appeared in: Umweltinformatik Aktuell, Metropolis Verlag, Marburg, Germany, 1996

Qualitative Model Composition and Transformation for Improving Environmental Decision Support

Ulrich Heller and Peter Struss[1]

Abstract

We present some intermediate results of a project that applies model-based reasoning techniques to support environmental decision making. Using a small example from the application domain of water quality management, we demonstrate our modeling formalism. Basic processes in the domain are described qualitatively as independent model elements. Their spatial distribution is also handled by means of a qualitative abstraction, so-called compartments. A model of a complex system is generated in two steps: first, the model elements of the relevant processes are aggregated under consideration of the compartmental structure. Then, this potentially overly detailed model is transformed into one that reflects the needs of a particular situation and task. The paper presents contributions to a theoretical foundation of such an automated transformation of dynamic models.

1 Outline

Our project is guided by a concept of a software framework that integrates modules for supporting different tasks in environmental decision making. This integration is mainly established by sharing a model of the environmental system at hand. Besides this, different problem solving modules are seen as different instances of a basic model-based problem solving paradigm, namely consistency-based diagnosis. We give an outline of this framework in the following section. Next, we briefly introduce the background of the application domain of hydro-ecological systems. Section 5 illustrates how models are described and composed and the need to simplify and transform the obtained compositional model. Finally,

[1] Dept. of Computer Science, Technical University of Munich, Orleansstr. 34, 81667 München, {heller, struss}@informatik.tu-muenchen.de

it is shown how this transformation can be performed through a set of generic syntactic operators acting on the model representation.

2 Integrated Model-based Environmental Decision Support

In a collaboration between researchers from Brazil, France and Germany, we started the development of an integrated model-based environmental decision support system called ΣIGMA ("System for InteGrated MAnagement of environmental resources", originally: "Sistema Integrado de Gerenciamento do Meio Ambiente" [Guerrin et al. 96]). The architecture is targeted to provide knowledge-based assistance for the tasks of model building, situation assessment, continuous monitoring, diagnosis of problem causes and planning of remedial actions ("therapy planning"). These tasks are closely related by sharing the knowledge base and by providing information and results to each other. Figure 1 shows the basic architecture with the modules, the knowledge base and their dependencies.

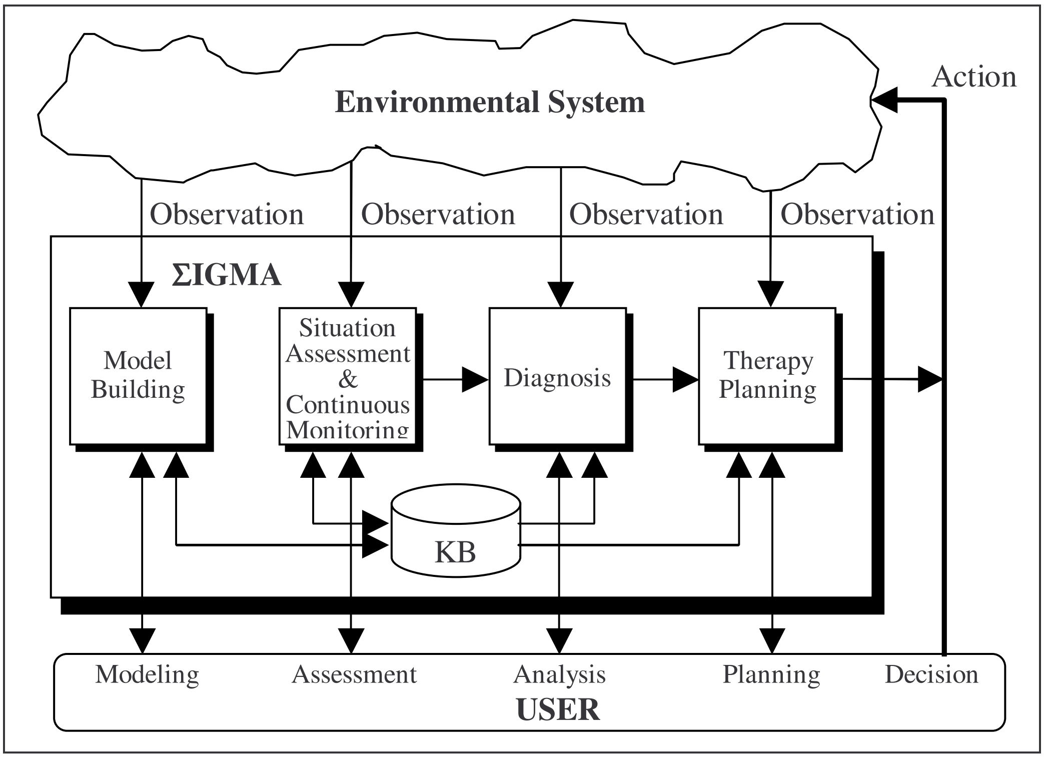

Figure 1: The basic architecture of the ΣIGMA system

So far, computer support for the various tasks has mostly tackled them separately. However, it is fairly obvious that a software framework integrating the various tools and supporting information flow between them would be highly beneficial to each individual step. Beyond this, in our approach to supporting these steps, there is a strong analogy of the four modules at a technical level: all of them rely upon the basic **revision loop** (compare Figure 2):

- In the **modeling phase**, tentative models have to be revised until they adequately represent the (knowledge about the) relevant fundamental processes in the specific ecosystem to be investigated.
- For **situation assessment**, the system model has to be adapted in order to capture the actual conditions. This is achieved by adaptation of parameters and/or disturbances of processes.
- For **diagnostic purposes**, the model has to identify the causes of undesirable states and effects.
- Finally, **therapy planning** can be regarded as a problem solving process that seeks to revise the model by incorporating remedial actions until it satisfies certain objectives.

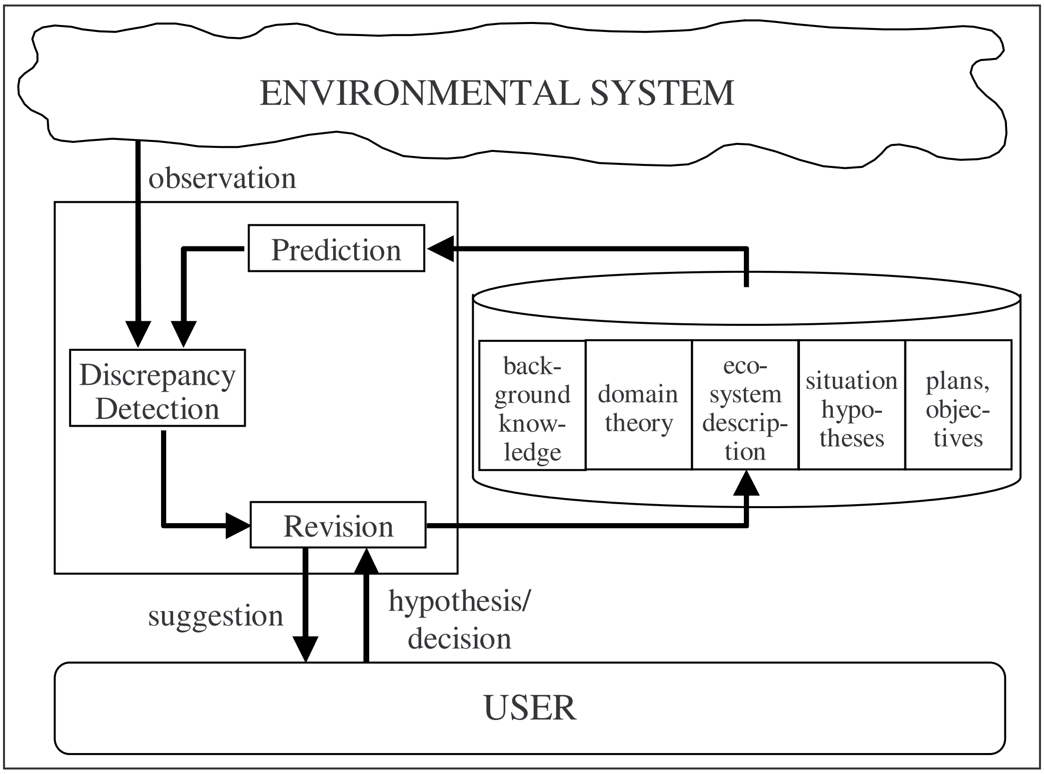

Figure 2: The situation assessment module

The internal representation of the environmental system, i. e. "the model" in the broadest sense is divided into five conceptual segments (compare Figure 2):

- the **domain theory**, i. e. a library of processes relevant for the class of ecosystems under consideration (e. g. hydro-ecological systems)
- the **description of the specific ecosystem** at hand, including the topology, the plant and fish populations present, and external influences
- a collection of **hypotheses about the current situation** of the ecosystem, i. e. parameter values and active processes
- **objectives and plans** , i. e. a representation of the desired state of the ecosystem and the intended actions to be taken in order to achieve it
- the domain independent **background knowledge**, e. g. specifying methods for the composition of the model elements in the other segments

The problem solving modules differ only with respect to the parts of the internal representation they act upon and the strategies employed. For illustration, a specialization of the basic revision loop for the situation assessment module is shown in Figure 2.

There is a well-established theory for this kind of model revision originating from the field of consistency-based diagnosis ([Dressler/Struss 95]). The principle is to predict the behavior of a device from a model and to find faulted components from the discrepancies to the actual observed behavior produced. A number of general and efficient implementations of such systems exist. We intend to employ the technique to a slightly more general application context than to finding faulted components of a technical device, but the formal foundations are largely identical:

The model serves as a basis for prediction of variables, parameters or future behavior, which will be compared to the actual behavior (as reflected in the measurements and observations) or the user expectations (in the model building phase). From the discrepancies the system will derive suggestions for revision of the model, in case of the situation assessment module the set of situation hypotheses in terms of active processes and process parameters. The user will control this revision process that eventually yields a model that captures the relevant features of the actual situation.

3 Application Domain

The ecosystem under consideration is the Rio Guaíba in Brazil. A river by convention, the hydrological system is better described as a small lake crossed by a river. A number of bays that broaden up to 40Km further contributes to the complicated hydrodynamics. Most of the phenomena taking place in the Rio Guaíba are of limited spatial extent or at least largely influenced by spatial distribution processes, the most important of which is the transport of matter by directed flow of water, called advection.

The river is heavily polluted by industrial and agricultural waste, as well as city sewage, while at the same time serving as the main source of drinking water for the 1.2 million inhabitants of the city of Porto Alegre. The municipal authorities are studying the impact on the usability of the water body for recreational and industrial purposes and as fishing grounds, but also become increasingly concerned with the development of the ecosystem as a whole.

The phenomenon we concentrated on in the initial phase of our studies is the occurrence of local algal blooms, i. e. a steep and significant increase in the biomass of phytoplankton ("algae"), typically limited to an area of some square kilometers and a duration of a few days. The main cause under examination is the excessive amount of nutrients for phytoplankton that is determined by advective transport, by chemical reactions and biological transformation mostly achieved by microorganisms.

Among the undesirable effects of algal blooms is the clogging of pumps, that are used to collect raw water for the drinking water supply. Also unpleasant taste and odor of the water even after standard treatment was observed. The long term effects of environmental stress factors like algal blooms impose on the plant and fish populations in the river include a dangerous decrease in biodiversity. Both for long term prevention of and for immediate reaction to blooms (e. g. by shutting off pumps in the vicinity) it would be desirable to be able to predict (the possibility of) such events. This is one of the main goals in the modeling efforts for the envisioned decision support system.

4 Modeling - Requirements and Problems

Obviously, the model plays a crucial role in our approach. From both the presented architecture of the decision support system and our experience in modeling a real ecosystem, the Rio Guaíba in Southern Brazil, four important requirements for the models arose:

- **compositional models**, i. e. the behavior model of a complex system is aggregated of model elements taken from a library. This modularity makes the modeling task feasible and provides computational support for model building and revision and creates re-usable model elements. In our work, we use process descriptions, e. g. of a chemical reaction, as model elements.
- **dynamic models**, i. e. representing the temporal evolution of systems while, particularly, taking into account different interacting process speeds. For example a chemical reaction will act much faster than a medium scale transport process in a river or the population dynamics of some species.
- **multiple models**, i. e. representing models on different levels of abstraction for different tasks to enable switching to the appropriate one in a systematic, possibly automatic, way, for instance in analogy to ecologists describing chemical reactions in many cases in form of their equilibrium equation.
- **qualitative models**, i. e. making the necessary distinctions only. A human expert, for example, is able to reason effectively about the chemical subsystem of an ecosystem without precise knowledge about the reaction speeds or influence coefficients.

The choice of qualitative models is motivated by several intrinsic properties of the domain and the task, e. g.

- lack of precise knowledge about some important phenomena,
- difficulties with or even impossibility of adequate measurements
- the requirement for tractable computation of predictions,
- the necessity for communication with unexperienced computer users (e. g. expert ecologists), and
- the preference for interpreted and qualitative output in many cases, e. g. in terms of the possibility of dangerous events or proposals for adequate counter-measures, rather than numerical values.

We propose an approach that achieves the creation of a model into two steps:

- **composition:** aggregating a system model based on the generic, detailed model elements in the library
- **transformation** of the composed model according to the needs of the specific task, e. g. by eliminating irrelevant distinctions and intermediate

variables, by (temporal) abstraction of subsystems, and by applying appropriate modeling assumptions.

Research on Qualitative Modeling (see [Faltings/Struss 92], [Weld/de Kleer 90]) has provided means for supporting or automating of both steps. The work we present in the remainder of this paper is an application of composition techniques and a contribution to (semi-)automated model transformation.

5 Compositional Model Representation

In the specific ecosytem we have modeled, the compositional approach, i. e. to specify context-free model elements that will be (automatically) instantiated and assembled according to a system description, is applied in two "dimensions": we use a spatial partitioning of the ecosystem into compartments and a conceptual separation of the ecosystem dynamics into individual (instances of) processes. Compartmental modeling is treated for example in [Ironi/Stefanelli 94], the theoretical basis of qualitative process-oriented modeling can be found in [Forbus 84].

5.1 Compartmental Spatial Representation

We use a topological partitioning of the water body into **compartments**, i. e. regions that are assumed to have similar parameter values and flow characteristics. This corresponds to the notion of "hydrodynamic regimes" employed by hydrologists. We use only topological information (i. e. neighborhood relations) and some individual characteristics (like volume) for the compartments. When treating for example the concentrations of chemical constituents as homogeneous inside each compartment, we obtain a qualitative abstraction of the spatial distribution of the respective substance.

One part of the system description (stored in the ecosystem-specific segment of the knowledge base) is the topological layout of the compartments. Processes (see next section) will be located either inside of a compartment (e. g. chemical reactions, having primarily local effects) or between adjacent compartments (e. g. a flow process achieving exchange of matter).

A simple example of a compartmental partitioning of a bay of the river is shown in Figure 3. The arrows indicate the water flow into a compartment "X" from upstream (compartment "In") and further downstream (into "Out").

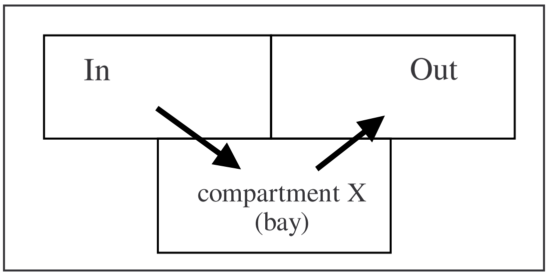

Figure 3: A simple compartmental structure

5.2 Process Models

The domain knowledge segment of the ΣIGMA knowledge base consists of a collection of generic process descriptions.

These process specifications as well as the composed behavior models will be visualized using a graphical representation of qualitative dynamic models we developed, that includes an explicit causal notation and serves as a basis for model transformations in order to simplify and abstract models on a syntactical basis. Some relevant primitive elements of the **qualitative influence diagrams** are shown in Figures 4 through 6.

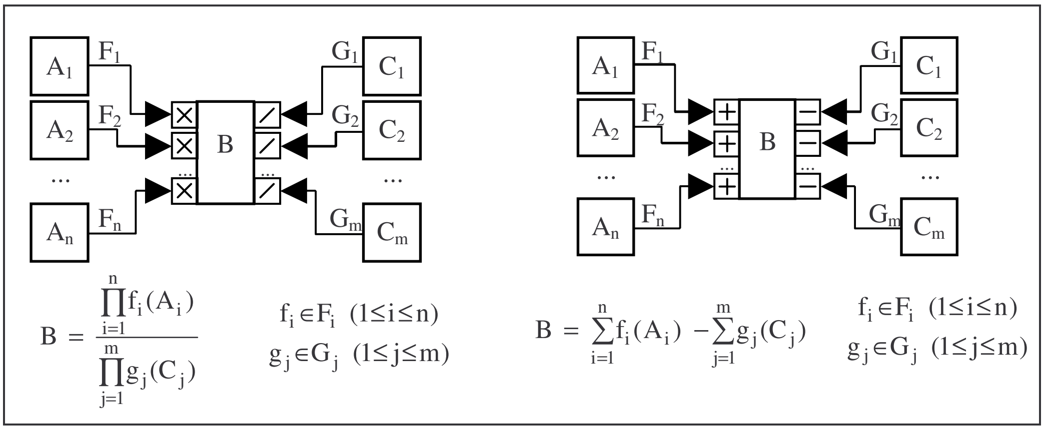

Figure 4: The basic elements of the influence diagrams,
with completely multiplicatively and additively decomposed influences

The large boxes stand for continuous real-valued variables (or any abstraction thereof). The arrows denote influences that combine additively or multiplicatively as given in the combination information (smaller boxes at the side of the variables).

The influence functions are restricted by a function class denoted by the arrow labels (F_i, G_j). The most general class we use is (strictly) monotonic functions (label "Mon"). Additionally, a Lipschitz condition (denoted "Lip") can hold, i. e. | f

$(x) - f(x') | \leq M \cdot | x - x' |$ for some $M \in \mathbb{R}^{+}$. The label "Lin" corresponds to linearity and the label "Id" refers to strict identity (this label will occasionally be left out).

Integrative influences (i. e. influences on the derivative of a variable) are depicted using circles or rectangles with rounded edges. The same function class restrictions are applicable. For a detailed description of the semantics, please refer to [Heller/Struss 96] or [Heller 95].

Figure 5 shows a generic process description for the effect of **advection** (i. e. the exchange of matter by directed water flow) between two compartments (*src* and *dest*) on the concentration of a single chemical constituent (*const*). This simplified version assumes constant volumes of the involved compartments. The corresponding (classes of) differential equations are given on the right hand side of the figure. These equations are written as if no other influences on the concentrations exist. When combined with other model elements, the influences are assumed to combine additively (a short discussion of this assumption is given in [Heller/Struss 96]).

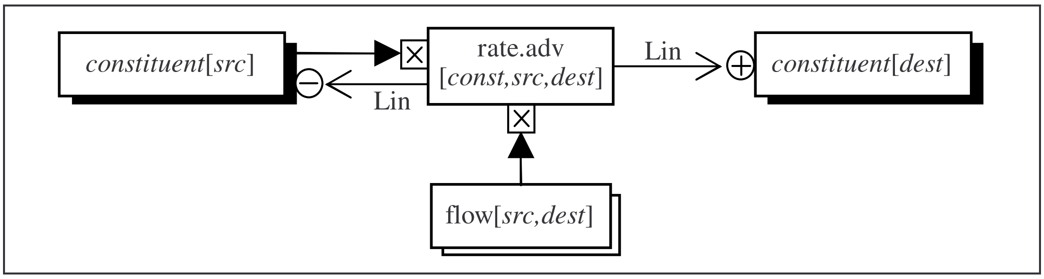

Figure 5: The advection process (simple version)

A model element for **dissociation of ammonia** (according to the reaction $NH_4^+ + OH^- \leftrightarrow NH_3 + H_2O$) is shown in Figure 6. This process is instantiated inside each compartment *location* that has a non-zero concentration of ionized ammonia (NH_4). Here, the feedback of the reaction on the concentration of NH_4 is neglected, since for the standard range of the pH we can treat NH_4 as total ammonia. Again, equations (for the isolated model element) are given.

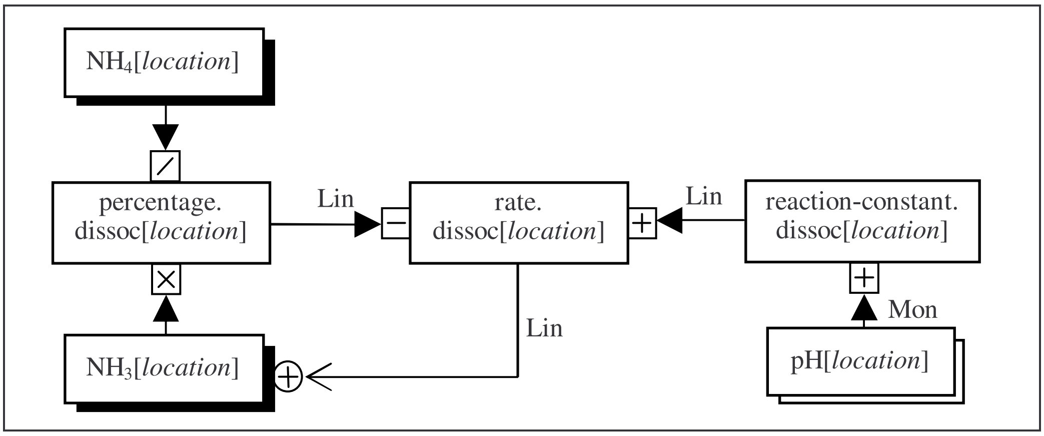

Figure 6: Dissociation process without feedback

5.3 Model Composition

The generic process description will be instantiated for each set of parameters and compartments, respectively, that meets some given instantiation conditions. An example instantiation constituting a simple behavior model of the interaction of advection and ammonia dissociation for the compartmental structure from Figure 3 yields the diagram in Figure 7.

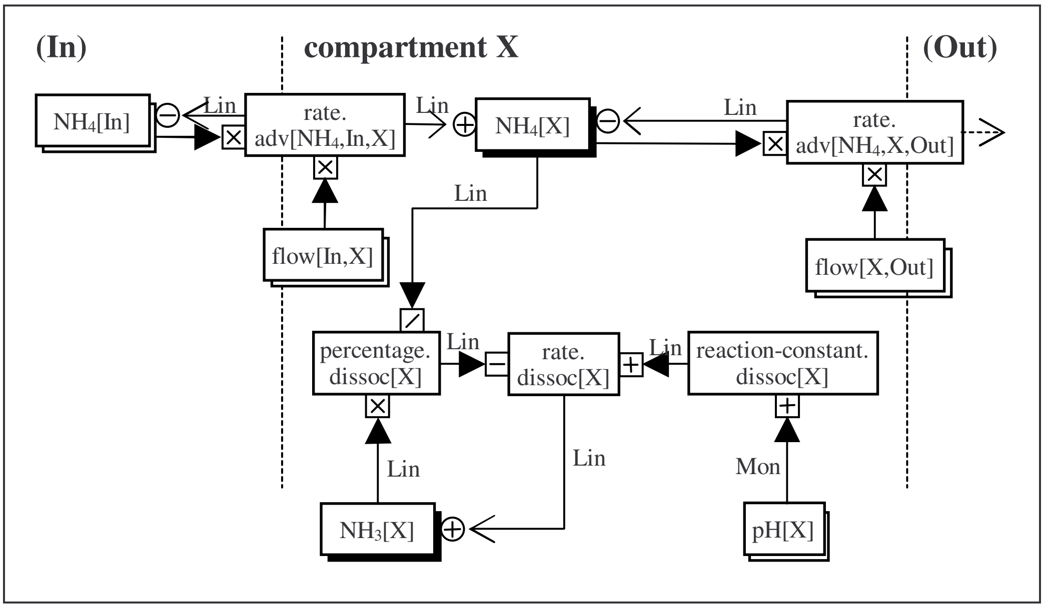

Figure 7: The interaction of dissociation and advection

However, in our test runs even with simple (sub-)models like the one presented here, we obtained unsatisfactory results in terms of the predictive power of our models. We discovered this to be a typical instance of a more general problem, based on the outcome of the composition phase running the risk of being

- too detailed, because of the local specification of model elements, e. g. there might be a number of unobservable variables or ones that can be determined to be constant only from the context of the composed model.
- too fine-grained, because of the interactions with widely different time-scales. The representation of chemical dynamics imposes large computational costs while usually only generating rapid but minimal fluctuation of concentrations that are basically held in chemical equilibrium, so the detailed prediction does not contribute significantly to long term prediction.

Qualitative simulators (e. g. QSIM, see [Kuipers 86]) sometimes even generate spurious solutions by misjudging the relative strengths of the influences. A typical instance of this is the (erroneous) assumption, that the effect of the advective transport is comparable to the chemical reaction rate, so that the two forms of ammonia are predicted to be significantly out of equilibrium.

The scheme of the basic revision loop in the ΣIGMA modules shows the importance of the predictive power of the models and the algorithms employed. Only if focused predictions can be derived, they can be profitably compared to actual observations in order to detect discrepancies.

6 Model Transformation

To address the problem of different process speeds, Kuipers proposed a technique called "time-scale abstraction" ([Kuipers 87]). The principle is the substitution of a fast dynamic subsystem, that tends to establish an equilibrium, by a functional relationship corresponding to the solution of the equilibrium equation. Thus, the deviations of the subsystem from equilibrium due to shifting context conditions are neglected, it is assumed that the equilibrium is achieved instantaneously. In the theory of model relations ([Struss 91]) this is characterized as approximation rather than an abstraction in the strict sense.

The benefit of this substitution is a reduction in the computational complexity of prediction, and at the same time the erroneously generated behaviors mentioned above can be ruled out.

We will demonstrate an approach to automate time-scale abstraction. The first step will be the identification of a feedback subsystem that could be subject to this kind of approximation. For this purpose, the model will be simplified (strictly speaking, a "view" of a given model will be created, see [Struss 91]) and then the identified structure can be substituted. Both steps will be achieved by making use of a set of syntactical operators acting on the introduced qualitative influence diagrams.

One of the basic simplification operators, that will eliminate intermediate variabes while preserving additive decomposition of influences is shown in in Figure 8. The conditions for the application and the calculation rules for the resulting function class restrictions are given. The maximum of the function classes is taken with respect to set inclusion (note that Id ⊂ Lin ⊂ Lip ⊂ Mon).

The proof that the local application of this operator actually creates a view, i. e. the resulting model strictly implies the original one, can be found in [Heller 95].

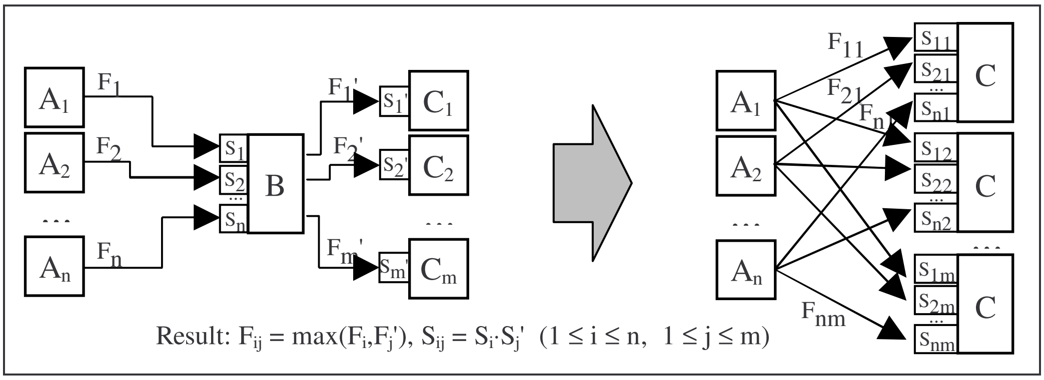

Figure 8: A basic abstraction operator for eliminating intermediate variables with completely decomposed influences

By applying this operator to the example model (Figure 7) the variable reaction-constant.dissoc[A] and then rate.dissoc[A] can be eliminated, which yields the qualitative influence diagram in Figure 9.

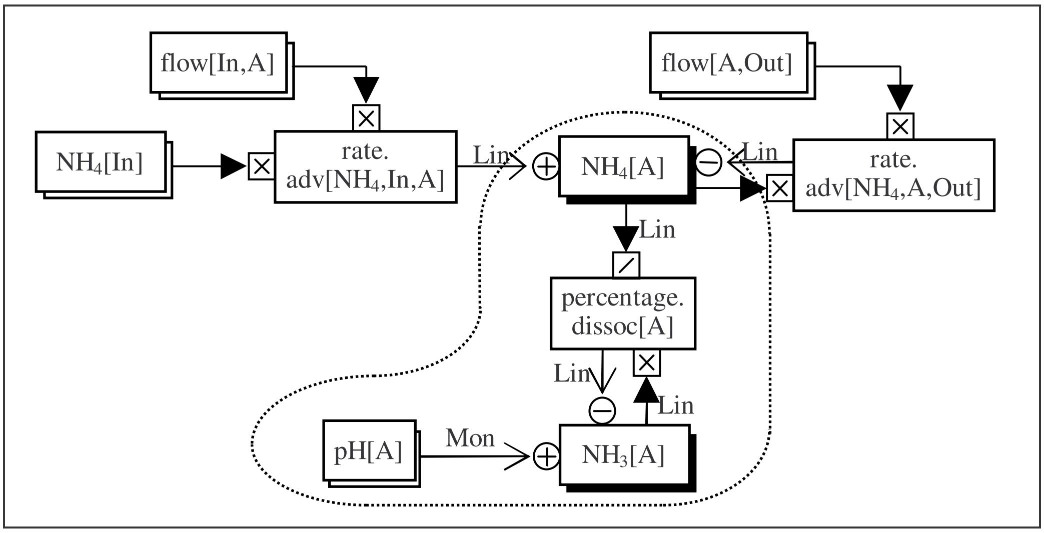

Figure 9: The example model after the elimination of the intermediate variables

The structure in the lower part of Figure 9 is identified as an instance of a multiplicatively mediated linear self-stabilization, which is a feedback system tending to establish and stabilize equilibrium. In Figure 10 we present a time-scale approximation operator, that can substitutes the subsystem by the solution of its equilibrium equation, i. e. a functional dependency instead of the integrative influences.

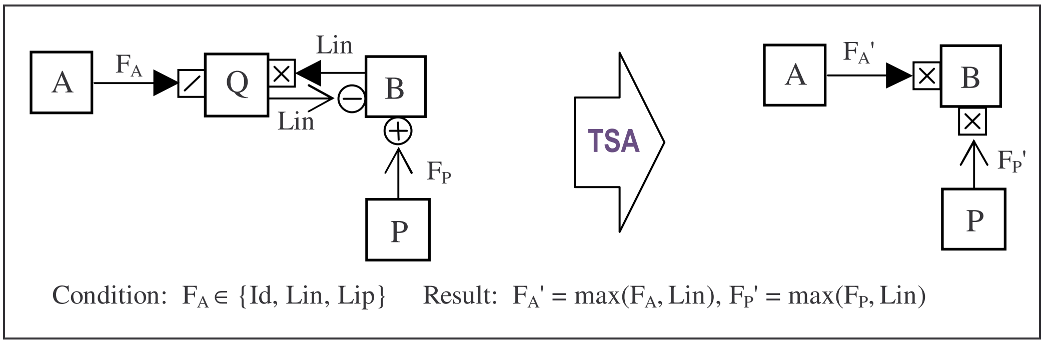

Figure 10: Two time-scale abstraction operators for influence diagrams

The approximation error is decreasing with the "separation of the time-scales" of the context and the feedback subsystem (for a detailed discussion of the underlying differential equations please refer to [Heller 95]). In our example we have very different process speeds of the chemical reaction and the advective transport, which can be determined locally from the linearity coefficients. Whether the operator is profitably applied can by decided automatically, e. g. by some threshold. If the approximation is employed, we obtain the diagram in Figure 11.

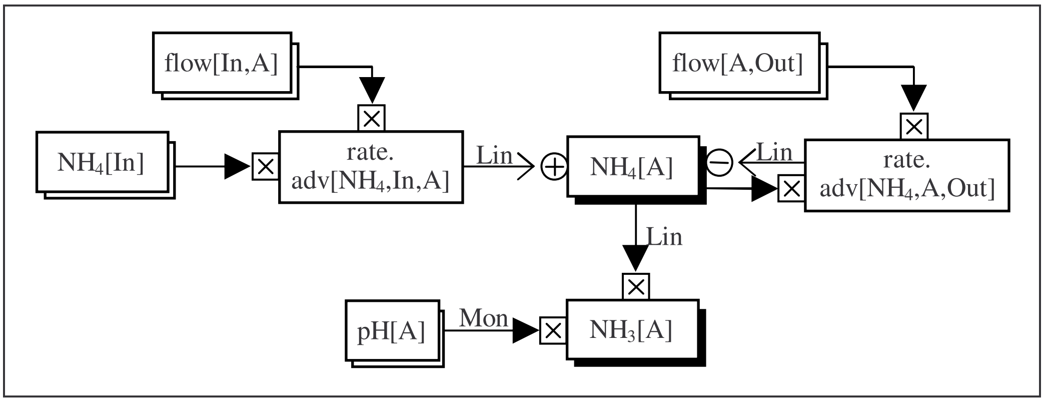

Figure 11: The resulting model for the interaction of dissociation and advection

In our test runs we obtained much more focused predictions with models transformed in this way, with a number of spurious behavior predictions ruled out.

So the operators not only increase the efficiency of the reasoning task, but also improve the predictive power of the models assembled from the library of generic process description by making use of the information about separated time-scales.

7 Current State and Future Work

The realization of ΣIGMA will start off with a monitoring and a diagnosis module, since they can be based on available implementations of consistency-based diagnosis techniques. The modular structure of the decision support system enables the incremental realization of the conceptual design, while the user can profitably work with single modules. Thus, the task of environmental resource management can be supported to a continuously increasing degree by computational tools. Simultaneously, the models of the ecosystem under consideration will be evaluated and validated so there is a driving real-world application for the development of the system.

It is intended to implement a tool that supports the creation of models based on the qualitative influence diagrams formalism and of operators on these models. The diagrammatic notation will prove useful for the communication with scientist without a background in computer science. What we have achieved at the present time, is an initial theory of transforming models of dynamic systems. The set of operators developed and proved so far is certainly not complete. We also intend to introduce more function classes, for instance, in order to determine dominant influences in a combination of counteracting ones (e. g. overlinear versus linear growth).

Acknowledgements

We like to thank Francois Guerrin and Waldir Roque, who participated in the design of the ΣIGMA system. This work was partially supported by the Brazilian Research Council (CNPq) and the German Ministery of Education and Research (BMBF).

References

[Crawford/Farquhar/Kuipers 90] James Crawford, Adam Farquhar, Benjamin Kuipers: QPC: A Compiler from Physical Models into Qualitative Differential Equations. AAAI 1990.

[Dressler/Struss 95] Oskar Dressler, Peter Struss: The Consistency-based Approach to Automated Diagnosis of Devices. In: Gerhard Brewka (ed.): Principles of Knowledge Representation. CSLI Publication, Stanford, 1995.

[Forbus 84] Ken Forbus: Qualitative Process Theory. AI Journal 24 (1-3), 1984.

[Faltings/Struss 92] Boi Faltings, Peter Struss (eds.): Recent Advances in Qualitative Physics, MIT Press, 1992.

[Guerrin et al. 96] François Guerrin, Ulrich Heller, Peter Struss, Waldir Roque: ΣIGMA: um sistema integrado de gerenciamento do meio ambiente. In: ECOS Journal, No 6, DMAE Publications, Porto Alegre, 1996. (in Portuguese)

[Heller et al. 95] Ulrich Heller, Peter Struss, Francois Guerrin, Waldir Roque: A Qualitative Modeling Approach to Algal Bloom Prediction. In: Workshop Notes of the International Workshop on "Artificial Intelligence and the Environment" (at IJCAI-95), Montreal, Canada, 1995.

[Heller 95] Ulrich Heller: Temporale Verhaltensabstraktion am Beispiel von Regelkreisen und hydro-ökologischen Systemen. Master thesis, Department of Computer Science, Technical University of Munich, 1995. (in German)

[Heller/Struss 96] Ulrich Heller, Peter Struss: Transformation of Qualitative Dynamic Models - Application in Hydro-Ecology. In: Tenth International Workshop on Qualitative Reasoning (QR-96), AAAI Press, 1996.

[Ironi/Stefanelli 94] Liliana Ironi, M. Stefanelli: QCMF: A Tool for Generating Qualitative Models from Compartmental Structures. QR-94, Nara, Japan, 1994.

[Kuipers 86] Benjamin Kuipers: Qualitative Simulation. AI Journal 29(3), 1986.

[Kuipers 87] Benjamin Kuipers: Abstraction by Time-Scale in Qualitative Simulation. AAAI 1987.

[Struss 91] Peter Struss: A Theory of Model Simplification and Abstraction for Diagnosis. QR-91, 1991.

[Weld/de Kleer 90] Dan Weld, Johan de Kleer (eds.): Readings in Qualitative Reasoning about Physical Systems, Morgan Kaufmann, San Mateo, 1990.