. That is, the system had to tackle the problem of person-environment fit -- how do the characteristics of the individual and his/her environment interact to affect personal well-being [CAPLAN].

The importance of the domain expert as a crucial source of domain knowledge required to build an expert system has been emphasized in the literature [HAYES-ROTH]. It is this type of person who, by virtue of his/her education and experience, is supposed to be singularly capable of transmitting requisite domain information to the knowledge engineer. The knowledge engineer is the person responsible for eliciting and then encoding that information in the expert system. One would have expected domain experts to be invaluable in identifying the requisite person-environment interactions.

However, there are serious methodological and philosophical problems associated with defining what constitutes a good 'person-environment fit' and what factors influence the achievement of such a fit. This problem has been a focus of study by psychologists, mental health professionals, sociologists, and the like. Definitive rules have yet to emerge, however, and defining person-environment fit remains an open social sciences research question [RAPPAPORT] [SEGAL]. This lack of consistent domain knowledge greatly influenced the match strategy that was adopted by FOCES. The system's design was also influenced by the nature of the information available from the agency.

The information typically maintained by a social services agency about its foster home resources and foster care clients consists of long, free-form, narrative case records, complemented by a limited amount of information coded on forms for use in the state's automated information system. For children and foster homes that have been known to the agency for a long time, the information contained in these records can be voluminous, confusing, and possibly contradictory. At the other extreme, it is not uncommon for social workers to require a placement for a child about whom they know little except his/her name, gender, race, and approximate age.

To systematize the data collection process, two data collection instruments were developed, one about the child and one about the foster home [see Appendix 1]. The child form recorded information on current circumstances. The only historical information collected was about prior placements and their outcomes. A child's special needs and notable behaviors were recorded in two ways -- by using codes taken from an agency form and by recording, verbatim, relevant case record comments.

The data collection form for a foster home was longer. Information was collected about the following:

- demographics
- family composition
- physical capacities
- placement preferences/prohibitions
- past experience/training
- strengths and weaknesses

Free-form text was also recorded about foster home strengths and weaknesses.

### 2.2.4 Change and Management of Information

A distinguishing characteristic of the information, maintained through social services agencies, is its volatility. The need to handle a large, rapidly changing knowledge base is typically not addressed in expert systems work. However, in domains such as foster care, changeability is the norm for most knowledge. Children come into foster care and leave foster care frequently. For example, in the Roanoke City Department of Social Services, approximately 50 placements were recorded during the months of December 1984 through February 1985. Similar statistics for case closings and child movements are not available but one would expect commensurate rates of activity.

Information about available foster homes is also subject to rapid change. As children are placed in a home, its capacity to take new children is diminished. Removal of a child from a home may increase its availability. Changes within the foster family, e.g. death or illness of the caregiver or her spouse, may also affect the home's capacity to serve certain kinds of children or to take any children at all. New foster homes are also being added continually and existing homes may close.

With this volume of data potentially going into and changing within a system's knowledge base, a fully functional expert system would have to incorporate good data maintenance facilities. Users would have to be able to easily add, delete and update foster child and home records. Without such support, the expert system's knowledge base would soon become obsolete and its usefulness as a basis for suggesting matches would decline.

The volatile nature of the data suggests an approach to expert system development that is a merger of management information systems capacities and artificial intelligence technology. This approach is beyond the scope of an experimental system such as FOCES. However, the system design did require support of some database management facilities. This level of commitment to database maintenance has not been required of most other expert systems in which the knowledge base is static.

# 3. SYSTEM IMPLEMENTATION

## 3.1 Knowledge Implementation

How knowledge about children and homes should be represented has been a critical concern for the designers of FOCES. Several factors influenced the selection of representational models:

- The descriptive, free-form nature of the 'real world' data available for analysis
- The use of the General pUrpose Expert System Shell (hereafter referred to as GUESS) [LEE]
- The decision to adapt information retrieval techniques as the guide for classifying children and homes and finding appropriate matches between the two types of entities

In order to satisfy the requirements imposed by each of these, FOCES has developed multiple knowledge representations for its data. Each is used at different stages of the system's processing.

### 3.1.1 Nature of the 'real world' data

Two data collection instruments were designed to capture information from child and foster home case records. These two forms are the first knowledge representations used by FOCES to describe the entities of interest. The decision to use this knowledge representation was made after approximately six months of using 'traditional' knowledge acquisition mechanisms, such as interviews with experts and searches in the domain literature. The developer had obtained as much domain knowledge as she could expect using those means, but they had not yielded sufficient information to proceed with system development.

The original proposal to develop a system capable of suggesting appropriate matches between children and foster care homes was made to Mr. James Ritchie, Director, Roanoke City Department of Social Services, in 1984. By agreeing to permit the project to be developed with cooperation from his agency, Mr. Ritchie hoped that such a system would eventually provide him with much needed data to identify service gaps. That is, he hoped that if the system consistently demonstrated that appropriate matches could NOT be suggested for certain kinds of children, then that data could be used to justify obtaining increased support for some kinds of child welfare facilities. He designated a key member of his administrative staff and a child welfare services supervisor to take charge of providing the project's developer with access to any information necessary to proceed. Whether FOCES was able to provide Mr. Ritchie with the kind of feedback he desired regarding foster care facilities will be discussed in Chapter 5 of this thesis.

The developer has a Masters in Social Work and had worked in a child welfare agency for many years prior to entering the field of computer science. She was able to use her social work background to build rapport with the designated experts and to structure multiple interviews with at least three experienced members of the agency's child welfare staff. One of those members was responsible for supervising child placements in foster care homes; two others were responsible for recruiting and approving foster care homes, and helping workers to select a home for a child. The following kinds of interview questions were asked:

- What is the 'typical' child in need of placement like? How does he/she behave?
- What qualities do you look for in a 'good' foster home?
- Are all homes suitable for all kinds of children? What makes a home suitable

for one kind of child and unsuitable for another kind of child?

- What factors do you consider when you have to choose a home for a particular child?

The social workers at the Roanoke City Department of Social Services have shown continuing interest in FOCES and have cooperated fully. Many at the agency who have devoted significant time to the project possess Masters' degrees in Social Work and many years of experience. Despite these credentials, the nature of the work they do ill-prepared them to serve in the 'expert' role – as it has been defined in the expert systems literature [HAYES-ROTH].

Social workers often deal with the 'squeakiest wheel', i.e., the most severe emergencies. The necessity to work in 'crisis mode' diminishes one's capacity to step back and analyze daily activities. As a result, one's capacity to serve as an 'expert' is considerably reduced.

An additional complication arises in the foster care domain. There is a long-term nationwide shortage of foster care homes [KADUSHIN]. As a result, homefinders rarely have the luxury of 'choosing' a home. This tendency to downplay the selection process is demonstrated by the lack of any professional literature devoted to the issue. A review of the last ten years of *Social Work Research and Abstracts*, which provides references to professional articles and books, reveals fewer than 20 citations to articles that address the selection problem.

The answers given to questions such as those listed above did not provide a sufficient base for building rules for matching. The social workers could identify some **general** characteristics of a 'good' foster home; however, they lacked the ability to list **distinguishing** factors -- that is, factors that make a home 'good' or 'bad' in particular circumstances. They did identify some 'generic' child and home attributes, such as 'child behaviors', 'special needs of children', and 'preferences of the foster home caregiver', which were important to investigate. Extensive case record analysis was required, however, to find the specifics needed to transform these attributes into suitable matching factors. For example, the records had to be scanned to identify which child behaviors were 'noteworthy', which preferences of foster home caregivers were routinely cited, and the like. The data collection forms that were developed (see Appendix 1) were used to standardize the analysis of case records. The attributes that were included on each were based on the observations of the domain experts.

Over the course of about six months, approximately 50 child case records and a like number of home case records were carefully read and detailed notes were taken on each. The results of this effort provided the basis for formulating the criteria that were subsequently used to select suitable homes for a child. Looking at the 'child behavior' data category as an example, the result of this effort was to identify eleven (11) notable behaviors that seemed to be characteristic of different types of children. These behaviors subsequently served as a basis for distinguishing whether a child better fit into one prototype or another for matching purposes.

On the average, reading and analyzing a single child or home case record took about 30 minutes. Figure 1 is an example of the kind of narrative text that was read along with the resulting summary that was produced.

### 3.1.2 Use of GUESS and Horn Clause (HC) Prolog Environment

The use of PROLOG for constructing expert systems has been suggested [CLARK]. Its use at Virginia Tech is especially appropriate, since the HC version of Prolog has been in use here for several years [ROACH]. Various tools for language and conceptual processing [SCHANK] have already been coded in HC Prolog. GUESS, a

## Sample Narrative

| | |
|---|---|
| 10/20/83 | Mrs. S. called re: Johnny's bedwetting. He is doing this about 2x a wk. ABC |
| 10/23/83 | Visited S. home today at request of Mrs. S. She is frustrated w/ Johnny's bedwetting and wants him moved. I will make appt. for him to be evaluated by Dr. X. She will not force Johnny to be moved until after that appointment ABC |
| 11/12/83 | Dr. X. found no physical cause for Johnny's bedwetting. Mrs. S. has agreed to work w/ Johnny for several more weeks ABC |

## Sample Summary

Notable behavior: bedwetting

# Figure 1. Example Child Case Record Narrative and Summary

shell or tool for developing expert systems, has also been developed at Virginia Tech and demonstrated on a number of problem domains [LEE] [VIRKAR]. FOCES is implemented using GUESS and HC Prolog.

The information encoded on the data collection instruments is represented in GUESS tables. Each table is used to represent a single attribute of a child or home as recorded on the individual's data form. For example, there is a table to represent the attribute identified as birthdate of a child and a second table to represent the attribute known as reason for placement. The former contains the birthdate of each child in the database along with that child's unique case identification number. The latter contains each child's unique case identification number and a list of reasons that a placement is being sought for a child. There is also a table to represent the attribute called race of caregiver and another to represent the age range of children that the caregiver would prefer to have placed in her home. Figure 2 portrays the first two levels of knowledge representation used by FOCES.

It was found, after a considerable number of case records were analyzed, that some attributes that had been identified as potential match factors could not be reliably recorded because relevance information was often missing from the case records. For example, it had been suggested that a foster home caregiver's participation in certain training courses might make her a better choice for placement of some kinds of children. However, a review of many foster home records showed that information about the kinds of training that had been taken by a caregiver was not accessible. All training was simply listed as 'agency training' and no details about the content of the course were given. Similarly, information regarding the types of behaviors or child characteristics that were prohibited by some foster home caregivers was often absent. Those attributes had to be eliminated from use as match factors. GUESS requires that each <key> <value> pair in a

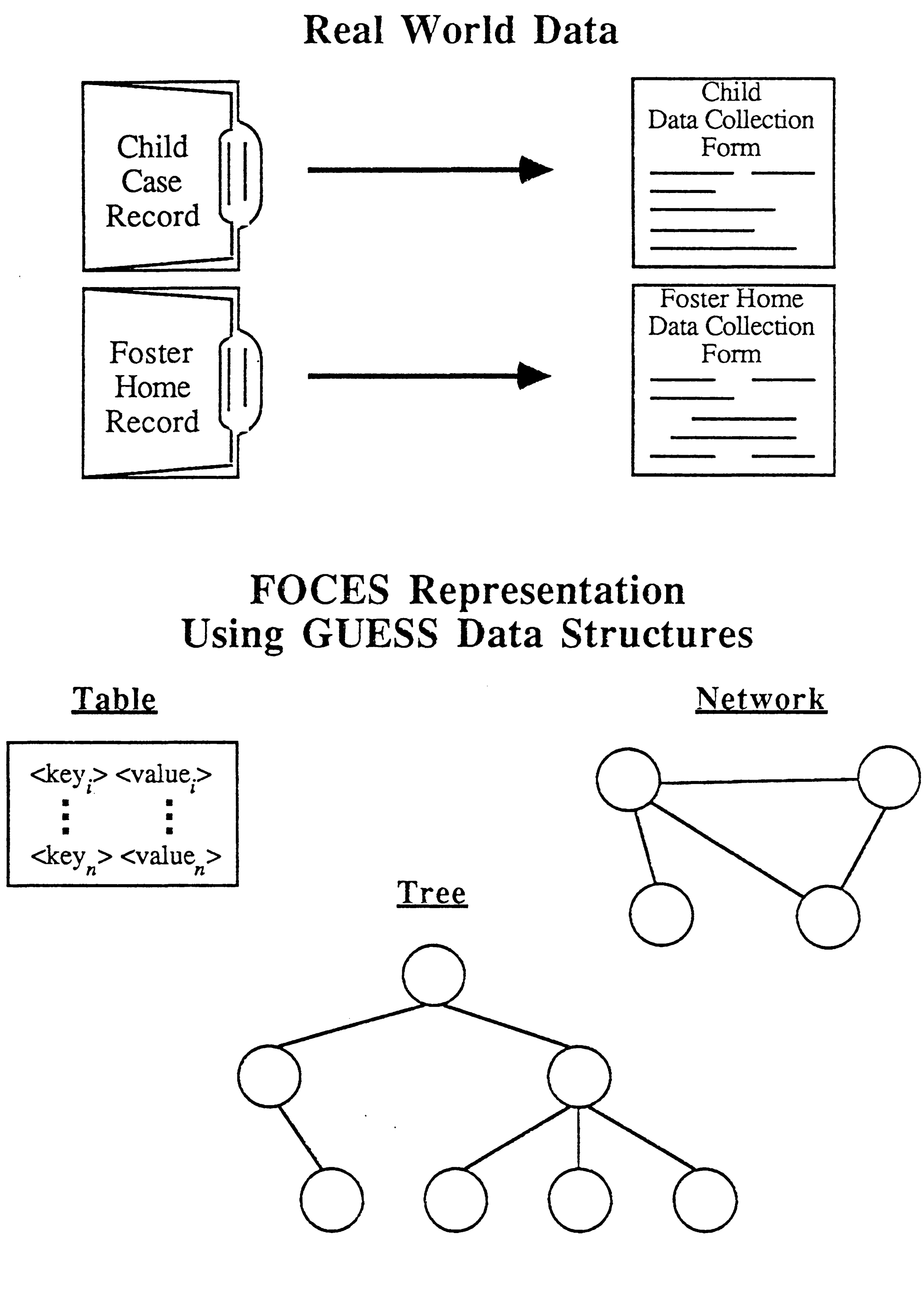


**Figure 2. First Levels of Knowledge Representation**

**Representation of a Home Attribute**

```
;  TABLE NAME      number of children approved for
;  USAGE           match factor:  number approved >= number to place
;  KEY             resource number of home
;  VALUE           a digit 1-9
;  SECURITY        unclassified
;  CATEGORY        read
;
;  Author          Sheila G. Winett
;  Date            March 1986
;
(  (TABLE "number of children approved for"   UNCLASSIFIED   (READ)
   (
                   ("000200"       4)
                   ("000201"       4)
                   ("000202"       6)
                   ("000203"       6)
                   ("000204"       6)
                   ("000205"       3)
   )
)  )
```

**Representation of Child Attribute**

```
;  TABLE NAME      birthdates of children
;  USAGE           stored birthdates of children in foster care
;  KEY             case number of child
;  SECURITY                          unclassified
;  CATEGORY                          read
;
;  Author          Sheila G. Winett
;  Date            January 1985
;
(  (TABLE   "birthdates of children"    UNCLASSIFIED        (READ)
   (
                 ("000001"           "00/00/77")
                 ("000002"           "00/00/73")
                 ("000003"           "00/00/71")
                 ("000004"           "10/00/73")
                 ("000005"           "07/00/76")
                 ("000006"           "00/00/70")
   )
)  )
```

**Figure 3. Sample GUESS Tables**

table must be unique. Its syntax allows a value to be a number, a character string, or a list. Figure 3 shows an example of a child attribute and a home attribute in a GUESS table.

For a single child, therefore, the information maintained about him/her is distributed across a series of tables, each representing information about a single child attribute. The same is true of information maintained about each foster home. Table 2 provides a complete list of the GUESS tables used to represent children and homes at this level of detail.

In addition to the tables described above, GUESS provides other data structures to represent knowledge. Trees may be used to represent hierarchically related attributes. Directed or undirected networks are also available to represent information. Nodes of a network denote objects or concepts and arcs between nodes represent a relationship among the objects or concepts. For each data structure, e.g. tree networks, GUESS provides appropriate data manipulation facilities.

The developers of FOCES found that GUESS tables were sufficiently flexible to represent all of the data that was being used for children and homes. However, the data manipulation routines provided by GUESS to manage those tables proved less useful for some system functions. Also, in order to use the information retrieval approaches being tested by FOCES, it was necessary to develop other knowledge representations to augment the GUESS structures.

### 3.1.3 Use of Information Retrieval Techniques

Vector representations of information about children and homes were constructed using the data stored in the GUESS data structures. Figure 4 defines these vectors.

# Table 2. List of GUESS Tables

## Child Tables

names of children
race of children
birthdates of children
current placements of child
prior placements of child
primary reason for placement
gender of child
special needs of children
behaviors of children

## Home Tables

name of home
date home opened
race of caregiver
type of housing
family members in foster home
date of birth of members
gender of members
number of children approved for
gender of children approved for
age of children approved for
special needs of children approved for
number of children preferred
gender of children preferred
age of children preferred
special needs of children preferred
number of children in care
notable strengths of caregiver

**Child**

$V_x^C$ = ( *Case #x* $(a_1, v_1, mf_1) \ldots (a_m, v_m, mf_m))$

where

$a_i$ = $i^{th}$ child attribute

$v_i$ = actual value of attribute $i$ for home $y$

$mf_i$ = real-valued membership function of attribute / value

**Foster Home**

$V_y^H$ = (*Resource #y* $(a_1, v_1, mf_1) \ldots (a_n, v_n, mf_n))$

where

$a_i$ = $i^{th}$ home attribute

$v_i$ = actual value of attribute $i$ for home $y$

$mf_i$ = real-valued membership function of attribute / value

**Figure 4. Vector Representation of GUESS Data Structure**

There is an almost 1:1 mapping between the information stored in the GUESS tables and the vector representations that were created. These vectors were later referred to by extended Boolean queries which were used to do the actual classification and selection of appropriate homes for particular children. The key difference between the GUESS representation and the vector representation lies in the assignment of a real-valued membership function to each attribute/value pair in the vector representation. This membership function signifies the degree tow which an individual child or home possess a particular attribute and may lie in the range of 0.0 – 1.0. However, for computational ease, the value of each membership function was stored as an integer 0 – 10 and later is normalized to a value between 0.0 and 1.0.

A membership function is capable of representing fuzzy membership in a set. However, the binary nature of many of the attributes currently used by FOCES only made it necessary to represent stiff set membership. For example, a child may only be in the set of female children if her gender is recorded as 'female'. In that case the value of the membership function would be one (1). If any other gender is recorded for a child, the membership function for that child in the set of females would be a zero (0).

## 3.2 DATA TRANSFORMATIONS

Data about children and homes are transformed in stages – summarization, indexing, and classification – so it will be possible to perform the match function that is the ultimate aim of FOCES. Figure 5 graphically represents these transformations.

The summarization phase organizes the 'real world' data into a form that usable to the system. The indexing phase changes the data to their vector representation to make them accessible to the extended Boolean queries that are constructed to determine the appropriateness of various prototypes. Finally, as a result of matching the vector

| | Real World | GUESS Data Representations | Vector Representations | Classification Lists |
|---|---|---|---|---|
| CHILD | $C_i$ | $D_i^C$ | $V_i^C$ | $L_i^C$ |
| HOME | $H_j$ | $D_j^H$ | $V_j^H$ | $L_j^H$ |

<u>Data Flow</u>

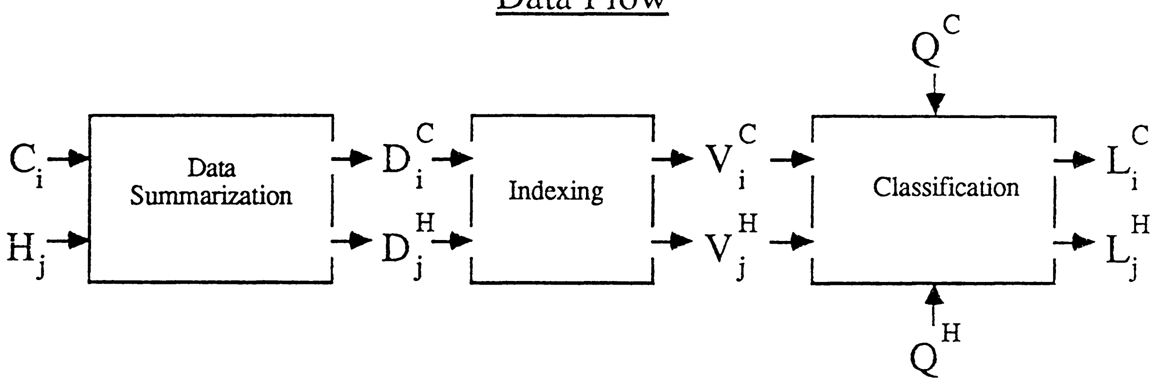


**Figure 5. Data Transformations**

representations against a series of p-norm queries, a representation of salient home and child qualities is developed. These can then be used to select relevant homes for consideration.

### 3.2.1 Data Summarization Phase

The objective of this phase was to find evidence in child or home records of the factors that had been identified as distinctive during previous case record analysis (see section 3.1.1). For example, the case records of children who would be in the FOCES database had to be analyzed to determine whether those children could be said to exhibit any of the eleven (11) previously defined 'child behaviors.' If they did, then that information had to recorded in a format that would make it amenable to the GUESS system and to the software used to do the matching.

### 3.2.2 Data Indexing Phase

The need to 'index' FOCES' child and home data, i.e., transform them to a vector form, was the result of incompatibility between GUESS data representations and the information retrieval techniques being used for matching. Each fact represented as a <key><value> pair within the GUESS shell had to be indexed to its vector analogue as either a 'cfact' or an 'hfact,' depending on whether it represented information about a child or a home, respectively. The data in these vectors were later classified with respect to the degree that each of a child's or a home's attributes fit each of the child or home prototypes. For one set of child data and one set of home data, figure 6 shows the correspondence between the GUESS and vector data formats.

## A. Sample Child Attribute

**GUESS Representation**

```
( (TABLE     "birthdates of children"     UNCLASSIFIED     (READ)
  (
              ("000001"        "09/00/77")
              ("000002"        "09/00/73")
              ("000003"        "09/00/71")
                        ......
  )
) )
```

**Vector Representation**

```
( (cfact "000001" age 107 10) )
( (cfact "000002" age 155 10) )
( (cfact "000003" age 179 10) )
          ......
```

## B. Sample Home Attribute

**Guess Representation**

```
( (TABLE   "age of children approved for"  UNCLASSIFIED     (READ)
  (
              ("000200"        (00 18))
              ("000201"        (01 14))
              ("000202"        (00 18))
                        ......
  )
) )
```

**Vector Representation**

```
((hfact     "000200"    youngest  0 10))
((hfact     "000200"    oldest 18 10))
((hfact     "000201"    youngest 1 10))
((hfact     "000201"    oldest 14 10))
((hfact     "000202"    youngest 0 10))
((hfact     "000202"    oldest 18 10))
                  ......
```

**Figure 6. Example of Relationship between GUESS & Vector Representations**

### 3.2.3 Data Classification Phase

It was during this phase that critical tasks requisite to matching a child with potential homes were done. Specifically, the characteristics of each particular child and home were correlated with the predefined prototypes of children and homes. The resultant 'cproto' and 'hproto' vectors served as immediate input to the match process.

"Cproto" and "hproto" represent the degree to which attributes of individual children/home fit each child or home prototype. So, for child X, a list of "cprotos" would show the similarity between the attributes of X and each of the system's predefined child prototypes. For example, if X's age was 6 months, she might have an associated real valued membership function of .5 in the "infant" prototype. Looking at home Y, as an example, a list of "hprotos" would show how well suited home Y was to serve as a placement for children in each of the prototype categories. Figure 7 shows how the "cproto" and "hproto" vectors were used to represent a single child's and a single home's fit within each of FOCES' prototype categories.

Each prototype represented a distinguishable class or category. Five of the prototypes were based on age; for example, there was an infant prototype and a preteen prototype. Other prototypes were based on race and distinctive patterns of need; for example, there was a medically needy prototype, an intellectually impaired prototype and a black prototype. In all there were 16 prototypes used. Each had its child representation – that is, a representation of what child attributes were associated with being in each prototype. Every prototype also had a companion home representation – that is, a representation of which home attributes were considered important for a home to have if a prototypical child was to be placed there. A result of using this representational system was that there were essentially sixteen prototype pairs. Table 3 lists the prototype categories that were used.

```
; "protos.hc"
      (assert
            ((cproto 1 "000016" 1))
            ((cproto 2 "000016" 0))
            ((cproto 3 "000016" 0))
            ((cproto 4 "000016" 0))
            ((cproto 5 "000016" 0))
            ((cproto 6 "000016" 1))
            ((cproto 7 "000016" 0))
            ((cproto 8 "000016" 1))
            ((cproto 9 "000016" 0))
            ((cproto 10 "000016" 0))
            ((cproto 11 "000016" 0))
            ((cproto 12 "000016" 0.4907))
            ((cproto 13 "000016" 0.0248))
            ((cproto 14 "000016" 0.0191))
            ((cproto 15 "000016" 0.0112))
            ((cproto 16 "000016" 0.0318))
            ((hproto 1 "000200" 0.0585))
            ((hproto 2 "000200" 0.0585))
            ((hproto 3 "000200" 0.0585))
            ((hproto 4 "000200" 0.0585))
            ((hproto 5 "000200" 0.0585))
            ((hproto 6 "000200" 0))
            ((hproto 7 "000200" 0.5773))
            ((hproto 8 "000200" 0))
            ((hproto 9 "000200" 1))
            ((hproto 10 "000200" 0.7))
            ((hproto 11 "000200" 0.1596))
            ((hproto 12 "000200" 0.1189))
            ((hproto 13 "000200" 0.0874))
            ((hproto 14 "000200" 0.1121))
            ((hproto 15 "000200" 0.1121))
            ((hproto 16 "000200" 0.1201)) )
```

**Figure 7. Example of "cproto" and "hproto" Representations**

# Table 3. List of Prototypes

## Home Prototypes

place an infant (0-1 yr)
place a toddler (1-3 yrs)
place a preschooler (3-5 yrs)
place a child of school age (5-13 yrs)
place a teenager (>=13 yrs)
place a female
place a male
place a black child
place a white child
place a child of 'other' race
place a dependent/neglected child
place a sexually or physically abused child
place a medically needy child
place an emotionally impaired child
place an intellectually impaired child
place a child with behavior problem(s)

## Child Prototype

infant (0-1 yr)
toddler
preschooler
child of school age
teenager
female
male
black child
white child
child of 'other' race
dependent/neglected child
sexually or physically abused child
medically needy child
emotionally impaired child
intellectually impaired child
child with behavior problem(s)

For M children in the FOCES database, this phase yielded 16 x M 'cproto' vectors which identified how similar each child's characteristics were to the characteristics identified as belonging to children in each prototype. The similarity measure was presented as a normalized, real-valued membership function between 0.00 and 1.00. Likewise, for N homes, this phase produced 16 x N 'hproto' vectors to indicate the similarity between each home in the database and each of the home prototypes.

## 3.3 Matching Strategies

During the early planning for development of FOCES, several strategies for matching a child with potential homes were considered:

- direct match
- child/home exemplar match
- similar child match
- tailored p-norm query match

However, given what was learned during the knowledge acquisition phase of FOCES' development and the experimental nature of the thesis work, the decision was eventually made to select the approach that seemed to best relate to the problem and test it.

The 'direct match' approach is so-named because it would have involved selecting the home(s) which seemed most appropriate for a child on the basis of a direct comparison between the attributes of the child in need of placement and the characteristics of each individual home. It would have required the development of many rules to guide the system's reasoning process. Typically, an expert system would have several hundred such rules. For example, one might develop a rule like the following:

<u>if</u> (reason for placement = "sexual abuse") and ((number of males in home = 0)
or (age of males in home < > age of abuser))
<u>then</u> consider home as a placement

Given the difficulties encountered in attempts to have the experts specify rules such as the example cited above (see section 3.1.1), this approach was deemed inappropriate for this domain.

For different, albeit related reasons, the 'similar child' match strategy was also dismissed as a viable approach. The 'similar child' match strategy is based on finding homes that currently have or have had children in placement who are similar to the child for whom a placement is being sought. The rationale for this approach is that if a home currently has a child successfully in placement and that child is similar to the child for whom a placement is being sought, then there is a likelihood that the home may also be appropriate for the new child. As FOCES developed, however, unanticipated problems led to elimination of this approach as a match strategy.

It became clear that as an experimental system, FOCES could not readily manipulate the vast quantity of data that would have been required to effectively test the similar child approach. For each home, a complete data representation for every child currently in care and every child that had ever been in care in that home would have had to be maintained. For some homes this would have meant maintaining data for as many as 50 - 75 children. The effort required would have been monumental, even if all those records could have been located – an unlikely possibility. Additionally, it was pointed out by the social workers that it is not always reasonable to expect similarity to be a virtue. For example, if a home is successfully caring for a child with multiple handicaps, the effort required to provide that child with sufficient care might preclude placement of any other children, especially another handicapped child. As another example, if a home has a sexually active, manipulative adolescent in the home, the presence of a like adolescent might worsen some negative behaviors.

The two remaining match strategies: tailored p-norm query match and child/home exemplar match, both require the construction of p-norm queries as the basis for match. Using the former requires that for each child for whom placement is sought a 'hand crafted' p-norm query be constructed to reflect that individual's unique constellation of personal traits and external circumstances. In many ways this approach is analogous to the process used when a library user makes a search request – the librarian must interpret that request and transform it to a unique query that will select references of interest from a document collection. The exemplar approach, by contrast, requires that a set of 'stock' queries be constructed. Then, when a child is in need of placement, his/her particular characteristics and circumstances can be mapped to the exemplar queries and a similiarity measure can be computed to reflect the degree of match between the child and each exemplar. A parallel mapping is also done between each potential home and the set of home exemplars.

The exemplar match approach was selected for testing. It was compatible with the kinds of data available about children and homes and allowed the developers of FOCES to test the viability of applying information retrieval techniques to solve an artificial intelligence problem. In addition it permitted the developers to skirt the issue of having to identify unique matching factors that would contribute to a good 'person-environment fit'. Rather, that fit was characterized as a multidimensional matching process.

The 'cprotos' and 'hprotos' developed during the Data Classification Phase (see section 3.2.3) provided the data needed to implement the exemplar match approach. The cprotos represent the degree to which the child fits each of the sixteen (16) child exemplars (queries). The hprotos represent the degree to which each potential home match the sixteen (16) home exemplars (queries).

The first step in the match process is to pair the child in need of placement with each of the potential homes to form a list of <child, home> pairs. For N potential homes, the list contained N <child, home> pairs. For each <child, home> pair a cosine similarity coefficient is computed which represents the 'match' between the child and a particular home.

The correlation coefficient is computed using the following formula:

$$\frac{\sum_{K=1}^{16} (L_{iK}^{C} \cdot L_{jK}^{H})}{\sqrt{\sum_{K=1}^{16} (L_{iK}^{C})^2 \cdot \sum_{K=1}^{16} (L_{jK}^{H})^2}}$$

Once this correlation is computed for each <child, home> pair, the values produced by this step are sorted in descending order so that the homes are presented in best to worst order. The results obtained from using this match strategy are presented in the next chapter of this thesis.

# 4. RESULTS

The p-norm queries (prototypes) developed to match a child with the most appropriate foster home were tested for seventeen (17) children. A database of fifteen (15) potential homes was available for matching. All of the data used, for children and homes, were extracted from actual case records of the Roanoke City Department of Social Services. Identifying information about all children and homes was deleted to protect individuals' privacy.

The children for whom the system's match strategy was tested were evenly divided between males and females. About 80% of the children were caucasian; the remainder were black. The children ranged in age from birth to 17 years old. There was one child in the birth to one year old category; three were in the 1-3 year old category; eight belonged in the 5-13 year old group; the remaining four were teenaged. The children had come into care for a variety of reasons, with parental neglect being the most common cause for placement. The cases selected for testing reflected a broad spectrum of notable personal characteristics, such as chronic medical problems, mental retardation, and the like. A variety of behavior patterns, including excessive fearfulness, depression, and others was also evident among the children included in the test database.

The fifteen (15) potentially available homes that comprised the home database were selected from among those homes in use by the Roanoke City Department of Social Services during 1986. Six of the home caregivers were black; the remaining nine caregivers were white. Without exception all of the homes were approved for placement of both male and female children. Fourteen (14) of the fifteen had no noted preference for children of one gender or the other.

Many of the homes were also approved to care for children of all ages (birth to 18 years old). Any restrictions placed on a home with regard to placement of children in specific age categories tended to be at the extremes of the age spectrum. For example, four of the homes were not approved to care for infants, that is, children less than one year old. A like number of homes were not approved to care for preteen or teenaged children.

The amount of available knowledge about the notable strengths of individual homes varied a good deal. In some instances, the case records had identified many home qualities, e.g., patience, capacity to meet physical needs, that were considered useful as match factors. In other cases, there was very little information from which to work.

For each child in the test database, FOCES computed a cosine similarity coefficient for each of the fifteen (15) homes and listed those in descending rank order (highest cosine coefficient ---> lowest cosine coefficient). Appendix 2 provides a full listing of those results. To establish the "goodness" of the results obtained from FOCES' match strategy, two social workers from the Roanoke City Department of Social Services with whom the author had worked during the development of FOCES were asked to participate in an evaluation exercise.

Each social worker, working independently of the other, was given a packet of placement assessment forms, one for each of the 17 children in the test database. Each placement assessment form included a natural language summary of all the information that FOCES was given about each child. The social workers were also given natural language home descriptions for each of the 15 homes in the test database. Each child and home description was given a code number to hide the identity of the individuals. The social workers were asked to rate each potential home as to its appropriateness for each child. An ordinal scale of 1 - 5 was used, with one representing an 'excellent' choice and five representing a 'very poor' choice for placement of the child.

The social workers agreed with each other, i.e., each gave a particular home the identical rating as a placement for a specific child, from 40 - 86.67% of the time, depending on the individual child. In most instances, however, their agreement was with respect to the 'worst' homes for a child. That is, most of the agreement occurred for homes that received a rating of '5' from both social workers as a placement for a child. When only ratings of '4' or better are considered, the percentage of agreement between social workers decreased dramatically. For the seventeen children tested, the highest rate of agreement between the two social workers was only 26.67% when ratings of '4' or better were analyzed. Figure 8 reflects these results.

In order to compare the social workers' ratings to the results achieved by FOCES, an equivalency rating scheme had to be devised to equate the cosine similarity coefficient computed by FOCES for each <child, home> pair and the ordinal rating scale used by the social workers to rate the homes. The following was devised, based on visual inspection of the results:

| Similarity Coefficient | Rating |
|---|---|
| <= 0.3000 | 5 |
| <= 0.5000 | 4 |
| <= 0.6000 | 3 |
| <= 0.6500 | 2 |
| > 0.6500 | 1 |

Figure 9 shows the rates of agreement between FOCES' ratings and the ratings given by the social workers. Looking at the findings displayed in Figures 8 and 9, one can see that, when all ratings were considered (1-5), the rate of agreement between FOCES' ratings and the ratings given by at least one of the social workers was at least as good as the inter-social worker rating reliability in 70.6% (11 of 17) of the instances. Further, Figure 9 shows that there was unanimity among FOCES' ratings and the ratings given by the two social workers in a large percentage of the cases tested. Agreement between the ratings given by FOCES and those given by at least one of the social workers decreased significantly when ratings of '5' were eliminated from consideration.

| Child Id | % Social Workers Agree (all ratings) | % Social Workers Agree (ratings of 4 or better) |
|---|---|---|
| 000001 | 53.33 | 13.33 |
| 000002 | 40.00 | 0.00 |
| 000003 | 66.67 | 26.67 |
| 000004 | 66.67 | 26.67 |
| 000005 | 53.33 | 13.33 |
| 000006 | 53.33 | 13.33 |
| 000007 | 60.00 | 20.00 |
| 000008 | 66.67 | 26.67 |
| 000009 | 60.00 | 20.00 |
| 000010 | 60.00 | 20.00 |
| 000011 | 86.67 | 26.67 |
| 000012 | 60.00 | 0.00 |
| 000013 | 66.67 | 6.67 |
| 000014 | 66.67 | 26.67 |
| 000015 | 40.00 | 0.00 |
| 000016 | 53.33 | 13.33 |
| 000017 | 73.33 | 0.00 |

**Figure 8. Results of Social Workers' Home Suitability Evaluation**

| Child Id | % Both SW and FOCES agree (all ratings) | % One or Both SW and FOCES agree (all ratings) | % One or Both SW and FOCES agree (ratings of 4 or better) |
|---|---|---|---|
| 000001 | 46.67 | 80.00 | 40.00 |
| 000002 | 40.00 | 60.00 | 13.33 |
| 000003 | 46.67 | 66.67 | 26.67 |
| 000004 | 53.33 | 60.00 | 26.67 |
| 000005 | 40.00 | 73.33 | 26.67 |
| 000006 | 40.00 | 73.33 | 33.33 |
| 000007 | 40.00 | 60.00 | 33.33 |
| 000008 | 46.67 | 53.33 | 13.33 |
| 000009 | 40.00 | 53.33 | 13.33 |
| 000010 | 53.33 | 73.33 | 26.67 |
| 000011 | 60.00 | 60.00 | 20.00 |
| 000012 | 60.00 | 93.33 | 33.33 |
| 000013 | 66.67 | 93.33 | 33.33 |
| 000014 | 46.67 | 53.33 | 20.00 |
| 000015 | 40.00 | 60.00 | 20.00 |
| 000016 | 46.67 | 53.33 | 13.33 |
| 000017 | 60.00 | 73.33 | 13.33 |

**Figure 9. Comparison between Social Workers' Home Suitability Ratings and FOCES' Home Suitability Ratings**

When only ratings of '4' or better were evaluated, the rate of agreement between FOCES' rating and the rating of at least one of the social workers ranged between 13.33 - 40.0%.

In addition to analyzing the rating concurrence between FOCES and the two social workers, recall and precision measures were calculated for each child. Both measures are customarily used to evaluate the performance of document retrieval systems. In view of the fact that the p-norm queries used for matching by FOCES were adapted from information retrieval research, it seemed appropriate to adapt similar evaluation measures.

Loosely speaking, recall examines the effectiveness of a document retrieval system, while precision evaluates a system's efficiency. The formulae used to represent each are given below:

$$\text{recall} = \frac{\text{No. of Relevant Documents Retrieved}}{\text{No. of Relevant Documents}}$$

$$\text{precision} = \frac{\text{No. of Relevant Documents Retrieved}}{\text{No. of Documents Retrieved}}$$

When applied to a document retrieval system, a high recall coefficient implies that the system's searching and matching strategies are retrieving all or most of the documents deemed relevant to a user's query. When applied to FOCES' results, a high recall measure shows that FOCES' match strategy is selecting all or most of the homes that are considered appropriate for placing a particular child.

Precision is a measure of a document retrieval system's efficiency. That is, it measures how many of the total number of documents retrieved by a system were relevant. If a system's precision measure is high, the system is bypassing irrelevant documents and selecting mostly documents that are germane to the query. As applied to FOCES, a high precision measure implies that the system's match strategy could effectively distinguish appropriate and inappropriate homes and select relevant, i.e. appropriate, homes for a particular child.

To evaluate FOCES' recall and precision, a listing of homes was compiled. Homes for which at least one social worker had given a rating of '1' or '2' for a particular child were included on the list. For each of those homes, the ranking that FOCES had assigned to that child-home match was noted. The results of this analysis were extremely positive.

In 16 of the 17 child cases tested, the home ranked by FOCES as most appropriate for the child was included among the homes rated as a '1' or '2' by at least one social worker. In eight of those instances, both social workers had given a rating of '1' or '2' to the home that had been top-ranked by FOCES. In ten of the child cases, the home ranked as second best for a child by FOCES was also among those identified with a rating of '1' or '2' by at least one of the social workers.

Overall, for the 17 children in the test database, 68.6% of FOCES' top three choices for placement of a child had been rated as '1' or '2' by at least one social worker. Approximately 20% of FOCES' top three choices for placement of each child had been rated as '1' or '2' by both social workers. Appendix 3 contains a complete listing of these recall/precision results. Figure 10 shows a summary of those results.

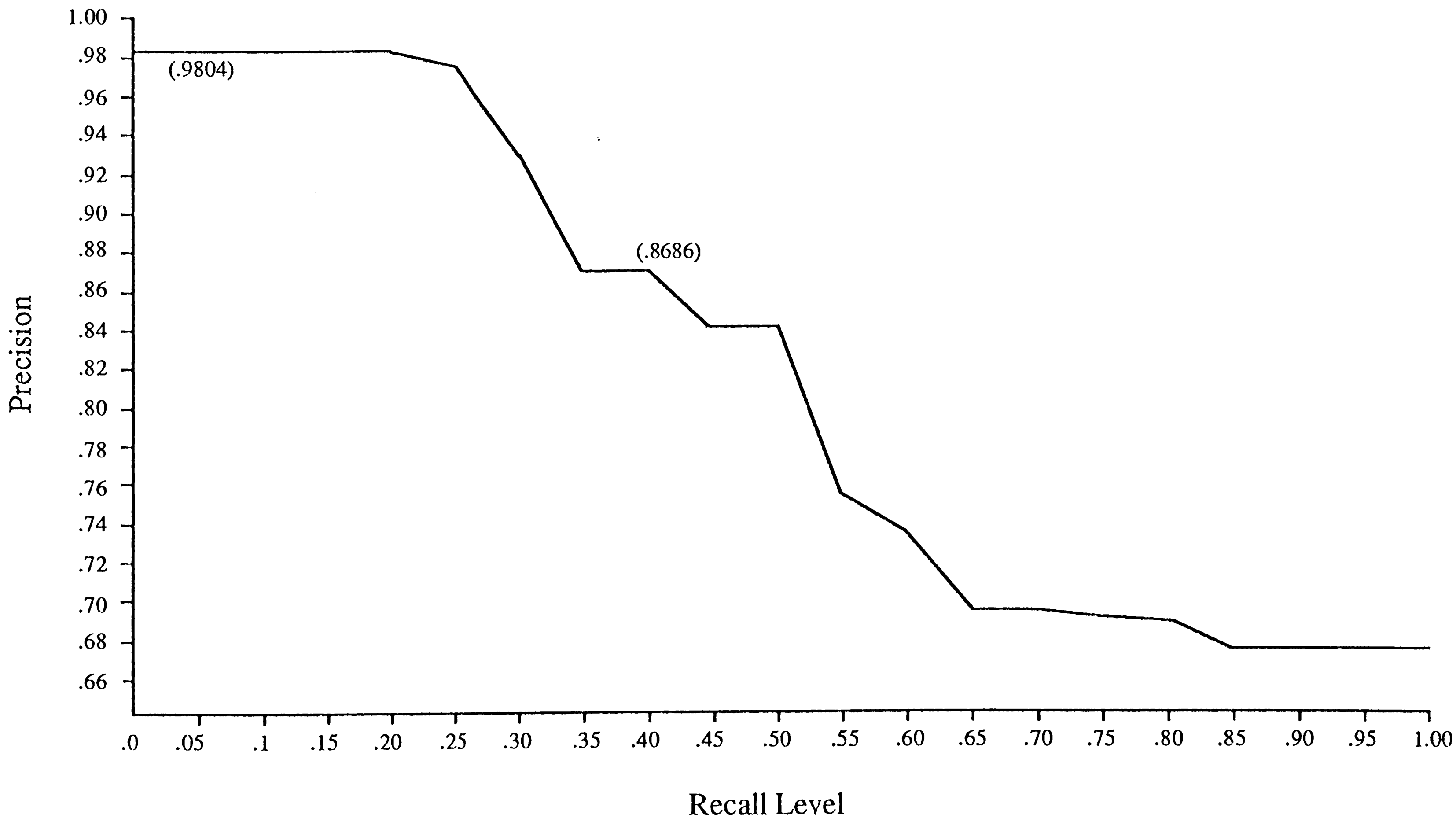


**Figure 10. Summary of Recall – Precision Results**

# 5. CONCLUSIONS AND FUTURE WORK

The development and implementation of FOCES during the past two years represents an effort to break ground in two new areas:

- application of an expert system approach to a new domain
- adaptation of information retrieval techniques to solve an artificial intelligence task

The social services domain offers some special challenges to accepted knowledge acquisition methods. As earlier chapters of this thesis have described, the accepted notion of the domain expert may not be appropriate in this domain. The social worker's tendency to direct his/her attention to the individual circumstances of each child and family, which may be the profession's greatest strength, makes focusing on underlying knowledge principles very difficult. The day-to-day exigencies of the work environment – the so-called 'squeaky wheel' approach to solving many casework problems – also contributes to this difficulty. As a result, the importance of the domain expert as a primary source of knowledge is significantly diminished.

In place of reliance on a single domain expert, the primary developer of FOCES utilized extensive case record analysis and an informal network of domain experts who, at least initially, served as a sounding board for the developer's impressions and intuitions about important matching criteria.

These knowledge acquisition methods proved to be timeconsuming and labor-intensive. Approximately six months were devoted to this phase of the projects developments. One data sufficient to adequately form child and home representations were collected and analyzed, construction of the p-norm queries was relatively simple. Each query took about 1-2 hours to prepare. If subsequent work in this domain confirms the experiences of this project with regard to knowledge acquisition, techniques like using p-norm queries for acquiring domain knowledge will need to be more fully developed and tested. Content analysis techniques currently used extensively in social sciences research to discern the meaning of prose sources may need to be adapted to extract expert knowledge from case records, policy manuals, and like sources. Automatic indexing techniques used in information retrieval systems may also be appropriate for this task.

The difficulties that the primary developer of FOCES encountered with respect to establishing a workable project goal point to a possibly different direction for some future expert systems work. Until now, the primary thrust of expert systems development has been to emulate well-defined, data-driven human decision-making processes in which there are a bounded number of complex paths leading to a solution. The social services domain, as well as perhaps other domains not yet explored, offers a large number of 'open' decision-making processes in which professionals are still contending key issues.

For example, the 'person-environment fit' question encountered during the goal establishment phase of FOCES' development remains unresolved by social scientists. It may well be that the development of experimental expert systems to model divergent reasoning approaches to such open questions can yield important research results. Rather than depending on preexisting knowledge about a domain, perhaps expert systems can prompt discovery of new knowledge.

The results of FOCES' evaluation, for example, indicate that even professional social workers do not consistently agree on which homes would be 'excellent' choices for a particular child. However, they can apparently screen out undesirable placements easily. Likewise, FOCES, using comparatively simple prototypes, could consistently discern poor child-home matches. Its selection of 'good' <child, home> pairs was at least as good as the performance of the social workers in most cases. Of course, only extensive, long-term follow-up could document the efficacy of either choice – the social worker's or the expert system's. Nevertheless, the results do point out the feasibility of using expert system generated results as a contribution to resolving open questions such as these.

FOCES' successful use of extended p-norm queries to represent its entities of interest and to guide the matching process demonstrates that such an approach is viable for addressing some artificial intelligence tasks. Just as the salient characteristics of a child and a home could be represented in FOCES, it would be possible to use a similar representation to model a patient's symptoms and relevant diseases, for example. The match process in that case would need to find the disease prototype(s) which most closely mirrored the patient's symptom query. Other problems in which the central task is to match a particular set of circumstances to some predefined model would be equally appropriate to this approach. P-norm queries offer an expressive representation scheme for many complex entities. Further, the availability of well-defined criteria for evaluation of p-norm query systems, i.e., recall and precision measures, is an added advantage.

In general, the question of how or whether one should evaluate the performance of an expery system is unresolved [HAYES-ROTH]. The human experts upon whose knowledge these systems depend are rarely objectively evaluated. The statistics available are often not applicable to the data available for evaluation. Further, these are arguments about ehat aspects of the expert system should be measured – its performance as conpared to the human expert, its ability to explain its reasoning regardless of the quality of its performance, etc. Despite these reservations, attention to evaluation is important if expert systems are to be seriously considered for use as decision support tools in important application areas.

FOCES represents an initial research effort and its results point to several interesting areas for future work:

- consolidation of management information systems and expert systems approaches
- development and testing of different prototypes
- development of more efficient matching algorithms

As earlier writings about FOCES have indicated [WINETT], the volume and volatility of the child and home data which a fully operational system would need to manipulate is immense. Some merger of management information systems techniques and expert systems techniques would be needed.

At the Roanoke City Department of Social Services, which is not a large public agency, there are approximately 300 children in care at any one time. These children could be placed in any of up to 150 open foster homes or any of several group homes and institutions. The sheer weight of data generated by this many individuals would overwhelm any current expert system. However, if expert systems are to progress toward dealing effectively with problems such as the child-home matching problem addressed by FOCES, means for manipulating large scale databases in real time will need to be addressed. Experience with FOCES indicates that a totally PROLOG based system is not sufficient for the task.

Finally, the prototypes used by FOCES were relatively simple. The weights assigned to different clauses and attributes were educated guesses, based on extensive analysis of case records and the author's own intuition as a former social worker. Clearly, future work should evaluate the relative performance of a system using different prototypes, including varying clause weights, adding new clauses, and the like. Future work might also evaluate the effects of using tailored p-norm queries to represent the individual circumstances and characteristics of each child rather than using prototypes, which impose a level of indirection.

While further comparative studies are thus clearly in order given the investment of the time and effort in working with problems of foster care placement, it is also hoped that the key ideas and methods of this investigation will be picked up and applied to a variety of similar AI type problems.

# Bibliography


[AIKINS] Aikins, J.S. *Prototypes and Production Rules: A Knowledge Representation for Computer Consultations.* Report No. STAN-CS-80-814, Memo HPP-80-17, Stanford Univeristy, Palo Alto, CA, 1980.

[BOBROW] Bobrow, D. G.; Winograd, T. "An Overview of KRL." *Cognitive Science.* 1(1): 3-46, 1977.

[CAPLAN] Caplan, R.D. "Person-Environment Fit: Past, Present, and Future." In C.L. Cooper (editor): *Stress Research: Issues for the Eighties.* John Wiley and Sons, Ltd., 1983.

[CLANCEY] Clancey, W.J.; Shortliffe, E.H.; Buchanan, B.G. "Intelligent Computer-Aided Instruction for Medical Diagnosis." In *Proc. of 3rd Symp. on Computer Application in Medical Care.* 1979.

[CLARK] Clark, K.; McCabe, F. "PROLOG: A Language for Implementing Expert Systems." In J. Hayes and D. Michie (editor): *Machine Intelligence 10.* Chesterfield, England: Ellis & Horwood, Ltd., 1982.

[DAVIS] Davis, R. "Interactive Transfer of Expertise: Acquisition of New Inference Rules." *Artificial Intelligence.* 12: 121-157, August 1979.

[DELAND] Deland, E. *Fluidmod.* Technical Report, Rand Corporation, 1974.

[DUDA] Duda, R.; Gaschnig, J.G.; Hart, P.E. "Model Design in the Prospector Consultant System for Mineral Exploration." In D. Michie (editor): *Expert Systems in the Microelectronic Age.* Edinburgh University Press, 1979.

[DYER] Dyer, M.G. *In-Depth Understanding: A Computer Model of Integrated Processing for Narrative Comprehension.* Cambridge, Mass.: MIT Press. 1983

[FAGAN] Fagan, L.; Kunz, J; Feigenbaum, E.A.; Osborn, J. "Representation of

Dynamic Clinical KNowledge: Measurement Interpretation in the Intensive Care Unit." In. *Proc of the Sixth Conf. on Artificial Intelligence*. 1979.

[FAIN] Fain, J.; Gorlin, D.; Hayes-Roth, F.; Rosenschein, S.; Sowiziral, H.; Waterman, D. *The ROSIE Language Reference Manual.* Technical Report, Rand Corporation, 1981.

[FEIGENBAUM] Feigenbaum, E.A.; Buchanan, B.G.; Lederberg, J. "On Generality an Problem Solving: A Case Study Using the DENDRAL Program." In B. Meltzer and D. Michie (editor): *Machine Intelligence 6*. American Elsevier, 1971.

[FIKES] Fikes, R.E. "REF-ARF: A System for Solving Problems Stated as Procedures." *Artificial Intelligence.* 1(1/2):27-120, March, 1970.

[FOX] Fox, E.A. *Extending the Boolean and Vector Space Models of Informaiton Retrieval with P-Norm Queries and Multiple Concept Types.* Dissertation, Cornell University, Ithaca, NY, August 1983.

[FRANCE] France, R.K. *An Artificial Intelligence Environment for Information Retrieval Research.* Masters Thesis, Virginia Polytechnic Institute and State University, July 1986.

[GELERNTER] Gelernter, J.L.; et al. "Empiricial Explorations of SYNCHEM." *Science.* 1041-1049, September 1977.

[HAHN] Hahn, U.; Reimer, U. "Condensation of Full Texts in the TOPIC System." ACM Transactions on Office Information Systems. 1985.

[HAYES-ROTH] Hayes-Roth, F.; Waterman, D.; Lenat, D. (editors): *Building Expert Systems.* Reading, MA: Addison-Wesley Publishing. 1983.

[HELLY] Helly, J.J. *A Distributed Expert System for Space Shuttle Flight Control.* Report No. CSD-840038, UCLA, Los Angeles, CA, 1984.

[HENDRIX] Hendrix, G.G. "Modeling Simultaneous Actions and Continuous Processes." *Artificial Intelligence*. 4: 145-180, 1982.

[KADUSHIN] Kadushin, A. *Child Welfare Services*. New York: MacMillan Publishing Co., Inc., 1980.

[KUNZ] Kunz, J.; et al. *A Physiological Rule-Based System for Interpreting Pulmonary Function Test Results*. Technical Report HPP-78-19, Stanford University, 1978.

[LEE] Lee, N.S.; Roach, J.W. *GUESS: A General Purpose Expert System Shell*. TR-85-3, Virginia Polytechnic Institute and State University, Blacksburg, VA, 1985.

[LEHNERT] Lehnert, W.; Black, J.B.; Reiser, B.J. "Summarizing Narratices." *Proc. of IJCAI 7*. August 1981.

[MCCUNE] McCune, B.P.; Tong, R.M.; Dean, J.S.; Shapiro, D.G. "RUBRIC: A System for Rule-Based Information Retrieval." In *Proc. of IEEE 7th International Computer Software and Applications Conference*. Chicago, IL, November 1983.

[MCDERMOTT] McDermott, J. "R1: An Expert in the Computer Systems Domain." In *Proc. of the First National Conf. on Artificial Intelligence*. 1980.

[MINSKY] Minsky, M. "A Framework for Representing Knowledge. " In P. Winston (editor): *The Psychology of Computer Vision*. New York: McGraw-Hill. 1975.

[NII] Nii, H.P.; Aiello, N. "AGE(Attempt to Generalize): A Knowledge-Based Program for Building Knowledge-Based Programs." *IJCAI*. 645-655, 1979.

[PAUKER] Pauker, S.G.; Gorry, G.; Kassirer, J.; Schivarty, W. "Towards the Simulation of Clinical Cognition." *The American Journal of Medicine*. 60: 981, June 1976.

[POLITAKIS] Politakis, P.; Weiss, S.M. "Using Empirical Analysis to Refine Expert System Knowledge Bases." *Artificial Intelligence.* 22: 23-48, 1984.

[POPLE] Pople, H.E. "DIALOG: A Model of Diagnostic Logic for Internal Medicine." *Proceedings of the International Conference on Artificial Intelligence.* 1975.

[POPLE] Pople, H. "The Formation of Composite Hypotheses in Diagnostic Problem Solving - an Exercies in Synthetic Reasoning." *IJCAI.* 1030-1037, 1977.

[RAPPAPORT] Rappaport, J. *Community Psychology.* New York: Holt, Rinehart, and Winston.,1977.

[RICH a] Rich, E. "Using Modeling via Stereotypes." *Cognitive Science.* 3: 329-354, 1979.

[RICH b] Rich, E. *Artificial Intelligence.* New York: McGraw-Hill, 1983.

[ROACH] Roach, J.W.; Fowler, G. The HC Manual: *Virginia Tech Prolog Technical Manual.* Virginia Polytechnic Institute and State University, Blacksburg, VA, 1983.

[SALTON] Salton, G.; McGill, M.J. *Introduction to Modern Information Retrieval.* New York: McGraw-Hill, 1983.

[SCHOECH] Schoech, D.; Jennings, H.; Schkade, L.; Hooper-Russell, C. "Expert Systems: Artificial Intelligence for Professional Decisions." *Computers in Human Services.* 1(1): 81-117, Spring 1985.

[SCHANK] Schank, R.C.; Riesback, C.K. (editor): *Inside Computer Understanding: Five Programs Plus Miniatures.* Hillsdale, NJ: Lawrence Erlbaum, 1981.

[SEGAL] Segal, S.P.; Baumohl, J.; Nelson, L.S. "Seeking Person-Environment Fit in Commmunity Care Placment." In R. Budson (editor): *Issues in Community Residental Care.* San Francisco: Jossey-Bass, 1987.

[SHORTLIFFE] Shortliffe, E.H. *Computer-Based Medical Consultations: MYCIN.*

New York: American-Elsevier, 1976.

[SWARTOUT] Swartout, W.R. "XPLAIN: A System for Creating and Explaining Expert Consulting Programs." *Artificial Intelligence.* 21: 285-325, 1983.

[VERE] Vere, S.A. "Planning in Time: Windows and Durations for Activities and Goals." *PAMI.* 5(3): 53-60, May 1983.

[VDSS] Virginia Department of Social Services. *Policy Manual 7.* 3(3), July 1984.

[VIRKAR] Virkar, R.S. *Development of POMME: The Pest and Orchard Management Expert System.* Master's Thesis, Virginia Polytechnic Institute and State University Computer Science Dept. December 1984.

[WEISSa] Weiss, S.M.; Kulikowski, C.A.; Safir, A. "A Model-Based Consultation System for the Long-Term Management of Glaucoma." *IJCAI.* 826-832, 1977.

[WEISSb] Weiss, S.M.; Kulikowski, C.A. "EXPERT: A System for Developing Consultation Nodes." *IJCAI.* 642-947, 1979.

[WINETT] Winett, S.; Fox, E.A. "Information Retrieval Techniques in an Expert System for Foster Care." *Proc. Expert Systems in Gov. Symp.* 391-396, Oct. 1985.

[WINETT] Winett, S.; Fox, E.A. "Using Information Retrieval Techniques in an Expert System." *Proc. Second Int. Conf. on Artificial Intelligence Applications.* Dec. 11-13, 1985.

[YIP] Yip, M. *An Expert System for Document Retrieval.* Masters Thesis, M.I.T., February 1981.

# Appendix 1. Data Collection Instruments

## DATA COLLECTION SHEET
### Child Record

CASE No.________________ NAME____________________________

GENDER_________ RACE__________ DOB__________________

No. OF SIBLINGS: _________

F.C STATUS OF SIBLINGS: ____________________________________________

SCHOOL INFO: ____________________________________________________

____________________________________________________

CURRENT PLACEMENT______________________

REASON(S) FOR CURRENT PLACEMENT_____________________________________

_____________________________________

IN FOSTER CARE IN PAST? ( Y N )

| PRIOR PLACEMENT | REASON FOR LEAVING |
|---|---|
| 1 ____________________ | ________________________ |
| 2 ____________________ | ________________________ |
| 3 ____________________ | ________________________ |
| 4 ____________________ | ________________________ |
| 5 ____________________ | ________________________ |
| 6 ____________________ | ________________________ |

SPECIAL NEEDS/NOTABLE BEHAVIORS

COMMENTS:

# DATA COLLECTION SHEET
## Foster Home Record

CASE No.__________________ NAME____________________________

DATE OPENED(approval date)_______________ RACE______________

TYPE OF HOUSING: SFD APT FARM INSTITUTION GRP.HOME OTHER

MEMBERS OF FOSTER FAMILY LIVING IN HOME:

| Relation to Head | DOB | Gender |
|---|---|---|
| ____(SELF)__________ | _______ | _____ |
| ____________________ | _______ | _____ |
| ____________________ | _______ | _____ |
| ____________________ | _______ | _____ |
| ____________________ | _______ | _____ |
| ____________________ | _______ | _____ |
| ____________________ | _______ | _____ |

APPROVAL INFORMATION

| NUMBER | GENDER | AGE RANGE | SPECIAL NEEDS |
|---|---|---|---|
| ------ | ------ | --------- | ------------- |

PREFERENCE INFORMATION

| NUMBER | GENDER | AGE RANGE | SPECIAL NEEDS |
|---|---|---|---|
| ------ | ------ | --------- | ------------- |

No. of Children Currently in Care____________________

**Children Currently in Care:**

| NAME | DOB |
| --- | --- |
| ______________________ | _______ |
| ______________________ | _______ |
| ______________________ | _______ |
| ______________________ | _______ |
| ______________________ | _______ |
| ______________________ | _______ |
| ______________________ | _______ |
| ______________________ | _______ |

PROHIBITION INFORMATION

| GENDER | AGE RANGE | SPECIAL NEEDS |
| --- | --- | --- |
| ------ | --------- | ------------- |

**Training**

| TYPE TAKEN | DATE |
| --- | --- |
| ---------- | ---- |
| ___________ | ___________ |
| ___________ | ___________ |
| ___________ | ___________ |
| ___________ | ___________ |
| ___________ | ___________ |

**Previously Successful Placements:**

| NAME | DOB |
|---|---|

**Previously Failed Placements:**

| NAME | DOB |
|---|---|

Notable
Strengths________________________________________________

Notable
Weaknesses______________________________________________

Comments:

# Appendix 2. Listing of Cosine Coefficients for Each <child, home> Pair

```
; "match.hc"
(assert
    ((match "000001" "000200" 0.6844))
    ((match "000001" "000201" 0.2636))
    ((match "000001" "000202" 0.5352))
    ((match "000001" "000203" 0.2212))
    ((match "000001" "000204" 0.5389))
    ((match "000001" "000205" 0.6212))
    ((match "000001" "000206" 0.2596))
    ((match "000001" "000207" 0.2610))
    ((match "000001" "000208" 0.6902))
    ((match "000001" "000209" 0.0387))
    ((match "000001" "000210" 0.6355))
    ((match "000001" "000211" 0.6181))
    ((match "000001" "000212" 0.2144))
    ((match "000001" "000213" 0.4567))
    ((match "000001" "000214" 0.6323))
    ((match "000002" "000200" 0.4547))
    ((match "000002" "000201" 0.2695))
    ((match "000002" "000202" 0.5378))
    ((match "000002" "000203" 0.2283))
    ((match "000002" "000204" 0.5548))
    ((match "000002" "000205" 0.6351))
    ((match "000002" "000206" 0.2602))
    ((match "000002" "000207" 0.2708))
    ((match "000002" "000208" 0.4569))
    ((match "000002" "000209" 0.2919))
    ((match "000002" "000210" 0.6342))
    ((match "000002" "000211" 0.6244))
    ((match "000002" "000212" 0.2285))
    ((match "000002" "000213" 0.7035))
    ((match "000002" "000214" 0.6414))
    ((match "000003" "000200" 0.6918))
    ((match "000003" "000201" 0.2555))
    ((match "000003" "000202" 0.5320))
    ((match "000003" "000203" 0.2145))
    ((match "000003" "000204" 0.5520))
    ((match "000003" "000205" 0.6326))
    ((match "000003" "000206" 0.2511))
    ((match "000003" "000207" 0.2586))
    ((match "000003" "000208" 0.6954))
    ((match "000003" "000209" 0.0445))
    ((match "000003" "000210" 0.6344))
    ((match "000003" "000211" 0.6291))
    ((match "000003" "000212" 0.2181))
```

```
((match "000003" "000213" 0.4553))
((match "000003" "000214" 0.6347))
((match "000004" "000200" 0.4629))
((match "000004" "000201" 0.2690))
((match "000004" "000202" 0.5250))
((match "000004" "000203" 0.2318))
((match "000004" "000204" 0.5672))
((match "000004" "000205" 0.6396))
((match "000004" "000206" 0.2520))
((match "000004" "000207" 0.2783))
((match "000004" "000208" 0.4462))
((match "000004" "000209" 0.3302))
((match "000004" "000210" 0.6135))
((match "000004" "000211" 0.6232))
((match "000004" "000212" 0.2440))
((match "000004" "000213" 0.6979))
((match "000004" "000214" 0.6444))
((match "000005" "000200" 0.4525))
((match "000005" "000201" 0.2638))
((match "000005" "000202" 0.5360))
((match "000005" "000203" 0.2217))
((match "000005" "000204" 0.5471))
((match "000005" "000205" 0.6301))
((match "000005" "000206" 0.2588))
((match "000005" "000207" 0.2620))
((match "000005" "000208" 0.4521))
((match "000005" "000209" 0.2785))
((match "000005" "000210" 0.6358))
((match "000005" "000211" 0.6261))
((match "000005" "000212" 0.2177))
((match "000005" "000213" 0.6959))
((match "000005" "000214" 0.6400))
((match "000006" "000200" 0.6917))
((match "000006" "000201" 0.2638))
((match "000006" "000202" 0.5360))
((match "000006" "000203" 0.2217))
((match "000006" "000204" 0.5471))
((match "000006" "000205" 0.6301))
((match "000006" "000206" 0.2588))
((match "000006" "000207" 0.2620))
((match "000006" "000208" 0.6921))
((match "000006" "000209" 0.0486))
((match "000006" "000210" 0.6358))
((match "000006" "000211" 0.6261))
((match "000006" "000212" 0.2177))
((match "000006" "000213" 0.4593))
((match "000006" "000214" 0.6400))
((match "000007" "000200" 0.6162))
((match "000007" "000201" 0.2486))
((match "000007" "000202" 0.7410))
((match "000007" "000203" 0.4664))
```

```
((match "000007" "000204" 0.7114))
((match "000007" "000205" 0.5594))
((match "000007" "000206" 0.2453))
((match "000007" "000207" 0.2463))
((match "000007" "000208" 0.6213))
((match "000007" "000209" 0.0539))
((match "000007" "000210" 0.6189))
((match "000007" "000211" 0.6033))
((match "000007" "000212" 0.4580))
((match "000007" "000213" 0.4180))
((match "000007" "000214" 0.5499))
((match "000008" "000200" 0.7022))
((match "000008" "000201" 0.2534))
((match "000008" "000202" 0.5233))
((match "000008" "000203" 0.2132))
((match "000008" "000204" 0.5639))
((match "000008" "000205" 0.6487))
((match "000008" "000206" 0.2470))
((match "000008" "000207" 0.2548))
((match "000008" "000208" 0.6803))
((match "000008" "000209" 0.0790))
((match "000008" "000210" 0.6213))
((match "000008" "000211" 0.6435))
((match "000008" "000212" 0.2177))
((match "000008" "000213" 0.4523))
((match "000008" "000214" 0.6520))
((match "000009" "000200" 0.7022))
((match "000009" "000201" 0.2534))
((match "000009" "000202" 0.5233))
((match "000009" "000203" 0.2132))
((match "000009" "000204" 0.5639))
((match "000009" "000205" 0.6487))
((match "000009" "000206" 0.2470))
((match "000009" "000207" 0.2548))
((match "000009" "000208" 0.6803))
((match "000009" "000209" 0.0790))
((match "000009" "000210" 0.6213))
((match "000009" "000211" 0.6435))
((match "000009" "000212" 0.2177))
((match "000009" "000213" 0.4523))
((match "000009" "000214" 0.6520))
((match "000010" "000200" 0.4524))
((match "000010" "000201" 0.2681))
((match "000010" "000202" 0.5341))
((match "000010" "000203" 0.2279))
((match "000010" "000204" 0.5516))
((match "000010" "000205" 0.6286))
((match "000010" "000206" 0.2573))
((match "000010" "000207" 0.2699))
((match "000010" "000208" 0.4547))
((match "000010" "000209" 0.2933))
```

```
((match "000010" "000210" 0.6291))
((match "000010" "000211" 0.6192))
((match "000010" "000212" 0.2332))
((match "000010" "000213" 0.7004))
((match "000010" "000214" 0.6369))
((match "000011" "000200" 0.0270))
((match "000011" "000201" 0.6381))
((match "000011" "000202" 0.2083))
((match "000011" "000203" 0.5289))
((match "000011" "000204" 0.2144))
((match "000011" "000205" 0.2471))
((match "000011" "000206" 0.6402))
((match "000011" "000207" 0.6356))
((match "000011" "000208" 0.0283))
((match "000011" "000209" 0.6671))
((match "000011" "000210" 0.2483))
((match "000011" "000211" 0.2452))
((match "000011" "000212" 0.5243))
((match "000011" "000213" 0.2709))
((match "000011" "000214" 0.2481))
((match "000012" "000200" 0.0266))
((match "000012" "000201" 0.6373))
((match "000012" "000202" 0.2080))
((match "000012" "000203" 0.5278))
((match "000012" "000204" 0.2126))
((match "000012" "000205" 0.2460))
((match "000012" "000206" 0.6403))
((match "000012" "000207" 0.6340))
((match "000012" "000208" 0.0269))
((match "000012" "000209" 0.6646))
((match "000012" "000210" 0.2484))
((match "000012" "000211" 0.2452))
((match "000012" "000212" 0.5218))
((match "000012" "000213" 0.2690))
((match "000012" "000214" 0.2482))
((match "000013" "000200" 0.0341))
((match "000013" "000201" 0.6461))
((match "000013" "000202" 0.2202))
((match "000013" "000203" 0.5385))
((match "000013" "000204" 0.2184))
((match "000013" "000205" 0.2474))
((match "000013" "000206" 0.6425))
((match "000013" "000207" 0.6436))
((match "000013" "000208" 0.0393))
((match "000013" "000209" 0.6707))
((match "000013" "000210" 0.2554))
((match "000013" "000211" 0.2403))
((match "000013" "000212" 0.5320))
((match "000013" "000213" 0.2887))
((match "000013" "000214" 0.2541))
((match "000014" "000200" 0.4444))
```

```
((match "000014" "000201" 0.2551))
((match "000014" "000202" 0.5253))
((match "000014" "000203" 0.2163))
((match "000014" "000204" 0.5511))
((match "000014" "000205" 0.6236))
((match "000014" "000206" 0.2461))
((match "000014" "000207" 0.2645))
((match "000014" "000208" 0.4502))
((match "000014" "000209" 0.2848))
((match "000014" "000210" 0.6234))
((match "000014" "000211" 0.6147))
((match "000014" "000212" 0.2300))
((match "000014" "000213" 0.6908))
((match "000014" "000214" 0.6252))
((match "000015" "000200" 0.4627))
((match "000015" "000201" 0.2550))
((match "000015" "000202" 0.5187))
((match "000015" "000203" 0.2178))
((match "000015" "000204" 0.5680))
((match "000015" "000205" 0.6423))
((match "000015" "000206" 0.2424))
((match "000015" "000207" 0.2652))
((match "000015" "000208" 0.4426))
((match "000015" "000209" 0.3208))
((match "000015" "000210" 0.6120))
((match "000015" "000211" 0.6312))
((match "000015" "000212" 0.2353))
((match "000015" "000213" 0.6879))
((match "000015" "000214" 0.6440))
((match "000016" "000200" 0.2960))
((match "000016" "000201" 0.6232))
((match "000016" "000202" 0.4988))
((match "000016" "000203" 0.8101))
((match "000016" "000204" 0.5033))
((match "000016" "000205" 0.2819))
((match "000016" "000206" 0.6173))
((match "000016" "000207" 0.6245))
((match "000016" "000208" 0.2681))
((match "000016" "000209" 0.4800))
((match "000016" "000210" 0.2929))
((match "000016" "000211" 0.3258))
((match "000016" "000212" 0.8136))
((match "000016" "000213" 0.0504))
((match "000016" "000214" 0.2577))
((match "000017" "000200" 0.2694))
((match "000017" "000201" 0.6312))
((match "000017" "000202" 0.2088))
((match "000017" "000203" 0.5264))
((match "000017" "000204" 0.2246))
((match "000017" "000205" 0.2471))
((match "000017" "000206" 0.6274))
```

```
    ((match "000017" "000207" 0.6383))
    ((match "000017" "000208" 0.2717))
    ((match "000017" "000209" 0.4459))
    ((match "000017" "000210" 0.2458))
    ((match "000017" "000211" 0.2409))
    ((match "000017" "000212" 0.5331))
    ((match "000017" "000213" 0.0446))
    ((match "000017" "000214" 0.2460))
)
```

# Appendix 3. Recall – Precision Results for All Children

*** QUERY 1 ***

| rank | did | recall | precision |
|---|---|---|---|
| 1 | 9 | 0.1429 | 1.0000 |
| 2 | 1 | 0.2857 | 1.0000 |
| 3 | 11 | 0.4286 | 1.0000 |
| 4 | 15 | 0.5714 | 1.0000 |
| 6 | 12 | 0.7143 | 0.8333 |
| 9 | 14 | 0.8571 | 0.6667 |
| 10 | 3 | 1.0000 | 0.7000 |

*** QUERY 2 ***

| rank | did | recall | precision |
|---|---|---|---|
| 1 | 14 | 0.5000 | 1.0000 |
| 6 | 5 | 1.0000 | 0.3333 |

*** QUERY 3 ***

| rank | did | recall | precision |
|---|---|---|---|
| 3 | 15 | 0.1667 | 0.3333 |
| 4 | 11 | 0.3333 | 0.5000 |
| 6 | 12 | 0.5000 | 0.5000 |
| 7 | 5 | 0.6667 | 0.5714 |
| 8 | 3 | 0.8333 | 0.6250 |
| 9 | 14 | 1.0000 | 0.6667 |

*** QUERY 4 ***

| rank | did | recall | precision |
|---|---|---|---|
| 1 | 14 | 0.2000 | 1.0000 |
| 2 | 15 | 0.4000 | 1.0000 |
| 3 | 6 | 0.6000 | 1.0000 |
| 6 | 5 | 0.8000 | 0.6667 |
| 9 | 9 | 1.0000 | 0.5556 |

*** QUERY 5 ***

| rank | did | recall | precision |
|---|---|---|---|
| 1 | 14 | 0.2000 | 1.0000 |
| 2 | 15 | 0.4000 | 1.0000 |
| 4 | 6 | 0.6000 | 0.7500 |
| 7 | 3 | 0.8000 | 0.5714 |
| 8 | 1 | 1.0000 | 0.6250 |

*** QUERY 6 ***

| rank | did | recall | precision |
|---|---|---|---|
| 1 | 9 | 0.1429 | 1.0000 |
| 2 | 1 | 0.2857 | 1.0000 |
| 3 | 15 | 0.4286 | 1.0000 |
| 4 | 11 | 0.5714 | 1.0000 |
| 6 | 12 | 0.7143 | 0.8333 |
| 8 | 3 | 0.8571 | 0.7500 |
| 9 | 14 | 1.0000 | 0.7778 |

*** QUERY 7 ***

| rank | did | recall | precision |
|---|---|---|---|
| 1 | 3 | 0.2000 | 1.0000 |
| 3 | 9 | 0.4000 | 0.6667 |
| 4 | 11 | 0.6000 | 0.7500 |
| 5 | 1 | 0.8000 | 0.8000 |
| 6 | 12 | 1.0000 | 0.8333 |

*** QUERY 8 ***

| rank | did | recall | precision |
|---|---|---|---|
| 1 | 1 | 0.1667 | 1.0000 |
| 2 | 9 | 0.3333 | 1.0000 |
| 5 | 12 | 0.5000 | 0.6000 |
| 6 | 11 | 0.6667 | 0.6667 |
| 8 | 3 | 0.8333 | 0.6250 |
| 9 | 14 | 1.0000 | 0.6667 |

*** QUERY 9 ***

| rank | did | recall | precision |
|---|---|---|---|
| 1 | 1 | 0.1667 | 1.0000 |
| 2 | 9 | 0.3333 | 1.0000 |
| 5 | 12 | 0.5000 | 0.6000 |
| 6 | 11 | 0.6667 | 0.6667 |
| 8 | 3 | 0.8333 | 0.6250 |
| 9 | 14 | 1.0000 | 0.6667 |

*** QUERY 10 ***

| rank | did | recall | precision |
|---|---|---|---|
| 1 | 14 | 0.2000 | 1.0000 |
| 2 | 15 | 0.4000 | 1.0000 |
| 4 | 6 | 0.6000 | 0.7500 |
| 6 | 5 | 0.8000 | 0.6667 |
| 8 | 9 | 1.0000 | 0.6250 |

*** QUERY 11 ***

| rank | did | recall | precision |
|---|---|---|---|
| 1 | 10 | 0.3333 | 1.0000 |
| 4 | 8 | 0.6667 | 0.5000 |
| 5 | 4 | 1.0000 | 0.6000 |

*** QUERY 12 ***

| rank | did | recall | precision |
|---|---|---|---|
| 1 | 10 | 0.2000 | 1.0000 |
| 2 | 7 | 0.4000 | 1.0000 |
| 3 | 2 | 0.8000 | 1.0000 |
| 5 | 4 | 1.0000 | 1.0000 |

*** QUERY 13 ***

| rank | did | recall | precision |
|---|---|---|---|
| 1 | 10 | 0.1667 | 1.0000 |
| 2 | 2 | 0.333 | 1.0000 |
| 3 | 8 | 0.5000 | 1.0000 |
| 4 | 7 | 0.6667 | 1.0000 |
| 5 | 4 | 0.8333 | 1.0000 |
| 6 | 13 | 1.0000 | 1.0000 |

*** QUERY 14 ***

| rank | did | recall | precision |
|---|---|---|---|
| 1 | 14 | 0.5000 | 1.0000 |
| 6 | 5 | 1.0000 | 0.3333 |

*** QUERY 15 ***

| rank | did | recall | precision |
|---|---|---|---|
| 1 | 14 | 0.2500 | 1.0000 |
| 3 | 6 | 0.5000 | 0.6667 |
| 6 | 5 | 0.7500 | 0.5000 |
| 9 | 9 | 1.0000 | 0.4444 |

*** QUERY 16 ***

| rank | did | recall | precision |
|---|---|---|---|
| 1 | 13 | 0.2500 | 1.0000 |
| 2 | 4 | 0.5000 | 1.0000 |
| 3 | 8 | 0.7500 | 1.0000 |
| 4 | 2 | 1.0000 | 1.0000 |

*** QUERY 17 ***

| rank | did | recall | precision |
|---|---|---|---|
| 1 | 8 | 0.2500 | 1.0000 |

| | | | |
|---|---|---|---|
| 4 | 13 | 0.5000 | 0.5000 |
| 5 | 4 | 0.7500 | 0.6000 |
| 6 | 10 | 1.0000 | 0.6667 |

| Recall – Precision | Level 1 |
|---|---|
| 0.00 | .9804 |
| 0.05 | .9804 |
| 0.10 | .9804 |
| 0.15 | .9804 |
| 0.20 | .9804 |
| 0.25 | .9706 |
| 0.30 | .9314 |
| 0.35 | .8686 |
| 0.40 | .8686 |
| 0.45 | .8392 |
| 0.50 | .8392 |
| 0.55 | .7510 |
| 0.60 | .7314 |
| 0.65 | .6995 |
| 0.70 | .6995 |
| 0.75 | .6884 |
| 0.80 | .6851 |
| 0.85 | .6761 |
| 0.90 | .6761 |
| 0.95 | .6761 |
| 1.00 | .6761 |

Average precison for 3 intermediate points
Prec. = 0.8327

| Statistic | 1 |
|---|---|
| Norm Recall | 0.9986 |
| Norm Precision | 0.9571 |
| Rank Recall | 0.4588 |
| Log Precision | 0.7595 |
| Precision after 10 docs | 0.4824 |
| Precion after 30 docs | 0.1608 |
| Recall after 10 docs | 1.0000 |
| Recall after 30 docs | 1.0000 |
| E, 0.5, 10 docs | 0.4668 |
| E, 1.0, 10 docs | 0.3634 |
| E, 2.0, 10 docs | 0.1977 |
| E, 0.5, 30 docs | 0.8075 |
| E, 1.0, 30 docs | 0.7261 |
| E, 2.0, 30 docs | 0.5225 |
| Number queries | 17 |

**The three page vita has been removed from the scanned document. Page 1 of 3**

**The three page vita has been removed from the scanned document. Page 2 of 3**

**The three page vita has been removed from the scanned document. Page 3 of 3**