ELSEVIER

Artificial Intelligence in Medicine 27 (2003) 335–366

Artificial Intelligence in Medicine

www.elsevier.com/locate/artmed

# A multi-agent intelligent environment for medical knowledge

Rosa M. Vicari[a,*], Cecilia D. Flores[a], André M. Silvestre[a], Louise J. Seixas[b], Marcelo Ladeira[c], Helder Coelho[d]

[a]Informatics Institute, Federal University of Rio Grande do Sul, Caixa Postal: 15064 91501-970, Porto Alegre, Rio Grande do Sul, Brazil

[b]Post-graduate Course on Computer and Education, Federal University of Rio Grande do Sul, Porto Alegre, Rio Grande do Sul, Brazil

[c]Computer Science Department, University of Brasilia, Brasilia, Brazil

[d]Informatics Department, Faculty of Science, University of Lisbon, Lisbon, Portugal

Received 15 October 2002; received in revised form 15 October 2002; accepted 18 October 2002

## Abstract

AMPLIA is a multi-agent intelligent learning environment designed to support training of diagnostic reasoning and modelling of domains with complex and uncertain knowledge. AMPLIA focuses on the medical area. It is a system that deals with uncertainty under the Bayesian network approach, where learner-modelling tasks will consist of creating a Bayesian network for a problem the system will present. The construction of a network involves qualitative and quantitative aspects. The qualitative part concerns the network topology, that is, causal relations among the domain variables. After it is ready, the quantitative part is specified. It is composed of the distribution of conditional probability of the variables represented. A negotiation process (managed by an intelligent *MediatorAgent*) will treat the differences of topology and probability distribution between the model the learner built and the one built-in in the system. That negotiation process occurs between the agents that represent the expert knowledge domain (*DomainAgent*) and the agent that represents the learner knowledge (*LearnerAgent*).

© 2003 Published by Elsevier Science B.V.

*Keywords:* Multi-agent system; Intelligent learning environment; Bayesian networks

* Corresponding author. Tel.: +55-51-33166801; fax: +55-51-33166160.
*E-mail addresses:* rosa@inf.ufrgs.br (R.M. Vicari), dflores@inf.ufrgs.br (C.D. Flores), andre.silvestre@terra.com.br (A.M. Silvestre), seixasl@terra.com.br (L.J. Seixas), mladeira@cic.unb.br (M. Ladeira), hcoelho@di.fc.ul.pt (H. Coelho).

0933-3657/03/$ – see front matter © 2003 Published by Elsevier Science B.V.
doi:10.1016/S0933-3657(03)00009-5

## 1. Introduction

People simplify things when they want to solve problems and take decisions in real life, where information is partial (not complete) or approximated (not exact). For that kind of problem, we only get solutions full of uncertainties. A physician can diagnose a disease based on some symptoms, but the diagnosis will be only a hypothesis, which can be wrong. A mistake may come from the incomplete knowledge about the pathology in question, from determinant symptoms not detected yet, due to the initial evolution phase of the disease, or from the lack of complementary tests. However, this diagnosis has higher reliability than a simple guess, because the physician who did it has a notion on the value of the formulated model. Currently, a way of handling the uncertainty mentioned above is to offer physicians medical software based on probabilistic reasoning. These systems may be used in teaching/learning processes and in the physician's daily practice.

In this paper, we introduce AMPLIA. The proposal here is to present a multi-agent intelligent learning environment according to the following three criteria. First, the learner elaborates his/her own knowledge model and the system will continuously ask him/her about his/her actions and decisions. Second, feedback and additional information is available all the time. Third, if necessary a negotiation process among the agents and the learner will take place as a way of making him/her review his model.

For the purpose of medical education, medical students should practice two skills: hypothetical model construction and diagnostic reasoning, because both are problem-solving tasks. Firstly, apart from the diagnosis, the learner should have an opportunity to actively construct models of diseases, including the diseases possible causes, associated symptoms, and, finally, evaluate the model application. This way, the learner can acquire and use knowledge necessary in diagnostic reasoning. Secondly, the learner should have the opportunity to actively apply strategies while performing diagnostic reasoning.

AMPLIA is a multi-agent computational environment aimed at supporting learning using a constructivist approach to perform diagnostic reasoning. It is not fully developed yet, once it was initially designed only to allow knowledge modelling for decision support systems construction. Now it has received a larger function: to put available a tool that allows the learner to learn how to build a model through probabilistic networks. AMPLIA compares its domain built-in model against the one the learner has built, and if they are different, the environment starts a negotiation process, based on pedagogical strategies, in order to induce the learner to review his/her model.

The learner may further specify, evaluate, and review the model at qualitative and quantitative level, under AMPLIA advice. Training of strategies for diagnosis will also be supported qualitatively (i.e. which information is necessary in order to support an hypothesis or differentiate between two different ones?) and quantitatively (i.e. how does information gathered affect a diagnostic hypothesis? Which is the most important information to be acquired to go on?).

Besides the general criteria of this project, we chose Bayesian network approach to deal with uncertain knowledge because it is mathematically principled. Bayesian networks have been widely used all over the world to model uncertain domains [2]. Uncertainty is represented by probability and the basic inference is the probabilistic reasoning, that is, the probability of one or more variables assuming specific values giving the available evidence. Another important reason for choosing the Bayesian network approach is its two-fold feature that enables qualitatively and quantitatively domain modelling. The qualitative side is represented by the set of variables and their causal relationship. The expert can easily construct this qualitative domain model, using a Directed Acyclic Graph. The quantitative side expresses the strength of this causal relationship. It is represented by conditional probability distributions. Our group has opted for probabilistic networks to represent knowledge since 1996 [24]. This work resulted in a Ph.D. thesis in 2000 that brought out the multiple sectioned influence diagram formalism.

We follow the hypothesis that a physician engaged in medical diagnosis implicitly performs probabilistic reasoning. The physician's practice corresponds to take full advantage of the probabilistic relationship between the variables present in a Bayesian network that models the medical domain of interest. Reviews of published case studies in the domain of environmental medicine support this hypothesis [2,12,17,32,37]. More generally, there is empirical evidence that the probabilistic reasoning, when supported by Bayesian networks, corresponds closely to the human reasoning pattern [33].

**Definition (Bayesian network).** A Bayesian network is a direct acyclic graph where nodes are random variables, and arcs represent direct probabilistic dependence relations among the nodes they connect. The strength of the relationship of $X_i$ with $\mathrm{pa}(X_i)$, its parents (nodes with arcs that arrive in $X_i$), is given by $P(X_i|\mathrm{pa}(X_i))$, the conditional probability distribution of $X_i$ given its parents. The $P(X_1, \ldots, X_n)$ is the joint probability distribution of all variables. If $\mathrm{pa}(X_i)$ is an empty set, $P(X_i|\mathrm{pa}(X_i))$ is reduced to the unconditional distribution of $X_i$.

The systems mentioned below are related to our work. They represent the application knowledge domain through Bayesian networks. Concerning the teaching area we can cite CAPIT, a normative constraint-based intelligent tutoring system (ITS) that uses Bayesian networks for long-term student modelling and decision theory to select the next tutorial action [26]. Pathfinder, for lymphatic disease diagnosis, is an example of one of those systems that have success, both in professional practice and physicians training, [17,32]. Munin used to get a preliminary diagnosis of muscles and nerves diseases [2]. Child [12] was developed to support physicians of the London Great Ormond Street (GOS) Hospital making distance diagnosis of congenital cardiac diseases in new-borns. As an example of teaching environments using Bayesian networks based on learner's model, we cite the expert system for the mechanics physics, ANDES [15]. ANDES aim is to know learner's intentions and knowledge in order to help him/her in problem-solving tasks. Table 1 presents more recent examples of decision support systems that make use of Bayesian networks for medical diagnosis. Notice that some of the systems analysed are composed only of knowledge bases (like Community-Acquired Pneumonia and MammoNet [22]), while others (like Internist/QMR [27] and AMPLIA) are knowledge bases coupled with Bayesian inference engines.

In the examples cited above, the goal is to find the problem solution. The underlying educational strategy is based on how a solution involves an explanation, the identification of the problem's reason leads to the proposal of its solution. This educational strategy attends those students who like to integrate theory and practice to solve real problems. With

Table 1
Systems that make use of Bayesian networks for medical diagnosis

| Criteria | Community-Acquired Pneumonia [4] | MammoNet [22] | Internist/QMR [27] | Medicus [11,28] | AMPLIA[a] |
|---|---|---|---|---|---|
| Qualitative development | Variables extracted from retrospective study with 32,662 patients; interviews with medical experts | Medical literature reviews; statistic medical databases; expert interviews | Cases abstracted from SAM[b] database | Built-in knowledge acquisition facilities | Medical literature reviews; expert interviews |
| Quantitative development | Machine learning algorithm for Bayesian networks | – | Cases abstracted from SAM database | Built-in knowledge acquisition facilities | Interview with experts |
| Accuracy testing method | Case-control study with real patients | Case-control study with patients from mammographic atlas | Sample from SAM cases | NA | Case-control study with real patients[c] |
| Metrics for testing | Sensibility, sensitivity, ROC curve, positive predictive value | Sensibility, sensitivity, ROC curve, positive predictive value | Wilcoxon signed-rank test | NA | Sensibility, sensitivity, ROC curve, positive predictive value[d] |
| Pedagogic tool | No | No | No | No | Yes |
| Diagnostic purpose | Yes (knowledge base only) | Yes (knowledge base only) | Yes | Yes | Yes |
| Agent oriented | Third party inference engine | Third party inference engine | No | No | Yes |

NA: not applicable; (–) information not available.

[a] Four networks in database.

[b] Scientific American Medicine Continuing Medical Education Service.

[c] Only Heart Failure network validated by case-control study.

[d] Only Heart Failure network with calculated metrics.

a more pedagogical approach, the teacher can enlarge the learner's learning field by inquiring him/her in such a manner that he/she has to answer not only "how", but also "what", "why" and "what if" [6]. This strategy forces the learner to interact more with the environment and moves the focus of studies on Intelligent Learning Environments (ILE) [3] and Multi-Agents Systems (MAS) [25].

The migration from ILE to multi-agent ILE is not a simple task. In fact, the Artificial Intelligence Research Group at UFRGS has been carrying out research on teaching–learning environments since 1990, through the development of ITS [35,39,43]. Recently, this group has employed MAS approach to develop multi-agent ILE [1,7,9,16,39,42]. The group's intention of developing learning environments with emphasis on pedagogical approach had motivated this move.

Through the adoption of MAS paradigm and under the knowledge representation viewpoint, it was possible to develop more powerful systems, for example, student model and ITS both based on Belief, Desire, and Intention (BDI) architectures [8,30]. The use of Distributed Artificial Intelligence (DAI) techniques enabled both the development of CPU intensive teaching–learning systems and the component reuse [40].

The project herein described is the result of a partnership between a private enterprise, ARQ Systems, the Informatics Institute at the Federal University of Rio Grande do Sul (UFRGS), and the Computer Science Department at the University of Brasilia (UnB).

As it follows, Section 2 describes AMPLIA environment. Section 3 presents AMPLIA architecture, Section 4 brings some examples of AMPLIA use, and Section 5 brings conclusions and future works.

## 2. The AMPLIA environment

AMPLIA is a multi-agent[1] computational environment aimed at supporting learning using a constructivist approach. We will focus on the medical reasoning to describe AMPLIA. The development of this environment is in accordance with the physician process of technical education and specialisation, which, in general, happens through the following activities: medical appointments, classes attendance, and round sessions. Medical students and instructors discuss real cases and current topics of their specialities in round sessions. They also use some classes to discuss papers previously handed out by the teacher and read by learners. The medical student can use AMPLIA as a complementary tool to ease his/her technical skill development on formulated diagnoses, at his/her own pace.

In short, the process of formulating a definite medical diagnosis can be seen as composed of the following steps: medical interview, Current Disease History (CDH), formulation of a differential diagnosis, formulation of a preliminary diagnosis, and definite diagnostic formulation. If suitable, after the formulation of a preliminary diagnosis and before the definite diagnostic formulation, the physician can review the technical literature and request complementary lab tests. In medical interviews, the physician interviews the patient to know the history of his/her last diseases. To obtain the CDH, the physician

[1] An agent is a program that works autonomously for a while, interacts with other agents and services, lives in an environment, and performs a task in the name of a person or an organisation [29].

questions the patient about what his/her main complaint is. The physician visually examines the patient in order to determine his/her condition, searching for signs on the body, and notes down the symptoms the patient mentioned. Then, the physician examines the patient physically, oriented by signs and symptoms already detected. With all these information, the physician makes the differential diagnosis, that is, selects a set of pathologies (diseases) compatible with the data collected, and tries to get new results that can exclude some hypothetical pathology. Through reducing the set of hypothetical pathologies, it is possible to establish the preliminary diagnosis, that is, to determine the most probable pathology. If there is need of confirming the preliminary diagnosis, the physician requests supplementary investigations. While the physician waits for the lab results, he/she can review the technical literature about the pathology he/she suspects the patient is suffering from. The lab tests analysis can confirm the preliminary diagnosis, making it definitive, or supply new information for a new preliminary diagnostic formulation.

The use of probabilistic systems to support this medical procedure can make easier the differential, preliminary, and definitive diagnostic stages. It can reduce the need for technical literature review and complementary lab tests but does not reduce the definitive diagnostic quality. Therefore, it can improve the medical diagnostic quality, reducing both the time needed to make a diagnosis formulation and costs. This reduction of costs is associated to the expectation of a change in the physician's behaviour. We expect that, if the physician has a probabilistic model that helps him/her with the diagnose process then the lab tests requests will decrease. Lab tests will be requested only in cases where it is not possible to get the definitive diagnosis based on interviews and analysis of the hypothetical disease the physician is considering possible, having in mind the patient's main complaint.

AMPLIA is composed of three agents: *LearnerAgent*, *DomainAgent*, and *MediatorAgent*. This section describes the AMPLIA major functions not focusing on agents' details, while Section 3 presents the AMPLIA multi-agent architecture. The *DomainAgent* represents the expert knowledge domain and the *LearnerAgent* represents the learner knowledge. The learner will learn with the AMPLIA support by building probabilistic models, and evaluating diagnostic hypothesis. If the model the learner has built is different from the built-in model, the system's *MediatorAgent* motivates the leaner to review his/her model qualitatively or quantitatively. The *MediatorAgent* guides the learner based on pedagogical strategies.

### *2.1. AMPLIA: phase I—model construction*

#### *2.1.1. Qualitative review of the model construction*

The learner constructs a qualitative model using a Directed Acyclic Graph (DAG). After the initial formulation of the model, it has to be checked on a qualitative level, represented by the causal relationships of domain variables. The *DomainAgent* (see Section 3), at first, verifies if the network is acyclic and connected. Secondly, it verifies if the learner considered all the main variables, and if the dependence and independence implied by the DAG correspond to the expert net. Formally, conditional independence is described by the d-separation criterion presented in [33]. The variables are categorised as findings or diagnosis variables. The findings represent deterministic values to variables obtained during the medical interview process (e.g. signs, symptoms, CDH) or complementary lab tests, etc. The diagnosis variables represent diagnostic hypothesis. The acquiring of findings changes the diagnosis variables' probabilities.

The findings are of the type trigger, essential, complementary, or excluder. *Trigger finding* singles out the diagnosis as a potential solution to the problem, when present. *Essential finding* must be present to assure the diagnosis identification. *Complementary finding* might be present to increase the probability of the diagnosis. *Excluder finding* indicates the diagnosis is improbable (i.e. it has a very low probability), if present. Of course, some diagnosis variables can be considered finding variables for other diagnosis variables if the primes are potential cause for the lasts. The *DomainAgent* will use this classification together with the domain built-in model to inform the *MediatorAgent* about the differences between the learner solution and the domain model. The *MediatorAgent* selects a pedagogical strategy from a database in order to mediate a negotiation process between the learner and the *DomainAgent*, as described in Section 3.

### 2.1.2. *Quantitative review of the model construction*

The quantification of the qualitative part of the network implies evaluating all conditional probability distributions of all variables represented. This codification of probabilities tends to be not the hardest but certainly the longest task in the modelling process. Despite sometimes, in medical domains, there is probabilistic data in abundance in literature; this kind of information can be seldom used directly in the probability distributions assessment. Sometimes, medical literature presents probabilities such as the frequency of symptoms given the occurrence of a disease, but almost never the frequency of such symptoms in the absence of that disease. Furthermore, medical literature does not concern with questions like: what is the probability of a pathological condition to determine the appearance of some symptom or lab finding? What we usually find are statements like "rare finding" or "common finding". Otherwise, if literature cannot provide reliable probabilistic information, estimates can be obtained from statistical data analysis or using machine learning techniques [18,27,32,34]. Experience shows, however, that even when there is good availability of data, they very rarely contribute to the quantification efforts [23]. In medical statistical data, for example, intermediate pathophysiological states of disease are not typically registered. Consequently, domain experts need to evaluate a great number of probabilities.

The decision analysis offers several techniques for eliciting arbitrary probabilities [31]. The simplest technique is to use a numerical probability scale. A probability scale is either a vertical or a horizontal line with endpoints 0 and 100%, and some few anchors, like 25, 50 and 75%. A second technique, used in conjunction with the one above, is the frequency technique. The expert is given the suggestion of envisioning 100 cases with some particular context. For example, the domain expert, a cardiologist, is asked to imagine a population with 100 patients with cardiomegaly (enlarged heart) and to determine how many of them present Chagas disease, as the cause of this pathology. Unfortunately, experience has shown that the use of the probability scale, together with the frequency method, provides experts little to go by and it may result in highly inaccurate probabilities assessments [13].

A natural evolution of these techniques above is due to [14]. They present a new numerical and verbal scale, with only verbal questions and no math notations (see Fig. 1). The marks on this scale are pairs of numerical divisions of probability (0, 15, 25, 50, 75, 85,

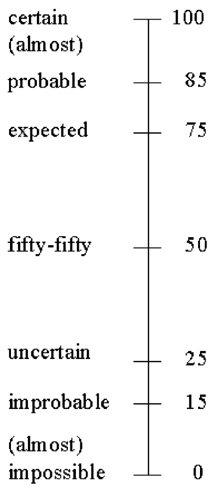


Fig. 1. The modified probability scale proposed by van der Gaag and Renooij.

and 100%) and commonly used expert expressions (impossible, improbable, uncertain, 50-50, expected, probable, and certain).

The researchers have studied how close is the common sense meaning of these expressions and their numerical counterparts approaching. Besides that, they realised that verbal questions were necessary, since many experts were uncomfortable with mathematics probability notations. Another fact noticed was that, the more specialised the experts were in a particular topic, the more prone they were to reason in terms of words. The verbal anchors in the scale then helped them to determine which probability they actually had in mind (Fig. 1). This method also groups together the questions related to the same probability distribution, so that the domain expert can consider them simultaneously. Experts are then encouraged to assess probabilities they are most certain, usually the extremes, and probabilities with unique causal influences, and then to interpolate the remaining probabilities.

Using the van der Gaag et al. [14] modified probability scale idea, AMPLIA offers the learner a graphical editor that allows him to assess the needed probabilities. The last AMPLIA advice in this phase is to verify whether the assessed probability distributions the learner realises are compatible with the probability axioms.

### *2.2. AMPLIA: phase 2—diagnostic hypothesis evaluation*

The diagnostic hypothesis evaluation starts after the construction model phase. It consists of the evidence input (i.e. findings) and in their propagation into the learner's Bayesian network. The SEAMED[2] tool supports both of them. The evidence propagation

[2] SEAMED [10] was the beginning of AMPLIA, and now it is incorporated as a computational tool of this extended challenge. SEAMED presents a graphical interface designed to ease the construction of decision-making support applications in any medical field. In order to analyse a specific case, the application user should enter the available evidence. The application propagates this evidence through the other random variables and updates their conditional probabilities giving the available evidence entered into the system.

yields the update of all the probability distributions inherent to such Bayesian network. AMPLIA uses the junction tree method to propagate evidence and to derive the marginal probability distribution (marginal for short) for each diagnosis variables of interest from the updated probability distributions [20]. The marginals are probability distributions conditioned to all the findings entered in the Bayesian network. Into the medical domain context, findings represent information gathered during the interview process, lab test results, etc. that the physician use to guide his/her reasoning during the process of diagnostic hypothesis evaluation. Findings can make some diagnostic hypothesis more probable and others less probable. The SEAMED represents this fact by increasing or decreasing some diagnostic hypothesis probability. The purpose of AMPLIA, during this phase is to help the development of the physician learner's skills in the diagnosis process, mainly in the differential diagnosis. AMPLIA supports it qualitatively and quantitatively.

### 2.2.1. *Qualitative diagnostic hypothesis evaluation*

The AMPLIA qualitative diagnostic evaluation helps the learner to answer questions like:

- Which are the most probable diagnostic hypotheses?
- Which is the most referential finding in order to give support to accept or reject a particular diagnostic hypothesis?
- What do I have to do in order to diagnose a specific disease, for example, hepatitis?
- Which are the essential findings for a specific diagnosis?
- Which findings are excluders (i.e. improbable) in a specific diagnostic hypothesis?

The learner has the choice of asking AMPLIA to present a textual explanation that describes a selected diagnosis and one report, which has intrinsic information about changes in the dynamic relationship between the variables as consequence of the evidence propagation. This information is included in the AMPLIA built-in model by a domain expert.

### 2.2.2. *Quantitative diagnostic hypothesis evaluation*

AMPLIA support for quantitative diagnostic hypothesis evaluation includes advice in questions like:

- How much does a certain piece of evidence (finding) contribute to the best diagnostic hypothesis?
- How can a particular finding gathering affect the learner's preliminary diagnosis?
- What evidence would be necessary in order to achieve a certain probability level for a diagnostic hypothesis?

In addition, AMPLIA provides sensitivity analysis features for studying how variables' probability distribution assessments affect the results of the model built by the learner, especially on a variable of interest. In short, for example, how a diagnostic hypothesis is affected by several finding's probability distribution assessments. A sensitivity analysis, in which a single assessment is varied, is termed an *one-way sensitivity analysis*. In a *two-way sensitivity analysis* of a Bayesian network, two probability assessments vary simultaneously.

This analysis reveals the joint effect of variation on a probability of interest. AMPLIA supports both of them.

## 3. The AMPLIA architecture

Our AI Group has accomplished research in cognitive agents [8,30], using logical model of the agent's mental states to represent its beliefs, desires, and intentions. In the present project, we have been investigating the use of Bayesian networks in the agent's beliefs modelling and mental states to guide the negotiation process. Three cognitive agents (*LearnerAgent*, *MediatorAgent*, and *DomainAgent*), two databases (expert knowledge built-in model database and pedagogical strategy database) and the interface module compose the AMPLIA architecture, illustrated in Fig. 2.

The *LearnerAgent* represents the student's beliefs in that domain, the confidence degree this learner has on the network model he/she has built. It also includes a steady part with basic information about the learner. The *LearnerAgent* elaborates and updates the student's model. From now on, we will use learner and student interchangeably.

The *MediatorAgent* makes decisions about when interfering during the learner's network model construction process, besides it can act by learner's request. It will select the most appropriate pedagogical strategy to query and help the learner.

The *DomainAgent* pairs the network built by the learner with the built-in model. The result is sent to the *MediatorAgent* to co-ordinate the negotiation process.

The *built-in model database* contains the networks built by the knowledge expert, the nodes classification (trigger, essential, complementary, or excluder), explanation resources and brief texts on the problem.

The *pedagogical strategy database* stores and put available strategies according to the learner's model (see Table 2). The negotiation process uses pedagogical strategies.

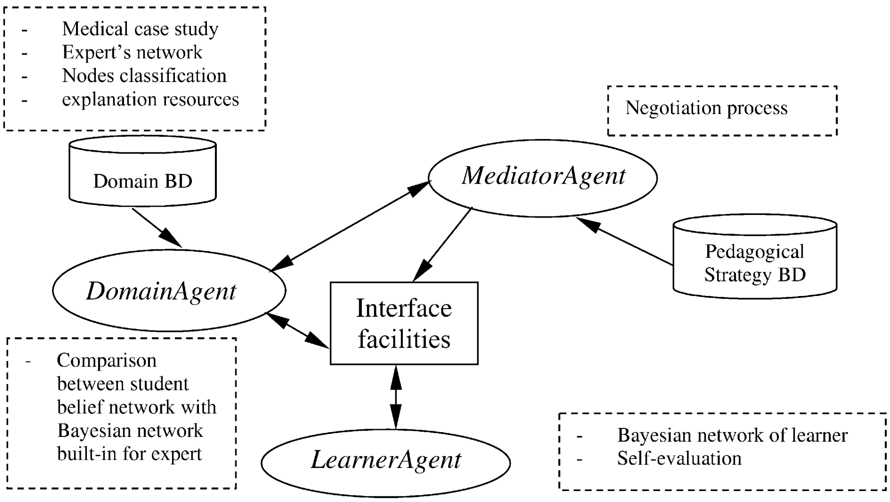


Fig. 2. The AMPLIA architecture.

Table 2
*MediatorAgent*'s negotiation strategies

| Strategy | Network model | Learner's confidence level |
|---|---|---|
| (1) Your diagnostic model is complete, matching the expert's network model | Complete | High |
| (2) Click on the nodes or causal relationship where your confidence is lower | | Moderate |
| (3) Click on the nodes or causal relationship where your confidence is higher | | Low |
| (4) Consider the following information: (...) and think about which nodes are necessary to incorporate findings | Feasible | High |
| (2) Click on the nodes or causal relationship where your confidence is lower | | Moderate |
| (3) Click on the nodes or causal relationship where your confidence is higher | | Low |
| (4) Consider the following information: (...) and think about which nodes are necessary to incorporate findings | Incomplete | High |
| (2) Click on the nodes or causal relationship where your confidence is lower | | Moderate |
| (3) Click on the nodes or causal relationship where your confidence is higher | | Low |
| (4) Consider the following information: (...) and think about which nodes are necessary to incorporate findings | Incorrect | High |
| (2) Click on the nodes or causal relationship where your confidence is lower | | Moderate |
| (3) Click on the nodes or causal relationship where your confidence is higher | | Low |
| (5) Your diagnostic model is not according to Bayesian network structures. Please, review the network concepts | Unfeasible | High |
| (5) Your diagnostic model is not according to Bayesian network structures. Please, review the network concepts | | Moderate |
| (5) Your diagnostic model is not according to Bayesian network structures. Please, review the network concepts | | Low |

The *interface* makes possible the learner's interaction. The student can access the brief text, the SEAMED tools, the menus for variable choices, the option to transmit his/her confidence degree in the network, and the option to save his/her network in a database folder.

### 3.1. Communication among agents

AMPLIA agents communicate over a FIPA-OS platform. The Foundation for Intelligent Physical Agents (FIPA) has put forward an agent communication language (ACL), based on the Speech Act Theory. FIPA assumes the existence of an agent management system, not part of the language, and abstracts the low-level communication details. Currently inform, request, query-if and query-ref acts are being used through FIPA's Request and Query Interaction Protocols. Bayesian networks are represented, for communication purposes, in a XML-based format (XBN) and FIPA-SL0 is used as content language for communicative act messages. To establish communication between agents there is a need for a common frame of reference or shared ontology. The shared ontology determines how particular message content is to be interpreted. Below are presented the types of messages exchanged among agents:

- *LearnerAgent*'s beliefs are modelled in a Bayesian network with a graphical editor, through which the learner builds hypothetical network models and diagnostics reasoning.

*LearnerAgent* sends request messages to *DomainAgent* with its beliefs asking for a review. The abstract syntax notation for action expressions carried by this message content type is:

$$\text{Review-beliefs}(\textit{LearnerAgent}\,\text{Goal}) = \{\text{XBN\_}\textit{LearnerAgent}\} \tag{1}$$

- *DomainAgent* has beliefs on medical domain knowledge depicted through Bayesian networks. It puts available to the student, knowledge domain and explanation resources enough to aid the learning process. *DomainAgents* send inform messages to the *LearnerAgent* with the case study to be modelled, and the list of variables the learner can use in the construction of the hypothetical model. Providing variables that are sensitive to the case study context will help to maintain common ontology between agents during the learning process. As the *DomainAgent* receives the learner's network model, it starts an action (program) that reviews the learner's belief, by comparing the expert network model against the student's. After identifying conflicts on the result of this comparison, *DomainAgent* sends inform messages to *MediatorAgent* presenting these conflict points. Abstract syntax notations for action expressions carried by case-study and conflict message types are, respectively:

$$\text{Case\_study}(\textit{LearnerAgent}\,\text{Goal}) = \{\text{Text\_case\_study, Variables\_list}\} \tag{2}$$

$$\text{Conflicts}(\textit{LearnerAgent}\,\text{Goal}) = \{\text{Classification\_model, Conflict\_points, Arguments}\} \tag{3}$$

- *MediatorAgent* mediates the interaction between *DomainAgent* and *LearnerAgent* aiming at helping in the conflicts solving process. After receiving the *DomainAgent* message, *MediatorAgent* sends query messages to *LearnerAgent* requesting information on the confidence level about the hypothetical domain built. Based on this information and on pedagogical strategies in use, *MediatorAgent* sends to *LearnerAgent* arguments (essentially simple inform messages) that will motivate the learner to review beliefs and modify his/her actions. Abstract syntax notations for action expressions carried by confidence discovery and strategy selection message types are, respectively:

$$\text{Confidence}(\textit{LearnerAgent}\,\text{Goal}) = \{\text{Confidence\_level}\} \tag{4}$$

$$\text{Strategy}(\textit{LearnerAgent}\,\text{Goal}) = \{\text{XBN\_model, Confidence\_level, Arguments}\} \tag{5}$$

Communication dialogue among agents will go on as far as the *LearnerAgent* wants to review his/her model. This can make him/her reach a network model identical to the *DomainAgent*'s or another feasible network model, although not identical to the *DomainAgent*'s.

### *3.2. The negotiation process*

In a multi-agent learning environment, where the learner is an active subject in the learning process, co-operation can be taken for granted. Co-operation must be planned for and achieved through communication and a sort of negotiation. Negotiation is based on argumentation.

Argumentation theory is an interdisciplinary field that calls the attention of philosophers, logicians, linguists, and speech acts researchers, for example. Much of the research accomplished in this area is not formal, but some organisation has been seen since the 1980s. Our approach to argumentation tries to explore possible relations between argumentation and learning, taking into account the point of view of researchers of cognitive sciences and education. A central mean of knowledge construction is reasoning, and the result of this reasoning is an argument, structure that consists of a conclusion and of a set of reasons that support it [38].

As well as AMPLIA, the empirical work of Baker [5] explored argumentation functions in collaborative problem-solving. According to his work, argumentation has three main functions: it works as an activator in the search for information, as a filter of flaw proposals and as a provider of interactive pressure to co-elaborate ideas. Another empirical approach of argumentation is Veerman's [41]. This work reports a study on collaborative learning through argumentation. According to Veerman, collaborative learning allows students to negotiate different perspectives by externalising and negotiating them. Through argumentation they can build and re-build knowledge with relation to other learning objectives. Veerman's empirical interest is on the relation between argumentation and production of constructivist activities.

In AMPLIA, the student expresses his/her argumentation through the Bayesian network modelling. Rolf and Magnusson [36] assert that the practice and teaching, and the reasoning and argumentation teaching are adequate to the use of diagrams—"a large number of textbooks present arguments in the form of boxes and arrows" [36]. We agree with this idea; this way the learner can build a domain model through a graphic editor where arguments are composed of nodes and links among them. Rolf classifies software that uses graphs to express arguments in three levels. This classification takes into consideration the calculation that has been used. The Belvedere system, for example, does not have any calculus, thus it is at the first level; Athenas and Reason!Able systems are in an intermediate level, containing some numeric designation and rules to filter the best arguments. These systems do not have a tutor or mediator as an aid in the learning task. AMPLIA is at the third level which is compounded by systems that contains advanced theory-based capacities for calculations. The AMPLIA calculus is based on expected utility and Bayesian probabilities.

The negotiation process follows an interaction/conversation protocol that is shown by the finite state machine of Fig. 3. In the initial state, the *DomainAgent* presents a case study to the learner. In the second state, the learner models his/her diagnostic hypothesis from the study of cases the *DomainAgent* has put available. Yet, in the second state, the *LearnerAgent* sends its model to the *DomainAgent* in order to be assessed. This evaluation results in a classification of points where the student model differs from the *DomainAgent*'s.

The classification is according to the importance of each finding of the model (see Section 2.1.1). In the third state, the student evaluates the message received from the *MediatorAgent* and makes his arguments through alterations accomplished in his/her diagnostic modelling. In this state, the *LearnerAgent* can also decide to abandon the learning process (because it feels satisfied or not). In the fourth stage, the *MediatorAgent*,

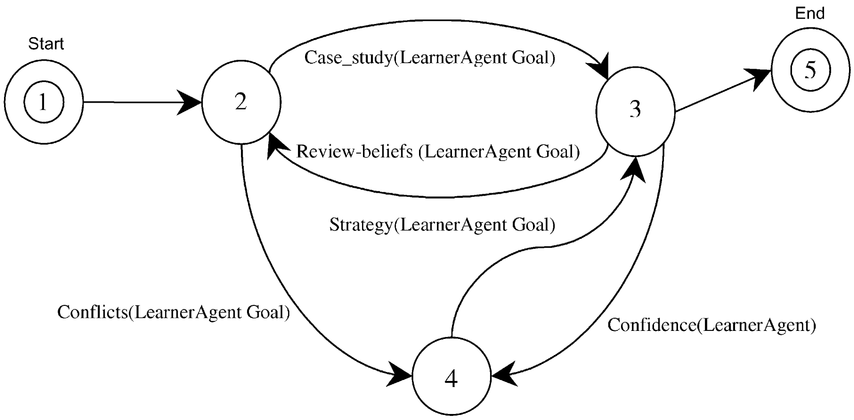


Fig. 3. Interaction protocol.

based on the result of the *DomainAgent* analysis and on the level of confidence provided by *LearnerAgent* triggers the best pedagogic strategy (Table 2).

The AMPLIA negotiation process happens through a dynamic choice of strategies. The parameters considered are linked to the learner's actions and to the assessment the *DomainAgent* carries out. In this negotiation process, only the student has the option of abandoning the interaction. However, when arguments the student used are adequate for a solution, the *MediatorAgent* may come to accept the student's modelling, even if it is different from the *DomainAgent*'s.

Negotiation is based on the student's own learning process. Presupposing that the student's ability of absorbing/considering all the findings of a diagnostic modelling be affected by his/her experience as a professional, the 'intuitive' negotiation process allows to change other agent's thoughts/ideas through argumentation (using a Bayesian graphic editor).

In a negotiation process, one must consider a confront space where there are several possibilities for the solution of an impasse. This happens during a diagnostic hypothesis modelling. That is, a diagnostic can be identified by different hypothesis (causal relation), however, the trigger and essential findings should be considered and the complementary findings would only help to better confirm the diagnosis, as well as the presence/absence of abstract nodes (see example of Section 4).

The negotiation process can be seen also like a game, where the agents are able of misleading the opponent. Therefore, in AMPLIA, the *DomainAgent* can confuse the learner including an excluder node in the variable list (see example of Section 4).

In this context, in what moment the *DomainAgent* could "yield" to the *LearnerAgent*'s opinions? When although the student modelling is not identical to the *DomainAgent*'s it solves the case-study problem.

The *LearnerAgent* could yield to the *DomainAgent* opinions when it realises the flaws on its modelling, considering (or not) strategies chosen by the mediator, as we can see in the following example.

- *DomainAgent* proposes a problem.
- *LearnerAgent* models and submits to evaluation.

- The result of this evaluation is sent to the *MediatorAgent* as specific comments for each region of the network. The *DomainAgent* informs also the main critical points or the most important ones (utility function).
- The *MediatorAgent* asks the student to tell his/her confidence level about the graph region.
- Based on these information, the *MediatorAgent* chooses the right strategy (which the student can consider or not).
- In the next phase of the process, the *MediatorAgent* has an impression about the student (if he/she considered or not the strategy presented) and also, based on this information, the *MediatorAgent* chooses again a new strategy (now taking into account also the modifications the student has done).

AMPLIA follows a constructivist line, intuitively the *MediatorAgent* should be condescending with the student's position, trying not to be aggressive with the student, trying to consider what the student knows more to help him/her to advance on his/her hypothesis.

The idea is then to maximise the student and expert's utilities. The *MediatorAgent* can soothe the conflicts that happen between both agents (*DomainAgent* and *LearnerAgent*), in a non-aggressive way, using pedagogical strategies that meet the constructivist objectives of the project.

AMPLIA pedagogical proposal follows the constructivist approach. In this approach, the student is an active subject in the learning process and the teacher undertakes the role of mediator to motivate this learner, proposing strategies for the reflection on the solution process. So, the AMPLIA mediation process is carried out by an intelligent *MediatorAgent*, which uses the pedagogical strategies during the learning process. Among the pedagogical strategies used in a constructivist environment, we cite: proposals of problems that involve hypotheses formulation, comparison and/or exclusion, data categorisation and models reformulation, searching for regularities and data reorganisation for effective actions. We also allude to the Kolb's learning cycle, which Belhot [6] discusses posing the following key-questions ''what?, how?, what if?, and why?'' in the knowledge approach, providing a formative (theoretical), prescriptive (practical), constructive (reflective) and prospective (critical) learning. Therefore, when using AMPLIA, the learner must be motivated to learn also the ''why'' and not only ''how'' to solve an issue.

In Kolb's schema, motivation is a key element in learning, and emotions have an important role when we talk about motivating the student [7]. Different from that, AMPLIA proposes that the student makes a self-evaluation and informs his confidence degree, which will be the variable considered by the *MediatorAgent*. From this pedagogical viewpoint, mistake and doubt are important moments for self-evaluation and reflection, therefore, the learner himself will determine his studies pace. Thus, the *MediatorAgent* intervention and the strategy used will always question the student's doubts and assertions.

*MediatorAgent* starts the negotiation process after the learner's beliefs review. This review is carried out by procedures that promote qualitative and quantitative assessment of learner's solution, classifying models as feasible, unfeasible, complete, incomplete, or incorrect.

*Unfeasible network model* is identified as a network that does not satisfy the definition of a Bayesian network. The model is not an oriented acyclic graph, and/or presents a disconnected network, and/or the probability distributions performed by the learner are not compatible with the probability axioms. This process of error identification involves a number of algorithms that will not be dealt in the present paper.

*Incorrect network model* is the network whose model is conceptually incorrect. That happens, for example, in the presence of an excluder node that should be in the model to refute the diagnosis, while its presence confirms it.

*Incomplete network model* is that network that presents the lack of some nodes or relations considered important (trigger, essential, complementary), whether they are diagnosis or findings. It is almost impossible to reach a correct diagnostic hypothesis evaluation in this case, even if the model is a complete and well-formed Bayesian network.

*Feasible network model* is a network different from the built-in model but it satisfies the case study proposed to the learner. The probabilistic (in)dependence relations expressed in a feasible model are equivalent to probabilistic relations of the built-in model, that is, causal relations represented in both models are equivalent. The *DomainAgent* will identify this classification through the qualitative and quantitative evaluation of the learner's models. In the qualitative evaluation, *DomainAgent* must identify in the student's model important variables (trigger, essential, complementary), whether they are diagnostic or findings.

*Complete network model* is identical to the model the expert built. The causal relationships of domain variables and the conditional probability distributions of all variables are identical to those of the built-in model.

The *DomainAgent* classifies the student model network and sends a message containing the conflict points to the *MediatorAgent*, together with a list of explanations that will be useful in the negotiation process. The *MediatorAgent* proposes that the learner makes a self-evaluation and informs his/her confidence degree in the model he/she has built. The confidence degree could be High, Moderate or Low. The *LearnerAgent* chooses the confidence degree and informs it to the *MediatorAgent*. The *MediatorAgent* sends to the *LearnerAgent* a message with arguments based on the model classification given by *DomainAgent* and the learner's confidence degree. This message will motivate the learner to review his/her beliefs, helping him/her to decide which will be the next actions. A generic example of message is:

> Your diagnostic model is (complete, feasible, unfeasible, incomplete) and your confidence is (high, moderate, low), (followed by the strategy itself).Table 2 presents strategies that could be used by the *MediatorAgent*.

Now, *MediatorAgent* will wait for the learner's next action which could be:

- Ask for additional information about nodes and/or causal relationship where the *LearnerAgent* has a low confidence degree.
- Ask for a review on Bayesian network concepts.
- Keep on with model changes to send it to the *DomainAgent* later.
- Leave the negotiation process.
- See expert's network model.

The negotiation process is an interactive method where the *MediatorAgent* will be continuously motivating the *LearnerAgent* to reach its objectives, which are: hypothetical model construction and diagnostic reasoning development. Over repeated encounters, *MediatorAgent* may analyse the student's patterns of behaviour to establish an analogy to the teacher's role in constructive approach. This may influence the evaluation of arguments, as we can see in situations such as an impasse. By observing the reactions to the arguments, the *MediatorAgent* can update and correct *LearnerAgent*'s model, thus refining its planning and argumentation knowledge.

An example of an impasse situation is when the student persists in the same error, even with all the initial motivational arguments provided. A new strategy should then be employed. For example, let us imagine that a student made an unfeasible model (the network presents cycles). The system, using a specific strategy, informs that the network structure is not adequate to Bayesian network concepts. However, this student insists in the mistake, and can not identify the problem, that could be a cyclic network, a disconnected network, or a network with incorrect probability distribution. With time, analysing the student's behaviour patterns, the *MediatorAgent* can take the decision of sending an argument right to the problem the student is facing, for example, "Your network presents cycles".

### *3.3. Implementation aspects*

At this moment, we have already implemented the following modules in AMPLIA's environment:

- Graphic interface editor the learner uses to create his/her network (developed with DELPHI 6.0 programming language).
- The probabilistic inference engine (in DELPHI 6.0).
- The *LearnerAgent* modelled through Bayesian networks.
- The domain module composed of the domain agent plus a medical case database with four different case studies and their corresponding diagnostic Bayesian networks (nets are modelled using XML).
- The agent communication mechanism (FIPA-OS).
- AMPLIA's main interface under development in Java.

To make possible for the system to execute over FIPA-OS platform the components that were not developed using Java, are currently being agentified using CORBA.

## 4. An example: Rheumatic Fever

An example of AMPLIA's SEAMED module in action is as follows. When accessing AMPLIA for the first time, the learner will fill a form with some data such as name, password, etc. aiming at creating a folder at the learner's database. In the next log-ins, she/he will only inform password. After identification, the *LearnerAgent* selects a medical case study from the *DomainAgent* database. A case study is represented by a textual diagnostic investigation problem, possibly enriched with illustrated graphical and/or sound files (see Fig. 4). The *DomainAgent* compares each case study with a built-in model managed.[3]

After this case-study presentation, the learner can start the construction model phase with the support of Bayesian network graphical editor. The *DomainAgent* presents a list with all nodes requested plus nodes not related to this current investigation to the learner. The learner is then encouraged to develop a new Bayesian network selecting all nodes he/she feels appropriate to the case, from the list nodes. The learner can elaborate both the qualitative and the quantitative parts of the model. Also, he needs to identify all finding variables and all diagnosis variables presented in the case and assesses the initial conditional probability distribution of the entire variables set.

Fig. 5 presents the built-in model previously developed by an expert for the example above. Note that it is only visible to learner under request.

Notice that the diagnostic node ‘‘Rheumatic Fever’’ represents the definitive diagnosis of this case. Rheumatic Fever is a systemic, autoimmune illness due to cross-reactivity with β-hemolytic streptococci of Group A. After a throat infection, for example, antibodies developed by the immune system against the bacteria (streptococcus) cross-react with the very own tissues of the patient affected (like joint tissues, and heart valves). In the susceptible 2% of the population, there may be permanent damage to heart valves and the risk of subsequent endocarditis (inflammation of the tissue that covers the walls of the heart) is increased. It usually follows a latent period of 2–4 weeks. Peak incidence: 5–15 years.

As stated previously, nodes are categorised as diagnostic and/or findings. Finding nodes are also sub-classified as trigger, essential, complementary, and excluder.

As an example of a *trigger* node (see Fig. 6), notice the ‘‘painless swellings over bony prominences’’.

A positive evidence for this node is enough to indicate a ‘‘true’’ posteriori distribution of ‘‘Subcutaneous nodules’’ (its sole parent—a diagnostic AND finding node). ‘‘Recent Streptococcal Infection’’, ‘‘Major Criteria’’, and ‘‘Minor Criteria’’ represent essential nodes in this network (see Fig. 7).

These nodes are classified as such because they are needed for the final diagnosis of ‘‘Rheumatic Fever’’. According to the revised Jones criteria, a diagnosis of Rheumatic Fever may only be established by the presence of evidence of previous streptococcal infection (represented by ‘‘Recent Streptococcal Infection’’ node) plus two Major Criteria (represented by ‘‘TwoOrMore’’ state of node ‘‘Major Criteria’’) or one major (state ‘‘OnlyOne’’ of this same node), and two Minor Criteria (corresponding to state ‘‘TwoOrMore’’ of ‘‘Minor Criteria’’ node).

Major and Minor Criteria, presented in Table 3, are symbolised by nodes. Notice that the nodes in Table 3 are classified as *complementary*, because they help to establish beliefs on other nodes (in this case, the essential nodes Minor and Major Criteria).

[3] Notice that the built-in network developed by the expert makes use of abstract nodes (like major and minor criteria) commonly seen in medical textbooks. It is important to state that the main purpose of AMPLIA is to serve as a pedagogical aid system (and not as precise diagnostic tool). Therefore, certain aspects of the subject being studied (like Rheumatic Fever and also Jones Criteria) should appear.

Identification: J.P.C., 9 years old, female, Caucasian, resident in Criciúma, state of Santa Catarina, Brazil.

Main Complaint: "Tender knee"

Current Disease History: A 9-year-old girl, accompanied by her mother, comes to her medical appointment complaining of a left tender, red knee. She refers that this pain started two weeks ago in her left ankle. As soon as the swelling in the ankle subsided, her knee started to redden. She feels the pain is now moving to her right hand too. Her mother informs that about 8 weeks ago she had a febrile illness, followed 2 days later by a severe sore throat. Also, she says that yesterday her daughter was also complaining of a burning sensation in her chest.

Vital signs:
- HR (Heart Rate): 56 bpm;
- BR (Breath Rate): 26 ipm;
- BP (Blood Pressure): 90/60mmHg;
- Temperature: 37.9°C

Physical Exam:

Ectoscopy :
- Pale ocular mucous membrane. Presence of a small, soft, pea-sized, painless to percussion swelling, over distal, anterior portion of her right forearm. The skin moves freely over it, and there's no apparent local change in color, or temperature. (Click here to see image - when clicking, an image of the arm will appear)
- Presence of a round 2 cm wide pinkish rash, not indurated on mandibular portion of face, that blanches on pressure.
- Pulmonary Auscultation: Uniformly distributed vesicular sounds heard, with crepitant rales over lung bases.

Cardiac Auscultation:
- Gallop rhythm, bradycardic, with presence of a high-pitched holosystolic blowing murmur over the aortic area. (Click here to listen - when clicking, the sound can be listened) Also audible is a pericardial friction rub, best heard over the left sternal border. (Click here to listen - when clicking, the sound can be listened)

Oral exam: No signs of mucous erythema, discrete hypertrophic tonsils.

Otoscopic exam: Transparent, shiny, tympanic membranes.

Abdomen Exam: Normal bowel sounds. No signs of visceromegalies.

Lab Tests: Ht: 31.8
Hb: 9.2
MCV: 90
HCM : 26,8
Leucocytes: 8000, Bast:4%, Seg: 69%, Mon:10%, Lymph:12%, Eo:5%
Platelets: 280000
ASOT : 600 (Normal range: up to 200)
ESR: 80 (Normal range: up to 20)
C-Reactive Protein: Reactive (+++). (Normal range: negative).

ECG: Sinus bradycardia, HR:52bpm, prolonged PR interval, elevation of ST segment on leads V1 to V3, compatible with epicardial injury and/or pericarditis. Echocardiography suggested.

Chest X-Ray: Cardiomegaly. Bilateral pericardial effusion. Kerling-B lines observed, and possible incipient signs of heart failure. (Click here to see image - on click

Fig. 4. A medical case study.

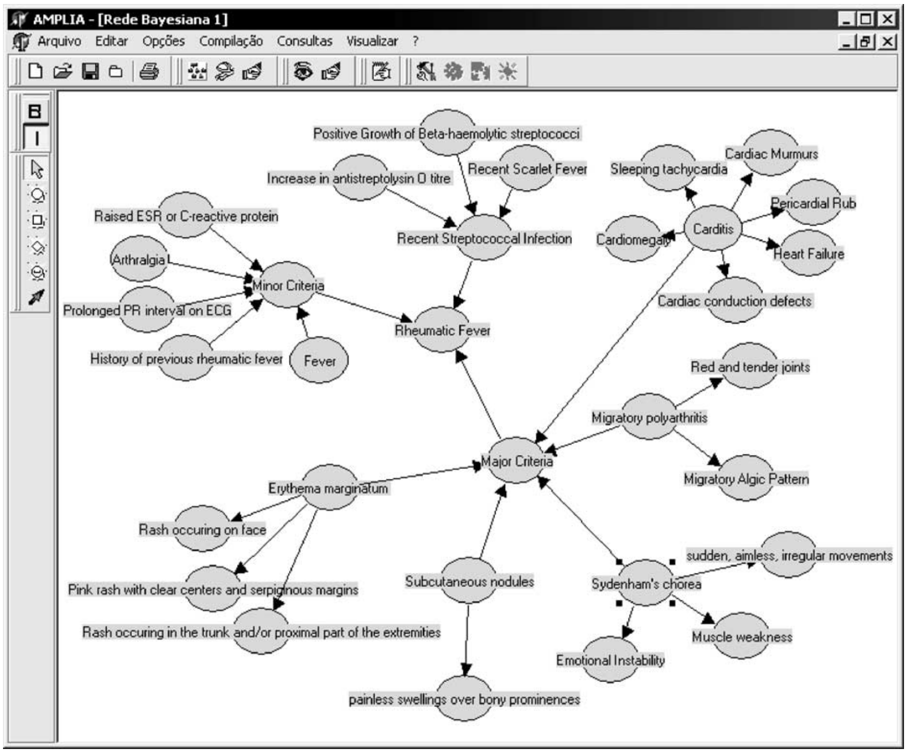


Fig. 5. The domain built-in model developed by an expert.

Table 3
Major and Minor Criteria

| Nodes | Presented by patient? |
|---|---|
| Major Criteria | |
| Carditis (heart inflammation) | Yes |
| Migratory polyarthritis (pain that ‘‘moves’’ from joint to joint) | Yes |
| Sydenham’s chorea (involuntary movements of extremities) | No |
| Subcutaneous nodules (swellings under skin) | Yes |
| Erythema marginatum (pink rash in skin) | No |
| Minor Criteria | |
| Raised ESR or C-reactive protein (unspecific signs of inflammation in body) | Yes |
| Arthralgia (tender joint) | No (because evidence was already marked for Migratory polyarthritis) |
| Fever | Yes |
| History of previous Rheumatic Fever | No |
| Prolonged PR interval on ECG | Yes |

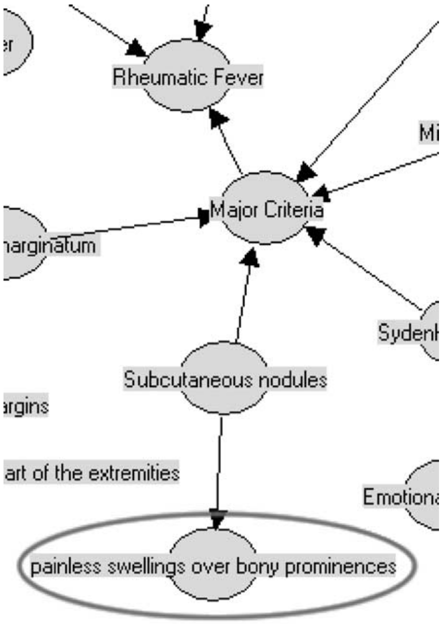


Fig. 6. A network with a trigger node.

Finally, as an example of *excluder* node (see Fig. 8), note the entity ''Rash occurring on face''. The *DomainAgent* included this node to confuse the learner trickily. ''Erythema Marginatum'', this pink rash never occurs on face! In case of a positive evidence of this node, the posteriori probability distribution of ''Erythema Marginatum'' node would need to be negative.

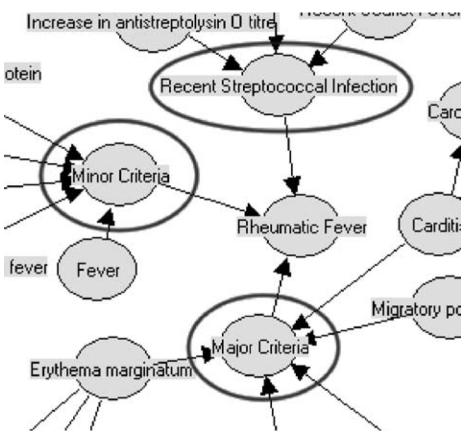


Fig. 7. A network with essential nodes.

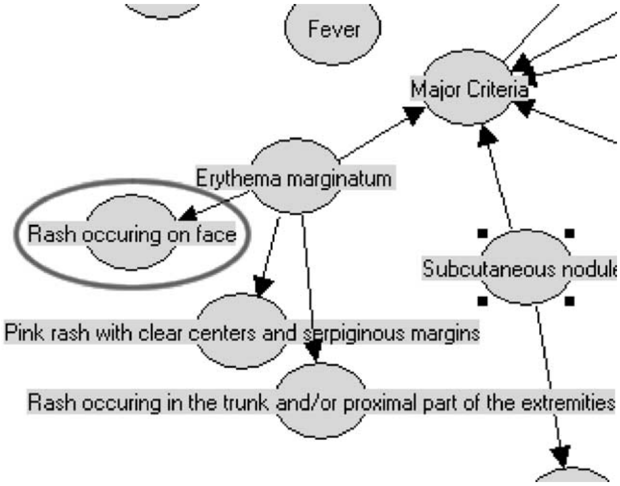


Fig. 8. A network with an excluder node.

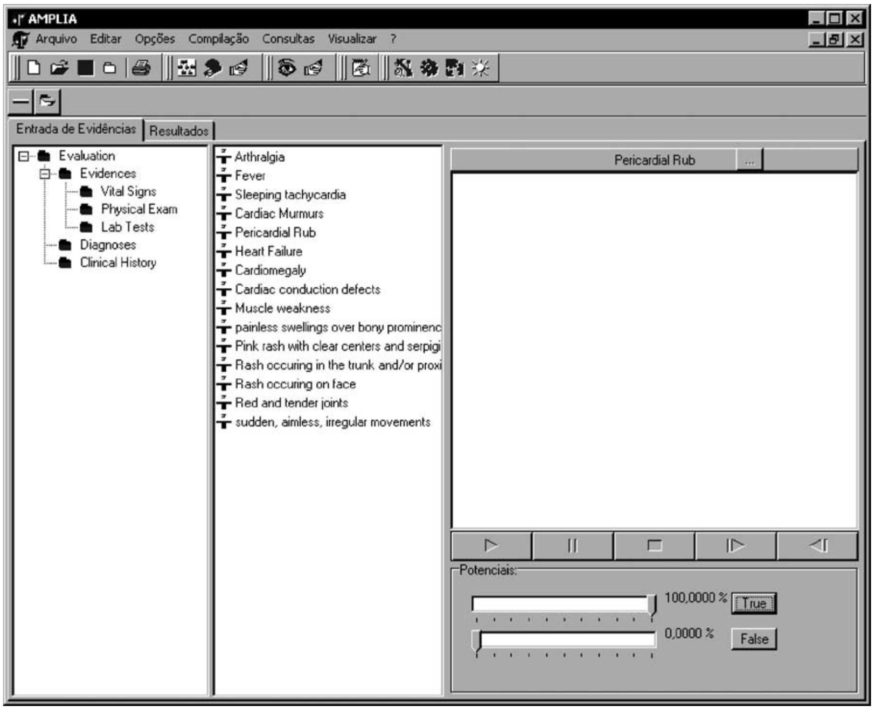


Fig. 9. Entering evidences.

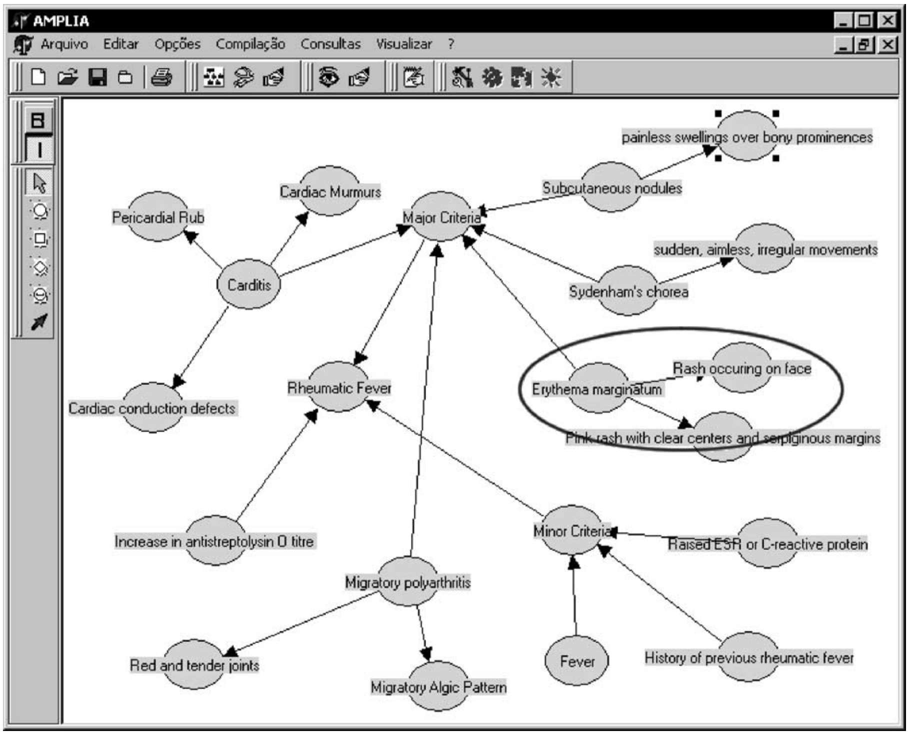


Fig. 10. A hypothetical learner's network.

After the learner finishes the network, he begins to enter the evidences from the textual diagnostic investigation, according to his own interpretation. Fig. 9 presents AMPLIA'S interface and shows this phase.

The *LearnerAgent* sends the qualitative part of the network (i.e. causal relationships) of the learners' network, together with the a priori and a posteriori probabilities distribution to the *DomainAgent*. In turn, the *DomainAgent* starts to ''compare'' these probabilities and causal relationships with the built-in model. Fig. 10 shows a hypothetical learner's network.

''Rash occurring on face'' has a correct relationship with ''Erythema Marginatum''. After inference tests, however, the *DomainAgent* discovers that both ''Erythema Marginatum'' and its excluder node ''Rash occurring on face'' have positive posteriori values. This fact triggers the *DomainAgent* to inform this error to the *MediatorAgent*. The *MediatorAgent* recognises what kind of error it is dealing with and explains to the learner (from the pedagogic strategy database) that ''Rash occurring on face'' is either not related to ''Erythema Marginatum'' or has a negative impact on this node. The *MediatorAgent* then provides the learner with information regarding the causative agents of ''Erythema Marginatum'' (notice that other nodes are also missing in the learner's network), what it is and to which it is related. This diagnostic reasoning is directed to the nodes classified as ''diagnostic'' (like ''Erythema Marginatum''). Otherwise, the system would not know whether to focus on ‘‘Erythema Marginatum’’ (a diagnostic node) or ‘‘Rash occurring on face’’ (a finding node).

Besides developing the domain network that resolves the medical investigation problem, the expert is also responsible for a brief argumentation concerning each node. The *MediatorAgent* eventually uses this data for pedagogical purposes. See Table 4 to observe the real data the *MediatorAgent* makes use during the learning process.

More examples of hypothetical learner models and their respective classifications are presented below.

**Example 1.** According to the classification that Table 2 presents, the hypothetical learner network seen in Fig. 11 falls under the *feasible* learner model. Although essential nodes like ‘‘Major Criteria’’, ‘‘Minor Criteria’’ and ‘‘Recent Streptococcal Infection’’ are missing, the learner managed to include all nodes present in the case. After performance analysis made by the *DomainAgent*, this network was able to correctly diagnose ‘‘Rheumatic Fever’’. Despite its limited use (i.e. the network could only be (safely) used to this case), the learner correctly synthesised the important aspects of this investigation challenge. See Fig. 11.

This example shows a common problem observed among expert and learner models: the difference in the presence/absence of abstract nodes. Both ‘‘Major Criteria’’ and ‘‘Minor Criteria’’ are logic entities. They summarise the presence/absence of their parent nodes. The *MediatorAgent* would discourage the learner, and the environment would lose

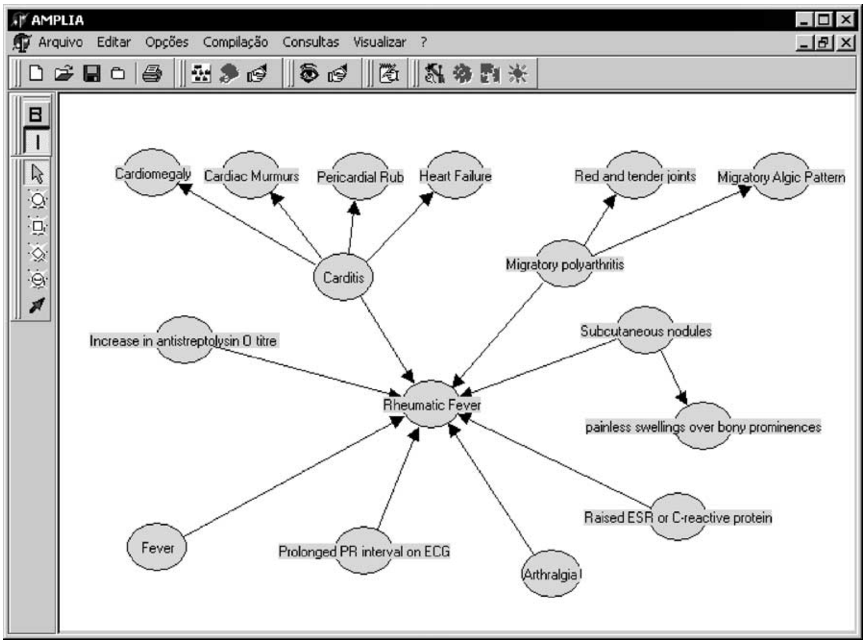


Fig. 11. A feasible network.

Table 4
Partial view of a real pedagogical argumentation table

| Node | Argumentation | Classification I | Classification II |
|---|---|---|---|
| Rheumatic Fever | This node is important because it holds the key to the solution of this case study. | Trigger | Diagnostic |
| Minor Criteria | This node is a logical node that summarises the presence/absence of unspecific signs or lab test results that correlate to this case. | Essential | Diagnostic |
| Major Criteria | This node is a logical node that summarises the presence/absence of important findings you must observe in order to establish a correct diagnosis for this case (tip: create different states to express the presence of one, two or more findings present). | Essential | Diagnostic |
| Recent Streptococcal Infection | This node is a logical node that summarises the presence/absence of parent nodes that make evident the history of a Recent Streptococcal Infection. Remember that, to establish a correct diagnosis for this case, either one of this two situations must be fulfilled.<br>Recent Streptococcal Infection plus two (or more) Major Criteria.<br>Recent Streptococcal Infection plus (at least) one major criterion and two (or more) Minor Criteria. | Essential | Diagnostic |
| Carditis | The heart is the site of the most characteristic and consequential involvement, and all its layers—endocardium, myocardium, and pericardium—may be involved. This generalised involvement gives rise to the term *rheumatic pancarditis*. The most characteristic and specific pattern of rheumatic inflammation is found in the *myocardial Aschoff body*, a submiliary granuloma. This lesion, when present in its classic form, is generally considered to be pathognomonic of . . . (by now you should have an understanding of what this investigation problem is all about). | Complementary | Diagnostic |
| Pink rash with clear centres and serpiginous margins | This finding should remind you of a major criterion called . . ., that is present in 2–10% of patients with . . .. | Complementary | Finding |
| Rash occurring on face | Perhaps your network is not working properly because this node excludes the possibility of a patient having . . ., a major criterion of . . .. | Excluder | Finding |

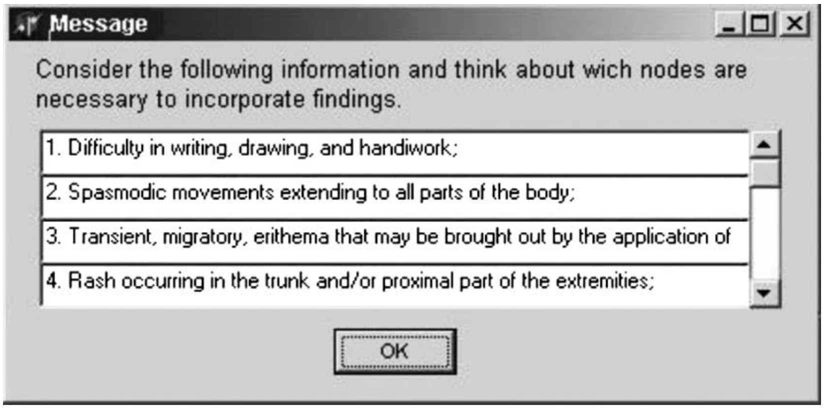


Fig. 12. A *MediatorAgent* intervention.

credibility in case only negative interventions were to be brought up. To identify situations like that, the *DomainAgent* searches for all missing nodes. If parent nodes of those missing are, nevertheless, present in the learner network (which is the case depicted in Fig. 11), the *DomainAgent* verifies (through extensive sensitive analysis) whether meaningful probabilistic impacts are still observed in each diagnostic node in the learner network. Because all diagnostic nodes "Rheumatic Fever", "Migratory polyarthritis", "Carditis", and "Subcutaneous nodules" (present in case) were correctly diagnosed, the *MediatorAgent* assumes, at first, that the learner falls under the "higher learner confidence" classification. To confirm this, the learner is inquired about his confidence in his model and only then he is inquired for the missing nodes. Fig. 12 displays an example of this kind of intervention.

It is important to note that complaints are obtained from the domain model's nodes database. Each entry listed represents a different (missing) node in the learner network (in this case, nodes "sudden, aimless, irregular movements", "Sydenham's chorea", "Migratory Algic Pattern", "pink rash with clear centres and serpiginous margins", "Rash occurring in the trunk and/or proximal part of the extremities" and "Muscle Weakness", respectively). These additional clues do not exhaust all missing nodes in the learner's model so that he/she does not get overwhelmed with too much information.

**Example 2.** Fig. 13 represents a complete model network because it includes all nodes necessary to establish a precise diagnose of "Rheumatic Fever". All possible Major and Minor Criteria are represented. Not only the topology is correct but also the parameter analysis confirms that the probabilities entered approximate to those of the expert model.

**Example 3.** Fig. 14 shows another learner model network classified as incomplete. Note that the diagnosis "Rheumatic Fever" cannot be established because some important nodes are missing. The learner is not considering "Sydenham's chorea" as a possible major criterion, although not present in case, this sign should always be sought and the oblivious of "Prolonged PR interval on ECG" as a possible finding that could be counted as a minor criterion (which is present in patient's ECG result).

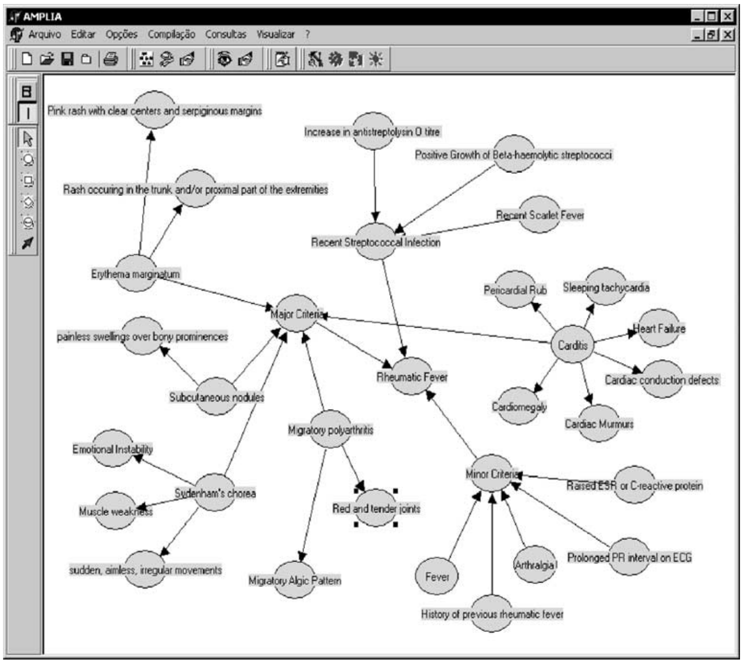


Fig. 13. A complete model network.

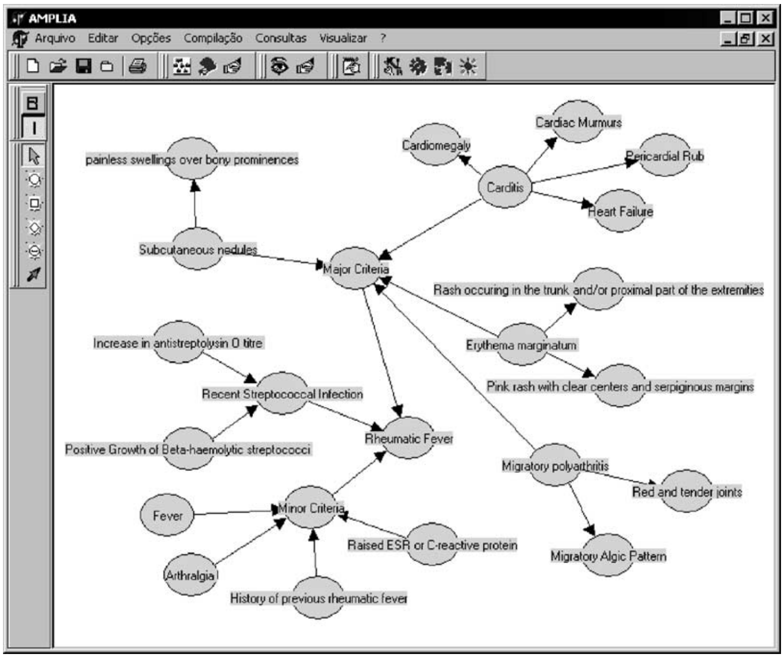


Fig. 14. An incomplete model network.

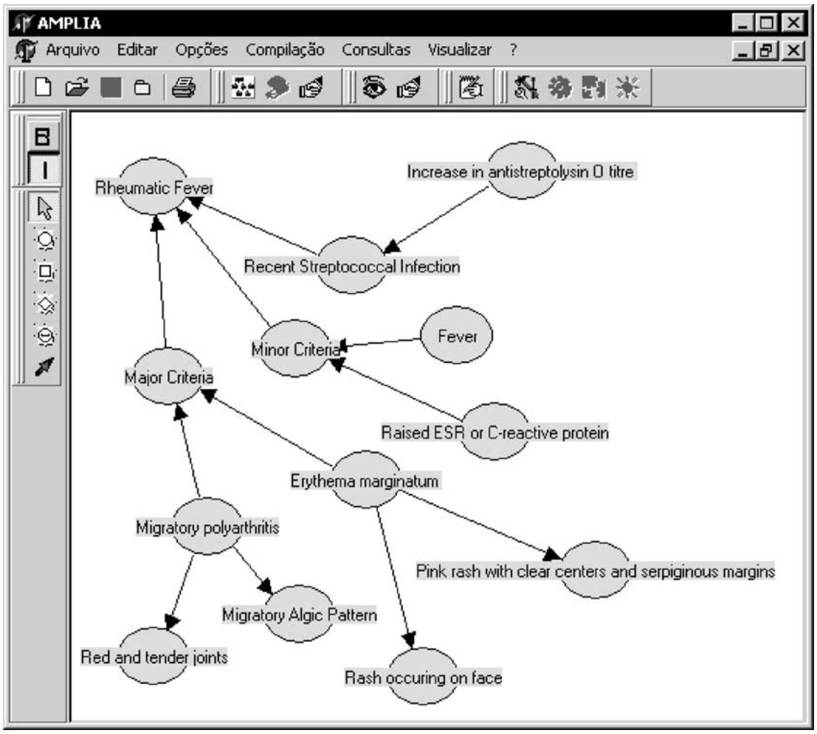


Fig. 15. An unfeasible model network.

**Example 4.** Fig. 15 is an example of an unfeasible model network. Notice that the learner is basing his ''Rheumatic Fever'' diagnosis on two Major Criteria: (1) ''Migratory polyarthritis'', presented by patient; and (2) Erythema marginatum, not present in case, besides, he is misusing ''Rash occurring on face'' as a confirmation node for ''Erythema marginatum''. As seen on Table 3, ''Rash occurring on face'' is an excluder finding that rules out this major diagnostic criterion.

## 5. Conclusion and future work

The intelligent probabilistic learning environment, AMPLIA, is designed to support the construction of explanatory models in complex, uncertain domains, and to support diagnostic reasoning. Our example application domains are from medicine. Unlike most existing systems based on the Bayesian network approach, AMPLIA is designed as a medical diagnostic learning tool. The learner may construct explanatory models; evaluate their consequences qualitatively and quantitatively. The learner may also state diagnostic hypotheses and receive feedback about the usefulness of diagnostic investigations.

With the use of probabilistic reasoning technology it is possible to improve the product performance. Currently, the probabilistic reasoning has a wide acceptance all over the world, because it is considered an efficient and correct way of representing and dealing with uncertainty. The junction tree technique is the current state-of-the-art of probabilistic inference.

Concerning applications of AMPLIA, one of our co-operations is aimed at generating realistic models with the help of case data. These models will serve for diagnostic training.

MAS have been successfully employed in the development of applications in a large number of domains [19]. In this context, multi-agents approach is an interesting alternative because it makes it easier integration of several components of the AMPLIA environment (some were agentified, e.g. the SEAMED facilities). This approach enables a better distance support to the learner, customised guiding, besides setting a real partnership among the several agents of the system, both human and artificial. The use of MAS helped also the development of systems with user's participation (learner and physician). The result is a flexible system, both in what concerns evolution of knowledge and teaching practices, and in terms of inclusion of new features whose necessity is realised while using the environment.

For the future, the following AMPLIA framework's developments are expected:

- to build new applications for decision support in medical area, through modelling of new domains within healthcare field;
- to expand this computational tool in order that it could also evaluate influence diagrams. An influence diagram is a probabilistic network that represents the formalisation of a problem of decision making in an environment with uncertainty [21]; and
- to expand this computational tool in order that it could also make inferences in multiply sectioned Bayesian network [44].

At present, the system is being submitted to an evaluation process concerning the following aspects: network accuracy is being tested by way of case-control studies with real patients related to specific medical fields. (In Brazil, there is not a certification organisation for that matter yet.) In these studies the physician follows medical interview protocols that list all variables of importance present in the particular net being tested and that should be analysed during the interview.

For future work, we intend to describe the methodological process that is in use to compare probabilistic networks including qualitative and quantitative aspects between student and expert models.

## Acknowledgements

The authors gratefully acknowledge the Brazilian agencies CAPES, FINEP and CNPq for the partial support to this research project.

## References

[1] Andrade AF, Vicari RM, Jung JL, Bordini R, Jaques P. A computational model of distance learning based on Vygotsky's socio-cultural approach. In: Vassileva J, editor. Proceedings of the Workshop on AI-ED 2001 Multi-Agent Architecture for Distributed Learning Environments, San Antonio, 2001. Saskatoon: IAIED Society, vol. 765, 2001. p. 33–40.

[2] Andreassen V, Jensen FV, Andersen SK, Falck B, Kjaerulff U, Woldbye M, et al. Munin—an expert EMG assistant. In: Desmedt JE, editor. Computer-aided electromyography and expert systems. Amsterdam: Elsevier Science; 1989. p. 255–77.

[3] Akhras FN, Self JA. Modelling learning as a process. In: Boulay B, Mizoguchi R, editors. Artificial intelligence in education: knowledge and media in learning systems. Proceedings of the AI_ED 97 World Conference on Artificial Intelligence in Education, Kobe, Japan, 1997. Amsterdam: IOS Press; 1997. p. 418–25.

[4] Aronsky D, Haug P. Diagnosing community acquired pneumonia with a Bayesian network. In: Chute C, editor. Proceedings of the AMIA 98 Annual Symposium. Philadelphia: Hanley&Belfus; 1998. p. 632–6.

[5] Baker MJ. The function of argumentation dialogue in cooperative problem-solving. In: van Eemeren FH, Grootendorst R, Blair JA, Willard CA, editors. Proceedings of the Fourth International Conference on Argumentation (ISSA'98), Amsterdam. Amsterdam: SIC SAT Publications; 1998. p. 27–33.

[6] Belhot RV. Searching for new ways of teaching. In: Proceedings of the International Conference of Engineering Education, ICEE 1998, Rio de Janeiro, 1998. CD-ROM. http://www.ime.eb.br/icee98/papers/232.pdf.

[7] Bercht M, Vicari RM. Pedagogical agents with affective and cognitive dimensions. In: Sanchez J, editor. Proceedings of the V Congreso Iberoamericano de Informática Educativa, 2000, Vinã del Mar. RIBIE 2000. V Congreso Iberoamericano de Informática Educativa, RIBIE 2000: CD-ROM. Universidade do Chile.

[8] Bordini RH, Lesser V, Viccari RM, Bazzan AL, Janone RO, Basso DM. AgentSpeak(XL): efficient intention selection in BDI agents via decision-theoretic task scheduling. In: Castelfranchi C, Johnson WL, editors. Proceedings of the AAMAS 2002—autonomous agents and multi-agents systems, Bologna, 2002. Bologna: ACM Press; 2002. p. 1294–302.

[9] D'Amico C, Vicari RM, Toscani LV. Agents in distributed intelligent learning environments. In: Lemaître C, editor. Proceedings of the Workshop on Distributed Artificial Intelligence and Multi-Agent Systems, Xalapa, Mexico, 1996. Mexico: MIA Universidad Veracruzana; 1996. p. 86–97.

[10] Flores CD, Viccari RM, Höher CL, Ladeira M. Una experiencia en el Uso de redes probabilísticas en el diagnóstico médico. Inform Med, Argent 2001;8:25–9.

[11] Folckers J, Möbus C, Schröder O, Thole H. An intelligent problem solving environment for designing explanation models and for diagnostic reasoning in probabilistic domains. In: Frasson C, Gauthier G, Lesgold A, editors. Proceedings of the Third International Conference ITS'96, Montreal. Intelligent tutoring systems. Berlin: Springer (LNCS 1086); 1996. p. 353–62.

[12] Franklin R, Spiegelhalter D, Macartney F, Bull K. Combining clinical judgements and statistical data in expert systems: over the telephone management decisions for critical congenital heart disease in the first month of life. Int J Clin Monitor Comput 1989;6:157–66.

[13] van der Gaag LC, Druzdel MJ. Elicitation of probabilities for belief networks: combining qualitative and quantitative information. In: Besnard Ph, Hanks S, editors. Proceedings of the 11th Conference on Uncertainty in Artificial Intelligence. San Francisco (CA): Morgan Kaufmann; 1995. p. 141–8.

[14] van der Gaag LC, Renooij S, Witteman CLM, Aleman B, Taal BG. How to elicit many probabilities. In: Laskey KB, Prade H, editors. Proceedings of the 15th Conference on Uncertainty. San Francisco (CA): Morgan Kaufmann; 1999. p. 647–54.

[15] Gertner AS, Conati C, VanLehn K. Procedural help in Andes: generating hints using a Bayesian network student model. In: Proceedings of the 15th National Conference on Artificial Intelligence AAAI-98, Madison, WI, July 1998. Cambridge (MA): MIT Press; 1998. p. 106–11.

[16] Giraffa LMM, Vicari RM. Modelling an interactive ITS using a MAS approach: from design to pedagogical evaluation. In: Proceedings of the Third International Conference on Computational Intelligence and Multimedia Applications, ICCIMA'99, Nova Delhi, 1999. Delhi: IEEE Computer Science; 1999. p. 153–8.

[17] Heckerman D, Horvitz E, Nathwani B. Towards normative experts systems. Part I. The pathfinder project. Methods Inform Med 1992;31:90–105.

[18] Henrion M, Pradhan M, Del Favero B, Huang K, Provan G, O'Rorke P. Why is diagnosis using belief networks insensitive to imprecision in probabilities? In: Horvitz E, Jensen FV, editors. Proceedings of the 12th Annual Conference on Uncertainty in Artificial Intelligence, UAI 96, Reed College, Portland, 1–4 August 1996. San Francisco: Morgan Kaufmann; 1996.

[19] Jennings NR, Sycara K, Wooldridge M. A roadmap of agent research and development. In: Autonomous agents and multi-agent systems, vol. 1. Boston: Kluwer Academic Publishers; 1998. p. 75–6,

[20] Jensen FV, Lauritzen SL, Olsen KG. Bayesian updating in causal probabilistic networks by local computations. Computational statistics quarterly, vol. 4. Heidelberg: Physica; 1990. p. 69–282.

[21] Jensen F, Jensen FV, Dittmer SL. From influence diagrams to junction trees. In: Proceedings of the 10th Conference on Uncertainty in Artificial Intelligence (UAI 94). San Francisco: Morgan Kaufmann; 1994. p. 367–73.

[22] Kahn CE, Roberts LM, Wang K, Jenks D, Haddawy P. Preliminary investigation of a Bayesian network for mammographic diagnosis of breast cancer. In: Proceedings of the MIDAS Consortium, SCAMC95, Milwaukee, Wisconsin, 1995. p. 208–12.

[23] Korver M, Lucas PJF. Converting a rule-based expert system into a belief network. Med Inform 1993;18(3):219–41.

[24] Ladeira M, Vicari RM, Coelho H. MSID: multiple sectioned influence diagram. In: Proceedings of the Workshop on Probabilistic Reasoning in Artificial Intelligence, Atibaia, SP, 2000. International Joint Conference IBERAMIA/SBIA 2000—Proceedings of the IBERAMIA/SBIA 2000 Workshops. São Paulo (Brazil): Tec Art Editora; 2000. p. 56–64.

[25] Lesser VR. Reflections on the nature of multi-agent coordination and its implications for an agent architecture. In: Autonomous agents and multi-agent systems, vol. 1, no. 1. Boston: Kluwer Academic Publishers; 1998. p. 89–111.

[26] Mayo M, Mitrovic A. Optimising ITS behaviour with Bayesian networks and decision theory. Int J Artif Intell Educ 2001;12:124–53 [Copyright 2001, The International AIED Society].

[27] Middleton B, Shwe MA, Heckerman DE, Henrion M, Horvitz EJ, Lehmann HP, et al. Probabilistic diagnosis using a reformulation of the INTERNIST 1/QMR knowledge base. II. Evaluation of diagnostic performance. Methods Inf Med 1991;30:256–67.

[28] Möbus C. Towards an epistemology on intelligent problem solving environments: the hypothesis testing approach. In: Greer J, editor. Proceedings of the AI-ED 95, World Conference on Artificial Intelligence and Education, Washington, DC. Washington (DC): Association for the Advancement of Computing in Education (AACE); 1995. p. S.138–45.

[29] Luiz M, Helder C. On tailoring teams of software agents tuned for tasks. In: Proceedings of the Workshop on Multi-Agent Systems: Theory and Applications (MASTA-01), 10th Portuguese Conference on Artificial Intelligence. Portugal: Porto; 2001.

[30] Mora MC, Viccari RM, Lopes JGP, Coelho H. BDI models and systems: reducing the gap. In: Muller JP, Singh MP, Rao AS, editors. Proceedings of the ATAL'98, Paris, July, 1998. Paris: Computer Science Laboratory, University Pierre et Marie Curie; 1998.

[31] Morgan MG, Henrion M. Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis. Cambridge: Cambridge University Press; 1990.

[32] Ng K, Abramson B. A sensitivity analysis of pathfinder: a follow-up study. In: D'Ambrosio BD, Bonissone PSPP, editors. Proceedings of the Seventh Annual Conference on Uncertainty in Artificial Intelligence, UAI 91, University of California at Los Angeles, 13–15 July 1991. San Mateo (CA): Morgan Kaufmann; 1991.

[33] Pearl J. Belief networks revisited. Artificial intelligence, vol. 59. Amsterdam: Elsevier; 1993. p. 49–56.

[34] Pradhan M, Henrion M, Provan G, Del Favero B, Huang K. The sensitivity of belief networks to imprecise probabilities: an experimental investigation. Artificial intelligence, vol. 85, no. 1/2. Amsterdam: Elsevier; 1996. p. 363–97.

[35] Oliveira FM, Vicari RM, Coelho H. A topological approach to equilibration of concepts. In: Proceedings of the XI Brazilian Symposium on Artificial Intelligence, Fortaleza, Ceará, Brasil, October 17–20, 1994.

[36] Rolf B, Magnusson C. Developing the art of argumentation, a software approach. In: Proceedings of the Fifth International Conference on Argumentation—ISSA 2002, University of Amsterdam, June 2002.

[37] Russel S, Norvig P. Artificial Intelligence: a modern approach. New Jersey: Prentice-Hall; 1995. 932 p.

[38] Schwarz BB, Neuman Y, Gil J, Ilya M. Effects of argumentative activities on collective and individual arguments. In: Proceedings of the European Conference on Computer-Supported Collaborative Learning—Euro-CSCL 2001, Maastricht, 22–24 March, 2001.

[39] Silveira R, Vicari RM. Distance learning: from the intelligent tutor paradigm to a multi-agents architecture. In: San Diego SCS, Bruzzone AG, Uhrmacher A, Page EH, editors. Proceedings of the

WEBSIM'99—International Conference on WEB-Based Modelling and Simulation, San Francisco, CA, 1999.
[40] Silveira RA, Vicari RM. JADE—Java agents for distance education framework. In: Proceedings of the DEC2001—Eighth Annual Distance Education Conference, Texas A&M University, 2001, CD-ROM.
[41] Veerman A. Computer-supported collaborative learning through argumentation. Analysis and modelling of collaborative learning interactions. Thesis. Enschede: Print Partners Ipskamp; 2000.
[42] Vicari RM, Giraffa LMM. The use of multi-agent system to build intelligent tutoring systems. In: Dubois D, editor. Int J Comput Anticipat Syst, vol. 1, no. 1. 2002. ISSN/ISBN: 0094243X.
[43] Vicari RM, Costa E. Pragmatic attachment devices for conversations with tutors. Applied artificial intelligence: an international journal. Amsterdam: Elsevier, vol. 2, no. 3. 1988. p. 277–84.
[44] Xiang Y. A Probabilistic framework for co-operative multi-agent distributed interpretation and optimization of communication. Artificial intelligence, vol. 87. Amsterdam: Elsevier; 1996. p. 295–342.