INTELLIGENT DECISION SUPPORT IN INTENSIVE CARE MEDICINE

Manuel F. Santos, Paulo Cortez
Dep. Sistemas de Informação
Univ. do Minho
Guimarães, Portugal
{mfs, pcortez}@dsi.uminho.pt

Pedro Gago
Escola Superior de Tecnologia e Gestão do Instituto Politécnico de Leiria, Portugal
pgago@estg.ipleiria.pt

Álvaro Silva, Fernando Rua
Instituto de Ciências Biomédicas Abel Salazar
Porto, Portugal
a.moreirasilva@mail.telepac.pt

Abstract – This paper introduces the INTCare system, an intelligent decision support system for intensive medicine. The system aims at the automation of the Knowledge Discovery Process by using autonomous agents that are responsible for the various constituent steps. The system enables automation of data acquisition and model updating avoiding human intervention. We present the first impressions after the deployment of INTCare in a real environment (Intensive Care Unit of the Hospital de Santo António, Oporto, Portugal) where it is supporting the physicians' decisions by means of prognostic Data Mining models. In particular, these techniques are used to predict organ failure and mortality assessment. The main intention is to change the current reactive behaviour to a pro-active one, enhancing the Quality of Service.

Keywords: Knowledge Discovery from Databases, Decision Support Systems, Agent Based Systems, Intensive Medicine

I. INTRODUCTION

One can easily find examples where Knowledge Discovery from Databases (KDD) is used, however almost all of them consist of retrospective studies. Indeed, a great part of the concepts need to be corroborated by real-world applications, in order to obtain a valuable feedback of its effective use [12]. Intensive medicine, due its high volume and complexity data, is a rich field where to perform those tests [5].

In this work the INTCare system is presented. It is an intelligent Decision Support System (DSS) based on the KDD and the Agent-Based paradigms, to support intensive care medical activities. In particular, the system will aid at the early detection of organ failure of six systems (liver, respiratory, cardiovascular, coagulation, central nervous and renal) and mortality assessment. INTCare is being tested in the Intensive Care Unit (ICU) of Hospital Geral de Santo António (HGSA), Oporto, north of Portugal. The system was built so that it could use knowledge discovered in previous research [14, 15] while allowing an easy deployment in the ICU. In order to reuse existing models, the PMML language [18] was used. Also, to allow for easier deployment and maintenance, data acquisition and model updating were automated.

The following sections will introduce the problem definition and the state of the art, followed by the INTCare system description in terms of a logic based formalism. Next, a discussion will be set over the critical aspects of the INTCare system. The chapter will be concluded, showing also perspectives of future work.

II. THEORETICAL BACKGROUND

A. *Decision Support Systems and Intensive Care*

Past experience has shown that knowledge retrieved from experts is not sufficient for the complex real-world problems [7]. The focus has thus shifted to the gathering of knowledge directly from the data, using intelligent data analysis. In terms of intensive care medicine, the application of Data Mining (DM) techniques is new, although possessing a huge potential [5]. However, the number of studies were these techniques are applied in a real environment is very scarce.

Currently, even though patients in an ICU are constantly monitored, the monitors are only used to sound alarms when anomalous values are collected. Moreover, if not directly used at the time of measurement, the monitored data is usually discarded. Physicians make decisions relying on their experience and the available clinical data (e.g. patients chart sheets) and no use is made of the vast amounts of data, which are automatically collected through bedside monitors. To the authors knowledge, although several prognostic scores have been developed, there are no intelligent DSS systems used in ICUs for outcome or organ dysfunction prediction. Yet, both patient and physicians could benefit if high-level, reliable and timely information is available. Hence, the INTCare system aims at filling this gap. Finally, it should also be stressed that success is also dependant on the ability to overtake physician resistance [1].

B. *Knowledge Discovery from Databases and Data Mining*

The interest in KDD and DM arose due to the rapid emergence of electronic data management methods. In particular, within the Medicine arena this lead to a storage of databases with large, complex and multi-source information (e.g. text, images or numerical data). However, human experts are limited and may overlook important details. Furthermore, the classical data analysis (e.g. logistic regression) breaks down when such vast amounts of data are present. An alternative is to use automated discovery tools to analyze the raw data and extract high level information for the decision-maker [4]. The above goals involve the application of a plethora of Machine Learning algorithms including Decision Trees (DT) [10] and Artificial Neural Networks (ANNs) [6]. More recently, there has been an emergent DM

research area that involves the use of ensembles for supervised learning, where a set of classification/regression models are combined in some way to produce an answer [2]. This interest arose due to the discovery that ensembles are often more accurate than individual learners.
Despite all advantages it presents, the KDD process still demands human intervention, especially in the data preprocessing and data mining steps [3]. This requirement makes evaluating new models more difficult (or more expensive).

III. INTCARE SYSTEM

INTCare is an agent based system, composed by several semi-autonomous agents in charge for the functionalities inherent to the system. The term agent is a metaphor allowing various definitions, interpretations and taxonomies [17]. In the context of this work, the AIMA definition prosecuted by Russel and Norvig was adopted, stating that an agent is an entity capable of perceiving the environment and actuating on that environment [11].Conceptually, the INTCare system can be viewed as set of four subsystems (Figure 1): Data Entry, Knowledge Management, Inference and Interface.

All data acquisition activities are performed by the Data Entry sub-system, which conveys the incoming data into a Data Warehouse in a format suitable for use by the agents responsible for knowledge maintenance. The Knowledge Management subsystem maintains the prediction models used by the Inference sub-system, assuring their validity and updating them when indicated by the assessment parameters stored in the Performance Database. Finally, the Interface sub-system allows the interaction between the doctors and the system. Formally, the INTCare system is defined as a tuple $\Xi \equiv \langle C_{INTCare}, \Delta_{INTCare}, q_{pp}, q_{cde}, q_{dm}, q_{pf}, q_{mi}, q_{dr}, q_{pd}, q_{sc}, q_{int} \rangle$:

- $C_{INTCare}$ is the context and corresponds to a logical theory, represented as a triple áLg, Ax, Dñ, where Lg stands for an extension to the language of programming logic, Ax is a set of axioms over Lg, and D is a set of inference rules;
- $\Delta_{INTCare}$ is the set of bridge rules defining the interaction among the systems' components (the agents);
- $q_{pp}, \ldots, q_{int}$ are the system's agents.

This formalism corresponds to a logical framework, suitable to specify agent-oriented systems based on the notion of context logic, and some properties of object-oriented design such abstraction, encapsulation, modularity and hierarchy [13]. In this work, the agents are represented as logical theories with a specific context (different agents may involve different contexts). Several agents (i.e., contexts) can be put together and be able to reason about the behaviour of the entire system as a (heterogeneous) logical theory. Special rules called bridge-rules are applied to provide the interface among agents and systems of agents. These rules describe the agent's reactions to events occurring in their environment. The agents include a set of event types, and a set of time points. Next, the overall system is described, explaining the more important agents in some detail.

A. *Description of the System*

Data quality is an essential requirement for good Knowledge Discovery results. In order to ensure it, INTCare includes a Clinical Data Entry agent that is responsible for the capture of clinical data from the

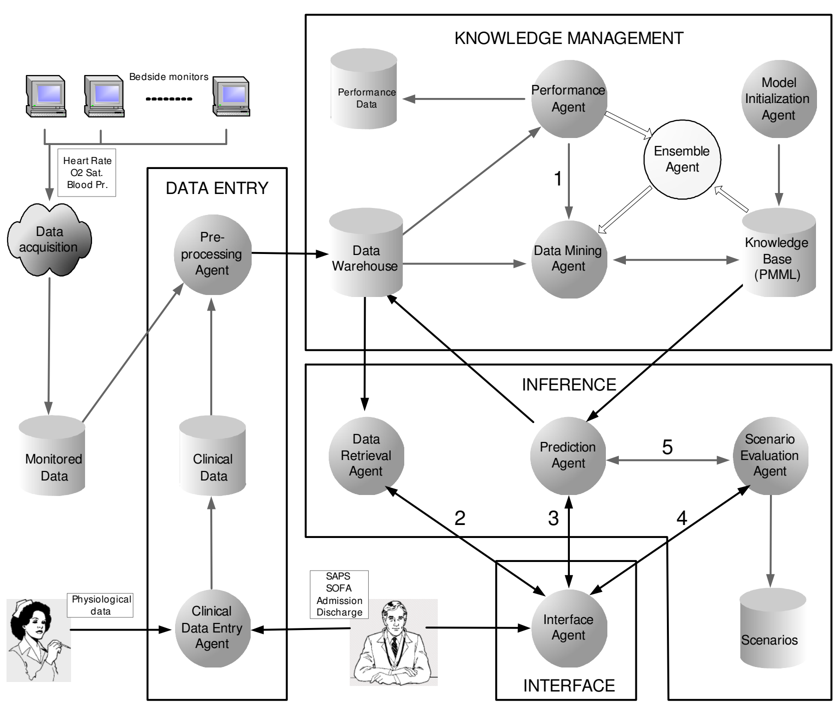

Figure 1– The INTCare System

medical and nursing staff. Also, monitored data is collected into a database making use of proprietary software provided by the equipment dealers. Nurses were asked to make sure the values automatically collected were as accurate as possible and were instructed on how to signal situations when readings are known to be inaccurate (e.g. when exams are being performed on the patient).

The data are then analysed by the Pre-Processing agent which is responsible for the correct linking of all the values in order to create a valid (even if limited in scope) medical record for the patient. It proceeds with the copy of the values entered by the medical and nursing staff (or recorded via the bedside monitors), examines them and derives new fields if necessary, such as Critical Events (that indicate that the recorded values are indicative of a possibly relevant medical condition) [14, 15].

Regarding the Data Mining agent, it is responsible for the application of artificial intelligence algorithms to train new models, whenever requested by the Performance agent, storing them into the Knowledge Base.

INTCare was specifically designed to work in a real environment. In order to enable adaptation to different situations its architecture allows not only off-line training but also incremental (i.e. update a model at each new record) and on-line learning (i.e. where past patterns are discarded if they are no longer relevant). Even though it increases the learning complexity, this added flexibility will allow the system to make full use of feedback information provided by physicians. Nevertheless, since in previous work interesting results were achieved, the tested DM algorithms (e.g. DT and ANN) are adopted as base concurrent models for organ failure detection and mortality assessment. The models are stored in the Knowledge Base, using the PMML specification language.

In a proactive way, the Performance agent continually consults the Data Warehouse for updates that allow statistics collection (e.g., discharge data that may or may not confirm a prediction made), as a base to calculate a set of assessment parameters maintained in the Performance Database. The evaluation metrics include classification accuracy, sensitivity and specificity values [15]. Whenever the collected statistics show that the performance has fallen bellow a predefined threshold a new model is requested. Most of INTCare's strength comes from the flexibility provided through this agent. Depending on the chosen learning strategy, the old model can be discarded or kept in the Knowledge Base.

This possibility will make it easier to conduct comparative studies. This characteristic will be further expanded through an Ensemble agent intended to enhance the system's predictive performance by combining several models in order to produce an answer. This agent is not yet implemented.

The Prediction agent answers user questions by applying the models in the Knowledge Base to the data stored in the Data Warehouse. The prediction is saved in the Data Warehouse so that the Performance agent is able to evaluate the model used.

What-if scenarios may be created and evaluating by using the Scenario Evaluation agent. Finally, the Interface agent allows interaction with the system by providing an easy way for doctors to request and receive prognostics and evaluate scenarios. The interaction among the INTCare's agents, is represented in terms of the bridge rules DINTCare described by Table 1 (where t stands for the system's time cycle).

IV. DEPLOYMENT AND EXPERIMENTATION

The INTCare is currently under use in the ICU of the Hospital Geral de Santo António in Oporto, Portugal. This unit contains ten monitored beds, treating around

	Bridge-Rule	Description
1	$\dfrac{C_{pf}: occurs(message\_dm, t)}{C_{dm}:[occurs(read\_dw, t) \wedge occurs(train, t) \wedge occurs(update\_models, t)]}$	When the *Performance* agent detects the need for the replacement of a model, the *Data Mining* agent is messaged, reads the available data in the Data Warehouse, trains a new model and finally saves it into the Knowledge Base.
2	$\dfrac{C_{int}: occurs(message\_dr, t)}{C_{dr}:[occurs(get\_data, t) \wedge occurs(send\_data, t)]}$	When triggered by the *Interface* agent, the *Data Retrieval* agent gets the patient's data from the Data Warehouse and sends it back to the *Interface* agent.
3	$\dfrac{C_{int}: occurs(re\_prediction, t)}{C_{pd}:[occurs(get\_models, t) \wedge occurs(predict, t) \wedge occurs(send\_data, t) \wedge occurs(update\_dw, t)]}$	When the *Interface* agent requests a forecast, the *Prediction* agent gets the adequate model from the Knowledge Base applies it and finally sends the results back to the *Interface* agent and updates the Data Warehouse.
4	$\dfrac{C_{int}: occurs(re\_scenario, t)}{C_{sce}:[occurs(re\_prediction, t) \wedge occurs(store\_scenario, t) \wedge occurs(send\_result, t)]}$	When the *Interface* agent requests a scenario evaluation it sends the necessary data to the *Scenario Evaluation* agent that then requests the forecast from the *Prediction* agent, stores the scenario in the Scenarios Database and finally returns the scenario results to the *Interface* agent.
5	$\dfrac{C_{sce}: occurs(re\_prediction, t)}{C_{pd}:[occurs(get\_models, t) \wedge occurs(predict, t) \wedge occurs(send\_data, t)]}$	When the *Scenario Evaluation* agent requests a forecast, the *Prediction* agent gets the adequate model(s) from the Knowledge Base applies it and finally sends the results back to the requesting agent.

Table 1– INTCare's Bridge Rules

400 patients each year. Deployment is planned to occur in three separate steps. First, the system is to be used with data acquired from a single bedside monitor. This allows for a less invasive introduction in the operating environment. Next, data will be acquired from all available monitors and the system will be used by all the doctors working in this ICU. Finally, INTCare will be extended to two other ICUs within the same region of influence.

In previous work, several DM models were applied to ICU data [14, 15]. The clinical data used was collected during the EURICUS II research programme [8], which involved a massive study in 42 ICUs from 9 countries during a period of 10 months, from 1998 to 1999. The database included thousands of daily records related to bedside measurements of critical ill patients, including features such as: the case mix - an information that remains unchanged during the patient's stay in the ICU (e.g. age or admission origin); the intermediate outcomes – being triggered from four monitored biometrics (e.g. the systolic blood pressure or urine output); and the patient's state – based on daily organ failure scores (e.g. SOFA index) and the final outcome (death/no death).

Organ dysfunction is measured by the SOFA (Sequential Organ Failure Assessment) score, which provides values within the range {0, 1, 2} (normal function) and {3, 4} (failure) [16]. Outcome is given by a 1 (death) or 0 (no death) condition, although intermediate values can be read as probabilities of occurrence. Both model's inputs include the patient's case mix and the intermediate outcomes.

This vast amount of information was modelled using a clustering framework to organ dysfunction diagnosis [14]; and an ANN approach [15], where easily acquired daily inputs were fed into Multilayer Perceptrons, in order to predict the failure of six organic systems (e.g. respiratory system). Artificial neural networks were also used to predict the final outcome. The resulting models were included as the base models in INTCare.

The current configuration includes one bedside monitor and two personal computers running Windows (one dedicated to data collection while the other has INTCare installed for use by the doctors). A TCP/IP network allows communication between the diverse components. Both the Clinical Data Entry agent and the Preprocessing agent were programmed using VB.Net. However, as the Xelopes library [9] was selected for model manipulation, Java is the programming language used in the remaining agents.

V. DISCUSSION AND TECHNOLOGY IMPACT

Currently, decisions related to end-of-life care (e.g. withdrawal of life support) are taken in a meeting where physicians exchange information and evaluate possible courses of action. Even though, in the majority of ICUs, patients are continuously monitored with several biometric sensors, this information is not fully taken into the decision process. This scenario changes with the introduction of the INTCare system, where all this data is stored and made available. Furthermore, INTCare is enriched with DM prognostic models, which can provide forecasts of both intermediate and final outcomes. In addition, the physicians may also evaluate different scenarios. The integration of INTCare in the ICU decision making procedure is straightforward as it makes available extra information to be considered in the group meeting.

Previous results obtained so far [14, 15] show the reliability of the DM prognostic models. Moreover, the availability problem is partly solved by using automatic data acquisition techniques. Regarding the physician resistance, it can be softened over time as this system is to be used as a support tool and its performance is expected to improve with time (as what happens with less experienced doctors). Adding to that, a computer based system has other advantages such as not being subject to between-physician variability in prognostic skill or beliefs regarding the construct of "futility". Furthermore model-based predictions may be more equitable because they do not incorporate value-based judgments regarding the ''worth'' of one life over another [1].

VI. CONCLUSIONS AND FURTHER WORK

This paper presented an Agent-Based System called INTCare, for ICU real time Intelligent Decision Support on clinical actuation using a KDD/DM approach for automatic knowledge acquisition. This approach implements all the KDD steps in an automatic way and is being tested in a real-world environment. INTCare enables a better overall predictive performance by the competition/combination of heterogeneous DM algorithms. As a consequence of the use of INTCare, a data warehouse is automatically filled with medical data that is of high value for future studies.

At this moment the INTCare system is being tested in the ICU of HGSA. The preliminary results even promising are not absolutely credible due the limited number of cases recorded. Further experiments will consider a greater volume of monitored data and studied cases and its impact in terms of: on-line, incremental and off-line learning; system tuning and decision process. Next, the ensemble agent will also be exploited towards a major performance.

REFERENCES

[1] Barnato, A. and Angus, D. (2004) Value and role of intensive care unit outcome prediction models in end-of-life decision making. Crit Care Clin 20:345-362

[2] Dietterich, T. (1997) Machine Learning Research: Four Current Directions. AI Magazine 18(4) 97-136

[3] Fayyad, U., Piatesky-Shapiro, G., Smyth, P. (1996) From Data Mining to Knowledge Discovery: An Overview. In Fayyad et al. (eds) Advances in Knowledge Discovery and Data Mining. AAAI Press / The MIT Press, Cambridge MA, pp 471-493

[4] Hand, D., Mannila, H., Smyth P (2001) Principles of Data Mining. MIT Press, Cambridge, MA

[5] Hanson, C., Marshall, B. (2001) Artificial intelligence applications in the intensive care unit. Crit Care Med 29(2) 427-435
[6] Haykin, S. (1999) Neural Networks – A Compreensive Foundation (2nd Edition). Prentice-Hall, New Jersey
[7] Lavrac, N., Keravnou, E., Zupan, B. (1997) Intelligent Data Analysis in Medicine and Pharmacology. Kluwer, Boston
[8] Miranda, D., Nat, R., Nijk, A., Schaufeling, W., Lapichino, G. (2003) Nursing activities score. Crit Care Med 31:374-382
[9] prudsys AG. (2004) Xeli's Intro - Introduction to XELOPES Version 1.2.3 (Xelopes Users Manual)
[10] Quinlan, J. (1986) Induction of Decision Trees. Machine Learning 1 (1) 81-106
[11] Russel S and Norvig P (2003) Artificial Intelligence – A Modern Approach 2nd Edition. Prentice Hall, New Jersey
[12] Saita, L., Neri, F. (1998) Learning in the "Real world" Machine Learning 30(2) 133-163
[13] Santos, M.F. (1999) Learning Classifier Systems in Distributed Environments. Ph.D. thesis, Universidade do Minho
[14] Silva, A., Pereira, J., Santos, M.F., Gomes, L., Neves, J. (2003) Organ failure prediction based on clinical adverse events: a cluster model approach. In: Hamza MH (ed) Proceedings of the third IASTED International Conference on Artificial Intelligence and Applications (AIA 2003). ACTA Press
[15] Silva, A., Cortez, P., Santos, M.F., Gomes, L., Neves, J. (2004) Multiple organ failure diagnosis using adverse events and neural networks. In: Seruca I et al. (eds) Proceedings of the 6th International Conference on Enterprise Information Systems (ICEIS 2004) Vol II, INSTICC, pp 401-408
[16] Vincent, J-L, De Mendonça, A., Cantraine, F. et al.(1998) Use of the SOFA score to assess the incidence of organ dysfunction/failure in intensive care units: results of a multicenter, prospective study. Crit Care Med 26: 1793-1800
[17] Weiss G (1999) Multiagent Systems A Modern Approach to Distributed Artificial Intelligence, MIT Press, Cambridge MA
[18] Data Mining Group, DMG; www.dmg.org, 2005