The Knowledge Engineering Review, Vol. 11:1, 1996, 1–25

An overview of approaches to qualitative model construction

CIS SCHUT and BERT BREDEWEG

Department of Social Science Informatics (S.W.I.), University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam, The Netherlands., Email: schut@swi.psy.uva.nl

Abstract

In qualitative reasoning research, much effort has been spent on developing representation and reasoning formalisms. Only recently, the process of constructing models in terms of these formalisms has been recognised as an important research topic of its own. Approaches addressing this topic are examined in this review. For this purpose a general model of the task of constructing qualitative models is developed that serves as a frame of reference in considering these approaches. Two categories of approaches are identified: model composition and model induction approaches. The former compose a model from predefined partial models and the latter infer a model from behavioural data. Similarities and differences between the approaches are discussed using the general task model as a reference. It appears that the majority of approaches focus on automating model construction entirely. Assessing and debugging a model in cooperation with a modeller is identified as an important topic for future research.

1 Introduction

In the last decade qualitative reasoning research has become an important area in artificial intelligence. This is evident from the number of articles on qualitative reasoning that have appeared in journals (with several special issues) and conferences, by a number of books devoted to the topic, and by a successful series of annual qualitative reasoning workshops. This research effort has resulted in several formalisms that reason with qualitative knowledge. The availability of these formalisms, however, has not led (yet) to widespread application of qualitative reasoning systems. To some extent this lack of applications can be attributed to the difficulty of formulating models in terms of the proposed formalisms, although this is certainly not the only reason. This modelling problem is currently being recognised as an important research problem of its own. As a result there is an increase in the effort being spent on developing techniques for automating or semi-automating the model construction process. This article gives an overview of the different approaches that address the modelling problem.

To assess the contribution of each of these approaches, a frame of reference is required that allows a comparison of individual approaches. Section 2 provides such a framework. It discusses the characteristics of models in general and of qualitative models in particular, and gives a general model for the task of constructing a qualitative model. This general task model provides the frame of reference for considering each of the proposed approaches. These approaches are divided in two categories according to the particular circumstances in which they address model construction. Section 3 describes model composition approaches which tackle the modelling problem by composing a model from predefined parts. Section 4 describes model induction approaches which apply machine learning techniques to generate models from examples. In discussing the individual approaches it is indicated how they relate to the general task model. Finally, section 5 summarises the article, states the main results of considering the approaches from the perspective of the general task model, and identifies a number of issues for future research.

2 The nature of modelling

This section describes a frame of reference that will be employed to compare and evaluate a wide variety of techniques that address the task of constructing a qualitative model. In developing such a frame of reference, the characteristics of models in general are discussed. Because *qualitative* models have specific characteristics, a separate section is devoted to these characteristics. These discussions clarify the different ways in which the term *modelling* is employed in literature. Sometimes it refers to the process of defining a representation and reasoning formalism, other times it refers to the process of depicting a particular system in terms of such a formalism. The main focus of the article is on the latter: it excludes the formulation of a modelling framework. This is reflected in the term *model construction* that is used throughout the article. We adopt an empirical design perspective on model construction. Using such a perspective, we develop a general model of the task of qualitative model construction that should encompass all specific approaches. This general model will serve as the frame of reference for considering these individual approaches.

2.1 What is a model?

A general observation of models[1] is that they have a relation with something else, i.e., they refer to something. For example, a map can be considered a model that refers to some area in the world. The *referent* of the model is in this case the area in the world. The relation between a model and its referent is a representation relation. That is, a certain correspondence must exist between the model and the referent, at least for some *distinguished aspects*. For example, the distances between cities on a map correspond with the distances of the real cities. The reason why the relation between a model and its referent has a representational character, is that a model is meant to be used instead of the referent itself. Typically, a model is used to infer information that cannot be revealed from the referent directly, or only at the expense of much greater cost, time, danger, etc.

Which aspects of the referent the model is expected to infer (and also which aspects of the referent will be used as inputs for this *inference procedure*), is determined by the *purpose* (or intended use) of the model. A map, for example, may be used for inferring distances between cities, for inferring one's current position in a city, for determining the percentage of woodland in some area, and so on. It may not be intended for computing the number of trees in the referred area, monitoring traffic flows in that area, and so on. Therefore, most maps neither represent individual trees, nor the traffic on the roads. Since different users of a model may impose different requirements on the model, such differences may be reflected in the aspects considered relevant. This illustrates how the purpose of the model is crucial in *conceptualising* the referent.

More specifically, the purpose of a model lays down the particular *perspective* that is adopted in considering the referent. This perspective will have its repercussions on the *modelling framework* that is used for formulating a model. A perspective can be viewed as a pair of selective glasses that restrict the perception of the referent to the features that are relevant to the modellers purpose: some aspects of the referent are neglected, whereas others are considered explicitly. For example, in the map as a model of an area of the world, that area is perceived from a topographical perspective. Because rivers are only relevant in a topographical sense, the biological and chemical processes in it, for instance, are neglected. The implications of the purpose of a model are twofold: (i) it determines which aspects of the referent are represented in the model and which aspects are neglected, and (ii) it divides these relevant aspects in two groups, the ones that should be inferred from the model and the ones that are input for the inference procedure (for inferring the others!).

The modelling framework that results from a specific conceptualisation of the referent can be regarded as consisting of a *conceptual framework* (an ontology) and an *inference procedure*. Both are abstract descriptions: the conceptual framework describes the aspects of the referent that are

[1]For general discussions of modelling and numerous references see, for example, Rothenberg (1989) and Zeigler (1976).

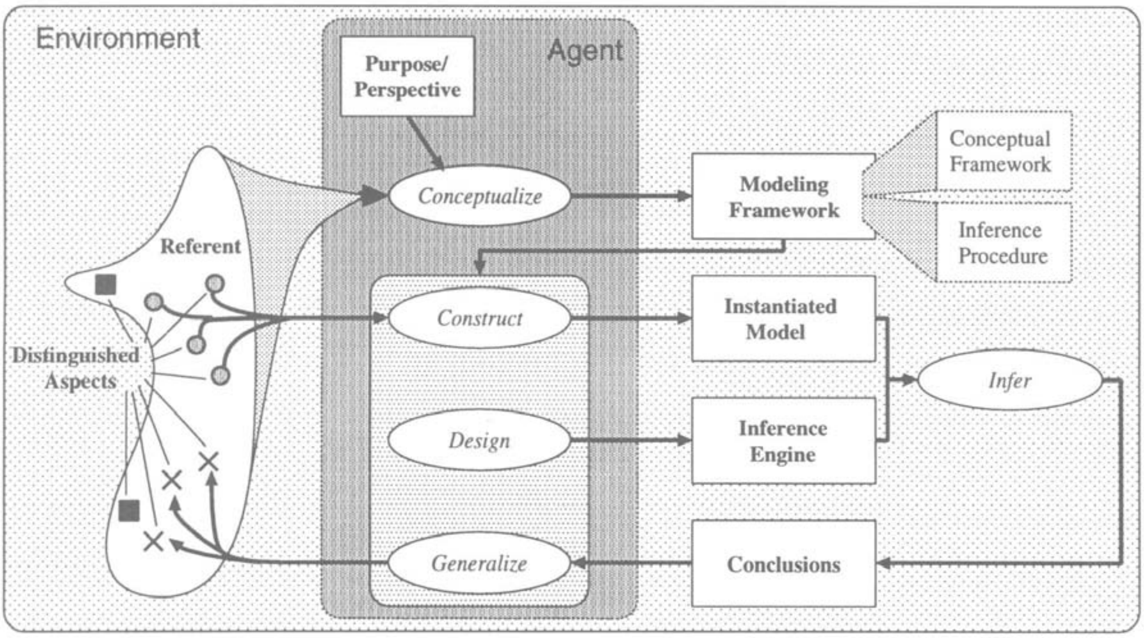

Figure 1 The activities in modelling

distinguished, the inference procedure describes how the concepts in the conceptual framework are employed to derive new information. They are mutually dependent in the sense that certain concepts allow certain derivations, and certain derivations require certain concepts. Because a model and its referent exist side by side, a model must be constructed from its own "material". The conceptual framework is materialised in a set of *representation primitives*. In the map example, the conceptual framework consists of cities, rivers, roads, etc., and the representation primitives correspond to the legend of the map: the different symbols for cities, rivers, roads, and so on. In constructing a model for a particular referent, the representation primitives are instantiated according to the characteristics of the referent. For instance, a map of Portugal will have city symbols for Lisbon, Coimbra, Fátima, and so on. The inference procedure is materialised in an *inference engine* that actually uses the instantiated model to derive new information about the referent. Obviously, the process of designing the inference engine is influenced by the conceptualisation.

The process of model construction entails the instantiation of a set of representation primitives. This instantiation step concerns the translation of measurements[2] of the system into the representation primitives. This step is by no means a trivial step: a model is only as good as the information on which it is built. General discussions of the difficulties and solutions with respect to *measurement interpretation* can be found in Forbus (1983, 1986), De Coste (1991), and Hau and Coiera (1995). If the model is instantiated properly and the inference engine designed correctly, then the results of using the model, the *conclusions*, can be *generalised* to hold for the referent as well. Thus, the conclusions *represent* certain aspects of the referent. For example, if the map is correct then its scale and the distance between two cities on the map can be used to compute their actual distance. The inference engine must be general enough to cover all questions that the modeller expects the model to answer. That is, it should not only be able to compute the distance between Lisbon and Coimbra, but between any two geographical items on the map. Therefore, the inference engine is defined with respect to the representation primitives, not with respect to how they are instantiated.

Figure 1 shows the activities that are required for constructing a model. The referent has to be conceptualised in a certain way, depending on the purpose of the model, i.e., the goals of the

[2]Observations are also considered measurements.

modeller. Conceptualising should result in a particular modelling framework, consisting of a conceptual framework and an inference procedure. The modelling framework is materialised in a set of representation primitives (not shown in the figure) and an inference engine (design inference engine). The second major activity concerns the actual model construction: the properties of the referent have to be depicted in terms of the representation primitives. This process, which is directly influenced by the properties of the modelling framework, results in an instantiated model. The instantiated model and the inference engine must be able to infer conclusions, i.e., additional information that was not present in the model. Finally, the conclusions have to be generalised so that the conclusions can be interpreted as aspects of the referent.

Summarised, a model has the following characteristics:

- A model represents a referent that exists independently.
- A model has a purpose that determines the perspective with which the referent is conceptualised.
- The conceptualisation determines which aspects of the referent are considered and which are neglected. It also determines which aspects form the model and which must be inferred from the model. Finally, it determines how such inferences are made.
- The conceptualisation is materialised in a set of representation primitives and an inference engine.
- A model is formed by a set of instantiated representation primitives.
- The inference engine uses the model to infer new information.
- Through a generalisation process the new information can be related to the referent.

This results in the following working definition of a model:

> *A model is a representation of a phenomenon, conceptualised according to a particular goal, for which an inference procedure exists that allows the derivation of new information about that phenomenon.*

2.2 *Qualitative models*

The general characterisation of a model that was provided above can be specialised for the particular type of models of our concern: qualitative models. This section discusses the features that are specific for qualitative models.

First, the referents of qualitative models have traditionally been the range of physical systems: the research area is sometimes considered equivalent with qualitative physics. Although physical systems have been the major focus of qualitative research, other domains have been investigated as well, for example, economics, chemistry, biology (see Kuipers (1994) for an overview and references).

The predominant purposes for which qualitative models are used, can be categorised into predicting system behaviour and explaining system behaviour. The "qualitativeness" of the models stems from the situations in which they are intended to be employed. These situations occur in several situations.[3] They concern, for example, situations in which only incomplete knowledge, i.e., qualitative knowledge, is available, or in which it is simply cheaper to employ only qualitative knowledge (i.e., when appropriate answers are satisfactory). Especially relevant in an explanation context, qualitative knowledge is often more in line with the knowledge of human subjects. The predictive purpose of qualitative models is often derived from a more general goal. For example, in

[3]We will not discuss in detail *why* qualitative models are useful at all. See Forbus (1988) for a discussion of situations where system behaviour cannot be captured quantitatively or where a qualitative model is more appropriate than a quantitative model.

designing a system the consequences of particular design decisions may be predicted by simulating a qualitative model of the system. As another example, in diagnosing a malfunctioning system the consequences of hypothesising certain errors might be predicted using a qualitative model. The purposes of prediction and explanation can also be present at the same time. For instance, if a model of a system serves as a part of a tutoring system, then it can be used by a pupil to perform experiments, but at the same time it can be used by the tutoring system to explain the system behaviour. As another example, in a diagnosis system the model may be used both for predicting consequences of hypothesised errors, and for explaining why a hypothesis explains the system's malfunctioning.

In the previous section we saw that the purpose of a model is crucial for determining which distinctions to make in considering the referent. The specific purposes of qualitative models, prediction and/or explanation of system behaviour over time without resorting to numeric values, dictate that certain aspects of the referent should be captured by the model. As a consequence, corresponding representation primitives have to be available as well.

First, the time-varying aspects of the referent have to be covered. Usually, these are found in the conceptualisation as a sequence of qualitative states. Each qualitative state is characterised by a number of *variables* that have a particular qualitative *value* and *derivative*, i.e., the direction in which the value is changing. Since qualitative models do not refer to numeric values, the precise value and derivative are abstracted into symbolic values and derivatives. The latter are usually abstracted into the range of *increasing*, *steady* and *decreasing*. The former are organised into ordered sets of symbols that abstract a ratio scale into an ordinal scale. The elements of the ordered set of symbols correspond to relevant points and intervals (qualitatively identical values). Whenever a variable takes on a different qualitative value, a transition to another qualitative state occurs.[4]

The *dependencies* between the variables that govern the behaviour over time have to be captured too. For different types of dependencies different types of representational analogs are required. In cases where the model has an explanation purpose, the causality in the workings of the system has to be captured. This can be achieved by including *a priori* causal dependencies in the model, or by applying a procedure that infers the causal structure.

Additional purposes may demand a different conceptualisation and corresponding representation primitives. For instance, if the model is meant to provide answers to questions about configurational properties of the system, then the system components and their layout should be represented explicitly in the model. In other cases, spatial properties might have to be included, and so on.

Finally, an inference engine, appropriate for the representation primitives is required. It must be capable of deriving the information that suits the particular purpose of the model. In the case of qualitative models, the inference engine must predict the qualitative changes occurring in the system and/or explain why these changes in the system occur.

It must be noted here that we are referring to *types* of representation primitives. Each model in a particular formalism is built from the same representation primitives. However, the instantiation of these primitives differs according to the particular system that the model refers to. For example, in modelling a refrigerator the variable temperature may be considered relevant and mass irrelevant, whereas in modelling an elevator, mass is relevant and temperature is usually not.

2.3 Two stages of qualitative model construction

Having investigated the properties of models in general, and the properties of qualitative models in particular, the task of constructing a qualitative model is now addressed. The discussion in section

[4]Qualitative reasoning may also encompass inequality reasoning so that whenever some relevant inequality between variables changes, a transition to another state occurs too.

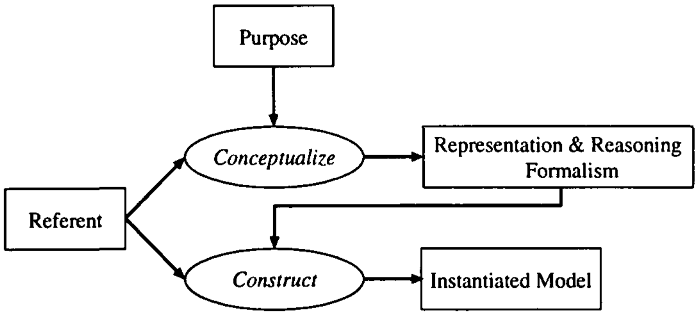

Figure 2 The two stages of model construction

2.1 illustrates that modelling can be considered a two-stage enterprise. Figure 2 illustrates this simplified view on the modelling process. The first stage, conceptualisation, consists of the formulation of a representation and reasoning formalism. The specific purpose of the model determines which aspects of the referent need to be modelled. Then a set of suitable representation primitives has to be selected and a reasoning formalism needs to be defined. The reasoning formalism must comply with the modellers intention: it must derive the type of information that the modeller expects. The second stage consists of the actual model construction, in the sense that the referent is depicted in terms of the specific representation primitives.

In the last decade of qualitative reasoning research, the first stage of modelling, the formulation of a representation and reasoning formalism, has received a considerable amount of attention (Weld & deKleer, 1990). Overviews of such qualitative reasoning formalisms are given in Coiera (1992), Cohn (1989) and Iwasaki (1989). The second stage, however, the process of depicting some systems in terms of a particular modelling framework, has only recently emerged as an explicit research area. The purpose of this article is to provide an overview of current approaches addressing qualitative model construction in the latter sense. Our aim is to survey the different answers that researchers give to the following dual question:

Given a particular representation and reasoning formalism on the one hand and a particular system on the other hand, how is a model of that system constructed and how can this process be (partially) automated?

2.4 Tasks in constructing qualitative models

Different researchers approach qualitative model construction from different circumstances and have different goals when building models. This should be acknowledged in assessing the similarities and differences between the model construction approaches they propose. Therefore, a global frame of reference should be developed for assessing such similarities and differences, and for identifying the conditions under which each approach can be usefully applied. In addition, this frame of reference can be utilised to identify conditions that are not addressed by existing approaches to qualitative model construction. In this way we can make recommendations for directions of future research.

To develop a global framework of reference, a task-oriented view is adopted. We consider qualitative model construction from the perspective of a knowledge intensive task that should be executed. Thus, a general task model of the model construction process itself will serve as the frame of reference. This section presents the model of the process of qualitative model construction from our conception.

The output of the model construction process is a model in terms of a particular modelling framework. Although the formulation of such a modelling framework (including a representation and reasoning formalism) on the one hand, and the construction of an instantiated model within

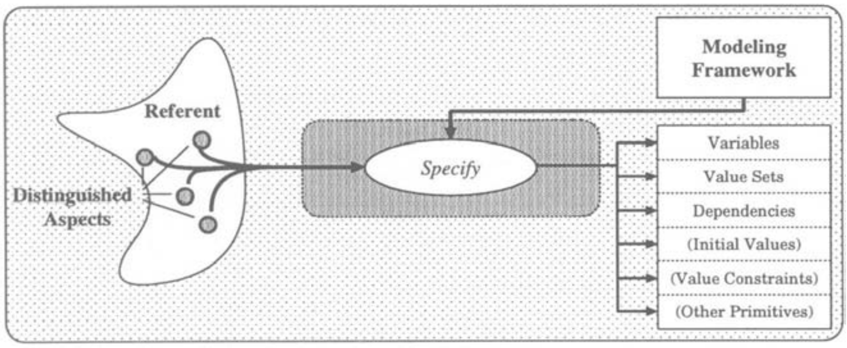

Figure 3 Decomposition of the first subtask in model construction: model specification

that modelling framework on the other hand, are different activities, this does not mean that the two are unrelated. Obviously, the characteristics of the representation formalism determine *which* aspects of a system can be represented at all. But the set of available representation primitives also determines *how* the relevant aspects are represented. Therefore, it is necessary to discuss the relevant modelling framework of the employed formalism for each qualitative model construction approach.

The input for model construction may vary widely. It may consist of observations of system behaviour, of a textbook description of the workings of the system, of an algebraic model, of a modeller's knowledge, and so on. Somehow the particular input has to be converted into an instantiated model, that is, into a set of instantiated representation primitives. Which are the steps that are performed to realise this transformation?

A very general perspective on model construction is to view the process as a hypothesise and test cycle. An initial model is specified and is subsequently modified repeatedly, based on the difference between the model's conclusions and observations of the referent. The task of specifying a model can be decomposed according to the different types of representation primitives that are employed in the formalism. Specification concerns the relevant variables, for each variable the set of values it can take, and the relevant principles that govern the system behaviour. If the successive behaviour of some initial state has to be predicted, then the known initial variable values should be specified too. Also an experimental frame may have to be selected. An experimental frame specifies the boundaries of the behaviours that have to be considered in the form of value constraints, that is, boundaries on the values for which the model is valid. For instance, the modeller might not be interested in the behaviour of a refrigerator whose cooling area is warmer than its surroundings. Finally, depending on the types of primitives employed in the formalism, other characteristics such as the components, their properties, their configuration, their geometry, and so on, may have to be specified. Figure 3 illustrates the subtasks that constitute the task of specifying an initial model.

When an initial model has been formulated, the model construction process enters another phase. Now that a model is available in terms of the employed modelling framework, it has to be assessed and debugged. Therefore, the model is processed by the inference engine to generate a behaviour prediction. The model is satisfactory if the conclusions that it generates, correspond to the observed features of the referent, in this case, if the predicted behaviour corresponds to the observed behaviour. Assessing the model is the process of verifying the desired correspondence between the behaviour predicted by the model and the real system behaviour. There are three reasons why this correspondence may not be perfect: first, the model may contain one or several errors; second, the inference engine may be incorrectly designed; third, the observations and/or measurements of the system were incorrect. Since this article only concerns model construction, we will assume that the inference engine is correct and that the measurements are appropriate. Therefore, if the behaviour prediction of the model does not match the actual system behaviour, the reason is assumed to be an error in the model. Debugging the model is the process of identifying

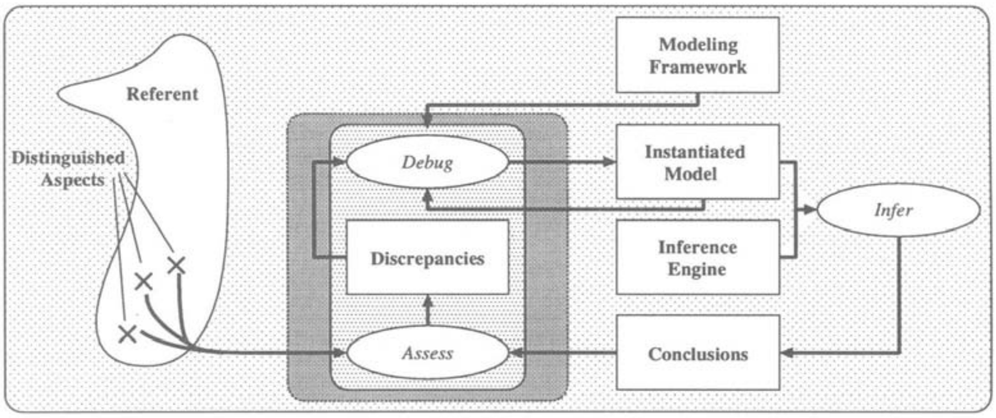

Figure 4 Decomposition of the second subtask in model construction: model assessment and debugging

one or several errors in the model and modifying them to bring the predicted and observed behaviour in line.

In assessing and debugging the model, the information with respect to the observed system behaviour can vary in completeness. In some circumstances, the system behaviour is entirely unknown, in others only the behaviour of some parts of the system is known, and sometimes there is just knowledge about behaviours that definitely cannot occur in the real system. It might even be the case that only a specification of the desired behaviour is known. Although the amount of information may vary, the model assessment task remains the same: predictions of the model have to be compared with the observed system behaviour. Assuming a correct inference engine and appropriate observations and measurements, any discrepancies between predicted and observed behaviour must be attributed to modelling errors: apparently something has gone wrong in representing the knowledge of the system and the physical principles that apply to it. These modelling errors have to be detected and repaired. In repairing modelling errors, the model construction process has become cyclic as the phase of specifying the model is entered again. Figure 4 summarises the subtasks in qualitative model construction.

This general task decomposition will serve as a frame of reference in describing individual approaches to qualitative model construction. The approaches can be partitioned into two categories according to the circumstances they presuppose: model composition approaches and model induction approaches. Both categories address the model construction process from specific circumstances.

3 Model composition approaches

The first category of approaches to qualitative model construction is characterised by the fact that a model is composed automatically from a set of predefined partial models. To realize this, *model composition* approaches make an explicit distinction between knowledge that is specific to a system and knowledge about (semi-) system-independent physical principles. The former are represented in a case model, usually referred to as *scenario*; the latter are represented in *model fragments*, the partial models. Besides facilitating automatic model construction, the distinction between case knowledge (the scenario) and domain knowledge (the model fragments) has the additional benefit that the domain knowledge becomes reusable. Because the physical principles are represented in a system-independent way, the model fragments can be reused in constructing models of different systems.

In the model construction process the modeller should specify a scenario that represents a certain system and its initial conditions. In addition, it is assumed that the domain knowledge has been represented in a library of model fragments. A separate procedure automatically selects those

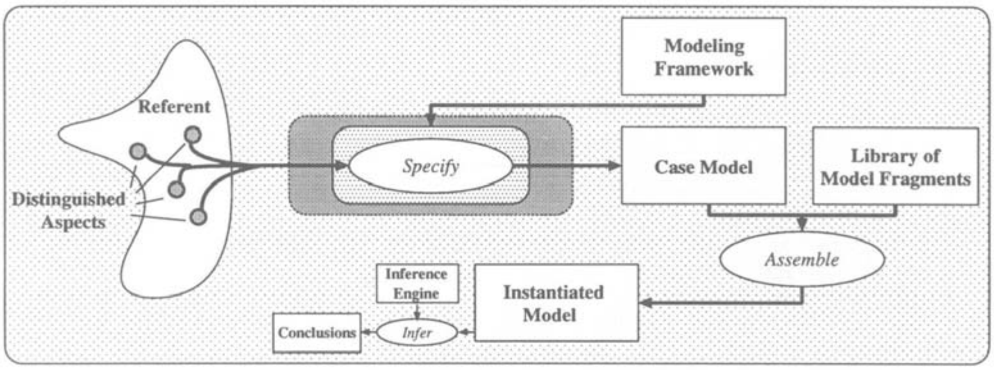

Figure 5 Model construction in model composition approaches

model fragments that represent the relevant physical principles applying to the system. The conjunction of the case model and the model fragments makes up an initial model of the system. Figure 5 illustrates this situation.

Any procedure for selecting the relevant model fragments from a library requires that the conditions under which the different physical principles are relevant, are explicit. Hence, model fragments are stated as structures with a conditional part and a set of givens that are added to the model if the conditions hold. However, appropriate vocabulary is needed for expressing those conditions. Therefore, in discussing individual model composition approaches, both the procedure for selecting the appropriate set of model fragments, as well as the modelling framework are presented. The category of model composition approaches can be subdivided into three major groups: elementary compositional modelling (section 3.1), advanced compositional modelling (section 3.2), both originating from qualitative reasoning research, and bond graph modelling (section 3.3), originating from system dynamics research.

3.1 Elementary compositional modelling

Although not referred to as such, compositional modelling was already present in the earliest qualitative reasoning formalisms of deKleer and Brown (1984) and Forbus (1984). DeKleer's component models and Forbus's individual views and processes can be considered partial models, each describing small parts of domain knowledge independent of their context. The use of such partial models, or model fragments as they grew to be known, relied on the *no-function-in-structure* principle. This principle states that the definition of a particular piece of domain knowledge should not refer to the functioning of the system as a whole. Put another way, the knowledge defined in one model fragment should not be confounded with knowledge of domain parts not referred to in the model fragment. DeKleer already pointed out that it is inherently impossible to define context-free model fragments: a description of a phenomenon at one level inevitably makes assumptions about phenomena at a more detailed level. Therefore, the notion of *class-wide assumptions* was introduced, operationalising the no-function-in-structure principle. Class-wide assumptions are those assumptions that are generic to a class of systems. For example, in modelling a fluid valve, the assumption that a fluid is incompressible is a reasonable assumption that applies to a whole class of hydraulic systems.

In DeKleer's ENVISION system a library of *component models* is employed. A component model describes the potential behaviours that a component may manifest. Behaviour is expressed as a set of *confluences*, i.e., dependencies between variables. A particular system is represented as a *device topology*. A device topology is a collection of components and conduits, i.e., objects that modify materials (e.g., motor, battery) and transport materials (e.g., pipe, wire). Model construction is equivalent to selecting those component models from the library, that match the *device topology*.

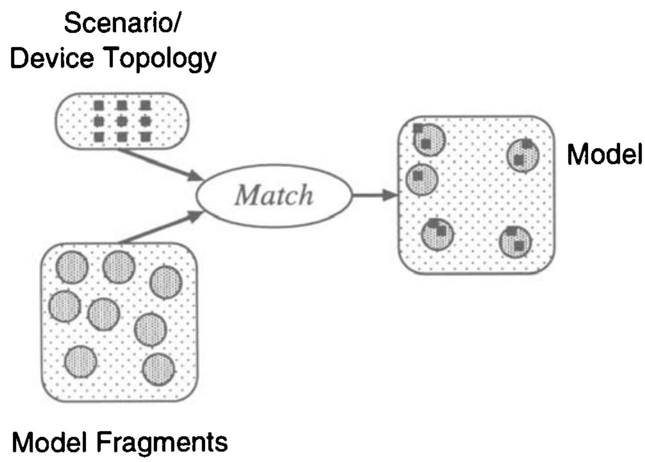

Figure 6 Model construction in ENVISION and QPE

A similar approach is employed in Forbus' QPT, although the considered aspects are different. QPT presumes a *scenario* and a *domain theory*. A scenario is a structural description of the system in terms of physical objects and attributes of those objects. For example, a U-tube consists of two containers, two fluids and a fluid-path that is aligned. The scenario may be augmented with initial values of some relevant variables. A domain theory is a collection of model fragments. Two types of model fragments are distinguished, *individual views* and *processes*. Individual views describe the static characteristics of objects, whereas processes describe dynamic aspects of objects. The difference between individual views and processes is that the former may only specify indirect influences and that processes may specify indirect *and* direct influences. Direct and indirect influences specify the causes of quantity changes. An indirect influence is called a *proportionality* and may be positively or negatively oriented. For example, quantity Q_1 being positively proportional to quantity Q_2 means that, all else being equal, a rise in Q_2 will induce a rise in Q_1. A direct influence is simply called an *influence* and may also be positively or negatively oriented. If a direct positive influence exists of Q_1 on Q_2, it means that if Q_1 has a positive value, then Q_2 tends to decrease. However, note that all direct and indirect influences on a certain quantity are qualitatively summed.

The model construction process is similar in both deKleer's and Forbus' approach (see Forbus (1990) for a description of QPE, the simulator that implements QPT). Both ENVISION and QPE assume a structural description of the system: ENVISION in the form of a device topology, QPE as a scenario. In addition, they assume a library of predefined model fragments that capture the relevant physics knowledge of the domain. The model is constructed by selecting those model fragments from the library that match the structural description of the system. For such a procedure it is essential that the terms used in the device topology or scenario correspond to the terms used in the library of model fragments. Figure 6 illustrates the model construction process in these approaches.

The scheme employed by deKleer and Forbus has been used by several other researchers. Bredeweg (1992) argues that the representation formalisms of deKleer, Forbus and of Kuipers (1986) (this formalism is not discussed because it does not address model construction issues) are complementary in some respects. The approach that he proposes, implemented as GARP, integrates primitives from all three approaches into one framework. Low and Iwasaki (1992) describe the DME (Device Modelling Environment) framework meant to integrate system modelling, simulation and explanation. The modelling phase is done in a compositional manner. The applicable model fragments give rise to an equation model, consisting of all the quantity relations specified by the model fragments. These are translated into QSIM constraints (Kuipers, 1986), and QSIM is used to simulate the model. A similar approach is adopted by Farquhar's QPC

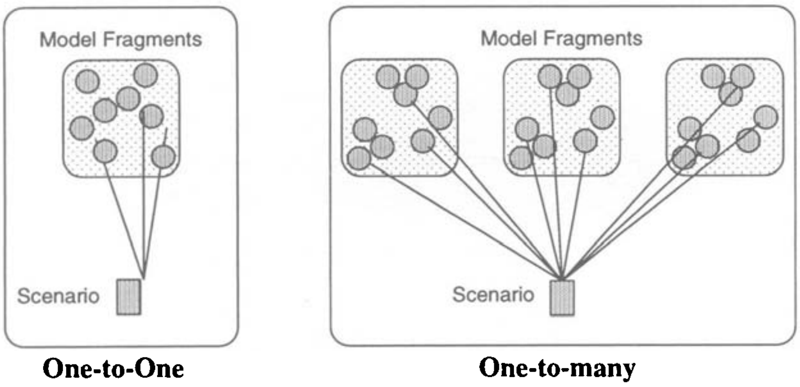

Figure 7 One-to-one and one-to-many mappings from scenario to model fragments

(Farquhar, 1993; Crawford et al., 1990). QPC combines the model construction features of the compositional modelling approach with the simulation features of Kuipers' QSIM formalism. The compositional modelling technique is used to determine which model fragments hold, given a scenario description. Next the closed world assumption is applied and a *qualitative differential equation* (QDE) is built. Note that a QDE usually consists of multiple constraints. In building the QDE the dependencies that are specified by the model fragments, which are in terms of Forbus's QPT (Forbus, 1984), are translated into constraints of the type used in QSIM. Next QSIM is used to generate an initial state and to predict the subsequent behaviour. Both during the generation of an initial state and during the subsequent prediction it may happen that a boundary of the QDE is reached as variable values are computed. This may cause additional model fragments to become active. Therefore, whenever a boundary condition (region transition) is reached, the QDE is recomputed. This entails the execution of a new modelling step.

3.2 Advanced compositional modelling

The approaches of deKleer and Forbus illustrate the emphasis on representation and reasoning issues in early qualitative reasoning research. The approaches were successful in the sense that they provided appropriate vocabularies for expressing qualitative knowledge. One of the shortcomings, however, was the impossibility to generate multiple models for one scenario. For example, a model of a solar heating system need not be very detailed if someone is interested in the effect of the intensity of sunshine on energy saving. However, it needs to contain a lot of detail for explaining why the temperature in one room increases more than in another. Elementary compositional modelling approaches are characterised by a one-to-one mapping from the scenario description to the library of model fragments, as illustrated in Figure 7. The modeller who builds the model fragment library determines how each phenomenon is represented, that is, which of its aspects are covered and which are not. For example, in representing the motion of an object, friction can be considered or neglected. These viewpoints are contradictory, and hence cannot hold at the same time. Once the library has been constructed, such choices in mapping from system to model are fixed. Therefore it can be considered a one-to-one mapping. Although it was recognised that different choices can be made in this mapping (see deKleer and Brown, 1984, section 2.5) such choices are implicit in the particular library that is used. All phenomena in a domain are represented with a fixed level of detail, a fixed perspective, a fixed accuracy, and so on.

To realize the ability to generate multiple models for one scenario description several approaches emerged that can be viewed as providing a one-to-many mapping from the scenario to multiple models of that system. Figure 7 illustrates this. Such models differ with respect to the perspective they take, the level of detail, the accuracy, the approximations entailed, and so on.[5] To

[5]Weld (1992) has proposed a number of *model dimensions* by which models can differ. These ideas were elaborated on by Shen et al (1994).

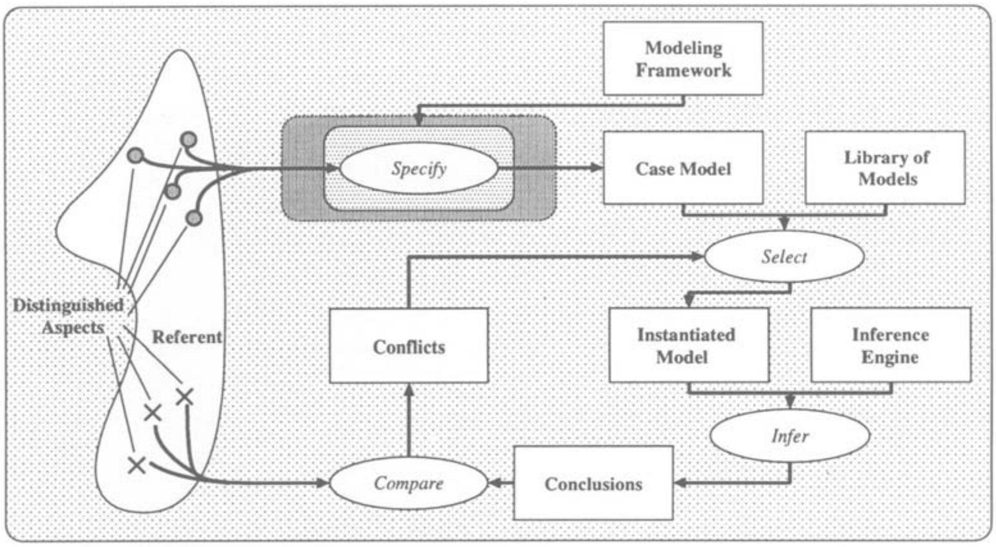

Figure 8 Model construction according to Addanki et al.

generate a specific model from such different viewpoints, additional information is required. The goal for which the model is used determines which of the multiple viewpoints is appropriate. For example, to explain the global working of a refrigerator, the detailed workings of the pump can be neglected. Additional information on the goal of the model can be derived from different sources. One source may be explicit knowledge about the function of the system, i.e., the purpose for which it was constructed. Another source may be found in a query about the system behaviour, i.e., a specific question about some aspect of the system behaviour that the model should answer.

The basic way in which several different viewpoints on the same phenomenon are captured, is by making explicit the assumptions that underly the different perspectives. Therefore the set of representation primitives is extended with elements for representing particular assumptions. Note that assumptions do not concern aspects of the referent of a model, but rather qualify the representation relation between an aspect of the referent and the model. Therefore, assumptions might be considered the modelling decisions that were made in representing the referent. Different viewpoints on the same phenomenon can be distinguished by the assumptions that these viewpoints make. Thus, the assumptions provide a handle to choose among these viewpoints.

The modelling approach proposed by Addanki et al (1991) is, strictly speaking, not a compositional modelling technique. It is discussed, nevertheless, since it first addressed the issue of multiple models and relies strongly on the notion of assumptions. The approach aims to deal with the issue of how to represent and reason with large amounts of physics domain knowledge. It is argued that the use of explicit assumptions permits the decomposition of such large domains into smaller knowledge bases, off-the-shelf models. Analysis can then occur within the simplest model that is an acceptable approximation of the system. The problem is of course how to select the appropriate model. The solution Addanki et al. propose, is to define *graphs of models* where the nodes are models and the edges are the assumptions that have to be changed when switching between models. Given a system description, a model is selected and a behaviour prediction is generated. Then the model is assessed with respect to its goal: it is verified whether the model predicts the set of measured values. Whenever a prediction does not match the measured behaviour, a so-called *empirical conflict* occurs and another model has to be selected. Another model can be selected by changing the set of assumptions and thereby the graph of models is traversed. Figure 8 illustrates this approach in terms of our general characterisation of models. A modeller specifies a case model, i.e., a system description. Then the technique selects a model that is simulated. Based on the conflicts between predicted and observed behaviour another model is selected. For details on the algorithm see Addanki et al. (1991).

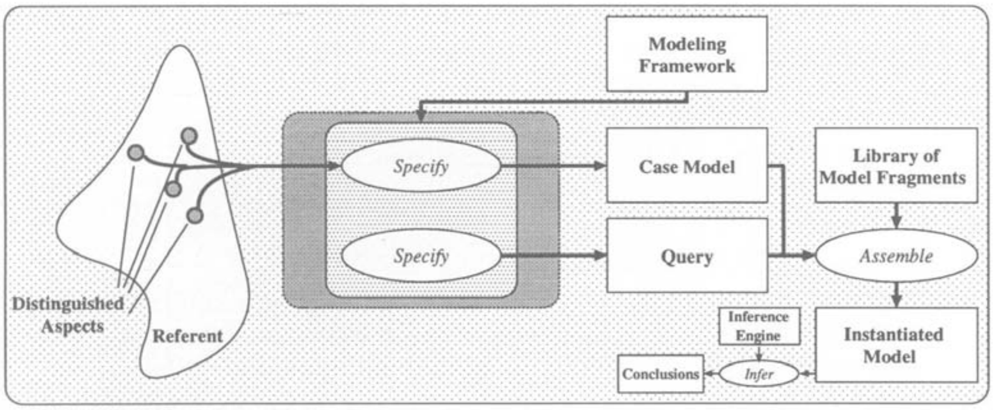

Figure 9 Model construction according to Falkenhainer and Forbus

A different approach is employed by Falkenhainer and Forbus (1991). They also use explicit assumptions but specify them for model fragments as opposed to Addanki et al. who define them for entire models. The employed representation primitives are based on QPT theory, but the distinction between *individual views* and *processes* is no longer made. All domain knowledge is represented in model fragments. Each model fragment, however, is made conditional on a (possibly empty) set of assumptions. In this way incompatible views on the same object(s) or process(es) can exist side-by-side in one library. Such clusters of mutually contradictory model fragments are called *assumption classes*. They capture the fact that selecting one assumption excludes other assumptions. For example, a fluid can be modelled as having zero viscosity, a standard non-zero viscosity, or a non-Newtonian viscosity (e.g., toothpaste), but only one alternative can hold. An assumption class is ordered internally with respect to the complexity of the assumptions. For example, including the zero-viscosity assumption in a model would yield a simpler model than including non-zero viscosity.

Two types of assumptions are distinguished: simplifying and operating assumptions. Both are used to control the instantiation of model fragments so that only the relevant aspects of a situation are considered. Simplifying assumptions make the underlying approximations, perspectives and granularity explicit. Operating assumptions state default assumptions about system behaviour to manage complexity. They serve to focus simulation by ruling out entire clases of behaviour. Both types of assumptions apply to individual instances in the scenario as opposed to the scenario as a whole. This specificity allows the construction of a finely tuned model in which, for example, for one motion friction is neglected whereas for another motion friction is considered.

Basically, the process of model construction consists of collecting a set of appropriate model fragments for a given scenario and a *query* (see Figure 9). A query is a question about a specific feature of the system that the model should answer, for example, *Does the temperature of the fluid reach the boiling point?*. The assumptions provide the handle to generate the set of appropriate model fragments. An ATMS (deKleer, 1986) is used to keep track of consistent sets of assumptions, referred to as *modelling environments*. First, all model fragments are collected whose structural conditions are satisfied by the scenario description. That is, the conditional individuals (i.e., components) of each model fragment are matched against the individuals specified in the scenario. This restricts the search space to the model fragments that might possibly apply. Subsequently, the relevant viewpoints on the system have to be selected from the set of applicable model fragments. The first step is the *analysis* of the *query*. The idea is that the query contains information about individuals, quantities and relations that must be included in the model. This is followed by the *object expansion* step. The intuition is that certain components of an object cannot be considered in isolation, but that enough of the situation has to be considered for all relevant components to be included in the model. For example, if two components of the same object are considered, then all of its other components have to be considered as well. After object expansion all relevant components are included in the model but it still has to be decided how each component should be modelled. In the *candidate completion* step all consistent ways of modelling the set of relevant components are collected. This step amounts to finding sets of consistent assumptions for all components. In the final step, *candidate evaluation and selection* one set of assumptions has to be selected. It is based on a preference for models including few components and simple assumptions.

3.2.1 Extensions to compositional modelling

Several researchers have proposed extensions and/or modifications to the approach of Falkenhainer and Forbus. Nayak (1992) formalises the notion of *simplicity* of a model and introduces *causal approximations* as the basis for identifying the simplest model. The formalisation of model simplicity captures the intuition that a model is simpler if it represents fewer phenomena and/or represents phenomena more approximately. For example, modelling a wire in an electrical circuit as a constant resistor neglects two possible effects: of temperature on resistance, and heat generation by the current flowing through it. Therefore, it results in a simpler model than modelling the wire as a temperature-dependent resistor. Nayak argues that the problem of selecting the simplest model can be solved in polynomial time if all different viewpoints on a phenomenon concern *causal approximations* (see Nayak (1992) for a precise definition).

The procedure that automates the generation of the simplest model for a system (Nayak et al.,1992), requires additional information in order to select the appropriate set of model fragments. For this purpose the structural and behavioural contexts of components, and the expected system behaviour are used. The former, the structural and behavioural context, capture two intuitions. The first intuition is that certain physical and behavioural properties, e.g., object is metallic, or pressure exceeds threshold, restrict the viewpoints that may apply to a component. The second intuition is that the choice of a particular viewpoint for some component implies certain viewpoints for other components. The latter, the expected behaviour, is operationalised by a pair of input and output quantities. An adequate model must be able to explain how the input quantity has an effect on the output quantity. Therefore, the model must include these quantities and a causal chain of intermediate quantities.

The initial phase in model construction concerns finding an adequate model. This phase consists of four steps. In the first step, a given system description is extended with the quantities that are specified in the expected behaviour of the system. In the second and third steps, the structural and behavioural contexts are used to collect a set of applicable model fragments. In the fourth step it is checked whether the model encompasses a causal chain of quantities that includes the quantities from the expected behaviour. If this is true then the second phase is entered. If not, the set of selected model fragments is adapted: less general model fragments are selected, until the model satisfies the expected behaviour. The second phase in model construction concerns simplifying the adequate model. Individual model fragments are removed from the model and/or replaced by more approximate model fragments until no simpler model can be found.

As a second extension of compositional modelling we discuss the work of Rickel and Porter (1992). They argue that, given a set of predefined model fragments, model construction involves two issues. The first is the identification of the relevant aspects of the scenario: the *scope* of the model.[6] The second is the identification of the appropriate *level of detail*. For example, take the question *How would a decreasing amount of soil water affect plant size and growth rate?* For answering this question the scope of the model must include the soil. Including all details of cell-metabolism in the model, however, would lead to an inappropriate level of detail with respect to the question.

To realise dynamical choice of scope and level of detail for different parts of the system that is modelled, Rickel and Porter expand the set of aspects of the referent that are considered relevant (and the set of representation primitives) to include *time scales*. Model fragments are annotated

[6]In more recent work (Rickel and Porter, 1994) referred to as the *system boundary*.

with time scale conditions. The search for a model with appropriate level of detail is based on finding a causal interaction path from some given quantity to a quantity of interest, where all interactions occur on the same time scale. The appropriate scope is determined by reasoning about the relevance of influences. All influences at the selected time scale that are connected (in)directly with a quantity of interest, are relevant.

Iwasaki and Levy (1993) propose another type of extension to the compositional modelling approach (see also Levy et al., 1992). This extension concerns the selection of model fragments on the basis of explicit *relevance* and *irrelevance* assumptions. These relevance and irrelevance assumptions are used to distinguish the different perspectives with which the same phenomenon can be considered. In model formulation they are used to generate a set of model fragments that is consistent with respect to the set of (ir)relevance assumptions that hold. See the original article for details on the algorithm.

3.2.2 Constructing the library of model fragments

The compositional modelling approach and its extensions reduce the model construction task to the problem of selecting the appropriate model fragments from a predefined library. Such an approach raises the question how a model fragment library was constructed in the first place. This question is not addressed by any of these approaches although a proper model fragment library is crucial for obtaining a satisfactory model. The assumption that a correct and complete set of model fragments is available also implies that model assessment and debugging need not be considered. However, if the correctness and completeness of a model fragment library is not guaranteed, compositional modelling and its extensions are inapplicable. In these cases a modeller must construct all missing model fragments and can only then revert to the model fragment selection techniques of the described approaches.

An approach that addresses model construction in situations where the necessary domain knowledge is not or incompletely available, is provided by Schut and Bredeweg (1994). They present a system that supports a modeller in specifying a model. It presents an analysis of the tasks that face a modeller in such situations. On that basis, a set of dedicated editors has been developed that support the formulation of a case model and a set of model fragments. The characteristics of the particular knowledge representation formalism that is used are employed in supporting the modeller. Besides model specification, they also address the subtask of model assessment and debugging. They describe techniques for modifying a model semi-automatically (in interactions with the modeller) in cases where the model predicts derivatives incorrectly (Bredeweg & Schut, 1993; Schut & Bredeweg, 1995). Another technique automatically finds model simplifications once a model has been constructed that predicts the system behaviour properly (Schut & Bredeweg, 1993). The coverage of these techniques is depicted in Figure 10.

3.3 Bond graph modelling

Another type of model composition approaches concerns bond graph modelling approaches. The bond graph concept has been developed in system dynamics (Paynter, 1961; Karnopp et al., 1990). Physical systems are conceptualised in terms of a limited set of energy processes (e.g., transformations, flows). Energy processes capture physical phenomena at a very abstract level. As a consequence, bond graph modelling is applicable to the entire range of physical domains.

Constructing a bond graph model proceeds by mapping the bond graph elements on the particular device, as Figure 11 illustrates. The use of bond graphs places certain demands on the knowledge possessed by the modeller. The modeller is not only expected to possess knowledge of the elementary bond graph concepts, but is also expected to possess substantial physics knowledge. The latter is required for applying the direct mapping from actual physical phenomena into bond graph concepts and for interpreting bond graph models. Such knowledge requirements are stronger than the requirements that qualitative reasoning approaches pose.

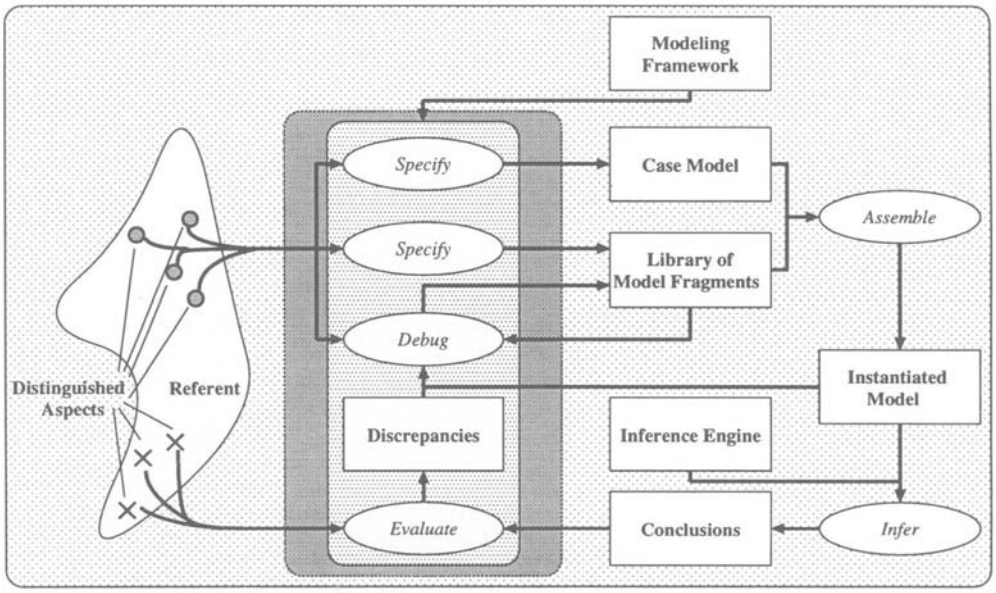

Figure 10 Model construction according to Bredeweg and Schut

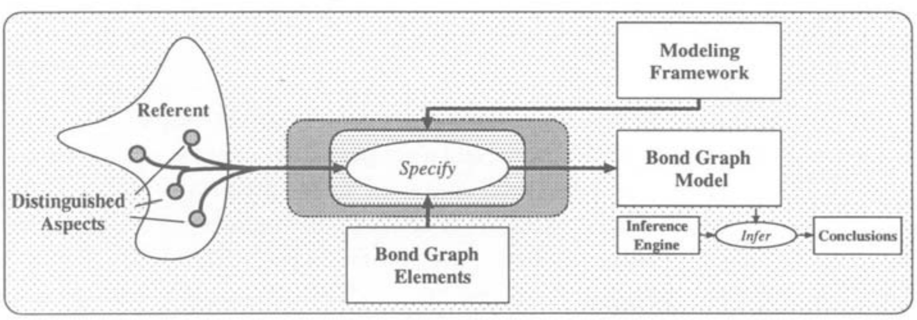

Figure 11 Model construction using bond graphs

The reason is that the conceptualisations employed in qualitative reasoning approaches are closer to the common sense conceptualisations that humans have. The knowledge requirements make bond graph models suitable in contexts where the modeller or model user is a physics expert.

If the modeller possesses the required knowledge, then the use of the bond graph language for representing physical phenomena has the advantage that it is a well-formalised language. As a consequence, checking the consistency and completeness of a model can be done very efficiently. In addition, causal explanations can be generated from the bond graph model directly (see, for instance, Top, 1993). The disadvantage is that bond graph models can not be used for generating behaviour predictions. The bond graph model has to be translated into a format that is suitable for a particular simulator. As such, bond graph models can be considered intermediate models.

In this fashion bond graphs are used by Top (1993), who describes the process of conceptually modelling physical systems. He argues that model construction should proceed by traversing four levels of description. These levels concern the functional components, the physical processes, the mathematical relations and the model data. The first level describes a system in terms of idealised functional components. The second level describes the bond graph model. It is generated from the component level through links from functional components to bond graph elements. The bond

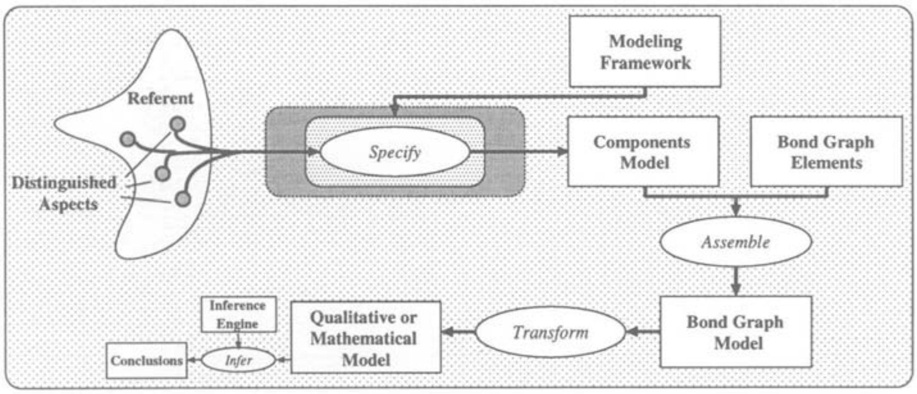

Figure 12 Model construction using a bond graph model as an intermediate model

graph model is used to assign the causality to the model. The third level constitutes the mathematical structure of the model. The bond graph model restricts the range of possible mathematical relations. Finally, the model data represent the variables of interest, their values and their accuracy.

The approach of Xia et al. (1993) is similar to Top's approach, in that bond graph models are used as intermediate models. At the highest level a device is represented in terms of generic components. The intermediate bond graph model links the generic components model to a qualitative model. The model construction technique they describe relies on a knowledge base of generic components that have an associated bond graph element. Some of the generic components have multiple associated bond graph representations. In their turn, relationships between bond graph elements are associated with qualitative relationships. The automatic model construction procedure takes two inputs: the generic components model and a set of (given) behavioural properties that the model should capture. Initially, the simplest bond graph elements associated with each component are selected. Analysis of the resulting bond graph and qualitative model may reveal that the model does not capture the specified behavioural properties. In such cases, for some components more complex associated bond graph elements are selected and the analysis is repeated. Eventually this results in the simplest satisfactory model. Figure 12 illustrates the use of bond graph models as intermediate models.

4 Model induction approaches

The second category of approaches to qualitative model construction presumes circumstances that differ strongly from those presumed by model composition approaches. This category concerns *model induction* approaches, which originate from the field of machine learning. These approaches aim to infer a model of a certain system from the behaviour it exhibits. The input for such systems always consists of observed system behaviours, usually referred to as positive examples. Individual techniques pose varying requirements in this respect. For instance, some require that *all* possible behaviours are provided, some require that *impossible* behaviours are provided too, and others that a description of the relevant components of the system is provided. The output of model induction approaches is a model of the system, where a model is taken to be a set of qualitative dependencies. This approach to qualitative model construction is illustrated by Figure 13. There is a vast body of literature dealing with this theme that is often denoted as *discovery learning*. Since we can not possibly cover this entire subfield of machine learning, we have chosen to discuss the work that is most relevant to qualitative model construction. The contributions are subdivided into *inductive logic programming* approaches and a number of miscellaneous approaches.

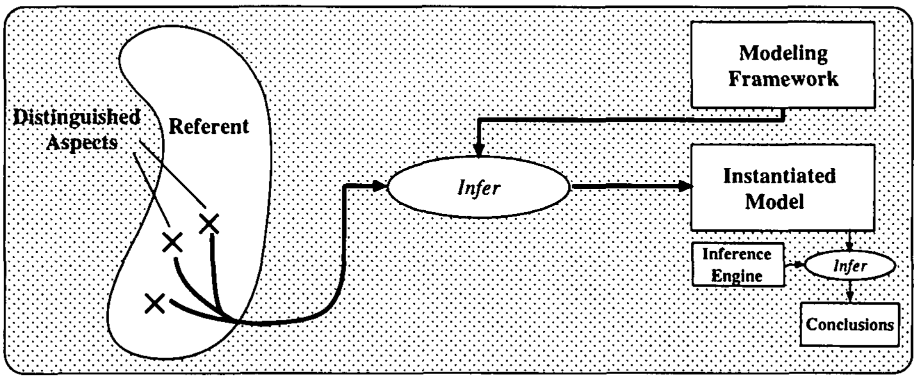

Figure 13 Model construction by induction

4.1 Inductive logic programming approaches

The earliest inductive logic programming (ILP) approach that was aimed at the generation of qualitative models, is that of Coiera (1989; Hau & Coiera, 1995). He developed GENMODEL, a system that learns qualitative models in terms of Kuipers (1986) QSIM formalism from a sequence of state descriptions. GENMODEL can be viewed as working in the opposite direction of QSIM. Whereas QSIM predicts qualitative system behaviour using a model that consists of a set of qualitative constraints, GENMODEL uses qualitative system behaviour to infer the qualitative constraints that permit that qualitative behaviour. The input for the algorithm is twofold. The first input is a sequence of qualitative state descriptions, each representing a single time-varying behaviour. A qualitative state is described by a number of system functions, characterised by their value and derivative. Note that behavioural data in the form of measurements of system variables should first be converted into qualitative behavioural data. This measurement interpretation problem, which already has been mentioned in section 2.1, is also addressed by GENMODEL. It will not be elaborated on here because it is a separate issue (see Forbus (1983, 1986) and DeCoste (1991) for discussions on this issue). The second input is background knowledge concerning the different types of QSIM constraints on system functions, and the conditions under which value combinations satisfy those constraints.

The general idea that is operationalised in GENMODEL is to identify all possible constraints, i.e., those constraints that are consistent with the initial state, and to prune this set gradually by considering subsequent qualitative behaviour states. The method that GENMODEL employs in learning a qualitative model is as follows. First, the initial state is selected. Then for each system function in the initial state all landmarks, i.e., the points at which a function reaches a different qualitative value, are collected. Also all combinations of function values are collected. Next, it is determined which constraints are consistent with those value combinations. The background knowledge is utilised in this step by having the legal QSIM constraints serve as templates. The set of candidate constraints is initially large, but is pruned as subsequent qualitative states are processed. Each of the next states may contain value combinations that are incompatible with some of the candidate constraints. When all states have been processed, any redundant constraints (e.g. $a + b = c$ is equivalent to $b + a = c$) are removed. The remaining constraints make up the most constrained model that allows the input behaviour. Any subset of the remaining constraints allows that behaviour too.

In later versions of the GENMODEL system (Hau & Coiera, 1995) two extensions are incorporated: dimensional analysis and fault tolerance. Dimensional analysis is employed to reduce the initial search space. Before accepting a constraint as a candidate for inclusion in the model, the compatibility of the units of the functions in the constraint is checked. Fault tolerance is incorporated to deal with noise in the learning data. It is operationalised as a simple threshold criterion for deciding when to prune a candidate constraint. Only if a constraint is incompatible with a certain fraction of the states, it is pruned.

The MISQ system developed by Richards, Kraan and Kuipers (1992) is an extension of the GENMODEL algorithm. It also induces models that have the QSIM format. MISQ can take hand or sensor generated quantitative data as input and converts these into qualitative behaviours. But if already available, qualitative behavioural data can serve as input too. Regardless of how they were obtained, by conversion or directly supplied, subsequently the qualitative behaviours are used to collect sets of QSIM constraints that are compatible with these behaviours. This step proceeds similar as in GENMODEL. Tuples of functions are generated and the values of these functions are checked against the satisfaction conditions of each type of QSIM constraint. In addition, the dimensions of the functions in the constraint are checked to be consistent. If complete behavioural information on the input behaviours is available, then this constraint generation process renders a unique model guaranteed to reproduce the input behaviours. However, if the behavioural information is incomplete then further processing is required. This occurs in several situations: when variables are missing in the behaviour descriptions; or when one or several qualitative values are missing; or when dimensional information is missing. In the latter case, several models are generated that all reproduce the input behaviours. The user has to select one. In the former cases, relational pathfinding is used to introduce new variables into the model. This feature is the major extension of the GENMODEL system.

Another example of the inductive logic programming approach is the work of Bratko, Muggleton and Varšek (1992). They use GOLEM (Muggleton & Feng, 1990) to infer a qualitative model. Like the above approaches, the generated model is in terms of the constraints employed in Kuipers' QSIM formalism. To realise this model induction capacity, the QSIM theory is reformulated in logic as a set of ground facts. These provide the background knowledge for learning qualitative models. A set of observed qualitative behaviour states of a system serves as examples for the learning system. These states are converted into a format acceptable for GOLEM, i.e., as *legal states* of variable value and derivative pairs. Next to these positive examples, the system requires a number of negative examples (near misses). Negative examples are qualitative behaviour states that do not occur in the behaviour of the system, but differ only minimally from legal states. These negative examples must be constructed by hand. GOLEM generates from these positive and negative examples, i.e., state descriptions in terms of values and derivatives, a model in terms of QSIM constraints. Like MISQ, in model generation new variables can be created. Unlike the above two approaches, however, GOLEM does require hand-generated carefully chosen negative examples and cannot guarantee that a model will be found.

4.2 *Miscellaneous approaches*

The work of Mozetič et al. is one of the earliest examples of qualitative model induction (Mozetič, 1987; Mozetič et al., 1990). The models that form the output of their system are based on the component oriented approach of deKleer (1984; see also section 3.1). Qualitative models are specified on different levels of abstraction. On a single level a model is specified by its structure (i.e., a set of components and their connections), the functions of the components (i.e., their qualitative states), and some utility predicates. The latter are used, for example, to specify the ordering of values. Each level of abstraction represents a system on a different level of detail: components may be decomposed, values may be refined, or additional variables may be introduced as the model becomes more detailed. The learning procedure consists of three steps. First, the user provides some instances of the behaviour of components or of impossible behaviour. The *learner* induces hypotheses for the function of the components. The algorithm that is employed is an extension of the AQ algorithm (Michalski, 1983). The hypotheses are input for the *interpreter* that tests the hypothesised model on a new component behaviour. If the model does not behave as intended, the *debugger* is invoked. The debugger employs the more abstract model to restrict the search space for generating new hypothesised functions.

An entirely different approach is adopted by Varšek (1991). His QME system employs a genetic algorithm (see Holland, 1975, and Goldberg, 1989) for learning a qualitative model in terms of

QSIM constraints. Input for the learning process is formed by a set of positive and negative examples which correspond in QME to legal and illegal system states. The basic idea in QME is the following. Throughout the learning process a *population* of *candidate solutions* is maintained. In this case, the candidate solutions are QSIM-based models. Given an initial population, the process operates in cycles called *generations*. Each generation consists of three steps. First, all members of the population (candidate models) are evaluated according to a *fitness function*. The fitness function assigns each member a value that represents how well it covers the set of examples. Second, the population is subject to *reproduction* in several iterations. In reproduction, parents (models) are selected from the population and *genetic operators* are applied. The fitness value determines the probability of selection. In applying a genetic operator to a parent, it is copied to produce offspring and subsequently the model that constitutes the offspring is modified. For example, parts of two models are interchanged, which is called crossover, or in individual models certain parts are changed randomly, called mutation. In the final step of a generation cycle the offspring are inserted into the population, and a certain fraction of the population is replaced. Eventually this process generates a set of models with different specificity, that explain all examples.

Falkenhainer has addressed both the issue of learning qualitative models, as well as integrating the discovery of quantitative and qualitative laws. His ABACUS system (Falkenhainer, 1986) addresses the latter issue. ABACUS uses qualitative proportionalities between system variables to focus on certain variables in the search for quantitative equations. A different approach is adopted in the Phineas system (Falkenhainer, 1987). Whereas ABACUS and related approaches (e.g. Langley et al., 1986) summarise behaviour, they do not explain nor verify the laws they discover. Phineas is aimed at covering the full range of tasks found in scientific theory formation. It does so by reverting to *analogical reasoning* for discovering new models and to assess the new models according to different types of *verification*. The system is based on the knowledge representation of Forbus' QPT.

Phineas takes as input a novel situation, that is, a set of variable measurements at different time instances that cannot be accounted for by current models.[7] The system generates a new model by the formation of two mappings based on analogy. First, the system scans previously observed behaviours, recorded in a history library to find a similar behaviour, i.e., quantities that behave in the same manner. For example, suppose a model of liquid flow is already known and a new situation is encountered where two objects of unequal temperature are placed in contact. The system now discovers that the behaviour of the new situation is similar to that found in liquid flow. Next, the model fragments that were employed in the analogical behaviour, are mapped into the new domain, yielding a new model (e.g., heat flow). If the new model adequately explains the phenomenon, the consistency of the model has been verified. For establishing this, the behaviour that the new model predicts is compared with the original measurements. If a behaviour can be found that corresponds to the measurements, then the model's consistency has been established. Then, the other behaviours that the new model predicts are considered. By recollecting earlier experiences from the history library, or conducting experiments, the additional predictions of the model are verified. If the analogy still proves useful, the limits of the analogy should be explored. This proceeds by considering the consequences of the additional predictions. If these consequences prove correct, then the extension of the new model has been verified.

Nordhausen and Langley (1993) have developed IDS, a system that induces both qualitative and quantitative models. IDS uses an isa-hierarchy of objects as background knowledge and a number of sequences of qualitative states as data (examples).[8] Qualitative states consist of a number of objects, their structural relations, and the changing features (i.e., variable derivatives). Sequences

[7]Here again the problem of measurement interpretation crops up. In Phineas it is tackled by Forbus' ATMI (Forbus & Gentner, 1986).
[8]The IDS system also deals with quantitative data and generates quantitative relations between variables, but these are less relevant for our purpose.

of qualitative states are formed by linking them temporally as successor links. The output of IDS consists of a taxonomy of qualitative states and a set of conditional qualitative laws. A taxonomy of qualitative states is a hierarchy of states with input states as terminal nodes and abstract or generalised states as internal nodes. Qualitative laws are relations between pairs of abstract qualitative states. As such, they give the equilibrium changes that follow changes in the structure of the objects. For example, connecting two containers filled with water of unequal levels will cause one level to decrease and the other to increase.

5 Summary and conclusions

In this article we have described a variety of approaches to qualitative model construction. The article began with a characterisation of the nature of modelling. We identified the main features of models in a broad sense, and gave an account of how these features manifest themselves in the specific case of qualitative models. Qualitative models traditionally refer to physical systems, although this has more historical then principal reasons. The purpose of qualitative models is to predict and/or explain system behaviour over time in situations where, for example, numerical information is lacking, qualitative conclusions are cheaper and still satisfactory, or where qualitative knowledge is simply more appropriate. These specific purposes determine the particular modelling framework that is chosen: particular aspects of the referent are distinguished which have corresponding representation primitives. In addition, these purposes determine which aspect, namely the time varying behaviour, should be derived by the inference procedure, and which aspects are used as input for the inference process. The input must consist of a set of qualitative dependencies between variables, and may be accompanied by structural information, spatial information, and so on. By instantiating the representation primitives for a particular referent, a model of that referent is constructed.

Given these features of qualitative models, a general task model for qualitative model construction was presented. Qualitative model construction can be viewed as a cycle of specifying an initial model and assessing and debugging that model. This general task model served as a frame of reference for considering the various approaches to qualitative model construction. Two broad categories of approaches were discussed: model composition and model induction approaches. These approaches were discussed from the perspective of the general task model. This perspective induced us to focus on four issues: the *conceptualisations* that are employed, the manner in which an initial model is *specified*, the manner in which the appropriateness of the models is *assessed*, and the manner in which models are *debugged*.

5.1 Conceptualisation

The different conceptualisations employed by the various approaches are determined by the purpose of the model. Typically a model can be used to predict the behaviour of a system, to support the generation of explanations, and to automate the model construction process itself. As the purpose of the model becomes more complex, richer conceptualisations are required.

If only prediction is the issue then a restricted conceptualisation suffices (e.g. QSIM). If explanation is included in the purposes of the model, and especially if automated model construction facilities are provided, then the conceptualisation needs to be more elaborate. In addition, a separation of the case and the domain knowledge is required. Exemplary is the work of Farquhar (see section 3.1). QPC aims to augment QSIM with automated model construction facilities. This is dealt with in a similar fashion as in the compositional modelling approaches: physical principles are generically described and made conditional on structural properties that may be found in different systems. To express such conditions, additional distinctions in the conceptualisation have to be made: the structure of a system in terms of its components, their attributes and configuration must be represented too. A similar shift to incorporate additional vocabulary for automating the model constructing process can be observed in the bond graph approaches.

To model a system in multiple ways (depending on the information that the model is expected to generate), the assumptions underlying the different perspectives on physical phenomena are made explicit. The description of these different perspectives requires additional vocabulary too. Compositional modelling and its extensions specifically address this issue, and therefore they require more complex conceptualisations. In general it seems fair to conclude that increasing automated model construction facilities requires extension of the vocabulary for representing systems and for representing modelling decisions.

5.2 *Specification*

The compositional modelling approach and its extensions address the model specification task only in a restricted sense. It is up to the modeller to provide an initial description of the system (e.g., device topology, scenario). In addition, it is assumed that a library of model fragments is available that describe all potentially relevant domain principles. Given an additional system description and a library of domain knowledge, these approaches provide facilities to select the appropriate model fragments from this library. With respect to this automatic selection procedure these approaches are very successful. However, the problem of how to construct and maintain such a library in the first place, is almost entirely neglected. In other words, there is no insight in the process of constructing a library of model fragments that is free of errors and complete enough to be used in modelling a variety of systems.

Model induction approaches inherently address the model specification task but can only provide leverage in some cases. The induction of dependencies between quantities that underly the behaviour of the system only suffices for representation formalisms that do not distinguish case and domain knowledge. These approaches cannot (yet?) induce the conditions under which these dependencies hold, and therefore cannot be used in generating a domain theory as used in compositional modelling. It is an open question whether model induction approaches can be designed that can come up with such conditional pieces of domain knowledge.

5.3 *Assessment and debugging*

Model assessment and debugging are common in model construction methods aimed at other types of models (Zeigler, 1976). Surprisingly however, they are not addressed by the majority of approaches that we have discussed. Certain model induction approaches (e.g., GENMODEL, MISQ, Phineas) perform this task inherently: the model is based on behavioural information, so the validity of the model is guaranteed. The bond graph approaches and the approach of Addanki et al. incorporate model revision on the basis of model output and behavioural criteria. In the compositional modelling paradigm only the approach described in section 3.2.2 addresses model assessment and debugging explicitly. This indicates that more effort could be invested in the development of methods for assessing the performance of models, and for automatic repair of models.

5.4 *Concluding remarks*

Researchers in artificial intelligence generally agree that building computational models is a difficult task. Constructing qualitative models is not an exception in this respect. Although rich conceptualisations have been introduced for the formulation of qualitative models, the process of model construction itself has not been addressed in great detail. This omission constitutes a bottleneck for the use and application of qualitative reasoning formalisms. Recent years, however, have shown a growing awareness in the qualitative reasoning community that the problem of qualitative model construction should be tackled explicitly. In this article we have presented an overview and analysis of the achievements to date. This brings up the question which direction future research should take.

In our opinion, the aspect that seems to be lacking most in the current approaches to model construction is a *task-oriented* view on the process. In the model induction approaches the task of model construction is almost completely automated. That is, the model is constructed by an artefact, almost without any help of a modeller. As a consequence, the modeller has no means at all to influence the outcome of the model construction task. There are no facilities for tuning the model to the specific goals that the modeller may have. The same applies to the compositional modelling approaches. In these approaches the modeller is involved in the modelling process, but only in a very limited sense. Again the model construction process proceeds beyond the control of the modeller. The real modelling problem here (i.e., how to construct and maintain a library of model fragments) is not addressed. In fact, it is unclear who should construct a library in the first place.

In our conception, current approaches are too restricted in their focus on automating the construction of qualitative models. Maybe the time has arrived for the qualitative reasoning community to incorporate ideas from the knowledge acquisition field. Instead of automating the model construction process entirely, the focus should be on the development of model construction tools that deliver models in cooperation and interaction with the actual user of the model. Considering the modelling problem from this perspective, it seems evident that more effort should be aimed at developing *model building environments*, dedicated to the cooperative construction of qualitative models. Such environments should permit man and machine to work together, both providing the capabilities that they are best at. In section 3.2.2 an initial attempt in this direction is discussed. However, it is still a small step onto a long road.

Acknowledgements

We would like to express our gratitude to Bob Wielinga and Kees de Koning for proofreading various drafts of this article and providing us with valuable comments. In addition, we thank the anonymous reviewers for their insightful comments and suggestions. This research was supported by the Foundation of Computer Science in the Netherlands (SION) with financial support from the Netherlands Organization for Scientific Research (NWO), project number 612-322-307.

References

Addanki, S, Cremonini, R and Penberthy, JS, 1991, "Graphs of models" *Artificial Intelligence* **51** 145–177.

Bratko, I, Muggleton, S and Varšek, A, 1992, "Learning qualitative models of dynamic systems" In: S Muggleton, editor, *Inductive Logic Programming*, Academic Press, pp 437–452.

Bredeweg, B, 1992, "Expertise in qualitative prediction of behaviour" PhD thesis, University of Amsterdam.

Bredeweg, B and Schut, C, 1993, "Reducing ambiguity by learning assembly behaviour" In: *Proceedings of the 13th International Joint Conference on Artificial Intelligence* pp 980–985, Chambéry, France.

Cohn, AG, 1989, "Approaches to qualitative reasoning" *Artificial Intelligence Review* **3** 177–232.

Coiera, E, 1989, "Generating qualitative models from example behaviour" Technical Report DCS 8901, Department of Computer Science, University of New South Wales.

Coiera, EW, 1992, "The qualitative representation of physical systems" *The Knowledge Engineering Review* **7** (1) 55–77.

Crawford, J, Farquhar, A and Kuipers, B, 1990, "A compiler from physical models into qualitative differential equations" In: *Proc. of AAAI-90*, 365–372, Boston, MA.

de Kleer, JH, 1986, "An assumption-based TMS" *Artificial Intelligence* **28** 127–162.

DeCoste, D, 1991, "Dynamic across-time measurement interpretation" *Artificial Intelligence* **51** 273–341.

deKleer, J and Brown, JS, 1984, "A qualitative physics based on confluences" *Artificial Intelligence* **24** 7–83.

Falkenhainer, B, 1987, "Scientific theory formation through analogical inference" In: *Proceedings of the 4th International Workshop on Machine Learning* 218–229, Irvine, University of California.

Falkenhainer, B and Forbus, K, 1991, "Compositional modelling: Finding the right model for the job" *Artificial Intelligence* **51** 95–143.

Falkenhainer, BC and Michalski, RS, 1986, "Integrating quantitative and qualitative discovery: the abacus system" *Machine Learning* **1** 367–401.

Farquhar, A, 1993, "Automated Modeling of Physical Systems in the Presence of Incomplete Knowledge" PhD thesis, University of Texas, Austin.

Forbus, KD, 1983, "Measurement interpretation in qualitative process theory" In: *Proc. of 8th IJCAI* 315–320, Karlsruhe, FRG.

Forbus, KD, 1984, "Qualitative process theory" *Artificial Intelligence* **24** 85–168.

Forbus, KD, 1986, "Interpreting measurements of physical systems" In: *Proceedings of the AAAI* 113–117, San Mateo, CA, Morgan Kaufmann.

Forbus, KD, 1988, "Qualitative physics: Past, present, and future" In: HE Shrobe, editor, *Exploring Artificial Intelligence* 239–296, Morgan Kaufmann.

Forbus, KD, 1990, "The qualitative process engine" In: DS Weld and J deKleer, editors, *Readings in Qualitative Reasoning about Physical Systems* 220–235, Morgan Kaufmann.

Forbus, KD and Gentner, D, 1986, "Learning physical domains: Toward a theoretical framework" In: RS Michalski, JG Carbonell and TM Mitchell, editors, *Machine Learning: An Artificial Intelligence Approach: Volume II* 311–348, Morgan Kaufmann.

Goldberg, DE, editor, 1989, *Genetic Algorithms in Search, Optimization, and Machine Learning* Addison-Wesley.

Hau, DT and Coiera, EW, 1995, "Learning qualitative models of dynamic systems" Technical Report HPL-95-31, Hewlett Packard Laboratories.

Holland, JH, editor, 1975, *Adaptation in Natural and Artificial Systems* University of Michigan Press.

Iwasaki, Y, 1989, "Qualitative physics" In: A Barr, P Cohen and E Feigenbuam, editors, *The Handbook of Artificial Intelligence vol. IV* 323–413, Addison-Wesley.

Iwasaki, Y and Levy, AY, 1993, "Automated model selection for simulation" In: *Proceedings of the Seventh International Workshop on Qualitative Reasoning about Physical Systems* 108–116, Seattle, WA.

Karnopp, DC, Margolis, DL and Rosenberg, RC, 1990, *System Dynamics: a Unified Approach* John Wiley & Sons, second revised edition.

Kuipers, B, 1986, "Qualitative simulation" *Artificial Intelligence* **29** 289–388.

Kuipers, B, 1994, *Qualitative Reasoning, Modelling and Simulation with incomplete Knowledge* MIT Press.

Langley, P, Zytkow, JM, Simon, HA and Bradshaw, GL, 1986, "The search for regularity: Four aspects of scientific discovery" In: RS Michalski, JG Carbonell and TM Mitchell, editors, *Machine Learning: An Artificial Intelligence Approach, volume 2* 425–469, Morgan Kaufmann.

Levy, AY, Iwasaki, Y and Motoda, H, 1992, "Relevance reasoning to guide compositional modelling" In: *Proceedings of the Sixth International Workshop on Qualitative Reasoning about Physical Systems* 7–21, Edinburgh, Scotland, August.

Low, CM and Iwasaki, Y, 1992, "Device modelling environment: an interactive environment for modelling device behaviour" *Intelligent Systems Engineering* **1** (2) 115–145, Winter.

Michalski, RS, 1983, "A theory and methodology of inductive learning" In: RS Michalski, JG Carbonell and TM Mitchell, editors, *Machine Learning: an Artificial Intelligence Approach* Tioga.

Mozetič, I, 1987, "The role of abstractions in learning qualitative models" In: *Proceedings of the 4th International Workshop on Machine Learning* 242–255, Irvine, CA.

Mozetič, I, Bratko, I and Urbančič, T, 1990, "Varying levels of abstraction in qualitative modelling" In: JE Hayes, D Michie and E Tyugu, editors, *Machine Intelligence 12* 259–280, Clarendon Press.

Muggleton, SH and Feng, C, 1990, "Efficient induction of logic programs" In: *Proceedings of the First Conference on Algorithmic Learning Theory* Tokyo, Japan.

Nayak, PP, 1992, "Causal approximations" In: *Proceedings of the Tenth National Conference on Artificial Intelligence AAAI-'92* 703–709, San Jose, CA, July.

Nayak, PP, Joskowicz, L and Addanki, S, 1992, "Automated model selection using context-dependent behaviors" In: *Proceedings of the Tenth National Conference on Artificial Intelligence AAAI-'92* 710–716, San Jose, CA, July.

Nordhausen, B and Langley, P, 1993, "An integrated framework for empirical discovery" *Machine Learning* **12** 17–47.

Paynter, HM, 1961, *Analysis and Design of Engineering Systems* MIT Press.

Richards, BL, Kraan, I and Kuipers, BJ, 1992, "Automatic abduction of qualitative models" In: *Proceedings of the Tenth National Conference on Artificial Intelligence AAAI-'92* 723–728, San Jose, CA, July.

Rickel, J and Porter, B, 1992, "Automated modeling for answering prediction questions: Exploiting interaction paths" In: *Proceedings of the Sixth International Workshop on Qualitative Reasoning about Physical Systems* 82–95, Edinburgh, Scotland, August.

Rickel, J and Porter, B, 1994, "Automated modeling for answering prediction questions: Selecting the time scale and system boundary" In: *Proc. of AAAI-94* 1191–1198, July.

Rothenberg, J, 1989, "The nature of modelling" In: LE Widman, KA, Loparo and NR Nielson, editors, *Artificial Intelligence, Simulation, and Modeling* 75–92, John Wiley & Sons.

Schut, C and Bredeweg, B, 1993, "Automatic enhancement of model parsimony" In: *Proceedings of the 7th International Workshop on Qualitative Reasoning* 194–203, Seattle, WA, May.

Schut, C and Bredeweg, B, 1994, "Supporting qualitative model specification" In: *Proc. of the Second International Conference on Intelligent Systems Engineering* 37–42, Technical University of Hamburg-Harburg, Germany.

Schut, C and Bredeweg, B, 1995, "Supporting qualitative model construction: Eliminating incorrectly predicted derivatives" In: *Proceedings of the 9th International Workshop on Qualitative Reasoning* 163–172, Amsterdam, The Netherlands, May.

Shen, Q, Leitch, RR and Steele, AD, 1994, "A generic harness for the systematic generation of multiple models" In: *Proc. of the Second International Conference on Intelligent Systems Engineering* 49–55. Technical University of Hamburg-Harburg, Germany.

Top, J, 1993, "Conceptual Modelling of Physical Systems" PhD thesis, Twente University, September.

Varšek, A, 1991, "Qualitative model evolution" In: *Proc. of the 12th IJCAI* 1311–1316, Sydney, Australia.

Weld, DS, 1992, "Reasoning about model accuracy" *Artificial Intelligence* **56** 255–300.

Weld, DS and deKleer, J, 1990, *Readings in Qualitative Reasoning about Physical Systems* Morgan Kaufmann.

Xia, S, Linkens, DA and Bennett, S, 1993, "Automatic modelling and analysis of dynamic physical systems using qualitative reasoning and bond graphs" *Intelligent Systems Engineering* **2** (3) 201–212, Autumn.

Zeigler, BP, 1976, *Theory of Modelling and Simulation* Wiley-Interscience.