

# Agent-Based Decision Making through Intelligent Knowledge Discovery

Marina V. Sokolova<sup>1,2</sup> and Antonio Fernández-Caballero<sup>1</sup>

<sup>1</sup> University of Castilla-La Mancha, Polytechnical Superior School of Albacete, Campus Universitario s/n, 02071-Albacete, Spain  
smv1999@mail.ru, caballer@dsi.uclm.es

<sup>2</sup> Kursk State Technical University, Kursk, ul.50 Let Oktyabrya, 305040, Russia

**Abstract.** Monitoring of negative effects of urban pollution and real-time decision making allow to clarify consequences upon human health. Large amounts of raw data information describe this situation, and to get knowledge from it, we apply intelligent agents. Further modelling and simulation gives the new knowledge about the tendencies of situation development and about its structure. Agent-based decision support system can help to foresee possible ways of situation development and contribute to effective management and planning.

## 1 Introduction

Humans, as the active element of environment, interact with it and receive the feedback, that is not always positive though, and sometimes appears to be unhealthy or even harmful. Negative environmental impact upon human health is a complicated one and can be, summed up from permanent effects, such as sun radiation, Earth's geomagnetic field; from human's life and work conditions (characteristics of profession, quality of portable water, air, ) which also include characteristics of the place they live in - transport and industrial activity, use of vehicles; and from instable characteristics like meteorological conditions.

Humans adapt to natural unhealthy factors, often, at the cost of the organism's reversible and irreversible disfunctions, which can be then inherited and came to light as birth misfunctions and defects.

During a long time an expert-system paradigm remained a standard solution for environmental assessment issues, as this approach supplied researches with all the necessary instrumentation: sets of expert information and analytical support data. Later, appeared multi-agent system based applications, called by the requirements to real-time decision making and intelligent data mining.

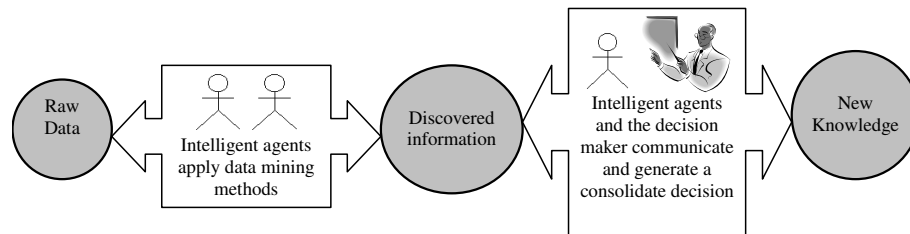
Some authors apply agent-approach to emergency assessment system development [1] related to traffic accidents response by allowing fast situation description, generation possible solutions and selection the decision maker for that case. MAS approach was utilized to increase the efficiency of hospital processes and to foresee possible effects of changes that can happen when expenses are reduced and the health-care budgets are restricted and stabilized. Different research institutes and organizations confirm the ecological contribution to illness

and death in adults and children [3]. The World Health Organization (WHO) declared the necessity of broader analysis of environmental risk factors over public health, with a strong emphasis on mortality [4].

## 2 The DSS Structure

Large amounts of raw data information describe the "environment - human health" system, but not all the information can be of use though. For the situation modelling we orient to factual and context information, presented in data sets and we use intelligent agents to extract it.

So, the information transforms from the initial "raw" state to the "information" state, which suggests organized data sets, models and dependencies, and, finally, to the "new information" which has a form of recommendations, risk assessment values and forecasts. The way the information changes is given on the figure below:



**Fig. 1.** The Information Change in ADSS

The hidden information is discovered by agents, but for new information construction not only intelligent agents, but knowledge of decision maker or expert are involved. The agent-based decision support system (ADSS) we are creating, provides these information changes [10].

The process of information change, shown on the Fig.1, corresponds to the MAS life cycle flow, which, in case of MAS counts the following steps: (1) domain and system requirements analysis; (2) design; (3) implementation; (4) verification; (5) maintenance [5].

The life cycle stages are supported as with theoretical methodologies, so with software tools. To determine the system and domain requirements, we suggest using an OWL-language based toolkit, as OWL has become a standard for ontologies description [6]. The Protege Ontology Editor and Knowledge-Base Framework [7] complies a set of procedures for ontology creation and analysis, offering a set of plug-ins covering viewers, problem-solving methods, knowledge trees, converters, etc. According to our proposal, ontologies can be represented by means of Protege and later may be incorporated into MAS.

In order to provide the system design to use the Prometheus Development Tool (PDT) [8], which provides a wide range of possibilities for MAS planning and implementation: the system architecture, the system entities, their internals and communications within the system and with outer entities. The most important advantages of PDT are an easy understandable visual interface and the possibility to generate code for JACK<sup>TM</sup> Intelligent Agents [9], which is used for MAS implementation, varification and maintenance.

### 3 Agent-Based Decision Making

As it comes from the agent nature, the agent 's automony means decision making [15]. Actually, the agents we use in the ADSS, belong to BDI agent model are intelligent by their nature, and have to make desicions every time while executing. Being acting within three logical levels of the ADSS, agents solve a set of problems. So, data search, fusion and pre-processing is being delivered by two agents, which do a number of tasks, following the workflow:

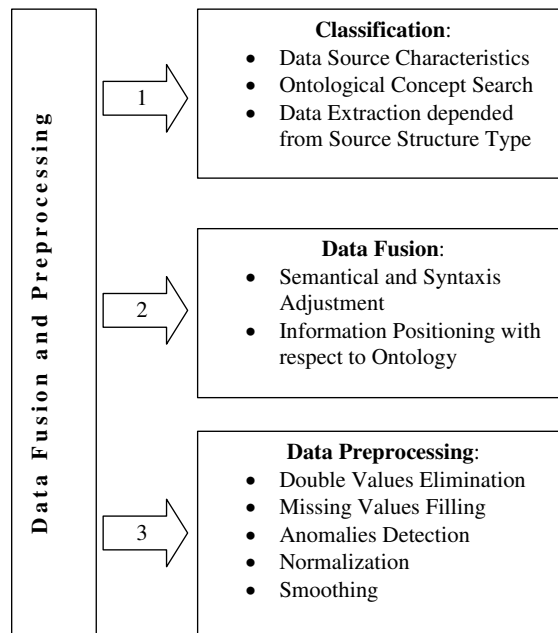
Information Search → Data Source Classification →  
Data Fusion → Data Preprocessing → Believes Creation

**Information Search** obliges agents to search for data storages which can obtain the necessary information, and then classify the found sources in accordance with their type, presence of ontology concepts and the file structure organization. After these tasks have been solved, the next work to do is to search the necessary values and their characteristics, agree with the domain ontology. The crucial task here is to provide the semantical and syntaxis identity of the retrieved values, saying they have to be preprocessed before being placed into ontology and agent 's belief. The properties for the "pollutant" concept include scale, period of measurement, region, value, pollutant class and "disease" include age, gender, scale, period of measurement, region, value, disease class.

If recovered values satisfy all the requirements imposed or have been adjusted properly, they are placed in ontology. Agent then pre-process the newly created data sets (they are checked for anomalies, double and missing values, normilyzed and smoothed) and create global belief, utilysed for futher calculations.

The second logical level is compleately based on autonome agents, which decide how to analyse data and use their abilities to do it. The prinipal tasks to be solved here are:

- to state the environmental pollutants which impacts on every age and gender group and determine if they are accociated with examined diseases groups;
- to create the models which explain dependencies between diseases, pollutants and groups of pollutants.



**Fig. 2.** The Tasks solved at the First Logical Level

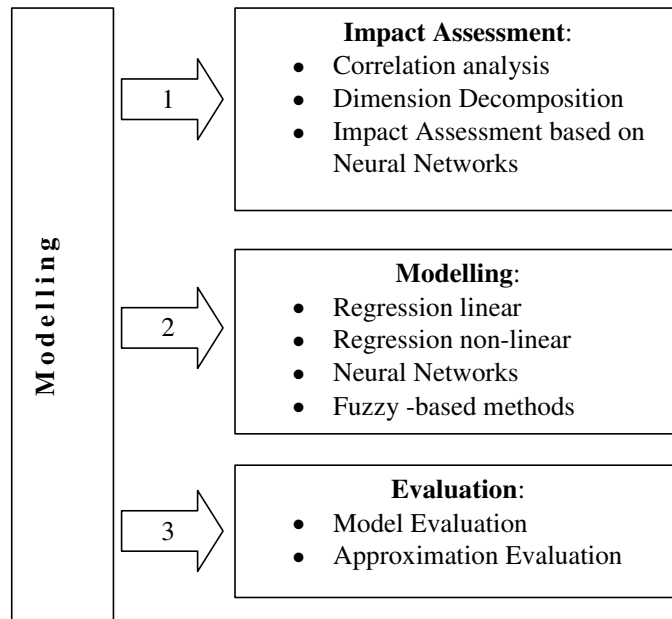
Here we are aimed to discover the knowledge in form of models, dependencies and accociations from the pre-processed information which comes from the previous logical layer. The workflow of this level includes the following tasks:

State Input and Output Information Flows → Create models → Assess Impact → Evaluate models → Select Models → Display The Results

Preliminary, we state the inputs (the pollutants) and the outputs (the diseases) for every model. The princpal errors to be avoided here are to include in the model input variables which are highly correlated to each other, and to include the variables which correlate with the dependent output variable, as in this case, we would not receive independent components and the model would not to be adequate. These difficulties are anticipated and warned by correlation analysis and factor dimension decomposition, which is based on neural-network approach [16].

The set of created models is wide and contain linear and non-linear regression models, neural-networks based and fuzzy-models and their modifications. The models are validated. The selection of the best models for every disease is realized by statistical estimators, which validate the approximation abilities of the models.

The third level of the system is dedicated to decision generation, so, both the decision making mechanisms and human-computer interaction are both important here. The system works in a cooperative mode, and it allows decision



**Fig. 3.** The Tasks solved at the Second Logical Level

maker to modify, refine or complete the decision suggestions, provided by the system and validate them. This process of decision improvement repeats until the consolidated solution is generated. The workflow is represented below:

```

State Factors for Simulation → State the Values of Factors →
Simulate → Evaluate Results → Check Possible Risk →
Display The Results → Receive Decision Maker Response →
Simulate → Evaluate Results → Check Possible Risk →
Display The Results

```

First, the decision maker selects the factor or factor, then states pollutants and the initial values for simulation. Agents simulate models, calculate risk assessment values and check for alarm. It is more than one recommendation generated.

Then the agents calculate the decision making criterion, select the best decision and recommends it to the system user. If he accepts it, his session of work with this factor is finished, and the recommendation is accepted. If not, the sequence of steps from **Receive Decision Maker Response** till **Display The Results** repeats until the consolidated decision is generated. As a result created models and dependencies become the foundation to receive new knowledge in form of variants of situation development and possible advantages and losses.

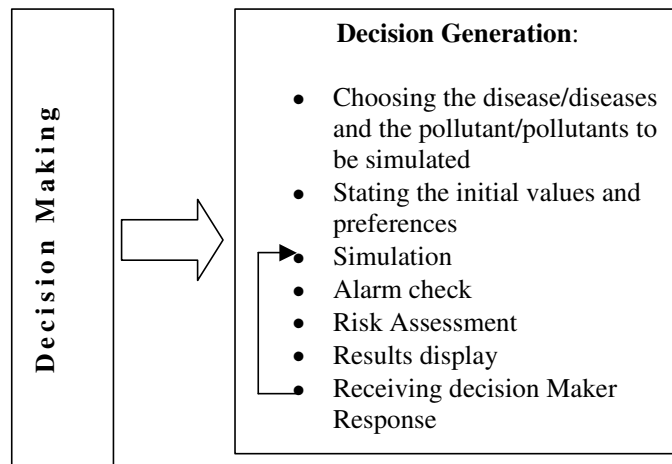


Fig. 4. The Tasks solved at the Third Logical Level

## 4 Conclusion

The agent-based Decision Making problem is a complicated one, especially for a general issue as environmental impact upon human health. Though supposing it to be a tractable problem, we should note some essential advantages we have reached, and some directions for future research. First, the ADSS supports decision makers in choosing the behaviour line (set of actions) in a such a general case, which is potentially difficult to analyse and foresee. As for any complex system, ADSS allows pattern predictions, and the decisioner's choice is to be decisive. Second, in spite of our time consuming modelling work, we are looking forward to both revise and improve the system and deepen our research. Third, we consider to make more experiments varying as with data structure, trying to apply the system to the other, but similar application field. Consequently, our future work can be drawn on various levels.

## Acknowledgements

Marina V. Sokolova is the recipient of a Postdoctoral Scholarship (Becas MAE) awarded by the Agencia Española de Cooperación Internacional of the Spanish Ministerio de Asuntos Exteriores y de Cooperación.

## References

1. Smirnov A., Pashkin M., Levashova T., Shilov N., Kashevnik A. Role-Based Decision Mining for Multiagent Emergency Response Management. AIS-ADM 2007, St. Petersburg, Russia, June 3-5, 2007, Proceedings. LNCS 4476 Springer, pp.178-191.

2. Herrler, R., Heine, C. Approaches and Tools to Optimize and Manage Clinical Processes. In: Application of Agents and Intelligent Information Technologies. IGI Publishing, 2007, USA, pp.39-65.
3. Kaiser R., Romieu I., Medina S., Schwartz J., Krzyzanowski M., Künzli N. (2004) Air pollution attributable postneonatal infant mortality in U.S. metropolitan areas: A risk assessment study. Environmental Health: A Global Access Science Source, vol.3, no. 1, p. 4.
4. World Health Organization. <http://www.who.int/en/>
5. ISO/IEC 12207 home page. <http://www.iso.org/iso/>.
6. OWL Web Ontology Language home page. (2004). <http://www.w3.org/TR/owl-features/>.
7. Prometheus Design Tool home page. <http://www.cs.rmit.edu.au/agents/pdt/>.
8. Padgham, L., Winikoff, M. (2002). Prometheus: A pragmatic methodology for engineering intelligent agents. Proceedings of the Workshop on Agent Oriented Methodologies (Object-Oriented Programming, Systems, Languages, and Applications), pp. 97-108.
9. Jack<sup>TM</sup> Intelligent Agents home page. <http://www.agent-software.com/shared/home/>.
10. Sokolova, M.V., Fernández-Caballero, A. (2007). A multi-agent architecture for environmental impact assessment: Information fusion, data mining and decision making. 9th International Conference on Enterprise Information Systems, ICEIS 2007.
11. Weiss, G. (2000). Multi-agent Systems: A Modern Approach to Distributed Artificial Intelligence. The MIT Press.
12. Sokolova M. Thesis for Doctorate Degree. Kursk, Kursk State Technical University press, 2004.