3.3. Information Technology Review

Introduction; Criteria for selection of IT; The recent IT developments considered; Technology at the service of public participation; Knowledge representation and intelligent multimedia systems; Levels of information systems for impact assessment.

3.3.1. Introduction

The review of public participation research (previous chapter) shows the privileged status of public participation in environmental impact assessment (EIA), making it the favored ground for my thesis research. In this chapter I discuss the criteria for narrowing down the information technologies (IT) that are the focus of this thesis; I review the recent IT developments in question, in particular those that best serve public participation; I discuss more in detail knowledge representation models, based both on literature review and my previous work in this area; and finally I suggest a classification of information systems for impact assessment, according to their role and use level.

3.3.2. Criteria for selection of IT

The choice of technology to introduce in the EIA review process was a critical factor in the whole thesis experiment.

In this thesis I argue that a specific set of recent information technology developments represent a qualitative jump in IT potential for impacting public participation in EIA. Although I present this argument at a later stage, I must identify such IT developments here, since I need obviously to select elements of these IT to use in the experiment.

The choice of IT for the experiment is further narrowed down by my formulation of the thesis experiment expected evidence:

"T.1) That new IT can help lay, common, citizens to play a more knowledgeable and effective role, in public consultation concerning decisions involving technical arguments."

This suggests the choice of knowledge-based IT, applicable in the context of EIA.

"T.2) - That new IT impacts on decision-making procedures: including and up to the point where many of the current procedures become inadequate and require a new regulatory framework."

This suggests the choice of technologies that are the base of modern decision support systems; and of new information systems that offer a reasonable expectation of helping the EIA review process.

"T.3) - That you need specific IT to best support a specific kind of public participation; and that IT solely promoted by the so-called "free market forces" does not satisfy this need, neither fulfills all the potential that new IT has in this domain."

This suggests the comparative use of IT available on the market, and an IT prototype specially developed and customized for public consultation.

"T.4) - That the presence alone (or even introduction) of new IT does not necessarily promote better public participation nor improve decision-making procedures favoring public participation and is actually unlikely to do so, unless a) there is a good understanding of the underlying planning paradigms in presence, and b) an effort is made to shape both new IT and a new institutional framework in order to build bridges between these planning paradigms."

This suggests the choice of IT and IT-based planning support systems that can be used by most, if not all, actors in the EIA review process and facilitate networked communication.

3.3.3. The recent IT developments considered

Among the significant IT developments relevant to the thesis experiment, I include:

3.3.3.1. Hardware:

a) The emergence of microcomputers (and personal computing) as a mainstream technology, enabled by the development of the integrated circuit, from a period where "real" computing implied mainframes and a mandatory MIS department. Notable component is also the computing power available in relatively cheap, portable computers.

b) Internet infrastructure (wire and wireless network, based on cable and satellite IT), together with digital telephone, with increased bandwidth for data transfers over the large net of telephone lines.

c) The massive distribution spread of CD-ROM readers (mass distribution of CD-RW "burners" only came by in late 90s, not really an option in 1996, but CD-R readers were at the time much more common in Portugal than internet access)

d) Other support IT, such as satellite-based remote sensing, low cost scanners, etc.

3.3.3.2. Software:

a) Modern operating systems (Unix, Mac OS, Windows), supporting desktop and portable "personal computers" (PC), as well as terminal distributed interactive access vs. batch process of mainframe-based OS (VMS, etc.);

b) TCP/IP (Transfer Communications Protocol / Internet Protocol), giving birth to an Internet where any kind of computer or operating system can connect to each other;

c) Hypermedia, multimedia;

d) Markup Languages Standards such as HTML (Hyper-Text Markup Language), corresponding multimedia server protocols such as HTTP (Hyper-Text Transfer Protocol) and other machine independent data representation (as opposed to word files incompatibility nightmare);

e) Artificial Intelligence applications (in particular knowledge representation, knowledge bases, inference engines, expert systems), and spin-off object-oriented languages with class inheritance, message/event driven software (scripting, automated metadata maintenance);

f) Direct Manipulation Computer User Interfaces, mouse-based, with new user interface paradigms such as cut-and-paste, drag and drop;

g) GIS (Geographic Information Systems) and spatial analysis tools.

The full discussion of why the particular relevance of these IT developments is left to a later chapter; here, I will lay down the general foundation.

In my view, the most adequate and promising IT for public consultation cannot be identified only from the point of view of the end user (either expert or "lay" citizen), but also and foremost from the point of view of the knowledge input and maintenance model. If data / knowledge input and maintenance is complex then it becomes expensive (time wise, expertise wise, equipment wise), it implies a specialized body of professionals (as at the early stages of computing: analysts, programmers, card punchers, operators, separated from user), and therefore such model is not likely to succeed. I will argue that the "IT qualitative jump" includes precisely the development of microcomputer, having as consequence the direct access of the end user to the machine, together with the control of its use, and even a certain level of programming (typically interpreted languages, vs. compiled, like macros and scripting languages). Therefore, the data structure, metadata, and mechanisms for data classification and metadata input are critical to a model where direct data input and classification is done by the end user.

This emphasizes the importance of metadata sustainable strategies and models, to which I dedicated previous work, and the concern about developing collaborative and automated classification tools (e.g. script events for meta classification, etc.) for the thesis experiment, as it will be further elaborated.

In table 3.3.3.1 I present a brief chronology of some of the significant landmarks in information technology developments:

Table 3.3.3.1 - Chronology of IT landmarks

(Global Reach 2002) (Boncheck 1996) (Hardy 1993) (Kurzveil 1990) (Owens 1986) (Panati 1984) (Langley 1968)

>600 BC	The abacus (resembles the arithmetic unit of modern computers) is invented in China
387 BC	Foundation of Plato's Academy, development (among others) of mathematical theories
334 BC	Foundation of Aristotles' Lyceum, consolidation of the work of the Academy
59 BC	First regular daily newspaper, "Acta Diurna", Julius Caesar
1450	Printing press invented (Johannes Gutenberg)
1642	Pascaline, a machine that can add and subtract, is invented by Blaise Pascal
1694	Liebniz computer, multiplies by repetitive additions, algorithm still used (Gottfried Wihelm Liebniz)
1728	Automatic weaving with punch cards. (Joseph-Marie Jacquard)
1822	Difference Engine, first computer built, calculated functions (Charles Babbage)
1835	Analytical machine, with punched paper band, first programmable computer designed although never built (Charles Babbage).
1844	First long-distance telegraph , Washington-Baltimore, USA (Samuel Morse)
1847	Boolean algebra ("Mathematical Analysis of logic", George Boole)
1867	First typewriter (Christopher Sholes)
1876	First telephone patent (Alexander Bell)
1879	Notation system for mechanical reasoning, precursor of predicate calculus and knowledge representation. (G. Frege)

1888	First experiment with radio wave emission. (Heinrich Hertz)
1897	Radio emission with antenna (Alexander Popov)
1897	First patent for radio (Marconi)
1906	**First broadcast of human voice,** AM radio (Reginald Fessenden)
1927	First version of the “Differential Analyzer” (MIT), a “thinking machine for high mathematics (Vannevar Bush)
1930	18 million radios owned by 60% USA households
1936	Regular TV broadcast begins in UK
1936	Binary calculus for programming - Turing machine (T. Turing, Louis Couffignall)
1940	First fully electronic computer, ABC (Atanasoff-Berry Computer)
1944	Mark I , fully electronic computer (Howard Aiken)
1951	**First electronic computer commercialized,** UNIVAC-1 (Eckert, Mauchly)
1955	First AI language, IPL-II information processing language (Newall, Shaw and Simon)
1955	First transistor-based calculator
1956	72 % USA households own a TV
1956	First **Artificial Intelligence** conference is held
1958	First integrated circuit (Jack St. Clair Kilby)
1960	6000 computers in USA
1965	Bell Labs produce integrated circuits (W.Hittinger, M. Sparks)
1968	First ARPANET Information Message Processor (IMP), installed at UCLA (precursor to INTERNET)
1971	**First microcomputer** in USA
1971	**First pocket calculator**
1972	Created the InterNetwork Working Group (INWG), giving birth to the **INTERNET**
1974	Marvin Minsky publishes “A framework for representing knowledge”, a landmark creating the sub field of **Knowledge Representation**

1975	**First Personal Computer (PC) introduced**
1975	5000 micro-computers sold in USA
1977	**First Apple PC (**Steven Jobs, Sthephan Wosniak)
1981	IBM introduces its PC
1981	212 Internet servers in operation
1982	**First Compact Disc (CD)** Players in market
1983	90% USA households own a TV
1983	6 million PC sold in USA
1986	700 expert systems in operation
1987	1900 expert systems in operation, mostly finance and manufacture control
1989	Developed HTTP (hypertext transfer protocol) at CERN, Switzerland
1991	**First Internet Web Server and Web Browser** (CERN)
1993	1,776,000 Internet servers in operation
1993	120 web sites on-line
1996	230,000 web sites on-line
2000	25,675,581 web sites on-line
2001	**529 million people on-line (Internet)**

3.3.4. Technology at the service of public participation

In the chapter reviewing public participation, I discussed the different objectives that are pursued, from different perspectives. How does each variety of computer tool relate to each kind of public participation objective? A multimedia tool such as an "Interactive Kiosk" may clearly play an important role in education, and (maybe less important role) in information exchange and support building. As for supporting citizen input and decision-makers, there lies a bigger challenge, since it requires a qualitative jump in interactivity (support user input and non-structured search), adaptability (to different kinds of users, expert and lay), versatility (support multi-domain conceptual links) and robustness (integrate user input with system knowledge and keep the whole consistent). After all, those Kiosks are essentially a one-way street for conveying information, where there is no questioning of the contents, no feedback, no possibility of correcting or adding contradictory views to the multimedia data base. Any computer tool developed having in mind public participation should be designed to clearly respond to one or more of these needs.

Given the complexities of an impact assessment, information systems play an important role as aids for gathering and structuring related information: for analysis, and for experimenting with different hypothesis through simulation. If we take the example of evaluating impacts in infrastructure planning, a Decision Support System (DSS) may help national agencies and local governments to make strategic choices, such as: between different users of the infrastructure services (e.g. residential vs. commercial vs. manufacturing); between capital investments and maintenance of existing services; between different infrastructure sectors; between different city and regional priorities; and between different institutional and regulatory arrangements. By the same process, a DSS can help public participation, by fostering understanding of the implications of each alternative.

Different kinds of information systems play different roles. Ortolano refers to several model-based systems to study the impact of infrastructure on land use: conventional multiple regression models, dynamic simulations, multiple-market equilibrium models (Ortolano 1988). Krueckeberg suggests that different land uses or activities have typical data found repeatedly associated with them in information systems (Krueckeberg 1974).

For cases in the domain of environmental impact assessment, government agencies have accumulated some experience with specialized IT, within the techniques of information they use: press reports, newspaper ads, custom-made newsletters and, more commonly, printed versions of non-technical summaries distributed or made available in public sites, sometimes together with more detailed technical dossiers (Sapienza 1993). Less frequently, it is cited the use of presentations to groups of experts and citizens using audio-visual technology, even if it is recognized to be the only technique (from all the above) that does not present any known disadvantage (EPA 1990) (Costa 1993) (Joanaz de Melo 1993) (Rua 1993). Significantly, most of the disadvantages associated with each technique refer to its high cost, in terms of required experts and time spent (EPA 1990) (Joanaz de Melo 1993).

These are conditions that at first glance point to expert systems as the most promising IT for EIA. So why don't we observe an explosion of development of such AI systems applied to public participation?

Environmental Impact Assessments are typically multi-disciplinary: they usually require experts from several domains (environment, transportation, economy, law, city planning, etc., etc.) and frequently involve multiple institutions. This leads to certain difficulties. Besides the difficulties of institutional integration, problems arise from the need to interface not only different bodies of knowledge, but also different value systems.

Expert Systems succeeded mainly in either highly focused and specialized domains, or in domains of taxonomic nature (Winston 1988) (Han 1989) (Chen 1991) (Wright 1993). In other words, in domains where knowledge can be easily represented in one single or dominant form. It seems then that, in order to successfully apply this IT to public participation, we need to tackle the problem of allowing different kinds of knowledge to be represented in the most adequate form, without imposing a dominant paradigm of representation; and we need some metaknowledge that will help to choose the best representation formalism. By the same token, a "public-participation-friendly" system should allow different kinds of data to be incorporated and visualized in the most adequate media. The criteria of adequacy, relating kinds of data (or knowledge) with the choice of media (sound, text, picture, map, video, etc.) may be not self-evident, and also require some expert knowledge included in the system - and, naturally, some kind of inferencing ability.

This leads us to discuss more in detail the information technology developments that address knowledge representation options, and in particular those able to handle multimedia formats.

3.3.5. Knowledge representation and intelligent multimedia systems

Among the multiple IT recent developments, it is of special relevance the progress done by a sub-field of artificial intelligence: knowledge representation. Why this relevance? I indicated above a specific motivation for a specific domain: the multidisciplinary nature of EIA and EIA reviews. But we can generalize this relevance to a broader domain. Any planning process, most particularly a decision making one concerning technical-dependent options, is supported on specialized knowledge, and not just the technical data *per se*. Hence the importance of a system able to represent "planning knowledge", elements of expertise and experience that can then be captured and stored in digital form and feed some form of computer-based support tool, usable by other experts and non-experts.

In this sub-chapter I analyze the different models of knowledge representation and their limitations; I then proceed to discuss the implementations that may have a direct bearing with the thesis experiment, based on specialized literature and my own earlier work.

3.3.5.1. - The limitations of knowledge representation models

One problem that persists in the design of systems that are not only knowledge-intensive but also must support multiple domains, is the choice of a suitable knowledge representation format. The problem lies in many fronts:

• Different types of knowledge require different types of representation. This is addressed by hybrid representation systems (Heylighen 1991). (Minsky 1981) (Winograd 1975) (Woods 1975);

• Different types of knowledge require different kinds of reasoning. This is addressed by the use of multiple inference engines, and intelligent "dispatching" systems (Carroll 1987) (Gleiz 1990);

• Knowledge acquisition and maintenance modules of the system are usually so hard-coded to a specific application (with pre-defined knowledge and knowledge types) that sustainability of the system is put in question. This is addressed with intelligent user interfaces (Ferraz de Abreu 1989) (Rissland 1984);

• Knowledge management usually implies the "internalization" of knowledge and data files, that is, any bit of information must be reformatted, re-classified and some times stored for private use of the system, creating a high impedance between the system and the outside world that further limits sustainability. This is addressed by non-obtrusive metadata strategies (Davis 1977) (Ferraz de Abreu 1992).

In Table 3.3.5.1-1 , I present a summary of my compilation of the different knowledge representation models, the kind of inference (reasoning) engine usually associated with each, and the more suitable system dynamic context (system control mechanism).

Table 3.3.5.1-1 - Knowledge Representation Models

(Heylighen 1991) (Ferraz de Abreu 1989a) (Winston 1988) (Brachman et al 1985) (Minsky 1981) (Maruyama 1973)

Representation	**Inference / Reasoning**	**System Dynamic**
Expressions (equations)	Algebra	attribute driven
Rule-Based	Production Rules (forward/backward chaining)	event or attribute driven
Regular Grammars (Automata)	Production Rules (expansion)	event or attribute driven
Semantic Networks	Relational Rules	relationship driven
Object-Oriented	Inheritance (Z,N)	attribute driven
Script/Procedural	Dispatcher	event driven
Frames	Daemons	event driven
Intelligent agents	Blackboard	event driven
Case-Based descriptors	Pattern-Matching	attribute driven

Reflecting the earlier "general problem solving" orientation that prevailed within artificial Intelligence, many authors favor this or that model of representation as the most promising for any domain. The discussion concerning the relationship between representation and the world of applications is still going on (Pearce 1992) (Aiken 1991) (Davenport 1991) (Gleizes 1990) (Jaffe 1989), and it remains as an open question.

My own approach, applied to my area of concern (EIA), was to consider building a library of default representation formats for each kind of "knowledge unit", in the domain of impact assessment considered by the system.

For instance, knowledge about primary and secondary consequences of infra-structure shortfalls and of each alternative action, is more about causal relationships (*if truck traffic and weak pavement than new road is needed*) than about knowledge in depth about entities or objects (*roads, trucks*); this points towards a rule-based representation and reasoning. Other knowledge domains may depend on much weaker cause-effect relationships and be instead more based on precedent experience (like border cases in environmental law applications), pointing towards a case-based representation and reasoning. Yet other domains may be based on in-depth knowledge about entities, or objects (like land uses, or parametric description of water treatment systems), hence pointing towards the use of object-oriented or frame-based representation and reasoning (Booch 1991).

To build a library of links between domain and representation, one needs to associate with each knowledge unit a descriptor about itself, or "*metaknowledge*" descriptor (Davis 1977). For the sake of tradition, I will use in this thesis the term *metadata* with the wider definition that include the *metaknowledge* concept.

Although my earlier work in this area targeted other application areas (such as infrastructure shortfalls and natural resource management), I can draw upon this experience for this thesis research, as I discuss next.

3.3.5.2. - Rule-based representation (expert system for infrastructure shortfalls)

Rule-based representation is usually associated with knowledge expressed in cause-consequence relationships, or "causal reasoning". Expert systems are the most typical approach to handle rule-based representation and use it to infer reasoning chains. There are many examples of successful expert systems in areas like finance and diagnosis. MYCIN (medical diagnosis), developed at MIT, is one of them (Kurzveil 1990).

Applying this representation paradigm to deal with planning knowledge, I developed a prototype of an expert system dedicated to explore the cycles of cause-consequence in relation to infrastructure shortfalls (Ferraz de Abreu 1991b). This system in particular uses a forward chaining inference engine, that I developed and programmed myself based on my previous work on intelligent graphic interfaces (Ferraz de Abreu 1989a), and 5 classes of rules: definition, qualitative, quantitative, spatial, and question. Fig. 3.3.5.2 - 1 shows an index of the rules and classes in this expert system.

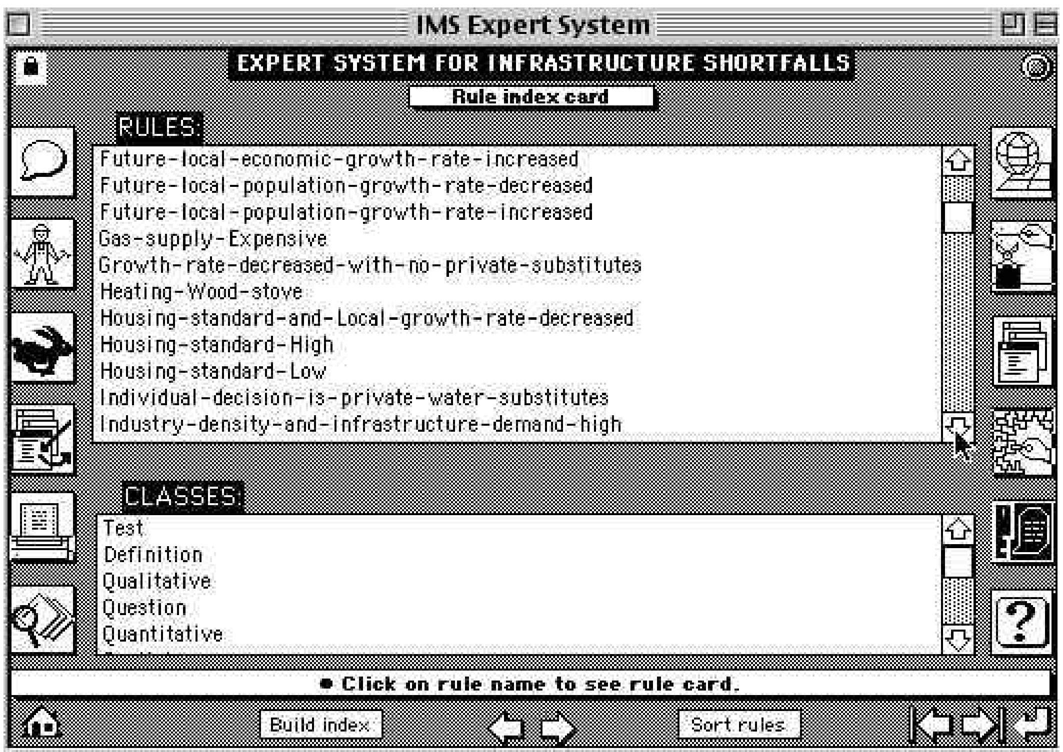

Fig. 3.3.5.2 - 1 - Rule Index card in the Expert System for Infrastructure Shortfalls

It is useful to consider a brief example of the correspondence between the issue (or reasoning) and its rule representation:

Suppose we have a great number of low-income households, therefore with very low housing standards, and that there is no service providing gas or other cooking / heating fuel (a shortfall). These houses are likely to have poorly ventilated wood stoves. This will cause indoor pollution (a primary consequence). Then, this will cause high rates of children suffering from chronic lung disorders; then, this will cause their mothers to lose hours of work time caring for them (secondary consequences); then, this will bring low productivity; if an epidemic arises, increased public health costs (aggregated secondary consequences).

Representing this reasoning with rules is fairly straightforward:

```
IF      household IS low-income
THEN    house-infrastructure IS low-standard
        ventilation IS poor

IF      house-infrastructure IS low-standard AND
        heating-fuel IS-NOT available
THEN    house-heating IS wood-stove

IF      house-heating IS wood-stove AND
        ventilation IS poor
THEN    indoor-pollution IS high

IF      indoor-pollution IS high
THEN    rate-of-children-lung-disorder IS high

IF      rate-of-children-lung-disorder IS high
THEN    mothers-productivity IS low
        public-health-costs IS high
```

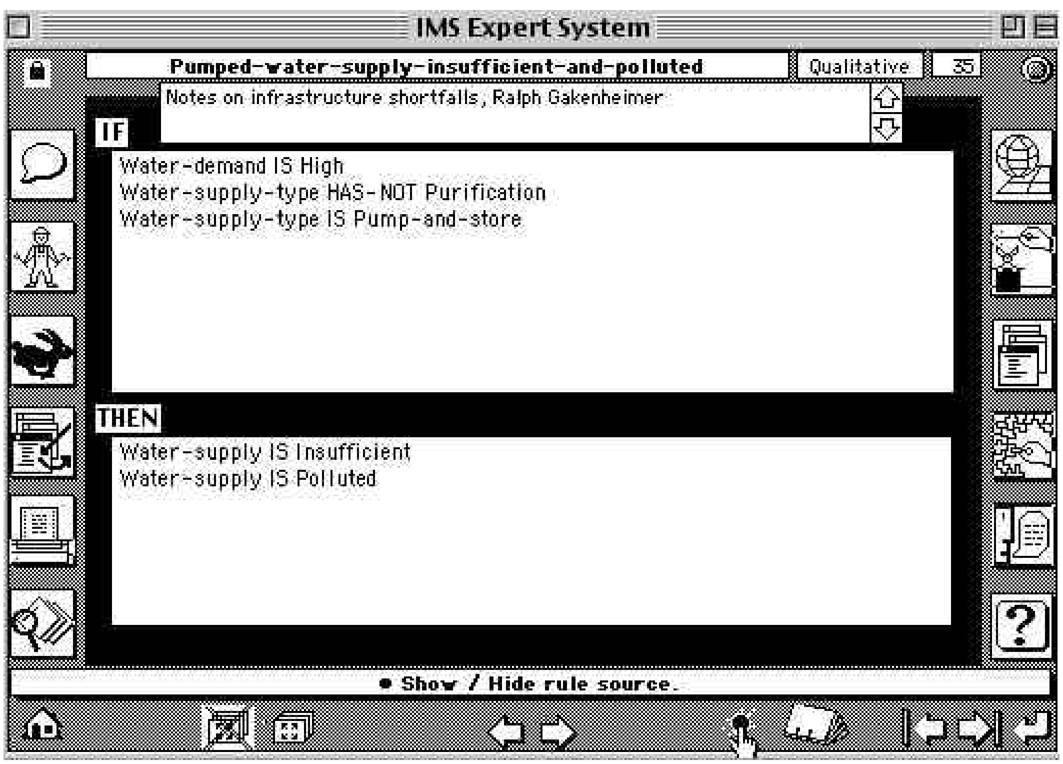

Fig. 3.3.5.2 - 2 - Rule example in the Expert System for Infrastructure Shortfalls

Fig. 3.3.5.2 - 2 shows how one of these rules is represented in the system.

The rule representation of the above reasoning is therefore adequate and simple. However, if we consider now that low productivity and increased costs are likely to cut on salaries and on health subsidies, which will perpetuate the low-income of the original families considered, we have a *positive feedback* or reinforcement of secondary consequences over the primary consequences. Representing these facets of causal reasoning with a rule-based system is not so trivial.

Because of the cyclical nature of the inference net, that is, a graph with cycles instead of a tree-graph, I implemented the inference engine in such a way that the user can visualize (Fig. 3.3.5.2 - 3) the intermediate steps of the inference process, and not just the final inference set (as it is more common). The output of this system can be extended to suggest policy recommendations, or estimate costs of shortfall situations. However, rule-based representation is clearly more suited to knowledge that can be expressed in tree-like inference nets.

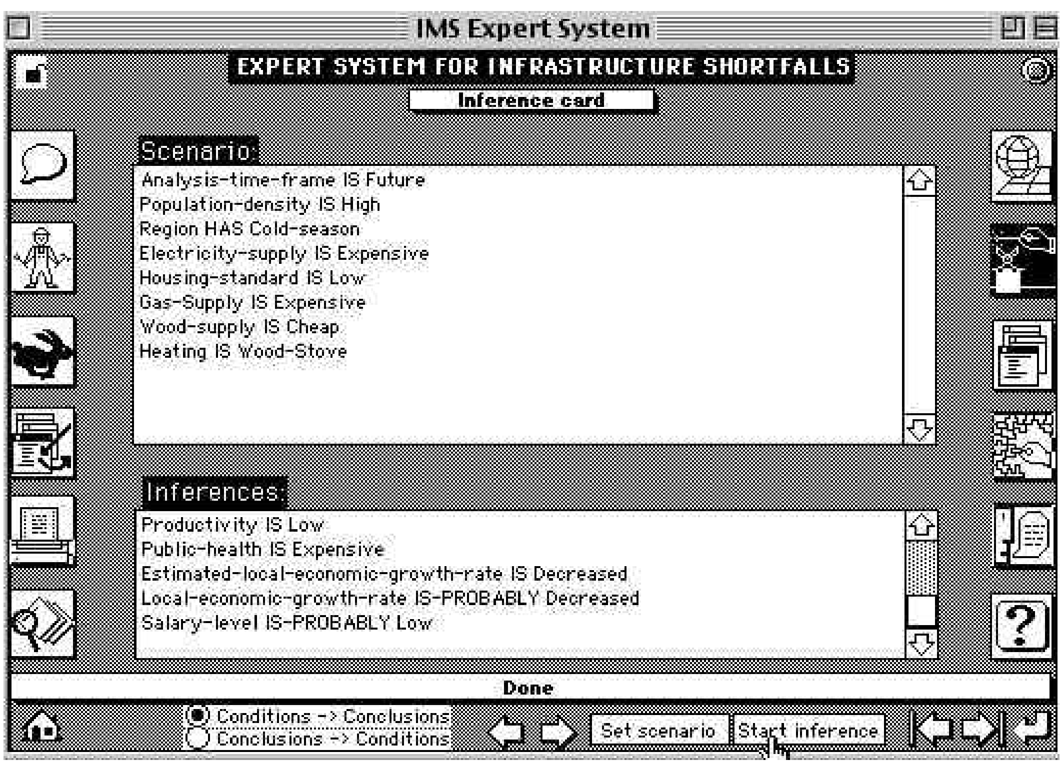

Fig. 3.3.5.2 - 3 - Expert system inference showing intermediate steps

3.3.5.3. - Rule-based vs. regular grammar representation and reasoning

Environmental impact assessments is a domain that, at first sight, seems to suit itself well to a rule-based representation model, since it is frequent to listen to experts arguing for cause-consequence relationships, using a "causal reasoning". But instead of the usual tree of inference, many problems in impact assessment demand also other forms (like a non-tree graph, or graph with loops) able to capture cycles and feedback. Representing cycles is important because consequences of impacts, just like the infrastructure shortfall example, may affect individuals, activities and the environment in general, cycling through all of them. A cycle imply that some kind of feedback is present, either positive (reinforcement) or negative (regulation). In such cases, a "regular-grammar" (state automata) representation model may be more adequate.

To clarify my application of the notion of positive and negative feedback's in modeling shortfall consequences, consider this more aggregated graph of inferences with the following factors:

In a city, there is a poor garbage collection service, resulting in the accumulation of garbage in the area (G). This will increase the number of bacteria present in the area (B). This will increase the number of diseases (D). All these are direct proportionality functions (if the number of G increases, B increases; if G decreases, B will decrease). Now consider that increasing diseases will induce people to leave the city (or will kill people), causing the reduction of the number of people in the city (P). This will cause the quantity of garbage to decrease, that is, a case of negative feedback or regulatory effect of the secondary consequences over the primary consequences.

In Fig. 3.3.5.3.-1 is a graph representation of this simplified model (adapted from (Maruyama 1973)), with other dimensions added: S for sanitary improvements (which will decrease directly both the number of diseases and bacteria); C for migration into the city (increasing the number of people in the city) and M for modernization of the city. In general, a + sign identifies a direct proportionality relationship, a - sign the inverse proportionality.

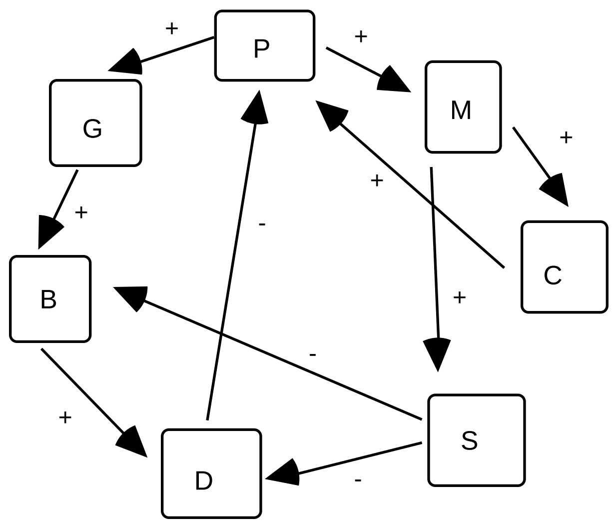

Fig. 3.3.5.3.-1 Graph representation of the inference net of shortfall consequences

This representation formalism is simple, yet very powerful. For instance, by counting the number of negative signs (inverse proportionality relationships) within a complete cycle, it is possible to forecast either a positive feedback - reinforcement (even number of minus signs) or a negative feedback - regulation (odd number of minus signs), for that cycle.

Several authors developed models of different aspects of these relationships that have some component relevant to the analysis of the shortfall implications. Laredo emphasizes the importance of the sectoral linkages of water services in its impact on agriculture, industry, health, and housing (Laredo 1990). Scenarios involving infrastructure shortfalls kind of problems can serve as a testbed for the potential of this representation formalism.

3.3.5.4. Case-based representation and reasoning issues

Environmental Impact Assessment (EIA) case-based reasoning presents issues that are similar to the ones faced in the domain of natural resource management, as I concluded from previous research (Ferraz de Abreu 2002b).

Case study materials collected for other purposes can be useful for "*crude hypothesis testing*" (Feeny 1992). They may be used to generate hypothesis inductively, as suggested by Elinor Ostrom (Ostrom 1992); or they may be used to test hypothesis derived from theory or from previous inductive reasoning. Just as within the EIA domain.

Examples of case studies to test hypothesis are the studies to examine the effects of group size on the performance of institutions managing common-property resources. Bullock, Baden and Feeny mention similar use of case studies (Baden 1977) (Feeny 1992). One advantage of this research approach is that it reveals patterns of variables or factors impacting on the outcome of the case. For instance, Feeny reports four factors that emerged from the referred study: cost of intragroup enforcement, cost of group exclusion, cost of decision making, and cost of coordination (Feeny 1992).

Representing case-based knowledge is not trivial either, and I did not find any example of a software implementation, other than adaptations from general-purpose data base management systems.

One common problem with domains that rely heavily on precedent experience, as commonly is the case in EIA, is the lack of a structured library of relevant cases. The problem is compounded by "syntactic" and "semantic" sub-problems:

On one hand, one needs more than written papers or reports to grasp the complexities and subtleties surrounding each case. For instance, dynamic visual data - typically recorded in videotapes, during series of field surveys - is often essential (Wiggins 1990). The sequential nature of the traditional analog video devices makes the search for the significant video segments a time consuming and tiring task, which further discourages the integration of that data in the analytical process.

On the other hand, case studies often provide conflicting evidence. No simple system can keep its consistency under these circumstances; for instance, it is not possible to use the already "traditional" approach of Truth Maintenance Systems in Database and Expert Systems.

Having in mind natural resource management, I designed an information system to make the most of a case-based approach: a "multimedia data base of research cases". Reviewing the data structure for this system is relevant, since it was one important step towards the system I prototyped to test the potential of "intelligent" multimedia technology in the context of EIA reviews.

a) Data structure:

The data unit of this multimedia data base is the *research case*. The body of this data unit is structured the following way:

- Case identifier (usually a name). Serves as index key;
- Context (resource type, geographic location, etc.);
- Initial status (conditions at a date defined as the beginning of the research period);
- Actions (deliberate, controlled human intervention impacting on the resource and its users);
- Events (non-deliberate, non-controlled natural or social changes impacting on the resource and its users);
- Final status (conditions at a date defined as the end of the research period, if past, or the current date);
- Outcome (degree of success or failure, which may be user defined);
- Experts (persons contributing with information).

b) Data model:

Modeling this kind of data (research case descriptor) in such a way that the system is comprehensive but at the same time simple to consult and update, is not trivial. The popular aphorism "there is no such thing of a free lunch" is particularly valid in the world of data base design. In this case, the more structured the data is, the better we can manipulate it; but also the greater loss of information content happens in the process.

In my approach, I intended to test a data model with two levels of abstraction (consequently, two levels of structure) to capture as much as possible the best of the two worlds; in this case, the trade-off is with redundancy. To illustrate this data model, consider Fig. 3.3.5.4.- 1:

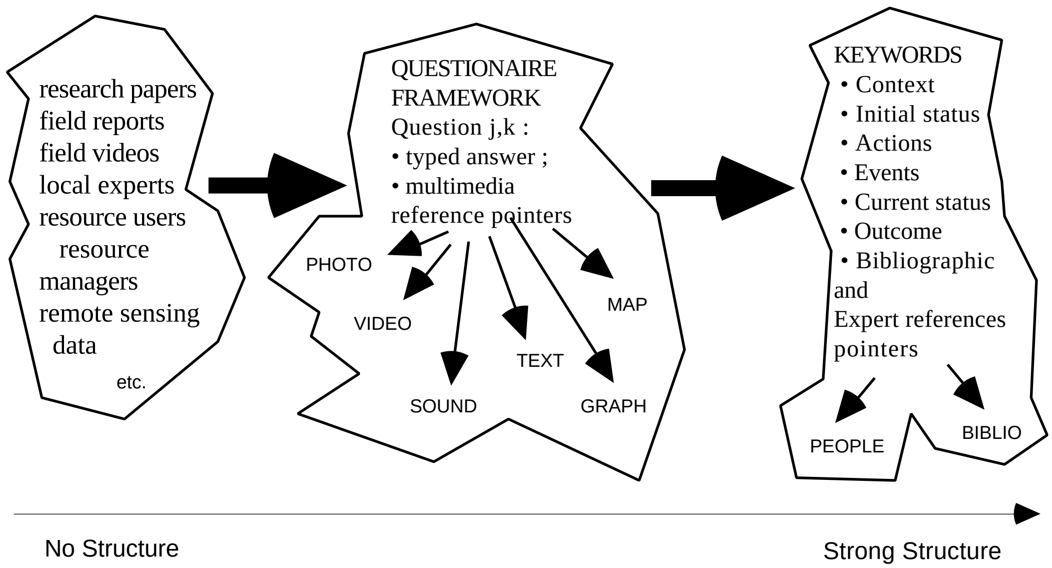

Fig. 3.3.5.4.- 1 - Data Model for Case-Based Knowledge Representation

Outside the data base, data is not constrained in any way by a particular data model structure. By bringing it in, through a pre-defined questionnaire, and then linking each answer with specific multimedia references (for instance, several discrete video segments), some structure is gained, which facilitates for instance comparative analysis between different cases. At the same time, some information that does not fit neatly in the questionnaire framework, will be lost. This is the first level of abstraction, which still allows a large degree of freedom, like free text directly typed into the data base, possible contradictory opinions and references, etc.

A second level of abstraction is then possible, by "summarizing" the characterization of the research case by sets of keywords. This allows for more sophisticated data analysis, such as cluster analysis, search by patterns of keywords (Pearce 1992), and deductive or inductive inferencing by generalization from the "nearest" matches among the data base cases (case-based reasoning). The price to pay is a more imperfect representation of the case - semantic loss - together with some redundancy - keywords may in some cases be a simple repetition of some of the sentences of the questionnaire's answers.

By adopting an object-oriented representation, it is possible to structure even more this information with recourse to a hierarchy of classes and class instantiations arising from the realm of the Environmental Impact Assessment. For instance, a class Industry has associated all the relevant information (relevant to impact assessment) that is shared by any and all industries; when a industry is added to the system, it is sufficient to declare it as belonging to the Industry class, in order to inherit automatically all that information. A taxonomy of industries can be represented under this class hierarchy (for instance Chemical industries, Textile industries, etc., for Industry class; Paint industries, Fertilizer industries, etc., for Chemical industry subclass, etc.). Problems may arise in some cases given the lack of rigorous consensus over the definitions and concepts.

The handling of conflicting evidence is a challenge, but in this data model it is possible to adopt Lenat's approach of co-existence of multiple belief or truth systems within the data base. This approach implies the introduction of an operator to detect conflict, and to call upon meta-rules to handle each conflict type.

An example of such meta-rules would be: if two cases (A, B) present all the same keywords identifying status, actions and events, and one of the keywords identifying outcome is different (not matched), we have a conflict of evidence. Then, search for all other cases in data base containing the conflicting outcome keywords; select among the cases those that contain the larger match of similar keywords defining status, actions and events; list the non-matching keywords defining status, actions and events; suggest to the user that the reason for conflicting outcome may be found in the fact that one of the keywords in this list is in reality present in case A, despite the fact that case A representation was not given that keyword. This way, the system has the means to infer best possible matches in conditions of conflicting truth systems, and give useful hints on analytical efforts to "break" the conflicting evidence.

3.3.6. Levels of Information Systems for impact assessment

One kind of system, or for that matter, one kind of IT, won't solve by itself the technological handicap presented by current systems when applied to public participation. It is therefore important to understand the context (of other systems and IT) in which it will play its best role.

In Fig. 3.3.6. - 1, I introduce a diagram modeling the role of different information systems in the quest for analyzing and correcting impact assessment problems. The diagram proposes four levels at which information systems may operate, and complement each other: source level, conceptual level, analytical level, and use level.

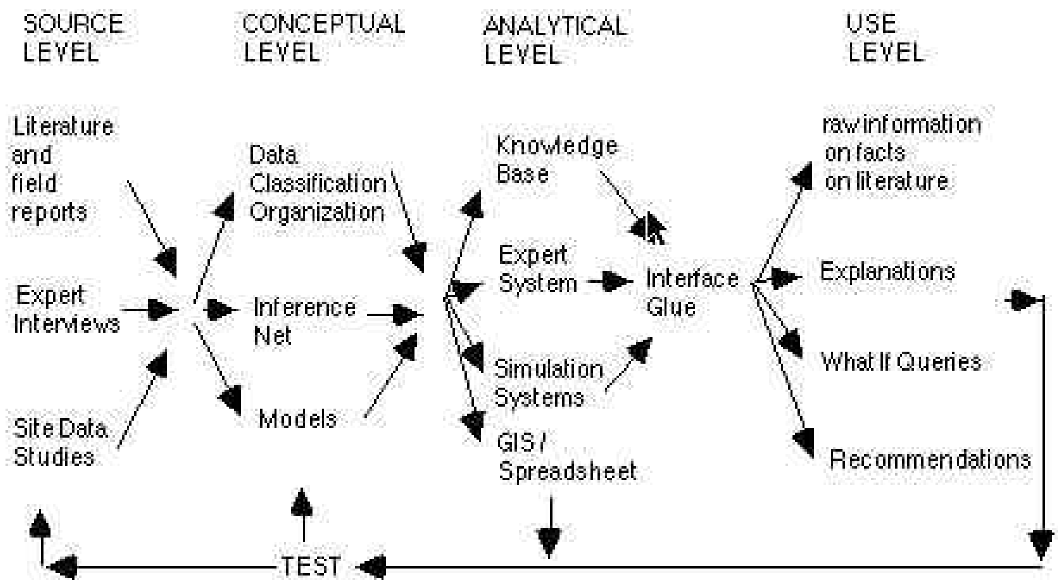

Fig. 3.3.6. - 1 - Role levels for information systems in impact assessment

An experimental prototype of an "Intelligent Multimedia Decision Support System" should be able to interact with any module at all these levels. However, targeting the use to public participation poses heavier requirements on the "Interface glue", to handle different levels of user domain expertise.