31

Biomedical Knowledge Engineering Using a Computational Grid

Marcello Castellano and Raffaele Stifini
Politecnico di Bari, Bari
Italy

1. Introduction

Bioengineering is an applied engineering discipline with the aims to develop specific methods and technologies for a better understanding of biological phenomena and health solutions to face the problems regarding the sciences of life. It is based on fields such as biology, electronic engineering, information technology (I.T.), mechanics and chemistry (MIT, 1999). Methods of Bioengineering concern: the modeling of the physiological systems , the description of electric phenomena or magnetic ones ,the processing of data, the designing of medical equipments and materials or tissues, the study of organisms and the analysis of the link structure property typical of biomaterials or biomechanical structures. Technologies of Bioengineering include: biomedical and biotechnological instruments (from the elementary components to the most complex hospital systems), prosthesis, robots for biomedical uses, artificial intelligent system, sanitary management systems, information systems, medical informatics, telemedicine (J. E. Bekelman et al, 2003).

Biomedicine has recently had an innovative impulse through applications of computer science in Bioengineering field. Medical Informatics or better the Bioinformatics technology is characterized by the development of automatic applications in the biological sector whose central element is the information. There are several reasons to apply the "computer science" in many fields, such as the biomedical one. Advantages as the turn-around time and precision are among the basically improving factors for a job. For example the identification of the functions of genes has taken advantage from the application of an automatically system of analysis of database containing the result of many experiments of microarray getting information on the human genes involved in pathologies (C. Müller et al, 2009). With a such approach regions with specifically activities have been identified inside the DNA regions, different regions exist in the genome, some stretches are the actual genes, others regulates the functions of the former ones. Other research have been made through computational techniques on the Functional Genomics, Biopolymers and Proteomics, Biobank e Cell Factory (M. Liebman et al, 2008).

This chapter explores a particularly promising area of systems development technological based on the concept of knowledge. The knowledge is useful learning result obtained by an information processing activity. The **Knowledge Engineering**, regards the integration of the knowledge in computer systems in order to solve the difficult problems which typically require a high level of human specialization. (M. C. Linn, 1993)

Whereas standalone computer system have had an important impact in Biomedicine, the computer networks are nowadays a technology to investigate new opportunities of innovation. The capacity of the networks to link so many information allows both to improve the already existing applications and introduce new ones; Internet and the Web are two well know examples. Information based processes involved in the research to discovery new knowledge take advantage from the new paradigms of distributed computing systems.
This chapter is focuses on the design aspects of the **knowledge-based computer systems** applied to the biomedicine field. The mission is to support the specialist or researcher to solve problems with greater awareness and precision. At the purpose, a framework to specify a computational model will be presented. As an example, an application of the method to the diagnostic process will be discussed to specify a knowledge-based decision support system. The solution here proposed is not only to create a knowledge base by the human expert (or by a pool of experts) but support it using automatic knowledge discovery process and resources enhancing data, information and collaboration in order to produce new expert knowledge over time.
Interoperability, resource sharing, security and collaborative computing will emerge and a computational model based on grid computing will be taken into account in order to discuss an advanced biomedical application. In particular in the next section it will be presented a framework for the Knowledge Engineering based on a problem solving strategy. In section 3 the biomedical diagnostic process will be analyzed using the knowledge framework. In particular the problem, the solution and knowledge resources will be carried out. In section 4 the design activity of the diagnostic process is presented. Results in terms of system specifications will be shown in terms of Decision Support System architecture, Knowledge Discovery and Grid Enable Knowledge Application. A finally discussion will be presented in the last section.

2. Method for the knowledge

Modeling is a building activity inspired by the problem solving for real problems which not have a unique exact solution. The Knowledge Engineering (K- Engineering) deals with the computer-system applications which are computational solutions of more complex problems which usually ask for a high levels of human skill. In this case the human knowledge must be encoded and embedded in knowledge based applications of computer systems. The K-Engineer build up a knowledge model useful for an algorithmic description by a structured approach or method. Three macro phases can be distinguished in the modeling process of knowledge:

1. Knowledge Identification (K- Identification);
2. Knowledge Specification (K- Specification);
3. Knowledge Refinement (K- Refinement).

These phases can be cyclical and times retroaction rings are necessary. For instance, the simulation in the third phase can cause changes in the knowledge model. (A. Th. Schreiber, B. J. Weilinga, 1992). Each phases is composed by specific activities and for each activities the K- Engineering literature proposes different techniques. The Fig. 1 shows the modeling of the knowledge based on the problem solving strategy. The proposed framework is applied at different levels of abstraction from high to low level mechanisms (top-down method).

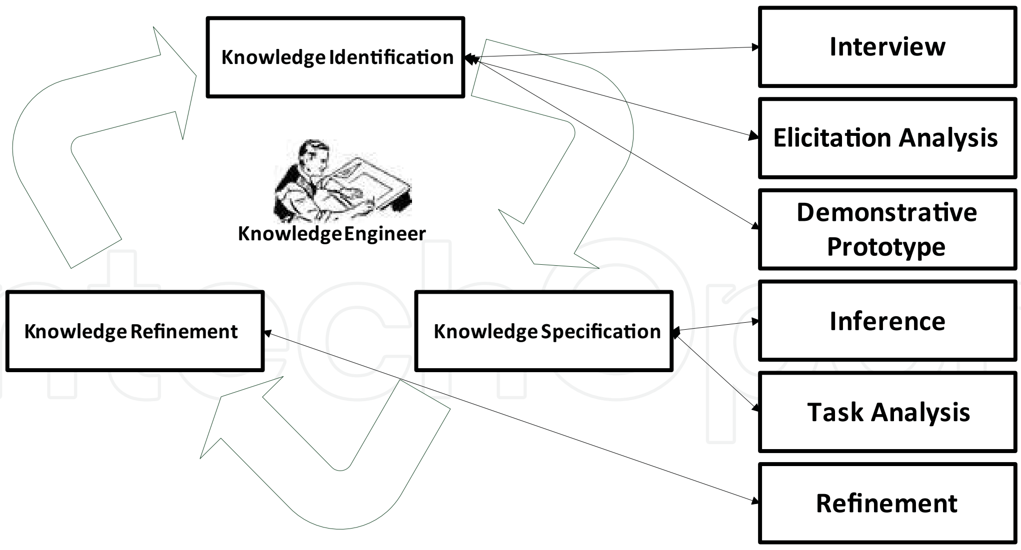


Fig. 1. The knowledge modeling framework composed by phases and activities.

The knowledge model is only an approximation of the real world, which can and must be modified during time.

2.1 Knowledge Identification

A Knowledge Based System (KBS) is a computational system specialized in applications based on knowledge, aiming at reaching a problem solving ability comparable with a human expert. The expert can describe a lot of aspects typical of his own way of reasoning but tends to neglect a significant part of personal abilities which cannot be easily explained. This Knowledge, which is not directly accessible, must, however, be considered and then drawn out. To mine the tacit knowledge an application on the elicitation techniques can be useful and it must be represented using a dynamic model.

The analysis must be inherent in the aims of the planner. On the other hand the representation of the complete domain is un-useful, so the effort is now to identify the problem, in order to finalize the domain analysis.

The approach here proposed is based on the answer to questions that must be taken in to account to develop the basic characteristics of the K-Model as shown in Table 1.

The phase of Knowledge Identification is subject to important considerations that go to better specify the system architecture. Most of a man knowledge or of a group is tacit and cannot be outspoken wholly or partly. Therefore, in a knowledge system, the human beings are not simple users, but an integrating part of the system. The representation is necessarily different from what is represented; it can capture only the most relevant aspects, chosen from the modeler. This difference can cause problems if one wants to use the model for differ purpose from the ones it quality allows. Moreover the difference between the real world and its representation should cause problems of uncertainty and approximation which will be solved showing the quality of the relevant knowledge in real way.

2.1.1 Interview

The interview is a conversation provoked by the interviewer, addressing to subjects choosen according ground of a plan of survey, aiming at the knowledge, through the conversation driven from the interviewer, using a flexible and not standardized scheme. *(Corbetta, 1999).*

What must be represented:	At the epistemological level identify what should represent aspects of knowledge that is necessary to consider the application to be addressed. In particular, what are its classes, patterns, what are the inferential processes involved and the quality of relevant knowledge.
Which is the problem:	To identify the problem to be solved is important to address the investigation about the relevant knowledge. It will be very important in the next modeling phases.
How the problem can be solved:	It indicates strategies for solving a given problem based on patterns bounded in the application domain.
How to represent:	Modeling derives from the subjective interpretation of the knowledge engineer with regards to the problem to be faced; a mistake is always possible and therefore the knowledge model must be made in a revisable way. Tools and processes for the knowledge management have been consolidated, this management can be expressed in several ways: rules, procedures, laws, mathematical formulae, structural descriptions.

Table1. Knowledge Identification guidelines.

The interviewee must have some characteristics related with his life or his belonging to a certain social group, the number of the people interviewed must however be consistent, so it is possible to obtain every possible information on the phenomenon.
The conversation between the two parts is not comparable to a normal conversation because the roles are not balanced: the interviewer drives and controls the interview respecting the freedom of the interviewer in expressing his opinions.
According to the different degree of flexibility is possible to distinguish among:

1. structured interview
2. semi-structured interview
3. not structured interview.

Usually the structured interview is used to investigate a wide phenomenon, the interview is carried out by a questionnaire supplied to a large sample of people; in this case the hypothesis must be well structured a priori.
The structured interview can be used in a standard way but at the same time the limited knowledge of the phenomenon does not allow the use of the multiple choice questionnaire. As the number of the interviewees decreases a semi-structured or a un-structured interview can be taken in to account.

2.1.2 Elicitation analysis

The elicitation analysis is an effective method used by the knowledge engineer to individuate the implicit mental patterns of the users, and categorizing them.
The knowledge engineer carry out this phase analyzing documentations about the domain under the investigation and match the information according to some well know mental model of him.
Knowing the mental patterns and the implicit categorizations makes possible the organization of the information so that it is more simple to use them, improving, in that way, the quality of the product.
Through the elicitation analysis it is possible to identify the classification criterion used by the users and to identify the content and the labels of categories they used. Possible differences in the categorization among various groups of interviewed can be seen and controlled.

2.1.3 A draft of the conceptual model

This activity establish a first formal representation of knowledge acquired up to now composed by elements and their relationships. The representation is used to check the correctness by the user. It is a formal scheme on which the K-specification phase will runs. The knowledge is represented using an high level description called conceptual model. This model is called conceptual because it is the result of a survey carried out by the literature and domain experts for the transfer of concepts considered useful in the field of study is concerned. Fundamental indications about “what is” and “how to build” the conceptual model are shown in Table 2. Some formalisms are proposed in the literature: the semantic networks are used to represent the knowledge with a graph structure; the frames are data structures which allow group, like inside a frame, the information about an entity; an object representation allows to join procedural aspects with declarative aspects, in a single formalism; and so on.

What it is not:	• it is not a basic of knowledge on paper/calculator • it is not an intermediate representation
What it is:	• it is a complete articulate and permanent representation of the structure of the knowledge of an applicative dominion both from a static point of view and a dynamic one • it is a representation independent from implementation • it is the main result of the activity of knowledge analysis and modeling.
How the conceptual model is built (some criteria)	• a formalism for the expressive conceptual representation allows to express powerfully all the concepts, the relationships and the link typical of the application • economic: synthetic, compact, readable, modifiable • communicative: easily intelligible • useful: a support for the analysis and modeling activities

Table 2. The Conceptual Model guidelines.

2.2 Knowledge specification

The goal of this stage is to get a complete specification of the knowledge. The following activities need to be carried out to build such a specification.

2.2.1 Inference

Now let is write about the inferential structures which make possible to know things starting from a codified knowledge. It is interesting, even for the inferential structures, to identify different types of structures, in order to focalize, during the construction of the conceptual pattern, the structures which are actually used in the specific application. The most general form is the one which turns into rules. However it is interesting to consider even more specific structures, as these ones help the identification of particular necessities inside an application. (Waterman [Wa86]). The main characteristics of an inference are the ability of specifying the knowledge, the ability of reasoning, the ability of interacting with the external world, the ability of explaining own behavior and the ability of learning. The inference architecture can be organized in object level and meta level. Each of them can be seen as an individual system, with an appropriate language of representation. The aim of the level object is to carry on reasoning in the domain of the system application, whereas the aim of the meta-level is to control the behavior of the object level.

2.2.2 Task analysis

The aim of the task analysis is to identify the “main task” by the analysis of the users involvement in order to understand how to they execute their work, identifying types and levels:

- How is the work carried on when more people are involved (workflow analysis);
- A single man work during a day , a week or a mouth (job analysis);
- Which tasks are executed by all the people who could use to product (task list);
- The order every uses of execution of the tasks (tasks sequences);
- Which steps and decision the user chooses to accomplish a task (procedural analysis);
- How to explode a wide task into more subtasks (hierarchies of task).

The Task analysis offers the possibility to view the needs, display the improvement areas and simplify the evaluation. It can be carried out according to: (Mager,1975; Gagnè,1970)

- rational analysis - Inside the theories of the Knowledge it is a procedure which divides a task into simpler abilities, up to reaching the activities that can be executed by every process the task is assigned to. The result of this procedure is a typical hierarchy of activities with a correspondent hierarchy of execution aims.
- empirical analysis - Inside the Knowledge Engineering it indicates a procedure which splits up the activity or task into executive process, strategies and meta-cognitive operations which the subject accomplishes during the execution of that task. The result is a sequence, not always ordered, of operations aiming at the realization of the task.

This is an activity about K-Specification. It works on the output of the K-Identification (see Table 1), that specify the resolution strategy of the problem. At the purpose the task analysis is carried out by the following specific steps: Problem specification; Activity analysis; Task modeling and Reaching of a solution.

In the **Problem Specification** phase the problem must be identified specifying one or more activities for its realization, at the conceptual level; these activities will be analyzed in the following step. In **Activity Analysis** a task is identified grouping activities which must be executed to achieve the aim of the task. There are different task hierarchies where the activities can be divided into subtask. This exercise on the task hierarchy means both to specialize every task and to study the task execution on the base of priorities and temporal lines. **Task Modeling** builds a model which precisely describes the relationships among tasks. A model is a logical description of the activities which must be executed to achieve the users goals. The model based design aims to specify and analyze interactive software applications from a more semantic level rather than from an implementative one. Methods for modeling the tasks are:

- Standard : analysis on how tasks should be made;
- Descriptive: analysis of the activities and tasks just as they are really made. Task models can be taken into accounts according to the following point of view:
- System task model. It describes how the common system implementation states the tasks must be executed;
- Envisioned task model. It describes how the users must interact with the system according to the opinion of the designer;
- User task model. It describes how the tasks must be done in order to reach the objects according to the opinion of the users.

Usability problems can arises when discrepancy occurs between the user task model and the system model. The last step in the task analysis (**Reaching of a solution**) is devoted to specify the tasks identified. That are conceptual building block of this analysis. Table 3 shows a formalism to specify the task aim, the technique used for the realization of it and the result produced on the task execution. Moreover a procedural description of the task must be carried out using conceptual building blocks based tools.

Task	Name of the task
Goal	Aim of the task
Techniques	Task descriptions and it implementation
Result	Output task

Table 3. A Task Description formalism.

2.3 Knowledge refinement

The aim of the Knowledge Refinement is to validate the knowledge model using it in a simulation process as much as possible. Moreover it is to complete the knowledge base by inserting a more or less complete set of knowledge instances.

3. Analyzing the biomedical diagnostic process building a model for knowledge based system

The case of study here presented, refers to the diagnostic process. This is a rich knowledge process prevalent in the biomedical field and to diagnostic pathologies starting from the symptoms.

3.1 Knowledge identification

The identification of knowledge in biomedicine has been here applied as described in table 4, using the framework proposed in the previous section.

Key Questions to drive the Knowledge	Methods
<What to represent>	Elicitation Analysis
<Which is the problem>	Interview
<How it can be solved>	Interview, Elicitation Analysis
<How to represent>	Semantic Network, Elicitation Analysis, Interview

Table 4. The K-Identification for the diagnostic process.

3.1.1 Elicitation analysis

In order to create the first reference model of the diagnostic process the K-Identification starts with the elicitation study. As the final aim is the development of an conceptual model could be useful to consider a process comparable where a living organism is like a perfectly working computer. If a computer problem rise, the operative system signal it to the user. To activate such a process, a warning is necessary, i.e. a message of mistake or wrong working. At this point a good computer technician put in action a diagnostic process based on the warning to individuate the problem, or in other words the error.

The diagnosis on an organism is similar at the described scenery: the occurring of a pathology is pointed out to the organism through the signal of one or more symptoms (as already described for the computer errors). The diagnostic medical process is exercised by the specialist (in analogy of the computer technician) that will study the symptom origin and its cause and hence the diseases. The arising of a problem can be due both endogen and exogenous causes and provokes an alteration which would not normally happen (for instance an alteration in the albumin level produced by the pancreas); this mutation causes a change, a working different from the mechanisms associated to that element (for example the mechanism, thanks to which the insulin makes the glucose enter into the cells for the production of vital energy, changes as an accumulation of glucose in the blood circle is met: subjects affected with diabetes). What it has been learned by the above application described on the elicitation analysis is shown in Table 5.

A stirring up cause…	Following a diminution of electric tension the computer is turned off while the hard disk is on	A strong and cold wind blows for several minutes on the patient neck
…modifies some values	The head fall and ruins the computer record	Not desired presence of Ag.VHI-A
…and a problem arises in the system	Not readable starting record	Aching eymphonodes
…and makes clear. Reveals itself.	The computer show the error	The patient has throat ache

Table 5. Some elicitation analysis results to know the Diagnostic Process.

3.1.2 Interview

Although the elicitation study supply important elements to build a conceptual model of the knowledge, the expert contribution is of the great important also. At the purpose the structured interview is here applied to extract the knowledge and the mental process of the experts when they work. In other words the interview application aims to describe the experiences of the specialists in terms of structured knowledge. By interview the Clinical Case emerges as a summary composed by: all the possible information about the patient, a list of symptoms and a set of objective and instrumental exam results. Table 6 summarizes the results carried out by interview and elicitation analysis, in terms of problem and solution.

Key Questions to drive the Knowledge	**Results**
Which is the problem	The diagnoses are not always easy to be identified from a set of the possible solutions, sometimes several possible investigations must be done up to individuate an optimal solution.
How it can be solved	A valuable support to the specialist in order to makes a decision is highly desirable. An automatic computer system based on biomedical knowledge should be able to support the specialist using updated information and supporting decisions using a computational formal approach.

Table 6. Interview and Elicitation Analysis Results for the Diagnostic Process.

The solution here proposed is not only to create a knowledge base directly from the human expert (or by a pool of experts) but especially to increase the own knowledge through the provision of advanced computational analysis tools capable of enhancing data, information and collaboration in order to become expert knowledge over time.

3.1.3 Semantic network

Here is applied the semantic networks representation formalism to complete the knowledge identification phase and to provide a conceptual model in input to the next phase. See Fig. 2.

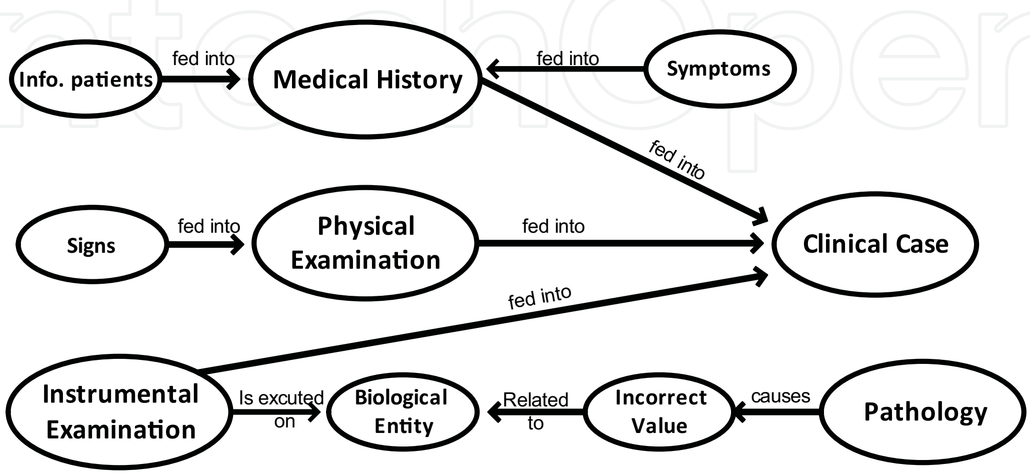


Fig. 2. Biomedical Knowledge representations using Semantic Annotations.

Nodes are objects, concepts, states while arcs, represent the relationships between nodes as carried out by knowledge identification phase. For example, the figure highlights how a pathology will lead to an alteration of the value of a biological entity and how this variation is detected by an instrumental examination. New relationships can be made clear, not only the evident ones but also those coming deducted from the father-son hereditariness. Fig. 2 suggests that the signs detected during physical examination can be directly annotated in the clinical case.

3.2 Knowledge specification

The second stage in the presented framework is to construct a specification of the k-model. In order to specify k-model of the biomedical diagnostic process it has been used an approach based on the inference and task knowledge.

3.2.1 Inference

Inference rules are used to capture the way in which a specialist reason according to the inference logical scheme. The inference rules so defined will be embedded into Decision Support System in order to develop a Knowledge-Specialist Based System. The summary of all these inferences help to develop activities that the system must make either for the development of a formal instrument describing the said inferences or the classes studied in the representative formalism. In Fig. 3 is represented the logical schema of the classes that describes some rules of inference such as "*the observation of a symptom strictly depends on the patient who is suffering from it*".

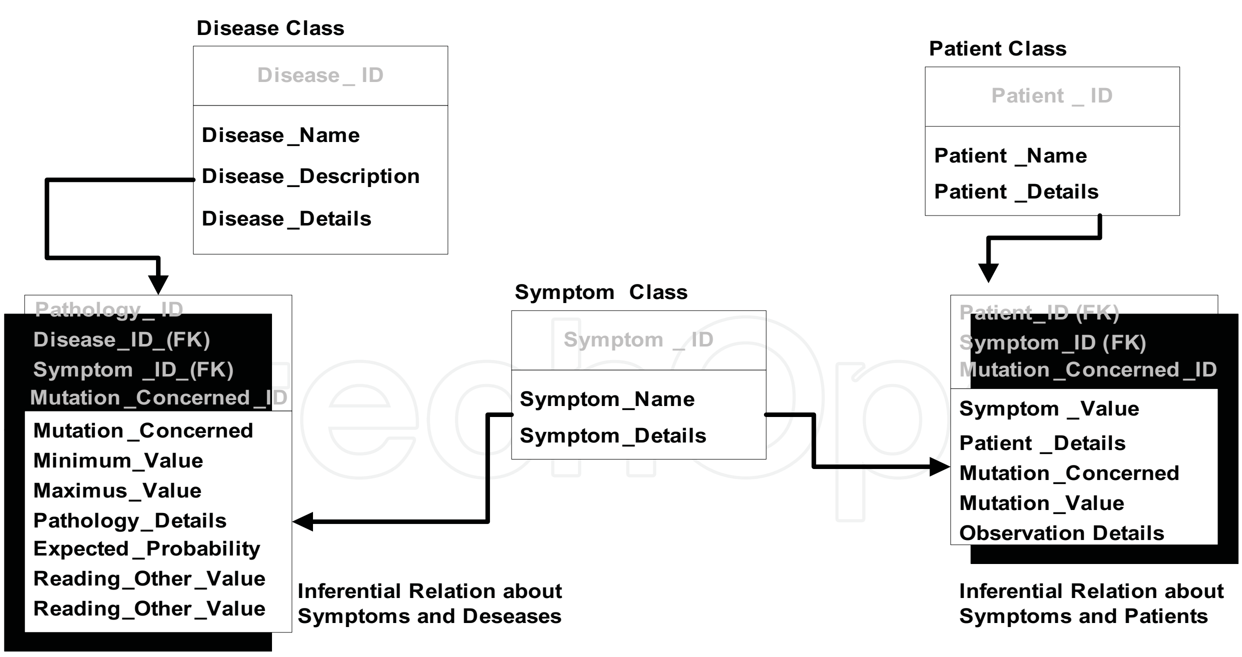


Fig. 3. Logical Schema based on Classes to represent the Inference component of the K-model.

The dependence of the activities of a system from the inferences rules can be translated both into obliged passages or into a series of guide-line on designing the activities of the system to be followed. Many informatics system translate it into an "algorithm" called "inferential engine" which can complete the design phase of the activities automatically. In some cases a prototype of the inferential engine is build up. In the implementation phase it will be possible to choose whether developing a suitable software or maintaining that engine at a conceptual level. Fig. 4 shows a conceptual description of biomedical diagnostic inferential engine, composed by the following elements:

- **Interpreter**: it decides the rule to be applied (meta-level);
- **Scheduler**: it decides the execution order of the rules (object level);
- **Job memory**: it contains a list of developed operations is (object level).

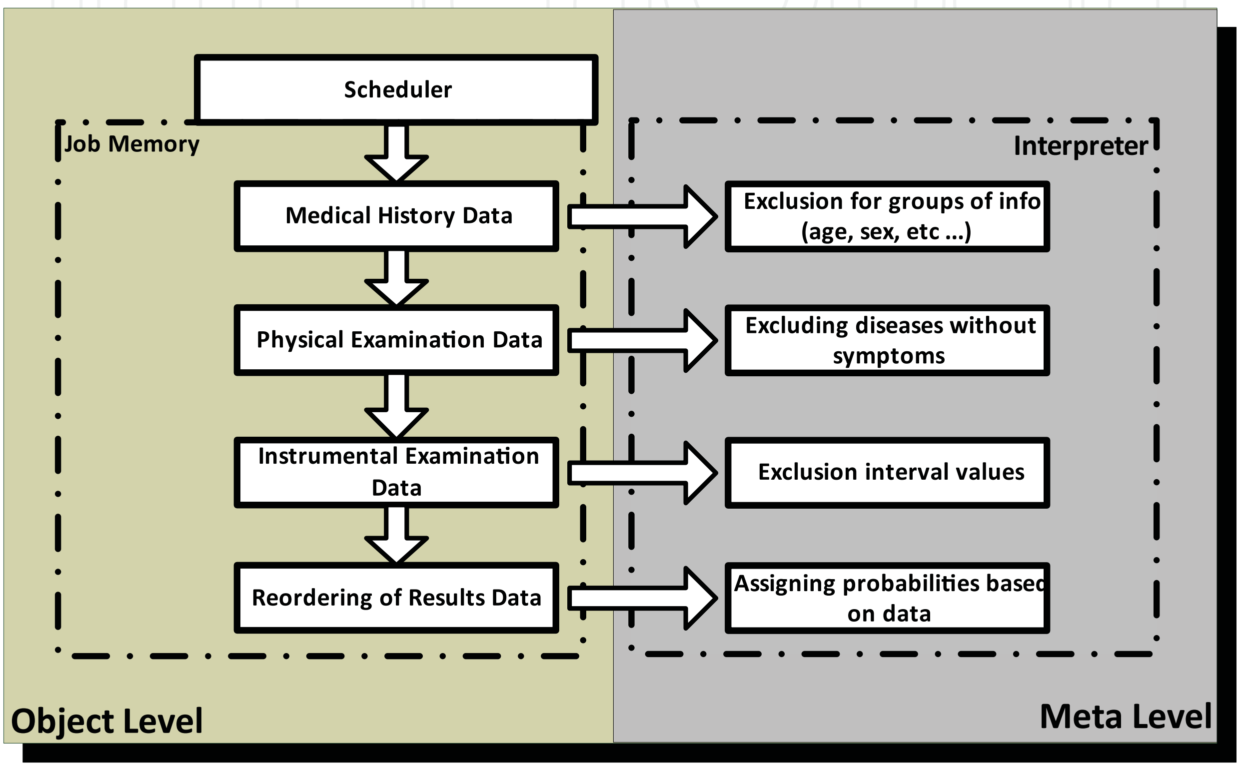


Fig. 4. A Conceptual Description of the Biomedical Diagnostic Inferential Engine.

3.2.2 Task analysis

To design an efficient and technologically advanced decisional system of support, the correctness of the applied methodology and rightness of the information must be considered. A good working system is based not only on accurately selected and organized data but also on a model epistemologically adherent to the everyday medical actions. The task design occurs at the purpose. Using the "top-down approach" the analysis and the task design starts from the macro-activity or main task designed for the solution about the problem that regard the support to the specialist decision.

The Fig. 5 shows the main task of the solution composed by both Central System and Central Db components. Different Database are referred by External and Internal Db components.

The kind of result proposes a **form** which gives back a list of pathologies with the relative probabilities from the letting in of a list of symptoms inserted by the specialist.

The solution is an automatic computer system based on biomedical knowledge that should be able to support the specialist using updated information and supporting decisions by

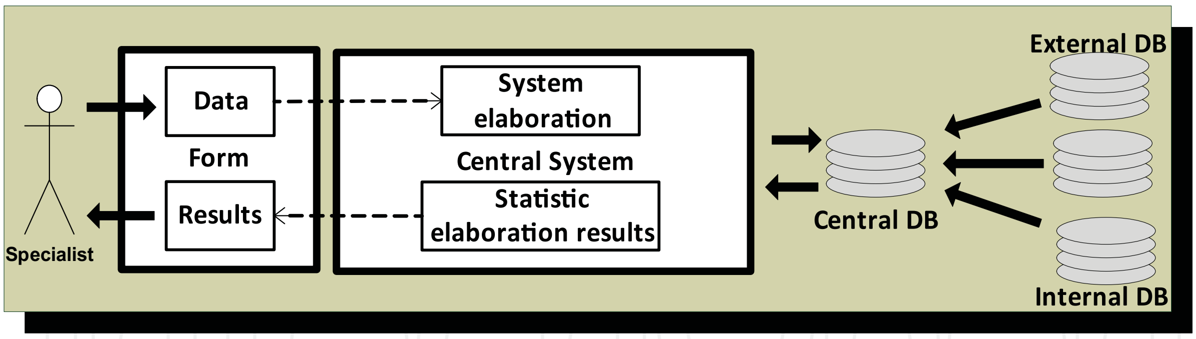


Fig. 5. A representation of the Main Task based on Functional Blocks and Components.

computational formal approach. The execution of a test implies a sequences of steps, each of them contributing to the achievement of the purpose. Task Analysis can be carried out using or the Descriptive approaches that describes the system organization in an analytical way or the Applicative approach to obtain single and simple elements, so that they can be studied. Task analysis starts from the study of the activities composing the main activity to the define the real phase, they will be translated into task system. Fig. 6 shows a scheme both to understand the succession of the activities and to determine the hierarchies.

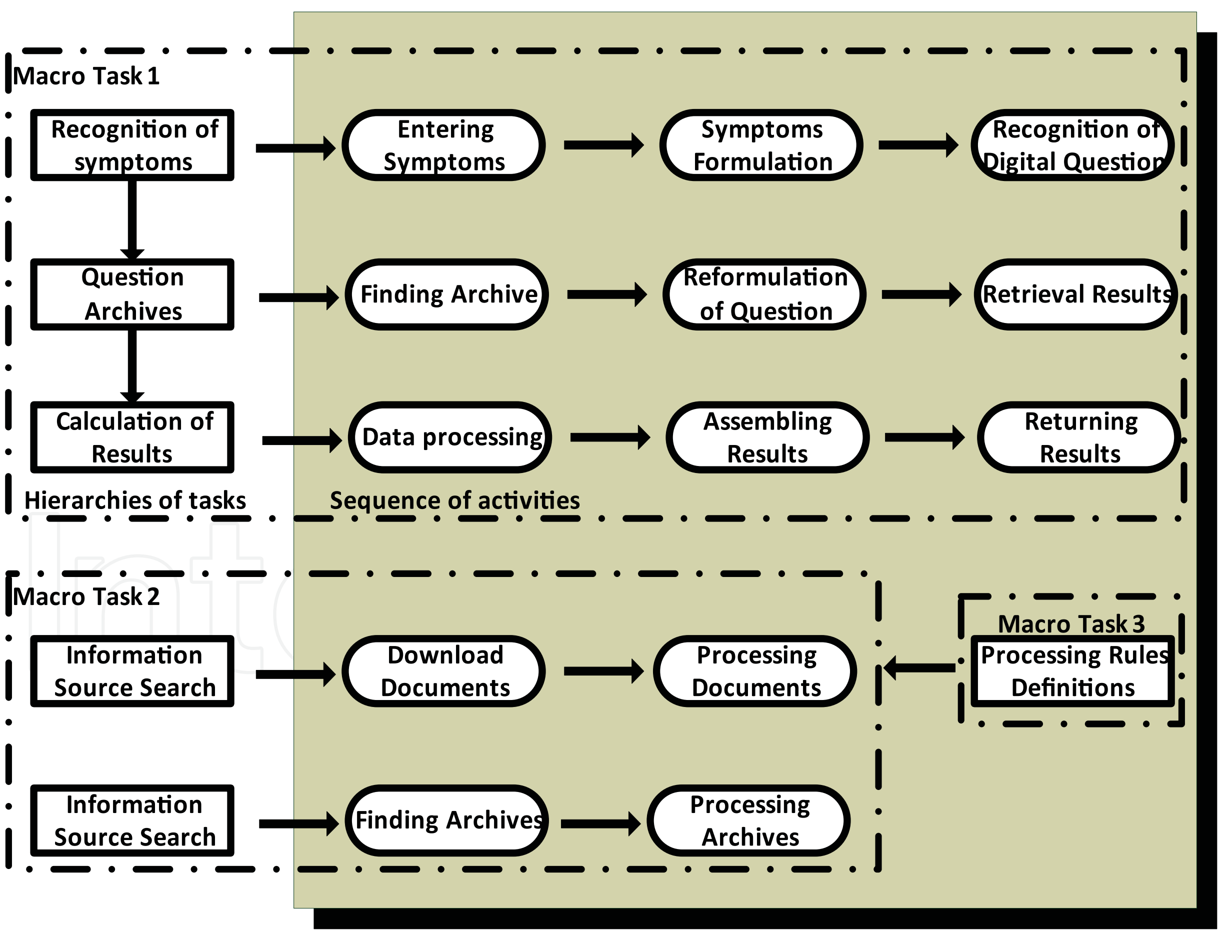


Fig. 6. Task Analysis and Design Results: hierarchies of tasks and task composition by sequences of Activities using building blocks based tools.

The results of the single task design is shown in Table 7.

Task Model	Knowledge Item Worksheet
NAME	Insert the symptom
POSSESSED BY	User terminal
USED IN	Main form
DOMAIN	It can be found both in the central system and in a remote terminal
PEOPLE	Users

Table 7. An example of Knowledge Specification for the Task Model in Diagnostic Process.

3.3 Knowledge refinement: analysis of information sources and data types

In this application the Knowledge Refinement phase is carried out to complete the Knowledge modeling. At the purpose an analysis of the additional sources of information is here presented as suggested by the task analysis. The biomedical knowledge is mainly organized in order to solve problems in the formulation of diagnosis and therapies. A problem that the specialist has to be solved can be classified according to the complexity of the diagnostic process: in the more simple case is available the useful knowledge that leads to a disease individuation directly; complex problem involves situations in progress about the disease knowledge so that not all the reasoning have been produced and some of these can be hidden in a large quantity of data. In this cases different thinking schools lead to different solutions and then the specialist makes a decision on heuristic knowledge. From the information point of view the scientific literature, medical documents, experts consultations and forums, can be cited as the main source of information. In particular the digital information can be organized in databases, glossaries and ontologies. Table 8 shows the different source of knowledge with their data types.

Information Source	Data Types
Archives of known and soluble cases	Structured data organized into databases.
Treatises of research and scientific publications	Not structured data in a textual form, with some metadata.
Social Networks and Forums	Heuristic data.

Table 8. Information Sources of Knowledge for the Diagnostic Process.

In information technology, the terms database indicates archives structured so as to allow the access and the management of the data (their insertion, delete and update) on account of particular software applications. The database is a whole of data parted according the arguments into a logical order and then these arguments are divided into categories. A database is much more than a simple list or table. It offers the possibility to manage the data allowing their recovery, arrangements, analysis, summaries and the creation of the **report** in a few minutes. However a more thorough consultation of the data can be carried out with the most advanced technologies such as data mining.

The activities of research provides large knowledge patrimonies. Several organizations works to organize patrimonies and accesses to them (i.e. **NLM, BMC**). They are available on different portals and catalogued into technical-scientific disciplines, geographical areas and, sometimes, into strategic programmers. Thanks to that cataloguing it is possible to make use of Information Retrieval techniques, which allows any system the automatically finding of scientific treatises. Nevertheless the information structures discussed up to now do not allow direct extraction of knowledge from data or among documents, at the purpose text mining technology could be used.

Another very interesting source of knowledge is the one coming from socialization of experiences in describing and persistent stores: like forum, social network, web pages, FAQ and any other place used for exchanging ideas and opinions. These tools have the ability to make explicit the tacit knowledge of individuals through socialization. In fact, these virtual places, users describe their experiences in making themselves available to compare with others, increasing their level of knowledge and acquiring new skills. The Social Knowledge is a flexible approach that support the specialist to make our decisions.

4. Designing the knowledge based system

Decision Support Systems are here considered to address the design of the solution to support the specialist in the diagnostic process.

4.1 Decision Support System

A system which aims is to support the user to make a decision, is called Decision Support System (DSS). That is a system which makes available tools in order to increase the efficiency of the decisional process (G Desanctis, RB Gallupe, 1987). Fig. 7 shows a DSS high level architecture.

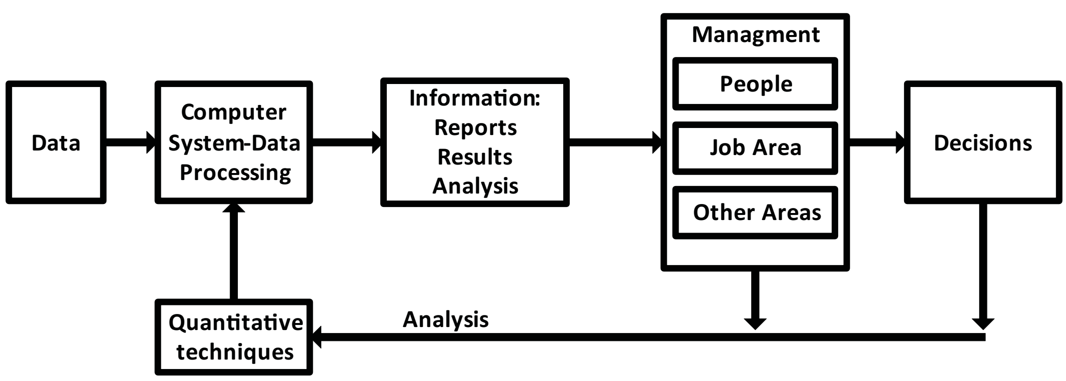


Fig. 7. A Scheme of a Computer Based DSS reference architecture.

Fig. 8 shows a structure of decision-making process over the different forms of knowledge possessed by the specialist. In particular, it describes the steps taken by decision makers from the most simple, based on data held, then the decision maker to proceed in its decision path will refer in order to other information that will acquire and then on explicit knowledge domain and finally on heuristic knowledge.

A "decision" is a definitive choice among incompatible options. The clinical decision are always temporary and can be modified as the clinical case alters. In many cases such a decision must be taken among different options and it can be often followed by further diagnostic or therapeutic decision, which have the same difficulty level of the first one.

Fig. 9 shows a decisional clinical tree. It is a schematic representation of the logical and temporal structure, of a clinical situation, in which it is necessary to take one or more decisions.

Based on the decision tree clinical model, different clinical strategies can be descrived by spefying the action which must be made for every possible situations. The analytical decisional structures so produced have a lot of advantages in comparison with the intuitive method to make clinical decisions.

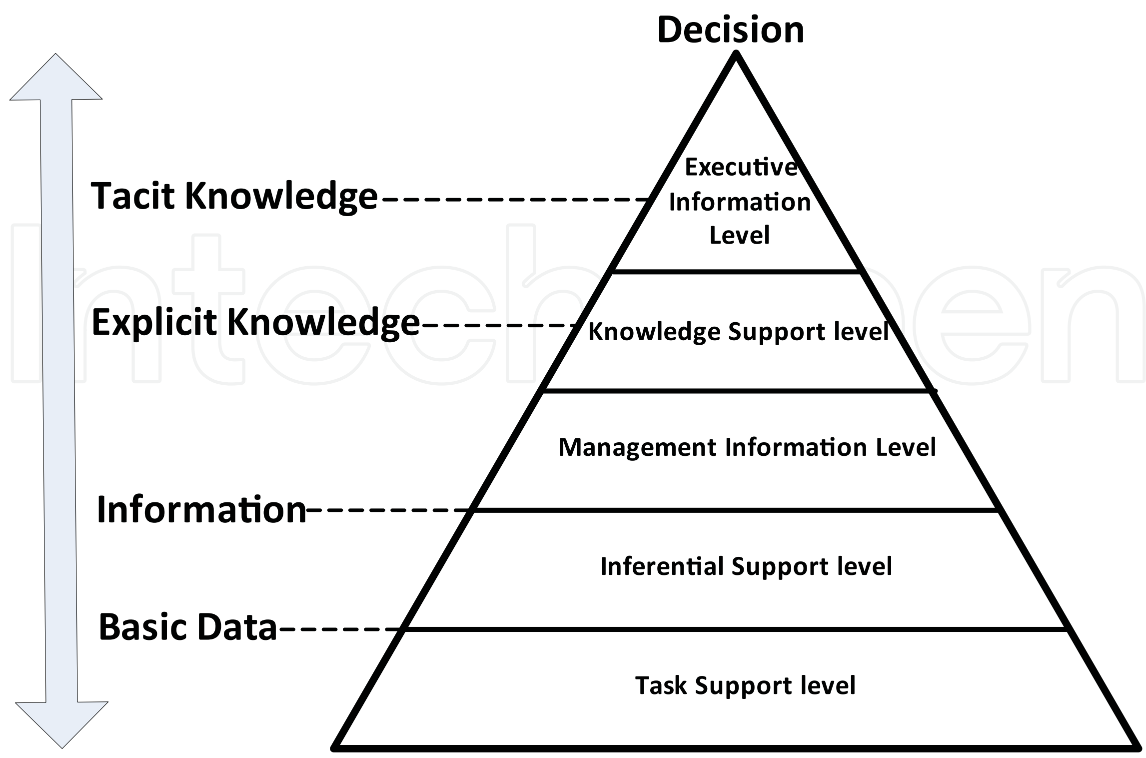


Fig. 8. Hierarchical Relationship among Information, Knowledge and Decision.

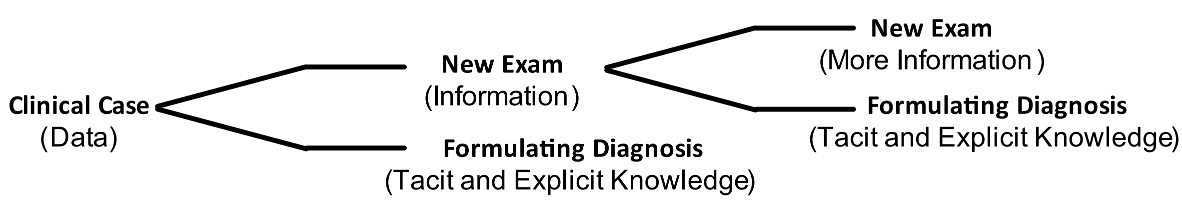


Fig. 9. A Decisional Clinical Tree Diagram to formulating Diagnosis.

Among these advantages there is the posibility to concentrate on a single aspect of the problem, always taking into account the whole. Another advantage is the analysis of the decisions which obliges the specialist to consider the relation between the acquired information and the subsequent decision. An example of analytical decisional structure is showed in Fig.10.

4.2 System structure

Fig. 11 shows the knowledge-based system architecture. The **K-Engineer** collects, analyzes and formalizes the knowledge producing a knowledge model, he is trained by the Domain Expert; the **User** (Biomedicine Specialist) will use the system and says what he would like and if he is satisfied and the **K-maintainer** provides the resources and use conceptual tools for the knowledge system updating.

The DSS examined has some functional elements which use conceptual instruments and resources, such described in Table 9. The **Knowledge Core Layer** contains all the procedures necessary to solve the problem of the user. To carry out a knowledge based system must be utilized precise procedures in order to identify the relationships between apparently

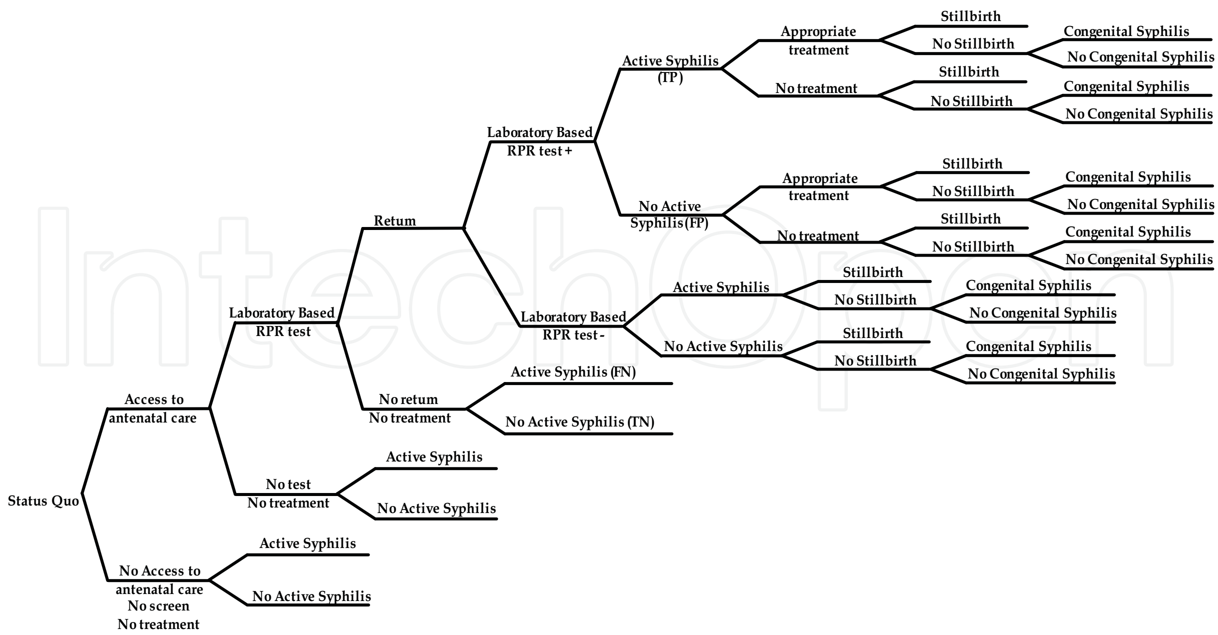


Fig. 10. An Analytical Decisional Structure.

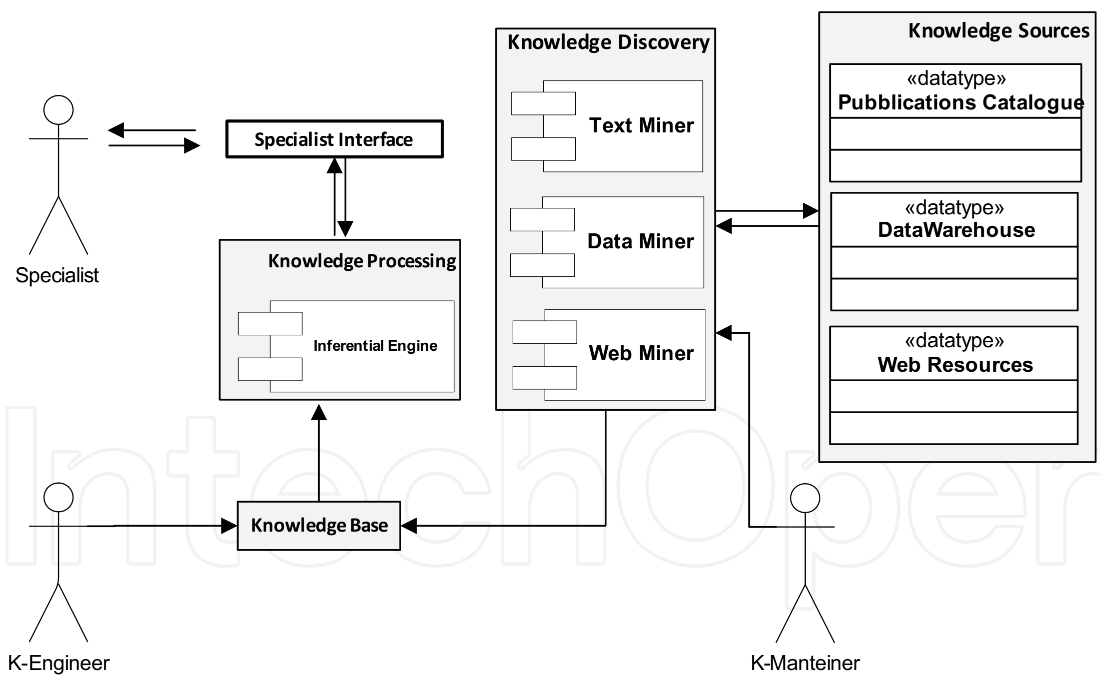


Fig. 11. DSS scheme of the organization.

independent variants, like sign and symptoms, and dependent variants like possible pathologies. The procedures can be obtained from various forms of decision tree. These relationships, stored in the **Knowledge Base,** use pre-arranged rules (if-else if-then) to drive the decision in the diagnostic process. Another functional elements are the **Knowledge Engine** whose perform automatic activities like the analysis of the texts and data; and finally the **Knowledge Sources**.

Functional elements	Conceptual Tools	Resources
Knowledge Discovery	Web mining, Data mining, Text Mining	Web Miner, Data Miner, Text Miner
Knowledge Processing	Statistical Procedures, Diagnostic Patterns	Inferential Engine
Knowledge Base	Knowledge Rules	
Knowledge Sources	Database Management Systems	Publication Catalogue, Web Resources, DataWarehouse

Table 9. List of functional elements.

The **Web Miner** is the element for the knowledge discovery from the web: it applies procedures similar to Data Mining to extract information from the resources present in the web. The computer paradigm is the Web Mining which derives from the Data Mining and deals with the discovery of resources on the web. It analyses web pages and from them extracts the resources of biomedical literature though query. The **Text Miner** has two objects: the transformation of the documents into a representation fit for the learning algorithm (the so called preprocessing phase) and the grammar analysis on the non structured text to extract the structured knowledge. As to processing phase the Text Miner draws out the meta information contained in the scientific texts and uses it to catalogue them in the Publication Catalogue where the document will wait for the analysis accomplishment according to the Text Mining algorithm . The grammar analysis starts after an K-Maintainer call which gives the Text Miner the parameters necessary to execute the analysis parameters contained in a document XML called Analysis Engine. The structured information will be memorized inside the Knowledge Repository. The **Data Miner** extracting the knowledge from the data the Text Miner got through automatic method, and elaborating them according to the K-Maintainer selected procedure. It is interfaced with the Publications Catalogue and makes use of the Data Mining Paradigm, which provides the instruments to locate, to extract and transform the data and the instruments for the management of data. The **Inferential Engine** is made by: an interpreter which decides the useful rules to request to the Knowledge Base, and a scheduler which organizes the rules to get developed and their execution order. The engine goal is to extract the rules for the problem solving according to an activation and recognition of the same, executing an exam of the rules on the ground of knowledge select the right rule. It is organized in two phases: a work memory (or blackboard) where the general plan is memorized and a workflow engine of what to do, besides a description of the solutions conjectured up to now. In the inferential engine there is also the consistence enforcer; this is a schedule which, once conjectured an hypothesis, looks for evidences to enforce or eliminate the hypothesis. Thus if the inference is a single deduction that the system can get from the premises (for example a single calculation), the inferential engine is the instrument fit to determinate the inferences complex. The **DataWarehouse** contains the information drawn out from the operations of text mining, whose structure is just like the semantic networks model one and is fit for the analysis algorithm of the Data Miner. It contains also the structured information from any external databases that the K-Maintainer wants to integrate into the system. The **Web Resources** is the database that contains the extracted information from the resources present in web. The **Publication Catalogue** is the repository of the scientific literature founded.

4.3 Grid-enabled knowledge application

In the field of the computer systems for the bio-medicine, the case studies have a long list of electronic documents about their health status, e.g. lab exams, diagnostic images, medical records, several special texts or even the letter of hospital discharge. Such data and information, since they are sector based, cannot be integrated or read by different systems tools even if digitalized, sometimes also because of the lack of appropriate communication infrastructures among the different systems. Even though they can be rebuilt or transferred on a similar computer platform, these documents cannot be reached from all researchers or specialists working on the same case study, just because such documentation has been kept in several autonomous computer systems. One of the major obstacles in their integration is the lack of a unique standard, yet also the different kinds of information products in the health world, the technological gap. The research and the agreement on a standard is even more important to the re-elaboration of the information acquired from the various information resources, balancing various elements (such as local resources, local networks, Internet) and employing various communication technologies, protocols and switching and filtering systems. Besides these drawbacks, the approaching techniques analyzed in the previous sections, such as the text mining or the statistic algorithm, when extended to a great amount of resources, require a huge computational power, a large broadcast band, and inner organization which can also make possible the interoperability. The solution proposed up to now can be really useful if an advanced technology based on computer systems is able to exceed computer networks and organizational problems that often make it difficult to share data and information. The Fig. 12 refers to a scenario of distributed resources all potentially very useful for continuous improvement of expert knowledge and indicates a solution based on grid computing. A worthy approach to use and share the resources and which includes computing power, mass memory, and a large broadcast band, comprise a network of information instruments called grid, whose main characteristic is its dynamism, i.e. the ability to re-configure itself in any moment, encompassing new knots and tolerating the loss of others, and capable to adapt to any required work every time it is necessary. All these operations are automatically performed, in a completely transparent way for the user, who will perceives the grid as a unique, stable, unchangeable, effective and always accessible system.

The grids (or computational grids) are advanced services which allow to the users to share computational resources, archival space, scientific data and instruments among different computers. They are then a valid instrument to automatically store the accomplished researches and spread the knowledge to the domain. The grid technology, indeed, facilitates data downloading, arranges documents, and periodically updates the researches.

The grids (or computational grids) are advanced services which allow to the users to share computational resources, archival space, scientific data and instruments among different computers. They are then a valid instrument to automatically store the accomplished researches and spread the knowledge to the domain. The grid technology, indeed, facilitates data downloading, arranges documents, and periodically updates the researches.

To define a grid architecture valid for the activities involved by the knowledge model, is necessary to take into account that it is not only a architectural model to provide computational resources for applications with great need of processing, but it is also an infrastructure where can connect and unify heterogeneous resources issued on local or geographical scale. The interoperability becomes the main problem to solve; indeed, it is necessary to allow the interaction among different platforms, languages, and research areas. In an interconnected zone, the interoperability implies common authentication protocols, therefore the grid to implement needs an architecture of protocols first of all. In this way it is possible to define a basic system through which users and knots can communicate.

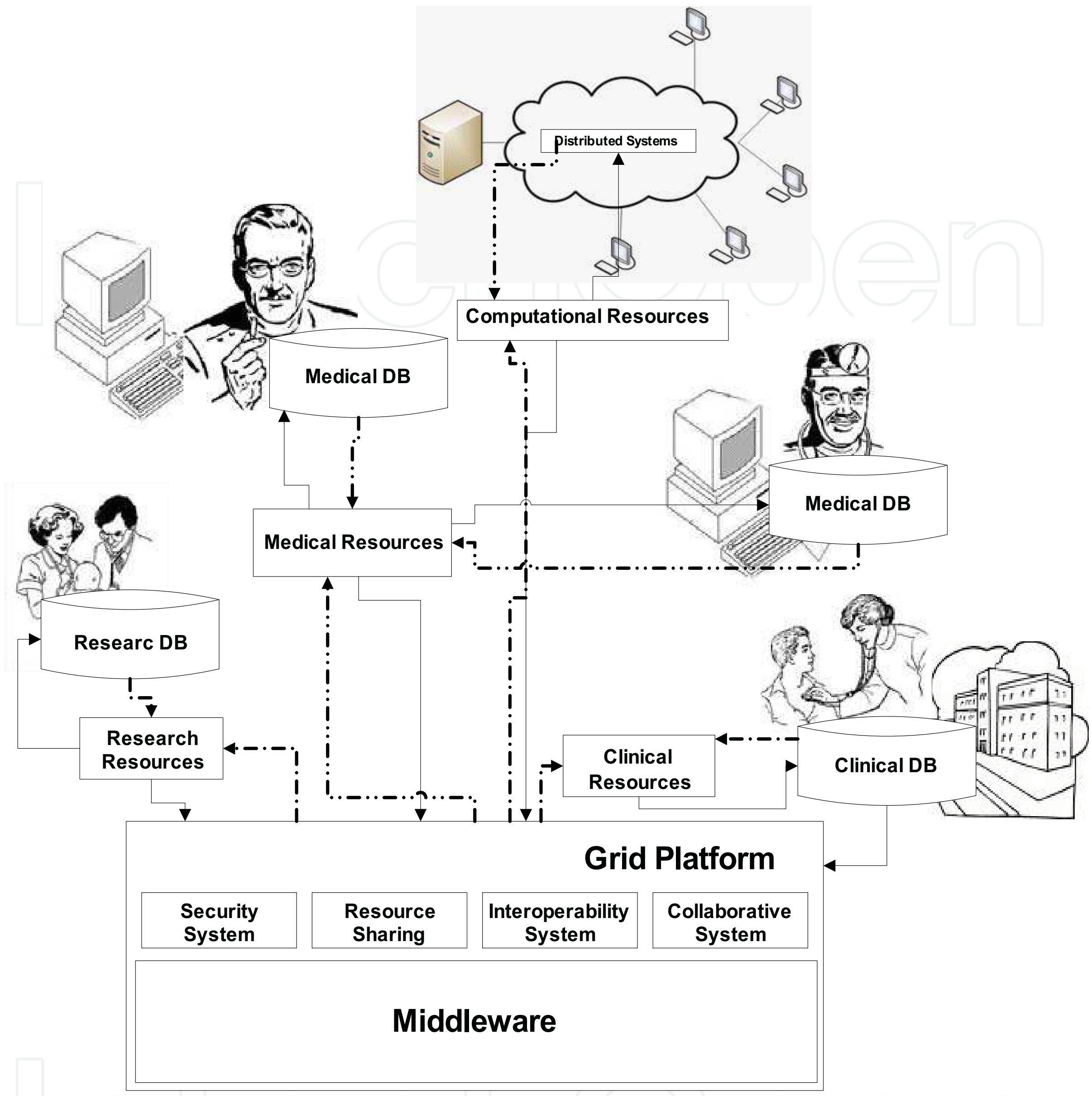


Fig. 12 Grid Computing.

A Grid system allows to manage its own resources to obtain different service qualities, such as response time, availability, security, and performance to satisfy users request. The quality of service (QoS) is a highly important parameter, but it is not the only one. An information system that has to solve highly complex problems, such as the management of the biomedical knowledge, can take other advantages of the grids. The scalability is one of these advantages: a grid can comprise millions of resources, implying a potential degradation of services. For this reason the grid-enabled application should be able to deal with connection latency on geographical distances. Dynamicity and adaptability are criteria that determine the reliability. In a grid a breakdown is a rule not an exception, since the great number of resources increases the chances of breakdown. This technology has been designed as a very promising application-level specialization and several proposals have been made of the same. (see for example Castellano at all, A bioinformatics knowledge discovery in text application for grid computing; A Workflow based Approach for Knowledge Grid Application; Intrusion Detection Using Neural Networks: A Grid Computing Based Data Mining Approach; Biomedical Text Mining Using a Grid Computing Approach)

5. Results and discussions

This chapter deals with a particular class of biomedical engineering systems such as Computer-Based Decision Support Systems. Different DSS applications are proposed in the literature to date. They highlight the system performances in the biomedical field, without examining the methodological aspects on which the investigation was based. Since these systems are of considerable complexity, their development and adoption are a problem in the field of biomedical engineering. In this chapter in response to the problem just mentioned a framework for analysis and design of these systems is discussed. It is based on a methodology of knowledge engineering. An application to a diagnostic process in the biomedical field is also presented.
Methods for problem solving vary from ad hoc to highly structured. Ad hoc approaches are often used to solve simple problems in a linear way. Highly structured methodology for studying complex problems become even more formal where it is usual to pre-specify all the system requirements. In this chapter has been used a problem solving highly structured life cycle based on the schema shown in Fig. 1. It has been used to carry out analysis and design phase: the analysis to gain an understanding of an application domain with its problem and the relative conceptual solution; the design to specify a system that meet these requirements. Although this chapter focuses on the analysis and design stages applied to the problem solving life cycle, it is said that there is a further phase of building systems in which one focuses on the actual construction of the system by an executive project.
Fig. 13 summarizes the results from the life cycle of problem solving both for the analysis and for design.

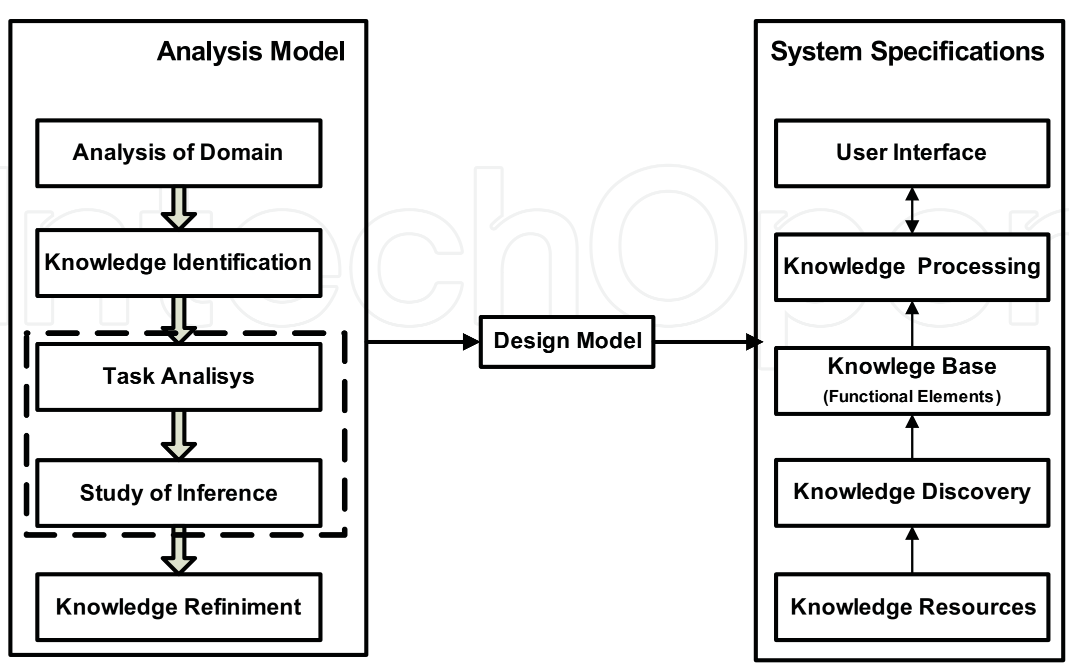


Fig. 13. Analysis and Design for the Biomedical Diagnostic Process.

The designed system is made up of any number of components. Each one carries out part of the system function. They are important because can help to handle systems complexity and then improve understanding of a system. Components communicate between them by passing messages. A good system will made up of highly independent components with minimal flows between them.

The components are identified according to the encapsulation strategy. It allows to hide the internal working of procedural elements so as to protect the other parts of the system from the changes which would occur in it, if this working was faulty, or if one decided to implement it in a different way.

From the technological solution point of view, the decision support system functional components aims to widen the human being capabilities. For example the automatic retrieval of scientific literature is important for the performances of the system. In fact the web suggest several scientific documents and that slows down the manual research. Moreover the use of automatic equipments is an advantage for the texts research, their readings and the knowledge extrapolation, which represent a study more similar to the one made by an expert, with reduced times, big quantities of examined texts and a lasting memorization. However when dividing the tasks into simple activities, it is necessary to respect the main rules of the biomedicine knowledge to develop a functional knowledge based automatic system. Algorithms and data will be effective only if the study about the knowledge domain has been clearly made.

Design decisions have specified functional blocks and their interaction needs to solve the problem. All of them can be regarded in terms of computational resources such as algorithms, data and computational power. As a consequence design phase must be specify also a computing model resource organization in order to satisfy user requirements and system performances. A grid computing based model has been discussed in order to support a knowledge based system development with high quality of services, based on interoperability, resource sharing, security and collaborative computing. Others computational aspects that are left open during analysis such as representational formats, computational methods used to computing inferences, dynamic data storage and the communication media, could be taken into account also.

Finally must be considered that the study presented in this chapter shows a methodological solution or framework focused on knowledge to develop a knowledge based system very useful in biomedical engineering and the application here presented is a case study to examine the way in which the framework works.

6. Conclusions

Different Decision Support System applications are proposed in the literature to date highlighting the system performances in the biomedical field, without examining the methodological aspects on which the investigation was based. Since these systems are of considerable complexity, their development and adoption are a problem in the field of biomedical engineering. In this chapter has been presented a tool for Biomedical Engineering based on a framework of Knowledge Engineering specialized into develop applications of decision support computer systems. The framework has been applied successfully to the Biomedical Diagnostic Process. The framework application explores technologies for knowledge applied to the processes and resources enhancing data, information, collaboration and computing models in order to produce new expert knowledge over time. The study presented in this chapter has shown that new knowledge in the digital age can also be produced automatically and that instead the knowledge produced by human processes can be effectively supported by specific digital resources (or knowledge resources). To this end, it was revealed that the application of distributed computing to these problems based on knowledge reveals structural characteristics such as interoperability, resource sharing, security and collaborative computing. At the aim the application here presented has proposed the grid computing approach.

7. References

Achuthsankar ,S. Nair, (2007). Computational Biology & Bioinformatics, A gentle Overview, Communications of Computer Society of India

Almeder, Robert, (1998). Harmless Naturalism, The Limits of Science and the Nature of Philosophy. Chicago and La Salle, Open Court.

Alston, William, (1989). Epistemic Justification, Essays in the Theory of Knowledge, Ithaca, Cornell University Press

Anderson MF, Moazamipour H, Hudson DL, Cohen ME. (1997). *The role of the Internet in medical decision making*, Int J Med Inform, University of California, VA Medical Center, USA

B. Kraft, N. Wilhelms, (2005). *Visual knowledge specification for conceptual design,* L. Soibelman, F. Pena-Mora (Eds.), Proceedings of the 2005 International Conference on Computing in Civil Engineering, pp. 1-14.

Baldi P. ,Brunak S., (2001). *Bioinformatics: The Machine Learning Approach*, MIT Press.

Barnes M.R., Gray I.C., (2003). *Bioinformatics for Geneticists, first edition.* Wiley

Berner E.S., Webster G.D. (1994) Shugerman A.A. et al. *Performances of four computer based diagnostic systems, Medical software's free future*, NEJM, pp. 1792-96

BMC, BioMed Central, The Open Access Publisher, www.biomedcentral.com

Bohem B.,(1988). *A Spiral Model of Software Development and Enhancement*, pp. 61-72.

Böhlen B., Jäger D., Schleicher A., Westfechtel B. (2005) *A Framework for Building Graph-Based Interactive Tools*, Electronic Notes in Theoretical Computer Science, Pages 91-101

Chittaro L., Guida G., Tasso C. e Toppano E., (1993). *Functional and Teleological Knowledge in the Multimodeling Approach for Reasoning about Physical Systems*, A Case Study in Diagnosis, IEEE Transaction on Systems, Man and Cybernetics

Daniel Brown. (2004), *Mastering Information Retrieval and Probabilistic Decision Intelligence Technology*, Chandos Publishing, Oxford

De Kleer J., Williams B., (1987). Diagnosing multiple faults. Artificial Intelligence pp. 97-130

Di Hai Hu,Michael Liebman,Richard J. Mural, Michael N. Liebman, (2008). *Biomedical informatics in translational research*

E. Gamma , R. Helm , R. Johnson , J. Vlissides, (1995). *Design patterns: elements of reusable object-oriented software,* Addison-Wesley Longman Publishing Co., Inc., Boston

Eisfeld M., Scherer, R. (2003) *Assisting conceptual design of building structures by an interactive description logic-based planner*, Advanced Engineering Informatics, pp. 41-57.

G Desanctis, RB Gallupe, (1987). *A foundation for the study of group decision support systems*, Management science, JSTOR

G. Carraro , Y. E. Kalay, (1994). *Knowledge-Based Computer-Aided Architectural Design*, Elsevier Science Inc., New York, NY

G. Schreiber, H. Akkermans, A. Anjewierden, R. de Hoog), N. Shadolt, W. Van de Velde, B. Wielinga. (2000), *Knowledge engineering and management: the CommonKADS methodology, he MIT Press, Cambridge, MA, 2000, ISBN: 0-262-19300-0, US*

I. Flood, (2003), *Proceedings of the Fourth International Symposium on Information Technology in Civil Engineering,* ICCC, ASCE, , pp. 1-12

K. Krauter, R. Buyya, M. Maheswaran, (2001). *A Taxonomy and Survey of Grid Resource Management System for Distributed Computing, Software, Practice and Experience*

Knowledge Engineering: Practice and Patterns, 16th International Conference,

Kraft B., Schneider G., Parameterized. (2003), *Specification of conceptual design tools in civil engineering, Proceedings of the Second International Workshop on Applications of Graph Transformation with Industrial Relevance*, AGTIVE, ser. LNCS, vol. 3062. Springer, Heidelberg. pp. 90-105.

Kraft B., Schneider G. (2007), *Semantic room objects for conceptual design support: a knowledge-based approach*, Proceedings of the 11th International Conference on Computer Aided Architectural Design Futures, pp. 207-216.

Kraft B., Wilhelm N. (2007), *Interactive distributed knowledge support for conceptual building design*

Beucke K. Firmenich B. Donath D. Fruchter R. Roddis K. (2007), *Interactive distributed knowledge support for conceptual building design*, Proceedings of the 10th International Conference on Computing in Civil and Building Engineering (ICCCBE-X), Bauhaus-Universität Weimar. pp. 1-14

M. Eisfeld, (2005). *Assistance in Conceptual Design of Concrete Structures by a Description Logic Planner*, Dissertation, TU Dresden, Kassel

Marburger, A. and Westfechtel, B., Behavioural (2003). *Analysis of telecommunications systems by graph transformations*, Proceedings of the Second International Workshop on Applications of Graph Transformation with Industrial Relevance, pp. 202-219.

M. Castellano, C. Digregorio, (2009). *A Workflow based Approach for Knowledge Grid Application*, KMIS 2009, pp. 230-235

M. Castellano, G. Mastronardi, G. Tarricone. (2009), *Intrusion Detection Using Neural Networks: A Grid Computing Based Data Mining Approach,* ICONIP, pp. 777-785

M. Castellano, G. Mastronardi, R. Bellotti, G. Tarricone. (2009), *A bioinformatics knowledge discovery in text application for grid computing*, BMC Bioinformatics

M. Castellano, G. Mastronardi, G. Decataldo, L. Pisciotta, G. Tarricone, L. Cariello (2008). *Biomedical Text Mining Using a Grid Computing Approach,*. ICIC, pp. 1077-1084

Marcia C. Linn, (1993). *Designing computer learning environments for engineering and computer science: The scaffolded knowledge integration framework*, Journal of Science Education and Technology

Maver T. (2006), *A number is worth a thousand pictures,* Automation in Construction, pp. 333-336.

Meniru, K., Bédard, C. and Rivard, H. (2002). *Early building design using computers*, Proceedings of the Conference on Distributing Knowledge in Building, Aarhaus School of Architecture, Dänemark.

Mora R. Bedard C. Rivard H. Beucke K. Firmenich B. Donath D. Fruchter R. Roddis K. (2007), *A framework for computer-aided conceptual design of building structures*, Proceedings of the 10th International Conference on Computing in Civil and Building Engineering (ICCCBE-X), Bauhaus-Universität Weimar. pp. 1-12.

NLM, U.S. National Library of Medicine, www.pubmed.org

O. Bodenreider, (2008). *Biomedical Ontologies in Action: Role in Knowledge Management, Data Integration and Decision Support*

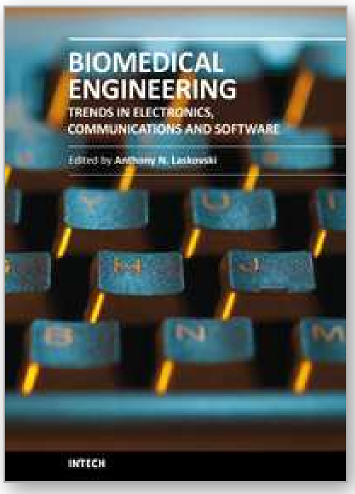
R.D. D. Coyne , M. A. Rosenman , A. D. Radford , M. Balachandran , J. S. Gero (1989) *Knowledge-Based Design Systems,* Addison-Wesley Longman Publishing Co., Inc., Boston, MA

S. Chakrabarti. (2003), *Mining the Web: discovering knowledge from hypertext data*, Elsevier, ISBN 1558607544

Stiny, G., Gips, J. (2006), *Shape grammars and the generative specification of painting and sculpture,* Information Processing. pp. 1460-1465.

Sulaiman M.J., Weng N.K., Theng C.D., Berdu, Z. (2002), *Intelligent CAD checker for building plan approval*, Christianson, P. (Ed.), Proceedings of the Conference on Distributing Knowledge in Building, Aarhaus School of Architecture, Dänemark.

William, Douglas, Clinger, (1981). *Foundations of Actor Semantics,* Lab Technical Report



Biomedical Engineering, Trends in Electronics, Communications and Software
Edited by Mr Anthony Laskovski

ISBN 978-953-307-475-7
Hard cover, 736 pages
Publisher InTech
Published online 08, January, 2011
Published in print edition January, 2011