Development of an Ecological Decision Support System[1]

Frits van Beusekom*, Frances Brazier**, Piet Schipper***, Jan Treur**

*International Plant Technology Service (IPTS), Bilthoven

**Vrije Universiteit Amsterdam
Department of Mathematics and Computer Science, Artificial Intelligence Group
De Boelelaan 1081, 1081 HV Amsterdam
Email: {frances, treur}@cs.vu.nl URL: http://www.cs.vu.nl/~{frances, treur}

***State Forestry Department, Dutch Ministry of Agriculture
and Nature Management, Driebergen

Abstract. In this paper a knowledge-based decision support system is described that determines the abiotic (chemical and physical) characteristics of a site on the basis of in-homogeneous samples of plant species. Techniques from the area of non-monotonic reasoning are applied to model multi-interpretable input information.

1 Introduction

Plants only grow in environments where conditions are appropriate. Knowledge of which set of factors is necessary for species to germinate and complete their life-cycle, has been acquired by experts over a large number of years. This knowledge of environmental preferences of plant species makes it possible to derive information about a terrain's abiotic (physical and chemical) characteristics on the basis of the plant species found. More specifically, experts are able to derive the abiotic conditions of a site in terms of acidity, nutrient value and moisture from the abiotic preferences of the species comprising the vegetation.

If knowledge on abiotic preferences of plant species is available, nature managers can use their knowledge of the plant species found in a specific terrain to determine its abiotic conditions. Often, however, these managers do not possess such detailed knowledge. An Environmental Knowledge-based System, EKS, has been designed to support them in this decision making process. Once the abiotic conditions of a terrain have been determined, measures can be proposed to maintain or to improve the quality of the site.

The specific domain of application in the current implementation is grasslands. The knowledge-based system, the development of which was funded by the

[1] In: A.P. del Pobil, J. Mira, and M. Ali (eds.), Tasks and Methods in Applied Artificial Intelligence (Proceedings of the 11th International Conference on Industrial and Engineering Applications of AI and Expert Systems, IEA/AIE'98, vol. II), Lecture Notes in AI, vol. 1416, Springer Verlag, 1998, pp. 815-825

organisations International Plant Technology Services (IPTS) and the State Forestry Department of the Dutch Ministry of Agriculture and Nature Management (Staatsbosbeheer), is based on knowledge acquired from experts in the fields of Plant Ecology, Eco-hydrology, and Soil Sciences. Acquiring consensus between experts on the response of individual plant species with respect to specific abiotic conditions is one of the main achievements of this project.

The observations made in the field, a sample, can often be interpreted in various ways. The meaning of the presence of a specific plant species depends on the view the expert takes. To be able to determine which conclusions can be reached, the expert needs to identify and interpret the different views possible. In this paper, the domain of application is introduced in Section 2. In Section 3 the knowledge-based system EKS is described. In Section 4 the reported results are discussed.

2 Domain of Application

Experts identify the abiotic conditions of a terrain on the basis of plant species they encounter. These conditions are expressed as values for each of the abiotic factors: *acidity* (basic, neutral, slightly acid, fairly acid, acid), *nutrient value* (nutrient poor, fairly nutrient rich, nutrient rich, very nutrient rich) and *moisture* (very dry, fairly dry, fairly moist, very moist, fairly wet, very wet).

In an abiotic homogeneous situation, a common set of abiotic conditions are found that are (by definition) shared by the plant species encountered on the site. A technique to determine the abiotic conditions in this case is described in Section 2.1. In practice, however, samples often include groups of plant species that, according to the knowledge available, could not possibly be found under the same abiotic conditions. It may also be the case that the sample has been taken from a site where the abiotic conditions vary (for instance, on a terrain forming the transition between dry and wet ground). But it also may happen that the heterogeneity can be explained by specific interaction between the species in the sample. In many cases it is not possible to eliminate heterogeneity from a sample. Therefore a method is required that respects heterogeneity and provides a basis for the analysis of the ingredients of the heterogeneity. A method to determine which compatible groups of plant species can be distinguished within such a sample is described in Section 2.2.

2.1. Homogeneous Sample: Greatest Common Denominator

In a homogeneous situation, at least one set of abiotic conditions can be found that is shared by all species inhabiting the site. An example of a group of plant species that can occur in a homogeneous situation is used to illustrate this technique. Examination of these plant species, depicted in Table 1, shows all possible values for each of the three abiotic factors, for each of the plant species. For example, the abiotic requirements of *Caltha palustris*, are:

- very moist or fairly wet,
- basic, neutral or slightly acid,
- nutrient poor, fairly nutrient rich or nutrient rich terrain.

For the species *Poa trivialis* a terrain needs to be

- fairly moist, very moist or fairly wet,
- basic or neutral,
- nutrient rich or very nutrient rich.

If both species occur in a terrain, this implies that the terrain can only be:
- very moist or fairly wet,
- basic or neutral,
- nutrient rich.

The occurrence of a single species restricts the possible abiotic conditions of the terrain, but the occurrence of species in combination restricts the possible abiotic conditions even further.

	Moisture						Acidity					Nutrient Value			
Species	vd	fd	fm	vm	fw	vw	bas	neu	sac	fac	ac	np	fnr	nr	vnr
Angelica sylvestris				x	x		x	x					x	x	
Caltha palustris ssp palustris				x	x		x	x	x			x	x	x	
Carex acutiformis				x	x		x	x					x	x	
Carex acuta				x	x	x	x	x	x				x	x	x
Deschampsia caespitosa			x	x	x		x	x	x				x	x	x
Epilobium parviflorum			x	x			x	x	x				x	x	
Equisetum palustre			x	x	x	x	x	x	x			x	x	x	
Galium palustre				x	x		x	x	x			x	x	x	x
Glyceria fluitans				x	x	x	x	x	x	x			x	x	x
Juncus articulatus				x	x		x	x	x			x	x	x	x
Lathyrus pratensis			x	x			x	x	x				x	x	
Myosotis palustris				x	x		x	x	x				x	x	x
Phalaris arundinacea			x	x	x	x	x	x						x	x
Phleum pratense ssp pratense			x	x			x	x						x	x
Poa trivialis			x	x	x		x	x						x	x
Scirpus sylvaticus				x	x	x	x	x	x				x	x	

Table 1. A homogeneous sample

Analysis of the abiotic conditions for all plant species presented in Table 1 shows that only a limited number of possibilities (but more than one) can be found in which all of these plant species can occur together. This *greatest common denominator* for the given plant species is defined by the following set of abiotic conditions:
- very moist
- basic or neutral
- nutrient rich

The combination of these plant species indicates that a site on which these plant species are found necessarily fulfils these conditions.

	Moisture						Acidity					Nutrient Value			
Species	vd	fd	fm	vm	fw	vw	bas	neu	sac	fac	ac	np	fnr	nr	vnr
Angelica sylvestris				x	x		x	x					x	x	
Anthoxanthum odoratum		x	x	x					x	x		x	x		
Caltha palustris ssp palustris				x	x		x	x	x			x	x	x	
Carex acutiformis				x	x		x	x					x	x	
Carex acuta				x	x	x	x	x	x				x	x	x
Carex nigra			x	x	x				x	x	x	x	x		
Carex panicea			x	x	x				x	x		x	x		
Carex riparia				x	x	x	x	x						x	x
Cirsium oleraceum				x	x		x	x					x	x	
Cirsium palustre				x			x	x	x			x	x	x	
Crepis paludosa			x	x	x		x	x	x				x	x	
Deschampsia caespitosa			x	x	x		x	x	x				x	x	x
Epilobium palustre			x	x	x				x			x	x		
Epilobium parviflorum			x	x			x	x	x				x	x	
Equisetum palustre			x	x	x	x	x	x	x			x	x	x	
Filipendula ulmaria				x			x	x	x			x	x	x	
Galium palustre				x	x		x	x	x			x	x	x	x
Glyceria fluitans				x	x	x	x	x	x	x			x	x	x
Juncus articulatus				x	x		x	x	x			x	x	x	x
Juncus conglomeratus		x	x	x					x	x		x	x		
Lathyrus pratensis			x	x			x	x	x				x	x	
Lotus uliginosus			x	x	x		x	x	x			x	x	x	
Lychnis flos cuculi				x	x		x	x	x				x	x	
Lysimachia vulgaris			x	x	x		x	x	x			x	x	x	
Myosotis palustris				x	x		x	x	x				x	x	x
Phalaris arundinacea			x	x	x	x	x	x						x	x
Phleum pratense ssp pratense			x	x			x	x						x	x
Poa trivialis			x	x	x		x	x						x	x
Scirpus sylvaticus				x	x	x	x	x	x				x	x	

Table 2. An inhomogeneous sample

2.2 Inhomogeneous Sample: Maximal Indicative Subsets

In the inhomogeneous case, a sample does not have a greatest common denominator of abiotic conditions. An example sample (from the Pommeren site) is shown in Table 2, together with the possible values for the three abiotic factors for each plant species. Focusing on the acidity of a terrain shows that *Angelica sylvestris* and *Carex acutiformis*, for example, only grow under basic or neutral conditions, whereas, for example, *Carex nigra* and *Carex panicea*, found in the same sample, only grow on a slightly or fairly acid site. These species, however, are all in the same sample. One common set of possible values of the abiotic factors for all plant species can not be derived.

A comparable analysis for all plant species in the same sample for the abiotic factors is required. To this purpose groups of plant species are identified: plant species that can all abide in the same abiotic conditions, defining a homogeneous group of plants. By grouping plant species into the largest possible homogeneous groups of plant species, *maximal indicative subsets* are derived.

These subsets are maximal with respect to compatibility of the plant species in the subset. In other words, all plant species in the sample that are compatible with the plants in a maximal indicative subset are in the subset. As shown in Table 3, in the example sample two maximal indicative sets of plant species can be distinguished. The *first maximal indicative subset* contains all plant species that can grow under

- very moist
- basic or neutral
- nutrient rich

conditions. The *second maximal indicative subset* contains all plant species that can grow under

- very moist
- slightly acid
- fairly nutrient rich

conditions.

Note that the two maximal indicative subsets share a number of plants (the intersection of the two subsets). These plants have a relatively broad amplitude of abiotic preferences. Note also that the conditions for the plant species that these two groups do not have in common are mutually exclusive with respect to acidity and (partially) nutrient value.

To interpret the abiotic heterogeneity indicated by the subsets additional knowledge is required. For example, in this case one expert interpretation is that the sample has been taken in a terrain that has a specific type of stratification (so-called rainwater lenses): thus providing an explanation for the two abiotic indicative sets of plant species.

	Moisture						Acidity					Nutrient Value			
Species	vd	fd	fm	vm	fw	vw	bas	neu	sac	fac	ac	np	fnr	nr	vnr
Angelica sylvestris				x	x		x	x					x	x	
Carex acutiformis				x	x		x	x					x	x	
Carex riparia				x	x	x	x	x						x	x
Cirsium oleraceum				x	x		x	x					x	x	
Phalaris arundinacea			x	x	x	x	x	x						x	x
Phleum pratense ssp pratense			x	x			x	x						x	x
Poa trivialis			x	x	x		x	x						x	x
Caltha palustris ssp palustris				x	x		x	x	x			x	x	x	
Carex acuta				x	x	x	x	x	x				x	x	x
Cirsium palustre				x			x	x	x			x	x	x	
Crepis paludosa			x	x	x		x	x	x				x	x	
Deschampsia caespitosa			x	x	x		x	x	x				x	x	x
Epilobium parviflorum			x	x			x	x	x				x	x	
Equisetum palustre			x	x	x	x	x	x	x			x	x	x	
Filipendula ulmaria				x			x	x	x			x	x	x	
Galium palustre				x	x		x	x	x			x	x	x	x
Glyceria fluitans				x	x	x	x	x	x	x			x	x	x
Juncus articulatus				x	x		x	x	x			x	x	x	x
Lathyrus pratensis			x	x			x	x	x				x	x	
Lotus uliginosus			x	x	x		x	x	x			x	x	x	
Lychnis flos cuculi				x	x		x	x	x				x	x	
Lysimachia vulgaris			x	x	x		x	x	x			x	x	x	
Myosotis palustris				x	x		x	x	x				x	x	x
Scirpus sylvaticus				x	x	x	x	x	x				x	x	
Anthoxanthum odoratum		x	x	x					x	x		x	x		
Carex nigra			x	x	x				x	x	x	x	x		
Carex panicea			x	x	x				x	x		x	x		
Epilobium palustre			x	x	x				x			x	x		
Juncus conglomeratus		x	x	x					x	x		x	x		

Table 3. Maximal indicative subsets within an inhomogeneous sample

3 The Decision Support System EKS

The above described expert knowledge on the determination of abiotic conditions on the basis of a terrain's vegetation, has been used to design a knowledge-based system to support ecologists in the management of nature reserves. The compositional development method DESIRE (see, e.g., [2]) has been used to design and implement this system, called EKS (Environmental Knowledge-based System).

3.1 The Compositional Development Method DESIRE

DESIRE is a compositional development method for the design and implementation of knowledge-based and multi-agent systems. A knowledge engineer is supported during all (iterative) phases of design: from initial conceptualisation to implementation, by the DESIRE method and the DESIRE software environment.

During all phases of design, the following types of knowledge are distinguished: process composition, knowledge composition and relations between process composition and knowledge composition. *Process composition* includes the identification of processes at different process abstraction levels (including task and role delegation) and process composition relations (defined by information exchange and task control). *Knowledge composition* includes knowledge structures at different knowledge abstraction levels and knowledge composition relations. *Relations* between processes and knowledge structures express which knowledge is used in which process.

Knowledge analysis focuses on the acquisition of a *shared task model*: an intermediary agreed model shared by both the expert and the knowledge engineer, in which these types of knowledge are made explicit (see [4], [5]).

Processes or tasks distinguished during conceptual design are modelled as *components*. Components can be primitive or composed: a component may encompass a number of other (either primitive or complex) components, or it may not. If not, the component is either a reasoning component with a *knowledge base* or a component with a so-called *alternative specification* (only its input and output are explicitly specified in the DESIRE modelling language, e.g., databases, OR-algorithms, neural networks, etc.). A knowledge-based system's behaviour, as a whole, is defined by the interaction between components, and between the system and its users.

DESIRE includes a software environment which consists of a graphical editor to support conceptual and detailed design, an implementation generator that translates DESIRE specifications into executable code, and an execution environment in which the translated code can be executed.

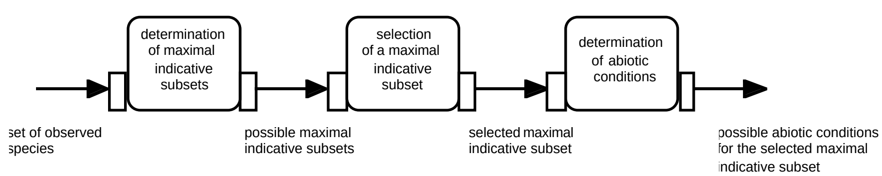

Fig. 1. The global design of EKS

3.2 Process composition

In Section 2, three tasks are distinguished: (1) grouping of plant species that "belong together", (2) selecting the set of plant species experts consider most "defining", and (3) identifying the related abiotic conditions. These three tasks are modelled by three components as shown in Figure 1. The first task, the *determination of maximal indicative subsets*, entails analysis of the plant species in the sample and the corresponding abiotic conditions to determine maximal indicative subsets of plant species. The choice of the most defining subset is performed by the component *selection of a maximal indicative subset*. The third task, *determination of abiotic conditions*, is relatively simple, and includes the presentation of the abiotic conditions of a maximal indicative subset. *Task delegation* is as follows: the first task and the third task are performed by the system; the second task is performed by the user.

The *information exchange* is depicted in Figure 1. The initial information needed by the system to determine the abiotic conditions of a terrain is a list of observed plant species. This is the input for the first component. The maximal indicative sets of plant species derived in the first task are the input for the second task. The result of the selection process (the second task), one of the maximal indicative subsets, in turn, is input for the third task (determination of abiotic conditions). The final output consists of the possible abiotic conditions for the selected maximal indicative subset.

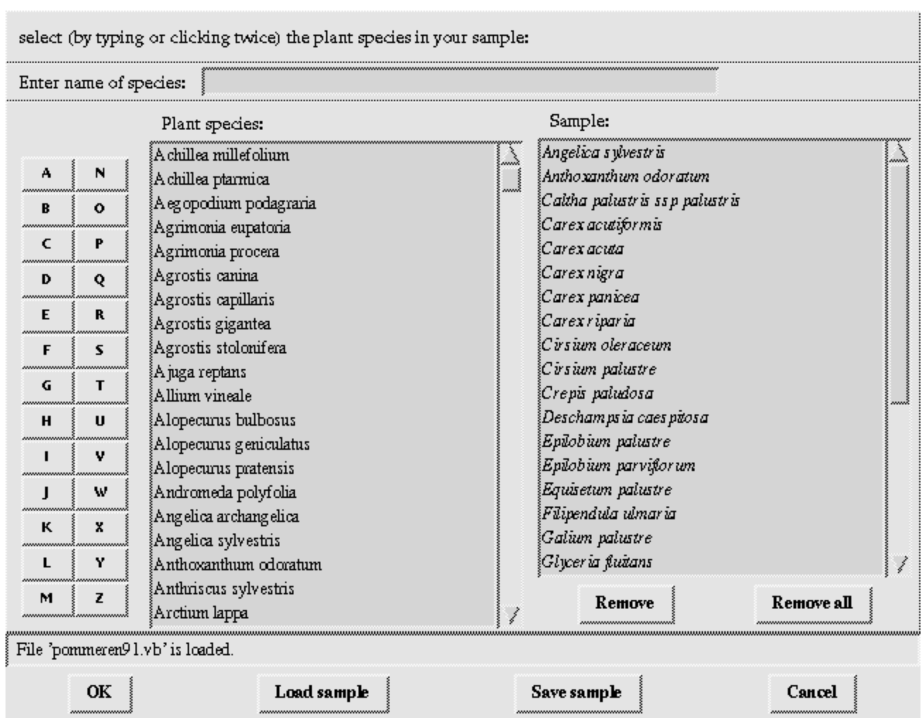

Fig. 2. Input window of EKS

Task activation is straightforward. Completion of the first task results in activation of the second. Completion of the second task results in activation of the third. Completion of the third task results in completion of the entire task.

The *knowledge* includes knowledge of plant species and the abiotic conditions in which they can abide, part of which is presented above in table format (see Tables 1 and 2). Each plant species has related values for each of the three abiotic factors. For reasons of efficiency, the first component is specified by an alternative specification.

EKS has been developed using the DESIRE method and software environment. In addition, a graphical user interface has been designed specifically for the system.

3.3 User-System Interaction

Initially a user is presented with a screen with which he/she can enter the plant species found on a terrain, as shown in Figure 2. The system analyses this information, resulting in this example in the two maximal indicative subsets of plant species. This information is presented to the user as shown in Figure 3. The overlap between the two maximal indicative subsets of plant species is presented on the screen as the list of *shared plant species*. The remaining plants are listed separately for each of the maximal indicative subsets as *abiotic indicative groups*. The user chooses which maximal subset he or she wants to analyze further. The final output of the system is a graphical presentation of the abiotic conditions for the site in question.

Fig. 3. Presentation of the maximal indicative subsets

4 Discussion

The multi-interpretability of samples of plant species has proven to be a central issue in this domain of application. Given the assumption that samples are always correct (the plant species named are indeed the plant species encountered), and that samples are only taken from sites which are homogeneous, the only reason for conflicting indicative information is that the specific domain knowledge on which conclusions are based is incorrect or incomplete. During the design of EKS this specific domain knowledge was continual subject of discussion between experts. The knowledge currently implemented in EKS is the result of consensus between experts, and is no longer a likely reason for conflicting indicative information.

The lack of homogeneity of a terrain is the cause of most conflicts, requiring additional expert knowledge to understand the nature of the heterogeneity. The reason for the lack of homogeneity can, for example, be vertical stratification, as in the example discussed in this paper. Another possibility is the development of a terrain over time: internal as well as external influences (pollution) can provoke changes and transitions in abiotic conditions, hence in vegetations. Inhomogeneous terrains are more common than initially proposed: multi-interpretable samples may not be the exception, but the rule. Different views of a sample are essential to the analysis of the plant species observed. EKS identifies these views and presents these views to the user. How these views can be interpreted requires additional knowledge, not (yet) included in the system.

The idea that information about the world can often be interpreted in different and conflicting manners was a central theme in the research reported in [8], [9], [12], and [14]. Using techniques to formalize non-monotonic reasoning, such as default logic (e.g., [2], [10], [11], [13]), often different (and often conflicting) possible outcomes of a reasoning process are obtained. In the area of research on non-monotonic reasoning, in general this is considered to be disturbing (e.g., called the *multiple extension problem*). In the literature often the non-monotonic inference operation defined by the intersection of all possible outcomes is taken (sceptical approach), or sometimes the union of all possible outcomes (credulous approach), to obtain one set of conclusions. For a particular domain, such as the ecological domain addressed in this paper, both approaches are unsatisfactory: the sceptical approach often does not lead to any possible conclusions on the abiotic conditions, whereas the credulous approach often leads to inconsistent information. For reasons like these, in [8] and [14] the multiple outcomes of a non-monotonic reasoning process are not considered to be disturbing, but are instead exploited as a useful feature that can provide an adequate formalization of the multi-interpretability often present in real-life information. In [6] and [7] a similar approach is developed, based on priority orderings between defaults.

In [8] the notion of belief set operator is introduced to formalize the multiple outcomes of a non-monotonic reasoning process, and a selection operator the make a choice between the different options. For the application domain discussed in this paper the latter approach is more suitable, because in this approach all alternative interpretations are generated. A formalisation of this expert reasoning task (and the system EKS presented here), based on belief set operators can be found in [3].

Acknowledgements

Within the EKS-project a number of persons have contributed their expertise: Frank Cornelissen and Edgar Vonk (programmers); Christine Bel and Rineke Verbrugge (parts of the analysis of the domain), Bert Hennipman (IPTS) and Wim Zeeman (State Forestry Department).

References

1. Besnard, P., An Introduction to Default Logic, Springer-Verlag, 1989.
2. Brazier, F.M.T., B. Dunin-Keplicz, N.R. Jennings, and J. Treur, Formal Specification of Multi-Agent Systems: a real-world case. In: V. Lesser (ed.), Proc. of the First International Conference on Multi-Agent Systems, ICMAS'95, MIT Press, Cambridge, MA, pp. 25-32. Extended version in: International Journal of Cooperative Information Systems, M. Huhns, M. Singh, (eds.), special issue on Formal Methods in Cooperative Information Systems: Multi-Agent Systems, vol. 6, 1997, pp. 67-94.
3. Brazier, F.M.T., J. Engelfriet, J. Treur, Analysis of multi-interpretable ecological monitoring information. In: A. Hunter, S. Parsons (eds.), Applications of Uncertainty Formalisms, Springer Verlag, 1998, to appear
4. Brazier, F.M.T., J. Treur, and N.J.E. Wijngaards, The Acquisition of a Shared Task Model. In: N. Shadbolt, K. O'Hara, G. Schreiber (eds.), Advances in Knowledge Acquisition, Proc. 9th European Knowledge Acquisition Workshop, EKAW'96, Lecture Notes in Artificial Intelligence, vol. 1076, Springer Verlag, pp. 278-289.
5. Brazier, F.M.T., J. Treur, and N.J.E. Wijngaards, Modelling Interaction with Experts: the Role of a Shared Task Model. In: W. Wahlster (ed.), Proc. European Conference on AI, ECAI'96, John Wiley and Sons, 1996, pp. 241-245.
6. Brewka, G., Adding Priorities and Specificity to Default Logic. In: C. MacNish, D. Pearce, L.M. Pereira (eds.), Logics in Artificial Intelligence, Proc. of the Third European Workshop on Logics in AI, JELIA-94, Lecture Notes in Artificial Intelligence, vol. 838, Springer-Verlag, 1994, pp. 247-260.
7. Brewka, G., Reasoning about Priorities in Default Logic. In: Proceedings of the AAAI-94, 1994.
8. Engelfriet, J., H. Herre and J. Treur, Nonmonotonic Reasoning with Multiple Belief Sets. In: D.M. Gabbay, H.J. Ohlbach (eds.), Practical Reasoning, Proc. of FAPR'96, Lecture Notes in Artificial Intelligence, vol. 1085, Springer-Verlag, 1996, pp. 331-344.
9. Engelfriet, J., V.W. Marek, J. Treur and M. Truszczynski, Infinitary Default Logic for Specification of Nonmonotonic Reasoning. In: J.J. Alferes, L.M. Pereira, E. Orlowska (eds.), Logics in Artificial Intelligence, Proc. of the Fourth European Workshop on Logics in AI, JELIA'96, Lecture Notes in Artificial Intelligence, vol. 1126, Springer-Verlag, 1996, pp. 224-236.
10. Makinson, D., General Patterns in Nonmonotonic Reasoning. In: D.M. Gabbay, C.J. Hogger, J.A. Robinson (eds.), Handbook of Logic in Artificial Intelligence and Logic Programming, Vol. 3, Oxford Science Publications, 1994, pp. 35-110.
11. Marek, V.W. and M. Truszczynski, Nonmonotonic logics; context-dependent reasoning, Springer-Verlag, 1993.
12. Marek, V.W., J. Treur and M. Truszczynski, Representation Theory for Default Logic, to appear in Annals of Mathematics and Artificial Intelligence, 1998.
13. Reiter, R., A Logic for Default Reasoning, Artificial Intelligence 13, 1980, pp. 81-132.
14. Tan, Y.-H., J. Treur, Constructive Default Logic and the Control of Defeasible Reasoning. In: B. Neumann (ed.), Proc. of the European Conference on Artificial Intelligence, ECAI'92, John Wiley and Sons, 1992, pp. 299-303.