# Intelligent Metadata Extraction For Integrated Coastal Zone Management

**Antoni Moore; Allan Jones*; Peter Sims *; Grahame Blackwell****

*Spatial Information Research Centre*
*Department of Information Science*
*University of Otago*
*Dunedin*
*New Zealand*
amoore@infoscience.otago.ac.nz

**Department of Geographical Sciences / **School of Computing*
*University of Plymouth*
*Drake Circus*
*Plymouth*
*PL4 8AA*
*U.K.*

**Abstract.** Integrated Coastal Zone Management (ICZM) is an application area of increasing importance, requiring much data, information and knowledge. The scale of this resource is acknowledged at an international level, and is known to be ever growing. Among other factors, the global repository of coastal data and information is known to be fragmented, is of different formats and occurs at various scales. Metadata, where it exists, is widely recognised as a valuable method of navigation for this resource. This paper demonstrates the successful use of an expert system to provide intelligent access to a set of coastal metadata (on the Fal Estuary, southwest Cornwall, UK), whilst taking a holistic approach. This approach is central to ICZM and is necessary when having to deal with the attendant multiple disciplines, scales and technologies, and when providing an integrated solution to managing coastal data, information and knowledge. Practically, the expert system employs the Dempster-Shafer theory of evidence and the associated concepts of *ignorance* and integration to tackle the uncertainty and complexity inherent in the holistic paradigm.

## 1 INTRODUCTION

Within Coastal Zone Management there are widespread calls for an integrated approach. To enable participatory decision making in coastal zone management, Deakin (1994) called for a 'rigorous framework' to oversee interpretation of data, dependent on quality and suitability for use. This is a holistic view, which Doody (1996) and the EC (1999) advocate as a solution to integrate fragmented and undocumented data. Kucera (1995) supports this but acknowledges that to adopt a holistic view of coastal environmental management would be a challenge, due to the diversity of the coast and feature interactions. Also, data relating to the coastal zone can confuse through its sheer quantity and variety, requiring a method of 'navigation', and an overall framework (Riddell, 1992).

It has been noted that land or water resources management are different from coastal resources management, in that a unified approach is essential for the latter (Clark, 1998). The alternative is a piecemeal approach, which has historically led to unsatisfactory results (i.e. one problem alone is addressed, causing other problems in turn). For example, coastal pollutants are seen to be rife where policies are narrow in approach (Kucera, 1995).

Integrated Coastal Zone Management (ICZM) is sustainable management of the coast, integrating the concerns of all stakeholders (in various activities and sectors) in relation to all goals (from local to international scale) (Clark, 1998; Scholten *et al*., 1999). ICZM employs a holistic approach, and really came to the fore through Agenda 21 of the Rio de Janeiro Earth Summit (Cecin-Sain *et al*., 1995, Drummond, 1998). Agenda 21 was an influence on the European Union (EU) Demonstration Programme on Integrated Coastal Zone Management (see EC, 1999). The Demonstration Programme was completed in 2000, and was followed by 'A strategy for Europe', a number of recommendations for ICZM, which has been taken up by a number of European nation states (EC, 2000).

This paper describes a use of the Coastal Management Expert System (COAMES), addressing the holistic paradigm. It is proposed here that metadata can be used to enable holism and therefore portability, as a way of navigating the huge amount of fragmented coastal data, information and knowledge that exists. To prove the power of metadata, COAMES will be tested on a coastal metadataset describing estuarine data (derived from the Atlantic Living Coastlines project, one of the EU Demonstration Projects on ICZM – Bayliss and Moore, 2000) encompassing both natural environmental themes, and at a variety of spatial and temporal scales.

The use of an expert system for ICZM originally came about as a response to the scarcity of such systems with a coastal application (but see Scheerer, 1993; McGlade, 1997; Houhoulis and Michener, 2000). This is despite expert systems and other types of coastal zone management information system (CZMIS) being seen as an integrating mechanism for coastal data and information, rectifying disparate formats, qualities, sources and disciplines (Ripple and Ulshoefer, 1987; Miller, 1994). Thus a strong potential niche in ocean and coastal science has emerged (Moore *et al*., 2001).

This paper will initially outline expert systems, with emphasis given to the inferencing mechanism, the Dempster-Shafer (D-S) Theory of Evidence (in Part 2). Part 3 will detail the expert system process, in addition giving an overview of the metadata to be used.

Five tests were set to prove the efficacy of the expert system, in terms of value for the coastal zone manager and capturing the essence of holism. The first investigates straightforward belief in a rule - if evidence supports the rule then the belief is updated. With the second test, conflicting belief in another, independent, rule is introduced, allowing a critique of linkages within the expert system. Furthermore, the new evidence is weighted with a large ignorance value (i.e. there is not much confidence in this rule), inviting pathways to holism (ignorance measures what is *not* known, an appreciable part of the whole picture). The third test adds further support for the weaker rule. The next test again assesses the effectiveness of links within the expert system, but in a hierarchical fashion. It assesses the effectiveness of a superset (consisting of two or more amalgamated rules to create another rule with its own belief) on its subsets, and vice versa. Finally, the explicit use of belief against a rule will be assessed, giving an idea of how links can be suppressed.

Finally, Parts 4 and 5 will present and discuss the results.

# 2 EXPERT SYSTEMS

## 2.1 Basics

Expert systems are "computer systems that advise on or help solve real-world problems requiring an expert's interpretation and solve real-world problems using a computer model of expert human reasoning reaching the same conclusion the human expert would reach if faced with a comparable problem" (Weiss and Kulikowski, 1984). The core of an expert system commonly consists of a domain-independent inference engine and a domain-specific knowledge base (Robinson, Frank and Blaze, 1986). The knowledge base comprises expert rules (which model behaviour of, and functions relating to, a theme) and facts (single values such as basic information and events).

COAMES (Coastal Management Expert System) is an object-oriented expert system (Figure 1), consisting of a user interface, a database, an object-oriented knowledge base (incorporating both the expert's factual knowledge and the process knowledge embodied in models) and most importantly an inference engine (Moore *et al.*, 1996). Models are one of four groups of functions, the others being data functions, rule functions and toolbox functions. Within the inference engine are algorithms to calculate belief through the Dempster-Shafer Theory of Evidence.

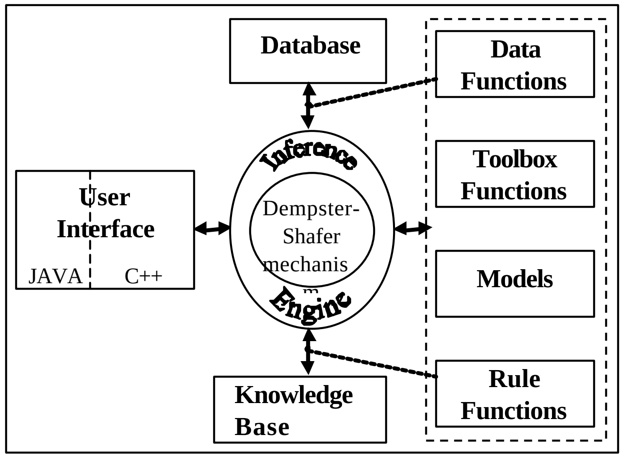


**Figure 1:** The configuration of COAMES containing the major components: user interface, database, functions, knowledge base and inference engine.

## 2.2 The Hierarchical Knowledge Structure

COAMES is underlain by an object-oriented knowledge structure, Figure 2 shows the form of the class structure of the prototype described here. For example, the geomorphology class is defined by the attributes and functions encapsulated as class members. This class inherits all the elements of the geography superclass. The broken line reveals another class from which inheritance is derived (through multiple inheritance), the physical class. This reflects the thematic overlap of geography and physics that geomorphology represents. Each instance of a given class is termed an object.

## 2.3 Dealing With Uncertainty

Expert systems, by their very nature, deal with a lot of uncertain data, information and knowledge. However, the treatment of uncertainty in expert systems has mainly been neglected, as some have found no significant difference between using uncertainty and an assumed certainty (Turban, 1995). Tversky and Kahneman (1974) quote empirical evidence that people are poor estimators of probability, and probabilities derived from published data were also found to be wide of the mark (Ben-Bassat *et al*., 1980). Even so, such deviation results in small changes for the value derived from Bayes theorem, for instance.

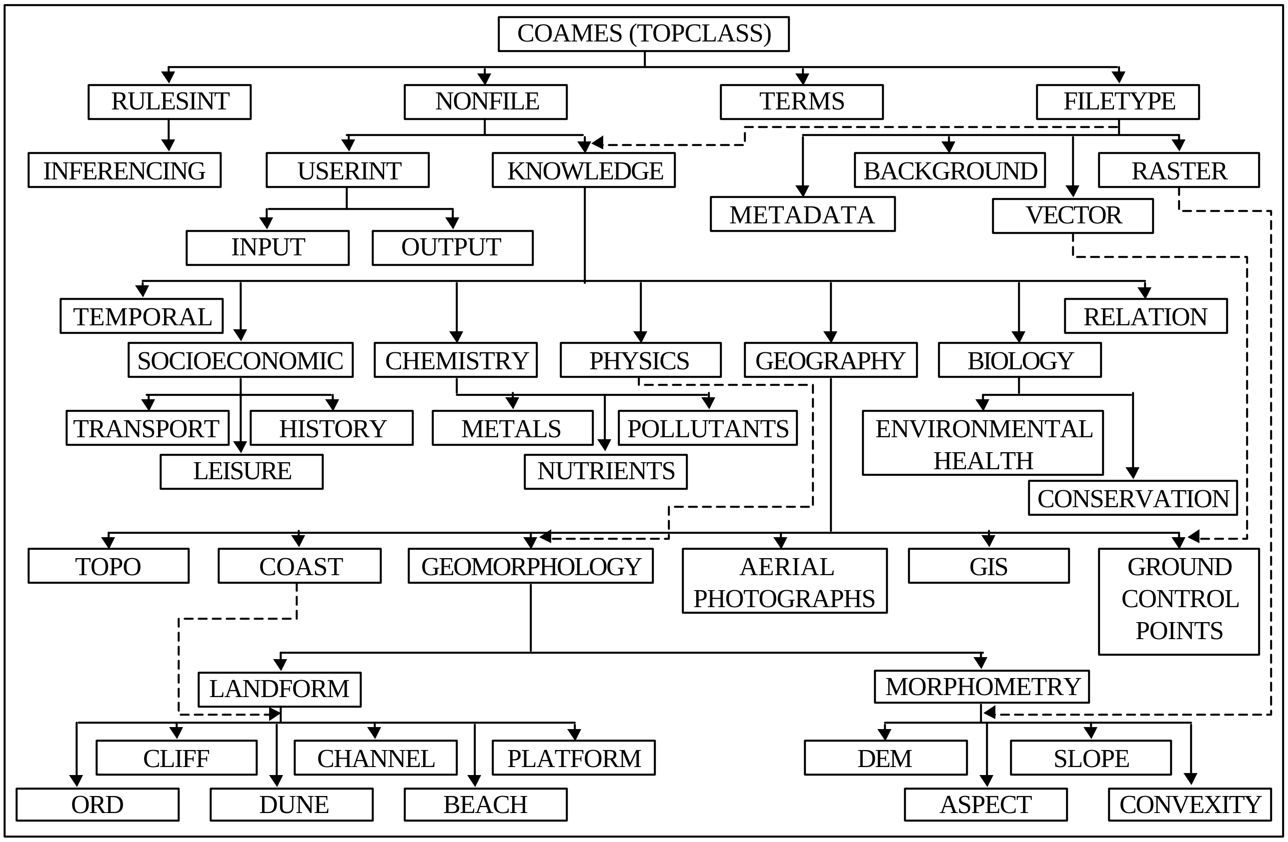


**Figure 2:** The object-oriented hierarchical structure of knowledge and data in the prototype (multiple inheritance links are dashed).

Despite this, several methods have been used to combine or integrate uncertain information for inference - these include Bayes Theorem, the Dempster-Shafer theory of evidence, certainty factors and fuzzy sets. The inferencing mechanism used by COAMES, Dempster-Shafer Theory, will be discussed here.

### 2.3.1 Dempster-Shafer Theory of Evidence

The Dempster-Shafer (D-S) theory of evidence (Dempster, 1967; Shafer, 1976) is an extension of Bayes theorem. It waives the need for exhaustive prior or conditional probabilities before calculation can take place and thus can be used where evidence is lacking or where evidence is based on vague perceptions. Most importantly, it introduces the representation of ignorance.

Normally, where probabilities are not known, maximum entropy means that equal prior probabilities are unrealistically assigned to each competing piece of evidence, and the sum of all assigned probabilities must equal one. With Dempster-Shafer theory, an ignorance value close to zero (ignorance = 0 represents complete ignorance) can be used to represent the lack of information, rectifying what would be erroneous with probability (Gordon and Shortliffe, 1984; Scheerer, 1993; Turban, 1995). Related to this is the fact that when belief is assigned to a particular hypothesis ($P(H)$), the remaining belief does not necessarily then support that the hypothesis' negation (i.e. $P(-H)$ does not necessarily equal $(1-P(H))$.

Other advantages of using D-S include the ability to use evidence supporting more than one hypothesis (a subset of the total number of hypotheses). Finally, D-S models the narrowing of the hypothesis set with the accumulation of evidence, which is exactly how experts reason (Gordon and Shortliffe, 1984).

However, singletons (single hypotheses) are assumed to be mutually exclusive and exhaustive. Likewise, evidence needs to be independent (Gordon and Shortliffe, 1984; Scheerer, 1993). Ling and Rudd (1989) offer methods to combine the dependent opinions of experts, which is a realistic option. Another drawback is the computational complexity of D-S (Gordon and Shortliffe, 1984). Barnett (1981) offers a routine that reduces the computations to linear time.

Gordon and Shortliffe (1985) developed an approximation method to D-S theory, this time tackling hierarchical evidence. Soon after, Shafer and Logan (1987) adapted D-S theory for hierarchies, producing a more robust method. This mirrored Pearl's (1986) efforts with Bayes' theorem and also reduced computational complexity.

Other examples of Dempster-Shafer use include Ferrier and Wadge (1997 - geological analysis of sedimentary basins), Srinivasan and Richards (1990 - remote sensing classification) and Kontoes *et al.* (1993 - classifying remotely sensed images for agriculture).

### 2.3.2 Dempster-Shafer Theory in Practice

The working area in Dempster-Shafer theory is called the Frame of Discernment (FoD) or $\Theta$. It is equivalent to the sample space in probability terms. The FoD contains a set of possible (mutually exclusive) answers or hypotheses aiming to resolve a question. For example (adapted from Gordon and Shortliffe, 1984), the question could be "What metals are polluting this estuary?". In response, the FoD may be:

$$\Theta = \{iron, nickel, lead, mercury\}.$$

Furthermore, any subset of this FoD is a hypothesis.

e.g. hypothesis of heavy metals $= \{lead, mercury\}$;
hypothesis of ferromagnetic metals $= \{iron, nickel\}$

The set of all hypotheses represented by a FoD is $2^{\Theta}$. If a FoD has four elements, as above, then the number of hypotheses $= 2^4 = 16$. The hypotheses are represented in Figure 3.

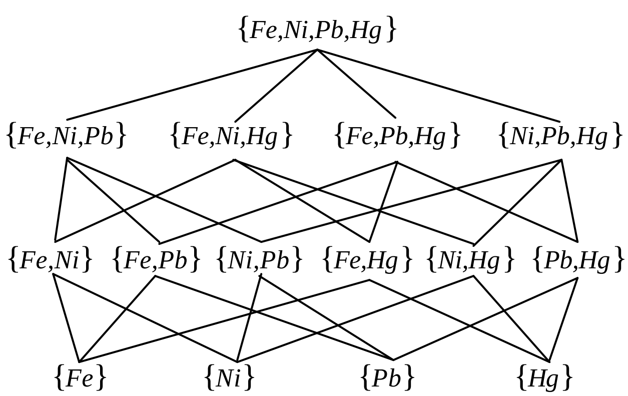


**Figure 3:** The hypotheses used for a four-member frame of discernment, consisting of the metals iron, nickel, lead and mercury.

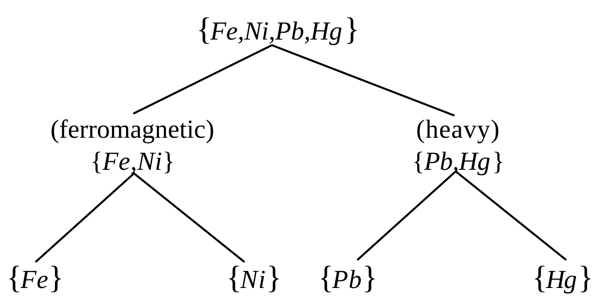


**Figure 4:** A representation of the important subsets in the four-member frame of discernment introduced in Figure 3.

The sixteenth subset is represented by the empty set $\emptyset$, which corresponds to a hypothesis that is known to be false. Out of these subsets, only some may be important, as shown in Figure 4, where metals are classified into ferromagnetic and heavy.

A basic probability assignment (BPA) is given to each subset $A$ where the BPA is represented by $m$ and $0 \leq m \leq 1$. All $m$ values in the FoD will add up to one. In the example above, the following values have been arbitrarily allocated:

$$m(\{Fe, Ni\}) = 0.2$$
$$m(\{Pb, Hg\}) = 0.4$$
$$m(\{Fe\}) = 0.2$$
$$m(\{Ni\}) = 0.1$$
$$m(A) = 0 \text{ for all other subsets}$$

Any probability not assignable to any subset is grouped under $m(\Theta)$ - this is the ignorance and is not used to negate any of the hypotheses, as in probability theory. In this case, $m(\Theta) = 0.1$.

The belief function (denoted $Bel$) corresponds to a specific BPA, $m$, and is the sum of the BPAs (or beliefs) committed to a subset $A$ (i.e. including any subsets beneath $A$ in the hierarchy).

e.g.

$$Bel(\{Fe, Ni\}) = m(\{Fe, Ni\}) + m(\{Fe\}) + m(\{Ni\})$$

$= 0.2 + 0.2 + 0.1 = 0.5$

Any observation against a hypothesis is viewed as evidence supporting its negation.

e.g. evidence against hypothesis $\{Fe\} \equiv$
evidence for hypotheses $\{Ni, Pb, Hg\}$
(i.e. $Ni$ OR $Pb$ OR $Hg$ )
Also, $Bel(-\{Fe, Ni\}) = Bel(\{Pb, Hg\})$.

*Combination of Belief Functions using Dempster's Rule*

This is analogous to the calculation of posterior probability from prior and conditional probability using Bayes' theorem. Consider two pieces of independent evidence represented by BPAs $m_1$ and $m_2$. Their combined probability assignment is $m_1 \oplus m_2$, which is the orthogonal sum of $m_1$ and $m_2$, as illustrated in the following calculations. It is based on the assumption that $\sum m_1(X)m_2(Y)$ always sums to 1 (where $X$ and $Y$ run over all subsets of FoD $\Theta$).

For a subset $A$:

$$m_1 \oplus m_2(A) = \frac{\sum_{X \cap Y = A} m_1(X) m_2(Y)}{1 - \sum_{X \cap Y = \varnothing} m_1(X) m_2(Y)}$$

The best way to effect this calculation is to draw an intersection table. Table 1 represents the evidence introduced as an example in the last section ($m_1$), and $m_2$ is as follows.

$$m_2(\{Fe\}) = 0.8$$
$$m_2(\Theta) = 0.2$$
$m_2(A) = 0$ for all other subsets.

**Table 1:** An intersection table, combining two groups of evidence.

| | **m₁** | | | | |
|---|---|---|---|---|---|
| **m₂** | $\{Fe\}$ (0.2) | $\{Ni\}$ (0.1) | $\{Fe, Ni\}$ (0.2) | $\{Pb, Hg\}$ (0.4) | $\Theta$ (0.1) |
| $\{Fe\}$ (0.8) | $\{Fe\}$ (0.16) | $\varnothing$ (0.08) | $\{Fe\}$ (0.16) | $\varnothing$ (0.32) | $\{Fe\}$ (0.08) |
| $\Theta$ (0.2) | $\{Fe\}$ (0.04) | $\{Ni\}$ (0.02) | $\{Fe, Ni\}$ (0.04) | $\{Pb, Hg\}$ (0.08) | $\Theta$ (0.02) |

$\kappa$ is a normalizing variable and is calculated as the sum of all values assigned to $\varnothing$.

$$\kappa = 0.08 + 0.32 = 0.4$$

Then the sum of evidence for each subset is divided by $(1-\kappa)$.

$$m_1 \oplus m_2(\{Fe\}) = (0.16 + 0.04 + 0.16 + 0.08)/(1 - 0.4) = 0.733$$
$$m_1 \oplus m_2(\{Ni\}) = (0.02)/(0.6) = 0.0333$$
$$m_1 \oplus m_2(\{Fe, Ni\}) = (0.04)/(0.6) = 0.0666$$
$$m_1 \oplus m_2(\{Pb, Hg\}) = (0.08)/(0.6) = 0.1333$$
$$m_1 \oplus m_2(\Theta) = (0.02)/(0.6) = 0.0333$$

Example of a belief calculation:

$$Bel_1 \oplus Bel_2(\{Fe, Ni\}) = m_1 \oplus m_2(\{Fe, Ni\}) + m_1 \oplus m_2(\{Fe\}) + m_1 \oplus m_2(\{Ni\})$$
$$= 0.733 + 0.0333 + 0.0666 = 0.833$$

*Belief Intervals*

For a subset $A$, the belief interval is of the form $[Bel(A), Plaus(A)]$.

$Plaus(A)$ is the plausibility of $A$, where $Plaus(A) = 1 - Bel(-A)$. It is the maximum amount of belief that *could* be committed to $A$. $Bel(A)$ is the amount of belief *currently* committed to $A$. The width of this interval is a measure of the belief that neither supports nor refutes $A$ - it is the amount of uncertainty associated with a hypothesis, given some evidence (Gordon and Shortliffe, 1984; Scheerer, 1993). The belief and plausibility have been called the lower and upper probability respectively by Dempster (1967).

## 3 PROCESS

### 3.1 The Study Area

The study area is the Fal Estuary, situated on the south coast of Cornwall (Figure 5). It is a prime candidate for Integrated Coastal Zone Management, as it is an area where many different coastal interests meet. To elaborate, the estuary is a drowned river valley (or ria), containing two major centres of commerce and population: Falmouth and Truro. Both settlements have ports able to accommodate freight (Truro less so) and recreational boating. Falmouth in particular is one of the largest ports in Devon and Cornwall, as well as being a major centre for tourism. A large proportion of tourism revenue for Falmouth derives from recreational boating. Outside the major settlements, the Fal and Helford Estuaries form a candidate Special Area of Conservation (cSAC), due to the presence of habitats such as subtidal sandbanks, intertidal sand / mudflats, saltmarsh and shallow inlets / bays (Bayliss and Moore, 2000).

One of the reasons for choosing the Fal Estuary as a case study area is an abundance of multidisciplinary data of different temporal and spatial magnitudes. Other reasons include applicability to the coastal zone manager and portability (an earlier case study for COAMES was set in a different, more localized coastal environment – this is discussed in Moore *et al*. 2001). Metadata itself is also prevalent in CZM information management across the globe, so the use of it in this case study is relevant.

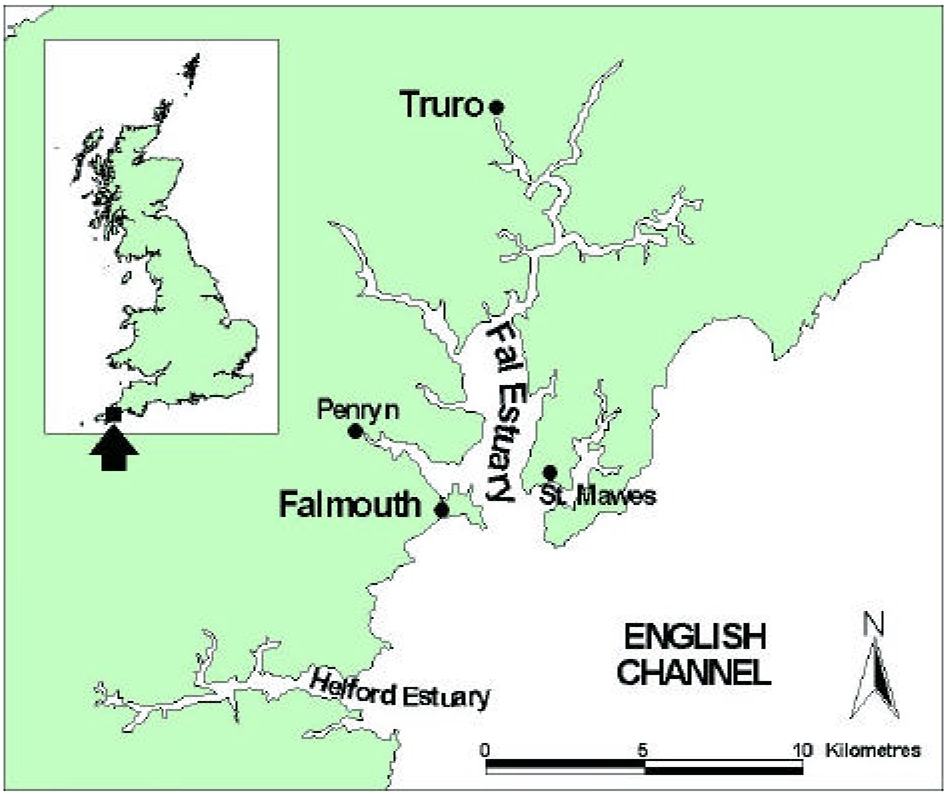


**Figure 5:** Location of the Fal Estuary, UK.

### 3.2 Metadata

#### 3.2.1 Definitions

Metadata is 'designed for description of the contents of a data set' (Goodchild, 1998) or put simply, it is 'data about data'. As such, it gives a general overview of the dataset it describes, and is therefore ideally placed to be the mode of linkage (in conjunction with stored knowledge) to provide a holistic capability to COAMES. Indeed, it is a way of bringing data together without physically integrating them. Metadata is universally accepted as being essential in data management, to make data useful through the description (Busby, 2000) and to aid data discovery (Payne, 2000). An indication of this importance is the repackaging of ESRI ARC/INFO in v.8 (1999), which includes metadata as a key component in its data management program, ArcCatalog (ESRI, 2001).

#### 3.2.2 Metadata Collection

The process of collecting the Atlantic Living Coastlines (ALC) metadata employed a short questionnaire which was sent to a large number of coastal zone managers in Devon and Cornwall (working in local authorities, fisheries, leisure, local wildlife trusts, national environmental institutions, as well as academics, estuary managers and harbour masters). The questionnaire responses were augmented by a series of meetings with the data holders, where data and information bases were described in further detail to fill in various categories of metadata (see next section).

The metadatabase was supplemented by the information database in the Rame Head to Lizard Point Shoreline Management Plan (Halcrow, 1999). The subject of the information described therein was on a broad range of coast-related topics, particularly if it was relevant to coastal defence planning. The fields in the database are similar to those in the metadata standard described in the next section: Title / Subject of data, Date, Area (in terms of sediment subcells - e.g. Fal Estuary = 6D-4), Format, Topic, Content, Author, Source, Availability.

#### 3.2.3 Metadata Standards

The standard used by ALC for their metadata was that of the Federal Geographic Data Committee (FGDC, 2001). The seven broad categories of the FGDC standard are Identification Information, Quality Information, Format Information, Projection Information, Attribute Information, Distribution Information and Metadata Reference. It is perceived to be normal for organisations to continue to establish their own standards, while aligning themselves in principal to the broad categories shown in Table 2 (adapted for ALC metadata). All metadatasets are stored in ASCII files.

**Table 2:** The metadata categories, fields and details, based on the standards of FGDC (2001).

| Category | | Field 1 | Field 2 | Field 3 | Field 4 | Field 5 |
|---|---|---|---|---|---|---|
| **Identification Information** | **Field** | Title | Geographical Coverage | Time Period | Level of Access | |
| | **Details** | Title of the dataset being described | The dataset's geographical extent (placename) | The period of time for which the data was collected | Level of access to the data for outside enquirers | |
| **Quality Information** | **Field** | Temporal Quality | Spatial Quality | Attribute Quality | | |
| | **Details** | Frequency with which data is collected | Spatial accuracy of the data | Accuracy of the attribute being measured (function of the measuring instrument) | | |
| **Format Information** | **Field** | Data Format | Functions | | | |
| | **Details** | Analogue or digital. If digital – GIS (raster / vector), spreadsheet, database etc. | If there is data behind the metadata, this field lists the functions that can be applied to the data. | | | |
| **Geographical Projection** | **Field** | Projection | | | | |
| | **Details** | About the geographical projection | | | | |
| **Attribute Information** | **Field** | Category | Attribute Name | Attribute Details | | |
| | **Details** | The discipline group to which the attribute belongs | The attribute name | Any details about the attribute (e.g. units of measurement) | | |
| **Distribution Information** | **Field** | Owner of Dataset | Charges | Supply Format | Restrictions | Originator of Dataset |
| | **Details** | Current person with the dataset | Any charges to be paid when acquiring the dataset | Range of supply formats | Any restrictions on the dataset's use | The person who originally created the dataset |
| **Metadata Reference** | **Field** | Logger | Last Update | | | |
| | **Details** | Person who collated and logged the metadata | Date on which the metadata was last updated | | | |

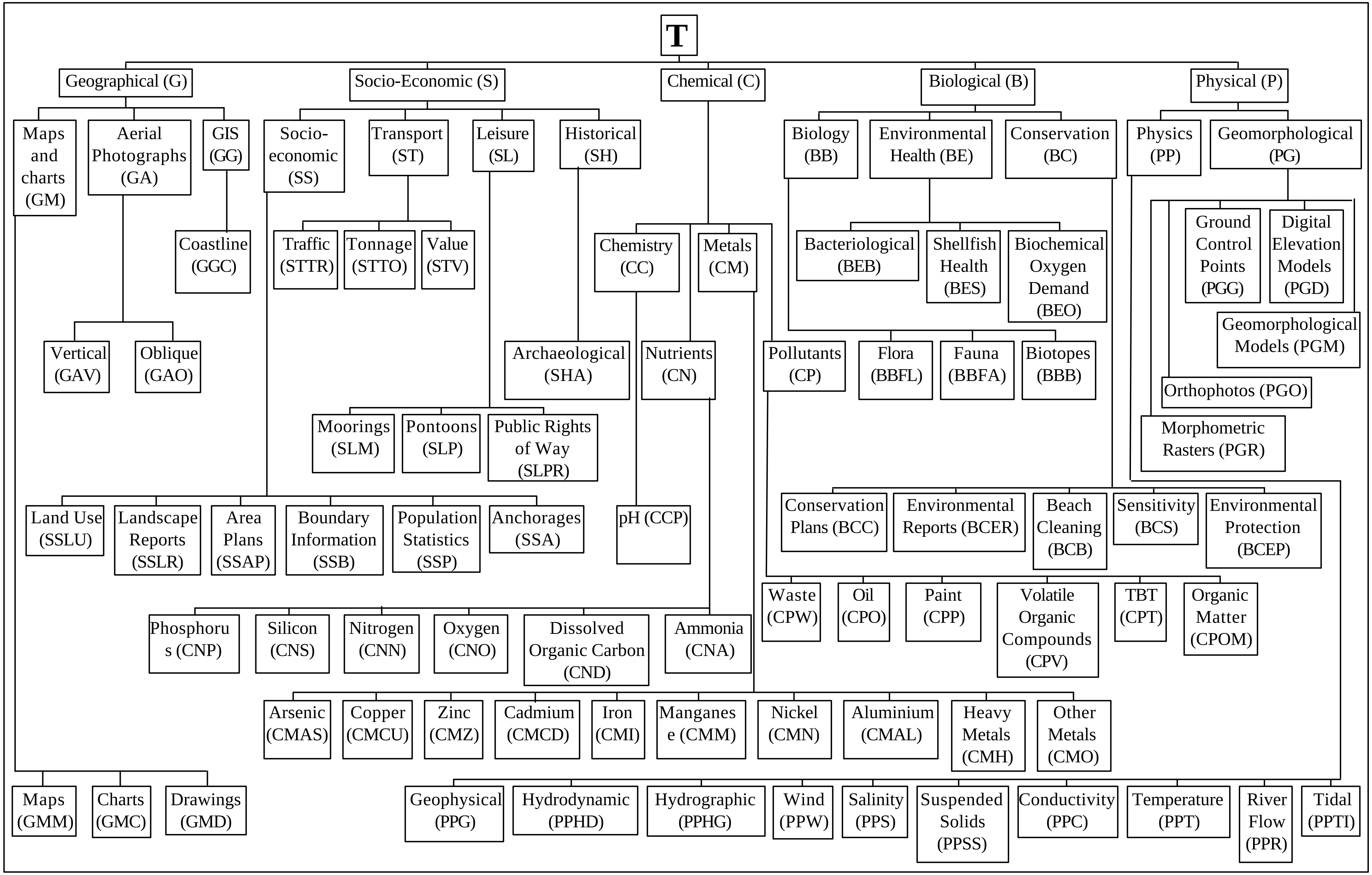


**Figure 6:** The full breakdown of metadata categories. Thematically, parts of this hierarchy are an extension of the object-oriented hierarchy shown in Figure 2

### 3.2.4 Metadata Groupings

Once collected, the metadata was grouped by theme or discipline into five major thematic groups that feed into coastal zone management - metadata describing chemical data, biological data, physical data, socio-economic data, and data with an explicitly geographical content such as coastlines or topographic maps. These five thematic categories were subdivided into more specific subcategories. The identity of these subcategories was limited to the scope of the collected metadata. For instance, the chemical thematic category was subdivided into nutrients, metals, pollutants and chemistry (i.e. pH, alkalinity). Within these subcategories, there are further divisions. For instance, the metals group is divided into specific metals such as copper and zinc. Again these are limited to the scope of the collected metadata. The full breakdown of the metadata categorisation is shown in Figure 6.

### 3.3 Use of the Expert System

The constituents of the expert system will be discussed in turn:

- User Interface
- Knowledge Base Structure
- Inference Engine

Data has already been discussed in the metadata section – for the aims of this case study, metadata is the sum total of data used by the expert system.

### 3.3.1 Java Graphical User Interface

The COAMES user interface in its initial state is shown in Figure 7. It has been constructed in the Java programming language (more specifically using the Java Swing classes) and accesses the C++ expert system code as a series of native functions.

The form of the COAMES interface can be divided into two parts: *dialogue* (on the left), through which the user responds to expert system output in the form of written queries and lists, and *display* (on the right), where the diagnostic information, metadata and output data are displayed (this paper is only concerned with metadata output). The interface tries to capture what is important in the expert system process: adequate explanation to the user and decision support output presented in a useful form.

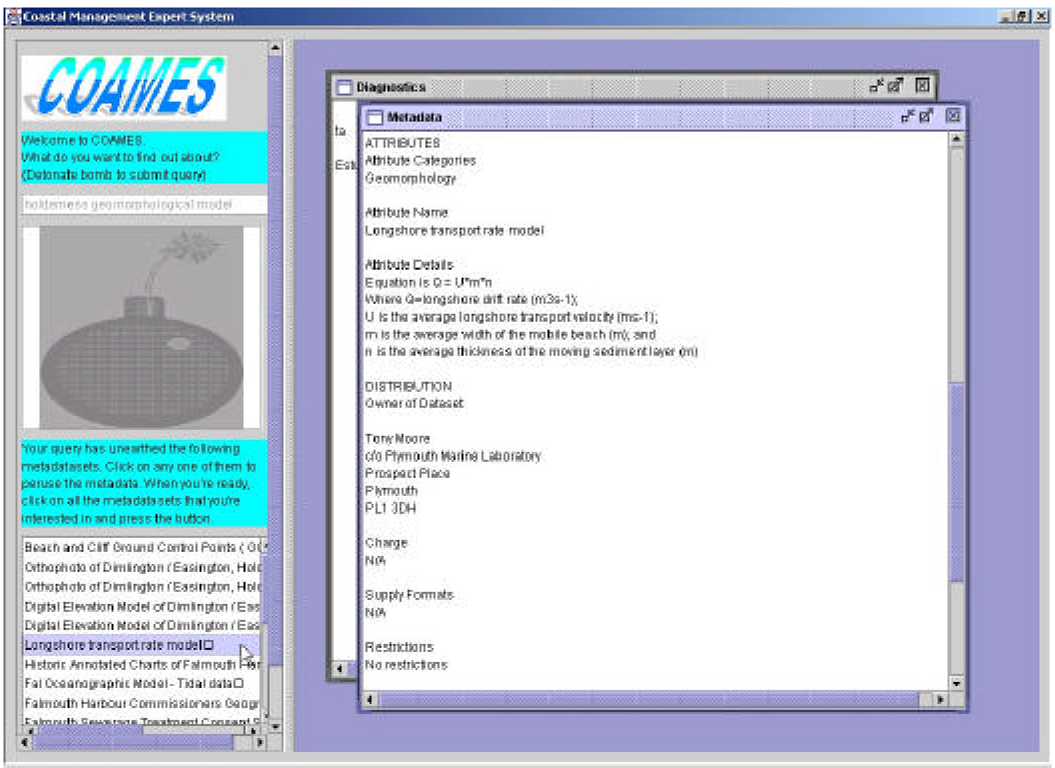


**Figure 7:** The COAMES user interface, with a dialogue partition on the left and a display partition on the right (with a metadata window).

### 3.3.2 Knowledge Base Structure

Table 3 describes the class members of 'rulesint' (rules) and Table 4 describes the 'inferencing' class (Frames of Discernment'). Tackling the 'rulesint' class first, each rule (which forms part of a rule chain) has a unique identifier ('ID') within their own FoD. The 'number of true rules' (t) relates to the number of function groups that can be implemented if the rule is proven true. The 'next rule' and 'last rule' refer to pointers to the next and previous rules in the rule chain respectively, forming a double-linked list. The 'set number' is used initially as a data file and FoD address. The 'start' and 'finish' members refer to the address range within a set of dictionary terms.

**Table 3:** The class members of the 'rulesint' class

| | | | |
|---|---|---|---|
| Rule ID | Number of true rules | Next rule | Last rule |
| Set number | Start | Finish | Number of functions |
| Function codes | Function: True or False | Number of Function Arguments | |
| Function Argument Codes | True Report | False Report | Endflag |

**Table 4:** The class members of the 'inferencing' class

| | | | |
|---|---|---|---|
| Frame of Discernment (FoD) ID | Number of members (rules) in FoD | Member codes | Number of belief values per member |
| Names of each item for which there is belief | Codes of members constituting belief items | Basic Probability Assignments (BPAs or mvals) | Belief Intervals |

The 'number of functions' (n) able to be accessed through a rule can be of various kinds, indicated by an array of n codes. The 'Function: true or false' class member indicates if a function is invoked when the rule is proved true or false. For each of the n functions the 'number of arguments' (m) must be set. The next member is a 2D array consisting of the 'function arguments' themselves, the size of which is delimited by n and m. For each of t groups of 'true' rules there are corresponding 'true reports', information that will inform the user of expert system progress. The same applies to the 'false report' if the rule is proved false. Finally, the 'endflag' indicates the end of a chain of rules (a value of one is assigned to the last rule).

The 'inferencing' class has eight members. An instance of this class corresponds to a Frame of Discernment (FoD), with a unique identifier ('FoD ID'). Each FoD has a 'number of members', or rules (n), and for each of these there is a 'member code', which corresponds to the rule IDs as introduced in the last paragraph. For each member, there may be a number of member subsets, or groupings for which there is belief. The is represented by $m_1$ for the first member, up to $m_n$ for the $n^{th}$ member. For example, inference of a member searching for evidence of iron may have three belief intervals, one for ignorance, the second for iron alone, and the third for ferromagnetic metals, a grouping of iron and nickel (which must be another FoD member). Each of these subsets are assigned names (e.g. "ignorance", "iron" and "ferromagnetic"), which are stored in a 2D n*m array. Alongside the names are their respective codes. In the example, "ignorance" and "iron" will have single member codes, stored as strings. The grouping of the "iron" and "nickel" members to form "ferromagnetic" is represented as a concatenation of the two member codes. Any two-digit codes have a zero inserted in between the two digits to avoid confusion with single-digit codes. This class member is a 3D array n*m*p, where p is the number of members in a grouping (e.g. for "ferromagnetic", p=2). Penultimately, for each grouping, there are Basic Probability Assignments (BPAs), stored in a n*m 2D array. Finally, the belief intervals for each subset of members (grouping) are stored in a 3D n*m*q array, where q is always equal to 2. This corresponds to the upper and lower probability values that form the interval (or belief and plausibility).

An account of how these classes work in practice is given in the next section.

### 3.3.3 Inference Engine

The inference engine is the heart of the expert system, assimilating user queries, and associated knowledge and data to provide meaningful output to the user. Knowledge processing is enabled through the knowledge structure via deduction, or forward chaining.

The initial query (for example "I am interested in metals, particularly arsenic") is passed in from Java to a C++ native function and broken up into the constituent words, ready for processing. The rule hierarchy for metadata is then started. The metadata category rules, FoDs and the metadata itself are loaded into the relevant structures (as defined by classes) from files. A note is also made of the current tier in the hierarchy. The rules within the FoDs are set up, with a linked list of rules being formed (through the class member nextrule) and the current rule pointer being placed at the first rule in the list, ready to traverse the hierarchy. As an example, the first Frame of Discernment contains the geographical, socio-economic, chemical, biological and physical themes (the rule structure is represented in Figure 8).

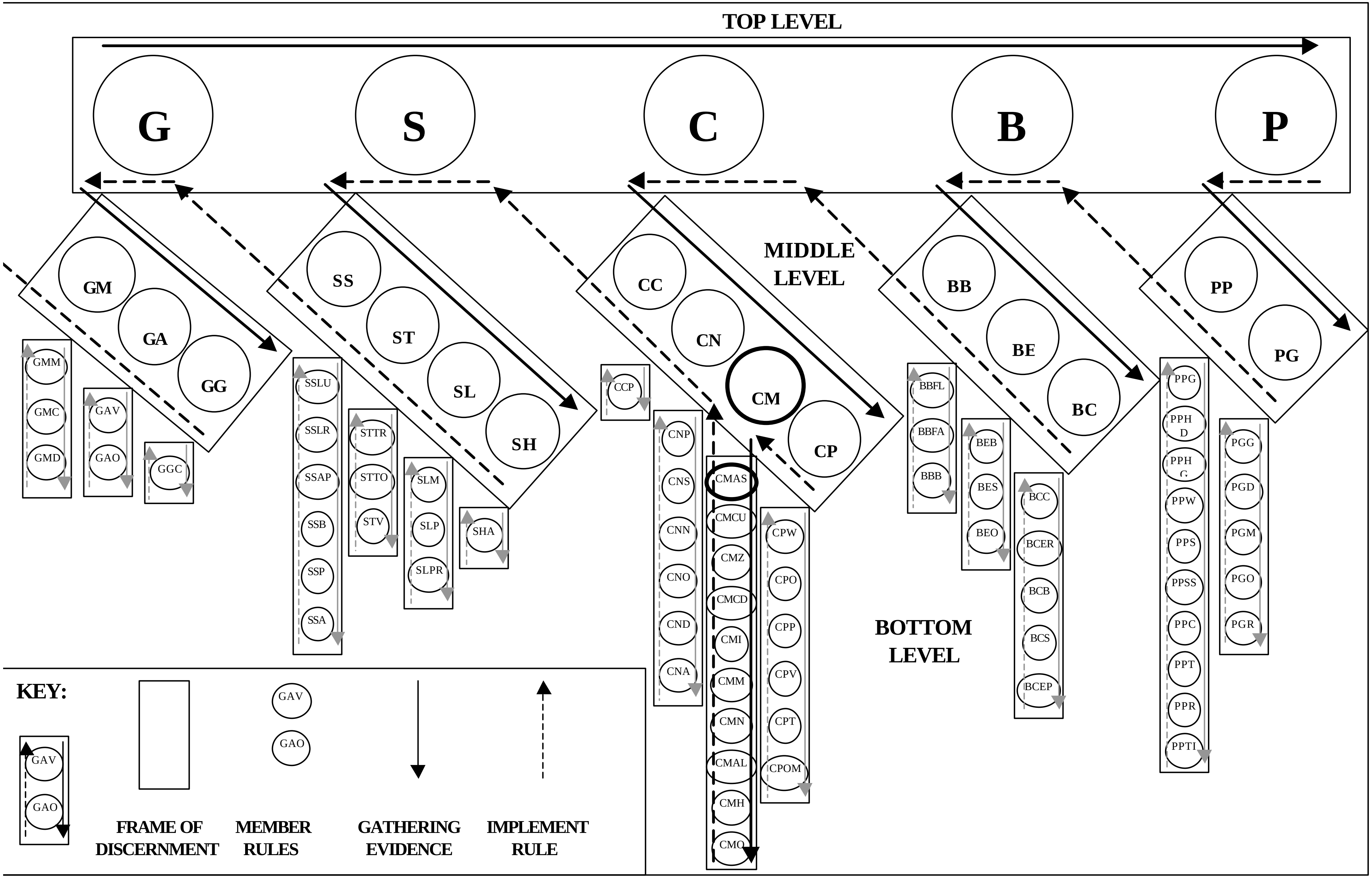


**Figure 8:** The arrangement of frames of discernment and rules, based on the hierarchy in Figure 6. A path through the hierarchy has been highlighted, based on a query concerning metals (CM) and arsenic (CMAS). The relevant rules have been fired. For a guide to the codes see Figure 6.

The first rule is passed to the inference engine, where it enters an evidence-gathering mode. All functions that are to be invoked before the belief calculation (these are coded within the rule), such as loading dictionaries, are run at this stage. The evidence gathering itself takes the form of comparison of the user query with a subset of the relevant dictionary terms (indicated by the rule's 'start' and 'finish' class members) to identify and extract important words (i.e. words of relevance to the content of the knowledge base). There is a dictionary corresponding to each of the rules on the top two tiers in Figure 8 (there is also a corresponding class definition). The comparison program is accessed through the toolbox functions. If there is a match, then the Basic Probability Assignments (BPAs) for the rule are updated, taking into account any other evidence that has been found. For example, in the case of the geographical theme, dictionary terms included 'geography' and 'spatial'. Belief in the geographical theme was zero (i.e. $Bel(\{geographical\}) = 0$), though the basic probability assignment (BPA) was $m(\{geographical\}) = 0.8; m(\Theta) = 0.2$. Should there have been evidence supporting the geographical theme, the belief would have been updated with the BPA.

In practical terms, this means that two 'inferencing' class structures (i.e. the current one, which has accumulated the belief for all evidence found so far, and the new evidence), or more specifically the 'mval' (or BPA) class member, have to be multiplied out into a matrix. This is Dempster's rule of combination, as explained in section 2.3.2 Having derived new BPAs, the belief interval is (re)calculated for each item of evidence (object), by considering if the other items constitute evidence for or against the object. The ignorance is also updated through this process.

Whether or not there is new evidence to be considered, the inference engine sets the rule pointer to the nextrule class member, and therefore the next rule in the chain. The lastrule class member for the new rule is set to the rule just processed, indicating which branch of the hierarchy has been traversed. Any belief intervals are also copied into the diagnostics array, ready to be displayed to the user when the time comes. This process of evidence gathering continues until the endrule is reached, at which time the chain of rules is traversed again, but from a reverse direction, repeatedly following the lastrule property back up to the first rule. The mode has also changed for this second traverse – all the evidence for the current FoD has been gathered, so each rule is now revisited with a view to implementing functions contained within the rule. The belief for each rule is compared with a threshold belief, and if greater, the functions and reports (copied into the diagnostics array) linked with the 'true' part of the rule are started. If less than the threshold, then their 'false' counterparts are started. Evidence for the current rule may also lie in subsets associated with other rules, so there is a check for this. The two-way traversal is necessary as evidence for one rule may affect another rule, so decisions made on actions to be implemented could only be made once all possible evidence has been combined.

Amongst the functions accessed through the true part of the rule are further rule chains, forming the middle and lower tiers of metadata themes. The upper two tiers make the initial comparison between query and dictionary terms; the third tier makes the same comparison but between the query and certain metadata fields (such as title, and attribute category, name and details). If there was a match in this case, the metadata set in question would be copied into an array ready to be passed to Java for display.

Once the top tier has been traversed, the tier below was accessed, starting with the physical theme rule. A new FoD was set up, consisting of the physics and geomorphological rules in the middle tier. The evidence gathering was started again, this time with the new rules. This process was repeated until the geographical theme rule was returned to (i.e. the first rule to be considered).

Figure 8 also shows the results of the featured query. It shows that belief was sufficient to fire rules where the important words in the user query (i.e. 'metals' and 'arsenic') matched the relevant dictionary terms. Attributes within the fired rules were subsequently used to extract metadata relevant to the original query.

## 4 RESULTS

Five progressively more challenging tests were put to COAMES. These tests are represented in Figure 9. The tests explored the use of the Dempster-Shafer inferencing mechanism for combining evidence provided from a user query, using the result to optimally and intelligently select coastal metadata related to the query. For the five tests, the 'metals' frame of discernment (FoD) will be used. There is a threshold belief that applies to all rules (e.g. arbitrarily set to 0.25 in this case) - this has to be exceeded for a rule to fire. The threshold can be raised or lowered, depending on how tight or loose the user wants the metadata search to be.

### 4.1 Metadata selection based on straightforward belief

This scenario uses a single rule based on a simple query. The members of the FoD are rules concerning the selection of arsenic, copper, zinc, cadmium, iron, manganese, nickel, aluminium, heavy metals and other metals. Each has a basic probability assignment (BPA) of $m(\{X\}) = 0.8$ (where $X$ is any FoD member) and $m(\Theta) = 0.2$ ($\Theta$ is ignorance).

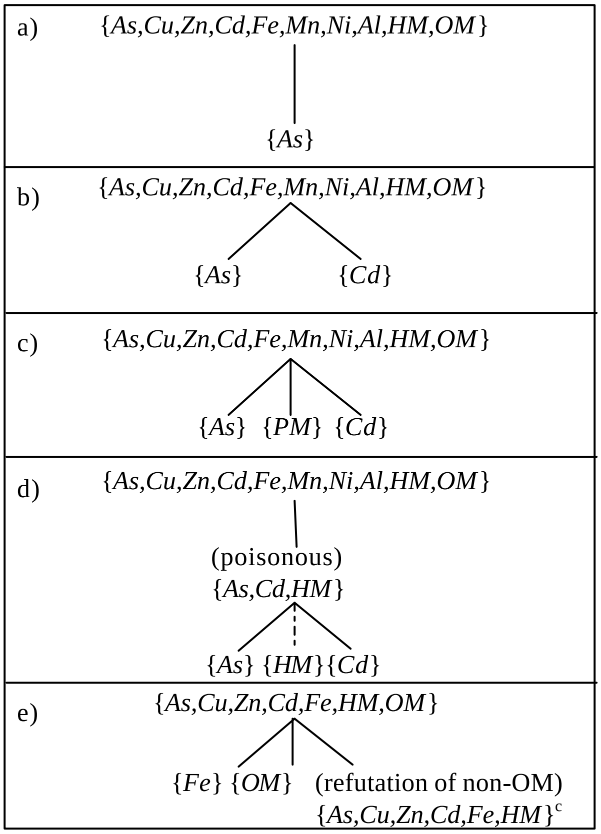


**Figure 9:** Arrangement of rules for which there is belief in the five tests. (a) straightforward belief; (b) competing rules; (c) competing rules (with further support for the weaker rule); (d) the influence of supersets; (e) belief against a rule (refutation). See section 3.9.1 for more details on calculation. (As = Arsenic; Cu = Copper; Zn = Zinc; Cd = Cadmium; Fe = Iron; Mn = Manganese; Ni = Nickel; Al = Aluminium; HM = Heavy Metals; OM = Other Metals; PM = Poisonous Metals)

Therefore, if the user query includes ‘arsenic’ then the BPA $m(\{arsenic\}) = 0.8; m(\{\Theta\}) = 0.2$. Since there is no other evidence to add, the calculation of the belief is straightforward - the belief interval of arsenic is [0.8 1.0]. The second number is the upper probability, which reduces as evidence against ‘arsenic’ mounts up. This is investigated in the next scenario. A full printout of the expert system explanation is shown in Box 1. Having exceeded the threshold belief, the arsenic metadata was extracted.

### 4.2 Metadata Selection with competing rules

This scenario uses two rules, one with a high BPA (as in the last scenario), and one with a low BPA (and correspondingly a large component of ignorance). This arrangement can be used to represent metadata that respectively have high and low confidence levels. The members of the FoD remain the same. The BPAs are also the same, except for arsenic, which has been reduced to $m(\{arsenic\}) = 0.25; m(\{\Theta\}) = 0.75$ to indicate dataset(s) of poor quality.

**Box 1:** Expert system output for straightforward belief.

Belief Interval of ignorance = [0.000 1.000]
No mention of 'Physics' category in the user query.
Belief Interval of ignorance = [0.000 1.000]
No mention of 'Geomorphological' category in the user query.
No mention of 'Physical' category in the user query.
No mention of 'Biology' category in the user query.
Belief Interval of ignorance = [0.000 1.000]
No mention of 'Conservation' category in the user query.
No mention of 'Environmental Health' category in the user query.
No mention of 'Biological' category in the user query.
No mention of 'Chemistry' category in the user query.
**Belief Interval of metals = [0.800 1.000]**
**Belief Interval of ignorance = [0.200 1.000]**
No mention of 'Pollutants' category in the user query.
**The 'Metals' category is mentioned in the user query.**
**Belief Interval of arsenic = [0.800 1.000]**
**Belief Interval of ignorance = [0.200 1.000]**
No mention of 'Other Metals' in the user query.
No mention of 'Heavy Metals' in the user query.
No mention of 'Aluminium' in the user query.
No mention of 'Nickel' in the user query.
No mention of 'Manganese' in the user query.
No mention of 'Iron' in the user query.
No mention of 'Cadmium' in the user query.
No mention of 'Zinc' in the user query.
No mention of 'Copper' in the user query.
**From searching available metadata 'Arsenic' are mentioned in the user query.**
**Metadata Extracted = Metals in Water, Sediments and Biota - Arsenic**
No mention of 'Nutrients' category in the user query.
No mention of 'Chemical' category in the user query.
No mention of 'Socio-Economic' category in the user query.
Belief Interval of ignorance = [0.000 1.000]
No mention of 'Historical' category in the user query.
No mention of 'Leisure' category in the user query.
No mention of 'Transport' category in the user query.
No mention of 'Socio-Economic' category in the user query.
No mention of 'Geography' category in the user query.

**Box 2:** Expert system output for competing rules.

……..tion of 'Chemistry' category in the user query.
**Belief Interval of metals = [0.800 1.000]**
**Belief Interval of ignorance = [0.200 1.000]**
No mention of 'Pollu…….. ……..n the user query.
**The 'Metals' category is mentioned in the user query.**
**Belief Interval of arsenic = [0.063 0.250]**
**Belief Interval of cadmium = [0.750 0.938]**
**Belief Interval of ignorance = [0.188 1.000]**
No mention of 'Othe …….. ……..n the user query.
**From searching available metadata 'Cadmium' are mentioned in the user query.**
**Metadata Extracted = Metals in Water, Sediments and Biota - Cadmium**
No mention of 'Zinc'

The BPAs of two different rules (the other is for cadmium metadata) now have to be combined, with resulting belief intervals of [0.063 0.250] for arsenic and [0.750 0.938] for cadmium, as well as an ignorance value of 0.188. The shortened printout for this scenario in Box 2 shows that metadata was only extracted for cadmium because of the poor belief in arsenic. The upper probability (also called the plausibility) for arsenic is also low due to the correspondingly high belief in cadmium, which in Dempster-Shafer terms is used as evidence to refute arsenic. The uncertainty in belief (the difference between the belief and plausibility) for the two rules is the same.

### 4.3 Metadata Selection with competing rules (with further support for the weaker rule)

A third rule is added to the previous scenario; this rule has an equivalent BPA to that of arsenic. The members of the FoD are the same as previously except for the 'heavy metals' rule, the scope of which has been expanded so that it now encompasses all 'poisonous metals' (which becomes the rule's new name). For this rule, there is a single BPA supporting arsenic to the same weak magnitude as in the last scenario (i.e. $m(\{arsenic\}) = 0.25; m(\{\Theta\}) = 0.75$ ). The aim is to see if the new evidence significantly increases the belief in arsenic. The results in Box 3 show that arsenic now has a belief of [0.135 0.308]. While this increase in belief (also plausibility) is not enough to select arsenic metadata when compared against the current threshold value, yet further evidence for arsenic may render the rule important enough to fire. Though still high, there is a corresponding drop in the level of belief in cadmium to [0.692 0.865]. The ignorance and uncertainty have also been reduced slightly by the introduction of the fresh evidence.

**Box 3:** Expert system output for competing rules (with further support for the weaker rule).

……..tion of 'Chemistry' category in the user query.
**Belief Interval of metals = [0.800 1.000]**
**Belief Interval of ignorance = [0.200 1.000]**
No mention of 'Pollu…….. ……..n the user query.
**The 'Metals' category is mentioned in the user query.**
**Belief Interval of arsenic = [0.135 0.308]**
**Belief Interval of cadmium = [0.692 0.865]**
**Belief Interval of ignorance = [0.173 1.000]**
No mention of 'Othe …….. ……..n the user query.
**From searching available metadata 'Cadmium' are mentioned in the user query.**
**Metadata Extracted = Metals in Water, Sediments and Biota - Cadmium**
No mention of 'Zinc'

### 4.4 Metadata Selection based on the influence of supersets

In Dempster-Shafer theory, groups of FoD members can be formed (supersets) to create a new hypothesis with corresponding BPA / belief. The FoD in this scenario is the same as for scenarios 1 and 2 (i.e. the 'heavy metals' member has been reinstated). However, in taking the 'poisonous metals' entity further, a new superset with that name has been formed, consisting of the 'arsenic', 'cadmium' and 'heavy metals' members. The BPAs assigned to arsenic are as follows:

$$m(\{arsenic\}) = 0.5$$

$$m\left(\left\{\begin{array}{l} poisonous = arsenic, cadmium, \\ HeavyMetals \end{array}\right\}\right) = 0.4$$

$$m(\Theta) = 0.1$$

The assignments for cadmium and heavy metals are the same as this, except for the first BPA, which relates specifically to either cadmium or heavy metals. The user query relates to both arsenic and cadmium, with an overall scenario aim of seeing whether or not the poisonous metals assignment linked with each will lead to the selection of heavy metals metadata. Box 4 displays the results of this scenario - both arsenic and cadmium have a belief interval of [0.333 0.667]. The 'poisonous metals' superset has a belief interval of [0.987 1.000], which is enough to extract metadata for all its members, including 'heavy metals'. The high belief and plausibility is because both arsenic and cadmium are subsets of poisonous metals and are therefore added onto the belief of that group. For the same reason they do not refute 'poisonous metals', leaving the plausibility at one. Note also that the belief from the superset does not get added onto the belief of each subset.

**Box 4:** Expert system output showing the effect of supersets.

……..tion of 'Chemistry' category in the user query.
**Belief Interval of metals = [0.800 1.000]**
**Belief Interval of ignorance = [0.200 1.000]**
No mention of 'Pollu…….. ……..n the user query.
**The 'Metals' category is mentioned in the user query.**
**Belief Interval of arsenic = [0.333 0.667]**
**Belief Interval of cadmium = [0.333 0.667]**
**Belief Interval of arsenic or cadmium or heavy metals = poisonous = [0.987 1.000]**
**Belief Interval of ignorance = [0.013 1.000]**
No mention of 'Othe …….. ……..n the user query.
**From searching available metadata 'Heavy Metals' are mentioned in the user query.**
**Metadata Extracted = Falmouth Inner Harbour Environmental Data - Heavy Metals**
No mention of 'Alu …….. ……..n the user query.
**From searching available metadata 'Cadmium' are mentioned in the user query.**
**Metadata Extracted = Metals in Water, Sediments and Biota - Cadmium**
No mention of Zin ……..n the user query.
**From searching available metadata 'Arsenic' are mentioned in the user query.**
**Metadata Extracted = Metals in Water, Sediments and Biota - Arsenic**

### 4.5 Metadata Selection using belief against a rule

It has already been seen that, in D-S theory, members that are not in common (e.g. {arsenic} and {cadmium}; {iron} and {arsenic,cadmium}) are used to refute one another. The last scenario deals with this refutation explicitly. The FoD in this scenario is the same as for scenarios 1 and 2, except that the manganese, nickel and aluminium members have been absorbed into 'other metals'. The aim is to show what happens when a member (i.e. other metals) is chosen to the exclusion of all others. This is based on the premise that evidence for 'other metals' negates belief in any other member in the FoD. The BPAs for 'other metals' are as follows:

$$m(\{OtherMetals\}) = 0.5;$$

$$m\left(\begin{Bmatrix} arsenic, copper, zinc, cadmium, \\ iron, HeavyMetals \end{Bmatrix}\right)^c = 0.4$$

(NOT arsenic or copper or zinc or cadmium or iron or heavy metals); and $m(\Theta) = 0.1$. In addition we have a competing BPA for the iron member: $m(\{iron\}) = 0.6; m(\Theta) = 0.4$. Box 5 shows the results of the user query (concerning iron and manganese as an example of 'Other Metals'). Iron has been suppressed effectively with a belief interval of [0.130 0.217], whilst the relatively high 'other metals' member ([0.435 0.870]) has a further influence by feeding into the 'NOT As/Cu/Zn/Cd/Fe/heavy metals' superset to give [0.783 0.870].

**Box 5:** Expert system output showing the effect of refutation.

……..tion of 'Chemistry' category in the user query.
**Belief Interval of metals = [0.800 1.000]**
**Belief Interval of ignorance = [0.200 1.000]**
No mention of 'Pollu…….. ……..n the user query.
**The 'Metals' category is mentioned in the user query.**
**Belief Interval of iron = [0.130 0.217]**
**Belief Interval of other metals = [0.435 0.870]**
**Belief Interval of NOT arsenic or copper or zinc or cadmium or iron or heavy metals = NOT As/Cu/Zn/Cd/Fe/heavy metals = [0.783 0.870]**
**Belief Interval of ignorance = [0.087 1.000]**
**From searching available metadata 'Other Metals' are mentioned in the user query.**
**Metadata Extracted = Metals in Water, Sediments and Biota - Manganese**
No mention of 'Hea……..

## 5 DISCUSSION

### 5.1 Results Discussion

The first test does little more than if there was no method used to inference with uncertainty. What is important here is the introduction of *ignorance*, as it is stated in the Dempster-Shafer theory of evidence. It is a truth that holism is an unattainable goal, since anyone modelling a domain cannot hope to totally represent that domain in all its complexity. Holism can be represented as a target, and the aim is to get as close as possible to that ideal. This, along with identifying pathways towards holism is one of the aims of COAMES.

What ignorance does is to introduce a means by which the unknown can be quantified, effectively providing one such pathway to holism. What the expert perceives is missing from the expert system rules and the metadata behind them can be represented by a proportion (a number between 0 and 1). The methods by which this figure is actually reached is another matter.

Another important facet of holism is integration. The way that Dempster-Shafer theory links different rules (via its combination method and hierarchy of supersets) satisfactorily attains this integration. Especially important in ICZM are causal relationships - the methods available here can effectively model these. For example, tests two and three have shown what happens when two rules of differing belief magnitude are combined. As expected, the weaker rule does not fire, even when further evidence is introduced. Eventually, with evidence mounting, the rule will be deemed important enough to fire, conforming to the deductive process.

Test four examines the other method of linkage, where specific rule members are grouped together to produce a new rule hypothesis. Although the BPAs assigned to the 'poisonous metals' superset and its subsets are probably too high (more testing is needed here), it shows clearly how effective this hierarchical link is. The firing of the 'poisonous metals' rule resulted in the firing of the 'heavy metals' rule beneath it in the hierarchy. This downward expression of belief would be a step towards the hierarchical passing of belief demonstrated by Shafer and Logan (1987). However, until full hierarchical capabilities are reached, inferencing will always be limited by the members of the FoD (i.e. a 'metals' member cannot directly be compared with a 'nutrients' member). This arrangement is a barrier to achieving full integration, and therefore a hindrance to the system being useful in aiding ICZM. A way round this could be to put every single rule in the same FoD, but this would be cumbersome if the number of rules rose above a certain amount, not to mention being computationally inefficient. The hierarchy is an efficient structure because it cuts time by omitting searches of whole chains if there is not enough evidence to support it.

The fifth test covers explicit refutation of a rule (refutation of rules has already occurred in tests 2-4 simply by there being two or more opposing rules present). Although the example given is highly unlikely, it would be easy to see where the power to refute would be useful in ICZM. With two rules concerning an increase in algae presence and fish productivity, for instance, a belief in the first would negate the other. Test five demonstrates this effectively.

The uncertainty element of D-S (i.e. the difference between the belief and plausibility) has been underplayed in this paper, as belief in the rules is not sufficiently complex for uncertainty to differ from rule to rule (i.e. in the featured tests it is normally equivalent to the ignorance). Uncertainty would come to the fore when using more FoD members in a variety of groupings to produce a range of supersets.

The setting of a threshold value is in this case quite arbitrary, but with increasing use of the system it can be tweaked with some accuracy to give a group of results that exactly meet the query, to a larger group of results that includes the previous group along with looser matches. It is up to the user to define this threshold level, which can change with the unique circumstances surrounding each query. In this way it is a very useful ability to have.

### 5.2 How COAMES faces the challenges of being holistic

The following discusses four major ways in which COAMES approaches holism:

- scale
- disciplines and institutions
- data, information and knowledge
- technology

Along with uncertainty, scale is one of the greatest challenges to adopting a holistic approach, according to the list provided by Bartlett *et al*. (1992) and Kucera (1995). Overcoming both these challenges is central to COAMES.

What has been seen in the case study is that the system can operate at a variety of spatial and temporal scales, without explicitly addressing the question of scale (the theme of the metadata has been the search criteria). Comparison of place names in a user query with those that may lie in the metadata will link the user and selected metadata on a common scale level (i.e. the place name itself is an indication of scale), but it is not intelligent and not in keeping with the principles of COAMES. It is suggested here that a further ruleset can be easily created, using a hierarchy of spatial scales, and using the 'Geographical Extent' field of the metadata as a point of comparison. Taken to its most extreme level, the 'world' is the group of all members within this 'spatial scale' frame of discernment, subdividing into various subsets (or supersets of the individual elements). For example, specific members may include 'Falmouth' or 'Truro' whilst the superset 'Fal Estuary' may include these members and more (refer to Figure 5). The number of members need only be as large as the metadata allows. In the same way, a hierarchy of temporal scales can be built (with the 'Period' field of the metadata as a point of comparison), ranging from the geological (i.e. millions of years) to fractions of a second.

The Fal Estuary case study was chosen to demonstrate holism across disciplines and institutions, not forgetting integration across the land-sea interface, which is central to ICZM. The use of metadata with the aidf of knowledge (see next paragraph) is the key to this, and has been shown to succeed in this case study.

Selected metadatasets effectively act as gateways to the data they are describing. A further role of COAMES is to use metadata to perform actions on the stored data (e.g. referring to an earlier application of COAMES [Moore *et al*., 2001], to derive morphometric measures such as slope and aspect from Digital Elevation Models). Though an in-depth explanation of this facility is beyond the scope of this paper, a note about the implied integration of data, knowledge and information would be worthwhile. Figure 10 shows how COAMES approaches the interplay of data, information and knowledge in an integrated manner. Metadata is taken to be a form of information, as it is a resource extracted from raw data; it has meaning in itself and gives meaning to the data. The user engages in dialogue with the system, and the input is processed by the inference engine, using knowledge to derive the relevant metadata. Metadata is seen as holding the key to the data and models situated behind it. Knowledge is the intermediate step in this case, as well as in providing output to the user. So, knowledge has an all-encompassing role.

COAMES achieves technological holism, as it brings together expert systems and GIS, as well as remotely sensed data and GPS measurements, as proved by a previous case study (Moore *et al*., 2001). It has already been said that tools such as ES and GIS, and the knowledge that the ES rely on, are essential to a computer-based holistic approach to coastal zone management (Davis *et al*., 1989; Ricketts *et al*., 1989; Riddell, 1996). Since expert systems store knowledge, they are useful in this case and also have the potential to place information into the hands of decision-makers in a useable form.

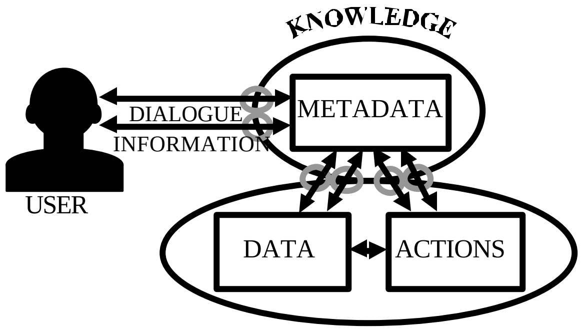


**Figure 10:** The current interplay of knowledge, information (metadata) and data in COAMES. The grey circles emphasize the role of knowledge in the system (ideally there would be use of knowledge between data and the actions applied to that data).

## 6 CONCLUSION

COAMES, backed up by the Dempster-Shafer inferencing mechanism, has proved adept at intelligently extracting metadata from the Fal Estuary metadataset. Regarding uncertainty and holism, the concept of ignorance is crucial as a pathway to holism. Also important to holism is integration, effected here by methods of rule combination and supersets. The capability of explicitly refuting a rule has also been proved. The Fal Estuary case study effectively demonstrates integration across disciplinary and institutional boundaries, as well as holism in terms of a parsimonious data-information-knowledge structure and use of various technologies. This underlines COAMES' potential value to coastal zone managers.

By way of limitations, the limited belief passing between tiers in the rule hierarchy is a severe handicap to the system's ability to handle holism. A simple solution would be to do away with the hierarchy altogether, which, although achievable with the limited number of rules presently in the system, would not be viable in the long term as the rule base grows and the inferencing process correspondingly becomes more inefficient. The current treatment of spatial and temporal scale, although adequate, is not intelligent and therefore not in keeping with the aims of COAMES. However, simple methods have been suggested that can address this issue.

## REFERENCES

Barnett, J A. (1981) "Computational Methods for a mathematical theory of evidence." In: *Proceedings of 7th International Joint Conference on Artificial Intelligence*, 868-875. Vancouver.

Bartlett, D, Devoy, R and Scalise, A. (1992) "Conceptual and methodological issues in creating a coastal zone GIS application." In: *Third Meeting on the Sea, Livorno, Italy*. Monaco: International Hydrographic Organization.

Bayliss, R and Moore, P. (2000) *Atlantic Living Coastlines - A Framework for Managing the Coast of Devon and Cornwall*. Devon County Council.

Ben-Bassat, M, Klove, K L and Weil, M H. (1980) "Sensitivity Analysis in Bayesian Classification Models: Multiplicative Deviations." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-2, 3, 261-266.

Busby, J R. (1999) "Principles, Practices and Current Directions in Information Management - An Overview." In: L Bridge (ed) *Info-Coast'99 Symposium Report*, 7-10. Coast*Link*, Coastal and Marine Observatory at Dover, UK and EUCC-UK, Brampton, UK.

Cecin-Sain, B, Knecht, R W and Fisk, G W. (1995) "Growth in capacity for integrated coastal zone management since UNCED: an international perspective." *Ocean and Coastal Management*, 29, 1-3, 93-123.

Clark, J R. (1997) "Coastal zone management for the new century." *Ocean and Coastal Management*, 37, 2, 191-216.

Davis, J R, Nanninga, P M, Hoare, J R L and Press, A J. (1989) "Transferring Scientific Knowledge to Natural Resource Managers Using Artificial Intelligence Concepts." *Ecological Modelling*, 46, 1-2, 73-89.

Deakin, R. (1994) "Data to Information: GIS and decision support for coastal management." In: A Verwey et. al. (eds.) *Hydroinformatics '94. Proc. 1st International Conference*, Delft, 2, 559-564.

Dempster, A P. (1967) "Upper and Lower Probabilities Induced by a Multivalued Mapping." *Ann. Math. Stat*., 38, 325-339.

Doody, J P. (1996) "The Coastal Directories Project." In: J H Barne, C F Robson, S S Kaznowska, J P Doody, N C Davidson and A L Buck (eds), *Coast and Seas of the United Kingdom: Region 10 – South West England – Seaton to the Roseland Peninsula*. JNCC: Peterborough.

Drummond, J E, Tait, D A and Zamlope, Z. (1997) "Building a coastal GIS using digital photogrammetry." *Photogrammetric Record*, 15, 90, 863-873.

Environmental Systems Research Institute (ESRI). (2001) *ArcInfo 8 - ArcCatalog*. http://www.esri.com/software/arcinfo/arccatalog.html [Accessed 10th April 2001].

European Commission (EC). (1999) *Lessons from the European Commission's Demonstration Programme* *on Integrated Coastal Zone Management*. European Communities: Luxembourg. 93 pp.

European Commission (EC). (2000) *On Integrated Coastal Zone Management: A Strategy for Europe*. European Communities: Luxembourg. 27 pp.

Federal Geographic Data Committee (FGDC). (2001) *Content Standards for Digital Geospatial Metadata*. Reston, VA: United States Geological Survey. Available from: http://www.fgdc.gov/standards/standards.html [Accessed 26 January 2001]

Ferrier, G and Wadge G. (1997) "An Integrated GIS and Knowledge-Based System as an Aid for the Geological Analysis of Sedimentary Basins." *International Journal of Geographical Information Systems*, 11, 3, 281-297.

Goodchild, M F. (1998) "Different Data Sources and Diverse Data Structures: Metadata and Other Solutions." In: P A Longley, S M Brooks, R McDonnell and W Macmillan (eds) *Geocomputation: A Primer*, 61-73. J Wiley and Sons: Chichester.

Gordon, J and Shortliffe, E H. (1984) "The Dempster-Shafer theory of evidence." In: B G Buchanan and E H Shortliffe (eds.). *Rule-based Expert Systems: the MYCIN Experiments of the Stanford Heuristic Programming Project*, 272-292. Addison-Wesley, Reading MA.

Gordon, J and Shortliffe, E H. (1985) "A method for managing evidential reasoning in a heirarchical hypothesis space." *Artificial Intelligence*, 26, 323-357.

Halcrow, Sir William. (1999) *Scoping Study for the Cornwall Shoreline Management Plan*.

Houhoulis, P F and Michener, W K. (2000) "Detecting Wetland Change: A Rule-Based Approach Using NWI and SPOT-XS Data." *Photogrammetric Engineering and Remote Sensing*, 66, 2, 205-211.

Kontoes, C, Wilkinson, G G, Burrill, A, Goffredo, S and Megier, J. (1993) "An Experimental System for the Integration of GIS Data in Knowledge-Based Image Analysis for Remote Sensing of Agriculture." *International Journal of Geographical Information Systems*, 7, 3, 247-262.

Kucera, G. (1995) "Object-oriented modelling of coastal environmental information." *Marine Geodesy*, 18, 3, 183-196.

Ling, X and Rudd, W G. (1989) "Combining Opinions from Several Experts." *Applied Artificial Intelligence*, 3, 439-452.

McGlade, J M. (1997) *SimCoast$^{TM}$ – an expert system for integrated coastal zone management and decision-making. NOAA – NMFS workshop on biological and physical changes within the Northeast shelf ecosystem of the USA*. Workshop Background Paper, University of Warwick.

Miller, D. (1994) "Coupling of process-based vegetation models to GIS and knowledge-based systems for analysis of vegetation change." In: W K Michener, J W Brunt and S G Stafford (eds), *Environmental Information Management and Analysis*, 497 - 509. Taylor and Francis: London.

Moore, A B, Morris, K P and Blackwell, G K. (1996) "COAMES - Towards a Coastal Management Expert System." In: *Proceedings 1st International Conference on GeoComputation*, 2, 629-646. University of Leeds: Leeds, UK.

Moore, A B, Morris, K P, Blackwell, G K, Jones, A R and Sims, P C. (2001, in print) "Using Geomorphological Rules to Classify Photogrammetrically-derived Digital Elevation Models." *International Journal of Remote Sensing*.

Payne, J L. (1999) "Coastal Decision-Making: The role of information, technology and capacity building." In: L Bridge (ed) *Info-Coast'99 Symposium Report*, 11-14. Coast*Link*, Coastal and Marine Observatory at Dover, UK and EUCC-UK, Brampton, UK.

Pearl, J. (1986) "Fusion, propagation, and structuring in Belief Networks." *Artificial Intelligence*, 29, 241-288.

Ricketts, P J, McIver, A R and Butler, M J. (1989) "Integrated Information Systems, the key to coastal zone management: The Fundy / Gulf of Maine / Georges Bank (FMG) project, a Canadian case study." In: *Coastal Zone '89*, 5, 4138-4150. ASCE: New York.

Riddell, K J. (1992) "Geographical Information and GIS: Keys to Coastal Zone Management." *GIS Europe*, 1, 5, 23-25.

Ripple, W J and Ulshoefer, V S. (1987) "Expert Systems and Spatial Data Models for Efficient Geographic Data Handling." *Photogrammetric Engineering and Remote Sensing*, 53, 3, 1431-1455.

Robinson, V B, Frank, A U and Blaze, M A. (1986) "Expert Systems Applied to Problems in Geographic Information Systems: Introduction, Review and Prospects." *Computers, Environment and Urban Systems*, 11, 4, 161-173.

Scheerer, D J. (1993) "Reasoning under Uncertainty for a Coastal Ocean Expert System." *Johns Hopkins APL Technical Digest*, 14, 3, 267-280.

Scholten, H J, Fabbri, K, Uran, O and Romao, T. (1999) "Understanding the role of Geo Information Communication Technology in building a desirable future for CZM." In: L Bridge (ed) *Info-Coast'99 Symposium Report*, 27-34. Coast*Link*, Coastal and Marine Observatory at Dover, UK and EUCC-UK, Brampton, UK.

Shafer, G. (1976) *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, NJ.

Shafer, G and Logan, R. (1987) "Implementing Dempster's Rule for Hierarchical Evidence." *Artificial Intelligence*, 33, 271-298.

Srinivasan, A and Richards, J A. (1990) "Knowledge-based techniques for multi-source classification." *International Journal of Remote Sensing*, 11, 3, 505-525.

Turban, E. (1995) *Decision Support Systems and Expert Systems*. Prentice-Hall, Englewood Cliffs, NJ.

Tversky, A and Kahneman, D. (1974) "Judgements under Uncertainty: Heuristics and Biases." *Science*, 185, 1124-1129.

Weiss, S M and Kulikowski, C A. (1984) *A Practical Guide to Designing Expert Systems*. Rowman and Allenheld: Totawa NJ.