Supporting Qualitative Model Construction: Eliminating incorrectly predicted derivatives*

Cis Schut Bert Bredeweg
Department of Social Science Informatics (S.W.I.)
University of Amsterdam
Roetersstraat 15
1018 WB Amsterdam (The Netherlands)
Telephone: +31-20-525 6765, Telefax: +31-20-525 6896
E-mail: {schut,bert}@swi.psy.uva.nl

Abstract: *Constructing a qualitative model of some device usually proceeds as a cycle of model formulation and model debugging. The latter is driven by discrepancies between the behaviour predicted by the model and the actual device behaviour. This paper describes how the elimination of one type of discrepancy, incorrectly predicted derivatives, can be supported. It provides an analysis of the knowledge that is required for succesfully adapting the model. Then an implemented procedure is described for supporting the elimination of incorrect derivatives. Heuristics are employed to generate plausible hypotheses, and in interaction with the modeler one model adaptation is selected.*

1 Introduction

A qualitative model is hardly ever constructed in one go. Usually the modeling process proceeds incrementally. It consists of cycles of model simulation and model debugging. Figure 1 depicts the cyclic nature of model construction. An initial model is formulated

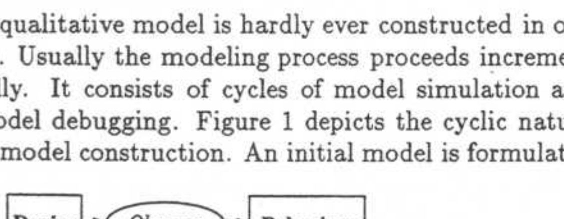


Figure 1: The incremental nature of qualitative model construction

and the qualitative simulator generates a behaviour prediction. Usually the behaviour predicted by the model is not completely in line with the real device behaviour. For example, the model predicts that a certain temperature increases, whereas in the real device it decreases. The modeler is faced with the task of finding out which modeling error(s) caused this incorrect prediction, and to change the model in order to resolve the discrepancy. We call the process of eliminating discrepancies between predicted behaviour and expected behaviour, *model adaptation*. Our aim is to support a modeler in this process. The paper focuses on one specific type of discrepancy: *incorrect derivatives* of quantities. We assume that a modeler identifies one or more derivatives that are incorrect. Then we describe how the support system finds, in interaction with the modeler, the model modifications that eliminate the incorrect derivatives. The contributions of the paper are threefold:

- It describes the modeling knowledge that is required for eliminating incorrect derivatives
- It gives an analysis of the reasoning process in eliminating incorrect derivatives
- It demonstrates how this process can be supported

First, section 2 briefly explains the knowledge representation that is employed in the qualitative reasoner that we use. Then, section 3 gives an example of the model adaptation problem and discusses the model adaptation task in a general way. Section 4 discusses the specific case of adapting a model in order to eliminate a discrepant derivative. Section 5 explains how the occurrence of multiple discrepant derivatives is treated. Finally, section 6 discusses some open-ends and relates model adaptation to model-based diagnosis.

2 Knowledge Representation

The process of constructing a qualitative model for some device is, among others, determined by the particular qualitative reasoning formalism. The model

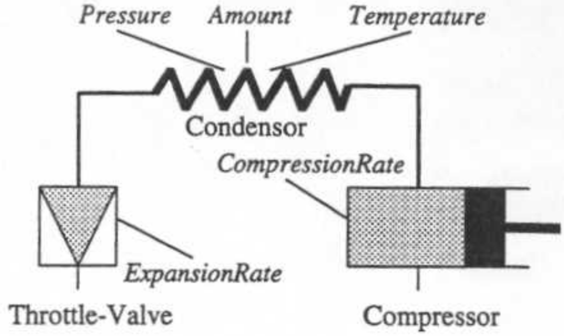
*The research presented in this paper was partially supported by the Foundation for Computer Science in the Netherlands (SION) with financial support from the Netherlands Organisation for Scientific Research (NWO), project number 612-322-307. We thank Bob Wielinga, Jaap Kamps, and Kees de Koning for their insightful comments.

primitives of the formalism dictate what can be expressed and how. The formalism we use, GARP (see [1]), employs *model fragments* to represent pieces of generic domain knowledge, and a *case model* for representing the particular device. Model fragments are expressed as a set of conditions and consequences. Both the case model and the model fragments are made up of the following set of model primitives: *entities* represent physical objects (e.g. a container); *attributes* represent properties of objects (e.g. connected_to); *quantities* represent variable properties of objects (e.g. temperature) that have a qualitative *value* and a *derivative*; *dependencies* represent how different quantities are qualitatively related (e.g. an inequality).

In order to predict the behaviour of some device, its structure and initial conditions have to be modeled in a case model. The qualitative simulator searches the library of model fragments and instantiates the ones whose conditions match the case model. This results in a set of instantiated dependencies that are used to compute the values and derivatives of all quantities, yielding a description of the device behaviour during a period of time: a qualitative state. Changes from one state to another are derived by applying transformation rules. A sequence of states represents the device behaviour over time.

3 Model Adaptation

First, we will discuss the model adaptation task in a general sense in order to illustrate which notions are important. The next section describes in detail how model adaptation proceeds in the case of incorrectly predicted derivatives. As an example of the model adaptation task, consider figure 2. The drawing represents the condensor part of a refrigerator. A compressor pumps substance into the condensor, and the substance flows out through the throttle-valve. Due to the increased amount of substance in the condensor, the pressure rises, and as a consequence, the temperature of the substance rises too. Now suppose that the modeler has constructed an incorrect model for which the behaviour is predicted. The table in figure 2 gives the expected and the predicted derivatives of *Pressure* and *Temperature* for a state in which the compressor starts working. *Pressure* is predicted to be decreasing, and *Temperature* is predicted to be steady, whereas they both should be increasing. The modeler's task is to identify what is wrong with the model, and how it has to be modified. The first step is to find out which elements in the model affect the derivatives of *Pressure* and *Temperature*. These *determinants* are depicted in the left side of figure 3. Next, the modeler has to determine how the set of determinants is incorrect and which modifications to the model might resolve the incorrect prediction. For example, the derivative of the *Amount* could be incorrect. If it were decreasing,



	Predicted Derivatives:	Expected Derivatives:
Pressure	*decreasing*	*increasing*
Temperature	*steady*	*increasing*

Figure 2: The condensor with predicted and expected derivatives

the derivative of *Pressure* would be correctly predicted to increase as well. Likewise, adding a positive influence of *CompressionRate* on the *Temperature* would resolve the incorrectly predicted derivative of the latter. This shows that a selection has to be made from the various ways of modifying the model. The right side of figure 3 shows the model that should be attained. From all possible candidate model adaptations the modeler has to select the one that realizes three things: it must change the negative proportionality between the *Amount* and the *Pressure* into a positive one, it must add a positive proportionality from *Pressure* to *Temperature*, and it must remove the influence of the *ExpansionRate* on *Pressure*.

This example illustrates four notions that are important for model adaptation in general:

Discrepancies. In general, discrepancies are characterized by the model primitive they concern, and by their type. Three types of discrepancies can be distinguished: i) inappropriate: a model primitive is instantiated in the model where it should not have been (e.g. the influence of *ExpansionRate* on *Pressure*); ii) incorrect: the primitive should indeed be instantiated but have another value (e.g. *Pressure* being negative proportional to *Amount* should be positive proportional); iii) missing, it is not instantiated where it should have been (e.g. the proportionality between *Pressure* and *Temperature*).

Current model state. The predicted behaviour provides the instantiated current model state. It allows the modeler to identify which elements in

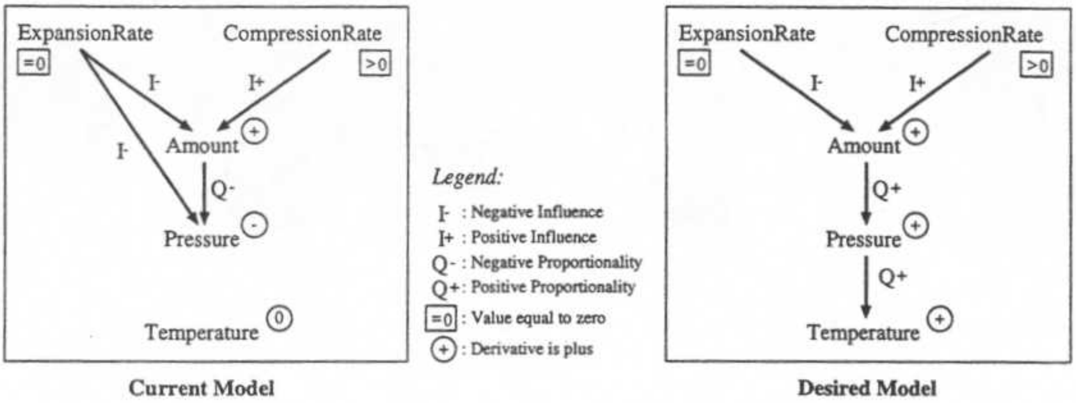


Figure 3: Current Model State and Desired Model State

the model were responsible for predicting a particular primitive instance. In the condensor example the determinants of the derivative of *Pressure* are: the influence of *ExpansionRate*, and the proportionality with *Amount*.

Modeling knowledge. Modeling knowledge can be thought of as a meta-view on the reasoning mechanism of the qualitative reasoner. For example, it comprises knowledge about the effect of adding a certain primitive to the model, knowledge of the reasons that might explain why a certain expected primitive is not in the model, and so on. The relevant modeling knowledge differs for each combination of primitive type and discrepancy type. The next section provides an account of the modeling knowledge that is used for eliminating incorrect derivatives.

Model modifications. A model modification is a set of one or more alterations in the model (including additions), that together eliminate the discrepancy. A model modification is the result of the model adaptation process.

The process of model adaptation consists of four main reasoning steps: *discrepancy detection, determinant collection, hypothesis generation*, and *hypothesis selection*. In the discrepancy detection step the differences between predicted and expected behaviour are identified. Determinant collection is the process of identifying which elements in the model are responsible for determining the discrepant primitive. It takes the discrepancy and the current model state as input and generates the relevant determinants as output. Hypothesis generation is the process of finding the candidate model modifications (hypothesis) that might be applied in order to resolve the discrepancy. The set of determinants and the modeling knowledge are used to come up with such candidate model modifications. Finally, one model modification has to be selected from the alternatives that were generated.

Note that modeling errors may propagate. For example, an incorrect derivative might be caused by an incorrect value of another quantity, which might be caused by a missing causal dependency, which finally can be traced to an omission in a model fragment. This illustrates that the cause of some discrepancy might be a discrepancy itself, possibly in another type of primitive. As a consequence, another discrepancy is discovered that requires model adaptation. This can go further until finally a discrepancy can be related to a change in a model fragment or the case model. Since the modeling knowledge used in model adaptation differs for each type of primitive, the model adaptation task can be decomposed naturally along the lines of the primitives involved in the discrepancies. Each combination of primitive type and discrepancy type can be dealt with as an independent part of the model adaptation task. The rest of the paper discusses one such combination: incorrect derivatives.

4 Eliminating Incorrect Derivatives

After discussing the model adaptation task in a more general way, we can now turn to the specific task of adapting the model in order to eliminate incorrect derivatives.

4.1 Detecting incorrect derivatives

It is assumed that the modeler is capable of assessing the appropriateness of the model's behaviour prediction. Hence, the modeler identifies the discrepancies between predicted and expected behaviour.[1] Discrepancies are stated in terms of the derivative value that is incorrect, and the derivative value that is expected. A predicted derivative value is one of: $+, 0$, or $-$. An expected derivative value is also one of $+, 0$ or $-$, but three ranges may be specified as well: $[-,0], [0,+]$, or $[-,0,+]$. If a range is indicated it means that the modeler expects different states in which the only difference is the derivative value of the particular quan-

[1] This step could be largely automated if an accurate description of expected device behaviour were available.

tity. In the condensor example the two discrepancies can be expressed as: discrepancy(incorrect, derivative(Pressure), +) and discrepancy(incorrect, derivative(Temperature), +), expressing that the type of discrepancy is incorrect, the model primitive involved is the derivative of Pressure and Temperature respectively, and the expected derivative is in both cases +. In this example two discrepancies are identified. There is a possibility that both discrepancies are caused by the same modeling error. Therefore, a distinction has to be made between situations in which only one discrepant derivative is identified, and situations where several are identified. This section discusses the case of a single discrepant derivative and the next section adresses the case of multiple discrepant derivatives.

4.2 Collecting determinants of derivatives

When the discrepant derivatives have been identified, the next step is to collect those elements in the model that have contributed in determining its value. These elements can be collected from the current model state, the set of instantiated model primitives in the state that exhibits the incorrect derivative. The model elements that have an effect on a quantity's derivative are the dependencies that hold for that quantity. However, not all types dependencies have an effect on the derivative of a quantity. Only three types of dependencies are used to compute derivatives:

Causal dependencies. The common way in which derivatives are determined is by one or more causal dependencies with other quantities. The causal dependencies occurring in the formalism we use, are influences, I+, I-, and proportionalities, Q+, Q- (see also [4]).

Constraints. If various causal dependencies interact then the relative effects cannot be determined from the dependencies themselves. Sometimes there is knowledge available about such relative effects. Constraints are the vehicle for modeling knowledge about relative effects. In addition, they can be used for pruning behaviours that are not of interest to the modeler. We use five types of constraints: $>, \geq, =, \leq, <$, each of them defined with respect to two quantities.

Assignments. If only certain behaviours of a device are of interest, c.q. only behaviours where some quantity has a fixed derivative, an assignment can be used. Also if it is exogenously determined that a quantity has one particular derivative throughout the device behaviour then it is assigned this value. Quantities can be assigned one particular derivative, $+, 0$, or $-$, or they can be assigned a range: $[0,+]$, $[0,-]$, resulting in the prediction of different states with different derivatives for the particular quantity. Assignments are defined as inequalities of derivatives with respect to zero: $>, \geq, =, \leq, <$.

The effect of a particular type of dependency is not only determined by the type of that dependency but also by the value or derivative of the quantity that exerts an effect through that dependency. Therefore, all determinants (causal dependencies, constraints, and assignments) that are present in the current model state, are annotated with the effect they have on the particular quantity's derivative. This is possible since the values and derivatives of the quantities involved in a dependency, are also available in the current model state.

Annotating a determinant with its effect differs for each type of determinant. For an assignment it is determined which derivatives it allows and consequently which are excluded. For example, if quantity A is assigned derivative +, then $-$ and 0 are excluded. Constraints are annotated in a similar way: the type of constraint and the derivative of the constraining quantity determine which derivatives are allowed and which are excluded. For example, suppose there holds an equality between the derivatives of A and B and the derivative of B is +. This means that for the derivative of A only + is allowed and hence $-$ and 0 are excluded. Causal dependencies are annotated in a slightly different way because their effects are qualitatively added. They are annoted according to the *direction of influence* (Doi) they have. The direction of influence is determined on the basis of the type of causal dependency and the value or derivative of the determining quantity. For example, if A is positively proportional to B and B has derivative +, the Doi of the dependency is +. Table 1 shows how the direction of influence for each causal dependency is determined. In the condensor example the determinants of *Pressure* are:

- $Q-(Pressure, Amount)$ with $Doi = -$ because $\delta Amount = +$
- $I-(Pressure, ExpansionRate)$ with $Doi = 0$ because $ExpansionRate = 0$

4.3 Generating candidate model modifications

Having collected all determinants and their effects, the next step is to generate candidate model modifications that resolve the discrepancy. The idea is that each of the identified determinants may involve a modeling error, or propagates the effect of another modeling error (the determinant may also be correct, of course). Besides the possibility of errors in the model, it may also be incomplete: determinants may be missing. Which determinants may be missing and how each individual determinant may be modified, constitutes one type of the modeling knowledge mentioned in section 3.

Dependency Type:	I-	I+		Q-	Q+
value V of determining quantity	*Doi*	*Doi*	derivative D of determining quantity	*Doi*	*Doi*
$V < 0$	+	−	$D = -$	+	−
$V = 0$	0	0	$D = 0$	0	0
$V > 0$	−	+	$D = +$	−	+

Table 1: Determining the *Direction of Influence (Doi)*

There is a limited number of ways in which a determinant can involve a modeling error, or in which it propagates one. As a consequence, the number of ways in which each determinant can be modified, is limited as well. This makes it possible to enumerate in advance all different modifications for each type of determinant. In addition, each modification can be annotated with the effect that it has on the quantity involved, similar to how determinants are annotated.

Modifications are expressed implicitly as the inverse of a discrepancy of another model primitive. For example, if a modification entails that a positive influence is replaced by a negative influence, this is expressed as: $\mathsf{discrepancy}(\mathsf{incorrect}(I+(A,B)), I-(A,B))$. The reason for expressing a modification in this way, is that, as already explained in section 3, modeling errors propagate. Therefore, one discrepancy may be explained by another discrepancy until one can be explained by an error in the case model or in the model fragments. For example, replacing a positive by a negative influence may resolve an incorrectly predicted derivative but introduces the problem of why the positive influence was present in the model: some model fragment may have been defined incorrectly, or it should not have been included in the model whereas another (specifying a negative influence) should have been included, and so on. For our purposes, we assume that the new problem is adressed by another part of the model adaptation task.

As an example of a set of determinant modifications, take the determinant $Q-$(*Pressure, Amount*) from the condensor system. This determinant may be incorrect or inappropriate. The incorrectness may concern the derivative of the *Amount* or the dependency itself. Suppose the annotation of the determinant, its Doi, is + in the current model state. Then the possible modifications are:

- incorrect $\delta Amount$, should be − (new Doi = −)
- incorrect $\delta Amount$, should be 0 (new Doi = 0)
- incorrect $Q-$(*Amount, Pressure*), should be $Q+$ (new Doi = +)
- incorrect $Q-$(*Amount, Pressure*), should be $I+$ (new Doi depends on value of *Amount*)
- incorrect $Q-$(*Amount, Pressure*), should be $I-$ (new Doi depends on value of *Amount*)
- inappropriate $Q-$(*Amount, Pressure*), (giving no Doi)
- correct $Q-$(*Amount, Pressure*), (Doi remains −)

As another example, suppose that one of the determinants of a quantity A is the constraint that the derivatives of A and B are equal, and also suppose that the derivative of B is 0. In this case the errors are:

- incorrect δB, should be −, realizes −
- incorrect δB, should be +, realizes +
- incorrect $\delta A = \delta B$, should be $\geq$, realizes $[0,+]$
- incorrect $\delta A = \delta B$, should be $>$, realizes +
- incorrect $\delta A = \delta B$, should be $<$, realizes −
- incorrect $\delta A = \delta B$, should be $\leq$, realizes $[-,0]$
- inappropriate $\delta A = \delta B$, realizes no derivative
- correct $\delta A = \delta B$, realizes 0

In this fashion it is possible to enumerate all different ways in which each determinant may be modified, annotated with the effect that the modification has. Recall that annotations are for causal dependencies in terms of the direction of influence, for constraints and assignments they are in terms of the derivative they realize. For constraints and assignments the sets of possible modifications can now be pruned: some modifications can be discarded already, since they are not compatible with the expected derivative.[2] Because the modifications are annotated it can be checked whether the annotation excludes the expected derivative. If it does, that particular alternative is no longer considered. An example may clarify this. Suppose an assignment is present in the current model state, realizing derivative +. Possible modifications are: incorrect assignment (annotation $[0,+]$); incorrect assignment (0); incorrect assignment ($[-,0]$); incorrect assignment (−); correct assignment (+). Suppose that the expected derivative is −. Now the first two modifications and the last one are not compatible with

[2]The reason why this can only be done for constraints and assignments, is that they determine a derivative in an absolute way, unlike causal dependencies whose effects are qualitatively added.

the expected derivative: they need not be considered in modifying the model because they can impossibly lead to a model realizing the expected derivative.

Having for each determinant a set of possible modifications, hypotheses are generated by picking for each determinant one of its possible modifications. Such a set of alternatives forms a potential hypothesis. Since all modifications for each determinant are enumerated in advance, eventually a combination is formed that is the desired model adaptation. Here it becomes clear why correctness of a determinant is included as one of the possible "modifications" of that determinant. Explicitly including correctness of some determinant makes it possible to generate hypotheses where, for example, only one determinant is modified.

Besides picking one alternative for each determinant, it is also possible to extend hypotheses by including one or more missing determinants in the hypothesis. Thirteen types of annotated determinants might be missing: a causal dependency with a Doi of either $-, 0$, or $+$; a constraint or an assignment realizing $+, [0,+], -, [-,0]$, or $-$. Based on the same argument as above, the alternatives for missing constraints or assignments that exclude the expected derivative, are not considered.

A set of modifications of determinants, possibly extended with errors of missing determinants, is only a hypothesis if the conjunction of modifications leads to a model that correctly predicts the expected derivative. How can this be checked? First, recall that each modification has been annotated with its effect on the quantity involved. Thus the set of modifications (and additions) gives a complete picture of the situation where a model including these modifications would be simulated. Second, an additional type of modeling knowledge can be employed. This knowledge captures the fact that a certain derivative is only predicted if particular combinations of determinants are present in a model. These combinations can be represented as so-called *ideal model states*. An ideal model state specifies which annotated determinants have to be present, and which must not be present in the model so that the model predicts a certain derivative. For example, one way to realize an expected derivative of $+$, is to have two causal dependencies, one with Doi $= -$, and one with Doi $= +$, and an assignment realizing $+$: the effects of the causal dependencies are ambiguous but the assignment excludes derivatives $-$ and 0. This type of knowledge is based on the semantics of the dependencies and on the manner in which the simulator uses them to compute derivatives.

It is possible to enumerate all ideal model states that realize each particular derivative or range of derivatives. Table 2 gives all ideal model states for some expected derivatives.[3] The upper row designates the expected derivative. The first two columns specify the types of determinants that may occur in the model and their effects. The following columns of the table represent each an ideal model state for some expected derivative. The entries of a column specify which annotated determinants must be present, marked $\exists$, which must not be present, marked $\emptyset$, and which may or may not be present in that particular ideal model state, not marked. For example, take in table 2 the second column for expected derivative 0 (see $\uparrow$). Now start reading from above: derivative 0 is realized by a model that has no causal dependencies with $Doi = +$, that has a caisal dependency with $Doi = 0$, that has no causal dependencies with $Doi = -$, that has no constraints and no asignments realizing $+$ or $-$.

Now the set of modifications with their effects can be checked against the ideal model states. A set of modifications is a hypothesis only if it matches one of the ideal model states for the expected derivative. If not, it can be rejected as a hypothesis. In the condensor example, a hypothesis for the incorrect derivative of *Pressure* is that the derivative of the *Amount* is incorrect (should be 0), and that a causal dependency with Doi $= +$ is missing. Another hypothesis is that the proportionality with *Amount* is incorrect (should be positive). Note that both hypotheses contain a discrepancy concerning the same determinant, that is, $Q-$(*Pressure*, *Amount*). However, the former concerns the determining derivative, $\delta Amount$, the latter concerns the dependency.

Selecting a model modification

Obviously, this way of generating hypotheses leads to large numbers of hypotheses, many of which are implausible. The number of hypotheses can be reduced by asking the modeler to verify whether some modification of a determinant is accepted or not. If the modeler accepts it, then for that determinant the proper modification has been found, ruling out many combinations that include the other, invalid modifications. If the modeler does not accept that modification, it is marked as an invalid modification, which still rules out many candidate hypotheses. But for effective support the interaction with the modeler should be minimized. In order to do so, the plausible hypotheses should be generated first. This calls for the formulation of a set of heuristics that operationalize plausibility of an error. The effect of applying a heuristic is that some of the ways in which each determinant may be modified, are initially not considered. Only if no solution is found, then the heuristic is discarded and the previously skipped modifications are considered. By building several of these filters on top of each other, the most plausible hypotheses are generated before hypotheses that are less plausible.

[3] For purposes of conciseness, the ideal model states for expected derivative $-$ and $[-,0]$ are left out of the figure. They correspond to the mirrored situations of expected derivative $+$ and $[0,+]$ respectively.

		Ideal Model States for Expected Derivative														
		+				0								[0, +]		[−, 0, +]
Determinant	*Annotation*															
causal dependencies	+	∅	∃	∃	∃	∅	∅	∃	∃	∃	∃	∃	∃	∃	∃	∃
	0	∅				∅	∃									
	−	∅	∅	∃	∃	∅	∅	∃	∃	∃	∃	∃	∃	∃	∃	∃
constraints	+	∅		∃		∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅
	[0, +]	∅				∅				∃	∃			∃		∅
	0	∅	∅	∅	∅	∅		∃						∅	∅	∅
	[−, 0]	∅	∅	∅	∅	∅				∃			∃	∅	∅	∅
	−	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅
assignments	+	∃			∃	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅
	[0, +]											∃	∃		∃	∅
	0	∅	∅	∅	∅	∃			∃					∅	∅	∅
	[−, 0]	∅	∅	∅	∅						∃	∃		∅	∅	∅
	−	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅	∅
							↑									

Table 2: Ideal Model States

The first heuristic is based on the observation that modeling errors may propagate: a quantity earlier in the causal structure may have an incorrect value or derivative. Therefore, many hypotheses are excluded if a determiner can be rendered incorrect. This gives rise to the following heuristic:

- **Prefer modifications of determining values and derivatives over others.**

A major presupposition that can serve as a basis for some heuristics, is that the modeler is not constructing a model in a randomly fashion. This means that it is safe to assume that the larger part of the model is correct. As a consequence, hypotheses involving few modifications are more plausible than hypotheses involving many modifications. These considerations are captured in the heuristic:

- **Prefer simpler hypotheses over more complex hypotheses.**

The same presupposition allows us to order the hypothesis types according to the probability of their occurrence. Recall that hypotheses may specify that something is missing, inappropriate, or incorrect. It is unlikely that some determinant is first specified and later on appears to be inappropriate. Similarly, it is more probable that some specified determinant appears to be incorrect than that the determinant is completely forgotten. Then again, forgetting something is more likely than specifying it inappropriately. With respect to modifications of type incorrect a further refinement can be made. The most common way of specifying some dependency incorrect is the so-called sign fault, e.g. specifying a positive proportionality instead of a negative one. Therefore, within the category of incorrect modifications, sign faults should be preferred over more radical errors. This gives the following heuristics:

- **Prefer modifications of type incorrect over modifications of type missing and inappropriate.**
- **Prefer modifications concerning incorrect sign over modifications of other incorrectness.**
- **Prefer modifications of type missing over modifications of type inappropriate.**

As a second guideline in formulating heuristics, the different types of determinants can be ranked according to the probability that they are involved in modeling errors. This ranking is based on the different aspects that are represented by the dependencies. Causal dependencies, for instance, represent general features of the domain knowledge. Constraints, however, usually represent features of a device, or of a number of physical objects in a particular configuration. Assignments are used to exclude behaviours that are not of interest. It can now be argued that a modeler is less prone to make errors in the domain knowledge than in device-specific knowledge. Also it is more likely that mistakes were made in trying to exclude some specific behaviour than that device-specific aspects were incorrectly modeled. Hence, the following heuristics are used:

- **Prefer modifications involving assignments over modifications involving causal dependencies or constraints.**
- **Prefer modifications involving constraints over modifications involving causal dependencies.**

In the condensor example, the model adaptation process proceeds as follows. First, hypotheses are generated where only one change is involved (prefer simpler hypotheses). In addition, the set of heuristics determines that only errors of type incorrect and concerning determining values and derivatives are considered. This gives: {incorrect $\delta Amount$, Doi = +} as the only hypothesis. The modeler is asked to verify whether $\delta Amount$ is incorrect. Since that is not the case, the set of modifications for the determinant $Q-(Amount, Pressure)$ reduced: this modification of the derivative of *Amount* is ruled out. New hypotheses have to be generated after relaxing the set of heuristics: now incorrect assignments are considered too. No hypothesis can be found so again the heuristics are relaxed so that now also incorrect constraints are considered. This still gives no hypothesis. Further relaxation makes that incorrect causal dependencies are considered, and even more specific only sign-faults. This gives the hypothesis: {incorrect $Q-(Amount, Pressure)$, should be $Q+(Amount, Pressure)$, Doi = +}. Asking the modeler reveals that this is the right model adaptation. In a similar fashion it is discovered that for the *Temperature* a causal dependency with effect sf plus is missing.

This procedure does not discover (yet) that the other causal dependency, from *ExpansionRate* to *Pressure* is inappropriate. The reason is that in this state of the predicted behaviour, the influence is not harmful. In other behaviour states, however, the harmful effects of this dependency may become apparent, and as a consequence its inappropriateness is only discovered later.

5 Eliminating multiple discrepant derivatives

If we assume that one modeling error causes multiple discrepancies, then the error must have been propagated through a branching path. Therefore, identifying such an error if multiple discrepancies are observed, has to proceed the other way around. Because an error can only be propagated along the paths of causally related quantities[4], the network of causal dependencies can be used to identify candidates that might explain different discrepancies. More specific, we use the notion of *causal units* (see also [2]) to find such candidates. Causal units are based on the idea that the set of quantities can be divided into clusters that influence each other, but are independent of other quantities. A causal unit starts with one or more quantities that influence a quantity ($I+$ or $I-$), and is further comprised of proportionally related quantities. It ends with a quantity that has no causal effect on any other quantity by means of a proportionality. A causal unit is essentially a graph that may have recursive loops, more than one starting point and more than one terminal node (note that a causal unit can consist of one or more causal paths). Causal units can be generated from the predicted behaviour. In the condensor example only one causal unit exists: [*ExpansionRate*, *CompressionRate*] $\rightarrow$ *Amount* $\rightarrow$ *Pressure*.

One modeling error may only propagate to errors in the derivative of several quantities, if they occur in the same causal unit. Therefore, the first step is to find a causal unit in which all discrepant quantities occur. If such a causal unit is found then it is checked whether there is one *discrepant* quantity that precedes all other discrepant quantities. If there is one, then it will probably have caused the discrepancies of the other quantities. This is depicted in figure 4 in case 1: quantity C may have caused quantities G and K to be discrepant. Therefore it is selected and a model adaptation is sought in the fashion as described in section 4. Even if no single *discrepant* quantity precedes all

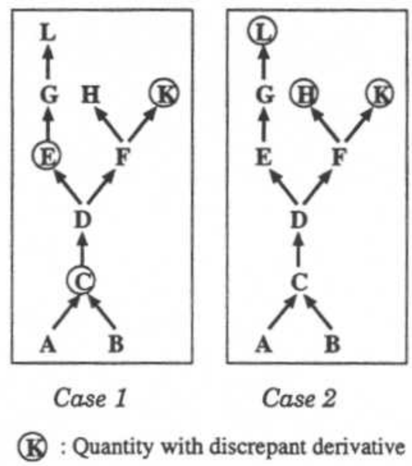


Figure 4: Causal unit with multiple incorrect derivatives

others, there might still be a single modeling error. If so, it must concern the quantities in the causal unit that precede *all* discrepant quantities (recall that the modeler might not have indicated all discrepancies). This stuation is depicted in the right part of figure 4 as case 2. The error might concern quantities A, B, C and D. The modeler is asked to verify the correctness of the derivative of the first quantity(ies) preceding all discrepant quantities (D). If the derivative is incorrect, it becomes input for model adaptation. If it is correct, however, then multiple errors must exist and as a consequence, incorrect derivative elimination must be repeatedly performed. Assuming now that two errors exist, the same process is repeated with the set of discrepant quantities divided in two independent groups. Because quantity L is the only one

[4] Although the knowledge representation formalism that we use, allows enforcing derivatives in a model *without* using causal dependencies (namely by constraints) doing so is considered improper modeling since it violates the causal view on how dynamic behaviour comes about.

in the first group, it is input for the first model adaptation procedure. In the other group are quantities H and K that both might be explained by an error in F. If the derivative of F is incorrect, it is input for the second model adaptation procedure. If it is correct, the process is repeated and there appear to be three modeling errors.

So far the case in which one causal unit is found in which all discrepant quantities appear. If no such causal unit can be found, then different causal units are sought that cover as many discrepant quantities as possible. The same process as above is repeated for each causal unit.

In the condensor example there is only one causal unit and it only contains *Pressure*. Therefore, model adaptations for *Pressure* and *Temperature* are selected consecutively.

6 Discussion and Conclusions

In this paper we have shown how to support the process of eliminating incorrectly predicted derivatives by adapting the model on which the prediction was based. It is assumed that the modeler indicates one or more incorrect derivatives, and the derivative(s) that are expected. The elements in the model that affect a derivative are identified and annotated with their effect. Hypotheses, sets of modifications and additions of determinants, are generated that eliminate the incorrect derivative *and* realize the expected derivative. They are based on explicit knowledge of how each determinant may be wrong, and on knowledge of which combinations of determinants realize a particular derivative. Hypotheses are generated in order of plausibility by using a set of heuristics. These candidates for model adaptation are presented to the modeler who decides whether the hypothesis is accepted or not. In this interaction eventually one model adaptation is selected.

The support system has been implemented and tested on a number of deliberately corrupted models. In all cases it came up with model adaptations that would realize the expected derivative. These model modifications were not always the inverse of the original model corruption. Sometimes a particular determinant was not modified or added. The reason is that, as we also saw in the condensor example, some adaptations are not necessary for realizing a particular derivative in a particular state. Their effects may become harmful in other states where another dependency structure may hold. From our tests it appeared that the procedure tends to become slow when six or more determinants are involved. In the models we are working with, however, the number of determinants seldomly exceeds four.

At first glance, model adaptation seems to consist of a *diagnosis* task that determines how the discrepancies have arisen, and a *repair* task, that resolves the underlying modeling error. Since diagnosis has been studied extensively in AI ([5]), the obvious approach would be to apply existing diagnosis techniques to model adaptation. However, closer examination immediately reveals that this is not feasible. The reason is that in device diagnosis the assumption is made that device structure is available and modeled correctly. The task amounts to identifying one or more components that function incorrectly. In model adaptation the correct model structure is not known, in fact the goal of model adaptation is to contribute to the formulation of a correct model. That means that often, a missing "component", or an inappropriate "component" explains the symptom. An undesirable effect is that many more hypotheses can be generated. This difference in starting points also blurs the distinction between a diagnosis and a repair. In device diagnosis one or more components are identified as faulty and a repair is simply its replacement. This implies that the fault has to be identified in one of the components that are present. Having identified the component(s) the repair is immediately known: replace that component. In contrast with this, in model adaptation the fault may lie in the fact that some component (e.g. a causal dependency), many times several components, are missing or are erroneous. Therefore, it is often not possible to mark one or several determinants as the cause of the discrepancy. However, viewed from the perspective of model construction, this is not a problem. Apparently the subtasks of diagnosis and repair are interwoven and executed in parallel. A diagnosis is determined by considering modifications for existing determinants *and* considering inappropriate and missing "components". Our paper demonstrates that model adaptation can effectively be supported. Although potentially the number of possible model adaptations can become very large, especially when there are many determinants of a derivative, we have shown how to deal with this. A set of heuristics is used for generating hypotheses in a step-wise fashion: the most plausible hypotheses are generated before less plausible hypotheses.

Other related work lies in the area of qualitative modeling. The prevailing research in this area is based on compositional modeling [3]. An important assumption in this approach is that *all* relevant domain knowledge has been represented in model fragments. Different viewpoints on the same phenomenon are represented in different model fragments. Here the modeling problem is to select the right set of model fragments. This is solved by incorporating the assumptions that underlie the different viewpoints in the model fragments and use these as handles in the selection process. Our starting points are complementary to compositional modeling: our focus is on the situation where the domain knowledge has not been modeled completely, i.e. where the modeler still has to define a set of model fragments. Later, when a particular domain has been modeled substantially, the techniques for selecting model fragments become applicable. Our aim is to provide a modeler with a set of tools for defining the domain knowledge in terms of model fragments. If a modeler has such a set of tools at his disposal then the whole domain modeling trajectory is supported. Some of our work in this area can be found in [2]. It describes a technique for eliminating incorrect derivatives as well, but applies if additional assumptions on the state of the model, are valid. In [6] we describe how the specification of the domain and device knowledge in terms of our qualitative reasoning formalism, is supported.

References

[1] B. Bredeweg. *Expertise in Qualitative Prediction of Behaviour.* PhD thesis, University of Amsterdam, 1992.

[2] B. Bredeweg and C. Schut. Reducing ambiguity by learning assembly behaviour. In *Proc. of the 13th IJ-CAI'93*, pages 980–985. Chambéry, France, August, 1993.

[3] B. Falkenhainer and K. D. Forbus. Compositional modeling: finding the right model for the job. *Artificial Intelligence*, 51:95–143, 1991.

[4] K. D. Forbus. Qualitative process theory. *Artificial Intelligence*, 24:85–168, 1984.

[5] W. Hamscher, L. Console, and J. de Kleer. *Readings in Model-based Diagnosis.* Morgan Kaufmann, San Mateo, CA, 1992.

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