FRANCQUI Chair Conferences
87-88
State University of GENT
(BELGIUM)

Fondation Francqui

Knowledge Based Systems

Towards Applications to Natural Resources Management and Biotechnology

Alain Pavé

-Revised 1992-

Foreword

This text was written mainly after conferences presented at the State University of Gent in April and May 1988 in the framework of the FRANCQUI Chair Foundation (International Francqui Professor). After these conferences, other ones on the same subject were presented particularly at the Central University and at the French Institute of Pondicherry (India). This version was terminated in March 1991 during a mission at Pondicherry and revised in March 1992. It is also the teaching consequences of two research contracts given, the first one, by the PIREN, an interdisciplinary research program on environment of the CNRS and, the second one, by the French Ministry for Research and Technology.

The author wishes to thank Professor G.C. Vansteenkiste who invited him in the State University of Gent and who shares widely his opinion on modelling approaches of biological and ecological systems, F. Rechenmann who was at the origin of his interest for Artificial Intelligence techniques and the participants of conferences for their useful remarks and comments.

The reader has to be informed that this text is not a fundamental course on Artificial Intelligence and applications. It is not the goal, and out of the author's skill, there are many good books in the domain. It is just a short introduction to basic principles of A.I. and applications in some domains of Natural Resources Management and Biotechnology. We consider aspects of knowledge bases, which can be viewed as models of real objects or situations then the text is presented in a **modeling approach** point of view.

Knowledge Based Systems - Towards Applications to Natural Resources Management and Biotechnology.

Alain Pavé
Professor at the University Claude Bernard - LYON 1 (France)
Titular of the FRANCQUI CHAIR 87-88

Contents

1- Overview of Knowledge Based Systems and Expert Systems.

Artificial Intelligence, and more precisely Knowledge Based Systems, has advanced considerably during the last few years. Some applications are now well known (MYCIN, DENDRAL, Prospector, OPS5, TOM…). The purpose of this text is to briefly present the state-of-the-art in Knowledge Based Systems and to show how formalisms and associated inference processes can not only aid and but also be a complement or an alternative to the mathematical modelling of complex systems or situations, particularly the "natural systems". Software connected to knowledge bases will certainly be important tools of the future in many fields of human activities ; particularly in the following domains :

- natural renewable resources management : for example, in agriculture and forestry, and even marine exploitation and management…
- biotechnology : from research laboratories to industrial applications.

The presentation is organized around a "modelling point of view", so Knowledge Based Systems are seen as modelling tools in the same way as mathematical formalism for "classical" modelling approaches, because some methodological aspects are common to these approaches. This is also the case for data structure modelling in Data Base Management Systems.

Some examples of works presently developed are presented, such as ECO and Edora for mathematical modelling assistance, PLANTER (Smith, 1985) a crop planting expert system…. Recent results and concepts (qualitative reasoning, spatial reasoning, taxonomic reasoning,…), which lead to new ideas of applications, are also discussed.

1.1. Knowledge modelling and representation.

All cognitive processes need a formal representation of the real world with associated processes to manipulate this representation.

To be handled and transmissible, the knowledge must be formalized, from hard formalisms (*e.g.* mathematical models) to soft ones (natural languages). Each has its own derivative rules and contains some implicit inference or deductive processes which enable the deduction of facts (or theorems), i.e. knowledge, from hypothesis or premises and other facts previously established (*cf.* Fig. 1). Mathematical formalism is particularly efficient to represent physical systems (*e.g.* simple electrical circuits), but there are many difficulties, biological as well as technological, in more complex and organized systems : (for example, for a complex rocket such as Ariane an expert system to aid the launch and control during the flight was seriously considered, although one could think that such a system can be represented by a mathematical model [Misiti *et al.*, 1987]).

Practically, only a part of the knowledge can be represented in a mathematical expression. This is an important point because we are generally interested in the control, prediction, diagnosis, etc… Secondly, it is well known that some "experts"

are particularly clever at practical situations analysis in their field of competence, using only their "internal representation" but without using of any precise formalism. Lastly Computer Sciences is devoted to do manipulation of different knowledge representations :

- mathematical representations : *numerical representation*, the most classical, is handled generally by programs in procedural forms (e.g. simulation programs which are often an extension of mathematical representation), and *symbolic representation* can be now taken into account by using formal calculus systems (*e.g*. Macsyma, Reduce and Mathematica),
- texts : are the traditional supports of knowledge, but we still do not have sufficiently efficient tools to analyze them. One of the biggest scientific problem to make computer systems to "understand" natural languages. In another way, *hypertext* concepts lead to a new manner to approach text processing.
- specific of computer sciences for information processing (e.g. Data Base Systems), or for organization and manipulation of elementary knowledge (e.g. Knowledge Based Systems, or more specifically Expert System).

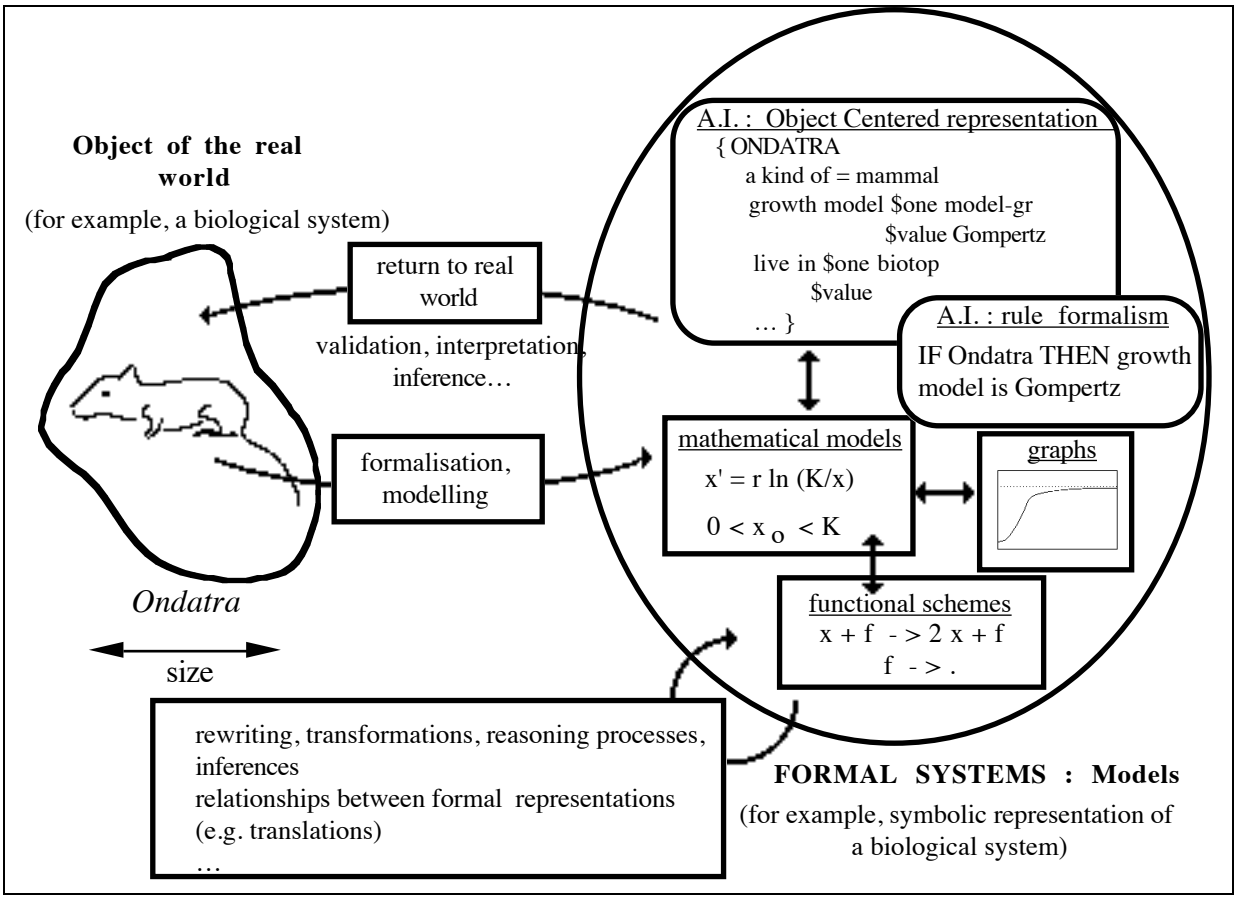


figure 1 - Relationships between the real word and some formal systems which are used to build models. Models are not only mathematical but also graphical, based on functional schemes, or on artificial intelligence formalisms…. Formal systems include specific rules for formula transformations and reasoning processes. There are also possibilities of relationships between different formalisms and in some cases translation is also possible (for example, the translation rules from functional schemes which look like formalism used in Chemistry to mathematical models have been very well presented by Garfinkel (1962). The translation from a differential system to a chemical like reapresentation has been proposed by Pavé and Pagnotte (1977)). This philosophy on models and modelling not restricted to mathematical ones are increasingly accepted in the scientific community (*cf*. for example Pavé, 1989).

In each case a formalism is used to represent knowledge. Perhaps we can speak of knowledge only when a formalism and the associated rules or procedures of manipulation and transformation exist, but even they are not explicit as in the human brain (following the physical symbol system hypothesis of Newell and Simon, [cf. Newel, 1976]). The general remarks and methodology of mathematical modelling approach can then be adapted to a more general knowledge modelling approach : (1) **definition of goals of modelling** (what are the expected results ?), (2) **system analysis** (identification of objects and relationships between objects), (3) **formalization by using schematic intermediate representations** (like entity-relations diagrams used in relational data base descriptions), (4) **implementation in an operational language or system** (the choice of a formalism which can be manipulated), (5) **tests, analysis of properties and validation** (are the goals reached ?) (*cf.* for example, Pavé, 1989).

Knowledge Based Systems are models of the real world. We speak of intelligent systems because some aspects of intelligent human behavior are simulated, particularly specific processes used by "experts" during their reasoning. These systems can also take into account different kinds of knowledge from scientific to empirical ones, in a declarative or procedural form. This apparent generality leads us to think that many fields of human knowledge can be approached in this manner. To day, we are at the beginning of important developments but there are still many problems to be solved, although some practical applications are already possible.

Expert Systems are well known essentially because applications were developed such as MYCIN, DENDRAL, PROSPECTOR, and certainly many domains of human activity which need "expertise" can now be considered in a computer aided system. Some specific problems encountered in natural resources management and in biotechnology, both at the fundamental and application levels, can be approached by this method. It is certainly not a red herring but it demands, like other modelling approaches, significant and hard work.

To summarize knowledge representation needs **formalisms and adapted processing mechanisms**. For examples :

- **Mathematical modelling** : formula and rules of transformations, (for example, symbolic representation of an analytic function and derivation rules which lead to a derivative which is also a function, *i.e.*, a formula correctly written following the mathematical syntax). Mathematical models are efficient but often too simple to represent the knowledge about a complex system.
- **Simulation modelling**. These models are generally large mathematical models completed by logical relations and empirical functions, but with limited processes (essentially numerical calculus).
- **Artificial intelligence modelling**. This approach offers a framework to represent large categories of fact and knowledge on the real world, particulary qualitative information.
- **Natural languages**. Texts are very general but complex to analyze (*i.e.* the syntactic aspect is already quite difficult to approach, and the semantic one is very difficult). However today the hypertext concepts enable high level manipulations .

1.2. Declarative and procedural forms

The difference between **declarative** and **procedural forms** (i.e**. What?** and **How?**) can be illustrated by taking the example of the definition of a circle :

- at the mathematical level one can say : *a circle is the set of equidistant points from a given point in the plane*.

From this definition, the following rule can be stated :
<u>if</u> a considered set of points is equidistant from a given point O on the plane
<u>then</u> this set is on a circle whose centre is O.
It enable to recognize a circle.

This is a **declarative** form and we can now answer the question : what is this set of points ? (a circle or not).

But how to draw a circle ? Then also, from the mathematical definition we can define a procedure to draw a circle:

begin
.take a compass;
.fix one leg;
.turn the opposite leg till the extremity comes back to the starting point;
end.

An equivalent form can be written in a procedural language to draw a circle on a screen. These are **procedural** forms of knowledge representation of the circle and these algorithms **generate equidistant points** on a plane. (a paper sheet or a screen).

Eventually, in computer science :
- procedural forms concern classical representations of knowledge (programs written in FORTRAN, PASCAL, … classical procedural languages),
- **declarative forms** are principally used in Knowledge Based Systems and also in Data Base System.

1.3. Structure of Knowledge Based Systems

The commonly accepted structure of Knowledge Based Systems today is roughly schematized in Fig. 2. For example, we can consider the architecture of the system developed within the framework of the EDORA project [*cf*. Pavé & Rechenmann, 1986, Rousseau & Rechenmann, 1988, Pavé, 1988]. The goal of the project is to build a computer system to aid the user in mathematical modelling approaches of biological phenomena (Fig. 3).

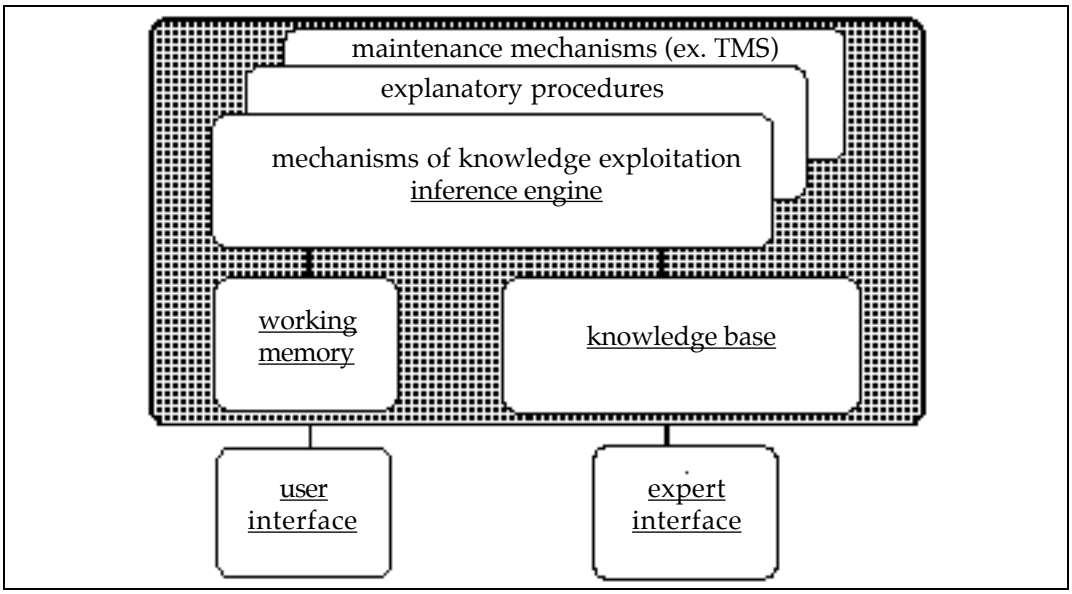


Figure 2. Commonly accepted basic structure of Knowledge Based Systems. The inference engine examines the knowledge base in a declarative form. It deduces new facts which are stored in the working memory. Additional modules can explain the steps of reasoning processes and to maintain 'truth' in the knowledge base when new hypotheses are proposed or deleted during reasoning processes.

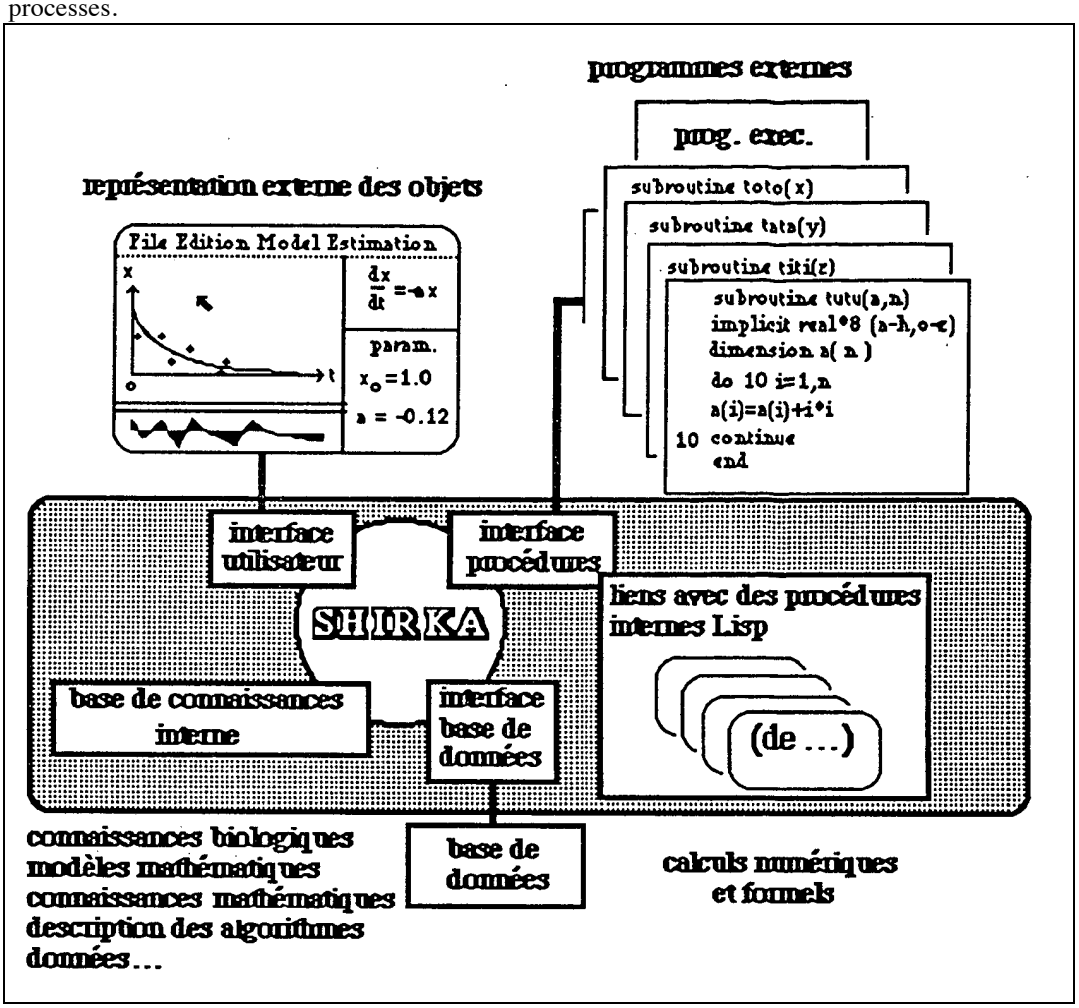


Figure 3. Example of Edora system organization. The originality is the management of declarative and procedural knowledge through an unique formalism (an object centered representation handled by the Knowledge Base Management System SHIRKA).

1.4. Principal knowledge representations and knowledge processing

First of all principal formalisms are presented briefly, and later simple examples are given to illustrate production rules and object centered representations and also the basic reasoning processes working on these representations.

1.4.1.Formalisms of representation

Currently three principal classes of representation can be distinguished :

a) Production rules

This representation is the better known and is based on elementary rules :

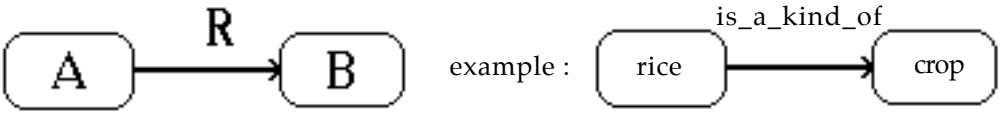
if <premises>|<conditions> then <conclusion>|<action>

They are defined on a logical basis (Horn clauses) with (first order logic) or without (logic of propositions) variable(s). The above definition of a circle illustrates this representation. Another example is presented in the next section (*cf.* 1.4.2.a))

b) Semantic networks

This representation has been studied on the basis of interpretation or significance of representations of objects and relationships between objects.

A and B are objects, R is the relationship between these objects



The next one, the object centered representation, takes into account the principal features of semantic networks, so we will not develop it.

c. Object Centered Representation (O C R)

It comes from the notion of frame (Minsky, 1975). The central concept is the *object* which is defined by its properties described in slots. The characteristics of *slots* and their values, or the way to find these values, are specified by *facets*. We then get a structured representation which is called, according to different authors as a frame, a schema, etc… An example from the Edora system is presented in section 1.4.2.b. Generally the formalism is based on a scheme which looks like the following one :

object_name		
	slot 1	*facet 1* specification or value 11.
		facet 2 specification or value 12.
	…	
	slot n	*facet 1* specification or value 21.
		facet 2 specification or value 11.

Objects are organized hierarchically in a knowledge base. Inheritance properties are built in (*i.e.* object inherits slots and corresponding features of upper levels

objects. Multiple inheritance is possible. For example, in the following figure object 1213 inherits objects 12 and 13 :

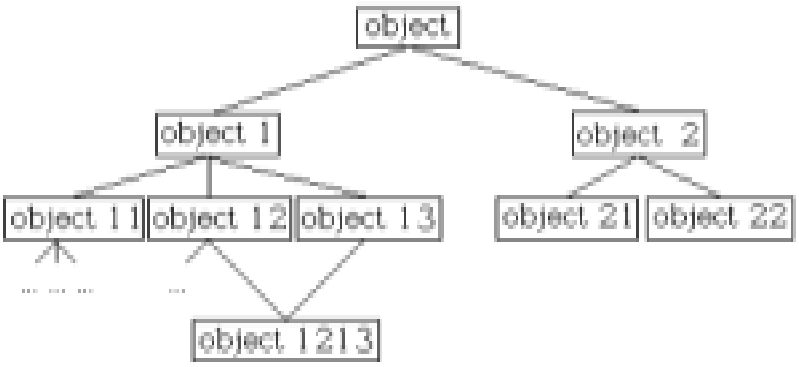


Fig. 4 and 5 illustrate this representation.

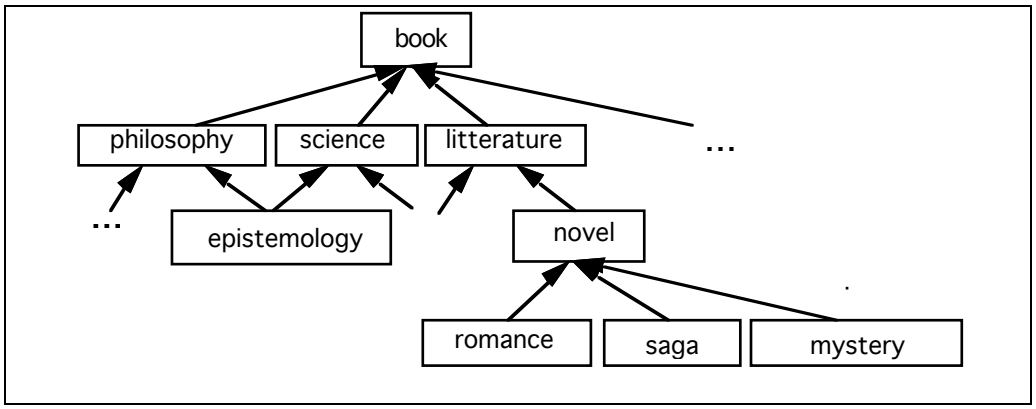


Fig.4 - Objects can be organized in a hierarchy or even parallel hierarchies. An object can inherit properties from more than one upper object. For instance, if we look at a hierarchy assumed to represent books, a book which is related to epistemology has to inherit from more general ones : philosophical and scientific.

1.4.2. Simple examples and basic reasoning mechanisms

a. Production rules

The example chosen is relative to crop management. Obviously it is schematic and even quite stupid, but clearly it is not possible to present a realistic one which would be too complex. Nevertheless this example is sufficient to illustrate the basic processes. These processes involve basic structures as presented in Fig. 1 : base of facts (basic facts which correspond to user observations or hypothesis and produced facts which result from application of rules), base of rules (which is built with human experts and corresponds to logic relationship between facts and objects). For example, consider the following case with one object in each structure :

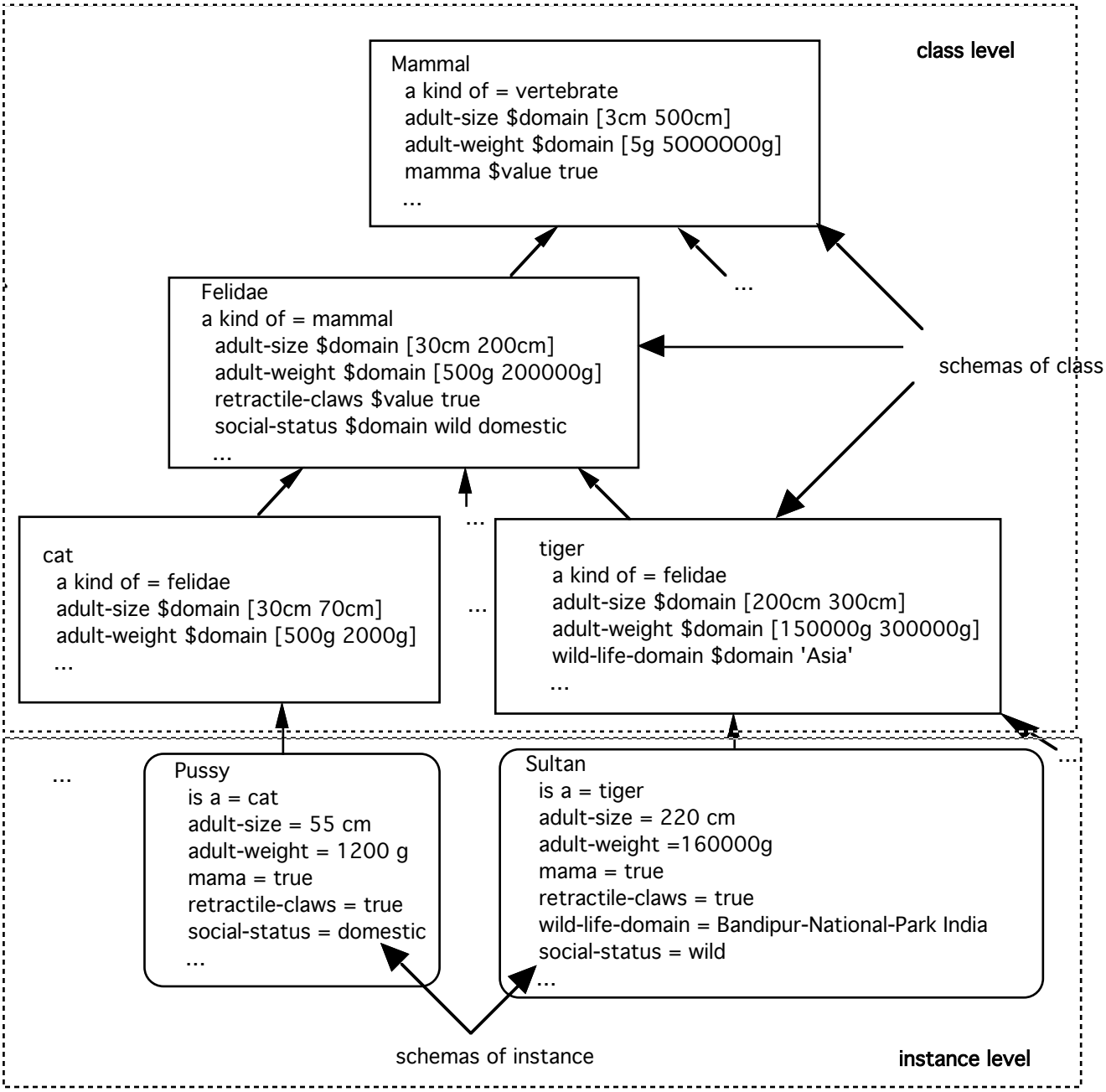


Fig 5 -This figure illustrates the object centered representation. Two levels are considered in this example. The level of classes where general objects are represented and organized in a hierarchy and the level of instances which represent an example or specimen of classes. Each schema at the class level denotes ensembles of objects which have the same properties (*i.e.* ensemble of tigers). At the level of instances particular elements are represented (*i.e.*, Sultan which is a tiger of a national wild life park). Each object is connected to some others which are more general in the hierarchy and inherits properties from more general ones (by arrows). For example, "tiger" and "cat "inherit from Felidae (*i.e.* Felidae is a more general object or concept), they both have retractile claws. A specialist might criticize this hierachy by observing that some Felidae, such as cheetah, do not have retractile craws. In fact this exception can be taken into account in such an object centered representation). Basically, schema formalism is used both for classes as well as for instances.

fact
rice is a crop
rule
<u>if</u> rice is a crop <u>then</u> rice is sensitive to pests

from this rule the *produced fact* is : rice is sensible to pests

This can be added to the facts base. We can note that more general rules can be defined, such as

if x is a crop then x is sensible to pests

or if crop(x) then pests_sensitive(x)

Variable x present in rules enable the representation of knowledge relative to sets of objects or situations. For example a set of crops having the same role as rice defined above can be defined in a data base or in a facts base.

Two basic exploitation (reasoning) mechanisms can be distinguish in such representations :

forward chaining or foretracking (or still data driven)

backward chaining or backtracking (or still goal driven)

They are presented in the following example with reference to the risk of pest infestation. Suppose the *rules base* is :

R1 : if	climate is warm		[Wa]
	climate is wet		[We]
	adult pests were present in the preceding year		[Ap(t-1)]
then	risk of development of a larval population		[Lp]
R2 : if	average temperature is greater than 20°C		[Tm]
then	climate is warm		
R3 : if	risk of development of a larval population		[Lp]
then	risk of development of adult population		[Ap(t)]
R4 : if	risk of development of adult population	[Ap(t)]	
	presence of crops		[Cr]
then	risk of infestation		[Ir]
R5 : if	relative humidity is greater than 90%		[Hu]
then	climate is wet		[We]
R6 : if	Adults of pest detected in the neighborhood		[Np]
then	risk of development of adult population		[Ap(t)]
R7 : if	risk of infestation		[Ir]
then	risk of crops destruction		[Cd]

Note that a more operational notation using simple symbols as defined above within brackets is practical :

R1 :	Wa & We & Ap(t-1) => Lp
R2 :	Tm => Wa
R3 :	Lp => Ap(t)
R4 :	Ap(t) & Cr => Ir
R5 :	Hu => We
R6 :	Np => Ap(t)
R7 :	Ir => Cd

Suppose the initial state of facts base are Tm, Hu, Ap(t-1). when the two basic reasoning processes are applied the differences between these processes are clearly demonstrated.

(i) *Forward chaining* or *foretracking* or still *data driven reasoning*. It consists of exploring the facts base and then activing every possible rule until no more rules can

be applied from the facts listed on the base. On this case, each rule is actived only once as otherwise redundant facts will be listed indefinitely).

Activated rule	state of the facts base
R2	Tm, Hu, Ap(t-1), Wa
R5	Tm, Hu, Ap(t-1), Wa, We
R1	Tm, Hu, Ap(t-1), Wa, We, Lp
R3	Tm, Hu, Ap(t-1), Wa, We, Lp, Ap(t)

From the initial state we can only deduce that there is a risk of development of an adult population. We can observe that an **explanation** of the reasoning process can be easily obtained from the list of activated rules.

(ii) *Backward chaining* or *backtracking* or still *goal driven reasoning*.

The question of a farmer is most probably : is there any risk of crop destruction ?

To reach the goal or "to demonstrate the theorem Cd", we use the rules base and the initial set of observations (e.g. Tm, Hu, Ap(t-1)). We also try to show if the wanted conclusion can be obtained from this information and knowledge and not by deducing all possible facts as was done earlier. So from the goal we can proceed from the goal and analyze the concluding parts of the rule to detect step by steps if the left hand side of the rule necessary to verify this final conclusion is in the fact base. In contrast to the previous mechanism, if a condition is not found, a program can ask the question to the user who can, or not, complete the list of verified facts. This kind of reasoning process is also called **abductive** reasoning, in opposition to the **deductive** one which characterizes the foretracking process. The reasoning mechanism can be schematized by a tree (cf. Fig. 6)

(iii) Inference engine.

Programs which analyze rules and facts bases are called inference engines. Some of them combine foretracking and backtracking mechanisms. The basic functions of inference engines are :

(1) selection of the rules to be applied : pattern-matching.

a problem may arise if more than one rule can be applied (conflict). A possible solution is to provide a rule for the rule to be applied (e.g. first rule in the base, degree of certitude in uncertain reasoning…)

(2) strategies (for backtracking) : depth or breadth of first search.

(3) structuring the rules base (meta-rules, context definitions)

(4) stop conditions :

. a particular fact (Foretracking)

. all possible rules were referenced (Foretracking),

. verification of all facts (Backtracking),

(5) Monotone and Non-Monotone reasoning.

. Monotone : inferred facts are definitives.

. Non-Monotone : inferred facts can be reconsidered, but this leads to problems in Truth Maintenance.

(6) Explanation of reasoning.

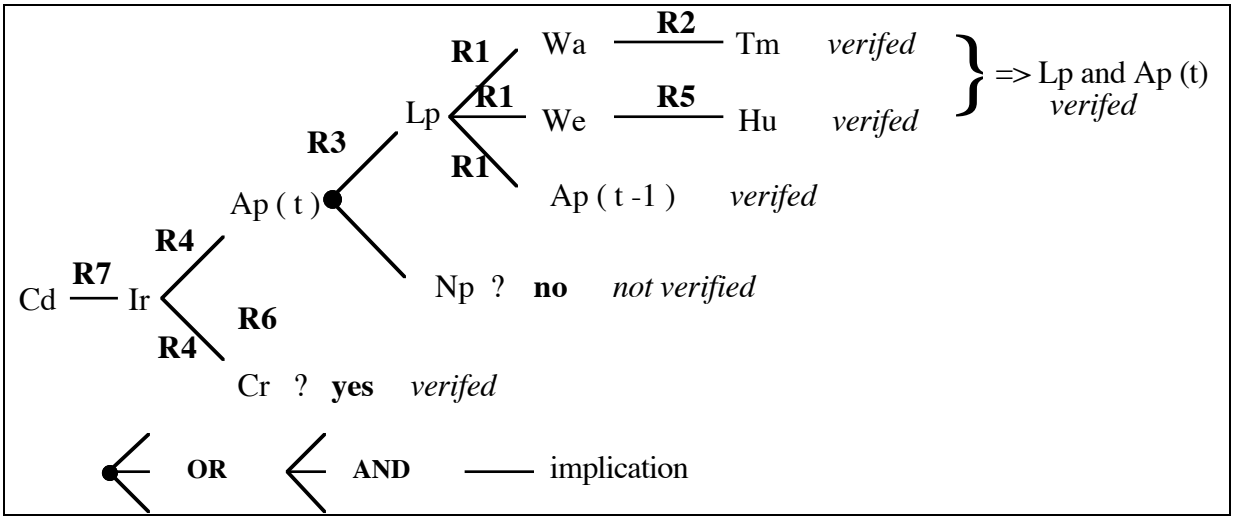


Figure 6 - Backtracking, or goal driven process : the program explores the rules from their concluding parts and tries to show if the proposed fact can be deduced from known information entered in the facts base. This example shows the interest of backtracking/foretracking : to demonstrate a proposition (e.g. Cd), the initial state of the facts base can be completed during the process by posing questions to the user (e.g. Cr ? means : are there crops ? The user answers yes, then this fact is considered as verified and enters the facts base).

b - Object-Centered Representation.

The origins can be found in the notions of frames introduced by Minsky (Minsky, 1975), of semantic networks, and of Object-Oriented Languages (e.g. Smalltalk). We present an adaptation of this representation used by a Knowledge Based system, SHIRKA, developed in France by F. Rechenmann (Rechenmann, 1988, Rechenmann & Uvietta, 1991). At first it was devoted to a specific application as mentioned above : to build a an aided modelling computer system within the framework of the EDORA project (Pavé & Rechenmann, 1986). But the general ideas implemented in the system envisage larger applications.

The general representation is the same as the one presented above, and the basic inference mechanisms working on this representation are principally :

.inheritance of slots (definition and values),

.default values (non-monotone reasoning),

.procedural attachment,

.pattern-matching,

.classification.

The organization of objects is hierarchical : classes, sub-classes and instances can be distinguished. We illustrate these points by an example from EDORA concerning mathematical models of growth. It is a simple example of a mathematical model base where a model can be obtained by pattern-matching following a biological interpretation :

(1) the growth can be limited by a growth factor (or not)

(2) this factor can be consumed by the biological system (or not),

(3) this factor can have a saturating effect on the growth rate (or not)

The assumed growth conditions are defined in a slot called *fact-cr*.

(i) for the **exponential model** : *fact-cr=fff* (the 3 conditions are false),

(ii) for the **logistic** one : *fact-cr=ttf* (the conditions (1) and (2) are true),

(iii) for the **Gompertz** model : *fact-cr=tff,*

(iv) for the model proposed by **Monod** : *fact-cr=ttt*.

The figure 7 presents the knowledge organization in the SHIRKA representation.

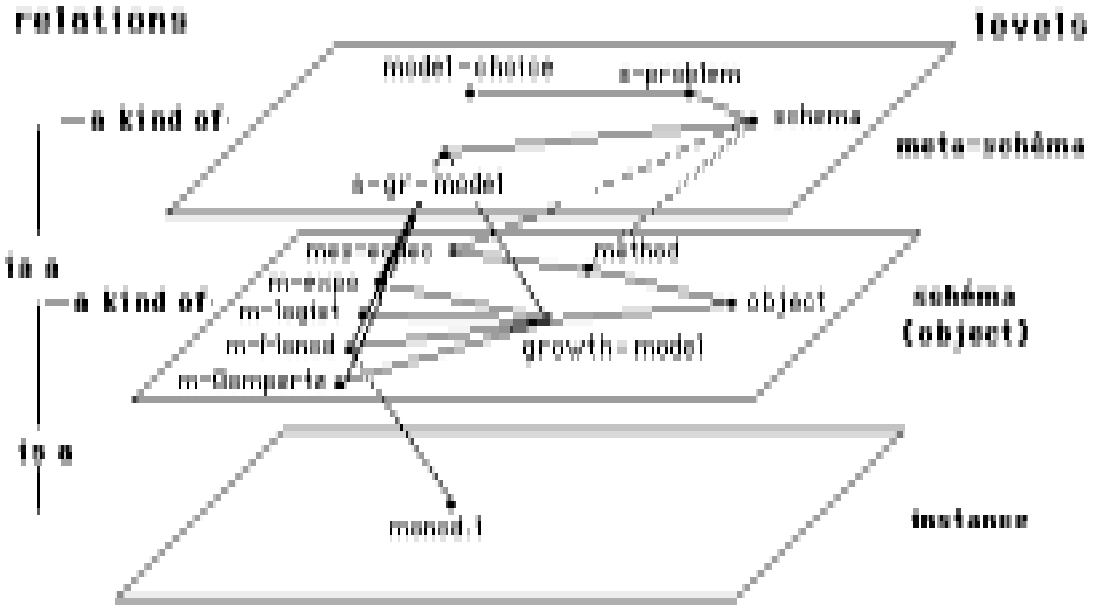


Figure 7. Organization of knowledge in SHIRKA : the simple example of growth models (Monod, Gompertz, logistic and exponential). Three levels can be distinguished, the level of instances (*i.e.* specimens of classes), of classes which describe objects in units called schemas. The meta level is necessary because **pattern-matching** can be applied on **instances of schemas**. When objects (at the object level) are wanted they can be obtained by pattern matching if they are instances of a meta-schema. As schemas at the object level can be considered as instances of schemas at the meta level (link is-a), a meta-schema corresponding to the problem to be solved must be constructed.

Here schemas corresponding to particular models are linked to the same meta-schema **s-gr-model** on which the pattern matching will be applied. This operation is defined in a schema called **s-problem.** At the object level, models are organized in a simple hierarchy where common slots and properties are stored in a generic schema : **growth-model.** The corresponding knowledge base has the structure described in Fig. 8.

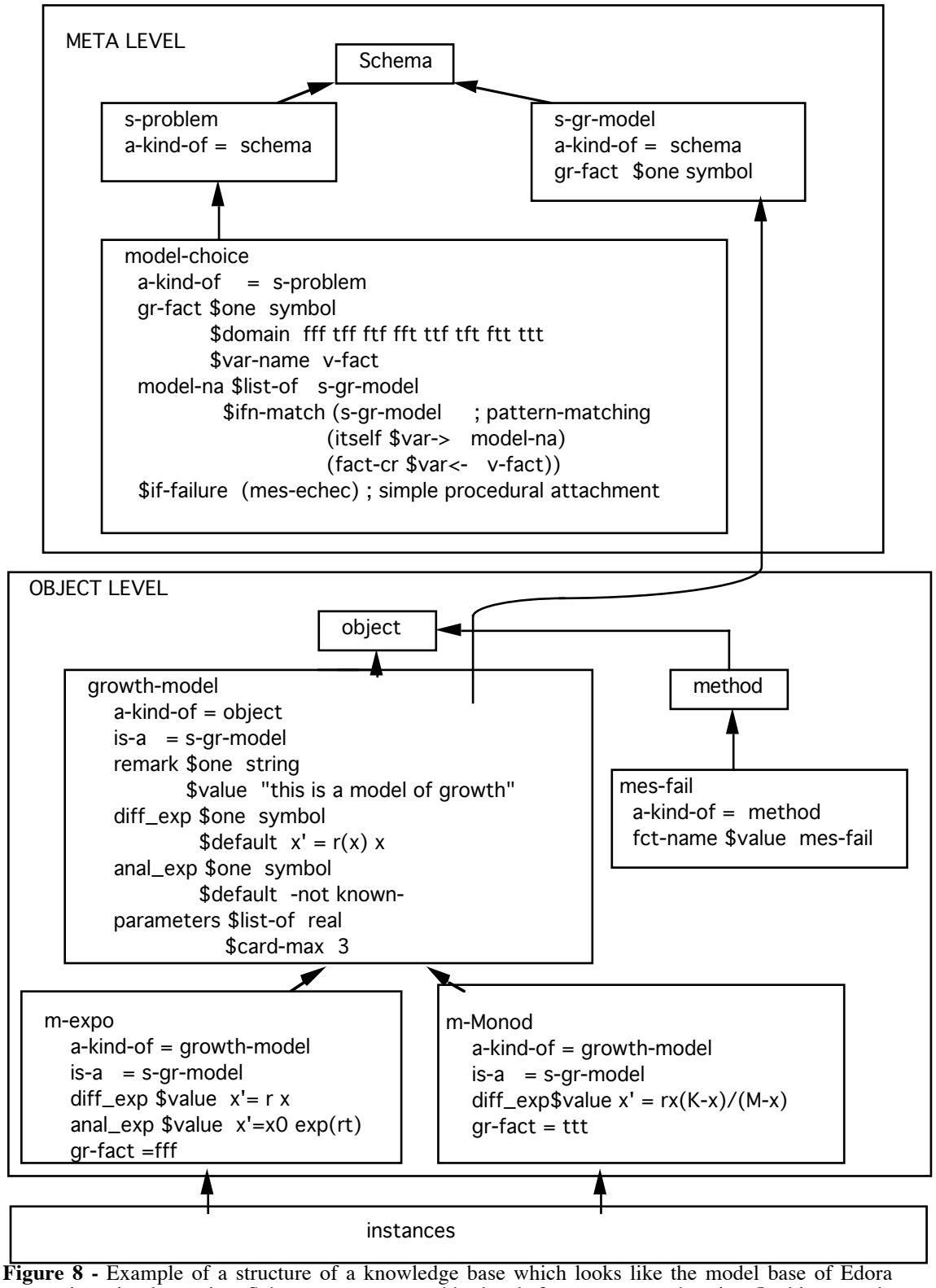


Figure 8 - Example of a structure of a knowledge base which looks like the model base of Edora system in a simple version. Schemas are presented by levels for easy comprehension. In this example basic inference mechanisms can be illustrated, besides classification which will be presented later.

Remark : mes-fail is a Lisp function which prints " I do not know any model adapted to these conditions".

Examples of utilization of the base given below illustrate the basic inference mechanisms. It must be remembered that the fundamental goal is to find a model which satisfies a biological description of growth.

(1) Example of instanciation of the schema m-Monod (the prompts and questions returned by SHIRKA are in italics). Remarks about commands and mechanisms are in small characters and preceded by Rem.

Shirka : cr-inst	Rem : command which enable the creation of an instance.
Class : m-Monod	Rem : name of the class schema to which the instance will be attached.
instance name : monod-1	Rem : name of instance.
anal-exp? -	Rem : - means no answer because it is unknown.
param? 0.1	Rem : values of parameters are given by the user.
param? 65.0	
param? 3.0	
-> monod-1	Rem : if instanciation is successful, then SHIRKA returns its name
Shirka : vi monod-1	Rem : command which permits the visualization of an instance
monod-1	
remark : "this is a model of growth"	Rem : this slot is inherited from growth-model schema
diff_exp : x' = rx(K-x)/(M-x)	Rem : this slot is inherited from m-Monod schema
anal_exp : -not known-	Rem : this slot and the default value come from growth-model schema
param : 0.1, 65.0, 3.0	Rem : values of parameters characteristic of this instance

(2) To search a model, an instance of the schema : model-choice (schema which describes the general problem of model choice from a biological description), obviously the user does not give any value to the slot model-na because it is not known (if it was known there would be no problem !). Finally the pattern-matching is activated when the user asks the system the value of this slot. The name of the model appears if the description corresponds to a known model (example 1), if not it prints a message which specifies that the model is unknown (example 2, where only answers are denoted).

Shirka : cr-inst	|	*Shirka* : cr-inst
Class : model-choice	|	*Class* : model-choice
instance name : choice-1	|	*instance name* : choice-2
gr-fact ? fff	|	*gr-fact ?* fft
model-na ? -	|	*model-na ? -*
-> choice-1	|	*-> choice-2*
Shirka : val? choice-1 model-na	|	*Shirka* : val? choice-2 model-na
-> m-expo	|	I do not know any model
	|	adapted to these conditions
	|	*-> search failure*
Example 1 : KNOWN MODEL		Example 2 : UNKNOWN MODEL

1.5 Sophisticated reasoning mechanisms - some problems in knowledge organization and processing *:*

At present there are many efficient mechanisms proposed in knowledge based systems, but there are also a lot of problems to be solved. So, among the sophisticated reasoning mechanisms and current problems, we can retain :

- reasoning explanation,
- reasoning involving time,
- reasoning involving space,
- qualitative reasoning,
- classification,
- points of view, contexts,
- learning,
- shallow and deep knowledge,
- construction and verification of knowledge bases

All these points are commented on briefly.

1.5.1. Reasoning explanation

The goal is to give an explanation of a proposed result, *i.e.* answers to questions such as "why … ?", "how… ?" or "why not…?"

Why…? : to justify, in backward chaining, a question asked by the system. In our simple example of risk on crops, to demonstrate the goal Cd (risk of crops destructions) the system could ask the question : are crops present ? The explanation of this question is simple and quite stupid : to have such risks crops are obviously a necessary condition. This kind of bad questioning can be avoided by adding a supplementary rule to the base which links "risk of crop destruction" with "presence of crops". Such difficulties are frequent and so a knowledge base must be constructed such that "intelligent, or at least no stupid, questions are asked the user".

However , still in the same example, Ap(t) : "risk of development of adult pest population" can be verified simply by answering the question "are there adult pests in the neighbourhood ?" which is not as stupid as the previous one… Then an explanation or suggestions can be proposed, such as: "for this pest an infestation may come by immigration from neighbouring infested patches" (a property which is not shared by all pests).

How…? : to explain a reasoning process (the most often a trace of activated rules).

Why not…? the negative explanation is very interesting, and often even more than a positive one (a good example is proposed in the classification mechanism in SHIRKA).

1.5.2. Reasoning involving time

Time is an important dimension to be considered in many applications. When events occur in an arranged (temporal) series, or when time appears implicitly or explicitly, as for example in rule R1 where time appears explicitly :

R1 :	<u>if</u>	climate is warm	[Wa]
		climate is wet	[We]
		adult pests were present the previous year	[Ap(t-1)]

then risk of development of a larval population [Lp(t)].

However, a special treatment of this variable is not needed here, at least in the first approach.

In other cases, this variable can be taken into account by **procedural approaches** (simulation models), as in COMAX/GOSSYM system where the simulation model is the central part of the expert system (cf. § 2.1.4.). But in some applications this approach is not necessary and sometimes is also not possible (too complex or if only qualitative knowledge is available…).

(1) Procedural approach : a simulation model is linked to an expert module which can help the user in the execution of the relevant procedure and in the analysis of simulation results (cf. COMAX/GOSSYM *op.cit*). However this approach is basically quantitative and needs accurate and precise information to obtain a "good" model (i.e. a validated model which simulates the reality well).

(2). Non procedural approach (for example, Charniak and McDermott, 1985) :

(i) the situation calculus is an attempt at a qualitative approach. This theory introduces the term **situation** : time interval over which a state does not changes its truth value (these intervals may be infinitesimal). Other concepts are also defined such as events, before, next…

(ii) TSA : Time System Analyzer : its goal is to predict what will happen, and explain what has happened. Such systems must contain laws of motion of systems.

(iii) TMM : Time Map Management. "Time Map" refers to a permanent data base of state and event tokens managed by a Time Map manager which establishes temporal relations.

The following questions are examples for which such a system is interesting:

- Did you plant corn before or after wheat ?
- Which forest did you visit most recently ?

These approaches can be envisaged themselves or linked with a simulation module.

1.5.3 Reasoning involving space

As in the preceding case, in this example we can also find a rule where space appears implicitly :

R6 : if Adult pests detected in the **neighborhood** [Np]
then risk of development of adult population [Ap(t)]

In fact the same remark is always true. In this example we do not need a specific treatment for space.

In other cases, two complementary ways can always be considered : procedural (most often quantitative) and knowledge based (most often qualitative).

The procedural approach needs a simulation model with space variable. Some simulation models, or general procedures, are linked to expert modules to help the user in the execution (cf. for example, FIDES a Fuzzy Intelligent (partial) Differential Equation Solver which constructs a finite element net for 2-D Elliptic PDE (Friedmann *et al*, 1987)), and analysis of the results. However, such approaches do not permit the handling of complex spatial situations. Problems in "Route Finding" can also be solved by procedural approaches (displacement within a graph structure).

In fact, spatial reasoning tries to imitate the good spatial apprehension of man which is very efficient to analyze topological and geometrical situations in complex (e.g. natural) structures. It consists " not only in making inferences from knowledge about space description, but also in infering new spatial knowledge". It contains descriptions of space, and laws of physical space (e.g. topography, presence of a field of gravity and circulating fluids: water, air, oil…, evaporation laws…) and some specific reasoning processes which consider properties, for example connexity in a landscape (i.e. connecting ways between same ecological units)…

Some examples are :

- ELFIN an expert system on oil circulation in soil is based on geological knowledge. From a knowledge of the geological structures, ELFIN tries to infer the migration of oil and then to detect the location of natural oil tanks (Martin-Clouaire, 1984).

- Analysis of avalanche sites: from the structure of an avalanche site the problem is to determine the risk of avalanche from the quality and quantity of snow and the meteorological conditions (temperature, direction of wind, humidity of air,…) (Buisson, 1987, 1991).

- Neurologist II : diagnosis of nervous system diseases using a geometrical description of the nervous system (Xiang and Srihari, 1985).

Combining time and spatial reasoning leads to **spatio-temporal** reasoning processes.

1.5.4. Qualitative reasoning

This type of reasoning leads to qualitatives results such as : " if a pest infests a crop, at first the vegetables will be degraded, but if an appropriate treatment is applied and if it is a success, then the crops can be saved", where no indications are given about quantitative aspects (intensity of treatment…). Other good examples can be found in "naive physics" [Kuipers,1982, 1986, McCloskey, 1983, Cross, 1983]. These approaches in A.I. try to simulate "common sense reasoning". Guerin (1990) has proposed a process to control a waste water treatment (based on an artificial lagoon) by using such a qualitative approach.

This is certainly an exciting way of research, one of the principal reasons being the nature of experimental information and results of observations which are often qualitative. In a next future we can envisage not only reconsidering some aspects of qualitative data analysis but also producing qualitative simulators or hybrids (symbolic and numerical). This problem will be dealt with in Chapter 2.1.5.).

1.5.6. Classification

Essentially applied to object centered representation, consists of finding objects in a classification which verifies some criteria or properties. For example, to find an animal in a zoological classification or a plant in a botanical classification from descriptions of morphological characters and other features (an example will be presented briefly in Chapter 2.4.1.). In fact when such problems are studied, it seems that different classifications are possible, from a determination key, oriented on practical identification of specimens, to phenetic or evolutionary or even a phylogenetic one. That is the objects of classification can be viewed from different angles and in different contexts and the corresponding descriptions can be useful.

1.5.7. Points of view and contexts

As mentioned above, an object can viewed from different angles or in different contexts. For example, an animal (or a plant) can be seen not only as an element of one or more taxonomical system (i.e. a species) but also as an element of an ecosystem (individuals in a population), or of a production system (a crop, a stock farm in an agrosystem), or even an economical point of view (wheat production in a country…). The management of knowledge bases including different points of view can be envisaged, particularly in an OCR (for example, the notion of family in SHIRKA is a first attempt in this direction, (Rechenmann & Uvietta, 1991)).

1.5.8. Learning

Today the knowledge in a Knowledge Based System is mostly given by an expert in a representation which is closed to internal representation and sometimes interactive aids are also furnished (e.g.TEIRESIAS which gives and aid to knowledge transfert from medical experts to the expert system MYCIN (Davis, 1976)). Major effort has to be made to enable direct learning by machines. A lot of work has been done in this direction particularly for pattern recognition, for example recognition of characters by some systems. The first attempt was probably the Rosenblatt's perceptron. There are many ways of studying this problem. This probably, and more basically, reflect the different processes involved in "natural" learning, because sudies on learning lead more or less to models of assumed "natural" processes. We mention only some examples :

- learning "by heart", information is *a priori* furnished to avoid treatment or computing (e.g. checkers game program developed by Samuel in 1959),

- learning by teaching (e.g. TEIRESIAS, *op. cit*)

- learning by analogy (or analogic reasoning). A typical example can be found in Evan's work were the program tries to answer classical "logic tests". For example, figures A and B are presented, where B is derived from A. A figure C is given and the next one as to be found in a set of alternative results D_1, D_2, D_3, such that the rule applied to transform C in D_i is the same as the one applied to transform A in B (cf. Figure 9).

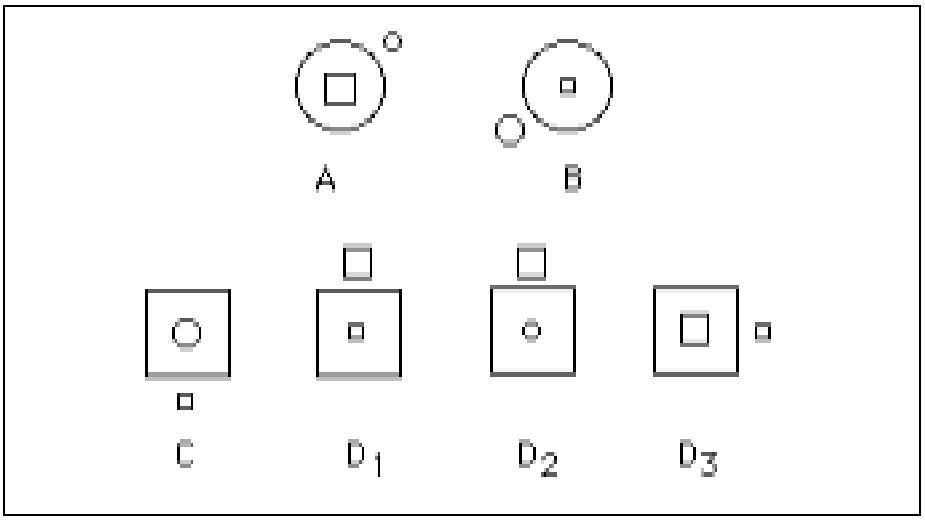


Figure 9 - Example of a problem in analogic learning or reasoning (from Evan's works on learning). The program has to infer the rule which enables the passage from A to B and apply it to find the right one in the set of D_i , that is the figure which results from the application of the inferred rule on C.

- learning by discovery : the system has to discover its surroundings. For example the programs AM and EURISKO written by Lenat (Lenat, 1982), which are oriented to dicover elementary mathematics results. AM is a frame based system : from a base of 115 concepts coded by frames and of 243 heuristic rules it can generate new concepts. An experiment showed that AM can generate interesting concepts : from 200 concepts, about 100 have been considered as reasonable by mathematicians and in this set some subtle concepts in fundamental arithmetics (such as the theory of divisibility, Golbach's conjecture…).

- learning by examples consists of generating general rules from two sets : one of examples of the concept to be learnt and the other of counter examples (Figure 10).

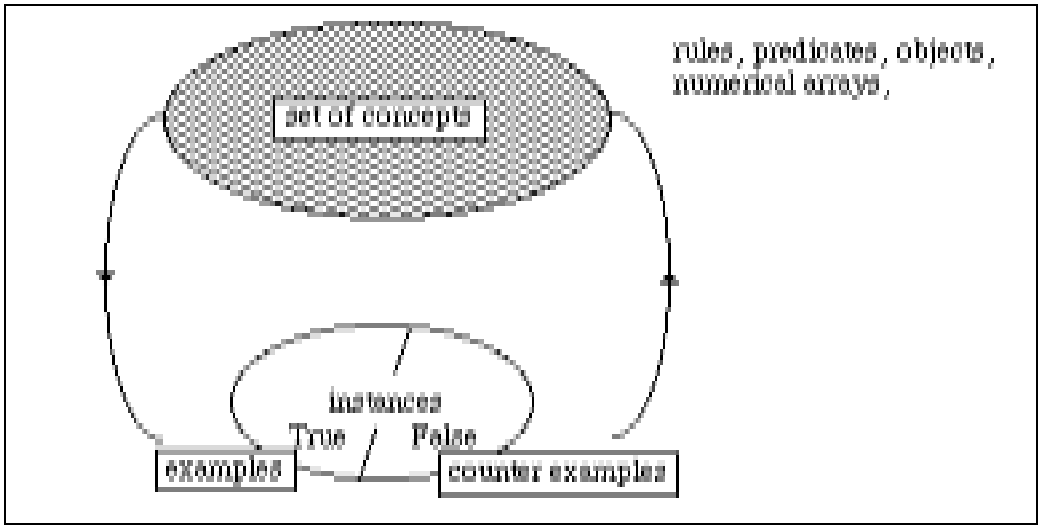


Figure 10 - Learning by examples and counter examples. From a set of examples and counter examples the system tries to find a set of concepts which characterizes the data. From the set of instances an iterative process enables us to verify if the referred concepts are valid.

Three main ways are often distinguished :

(i) *Connexionist paradigm*, where instances are coded numerically. The rules or concepts are coded as weighting tables (the learning consists of building these tables). Calculus formula or decision tables use this information. The best known

examples are : the perceptron (Rosenblatt), the hyperspheres or cluster analysis (Batchelor) and the Boltzmann's machine (which is a network of probalistic automats). In fact these methods are more or less connected to multivariate analysis (it is evident for cluster analysis), and one may ask question about the efficiency of these approaches over satistical ones.

(ii) *Stucturalist paradigm*, where the coding of rules and examples is similar and symbolic (e.g., predicates, rules, classes, objects…). Learning consists of infering the most general rules or objects by a repeated process : from an initial set of examples and counter examples a first "low level" set of rules (or objects) are infered. This set of rules (or objects) is then analyzed and a second level set is inferred, and so on… till there is more possible generalization (Figure 11).

This type of learning has been deeply studied, and a lot af algorithms are available (for example the algorithms of Mitchell, Winston, A* of Michalski & Chilausky…).

(iii) *Genetic paradigm*, where algorithms are constructed by analogy to the natural evolution of the living world, by applying genetic operators (mutation, inversions, crossing-overs…) and the selective values of rules are modified (i.e. their aptitude to generate "good concepts").

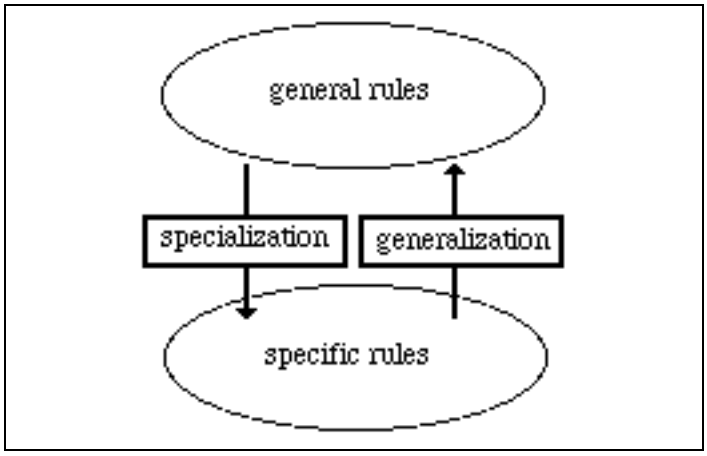


Figure 11 - Learning by examples. An initial set of examples and counter examples is furnished to the program. At first a low level set of rules is inferred and from this set a second set of higher level is inferred, etc…, till generalization is no more possible. The program works by applying repeatedly generalization and specialization procedures, which are closed to inference and deductive processes.

The problem of learning by examples is now envisaged in Knowledge Based Systems environments. For example, in Shirka environment, a work was developed by J.L. Aguire Cervantes (Aguire Cervantes, 1989) for automatic learning from examples by generalization-specialization, within the framework of the structuralist approach : class are defined from instances which are known as examples or counter-examples, and an attempt is made at an automatic construction of a hierarchy of classes (cf. Figure 12).

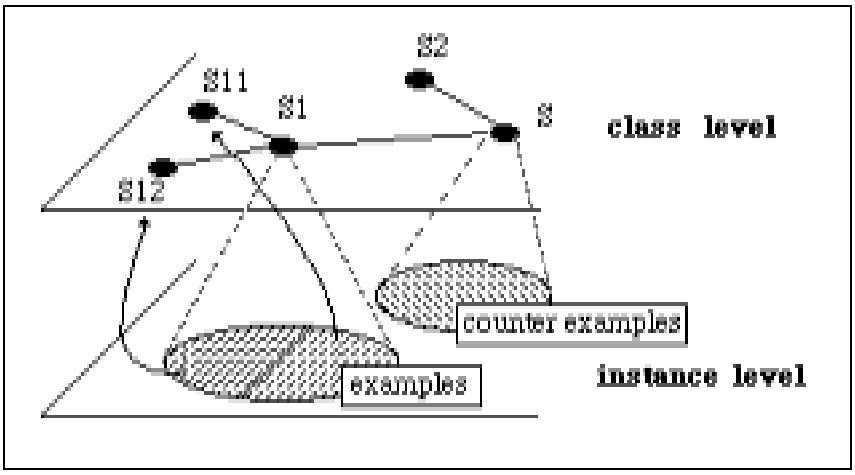


Figure 12 - Learning in a frame based system. Classes S1 and S2 are inferred from instances. A more generic one S , and also more specific ones (e.g. S11 and S12) , can be deduced in some cases, which could lead to the construction of a hierachy. Such approaches could be very interesting in symbolic taxonomy.

1.5.8. Shallow and Deep Knowledge

Shallow knowledge corresponds to observed phenomena (similar to the notion of blackbox in Cybernetics) and deep knowledge corresponds to the explanation level. When we open the blackbox to see smaller boxes which "explain" the behaviour of the big black box.

For example, the growth of some vertebrates (e.g. the musket rat) is well described in Gompertz's model (phenomenological level) :

$$\frac{dx}{dt} = a\,x\,(\ln(K) - \ln(x)) \qquad [1]$$

where x is a morphological variable (weight, total size…).

By introducing the supplementary variable $f = \ln(K) - \ln(x)$, this equation is equivalent to the system :

$$\begin{aligned} \frac{dx}{dt} &= a\,x\,f \\ \frac{df}{dt} &= -\,a\,f \end{aligned} \qquad [2]$$

which can be obtained from the following pseudo-chemical reactions :

$$x + f \rightarrow 2\,x + f \qquad \text{(i)}$$

$$f \rightarrow \ldots \qquad \text{(ii)}$$

So, the Gompertz's can be explained by two more "elementary" processes (Pavé et al, 1986) :

(i) growth controlled by a growth factor f,

(ii) simultaneous degradation of the growth factor .

The description by Gompertz's model [1] may be considered to correspond to the phenomenological level (shallow knowledge) and the expression [2] associated with the set of pseudo-reactions as corresponding to the deep knowledge (i.e. a decomposition in simpler, or "elementary", processes). Obviously, the notions of

shallow and deep knowledge are relative. On the same kind of example some authors consider that the shallow level corresponds only to the graphical description of data while the deep knowledge includes mathematical descriptions (e.g. Pierret-Golbreich, 1988).

Another example is provided by MYCIN (as noted by Cross, 1983). This expert system contains rules such as "if the patient is less than eight years old then do not administer tetracycline" (shallow knowledge) but does not represent the causal (or deep) knowledge that tetracycline can impair bone development in children. So if a physician asks why, MYCIN cannot answer.

It is important to establish links between the phenomenological level and the explanatory one and to enable reasoning at these two levels which would provide more satisfactory explanations and reasoning. Considering these relations would also be interesting for the knowledge on a particular subject itself. For instance, we have made a particular effort to analyze growth models in terms of implied processes when we were interested by a knowledge base construction concerning these models (Houllier, 1988, Pavé, 1988…).

1.5.9. Construction of Knowledge Bases

One of the biggest problems in Knowledge Based Systems is to establish the knowledge base itself. As already noted similar problems are encountered in mathematical **modelling**. Many problems can be identified , both at the technical and methodological levels, for which we have not general answers :

- organization (even in rule base, theoretically "not organized" efficiency leads to a structure of the base by using meta-rules). This organization is fundamental in OC Representations, particularly to manage different points of view.
- completion (is the base complete ?),
- rules which could lead to circular reasoning,
- redundancy (not very important but sometimes necessary),
- inconsistency (i.e. a fact and its negation obtained from the same premises in the base),
- "subtle" organization such as "intelligent" questions are asked of the user,
- validation : does the knowledge base permit an answer to the objective of its construction ?

2 - APPLICATIONS

Problems and examples connected to **Natural Resources Management** are mainly examined. That is, in sectors of applications (such as agronomy, fisheries, irrigation) and also in domains of a more general or fundamental interest : taxonomy, management of Geographical Information, simulation of ecological systems including sophisticated hypothesis on behaviours, etc… In **Biotechnology**, a general scheme of possible applications can be drawn, such as the aid in DNA sequences analysis, design of experiments, some industrial applications in process control and the use of data bases.

Finally some results are presented, concerning more or less both domains and of methodological approaches such as the aid to choose statistical methods, to construct a mathematical model, etc,… The degree of independance between such softwares and domains of utilization (e.g. what is the part of "context free" knowledge in a system devoted to mathematical modelling aid) are discussed. It is important to note that the major part of our references come more from research and systems under development rather than effective realizations. A.I. approaches are so interesting because of the new framework of modelling it gives even if no operating systems is buult.

2.1. Knowledge based systems in Natural Resources Management.

To day the two main goals of such systems are the **recognition** of typical patterns and aiding in **prediction** or in **decision making**.

2.1.1. Recognition

Recognition mainly concerns diagnosis (e.g. in plant pathology to find plant disease, in ecology to give the sensitivity of an ecological system and associated risks following probable evolution), taxonomy (e.g. to determine the species of an organism by intelligently exploring a determination key or an adapted representation in terms of knowledge base), localization (e.g. to find a natural oil tank in petroleum research such as in ELFIN (Martin Clouaire, *op. cit.*)).

2.1.2. Prediction and decision making

These problems can follow the recognition of a situation by defining the "best" action in the defined context, e.g. therapy and strategies following a diagnosis. For example, PLANTER (Smith *op.cit.*), gives planting recommandations from data about the history of a plot : previous crops, chemical or biological treatments… A neibouring class of problems concerns the analysis of strategies, or the control complex systems by projection in time, by using simulators. A.I. intervention is at the levels of model construction and result interpretation. The COMAX/GOSSYM set for cotton crop management is a good and elaborated example.

2.1.3. Limits of present approaches

Today the major part of A.I. approaches to Knowledge Based Systems (essentially Expert Systems), is based on production rules representations. So the limits discussed are essentially related to this type of knowledge modelling. However some topics are more general and emphasize the lack of human knowledge about knowledge itself :

- limitation in knowledge representation : as already mentioned we use essentially rule based systems. However there are some efficient systems which manipulate other kinds of representation (essentially Object Centered ones) or mixed formalisms (most often rules and frames) ;

- limited reasoning mechanisms. This question has already been discussed of and some developments presented (cf. 1.5.) ;

- difficulties in expressing a problem and formalizing it, this is also true for tactics and strategies ;

- relationships between shallow (i.e. analogies, associations, correlations) and deep knowledge (i.e. causal relations, explanation mechanisms…) ; for example rules are well adapted to represent shallow knowledge but not for deep ones or relationships between these two levels ;

- to take into account procedural and declarative knowledge within a same framework, i.e. objects and methods described following a similar representation ;

- management of points of view, of contexts, relationships between knowledge bases, and between knowledge bases and data bases. For example, in taxonomy for solving determination problems it often seems more convenient to contruct an analog of a identification key than to use phylogenetic classification. However, it is obviously interesting to have such classification and some other informations about objects in the same system which is devoted to identification (e.g. information about the biology and ecology of a recognized species (cf. Figure 13)).

Coulson [Coulson, 1987] speaks of Integrated Expert Systems or more generally of Integrated Knowledge Based Systems (I.K.B.S.) for the new generation of K.B.S. where solutions to these problems will be given (see next section and Figure 15). This idea is closed to the notion of Multi-Expert-Systems, or Multifacetted- Expert-Systems (following ideas of Zeigler in mathematical modelling (see Kerkhoff & Vansteenkiste, 1984)). Another example concerning irrigation problems in wet tropical zones is briefly described in Figure 14.

2.1.4. Some other examples in Natural Resources Management

a. Agronomy - Agriculture.

Advances in Knowledge Based Systems and Expert Systems will certainly be important in this area. Among present day applications, which are not still numerous, I have selected three which are characteristic of different approaches and problems to be solved : diagnosis of plant diseases, agronomical assistance for crop rotations, and crop management.

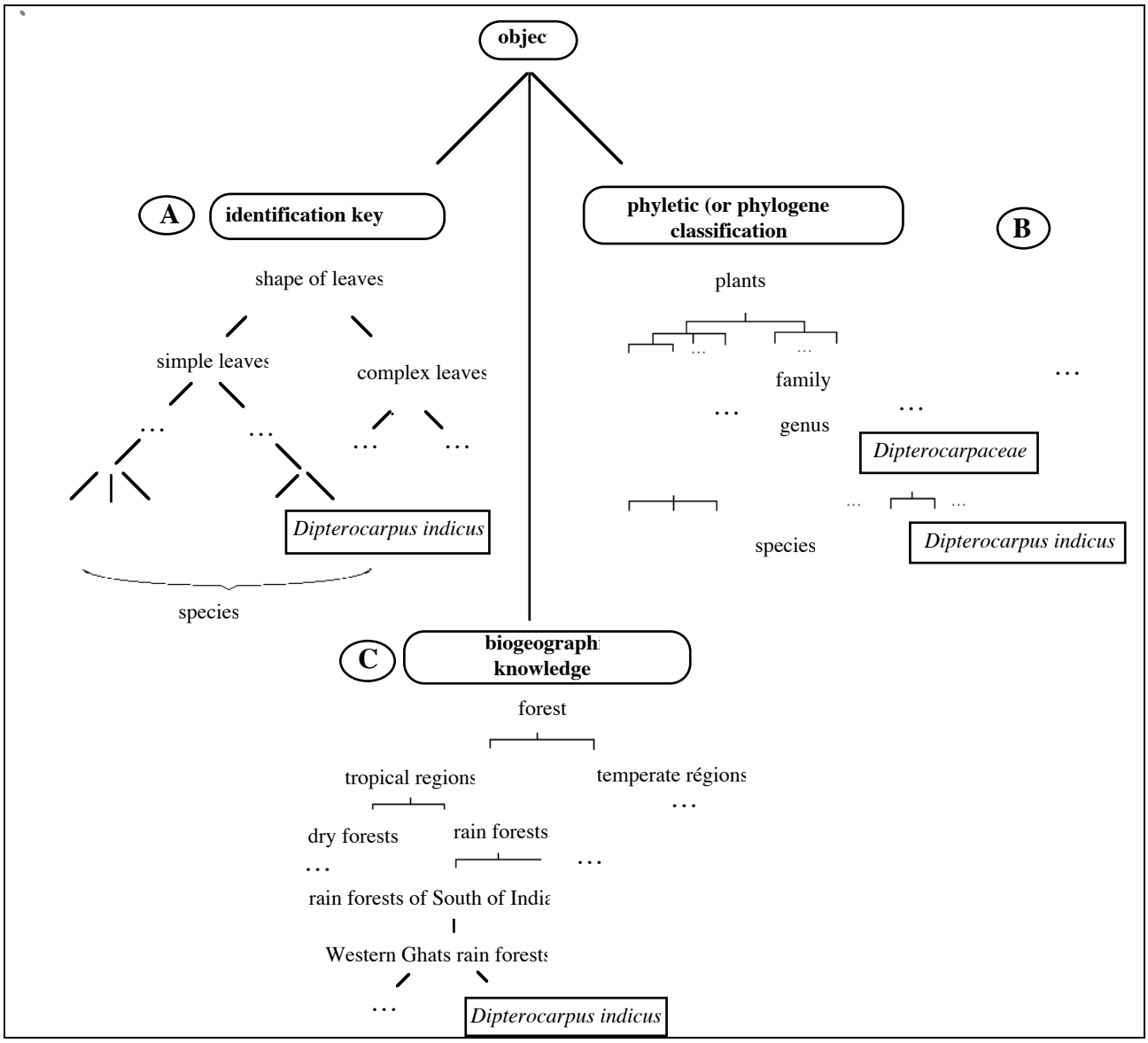


Figure 13 - An example of knowledge bases including different points of view about tree species : (A) is based on an identification key, (B) is a phylogenetic classification and (C) is related to biogeographical knowledge aof the species (species may be values of a slot or a specialization of the object "rain forest of Western Ghats", in the sense that these forests contain these species). So this figure shows that the same object (e.g. *Dipterocarpus indicus*) can be viewed as an element of differents bases from a particular point of view (or context). Obviously we need systems which enable such representations and to manipulate them. It would be also very intersting to connect them to data bases (for example to manage additional information about objects such as pictures, maps,...). For this purpose it seems that, to day, object centered representations would be a good solution.

- TOM and SEPV a general diagnosis system for plant diseases.

Developed by the French Institute For Agronomical Research (INRA), TOM is a diagnostic system of tomato diseases, adapted by a French Society (Cognitec™). This prototype was considered to be an accessible and useful tool not only for agronomists but also for farmers. About 30 analog systems have now been developed for a large variety of plants of agronomic interest. This set of system is called SEPV : "Système Expert en Pathologie Végétale" (Andro, Bachacou et al., 1985). They are rule based systems whose common inference engine is a rewritten

version of EMYCIN (Empty MYCIN, the inference engine of the well known system MYCIN).

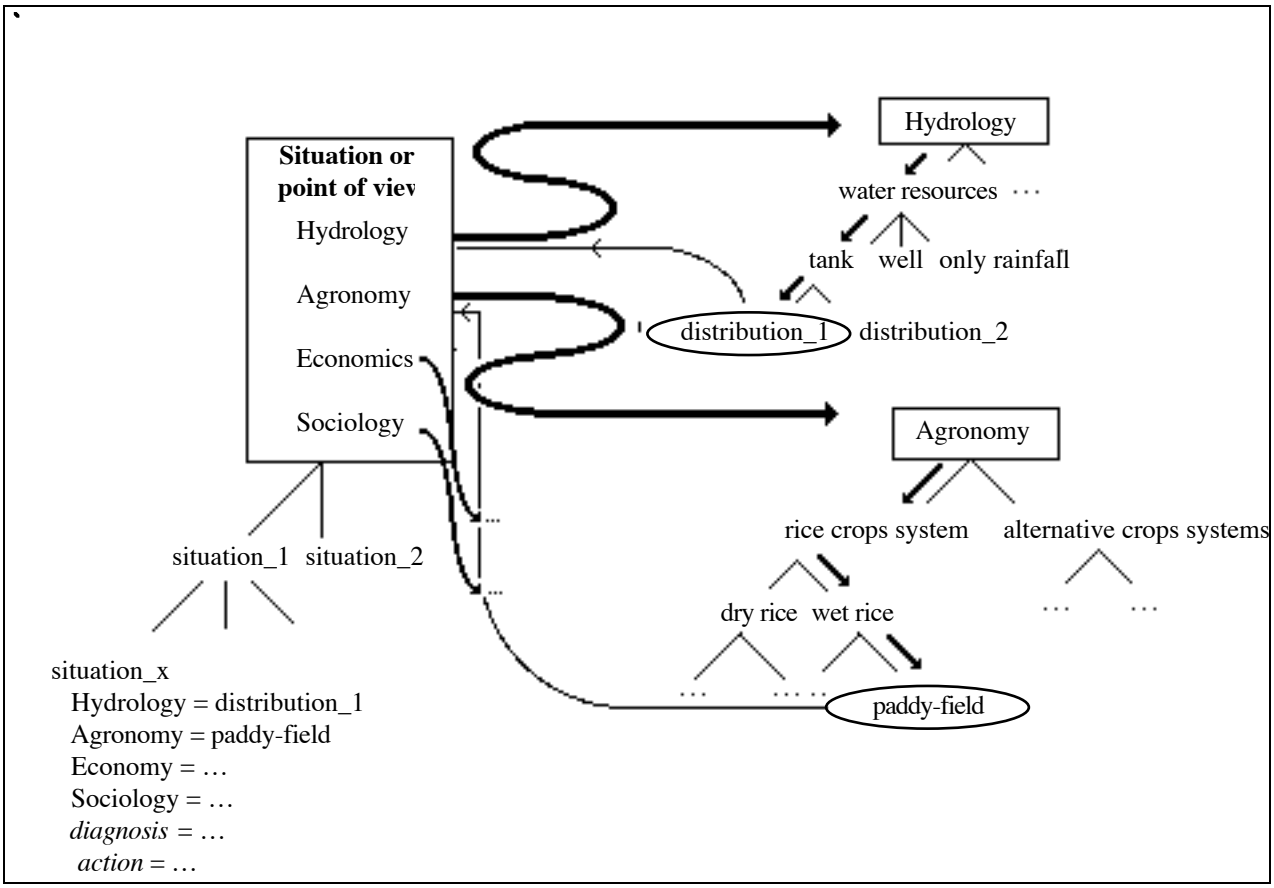


Figure 14 - To evaluate an irrigation system experts in various domains have to be consulted because it concerns not only technological aspects (e.g. Hydrology, Agronomy) but also has economic and social implications (e.g. Economics, Sociology). In fact, every time an irrigation system is examined all the experts are needed. This figure shows an example of a multi-expert system organization which enables irrigation system diagnosis. Once again it is based on an object centered representation (the study and feasability of such a system is well discussed in Oswald, 1991).

- PLANTER (*op. cit.*) : is a system which gives recommendations for selecting among crop planting and re-planting options from the history of a field : previous cultivations, treatments (particularly herbicides)… The Expert System is a complement of a Decision Support System for strategic studies and definition in agriculture (at the level of a farm). This system is a rule based system and the inference engine is Personal Consultant™ , which is also an adaptation of EMYCIN.

- COMAX/GOSSYM and MOVDEX Systems (McKinion *et al.*, 1987) : GOSSYM (GOS sypium SYMulator) is a simulation model for cotton crop management. It is a process level model which simulates the plant physiological, micrometeorological and physical processes in soil. In fact it can be considered as a procedural representation of deep knowledge on this agronomical system. COMAX (CrOp MAnagement eXpert) is an intelligent level which facilitates the use of GOSSYM for input data and also for interpretation of results of simulations. These joint systems enable us to take decisions about preplanting, planting date, row spacing, Nitrogen input, irrigation…

GOSSYM is written in FORTRAN 77 (about 3000 lines) while COMAX is written in a dialect of Common Lisp (about 6600 lines).

MODVEX provides an aid to model validation. In fact GOSSYM is a general framework which must be adapted to particular situations (parameter values and also modifications of some routines,which represent some particular submodels). As GOSSYM is used to aid cultivation control, it should have very good properties and so it must be validated. However GOSSYM is a big program difficult to adapt and modify. An expert system, MODVEX (MOdel Development and Validation EXpert), can therefore be used for model adaptation and development.

This system uses ART™ (Automated Reasoning Tool), an environment for Expert Systems development (also a rule based system).

Other examples of applications of artificial intellligence in agronomy and agriculture are given in a table at the end of the text (Annex).

b. Sea and Fishing

Catcur V1 (McGowan, 1987) is a fishery management expert system module which permits the analysis of catch curves. A catch curve (log(catch) vs age) of a fish species catched in a fishery can be examined for empirical interpretation of age of recruitment, mortality rate and apparent changes in fishing efforts and mortality with time and age. A microcomputer expert system was developed which asks for data, or obtains it from a data base and then uses a rule based knowledge base to calculate mortality, to describe possible confounding trends and to judge the confidence in the response. If the data is inadequate for a sound conclusion, the system advises the improvements that should be attempted. This system was foreseen as a module in a future comprehensive fishery management expert system.

c. Intelligent Geographical Information Systems

Important tools in Natural Resources management are thematic maps. However their classical presentations (paper) make it difficult to manipulate, compare and to consider simultaneously and finally to modify. If only the last problem is considered, we have to manage different time scales linked to the nature of represented objects (for example, geological data are more persistent than vegetal ones) and of the evolution of knowledge itself (even for geological data, as new facts are periodically taken into account, then maps have to be modified). The edition of maps by classical methods is long and so modifications cannot be made rapidly. Sometimes 3D representations whould be better than planar ones but have to be deduced from particular data, for example in landscape management from topographic data (*i.e.* dicretized isoclines). Such transformations are possible and 3D generation of a landscape is possible from these data by using classical interpolating methods. Finally the classical method limits the efficient use of geographical data and its development. So some systems have been proposed on data base structures to facilitate geographical data handling (*e.g.* ACT-Info). Now it seems that some extensions can be envisaged, including dynamics approaches via quantitative and qualitative simulation models and intelligent modules for handling declarative data and interfaces with users (*e.g.* to facilitate both the practical use of the systems and modifications of data bases, for example to aid in data validation before inclusion in the base…). Figure 15 shows the possible structure of such a system.

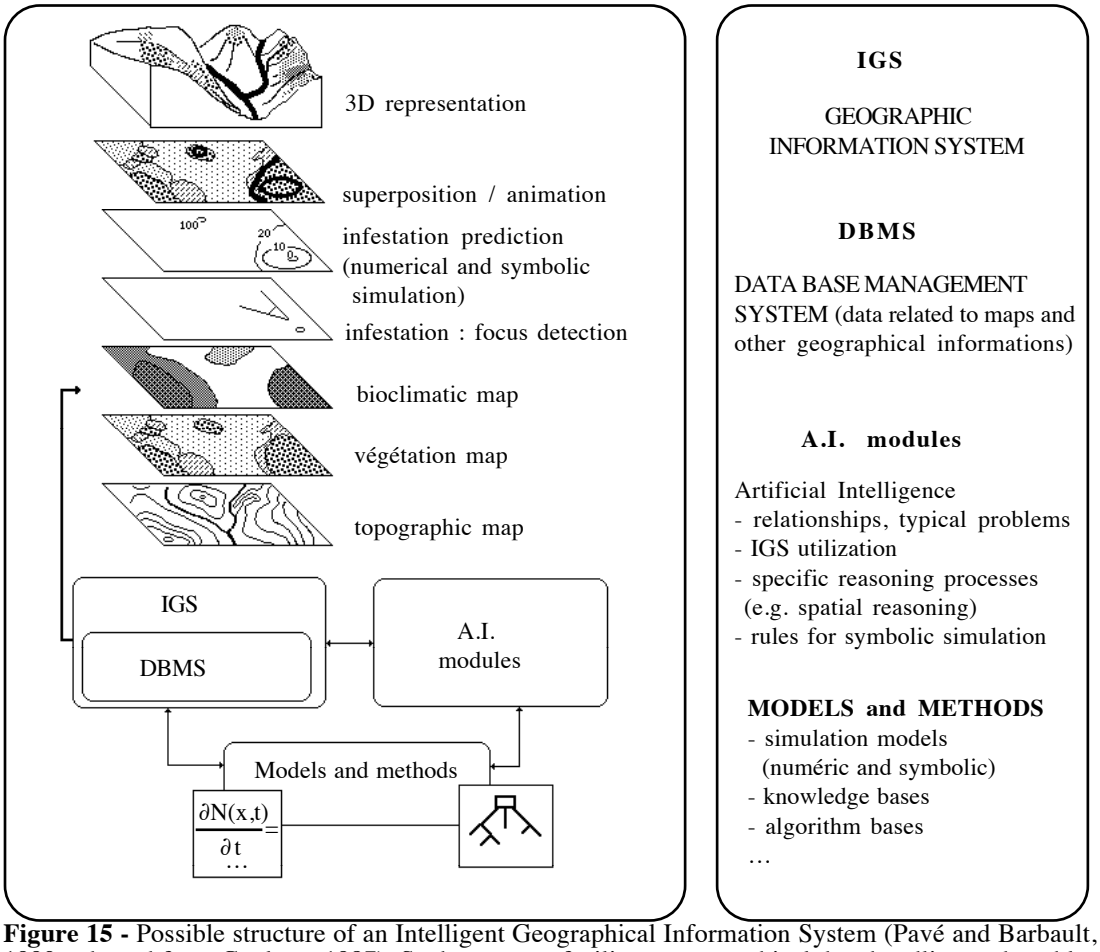


Figure 15 - Possible structure of an Intelligent Geographical Information System (Pavé and Barbault, 1990, adapted from Coulson, 1987). Such a system facilitates geographical data handling and enables the introduction of dynamic aspects. For example one can envisage to simulate a pest infestation by considering the biological, ecological and ethological aspects of the pest and then place this knowledge in the field context by using geographic data contained in data bases. We can imagine such systems which will include a large set of knowledge not only concerning geographic or ecological aspects, but also related to social aspects and even judicial ones.

Finally, such systems can help problem solving in spatial and temporal contexts, for which data and knowledge are qualitative and/or quantitative. A good example is Landscape Ecology where one has to study structures, functions and variations in a heterogenous medium of interacting ecological systems (Coulson, 1987). Many organizations and scientific departments are concerned with such problems (e.g. management of forests by forest departments, sea management by fisheries or oceanographic depatments…). Such systems have to take into account space and time dimensions, and views in a certain context. The time or space scales are different according to the scientific or technical domains as mentioned above : geological data are more persistent than vegetal one, for instance, but the nature of variables may differ. For example, from a vegetation map, we can deduce the vegetation in a particular place, if we go there we can certainly see the corresponding types of plants. Now if we consider a map of animals distribution, as animals are moving and have particular behaviour, we can only speak in terms of encounter probabilities (e.g. since many years the authors of the present text has been hoping to see a tiger in South India, he has regularly visited a natural park where tigers are numerous, but never he has never seen one up to now…).

2.1.5. Simulation of complex systems

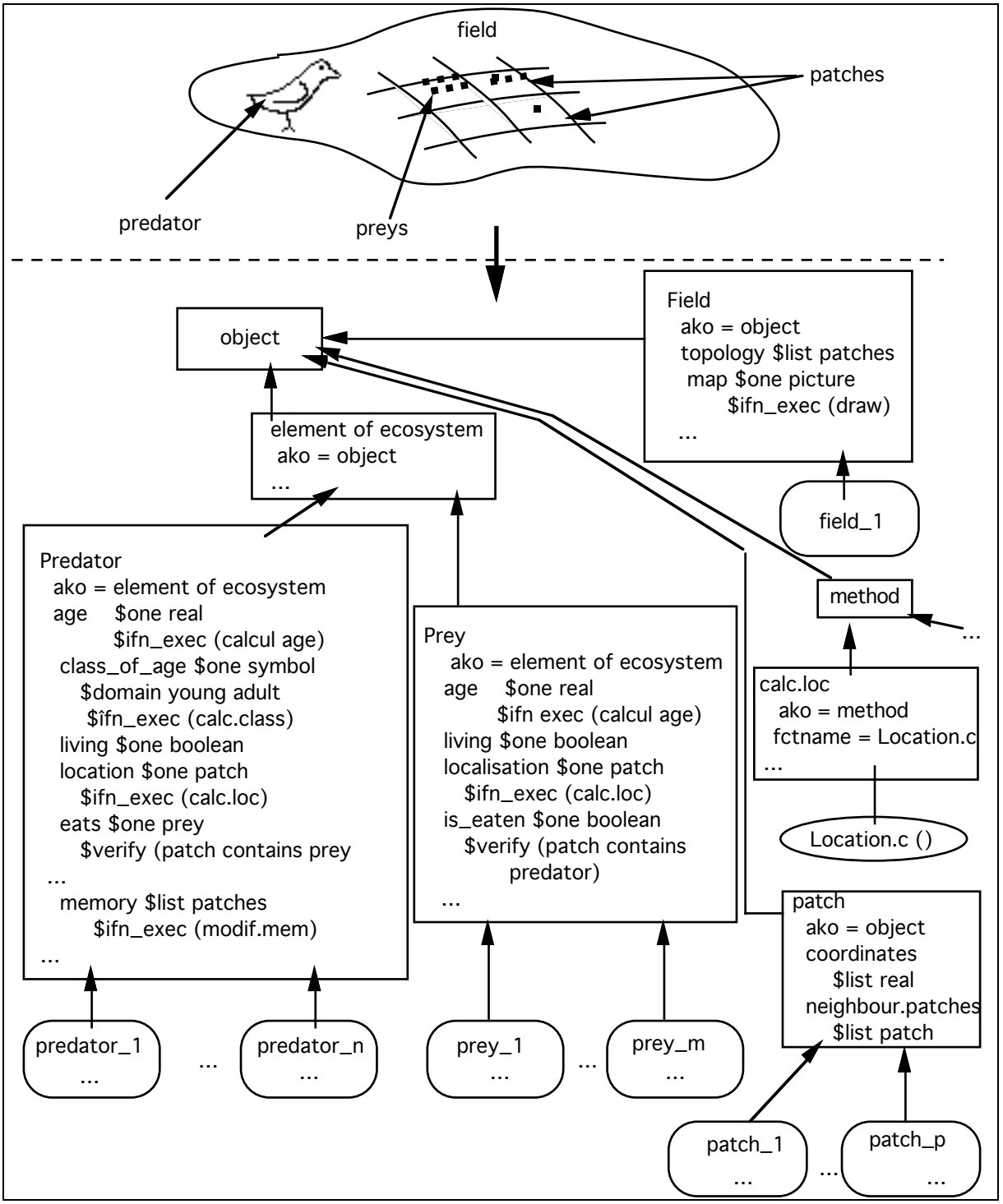


figure 16 - Example of a representation of a predator prey system in a complex environment (field) described by a list of homogeneous patches, by using an object centered representation. Predators, preys are instances of general schemas where common properties are described by slots. Similarly patches which caracterize a particular medium are instances of a class schema which gives the common and general properties of patches.

Very often we have to represent complex situations and complex systems. Classical numerical simulation programs permit such modelling approaches. However they are sometimes limited because they work only on numerical representation and on simple logical rules. For example, in ecological systems modelling, often we have to handle processes in space during time (spatio-temporal

processes), and also individuals of populations which may have sophisticated behaviours (for instance, a predator can learn how to catch a prey in a heterogeneous medium). Today we are mainly interested in the global population dynamics of in a complex and fluctuating environment where individuals could exibit sophisticated behaviours. The study of such a situation is not easy either at the experimental or at the modelling level. For example, let us consider a predator-prey system in a complex environment (e.g. heterogeneous environment relative to prey distribution, they are distributed in "patches"). The common method is to simulate the system numerically. Although the results are very interesting we are limited in introducing complexity (see, Bernstein *et al*, 1988). We can imagine an alternative to such simulation studies by using an object centered representation (representing a population of predators individually by instances of a predator class, preys by instances of a prey class, environment by a set of instances representing the space following parcels or patches. Figure 16 illustrates such a modelling approach).

2.3 Modelling and control of Closed and Artificial Ecological Systems

The Ecology is one of the Life Sciences disciplines which is the most formalized and where theoretical approaches are the most advanced. The basic mathematical models developed during the 30's are well known particularly in the community of mathematician. However, the problems in Nature are so difficult that besides conceptual approaches, the using of models to solve real and pratical ecological problems were still very limited during a long time. But, if we consider now artificial ecological systems, the problems may be simplified. In fact we have already some experiences, for instance some waste treatment processes use in fact an "ecological technology" (e.g. waste water treatment by microbial populations, or more complex processes which involve biomass production such as biological production of domestic sewage treatment lagoons [6], or still experiment in laboratories which involve "complex", generally microbial, systems).

The problem of synthesis, modelling and control of artificial ecosystems is posed by the research on life support systems for spatial exploration. How to ensure the life of a crew during a long flight, for example from Earth to Mars ? It is now admitted that classical techniques used around the Earth and based on physico-chemical processes and imported food cannot be retained and only a Close Ecosystems Life Support Systems (CELSS) are available in such a goal.

The problem is : how to control such systems, not only in an optimization point of view but sometimes uniquely to ensure a sustainable functionning ?

Mathematical models are efficient, but we have first to elaborate them, to verify their properties (qualitative behaviour, contollability, observability, indentifiability,...) and to include them in an operationnal control system. Theoretical results are more and more limited when the complexity increases (a system with more than two populations in an heterogeneous medium becomes difficult to study following a mathematical way, and only some of its properties cans be formalized related to quantitative variables). To help the engineers in the design

of an artificial ecological system and its associated models, it becomes necessary to help him with specialized softwares (aided modelling sofwares such as EDORA Project).

Still mathematical models would be efficient for simple components and to represent quantitative variables. Other variables (qualitative variables) and characteristics (e.g. data and general knowledge) related to an artificial ecological system (or more generally a complex system with biological components) are not easily represented by using a classical mathematical modelling approach. The present, and certainly, future solution is to follow a complementary way by using other formalism of **artificial intelligence** which enable, after the experience of particularly the works of the EDORA Group and some other recent results, already mentionned :

- In representing qualitative variables and associated reasonning processes,

- In elaborating computer systems which enable to manipulate together declarative knowledge, data, mathematical models and associated methods of exploitation. These kinds of approaches could lead to computer systems devoted to quantitative and qualitative control of Artificial (even Natural) Ecosystems. This kind of control system is also debated for classical technological systems, it is now considered as one of the future major application of A.I..

So we think that we have to do determinant advances in mathematical modelling and and their interesting properties in a control point of view, and particularly to do mathematical researches but also computer sciences domains to furnish Intelligent Aided Modelling Systems.

But as the knowledge introduced in a mathematical formula is limited to a little part of knowledge of a complex system, sometimes a mathematical model is local and in fact a set of mathematical models have to be used to describe a phenomenon is a sufficient domain (for example, we have studied the possibility to manage different denitrification process models following different field or experimental conditions).

We have also to manage parallely qualitative variables and associated reasonning processes, data sets and "general" knowledge (for example the "expert knowledge needed to make diagnosis in the case of disfunctions of an ecological system). So the second way of research, apart the aided modelling one, consists to elaborate the control systems architecture for ecological processes which include all types of necessary knowledge and adapted human interfaces (figure 17).

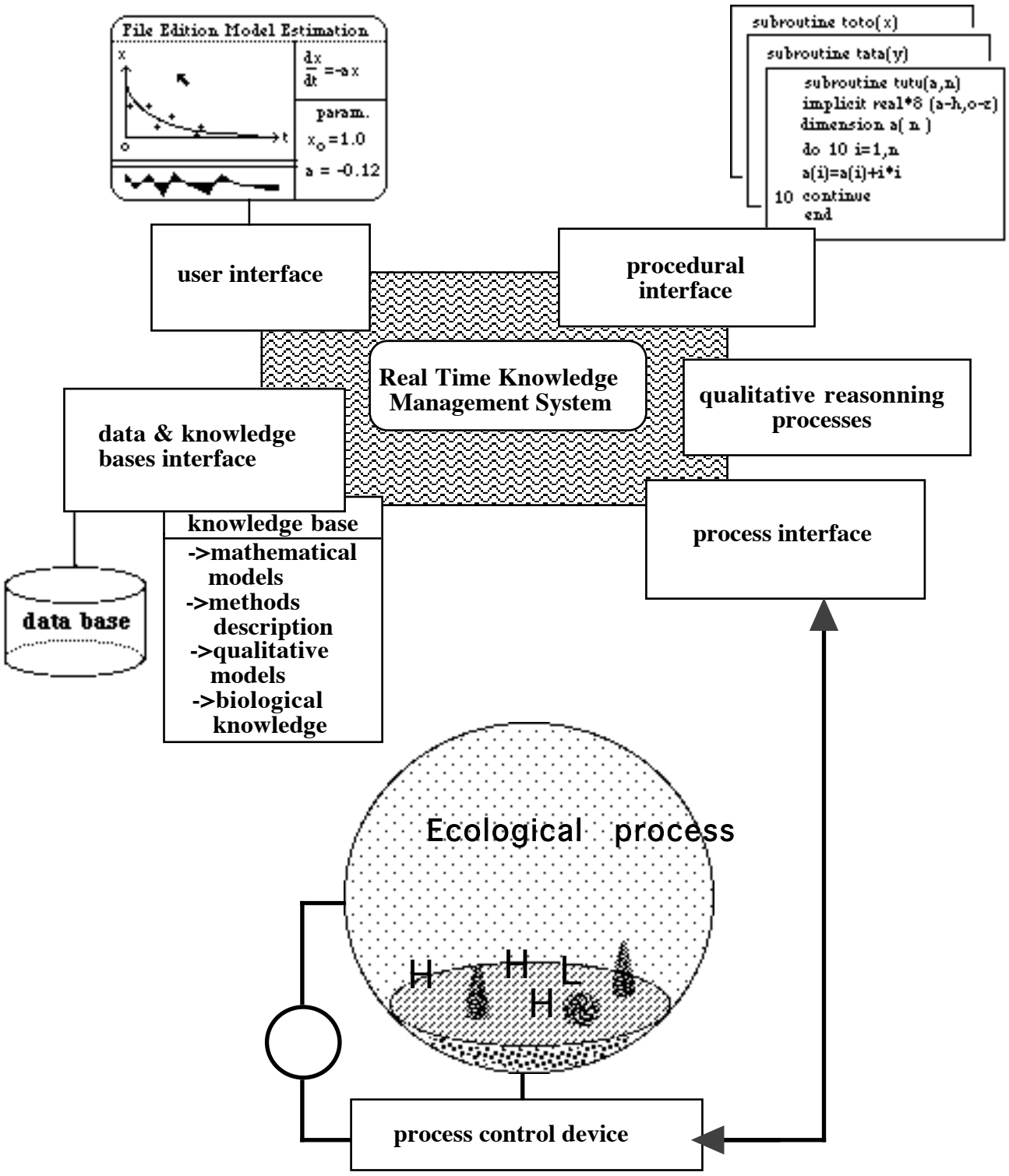


Figure 17 - An example of architecture of a Real Time Knowledge Based System to control complex systems with biological components (such as an artificial ecological systems and its connected devices). Such computer systems will include mathematical models and associated methods, reasonning processes to manipulate qualitative models, general knowledge and data about the system (e.g. knowledge which permit diagnosis in the case of a pathogenic behaviour of the system and associated information to repair it,…).

2.3. Biotechnology

Before examining applications in biotechnology we have to define precisely what is meant by biotechnology. In strict etymological sense this word refers all technologies where living systems are used. However, actually when one speaks of

biotechnology it is understood as "technologies which involve microorganisms and/or genetic manipulations". So the following presentation is restricted to this particular meaning, and Fig. 16 shows the main steps of such an approach. A.I. techniques could to solve some problems in this domain (e.g. analysis of DNA and protein structures).

- Learning and analysis of DNA sequences (Fichant, 1988). This.work shows how learning techniques can help in analyzing some signals (*i.e.* intron-exon separators) in DNA sequences. It is compared to statistical data analysis.

- Data Bases of Nucleic Acid Sequences (cf 2.4.3.). At present, most of the known sequences are stored in data bases (GENEBANK, ACNUC (Gouy *et al*, 1985…). Further, some softwares enabling analysis of sequences are connected to these banks (e.g. ANALSEQ for ACNUC). Then the intervention of A.I. can be seen at different levels:

. as new techniques for sequence analysis, G. Fichant's work cited above is an example,

. as an aid biologists in exploring data bases and in finding the desired information for him, which is the problem of intelligent interfaces of data bases,

. to complete the data from the knowledge associated with involved structures or biological processes. The problem is to define new computers systems which integrate both concepts of data base and knowledge base representations and management.

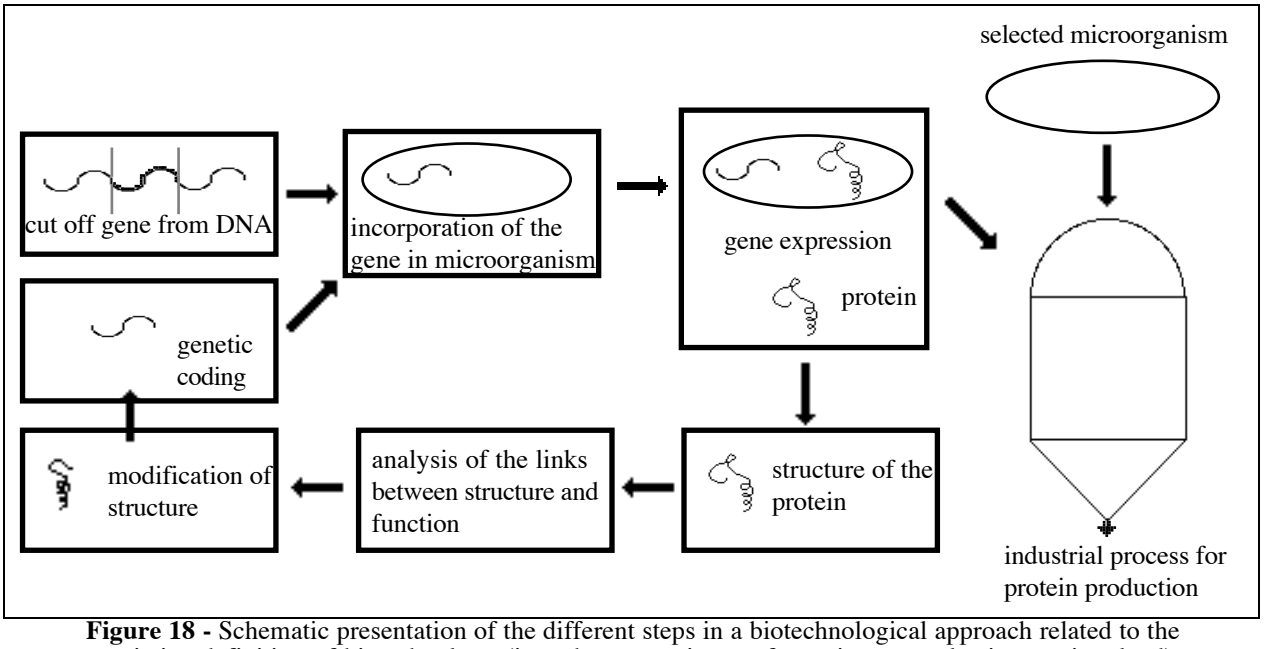


Figure 18 - Schematic presentation of the different steps in a biotechnological approach related to the restrictive definition of biotechnology (i.e. where genetic transformations, or selection are involved). A.I can aid biologists in some steps of this approach for example: (i) to aid in DNA structure analysis (cf. Fichant, 1988), (ii) optimization of experimental design (e.g. MOLGEN), (iii) for spatial structure of proteins (Arbanel, 1984). Some future developments can be easily envisaged, e.g. for process control in industrial production or even in the analysis of links between structure and function of proteins.

- Experimental Planning in molecular biology (MOLGEN, by Steifik, 1981). Very often the designing of experiments in molecular biology is difficult (gene bank elaboration, DNA cutting and analysis by using specific enzymes…). MOLGEN is one of the first system which can aid biologists in designing such experiments.

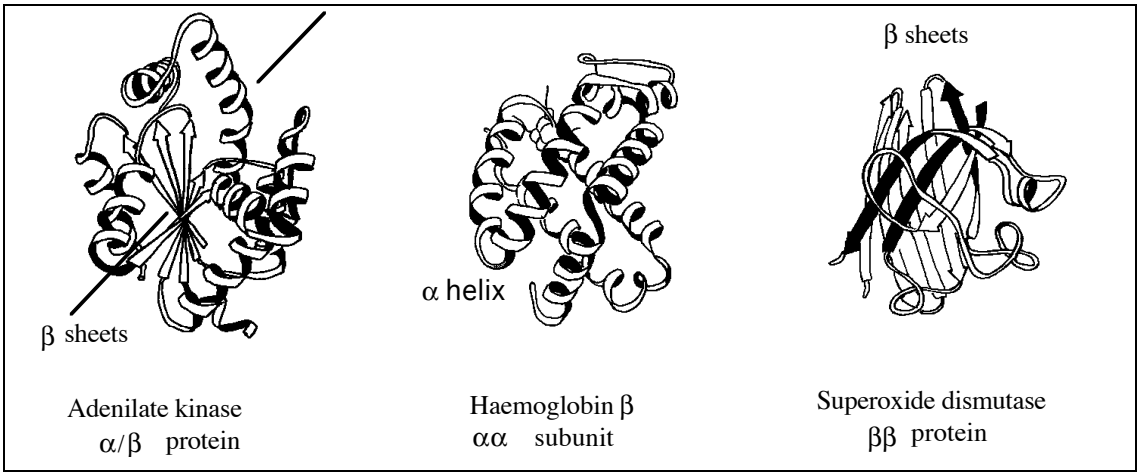


Figure 19 - Examples of typical spatial structures of proteins (from Abarbanel, 1984). The amino-acid composition charactizes sheets and helix. Transitions between helix and sheets correspond to some specific sequences of amino-acids. Rules can be proposed to recognize such sequences from the linear primary structuce and then to detect sheets, helix and transitions zones.

- Protein design [Abarbanel, 1984]. Determination of the spatial structure of a protein with the minimum of experiments, particularly form primary and secondary structures, is difficult. Some rules about the constitution and transition between characteristic subunits (sheets and helix) have been proposed. A computer system can then provide some solutions from a given primary structure (cf. figure 19).

- finally, as already mentioned A.I can be used for process control in biotechnical systems. For example, Guérin (1990) proposed to control a water treatment process based on lagoon exploitation by such a system.

2.4. General

Besides specific systems adapted to a particular situation, some more general tools have to be employed, which can manipulate a more fundamental knowledge (e.g. taxonomic concepts) or which could be an aid (with a philosophy closed to MODVEX one) to methodological approaches.

2.4.1. Taxonomy

Taxonomy is a scientific domain for which applications of artificial intelligence will be probably important in the near future, particularly in systematics of life systems. It is well known that bacteria, animals, plants, etc,..., can be classified. Some of these classifications are built on evolutionary principles while others have a practical interest for species recognition. The last ones are called determination keys. Systematist's work is devoted to constructing such classifications and consequently they are very clever to recognize organisms. Today systematists are not very numerous and this scientific field is often seen as one of the past. It is not our own opinion, but we have to remark that systematists are not always the best supporters of their discipline... One of the consequences of this "disapearing species" enhance, at least, the problem of knowledge conservation,

particularly for identification of specimens in nature. A possible solution is to use an identification key (i.e. a book which theoretically enables species recognition). It is well known that such keys are very useful specialy for specialists, and to anyone who has a good knowledge of life system classifications. Today it appears that artificial intelligence techniques are probably a good solution for nonspecialists and also as a tool for specialists (for example, to elaborate classifications based on learning techniques). We examine here some examples developed in our laboratory on the basis of the already mentioned knowledge base management system SHIRKA. Some references to other developments are also mentioned. It must be remembered that SHIRKA is based on an object centered representation which seems to be very convenient for represent ing taxonomy and to work on it.

Three prototypes were developed : classification of birds (Perennou, 1986) was the first, trees of tropical rain forests (Gautier and Pavé, 1991), fishes of Antarctic Ocean (Gautier *et al*, 1991). It seems that it is a good system to start for studying the problem of automatization in systematics (cf. Fig. 19 and 20).

The basic mechanism used is called **classification** one. In each schema of a level, discriminating slots are defined with their associated domain of values.The principle of this procedure is to create an instance at the upper level of a hierarchy. The system tries to go down as far as possible in this hierarchy following the values given to the slots (cf. Fig. 19 and Fig. 20). These values are obtained by posing questions to the user or by other inference processes. For example the value can be obtained by a procedural attachment. In systematics it can be useful if quantitative characters have to be considered and sometimes discrimination between species can be obtained by statistical comparisons. The system asked only significant questions, some answers cannot be given (reasoning process in an incomplete set of information). At the end of the process three sets of answers are given (trivalent logic) :

- the certain classes correspond to the schemas to which the instance can be attached (all slots have correct values for these classes, *i.e*. values of the domain),
- the possible classes correspond to the schemas for which there is no contradiction : the values of slots are correct or not determined (i.e. the user did not answer to all questions),
- the impossible classes correspond to schemas for which at least the value of a slot does not fit into the admissible domain.

Other inference processes are available : inheritance, default value, pattern-matching and procedural attachment. However, they must be used carefully because they are not always compatible with classification algorithms. In fact, the common way of obtaining slot values during this process is to ask the user or in some cases to compute this value by using an algorithm (e.g. the probability that an individual belongs to a statistical population from results of a discriminating criterion for quantitative characters which is often used to discriminate between species).

```
…
{ordinary-leaves
   ako = leaves;
   gland-on-lamina $domain false;
   kind-of-tree $one symbol
            $domain emergent high-tree medium-tree
                    little-tree liana}
                        |
_____________________________________________________________
|                       |                       |
```

```
                |                                          |                          |      …
                |                                          |
        {emergent                                  {lauracees-2
           ako = ordinary-leaves;                     ako = ordinary-leaves;
           kind-of-tree $domain emergent;             kind-of-tree $domain high-tree;
           smelling-leaves $one boolean}              itself $com "lauraceae cryptocarya
                |                                                  215.litsea 216a221"}
        ________________________________
        |                               |
{polyalthia-fragans                {dipterocarpacees
   ako = emergent-percurrent;         ako = emergent-percurrent;
   smelling-leaves $domain true;   smelling-leaves  $domain false;
   itself $com "annonaceae 28"}     itself $com "dipterocarpaceae dipterocarpus hopea vateria
                                                    91a99"}
```

Figure 20 - A simplified part of the knowledge base for the identification of tree species in tropical rain forests (South India). The main part of the identification is driven by characters attached to the shape of leaves but others have also to be considered such as the size of the tree. At a given level a schema is accepted if the answer to questions conform to the domain admitted. For example "emergent" is reached if the answer to the question "kind-of-tree ?" was "emergent". Then polyalthia-fragans and dipterocarpaceae are admitted. The value "faux" (false) given to the slot "smelling leaves" permits the selection of the genus dipterocarpaceae.

When a species is identified (or not) an explanation procedure can explain the reasons of the classification. It is particularly interesting when the result is negative (i.e. why an observed individual is not considered to belong to a species ?).SHIRKA give the first slot whose value is in opposition to the admissible domain.

In fact, the reasoning mechanism implied in the identification of an organism has very often been taken as examples to illustrate some aspects of Expert Systems. One can find such examples in many papers devoted to the formalisms of rules and related inference processes. Although some authors had long ago mentioned the interest of an A.I. approach to problems of systematics (e.g. Fortuner,1985), the concrete realizations are not very numerous. Among them, may be quoted :

- A prototype of an expert system based on rule representation developed by Woollet and Stone, 1987. This systems is adapted to the identification of a group of insects, more precisely species of the genus *Signiphora* (Hymenoptera: Signiphoridae).

- Another example concerns the identification of zooplankton organism, it was studied by Gandelin and Thonnat (1987). The system includes an automatic recognition of pictures and the A.I. level is based on an object centered representation used by CLASSIC, an environment for expert system development.

- A system for nematods recognition (Nemisys), which today is probably one of the most illustrative of the A.I. possibilities in systematics (Fortuner, 1989, Fortuner, Diederich and Milton, 1991).

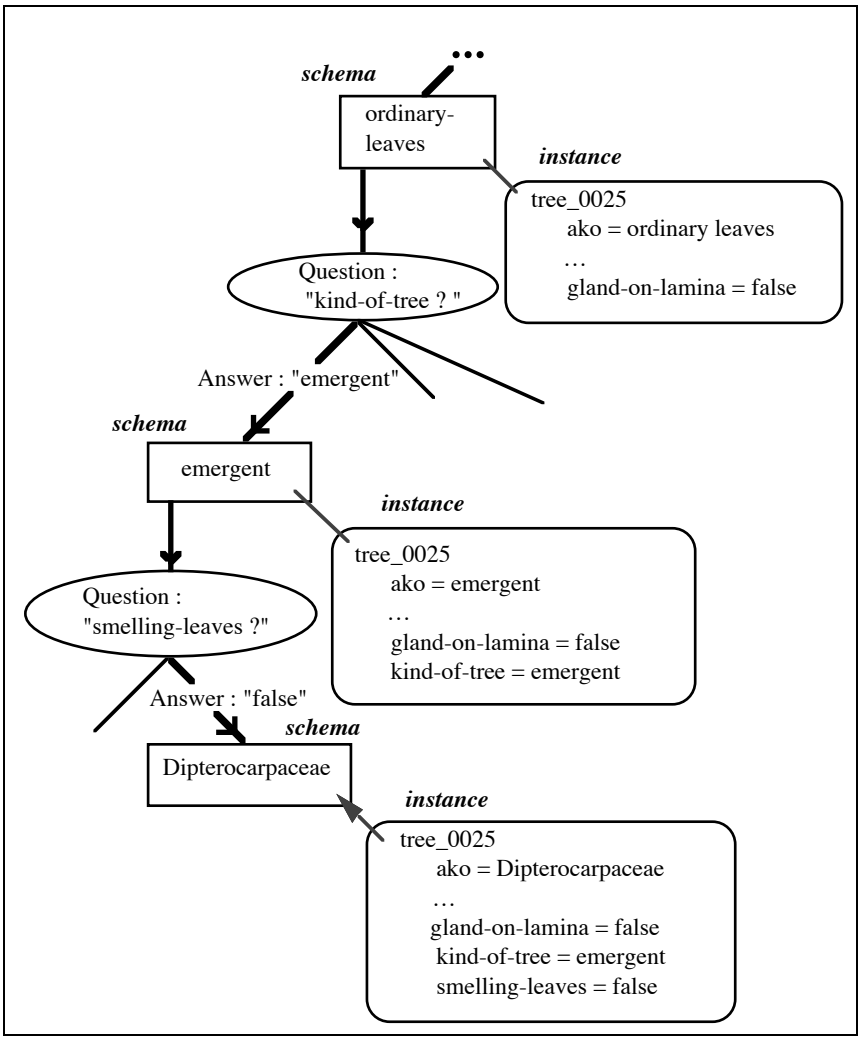


Figure 21 - This figure shows the simplified process of classification : at first, an instance is created, then values are set to slots, by posing questions to the user or by other inference processes, helping to keep this instance as low as possible in the hierarchy and to detect to which schema(s) it may be attached. (i.e. for which slot values are not inconsistent. If desired the attachment can be done by the user via a specific command if the instance is consistent for the class it is attached (note that ako means "a kind of").

2.4.2. Methodology

Considerable effort is made to enable the access of specific methodologies and techniques for all scientists, especially non specialists of these methods or techniques. All non statisticians are award of the dramatic situation, the psychological stress,… when they have to choose a statistical test to analyze their experimental data, and when their papers are returned with reviewers' notes, generally statisticians, who explain that the "*correct test is not xxx, but yyy in the adaptation proposed by X.& Y in Xmetrics, 1943…, and hence this article cannot be accepted in the proposed form for publication…*".

So, many statisticians, biostatisticians and bio-mathematicians are now convinced that, the teaching efforts by the classical lectures and literature, expertise must be proposed in other more convenient forms. Already some packages are used which make calculations but do not contain the expertise about the correct choice of

techniques. The logical continuation is to propose an **expert system to help in solving methodological problems**.

Many systems are now available among which may be mentioned :

- REX : an expert system developed by W. Gale and his collaborators, to aid definitions of problems and solving regression analysis. This system is based on a mixt knowledge representation : rules and objects (Gale, 1987).

- EDORA : a project for mathematical modelling approaches in biology and ecology (Pavé & Rechenmann, 1986,...). SHIRKA, an inference engine which manipulates an Object Centered Representation, was developed by F. Rechenmann for this purpose and has been widely used in other fields for example in systematics. A simple example from EDORA system is presented in section 1.4., it concerns the choice of a mathematical model of a growth process.

- ECO : an intelligent Front End to aid mathematical formulation of ecological problems (A. Bundy, R. Muetzelfeld *et al*, at the University of Edinburgh, 1986), on the basis of C-PROLOG language, the well known dialect of PROLOG developed at Edinburgh.

2.4.3.Data Bases and Knowledge Bases

It is not always easy to distinguish between data bases and knowledge bases. In fact, we have to consider not only the internal formalism but also the capabilities of corresponding management systems (DBMS and KBMS). With time it becomes quite confusing because there is a trend to mixt these concepts and then to propose new systems which manage in the same time knowledge and data. However, basice differences are :

- Data Bases store simple data representing facts in a *closed* "*world*". Relationships are external to data and basically Data Base Management Systems manage array type representations.

- Knowledge Bases store complex objects representing facts and also relationships between objects (e.g. rules) in an *open world*. Knowledge Base Management Systems manage facts (*i.e.* a simple data base), rules and/or more sophisticated representations such as objects. This last formalism which has already been discussed in detail tends to be increasingly used for DBMS, then Data Base and Knowledge Base concepts become more and more closer.

a. Example of a data base structure

A data base is a structured representation, storage and allocation of data, as opposed to flat unstructured files. In a way a data base is a model which exibits the type of data and relationships among them. We present briefly the example of ACNUC structure (Gouy *et al*, 1985). ACNUC (from ACides NUCleiques, *i.e.* Nucleic acids), was constructed to store genetic sequences and related information (reference of publication, type of organisms from which the sequence has been extracted, function of the corresponding protein if it is translated...). The basic information consists of a string of 4 letters (ACGU) which code elementary significant information : the nucleotides (A : adenosine, C : cytosine, G : guanosine, and U : uracile nucleotide). For example, "ACCUGGAGUUCAA..." is such a sequence. The structure of the data base is represented in Fig. 18.

A query language enables us to pose some sophisticated questions to the system. It consists of predefined keywords (SP, K, O, …), logical connectors (and, or…) and relations (<, >, ≠,…). For example the command :

"SP = HOMOSAPIENS" and "K = GLOBIN"

permits the selection of all known globin sequences of man.

It must be noted that in data base technology there is no normalized query language, but only an accepted standard : SQL.

b. Relational data bases

The example presented above is an example of a scientific data base and its associated management system, specifically elaborated for this application. To avoid this programming effort, computer scientists have developed general systems which theorically enable the management of all kinds of data. Historically the first systems were called *hierarchically organized systems* because they only permit the management of data organized following a tree structure. Now the most common systems are *relational data base systems*. They enable a fairly simple expression of the data structure (a model which exibits elementary information and relationships within them). This is a modelling approach and an analogy can be made between modelling data structures and compartmental system modelling (Fig. 22).

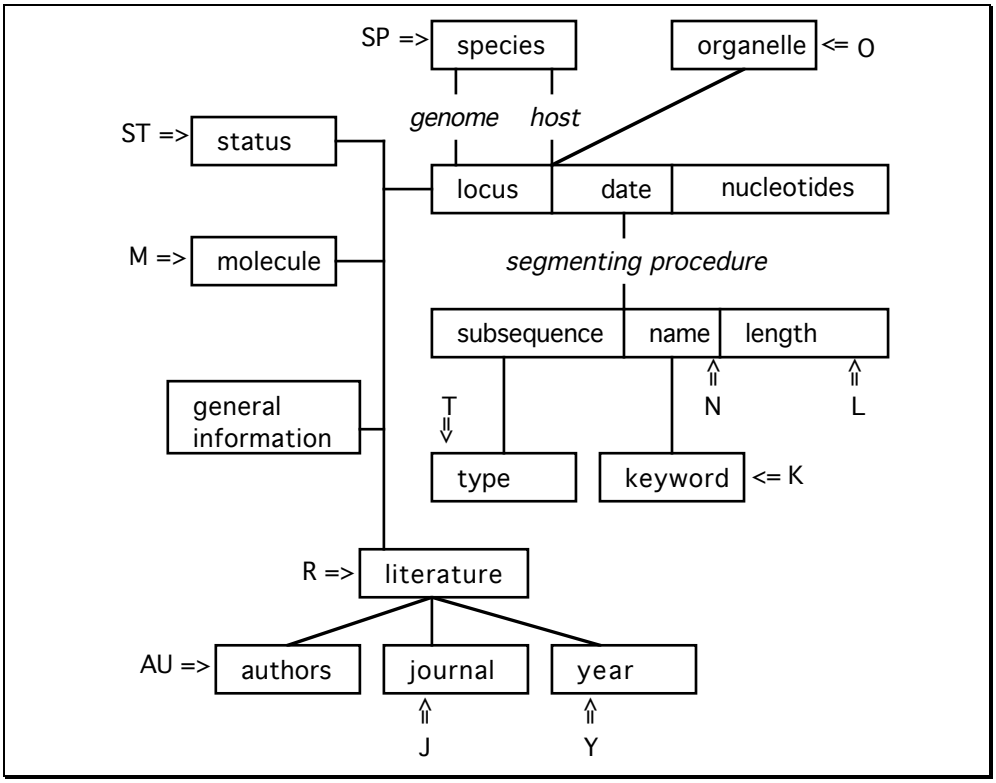


figure 22 - Data Base structure of ACNUC, a base for nucleic acid sequences management. Entry points are indicated by the symbol =>. Corresponding code in the query langage is associated to this symbol. Logical relations between substructures of the base are represented.

c. Problems and limitations of DBMS

Although DBMS are very useful, they have also limitations in their capabilities. Among the principal problems and limitations may be cited :

- difficulty use for a non specialist or non permanent user

- modifications are easy at the level of entities (contents) but difficult at the level of the structure (relationships between entities),

- limited to simple objects and relationships. The information is static and it is difficult to find implicit information … In fact DBMS was developed for specific users (libraries, buisness, administrative management…)and so they are not vell adapted to scientific data. However we must say that new systems based on object centered representations enable more complex object representations.

- despite the claims of developers only 46 % (in 1989) of DBMS is based on relational models.

d. Implicit information

Implicit information is information which is in the base but does not correspond explicitely to entities or relationships. It is generally not possible to find directly this information. However it is possible from a file extracted from the base at which a specific procedure is applied, it is not easy and implies a programming effort by the user).

In fact, what it is not kown in the base is assumed to be false. For example, consider the simple data base (in a Prolog form) :

person (Alain, Professor)
person (Jean-Luc, Assistant_Professor)
person (François, Research_Director)
person (Marc, Student)
person (Antoine, Schoolboy)

if the query is : *teacher (x) ?* the system replies *x = ()* which means there is no teacher in the basis. However it is well known that Professor and Assistant_Professor are teachers. In Prolog this can be avoided by adding a rule which permits the definition of the concept "teacher" :

teacher (x) -> person (x, Professor) or person (x, Assistant_Professor)

Now if you ask the question : *teacher (x) ?*

The system replies :

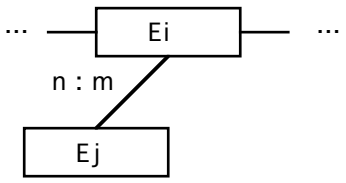
x = Alain

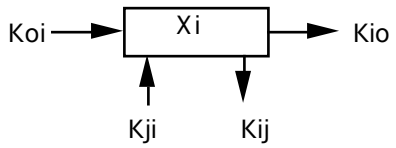
x = Jean-Luc

Very often when the structure of a data base is elaborated we do not have all the pertinent information. For example, when ACNUC was defined, some gene structures was not known : introns, exons… To introduce this new knowledge means redefining all the structures of the base which is very heavy work.

relational data base modelling **compartmental system modelling**

basic description of the structure
by using a graphical description language
(arrows and boxes schematization)

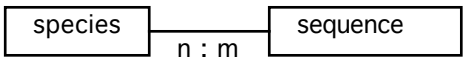




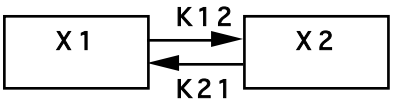
Ei and Ej are named entities corresponding to data in the base which have a particular meaning, type and format.
n is the number of elements Ej which correspond to one element of Ei
m is the number of elements Ei which correspond to one element of Ej

Xi is the state variable associated to the ith compartment
Koi is the flow constant from outside to the system
Kio is the flow constant to the outside from the system
Kij is the flow constant from i to j
Kji is the flow constant from j to i

examples



- several sequences (n) can correspond to a species
- the same sequence can be present in several species (m)



model of a two compartment closed system

code in an operationnal language

relational DBMS langage

```
type sequence : entity
      locus : string (10)
      date : integer (6)
      sequence : string (*)
end
type species : entity
      name : string (10)
      genus : string (10)
      ...
end
type gen_host : relationship
   between : sequences (*,*)
      and : species (*,*)
end
```

mathematical langage

$$\frac{dX_1}{dt} = -K_{12}\,X_1 + K_{21}\,X_2$$

$$\frac{dX_2}{dt} = K_{12}\,X_1 - K_{21}\,X_2$$

figure 23 - Comparison between relational data base modelling and compartmental analysis modelling

e. New developments in DBMS

Obviously the limitations of relational data base systems are well known and there are many developments, particularly in the following domains :

- **Multimedia data bases** enable tthe management of very different kinds of information : text, graphs and pictures, music…

- **Distributed data bases** : the information is located in different geographical locations. The problem is to find where the information is located, and how to obtain it.

- **Time management** : many data are related to time (day temperature, duration of a contract, climatic series, …). At present, time is entered (generally dates or intervals, periods…), as an entity defined by the user, but time has certain properties which could be inbuilt into the DBMS (*e.g.* types and permitted operations). For example:

types :

simple date (y,m,d) hour (h, m, s)

duration (n years, p months, q days)

permitted operations :

{addition : has a sense on the type duration but no sense on dates}

substraction : date_1 - date_2 gives the duration type (if date_1 ≥ date_2).

It would be interesting to define concepts as : history, version, photography,…

- To avoid the complete **revision of data base** structures it is envisaged to develop an intelligent layer to these systems which will permit the handling of new types of data, implicit information, new concepts (*e.g.* "time-expert" for INGRES, Overmyer *et al*, 1982), aid to query langage formulation,…

- **Deductive data bases**, in litterature this expression has two meanings :

. using data which are not only facts but also rules,

. formalisation of data base problems by using classical logics (*i.e.* predicate logics).

In fact there is no contradiction between the two meanings.

Today envisaged applications are :

- to bring DBMS and KBMS closer (by introducing rules and/or object representations and handling in DBMS and conversely to consider and manage efficientlylarge fact bases in KBMS)

- to interface DBMS with a logical programmation language (e.g. Prolog) or with a KBMS) (Kouloumdjian et al , 1986).

2.5.Towards future systems and developments.

Research in Knowledge Based Systems, Expert Systems and more generally in A.I. is very important and will certainly make rapid progress in the future. At present we do not have a paradigm in this domain, but it must not be an excuse to do nothing (*e.g.* not using the existing tools, being unaware of the principal ideas) : the existing tools can be good aids in several fields and conversely applications give ideas to develop new tools.

In the near future many systems will certainly be similar to Edora principles, where declarative and procedural knowledge can be handled in the same system,

with good user-friendly interfaces. Certainly, the object centered representation will also have a great future, after the rules based one.

Knowledge Based Systems enable the storing of knowledge in a particular domain, principally given by "experts". This approach enables to diffuse this knowledge more efficiency and widely than to consult experts who are generally very busy. It furnishes also a formal framework to discussed about the knowledge itself : what is known and also what is unknown. In some cases it can lead to a reconsideration of the expertise itself. This is typically a modelling approach.

Finally, the use of artificial intelligence representation and tools will be developed in qualitative modelling and hybrid modelling and simulation (the example illustrated in 2.1.5., foreshadows such an approach). It seems that often a qualitative approach is the most realistic perhaps the only possible one. Examples can be found in modelling individual behaviour of animals in a complex environment ; another example may be **the action of man on his environment**, by introducing not only adapted inputs in a numerical model but also by considering its individual and collective behaviour, the social structure and the economic constraints and needs. Now it can also be envisaged to construct models which take into account ina same framework the dynamics of natural systems and the behaviour and the decisions processes of human societies.

Annex : Some examples of A.I. and S.E. applications to agriculture

Name of the Application	Domain	Developing level	Formalism	Computer systems and tools	References
CHAMBER	diagnosis	?	rules	IBM-PC INSIGHT	Jones & Haldeman, 1986
TOM, SEPV	diagnosis of plant disease and treatment	17 expert systems developed	rules	GognitIF (Le_Lisp)	Andro et al 1987 Delhotal, 1987
POMME	management of apple trees orchard	prototype	objects	VAX11/780 Prolog	Drake et al.1987 Roach et al., 1987
COMAX-GOSSYM	management of coton crops	Large scale application	rules and procedures (math. models)	VAX MODVEX/ART Lisp	Mc Kinion et al., 1987, 1989
SMARTSOY	recommen-dation for soja crops, forecasting	1989	rules procedures (math. models)	IBM-PC INSIGHT 2+	Batchelord et al, 1989
FinArs	farm economics	1989	rules	IBM-PC INSIGHT 2+	Boggess et al., 1989
DHLES	milk production	prototype 1989	rules data base	IBM-PC (Turbo-Prolog)	Whittaker et al., 1989
CALEX	crops techniques plan	CALEX/Coton : 1989	rules and objects	IBM-PC CALEX (C, Fortran)	Plant, 1989
SELECT/CUE	Selection of crops varieties	validation	objects	IBM-PC SELECT (Prolog)	Morgan et al., 1989
GALAPLAN PLANIPORC	Technico-economics diagnosis of breeding	prototype 1990	rules	GURU	Bourgeat et Lapierre, 1990
OTELO	work management in large scale crops	test on field	rules and objects	?	Attonaty, 1990 and 91 Chatelin and Poussin, 1991
SERRISTE	greenhouse control	test and validation 1991-1992	rules and objects	IBM-PC KAPPA	Boulard et al, 1991 Cros and Martin-Clouaire, 1992

(from N. Girard report, 1992)

Literature : references and *suggested readings

Aguirre Cervantes J.L. - Construction automatique de taxonomies à partir d'exemples dans un modèles de connaissance par objets. Thèse de doctorat de l'Institut Polytechnique de Grenoble, 1989, 282p.

Andro T., Bachacou J., Delhotal P., Fayet J-C., Le Renard J, Rellier J-P, Thiolon C, Wahl V - SEPV : Un ensemble de systèmes experts pour l'aide au diagnostic des maladies des plantes cultivées en France. *Actes Coll "Expert Systems and their Applications"*, Avignon, France, 1987, p: 173-186.

Attonaty J.M., Chatelin M.H., Soler L.G., Poussin J.C. - Nouvelles orientations nouveaux outils de gestion. *Perspectives agricoles*, 147, 35-50.`

Attonaty J.M., Chatelin M.H., Poussin J.C., Soler L.G. - Advice and decision support systems in agriculture : new issues. Decision support systems : IFORS-SPC1, Bruges, 1991.

Batchelor W.D., Mac Clendon R.W., Adams D.B., Jones J.W. - Evaluation of SMARTSOY : an expert simulation system for insect pest management. *Agricultural Systems*, 31, 67-81.

Bensaïd A., Granier D., Rechenmann F. - SHIRKA : des systèmes experts centrés-objet. *Les Systèmes Experts et leurs Applications*. ADI, Avignon, 1984.

Bensaïd A, Rechenmann F., Simonet A., Vignard P. - Mécanismes d'inférences et d'explication dans les bases de connaissances centrées-objet. *Bases de Données et bases de Connaissances*, 8ème Journées Francophones sur l'Informatique, Grenoble, 1985.

Bernstein C. Kacelnik A., Krebs J. - Individual decisions and the distribution of predators in a patchy environment. The Journal of Animal Ecology, 1988, 57, 1007-1026.

Boggess W.D., Van Blokland P.J., Moss S.D. - FinARS : a financial analysis review expert system. *Agriculture systems*, 31, 19-34.

Boulard T., Jeannequin B., Martin-Clouaire R. - Analysis of knowledge involved in geeenhouse climate management. application to the determination of daily setpoints for a tomato crop. First IFAC/ISHS Workshop on Mathematical and Control Applications in Agriculture and Horticulture, Matsuyama, Japan, 1991.

Bourgeat E., Lapierre O. - Méthode pour le développement de systèmes experts de diagnotics technico-économiques des élevages. Actes du Coll. SFER : Nouvelles approches en gestion de l'entreprise agricole, Paris, Nov. 1990.

Buisson L.- Le raisonnement spatial dans les systèmes à bases de connaissances. Application à l'analyse des sites avalancheux. Thèse de Doctorat, Université Joseph Fourier, Grenoble, 1990, 176p.

Buisson L. - Elsa : a Problem Solving environment for avalanche Path Analysis. In "Artificial Intelligence in Numerical and Symbolic Simulation "(A. Pavé & G.C. Vansteenkiste Ed.) ALEAS, Lyon, 1991, 25-50.

*Charniak E. and Mc Dermott D. - **Introduction to Artificial Intelligence.** Addison -Whesley, 1985.

Chatelin M.H., Poussin J.C. - Diagnostic d'exploitation agricole : apport de la simulation. *Economie rurale*, 1991, 206.

Cros M.J., Martin-Clouaire R. - Programmation orientée-objet d'un logiciel de détermination de consignes climatiques. Summary of the "Congrès International d'Informatique Agricole", Paris, June 1992.

Cross S. - Qualitative Reasoning in an Expert System framework. Ph. D., University of Illinois, URBANA, 1983.

Comby S., Flandrois J.P., Pavé A. - A contribution to the study of discriminant capacity of expert systems. Application to the clinical bacteriology. *A.I. and Cognitive Sciences*, , Manchester Univ. Press, 1988, 305-322.

Coulson R.N., Folse L.J., Loh D.R. - Artificial Intelligence and Natural Resource Management. *Science*, 1987, 237, 262-267.

Diederich J., Milton J. - Nemisys : a computer perspective.In : *Advances in Computer Methods for Systematics Biology* (R. Fortuner Ed). The John Hopkins University Press, 1992.

Davis R. - Applications of meta-level knowledge to the construction, maintenance and use of large knowledge bases. In "**Knowledge based systems in artificial intelligence**" (Davis R. & Lenat D.B. Eds), McGraw-Hill, 1980.

Delhotal P. - Réalisation des systèmes experts d'aide au diagnostic. *C. R. de l'Académie d'Agriculture de France*, 73(7), 21-30.

Drake C., Roach J., Virkar R., Weaver M. - POMME : an expert system for helping apple growers. *Actes Coll "Expert Systems and their Applications",* Avignon, France, 1987, 207-230.

Fichant G. - Etude statistique des séquences d'acides nucléiques: reconnaissance et évolution des introns les gènes nucléaires codant pour les protéines. Thèse Université Claude Bernard (Lyon), 1988.

Fortuner, R. - Variabilité et identification des espèces chez les nématodes du genre Helicotylencus. Thèse de doctorat de l'Université Claude Bernard, Lyon, 1986.

Fortuner, R. (Ed) - **Nematode Identification and Expert System Technology**. NATO ASI Series, A, 162, Plenum Press, London, New York, 1988

Fortuner, R. - Problems in nematods identification and the Nemysis solution.In : *Advances in Computer Methods for Systematics Biology* (R. Fortuner Ed). The John Hopkins University Press, 1992.

Friedmann M., Schneider M., Kandel A. - FIDES : Fuzzy Intelligent Differential Equation Solver. *Actes Coll "Expert Systems and their Applications",* Avignon, France, 1987.

Gandelin M.H., Thonnat M. - Etude de l'auromatisation de la reconnaissance d'organismes zooplanctoniques à l'aide d'un système expert. *RFIA*, 1987.

Gautier N. et Pavé A. - Object Centered Representation for Species Systematics and Indentification in Living Systems in Nature. *Cabios*, 1990, 6, 383-386 .

Gautier N., Pavé A., Rechenmann F. - Object centered representation and fish identification in Antarctica. In : *Advances in Computer Methods for Systematics Biology* (R. Fortuner Ed). The John Hopkins University Press, 1992.

Girard N. - Intelligence Artificielle et Agriculture : quelles applications ? Report, DEA "Analyse et Modélisation des Systèmes Biologiques" UCB, Lyon 1992.

Guérin F. - Valorisation aquacole d'eaux usées par lagunage naturel, evaluation biotechnique et modélisation des connaissances. Thèse de doctorat de l'Université Paul Sabatier, Toulouse, 1990, 297p.

Guérin F. - Biological process interpretation using a qualitative reasoning approach. *IMACS annals on Computing and applied Mathematics proceedings*, MIM-S^2, 1990.

Gouy M., Gautier C., Milleret F. - System analysis and nucleic acid sequences banks. Biochimie, 1985, 67, 523-531.

*Hayes-Roth F., Waterman D.A., Lenat D.B. - **Builing expert systems**. Addison-Wesley, 1983, 444p.

Houllier F. - Construction et interprétation de modèles dynamiques : Exemples forestiers. *Cahiers d'Edora*, 1988.

Jones P., Haldeman J. - Management of a crop research facility with a micro-computer expert system. *Transaction of the ASEA*, 29(1), 235-242.

*Kerckhoffs E.J.H., Vansteenkiste G.C., Zeigler B.P. - **A.I. applied to simulation**. *SCS Simulation Series*, 1986, 18, 1, 205p.

Kuipers B.- Getting the Envisionement Right. Proceed. of the National Conf. on A.I., 1982, 209-212.

Kuipers B. - Qualitative simulation. *Artificial Intelligence*, 1986, 29, 289-274.

Lenat D.B. - The nature of heuristics. *Artificial Intelligenc*e, 1982,19, 189-249

Masini G., Napoli A., Colnet D., Léonard D., Tombre K. - **Les langages à objets**. InterEditions, Paris, 1989, 584p.

Martin-Clouaire R. - ELFIN, une approche Système-Expert et théorie des possibilités appliquée en géologie pétrolifère. Thèse de Doctorat, université Paul Sabatier, Toulouse, 1984, 180p.

McCloskey M. - Naive Theory of Motion. in "Mental Models", D. Gentner and A. Stevens (Eds). Lawrence Erlbaum Assoc. Publish., Hillsdale, 1983.

McGowan M.F. - Catcur V1 : a Fishery Management Expert System Module. *ORIA 87*, Marseilles, 1987.

McKinion J.M., Baker D.N., Lambert J.R., Whisler F.D. - Expert Systems for Model Development, and application In cotton Crop Management. *Actes Coll "Expert Systems and their Applications"*, Avignon, France, 1987, 187-206**.**

McKinion J.M., Baker D.N., Whisler F.D., Lambert J.R. - Application of the GOSSYM/COMAX system to cotton crop management. *Agriculture systems*, 31, 55-65.

Morgan O.W., Mac gregor M.J., Richards M.C., Oskou K.E. - SELECT : an expert system shell for selecting among decision or management alternatives. *Agricultural systems*, 31, 97-110.

Misiti Y, Poggi J.M., Astier R., Dupont J., Oppenheim G. Villemin F.Y. - Système expert en conception de lois de pilotage pour lanceurs spaciaux. *Actes Coll "Expert Systems and their Applications"*, Avignon, France, 1987, p: 1053-1075

Mouaddib N., Foucault O. - gestion des informations nuancées dans une base de connaissance en sciences naturelles. utilisation de la théorie de

possibilité/nécessité. 4ème journées "Bases de Données Avancées" PRC-BD3, INRIA Ed., 1988, 297-317.

Muetzelfeld R., Bundy A., Uschold M., Robertson D. - ECO, an intelligent front end for ecological modelling. In "*A.I. Applied to Simulation*" , Ed. Kerckhoffs, Vansteenkiste G.C., Zeigler B.P.. *SCS Simul. Serie*, 18, 1986, 67-70.

Oswald O. - An expert system for the diagnosis of tank irrrigated systems : a feasability study. Ph D. Thesis, Anna University, Madras, published by the French Institute of Pondicherry (India), 1991.

Pavé A. et Rechenmann F.- Computer aided modelling in biology : an Artificial Intelligence approach. In "*A.I. Applied to Simulation*" , Ed. Kerckhoffs, Vansteenkiste G.C., Zeigler B.P.. *SCS Simul. Serie*, 18, 1986, 52-66.

Pavé A., Corman A., Bobillier-Monot B. - Utilisation et interprétation du modèle de Gompertz. Application à l'étude de la croissance de jeunes rats musqués (*Ondatra zibethica* L.). *Biom. Praxim.*, 1986, 26 : 123-140.

Pavé A. - Systèmes experts : application à la modélisation en biologie. *MATAPLI*, 11, 5-19, 1987.

Pavé A. - Interprétation et construction de modèles différentiels de la dynamique des populations à l'aide de schémas fonctionnels. *Cahiers d'Edora,* 1988.

*Pavé A.(Ed) - Les Cahiers d'Edora. INRIA, 1988.

Pavé A., Barbault R. - Ecosystèmes et Environnement. *Nature-Science-Société*, 2, 1992.

Pavé A. - Biométrie, Modélisation et Intelligence Artificielle. "*Journal of integrated study of Artificial Intelligence, Cognitive Sciences and applied Epistemology*", 1989, 6, 2/3, 153-176.

*Pavé A., Vansteenkiste G.C. - **Artificial Intelligence in Numerical and symbolic Simulation**. Aleas, Lyon 1991.

Perennou C. - Systèmes Experts et Taxonomie. Report. DEA "Analyse et Modélisation des Systèmes Biologiques" UCB, Lyon 1986.

Pivot B., Prokop M. - Définition et réalisation d'une fonctionnalité de Classification dans SHIRKA. *Rapp Année Spéciale I.A., INP Grenoble,* 1987, 42p.

Plant R.E. - An integrated expert decision support system for agricultural management. Agricultural systemss, 1989, 29, 49-66.

Plant R.E. - An artificiel intelligence based method for scheduling crop management actions. Agricultural systemss, 1989, 31, 127-155.

Rechenmann F. - Intelligence Artificielle et construction de modèles dynamiques". *Intelligence Artificielle et Productique*, Paris, 1984.

Rechenmann F. - SHIRKA : mécanismes d'inférence sur une base de connaissances centrée-objet". 5ème Congrès-Exposition AFCET-ADI-INRIA "*Reconnaissance des formes et Intelligence artificielle*". Grenoble, 1985.

Rechenmann F., Rousseau B. - Edora : an Object Centered Knowledge-Based approach to Dynamical Modelling. in "AI and Cognitive Sciences", ed. Demongeot J., Proceed in non-linear science, Manchester University Press, 1988.

Rechenmann F., Uvietta P. - SHIRKA, manuel de présentation et d'utilisation. Grenoble, 1989.

Rechenmann F., Uvietta P. - SHIRKA, an Object Centered Knowledge Based System. In "Artificial Intelligence in Numerical and Symbolic Simulation "(A. Pavé & G.C. Vansteenkiste Ed.) ALEAS, Lyon, 1991, 25-50.

*Rich E. - **Artificial Intelligence**. Mc Graw Hill, 1983. Trad. Française : Intelligence Artificielle.. Masson, Paris, 1987.

Roach J., Virkar R., Drake C., Weaver M. - An expert system for helping apple growers. *Computers and Electronics in Agriculture*, 2, 97-108.

Rousseau B. et Rechenmann F. - Le projet Edora : Vers un poste de travail informatique pour l'aide à la modélisation des systèmes dynamiques en biologie. *Les Cahiers d'Edora*, 1988.

Rousseau B., Pavé A., Rechenmann F. et Landau M. - Edora project: Artificial Intelligence approach and work station concept to aid dynamic modelling in biology and ecology. *Suppl. Proceed. of the Summer Computer Conference*, Reno Nevada, 1986, SCS, 14-20.

Smith R. - Biomass storage, Decision Support System an Expert System in Crop Production and Processing. Ph. D., Purdue University, 1985, 278p.

Stefik M. - Planning and meta planning (MOLGEN : part 1). Artificial Intelligence, 1981, 16, 111-139.

Stefik M. - Planning and meta planning (MOLGEN : part 2). Artificial Intelligence, 1981, 16, 141-169.

*Trappl R. - **Impacts of Artificial Intelligence**. North-Holland, 1987, 279p.

Uschold M., Muetzelfeld R., Bundy A., Harding N., Robertson D.- An Intelligent Front End for Ecological Modelling. In "*A.I. Applied to Simulation*" , Ed. Kerckhoffs, Vansteenkiste G.C., Zeigler B.P.. *SCS Simul. Serie*, 18, 1986.

Whittaker A.D., Tomaszewski M.A., Taylor J.F., Fourdraine R., Van Overfeld C.J., Schepers R.G. - Dairy herd nutritional analysis using knowledge systems techniques. *Agricultural systems*, 31, 83-95.

*Winston P.H. - **Artificial Intelligence**. Addison-Weysley, 1984. Revision and French translation : Intelligence artificielle, Interédition, 1989.

Woolley J.B. and Stone N.D. - Application of Artificial Intelligence to Systematics : Systex - a prototype expert system for species identification. Syst. Zool., 36(3), 248-267, 1987.

Xiang Z, Srihari S.N. - Spatial Structure and Function Representation in Diagnostic Expert Systems. *Actes Coll "Expert Systems and their Applications"*, Avignon, France, 1985,191-206.