VICTORIA UNIVERSITY OF TECHNOLOGY

DEPARTMENT OF COMPUTER AND MATHEMATICAL SCIENCES

An Intelligent KBS
Learning Census Data from Examples

Richard Satur, Zhi-Qiang Liu and Mark Gahegan

(42 COMP 13)

August, 1994

(AMS : 68Q40)

TECHNICAL REPORT

VICTORIA UNIVERSITY OF TECHNOLOGY
(P O BOX 14428) MELBOURNE MAIL CENTRE
MELBOURNE, VICTORIA, 3000
AUSTRALIA

TELEPHONE (03) 688 4249 / 4492
FACSIMILE (03) 688 4050

Footscray Campus

An Intelligent KBS. Learning Census Data from Examples.

by Richard Satur, Zhi-Qiang Liu[♦] and Mark Gahegan[•]
Department of Computer and Mathematical Sciences
Victoria University of Technology, Victoria, Australia

Abstract:

In this paper we propose a Knowledge-Based System (KBS) that incorporates intelligence using methods of learning relational structures and evidential reasoning. Theories of causality in data and the role that learning paradigms play in generalising data, in particular evidential inductive learning schemes using fuzzy logic are discussed. This work is only the first step to providing an adaptive KBS that attempts to explain and categorise data. Census data are used to illustrate the effectiveness of this structure and the problems that human experts face when classifying data as a means to providing generalised explanations for subsequent human activity. It is envisaged that our approach will prove significant in application areas that involve processes for economic decision rationale and planning.

1. Introduction:

Human experts have for a long time played a significant role in the development of Expert Systems and the way in which expert systems are used to explain occurrences

---

[♦] Zhi-Qiang Liu is a Computer Science lecturer, in the Department of Computer Science, at The University of Melbourne, Victoria, Australia.
[•] Mark Gahegan is a GIS lecturer in the Department of Computer Science, at Curtin University of Technology, Western Australia, Australia.

that they themselves have to deal with on a daily basis [Bonissone 1985] [Parsaye 1988] [Turban 1988]. Often the finer structures that define the lower level organisation of their knowledge are obscured by years of practical and intuitive experience that have lead them to make decisions based on these 'built-up' generalisations. Human experts are capable of many levels of reasoning and, in particular high performance techniques of generalisations and associations. Even more important is to make decisions based on "*Context*", including both temporal and spatial, information that produces varying and in many cases different levels of reasoning for apparently equivalent instances with a collection of data sets.

It has been recognised that knowledge-based systems (KBS) need to model the real-world if they are to be used for real-world applications, especially in decision-making tasks. Modelling the real-world is fraught however, with enormous problems. The following major issues need to be addressed:

- developing a structure that is able to perform the representational tasks of a real-world system,
- capturing generalisations about the data such that the system is to a certain extent either non-committal about details that cannot be resolved and can explain minute, in many cases, differences, about apparently similar instances,
- dealing with contradictory (negative instances) knowledge,
- representing spatial and temporal references in data and
- developing dynamic knowledge structures that adapt to new data sets and expert intervention.

Woods succinctly describes the role of a knowledge representation system as a system "*called upon to support activities of perception, learning, and planning to act*" [Woods 1991]. To fulfil these primary tasks, there is a suggestion for the need of a formal taxonomic organisation. We propose a structure that:

- assimilates rules into such a taxonomic knowledge structure to facilitate discovery of interactions of objects (ie: concepts) at input time,
- formalises a compact structure in the specification of these rules,
- is a generalised description of the real-problem, representing contextual information, and
- is an adaptive, evolutionary system incorporating causality by methods of fuzzy evidential reasoning.

In our approach we use a multi-layered structure which possesses the following basic characteristics [Fike 1985]:

- expressiveness in its representation of the human knowledge and hence of the data itself,
- the structure can be interpreted by the human expert in a way that allows the human expert to interact with the system, (ie: reinforcement)

This paper endeavours to present a structure that also in part provides a learning scheme for building a multi-layered structure which represents data in a contextual form. This work is the basis for an Object-Oriented Context-Based System (OOCBS) which will be discussed. The structure is based on Fuzzy Cognitive Maps (FCMs) [Kosko 1986] and learning paradigms primarily concerned with supervised inductive learning models that also use processes of generalisation for reducing search spaces. We will discuss the interpretations formed by various machine learning techniques used in our system using census data collected by the Australian Bureau of Statistics†. Human experts will verify the development and validity of this KBS‡ .

2. FUZZY Cognitive Maps:

It has been recognised that human knowledge for the most part contains uncertainty. In dynamic environments such knowledge describes relevant concepts and causes for certain actions to be taken. In order to capture the adaptive nature of the human knowledge and to assist decision making, Axelrod [Axelrod 1976] proposed cognitive maps for representing social knowledge. However, such cognitive maps are based on a rigid structure with fixed measures of causality between different social events. As a consequence, it is unable to handle dynamically changing environments and is not effective to be used in adaptive reasoning systems.

Based on these concerns, in the late 1980's Kosko [Kosko 1986] introduced the fuzzy cognitive map (FCM) which incorporates fuzzy causality measures in the original cognitive maps. FCM provides a flexible and more realistic representation scheme for dealing with knowledge. This scheme is potentially useful in the problem domains that we are dealing with, which are commonly referred to as the soft knowledge domain where both system concepts and relationships and the meta-system knowledge can only

---

† Australian Bureau of Statistics, Level 1, Hyatt Centre, 30 Terrace Road, EAST PERTH, 6004. Phone: 61-9-3235140, FAX: 61-9-2212374. 1991 Census of Population and Housing. Basic Community Profiles.

‡ The finding and suggested outcomes of this work is not shared in any part by the Australian Bureau of Statistics and should not be used nor interpreted to mean anything more than just one representation of the data.

be represented to a certain degree. In addition, subtle (spatial and temporal) variations in the knowledge base can often result in completely different outcomes or decisions.

FCM provides a mechanism for representing such hazy degrees of causality between events/objects. This enables the constructed paths to propagate causality in a more natural fashion through the use of such techniques as forward and backward chaining. In this paper we propose a multi-layered FCM for knowledge representation and adaptive inference. In the following we briefly describe the FCM structure used in our research on context-dependent learning systems.

2.1 Representing a FCM:

Cognitive maps in their original forms are too rigid for knowledge-base construction since causality for the most part is fuzzy. FCMs on the other hand provide for a formal framework for representing fuzzy causality and the propagation of causality within FCMs. Figure 2 shows an FCM which is constructed based on the definition of a regional shopping centre (See Figure 1) obtained from interviews with human experts. The 'goal' for this FCM fragment and in fact the entire FCM (KBS) is to give answers to question *"Do we wish to build a Regional Shopping Centre in this region?"*

Definition of a Regional Shopping Center in Western Australia, Australia:

Location:	Middle/Outer Suburbs
Size:	>= 30000 sq. MNLA retail
Shops Included:	Discount department stores (1-2) Major chain supermarkets (1-2) (possibly) Department store approx. 50 small tenancies Free standing take away food/restaurants
Offices:	Government region-serving offices Client servicing agencies (Social security etc...) Other large offices in very large center (50,000 sqm +) Small private offices; possibly large organizations
Other:	Possibly associated with social and entertainment facilities (Cinema, hotel etc...)

Figure 1: Definition of a Western Australian Regional Shopping Centre (Australia).

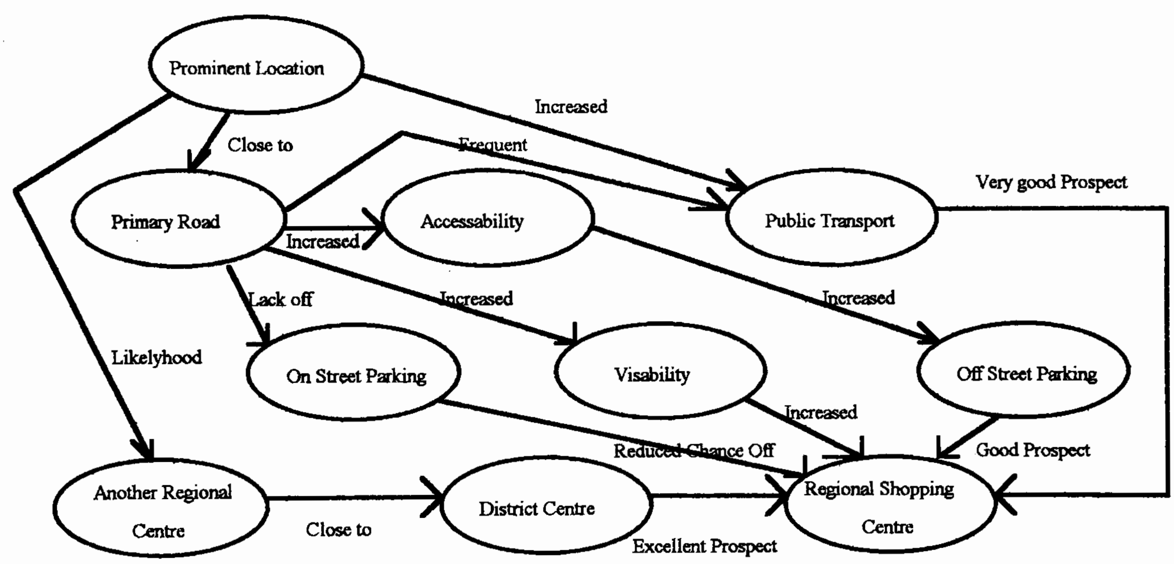


Figure 2: A fragment FCM of a larger FCM based on the Goal, "Should we build a shopping centre?"

In fuzzy causal algebra, it is not possible to state the obvious contraposition to "If X causes Y" (ie: "not-X causes not-Y" where 'X' and 'Y' refer to Objects in the real world) as would be the case for simple logical implication. One can refer however, to causally increased implication of X and Y. For example increasing X would cause an increase in Y if "X causes Y". This would also hold true for the causally decreasing relationships between fuzzy causal objects. If we think of causal objects as being represented as fuzzy subsets of some concept space (Ψ) where the fuzzy-set membership degree is the variance of the concept with respect to the concept space, then we can represent causal reasoning within a fuzzy logic set framework [Kosko 1986].

Definition 1: (a) Define fuzzy union or disjunction by $\cup$.

(b) Define fuzzy intersection or conjunction by $\cap$.

(c) Define fuzzy inclusion or logical implication by $\subset$.

Definition 2: A concept Ci is the fuzzy union of causally related fuzzy quantity sets Qi and respective dis-quantity set ~Qi, where ~Qi is the abstract negation or local fuzzy set complement of Qi and is the set index for (Qi,~Qi) which results in ~~Qi=Qi, or fuzzy causality for a concept Ci is expressed as $Ci = Qi \cup \sim Qi$.

Definition 3: We will use the symbols $\Leftrightarrow$, $\Rightarrow^{c}$, $\Rightarrow^{c\uparrow}$, $\Rightarrow^{c\downarrow}$ to represent 'iff', 'causes', 'causally increases' and 'causally decreases' respectively. Hence we say Ci causes Cj or $Ci \Rightarrow^{c} Cj$.

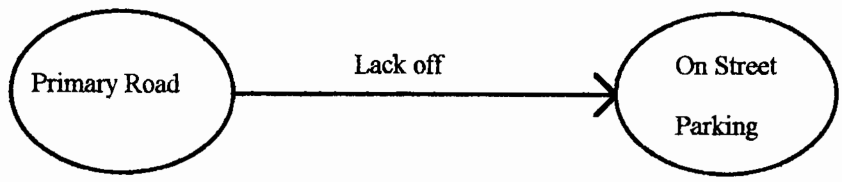
Definition 4: (a) $\{\exists Ci,Cj \in \Psi: (Ci \Rightarrow^{c} Cj) \Leftrightarrow ((Qi \subset Qj) \,\&\, (\sim Qi \subset \sim Qj))\}$

(b) $\{\exists Ci,Cj \in \Psi: (Ci \Rightarrow^{c\downarrow} Cj) \Leftrightarrow ((Qi \subset \sim Qj) \,\&\, (\sim Qi \subset Qj))\}$

Using definition 4(b) the negative causalities in the above FCM (see Figure 2) are replaced with positive correlates (See Figure 3). The general rule for causal link replacement in the FCMs construction, is:

Replacement Rule: Replace every Ci $\rightarrow_{-}$ Cj with Ci $\rightarrow_{+}$ ~Cj.

This FCM fragment:



Is equivalent (in the causal sense) to:

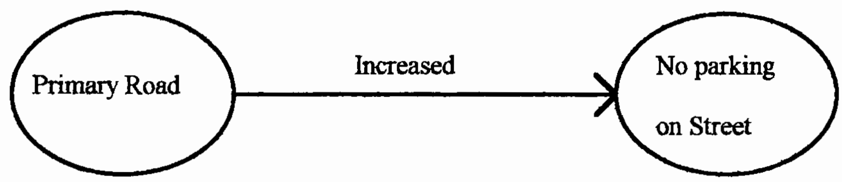


Figure 3: Replacing Negative Relationships with Positive correlates.

In the example above, (see Figure 3), the modifier equivalence ~~("no parking on street")=("On Street Parking") is true in the abstract negation sense. For real-world problems, unless we can guarantee the intervention of a human expert (supervised learning in the 'tutor' sense), absolute negation in the fuzzy sense, provides for correct link replacement. Therefore for some concept Ci there exists Qi, ~Qi and Mi that represent quantity, dis-quantity and modifier fuzzy sets, and:

Definition 5: Let U be the set universe, then initially Mi $\equiv$ U by default.

Definition 6: $\{\forall Ci \in \Psi: Ci = (Qi \cup \sim Qi) \cap Mi\}$

Definition 7: $\{\exists (Ci,Cj) \in \Psi:(Ci \Rightarrow^{c} Cj) \Leftrightarrow ((Qi \cap Mi) \subset (Qj \cap Mj))\&((\sim Qi \cap Mi) \subset (\sim Qj \cap Mj))\}$

Definition 8: $\{\exists (Ci,Cj) \in \Psi:(Ci \Rightarrow^{c\downarrow} Cj) \Leftrightarrow ((Qi \cap Mi) \subset (\sim Qj \cap Mj))\&((\sim Qi \cap Mi) \subset (Qj \cap Mj))\}$

Using the above rules and definitions in abstract fuzzy spaces, we obtain an abstract FCM framework and a fuzzy causal algebra, defined as follows;

Term	Definitions
FPS(X)	Fuzzy power set or the set of all fuzzy subsets of X
$d_s(A,B)$	Degree of subsethood of A in B [Bandler 1980]
$\lvert Q\rvert_x$	Quantity space on X
$\lvert M\rvert_x$	Modifier space on X
$\lvert C\rvert_x$	Concept space on X (contains the fuzzy nodes of the abstract FCM)
$\lvert \mathbf{F}\rvert_x$	Fuzzy cognitive space on X
$\lvert FCM\rvert_x$	Fuzzy cognitive map (fuzzy causal graph) on X
P	Partial ordered set (normally P = [0,1])
e	Fuzzy causal edge function

Table 1: List of Terms and Definitions for an abstract FCM framework and fuzzy causal algebra.

Let $\{X \in U : X \notin \varnothing\}$ & $FPS(X) = F(2^x)$

Then $\{A,B \in F(2^x) : d_s(A,B) = m_{F(2^B)}(A)\}$ [Bandler 1980]

Now call

$$((\zeta = |Q|_x) \subset F(2^x)) \Rightarrow \{\forall A \in \zeta) : \exists Q, \sim Q \in F(2^x) \;\&\; A = Q \cup \sim Q\}$$

Similarly

$$((\mathbf{M} = |M|_x) \subset F(2^x)) \Rightarrow \{X \in \mathbf{M}\}$$

and

$$((\xi = |C|_x) \subset F(2^x)) \Rightarrow \{\xi = \zeta \cap \mathbf{M}\} \text{ or } \xi = \{(Q \cup \sim Q) \cap M : Q \cup \sim Q \in \zeta,$$

$$M \in \mathbf{M}\}$$

Then ξ is causal if:

$$\{\forall (Ci, Cj) \in \xi, \exists (Qi \cup \sim Qi, Qj \cup \sim Qj \in \zeta) \;\&\; (Mi, Mj \in \mathbf{M}) :$$

$$(Qi \cap Mi \subset Qj \cap Mj \Rightarrow \sim Qi \cap Mi \subset \sim Qj \cap Mj) \;\&$$

$$(Qi \cap Mi \subset \sim Qj \cap Mj \Rightarrow \sim Qi \cap Mi \subset Qj \cap Mj)\}$$

Then if ξ is causal, $\mathbf{F} = (X, \xi)$ & $(\mathbf{F} \notin \varnothing)$,

Formally;

$$e : \xi \times \xi \rightarrow \mathbf{P} \text{ if } (e_{ij} = (Ci, Cj) = m_{F(2^{Cj})}(Ci)) \;\&\; (e_{ij} \neq 0) \;\&\; ((e, \xi)$$

is cycle free)

$$\Rightarrow (e, \mathbf{F}) = |FCM|_x$$

Term	Definitions
m-many causal paths from Ci to Cj	$(i,k_1^1,k_2^1,\ldots,k_{nl}^1)$ for $1 \leq l \leq m$
$I_l(Ci,Cj)$	The indirect effect of Ci on Cj on l causal paths
$T(Ci,Cj)$	The total effect of Ci on Cj over all m causal paths

Table 2: List of Terms and Definitions for fuzzy causal algebra.

[Kosko 1986] outlines a fuzzy causal algebra that governs causal propagation (forward and backward chaining) and causal combination on a FCM, and that depends only on the partial ordering of P,e and the general properties of fuzzy graph theory. For a FCM with m-many causal paths from concept Ci to Cj we can define the paths, indirect effect and total effect about any path from Ci to Cj (See Table 2) in terms of the minimum and maximum of all causal links from Ci to Cj respectively, and generally by t-norm and t-conorm respectively [Yager 1981].

Theorem 1: Given P then:

$I_l(Ci,Cj) = \min \{e(C_p,C_{p+1}): (p,p+1) \in (i,k_1^1,\ldots,k_{nl}^1,j)\}$ where p and p+1 are contiguous left-to-right path indices.

Theorem 2: Given P then:

$$T(Ci,Cj) = \max_{1 \leq l \leq m} I_l(Ci,Cj)$$

Fuzzy inputs may now be processed as if they were real valued inputs, the only trade-off is a fuzzy output, although adaptive defuzzification methods [Yager 1993] will prove useful when applied to a FCM's output.

2.2 A Multi-Layered Knowledge-Based System based on FCMs:

The structure we have used to represent human expert knowledge is a three layered structure. Each layer provides an intelligent index into the next layer, reducing the search space while providing similar reasoning procedures that might be used by human experts when considering sub-problems and contextual criteria.

The three layers of reasoning the KBS, in order of generalisation are (see Figure 4):

- **Goal Level:** Goals are desired outcomes linked by common criteria, exemplified by the shopping centre example, where commonality between the goals is derived from the similarities in the type of shopping centres that could be built. Linking these goals allows cross-referencing between goals that share commonality if in a way similar to analogous criteria in the KBS. This is the first layer and forms the largest generalised classes.
- **Context Level:** For each goal there are a number of generalisations or contexts. This layer is itself a FCM. In the example we considered earlier (the shopping centre FCM of Figure 2), the context was "access", in the sense of 'access to parking'. This layer provides the second level of generalisation.

- **Lowest FCM Level:** This is the final layer and is the layer of specialisation. For each context there can only be one FCM that links objects common in context, by causality. Although FCMs for different context might appear to be similar, it is the thread from goal to context to FCM that determine reasoning. The shopping centre example (see Figure 2) is just one example of this specialisation.

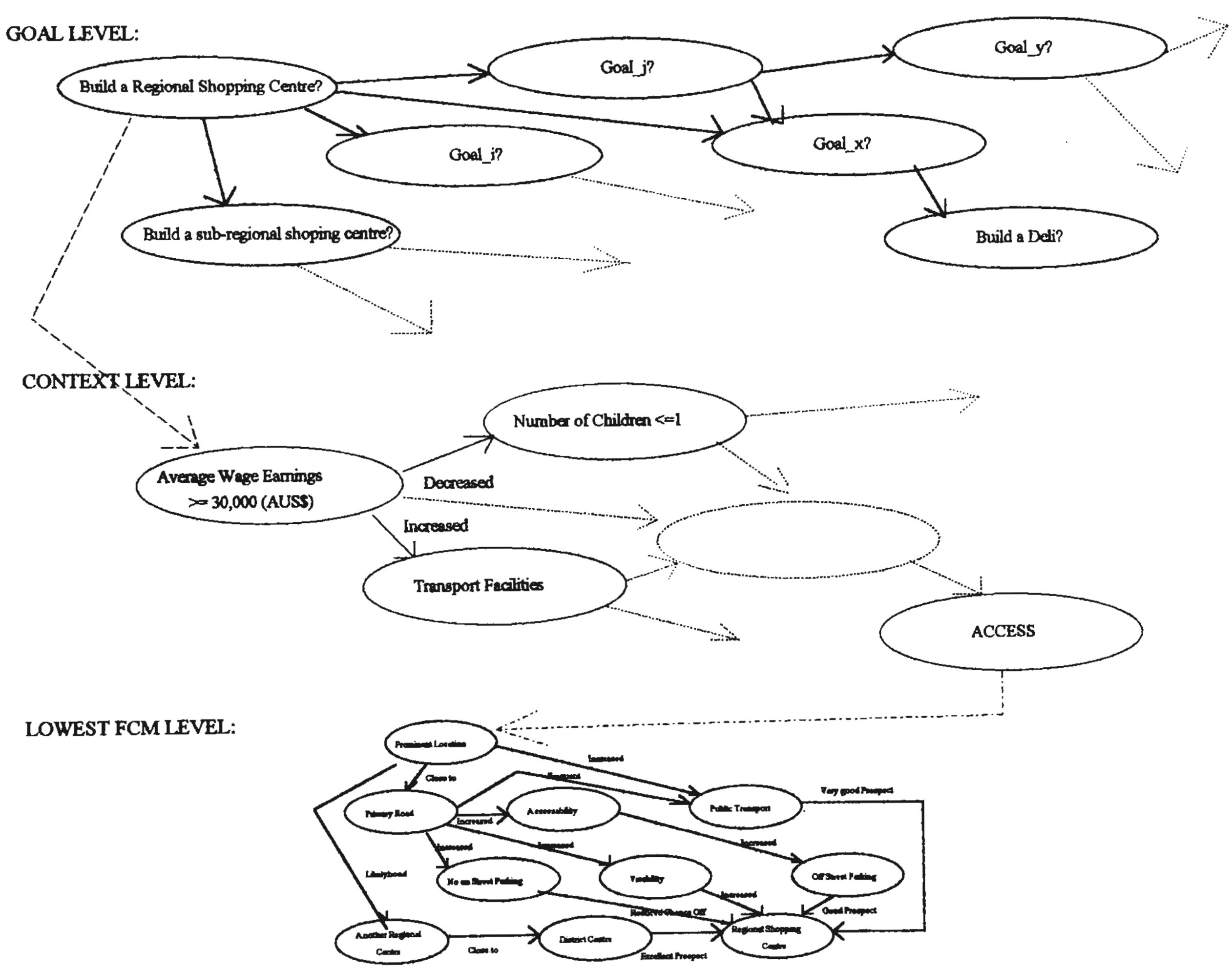


Figure 4: Part of the shopping centre Knowledge-Based System.

3. Learning From Examples:

The lowest level FCM level of this KBS is generated per context, using inductive concept acquisition from examples, or supervised learning. Learning by examples involves inferring descriptions, which may take the form of concept descriptions, classification rules, hypotheses or rules of a class (concept) from prototypical examples (related examples) for a class. The problem can be thus stated [Michalski 1983]:

- **Given:**
  * A set of prototypical examples (related examples, normally "attribute-value" pairs)
  * A initial inductive assertion (default null hypothesis)
  * Background knowledge including assertions, hypothesis and constraints.
- **Find:**
  * An inductive assertion (hypothesis) covering the following evaluation criteria:
    - Completeness (describes most/all of the positive examples),
    - Consistent (does not describe any of the negative examples), and
    - Satisfies the background knowledge.

The totality of the learning schedule is beyond the scope of this paper and is a combination of several learning paradigms, each is chosen for its robustness for a given task. It will surfice to indicate the learning paradigm used at each phase of the knowledge base construction.

3.1. The Overall System:

The overall architecture is a combination of human expert knowledge, a learning schedule that incorporate fuzzy logic for probabilistic classification and feedback to classify instances previously unclassified by human experts (see Figure 5).

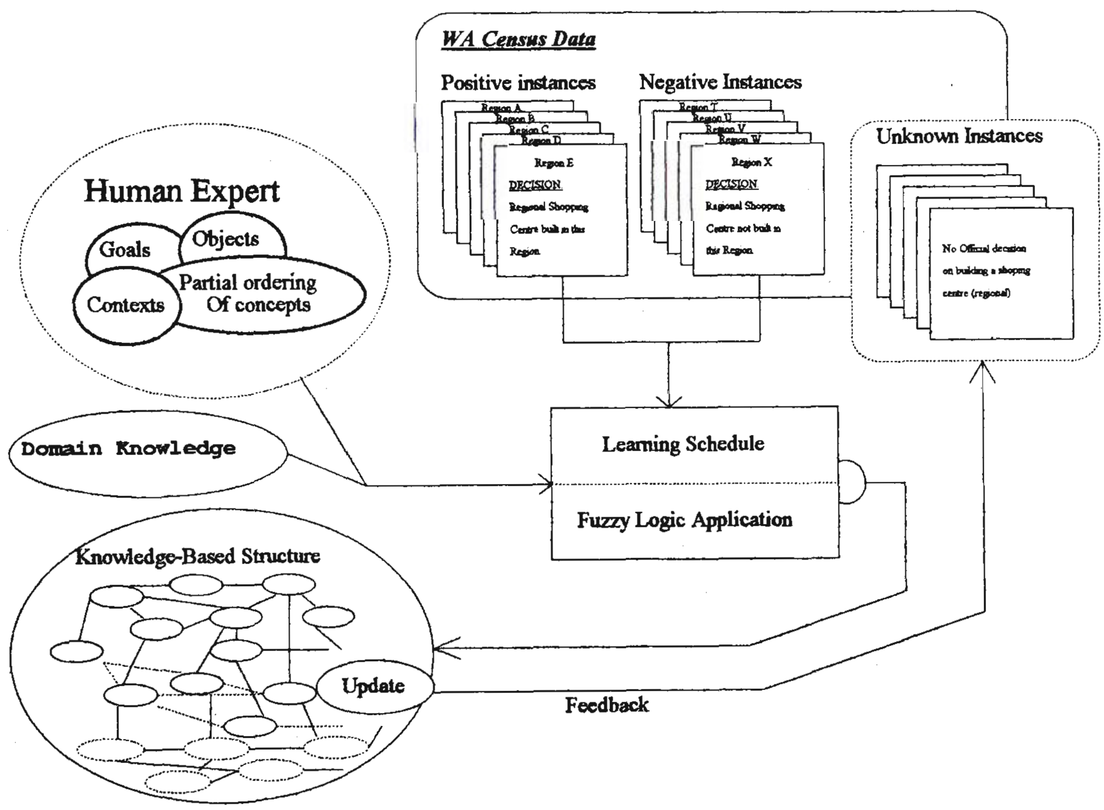


Figure 5: Components and interactions for the overall Knowledge base construction

The data sets (Western Australian Census Data for 1991) are dense with each record consisting of over 2700 fields (see Figure 6). The human expert provided much of the initial background knowledge, defining objects, contexts (objects per context) and the partial ordering to be used to label the causal links between these objects. This set of goals, contexts and concepts are treated as a minimal set of known facts and it is the role of the learning schedule and generalisation techniques to suggest others. The human experts identify a subset of new objects and contexts that seem to be 'useful' and 'interesting'

The data, initially symbolic, after some statistical pre-processing produced real valued bounds per concept. The data to some extent is already generalised about region boundaries (council boundaries) and other explicit generalisations were reflected in the groupings of the questions at the time of the initial survey. An example of regional boundaries is represented in Figure 7(a) and (b).

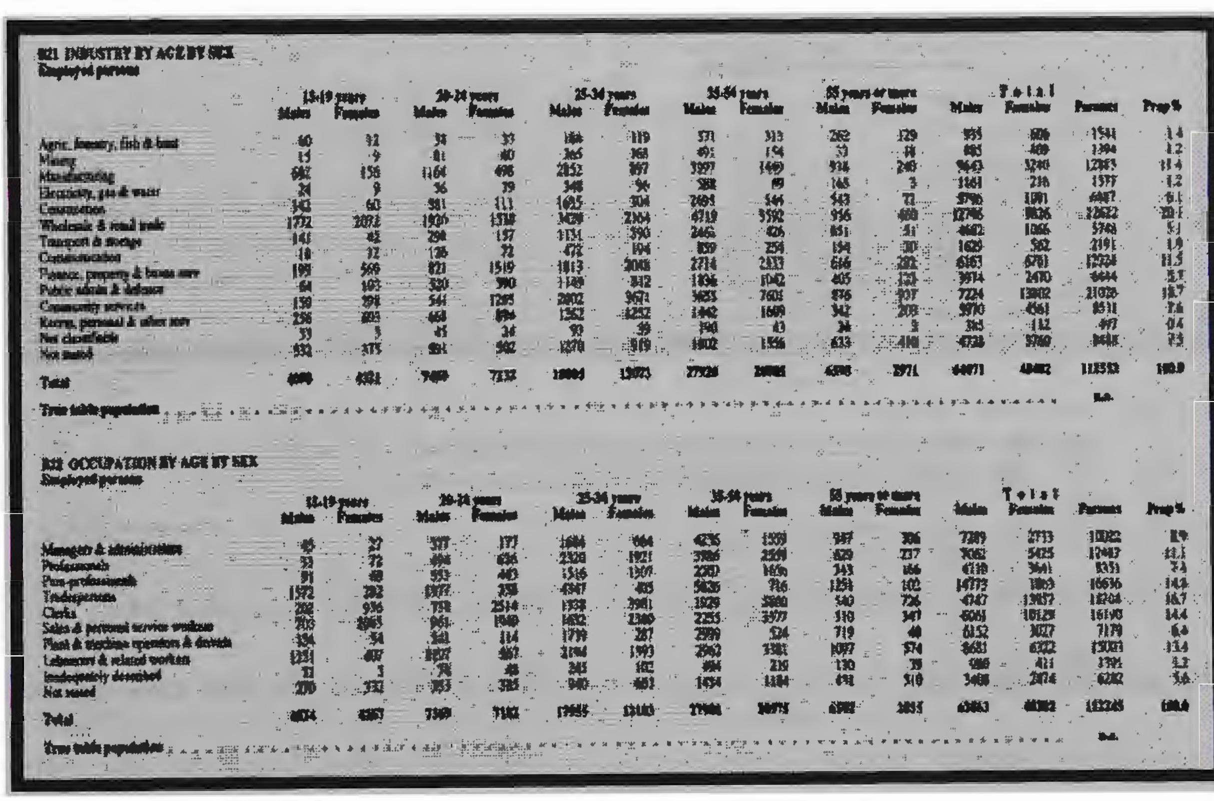
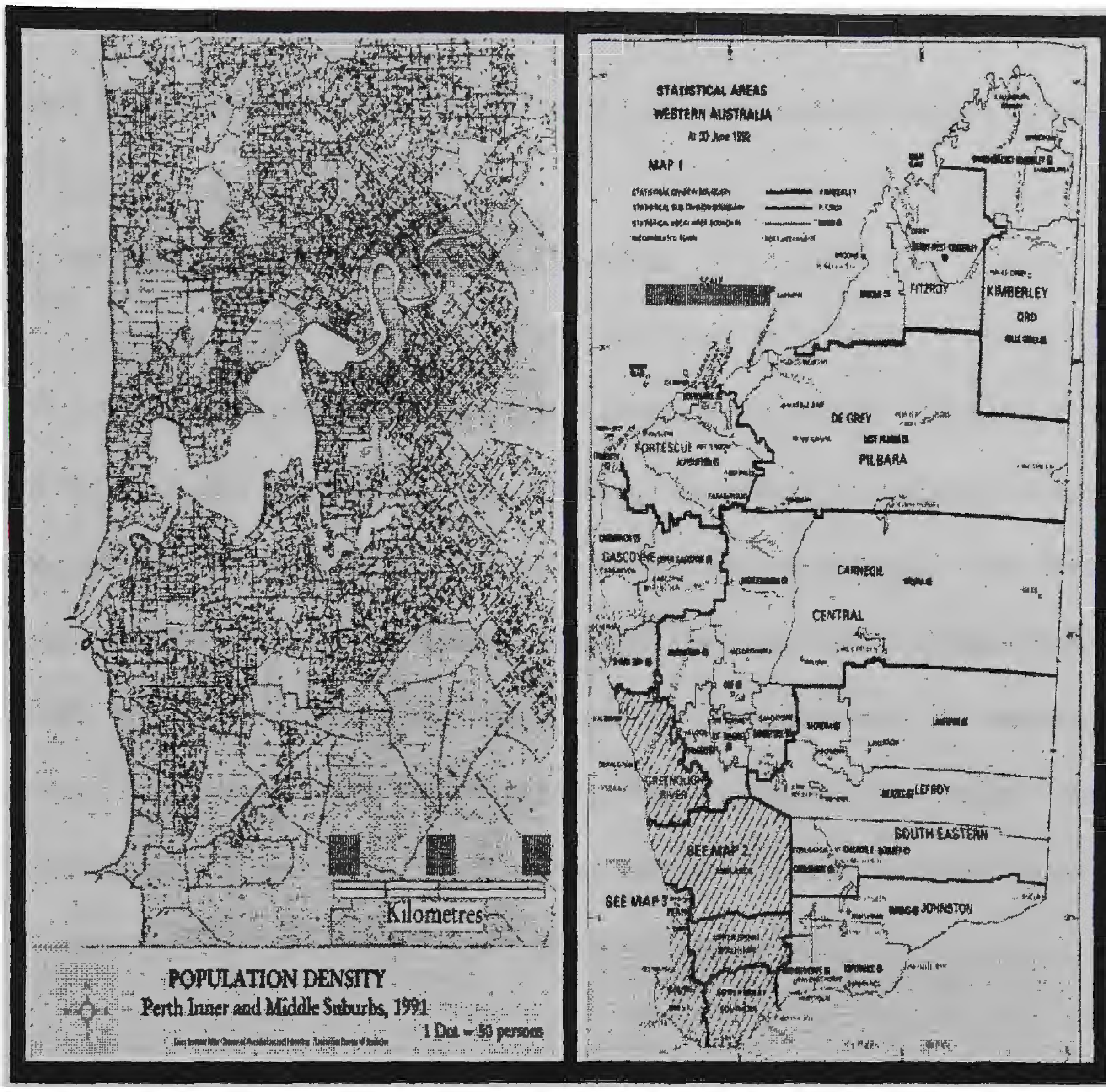


Figure 6: An example of one table of results form the 1991 Western Australian Census Data (after pre-processing).



(a) **(b)**

Figure 7: Examples of Smallworld representation (GIS representation) from dxf files. (a) Population density representation. (b) Regional boundary information.

The data is divided into positive, negative and unknown 'regional data sets'. Positive data sets are those that indicate a decision made on the belief that the data within that particular region was supportive towards a positive planning decision (ie: shopping centres were built in this region). Negative or counter examples, indicate that a planning decision process was undertaken and the outcome was negative. Unknown instances had no indication of positive or negative outcomes from any decision process known. Each example is similarly labelled positive, negative and unknown, per concept.

3.2. Learning Modifiers, Concepts and Causality:

The learning schedule is the result of three learning phases, each phase a pass over the entire data set. The first phase generalises descriptions for individual concepts, using positive and negative instances, per concept, from the entire data set. Version spaces [Mitchell 1978] with disjunction [Manago 1987] and inductive generalisations [Plotkin 1970] [Plotkin 1971] provide compact concept descriptions. For instance, to answer the question "Do we build a regional shopping centre in this region?" and a context goal ="Family type by region" then an example of a generalised description for the concept C("Is an old suburb") is, after the first learning phase:

C("Is an old suburb") =
[Industry((Mining <2%) & (Wholesale and retail trade(20% to 30%)))]
[Occupation by trade(Tradepersons(>12%))]
[Age by sex(Male 40 years of age and older(>40%) & Female 40 years of age and older(>40%))]
[Age by labour force status(Male unemployed(>30%) & Female unemployed(>30%))]
[Annual individual income(Persons(<20K))]
[Number of dependants living at home(None(>70%))]
[Structure of dwelling(Separate house & Not stated)]
[Monthly housing loan repayment by dwelling type(0-$200(>50%))]
[Landlord type by weekly rent(Commission/authority(>20%))]
[Method of travel(Bus(>50%) & Train(>50%) & Walk(>70%))]

Similarly there is a concept description for ~C("Is an old suburb") or C("Is a new suburb").

The second learning phase generates membership functions (e_{ij}) for each pair of concepts (Ci,Cj). It is also necessary to determine the membership functions for each combination of Ci, ~Ci, Cj and ~Cj. Since we do not allow negative causal links between concepts (see FCM framework discussed earlier), the membership values for all possible combinations of a pair of concepts will be used to determine replacement positive correlates to replace the negative causal links.

The fuzzy model used for determining membership functions between each pair of concepts is based on the calculus of linguistically quantified propositions [Zadeh 1983], [Zadeh 1989], where a linguistically quantified proposition and partial ordering of quantifiers, may be generally written as

Definition 9:

'*Q(x's $\in$ Ci) are Cj*' where Q is a linguistic quantifier. For instance:

Likely(x's=Region (Population age under 50 years of age) are (Is an old Suburb))

Definition 10:

Define P as the partial ordering of Q's. For instance

P[Poor, Unlikely, Average, Likely, Always] $\rightarrow$

P[[0-0.2),[0.2-0.4),[0.4-0.5],(0.5,0.8],(0.8-1]]

Alternatively each quantifier could be a function describing a fuzzy set in [0,1] exemplified by 'always' where:

$\mu_{(\text{'always'})}(x) = \{(1, \text{for } x \geq 0.8),(10x\text{-}8, \text{for } 0.8<x<0.9),(0, \text{for } x < 0.8)\}$

Extending the definition for linguistic quantified proposition to find concepts descriptions from examples (inductive learning) with "soft" positive and negative domains. Kacprzyk defines concept descriptions 'R' as [Kacprzyk 1990]:

Definition 11:

> '*Q+ Px's are R*' & '*Q- Nx's are R*' where P and N denotes soft positiveness and soft negitiveness respectively, and Q+ and Q- observe the criteria for completeness and consistency, within soft domains, respectively.

Applying definitions 10 and 11 to the second learning phase, we construct a FCM for the context, "family type by region" for fourteen nominated concepts (see Figure 8). This FCM is represented by a connectivity matrix (see Figure 9) and after some additional pre-processing (pruning) by human experts form our initial FCM (see Figure 10).

Since we do not allow negative causal links, we replace all negative links with their positive equivalents. For instance, for every negative link, (Ci,Cj), Cj is replaced by its modifier ~Cj and the new membership function '***new_eij = 1 - old_eij***'. Since this would affect the rest of the FCM from Cj, for every causal link from Cj to the terminating concept (ie: C14), we replace the subsequent link (Cj,Ck) (where Ci,Cj and Ck are contiguous left-to-right causally linked concepts in the FCM) with the membership function and causal link for (~Cj,Ck). Repeat the procedure for replacing negative links from ~Cj to the terminating concept. For example, for the above connectivity matrix and FCM, there exists a negative link between (C1,C3), hence:

C1 $\rightarrow_{0.2}$ C3 $\rightarrow_{0.6}$ C6 $\rightarrow_{0.6}$ C9 $\rightarrow_{0.8}$ C14 becomes C1 $\rightarrow_{0.8}$ ~C3 $\rightarrow_{0.1}$ C6 $\rightarrow_{0.6}$ C9 $\rightarrow_{0.8}$ C14 after the first replacement, since ($e_{C1\sim C3}=1-e_{C1C3}$) and $e_{\sim C3C6}=0.1$ from phase 2 of the learning schedule. Since (~C3,C6) is now a negative link, it will be replaced using the same procedure and the resulting FCM fragment is C1 $\rightarrow_{0.8}$ ~C3 $\rightarrow_{0.9}$ ~C6 $\rightarrow_{0.6}$ C9 $\rightarrow_{0.8}$ C14 where $e_{\sim C6C9}=0.6$ and so the replacement needs not to proceed any further.

(C1,..,C14) = (C1="Age under 50 years of age",
C2="Young children under 18 years of age",
C3="Is an old suburb",
C4="Car is major form of transport",
C5="Is a new suburb",
C6="Aging population(over 60 years of age)",
C7="Professional employment",
C8="Unemployment rate > 20%",
C9="Has a regional shopping centre",
C10="Hours worked by one member of the family (35-39 hours/week)"
C11="Average family wage earnings greater 20K",
C12="Two parent families",
C13="Other adjacent regional shopping centre",
C14="Good position for a regional shopping centre").

Figure 8: List of all 'useful' and 'interesting' concepts for the context, "family type by region". C9 and C13 have values of 1 or 0, determined from map references and C14 is the result of a decision made by a human expert (historical reference).

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
C1		0.6	0.2		0.7									
C2												0.5		
C3						0.6								
C4													0.6	
C5				0.8			0.8	0.2	0.3					
C6									0.6					
C7										0.6				
C8											0.5			
C9														0.8
C10														0.5
C11														0.9
C12														0.9
C13														0.1
C14														

Figure 9: Connectivity matrix for (C1,...,C14). Each null cell is a concept pair that is not considered and for each pair of concepts (Ci,Cj) there is also a value for (Ci,~Cj), (~Ci,Cj) and (~Ci,~Cj).

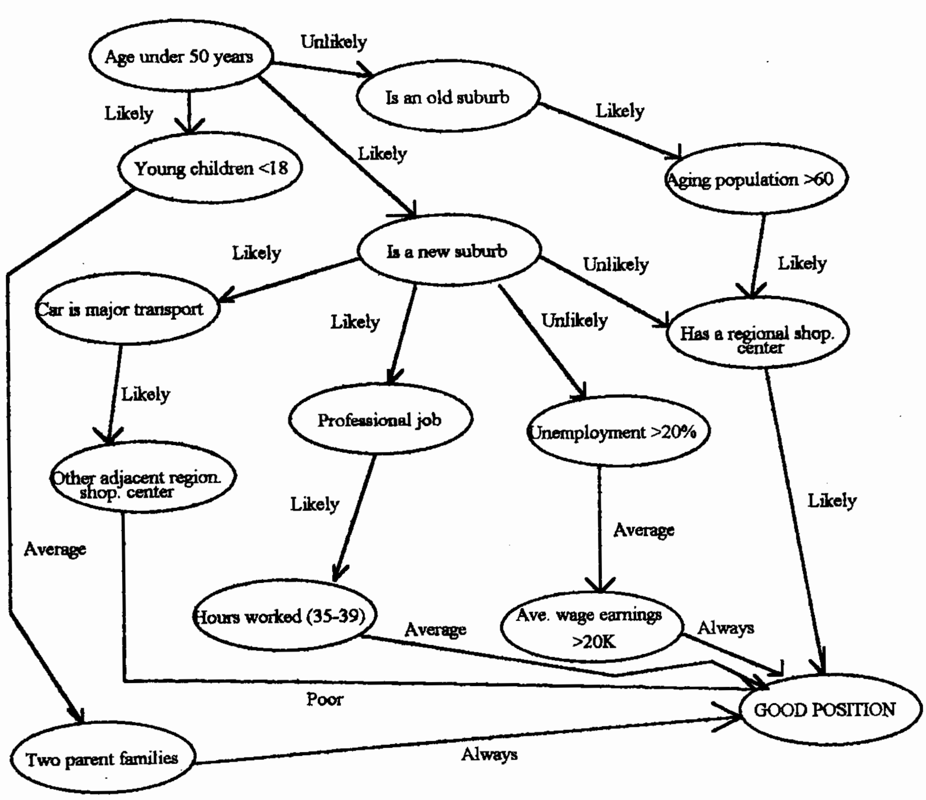


Figure 10: FCM for context, "family type by region".

The final FCM for the context, "family type by region" is shown in Figure 11 below:

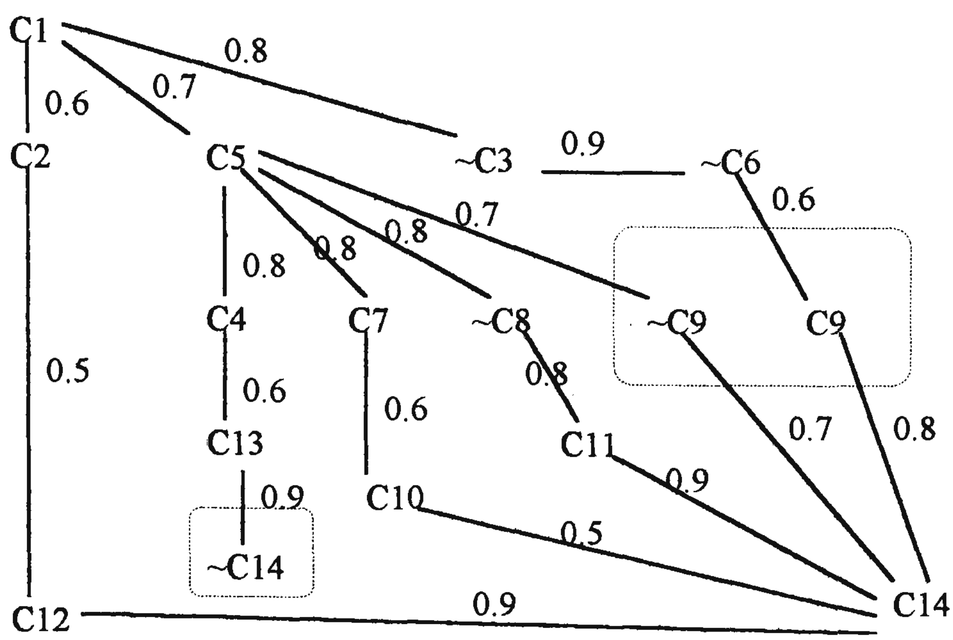


Figure 11: FCM for context, "family type by region", represented by positive causal connections.

Additional splits (see selected boxed sections in Figure 11) have occurred to cater for the inclusion of abstract negation of certain concepts.

This learning schedule is repeated for each context in the context layer of the KBS.

3.3 Feedback (Learning previously unlabeled instances):

The final phase of the learning system labels previously unclassified instances (regions). Since there exist descriptions for each concept and its abstract negation, classification of new instances can be achieved via direct mapping onto the knowledge base and confirmation from the human expert. From this case study, almost 80% of all unknown instances were deemed to be correctly labelled after consulting with the human expert. These correctly labelled instances, updated the counts for each of the concepts and context covered by these examples. The 20% that was classified incorrectly was divided into two categories; instances that could not be mapped onto the knowledge-base indicating an incomplete knowledge-base at the lowest FCM level, and those that were incorrectly classified pointing out subtle levels of context that were not covered by our general context categories (context layer). It is important to also point out that our findings are not necessarily held by the Australian Bureau of Statistics and should not be interpreted in any way other than one representation of the data.

4. Conclusion:

FCMs provides a flexible representation for soft knowledge domains. They are able to represent the fuzzy decision processes that human experts use and provide an adaptive structure based on evidential reasoning. We extended FCMs to include a three layered structure based on inductive learning, contextual indexing and human expert knowledge. This KBS has proved useful for representing the various decision processes that human experts use, and is the first step to a KBS that incorporate relational structures, both linguistically and spatially, contextual references and complex data sources. For the case study we have explored in this paper the initial research has revealed encouraging results.

Context representation in GIS is not a new concept [Gahegan 1988], [Gahegan 1993], [Roberts 1991] and we will investigate techniques for extracting spatial data to represent relational information between objects from GIS data, in an intelligent way using a model of our KBS. This would allow causality to be a function of both linguistically quantified variables and spatial representations [Flack 1993]. We also plan to explore and improve the feedback processes of this architecture, making use of the unknown instances and the human expert knowledge to refine and improve the KBS.

Acknowledgements:

Many thanks to Professor Terry Caelli for providing an atmosphere and environment for productive research and Jan Hurst for the many hours of explanation of GIS data and her expert knowledge. I would also like to thank Smallworld Systems (Australia) P/L and in particular Mike Bundock (Director), for providing me with an Object-Oriented GIS software package that endorsed efficient and productive coding hours. Without their help and software, this work would have taken many more months to realise its usefulness.

References:

1. [Axelrod 1976] Axelrod R. *"Structure of Decision: the Cognitive Maps of Political Elites"*, Princeton, NJ: Princeton University Press (1976).

2. [Bandler 1980] Bandler, W. and L. Kohout *"Fuzzy Power Sets and Fuzzy Implication Operators"*, Fuzzy Sets and Systems, Vol 4, pp 13-30 (1980)

3. [Bonissone 1985] Bonissone, P.P., and H.E. Johnson, Tr. *"Expert System for Diesel Electric Locomotive Repair."* Human Systems Management 4 (1985).

4. [Fike 1985] Fikes R., and T. *Kehler "The Roles of Frame-Based Representation in Reasoning"* Communications of the ACM, Vol 28, No 9, September 1985, pp 904-920 (1985).

5. [Flack 1993] Flack J. C., M. N. Gahegan and G. A. W. West *"The use of Sub-pixel measures to improve the Classification of Remotely-sensed imagery of Agricultural land"* To appear in Proc. Seventh Australasian Remote Sensing Conference, Melbourne, Australia (1993).

6. [Gahegan 1988] Gahegan M. N., and S. A. Roberts *"An Intelligent, Object-Oriented Geographical Information System"* International Journal of GIS, Vol 2, No 2, pp 101-110 (1992).

7. [Gahegan 1993] Gahegan M. N., and S. A. Roberts *"The use of Context and Significance within an Geographic Information System"* Proc. 1st Australia & New Zealand Conference on Intelligent Information Systems (ANZIIS), University of Western Australia, Perth, Australia (1993).

8. [Kacprzyk 1990] Kacprzyk J. and C. Iwanski *"Inductive Learning from Incomplete and Imprecise Examples"*, Lecture Notes in Computer Science, Vol 521, Uncertainty in Knowledge Bases, B. Bouchon-Meunier, R. R. Yager and L. A. Zadeh (Eds), pp 424-430 (1990).

9. [Kosko 1986] Kosko B., *"Fuzzy Cognitive Maps"* Int. J. Man-Machine Studies, No 24, pp 65-75 (1986)

10. [Manago 1987] Manago M. and J. Blythe *"Learning Disjunctive Concepts"*, Lecture Notes in Artificial Intelligence, Vol 347, K. Morik (Ed), Knowledge Representation and Organisation in Machine Learning, Springer-Verlag.

11. [Michalski 1983] Michalski R. S., J. G. Carbonell and T. M. Mitchell (eds) *"A Theory and Methodology of Inductive Learning"*, Machine Learning: An Artificial Intelligence Approach, Vol 1, Chapter 4, pp 83-134, Morgan Kaufmann, Los Altos, CA (1983)

12. [Mitchell 1978] Mitchell T. M., *"Version Spaces: An Approach to Concept Learning"* Doctorial dissertation, Computer Science Department, Stanford University (1978).

13. [Parsaye 1988] Parsaye, K., and M. Chignell, *"Expert Systems for Experts"* New York: John Wiley and Sons (1988).

14. [Plotkin 1970] Plotkin G. D., *"A Note on Inductive Generalization"* Machine Intelligence 5, pp 153-163 (1970).

15. [Plotkin 1971] Plotkin G. D., *"A Further Note on Inductive Generalization"* Machine Intelligence 6, pp 101-124 (1971).

16. [Quinlan 1992] Quinlan J. R., *"C4.5: Programs for Machine Learning"* Morgan Kaufmann, San Mateo, CA (1992).

17. [Roberts 1991] Roberts S. A., and M. N. Gahegan *"Supporting the Notion of Context within a Database Environment for Intelligent Reporting and Query Optimisation"* Eur. J. Inf. Systems. Vol 1, No 1, pp 13-22 (1991).

18. [Turban 1988] Turban, E., and P. Watkins, eds. *"Applied Expert Systems"* Amsterdam: North Holland (1988).

19. [Woods 1991] Woods, W. A., *"Important Issues in Knowledge Representation"* Knowledge-Based Systems: Fundamentals and Tools, O. N. Garcia and Y. Chien, IEEE Computer Society Press (1991).

20. [Yager 1981] Yager R. R. *"Textured Sets. An Approach to Aggregation Problems with Multiple Concerns"*, IEEE Transactions on Systems, Man and Cybernetics, Vol 11, pp 730-731 (1981).

21. [Yager 1993] Yager R. R. and P. F. Dimitar *"SLIDE: A Simple Adaptive Defuzzification Method"*, IEEE Transactions on Fuzzy Systems, Vol 1, No 1, February (1993).

22. [Zadeh 1983] Zadeh L. A. *"A Computational Approach to Fuzzy Quantifiers in Natural Language"*, Computers and Mathematics with Applications 9, pp 149-184 (1983).

23. [Zadeh 1989] Zadeh L. A. *"Knowledge Representation in Fuzzy Logic"*, IEEE Transactions on Knowledge and Data Engineering, Vol 1, No 1 (1989).