

AIR_POLLUTION_Onto: an Ontology for Air Pollution Analysis and Control

Mihaela M. Oprea

University Petroleum-Gas of Ploiesti, Department of Informatics,

Bd. Bucuresti nr. 39, Ploiesti, 100680, Romania, mihaela@upg-ploiesti.ro

Abstract The paper describes an ontology for air pollution analysis and control, AIR_POLLUTION_Onto, and presents its use in two case studies, an expert system, and a multiagent system, both dedicated to monitoring and control of air pollution in urban regions.

1 Introduction

The last decade has registered a strong challenge on the improvement of our environment quality under the international research framework of durable and sustainable development of the environment. The main concern of this challenge is to assure a healthy environment (air, water and soil) that allow the protection of ecosystems, and human health. One of the key aspects of this challenge is the air pollution control in urban regions with industrial activity [9]. In this context, more efficient tools has to be developed to solve the current environmental problems. Artificial intelligence provides several techniques that can solve efficiently such problems which have a high degree of uncertainty (see e.g [1] and [11]). The knowledge-based approach and the multi-agent systems (MAS) approach ([17]) offer ones of the best solutions to the environmental problems, as they reduce their complexity by structuring the domain knowledge from different sources in knowledge bases [10]. We have used the solution of expert systems [14], as the main sources of knowledge are given by the human experts as well as by the heuristic rules generated through machine learning techniques. The expert system DIAGNOZA_MEDIU was developed for the air pollution state diagnosis and control in urban regions with industrial activity. The solution of MAS was recently adopted by our research group in a postdoctoral research project running at our university. These types of approaches need to use an ontology specific to the expertise domain. Thus, we have developed an ontology dedicated to air pollution analysis and control, AIR_POLLUTION_Onto.

Please use the following format when citing this chapter:

Oprea, M.M., 2009, in IFIP International Federation for Information Processing, Volume 296; *Artificial Intelligence Applications and Innovations III*; Eds. Iliadis, L., Vlahavas, I., Bramer, M.; (Boston: Springer), pp. 135–143.

2 Air Pollution Analysis and Control

The main air pollutants are carbone dioxide (CO_2), carbon monoxide (CO), nitrogen dioxide (NO_2) and nitrogen oxides (NO_x), suspended particulates (particulate matters: respirable PM_{10} , and fine $\text{PM}_{2.5}$), sulfur dioxide (SO_2), ozone (O_3), lead (Pb), volatile organic compounds (VOC) etc. The concentrations of the air pollutants are measured in specific sites and compared to the standard values, according to national and international reglementations. The air pollutants have different dispersion models, and several mathematical models are used for the description of the relationships between environmental protection and meteorological factors (see e.g. an analysis of NO_2 and PM concentrations contribution to roadside air pollution [7]). Moreover, there are a lot of unpredictable factors that may influence the degree of air pollution, and it is quite difficult to establish with certainty which are the causes of an increase or of a decrease of an air pollutant concentration. The inclusion of most of the factors with their associated uncertainty degree would increase too much the complexity of the mathematical models, thus making them inefficient to solve real-time problems. The solution of a knowledge-based approach is an alternative to the mathematical models, as it allows the integration of multiple sources of knowledge in a knowledge base used by an inference engine that can deal also with uncertainty [12].

Prevention is an important step to air pollution control, and include different measures specific to each type of air pollutant and source of pollution. Some air pollution control strategies includes emission abatement equipment (e.g. wet and dry scrubbers, cyclones, bag filters), a policy of air pollution dispersion and dilution (e.g. a chimney of adequate height so that the pollution returned to ground level it poses no risk to health), change the process technology (e.g. fuel change, combined heat and power plant), change the operating patterns (e.g. alter the time that a process causes peak emissions), relocation (e.g. change the location of the process to have less impact on the urban and rural region).

3 Development of the Ontology AIR_POLLUTION_Onto

An ontology is an abstraction about the relevant aspects of the world. According to [4], an ontology is a specification of a conceptualization. The definition of an ontology involves the definition of a vocabulary with different terms for concepts (classes in the ontology), properties of concepts (slots in the ontology), relations between concepts, and instances. Also, a set of axioms must be defined. The axioms include the restrictions on the properties of the concepts (named facets of the slots). During the development of an ontology the following steps must be followed [16]: (1) define classes in the ontologies; (2) arrange the classes in a taxonomic hierarchy; (3) define slots and describe the allowed values for these slots; (4) fill in the values for the slots of the instances. A knowledge base is created by

defining individual instances of the classes filling in specific slot value information and additional slot restrictions. The basic types of relations that could appear between concepts are *ISA*, *AKO*, *HAS*, *PART_OF*, *SAME_AS*. They express the relations of membership, inclusion, composition and synonymity. Different names for these relations are *INSTANCE_OF*, *SUBCLASS_OF*, *SAME_AS*, *EQUIVALENT_CLASS*.

An ontology has two parts, the upper level ontology and the lower level ontology. The upper level ontology [5] is a set of general and well-known concepts shared by the most part of the human knowledge. The lower level ontology is a set of domain specific concepts that are shared by the human experts. The development of an ontology could be done by using several public available software. In our work we have used Protégé [15], a Java-based ontology editor.

To develop the air pollution domain ontology AIR_POLLUTION_onto, we have identified the terms specific to the analysis and control of air pollution. The ontology includes general terms (e.g. ENVIRONMENT, INDUSTRY, SOIL, WATER, WIND, WEATHER), and specific terms (e.g. AIR_POLLUTANT, POLLUTION_SOURCE, EMISSION, IMISSION, SO2, PM). After the identification of the general and specific terms, they are defined, characterized by their properties and constraints, and classified in taxonomies or hierarchies of classes. Also, the relationships between concepts are specified. Examples of relations between concepts are given below.

```
ISA(HUMIDITY, WEATHER_PARAMETER); ISA(FACTORY, POINT_SOURCE);
AKO(AIR_POLLUTANT, POLLUTANT); HAS(AIR_POLLUTANT,
CONCENTRATION);
```

The ontology has to include some meteorological terms as they are important for the analysis and control of air pollution. Such meteorological terms are RAINFALL, WIND, TEMPERATURE, RELATIVE HUMIDITY, ATMOSPHERIC PRESSURE. Each meteorological factor has a variability and seasonality.

Figure 1 shows a sequence from the ontology classes hierarchy in Protégé. Figure 2 shows a part from the ontology hierarchy with terms specific to air pollution.

The slots of each term are defined. In Protégé this means to specify for each slot the name, the cardinality (single, multiple), the value type (integer, float, string, symbol, boolean, class, instance, any) and other facets. Figure 3 shows the slots for class PM (term Particulate Matters). These slots are *Concentration level*, *MAC* (maximum admissible concentration for PM), *Network Site* (the site number in the monitoring network), *Period* (the period of concentration level measurement), *Position of the site* (position given on a map, e.g. A3), *Prevention measure*, and *Type of pollution source*.

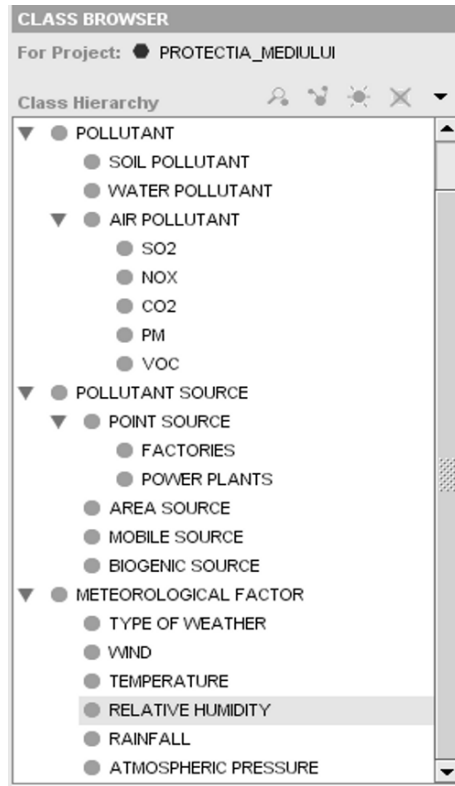


Fig. 1. Ontology hierarchy of classes in Protégé.

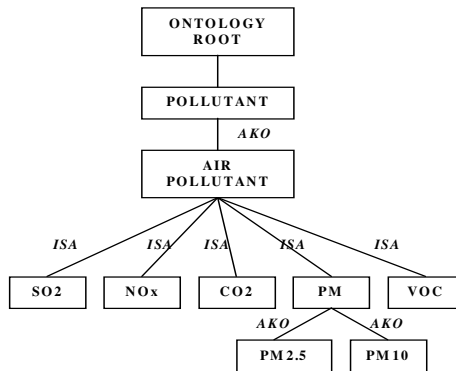


Fig. 2. Hierarchy of air pollution terms.

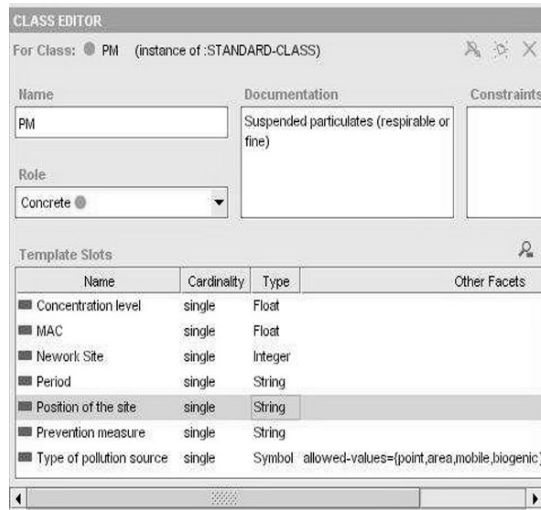


Fig. 3. Slots of PM class (Protégé screenshot).

Restrictions could be applied to the slots of a class. Examples of generic restrictions are given below.

$Concentration_level \leq MAC$ (admissible concentration);

$Concentration_level > MAC$ (inadmissible concentration);

$Rainfall_level > 1$ mm (it rains); $Rainfall_level < 1$ mm (it does not rain)

where MAC is the maximum admissible concentration for a specific air pollutant.

AIR_POLLUTION_Onto could be extended with new terms to be used by different types of intelligent systems such as expert systems and multiagent systems.

4 Case Studies

We have used AIR_POLLUTION_Onto in two systems, DIAGNOZA_MEDIU, an expert system for air pollution analysis and control, and MAS_AirPollution, a multi-agent system for air pollution monitoring and control, both systems being applied in urban regions.

4.1 DIAGNOZA_MEDIU

The main purpose of DIAGNOZA_MEDIU is to provide qualitative information to a decision support system that is used in the environmental protection management. The system was implemented in VP-Expert, a rule-based expert system

generator. The architecture of the prototype expert system is composed by a knowledge base, a backward inference engine, an explanation module, a knowledge acquisition module, and a user interface. Figure 4 shows the architecture of DIAGNOZA_MEDIU. The knowledge base has three main components, a rules base and a facts base, permanent and temporary. The expert system uses data from the environmental and meteorological databases as well as some forecasting data that are provided by a feed-forward artificial neural network (previously developed and presented in [13]).

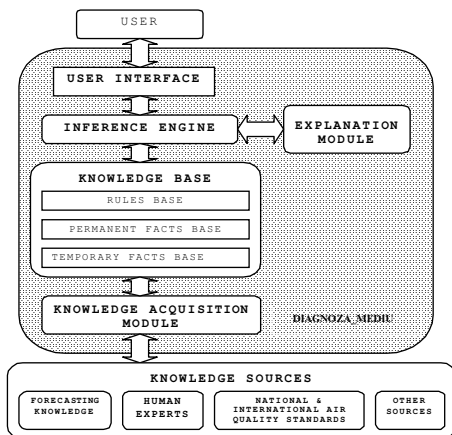


Fig. 4. The architecture of DIAGNOZA_MEDIU.

The knowledge base is generated from the ontology AIR_POLLUTION_Onto. Production rules are used as a knowledge representation method. To express the uncertain knowledge we have used a set of terms with linguistic certainty values that constitutes the verbal scale that the human experts and the users will use to express their degree of confidence in rules and facts. The fuzzy interval (similar with that used in [3]) includes nine ordered values for the goal variable **risk of air pollution**: *impossible*, *almost impossible*, *slightly possible*, *moderately possible*, *possible*, *quite possible*, *very possible*, *almost sure*, and *sure*. The ontology was extended with terms related to all these fuzzy values.

The rules base has mainly three types of rules: behaviour, decision and control. Examples of such rules are given bellow.

RULE BT_15 (*Behaviour rule*)

IF DT \geq 7 AND T \geq 37 AND MF = not_changed THEN TPRED = much_higher;

RULE CPM_12 (*Control rule*)

IF TPRED = much_higher AND $IP_{PM} \geq MAC_{PM} - 0.005$ THEN

Risk_pollution = quite_possible; Pollution_control =

Control_measure_for_PM_reduction;

// due to higher temperature and very close value of IP to MAC for PM

RULE DPM_7 (Decision rule)

IF Risk_pollution = quite_possible AND Air_Pollutant = PM THEN

Warning = “Area with potential air pollution risk – quite possible exceed of PM”

Prevention_measure = “Change the chemical process technology / Change the filter”;

where DT is the duration (in days) of air temperatures greater than 37°C, T is the minimum value of the maximum temperatures measured in the last seven days, MF represents a global parameter that refers to the evolution of the meteorological factors (wind, rainfalls, etc) in the next period of time (i.e. next 2 days minimum), TPRED is the symbolic value of the predicted temperature, IP is the predicted value for an air pollutant indicator (e.g. the concentration level), and MAC is the maximum admissible value for an air pollutant indicator.

The three rules given above are applied to specific sites, and thus the chemical plants, and other air pollution sources are known, and the warning as well as the prevention and counter measures specified in the generic rule DPM_7 are directly related to them. The instances specific to the urban region where DIAGNOZA_MEDIU is applied are included in the ontology.

4.2 MAS_AirPollution

AIR_POLLUTION_Onto is currently used in MAS_AirPollution, a multi-agent system that is under development as a simulation. The goal of the MAS is to monitor and control air pollution in urban regions.

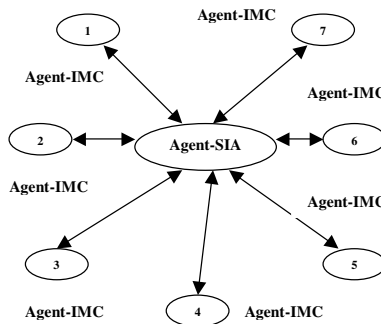


Fig. 5. The architecture of MAS_AirPollution system.

Figure 5 presents its architecture. The system is composed by a set of intelligent monitoring and control agents (Agent_IMC, specific to each site of the monitoring network) and a supervisor intelligent agent (Agent_SIA) in a star like architecture. AIR_POLLUTION_Onto is shared by all agents. In case some agents have dif-

ferent ontologies we are using ontology mapping methods (see e.g. [6]). Some of these methods identify maximal similarity between concepts definitions that are terminologically heterogenous, others are using structural methods. As most of the discrepancies of ontologies might be due to different terminology we apply in our simulated MAS a lexicon-based method similar to that given in [8]. Still, in case of real time systems, as it shall be the case of our MAS, a better method is much proper, QOM (Quick Ontology Mapping, [2]).

The agents communicate by using FIPA ACL. Figure 6 shows an example of a FIPA ACL message exchanged between two agents.

```
( inform
  :sender Agent-IMC_3
  :receiver Agent-SIA
  :content
    (warning (increase Temperature)
      (Concentration-level PM 0.14651)
      (meteo_factor wind weak)
    )
  :ontology AIR_POLLUTION_Onto
  :language sl
)
```

Fig. 6. Example of a FIPA ACL message exchanged in the system MAS_AirPollution.

In case one of the agent discovers during conversation that some concepts from the current message are unknown, it will generate a particular feedback message to establish a mapping of its private ontology with that of the sender agent. Usually, only a part of the ontology is mapped.

5 Conclusion and Future Work

The application of knowledge-based approaches and multi-agent systems to air pollution analysis and control has to include as a mandatory step in the system development methodology the design and implementation of a specific ontology. In this paper we have focused on such an ontology, AIR_POLLUTION_Onto, that was initially developed for DIAGNOZA_MEDIU, an expert system for air pollution analysis and control, and later was extended and applied to MAS_AirPollution, a MAS for monitoring and control of air pollution in urban regions. The use of the ontology in an expert system helps the generation of the knowledge base that should be coherent, non-redundant and complete, while in case of a MAS it is an important support for inter-agents communication. In an open multiagent system, it is important to apply ontology mapping to achieve interoperability. As a future work we shall extend AIR_POLLUTION_Onto for other applications in the area of environ-mental protection.

Acknowledgement The research reported in this paper was partially funded by the Romanian Education and Research Ministry and the National Council of Academic Research (CNCSIS) under the research grant AT429/2003.

References

1. Buisson, L., Martin-Clonaire, R., Vieu, L., Wybo, J.-L.: Artificial Intelligence and Environmental Protection: A Survey of Selected Applications in France, *Information Processing 92*, vol. II, Elsevier (1992) 635-644.
2. Ehrig, M., Staab, S.: Efficiency of Ontology Mapping Approaches, University of Karlsruhe, research report (2004).
3. Godo, L., Lopez de Mantaras, R., Sierra, C., Verdaquer, A.: MILORD: The architecture and the management of linguistically expressed uncertainty, *International Journal of Intelligent Systems*, 4 (1989) 471-501.
4. Gruber, T.: Towards principles for the design of ontologies used for knowledge sharing, *International Journal of Human-Computer Studies*, 43(5/6) (1995) 907-928.
5. Guarino, N., Giaretta, P.: Ontologies and Knowledge Bases: Towards a Terminological Clarification, in *Towards Very Large Knowledge Bases: Knowledge Building and Knowledge Sharing*, N. Mars (Ed), IOS Press (1995) 25-32.
6. Kalfoglou, Y., Schorlemmer, M.: Ontology mapping: the state of the art, *The Knowledge Engineering Review*, 18(1) (2003) 1-31.
7. Lam, G.C.K., Leung, D.Y.C., Niewiadomski, M., Pang, S.W., Lee, A.W.F., Louie, P.K.K.: Street-level concentrations of nitrogen dioxide and suspended particulate matter in Hong Kong, *Atmospheric Environment*, 33(1) (1999) 1-11.
8. Li, J.: LOM: A Lexicon-based Ontology Mapping Tool, Teknowledge Corporation, research report (2004).
9. Moussiopoulos, N. (Ed): *Air Quality in Cities*, Springer, Berlin (2003).
10. Oprea, M.: A case study of knowledge modelling in an air pollution control decision support system, *AiCommunications*, IOS Press, 18(4) (2005) 293-303.
11. Oprea, M., Sánchez-Marrè, M. (Eds): Proceedings 16th ECAI 2004 Workshop Binding Environmental Sciences and Artificial Intelligence (BESAI-4) (2004).
12. Oprea, M.: Modelling an Environmental Protection System as a Knowledge-Based System, *International Journal of Modelling and Simulation*, ACTA Press, 24(1) (2004) 37-41.
13. Oprea, M.: Some Ecological Phenomena Forecasting by using an Artificial Neural Network, *Proceedings 16th IASTED International Conference on Applied Informatics AI98*, Garmisch-Partenkirchen, ACTA Press (1998) 30-32.
14. Page, B.: An Analysis of Environmental Expert Systems Applications, *Environmental Software*, 5(4) (1990) 177-198.
15. Protégé-2000: <http://protégé.stanford.edu>
16. Uschold, M., King, M.: Towards a Methodology for Building Ontologies, research report AIAI-TR-183, University of Edinburgh (1995).
17. Wooldridge, M.: *Introduction to Multiagent Systems*, John Wiley and Sons, New York, (2002).