

## Chapter 9

---

# Defaults in Specifications:

## Distance Functions Between Temporal Models

SOFIA GUERRA

### ABSTRACT.

Default institutions were defined in [Schobbens,1993] as an extension of the concept of institution [GoguenBurstall,1992], where there is a notion of distance between interpretations. In this extension, we can have default modules that may be overridden by more specific properties. The semantics of a specification with an exception are given by the models of the exception that are as close as possible to the models of the specification according to the given notion of distance.

In this paper we present a formalisation of non-monotonicity in temporal logic by defining a temporal instantiation of a weak form of default institution. In this way, using temporal logic as a specification language, we get a way of handling defaults in specifications of dynamic systems.

## 1 Introduction

Default reasoning concerns the logic of assertions which should be assumed to be true but for which there might be exceptions. Although it first appeared as a field within Artificial Intelligence, the need to be able to express default assertions in specifications of software systems is well established. For example, using defaults we can specify *fault-tolerant systems*, exhibiting a normal behaviour to which exceptions arise when a fault develops. In this case, an exception should be dealt with within the specification, giving rise to appropriate corrective actions.

Another application of non-monotonicity in specifications is the ability to override constructs in specifications. Since the development of a specification often starts with a simplified model that is enriched and modified, the understanding of a system is not only obtained by syntactically enriching a simple model, but also by modifying or contradicting part of it. The formalisation of these modifications that contradict the initial specification is

one of the applications of defaults in specifications. In this paper we describe an example where a specification is enriched by adding a feature that contradicts the initial specification.

There are several different approaches to default reasoning, including model preference logics. This branch of approaches is based on the semantic notion of choosing a preferred set of models of a theory. The main idea is to consider instead of arbitrary models of a given axiom set, only models that satisfy a certain minimality condition. This minimality condition is related to the way models satisfy the defaults. One way of doing this is by defining an ordering between models, where interpretations of a logical language are ordered according to how well they satisfy some given default information; this framework is called *preference relations* [Kraus et al.,1990, Makinson,1994].

Another natural way of comparing interpretations is by the use of distance functions: if  $\mathcal{M}$  is a set of interpretations and  $m, n \in \mathcal{M}$  are interpretations, then  $d(m, n)$  is some notion of distance between  $m$  and  $n$ . In [Schobbens,1993], this is generalised by considering morphisms between interpretations and an ordering between morphisms. These morphisms correspond to several ways of comparing two interpretations; distances are the particular case of a single morphism between any two interpretations. The semantics are based on this notion of ordering and then selecting the models of the axioms that are as close as possible to the models of the defaults, according to the given definition of distance between models.

In order to handle defaults in specifications of dynamic systems, we define an instantiation with temporal logic of a weaker version of the default institution proposed in [Schobbens,1993]. In this paper we briefly describe the notion of default institution and then why we need a weaker version of the default institution for the temporal case, and the temporal instantiation. This instantiation is motivated by an example of a specification where defaults are used.

## 2 The default institution

Default institutions were proposed in [Schobbens,1993] as a way of extending the notion of institution [GoguenBurstall,1992]. Institutions provide a way of talking about an arbitrary logical system and provide a way of structuring specifications. However, with this structuring an existing specification can be enriched but not modified. Default institutions allow partial reuse of existing specifications, where the ‘default’ module can be modified by introducing exceptions to suit the application at hand.

The concept of default institution is an extension of institutions by adding a

notion of distance between interpretations. The semantics of a specification with an exception are given by the models of the exception that are as close as possible to the models of the specification according to the given notion of distance. This is non-monotonic since by adding exceptions it might be needed to retract some of the previous conclusions.

In general, we want to compare interpretations that may be different in nature. Therefore, we need a way to relate elements of different nature that play a similar role. This is done by morphisms between interpretations, the morphisms of the category  $Int(\Sigma)$  (here and in the rest of the paper  $\Sigma$  is a signature). In order to compare morphisms and give a precise meaning to ‘closest’, a pre-order  $\leq_{\Sigma}$  among morphisms is needed. These define the comparison category  $Comp(\Sigma)$ : the objects are the morphisms in the category  $Int(\Sigma)$  of interpretations, and the morphisms are defined by the pre-order  $\leq_{\Sigma}$ . This means that there is one and only one morphism between  $h : m \rightarrow n$  and  $h' : m' \rightarrow n'$  iff  $h \leq_{\Sigma} h'$ .

When instantiating a default institution with a logic, the choices we make are the morphisms between interpretations and the pre-order on these morphisms, i.e. we have to decide what are the components of the category  $Comp(\Sigma)$ . The other components of the default institution are the usual ones, the same as in the corresponding institution. This is exactly what we do in the next section for the temporal case. The formal definition of a default institution is the following:

**2.1 Definition** [Schobbens,1993] *A default institution consists of*

- a category *Sign* of signatures;
- a functor  $Sen : Sign \rightarrow Set$ , giving languages linked by translations;
- a contravariant functor  $Int : Sign \rightarrow Cat^{op}$ , giving interpretations and their morphisms; the class of interpretations should not be empty;
- a family of satisfaction relations  $\Vdash_{\Sigma}$  between the interpretations of  $\Sigma$  and its formulae ( $\Vdash_{\Sigma} \subseteq |Int(\Sigma)| \times Sen(\Sigma)$ );
- a functor  $Comp : Sign \rightarrow Cat^{op}$ , such that
  - $Comp(\Sigma)$  is a pre-order;
  - the objects of  $Comp(\Sigma)$  are the morphisms of  $Int(\Sigma)$ ;
  - the identities of  $Int(\Sigma)$  are initial (minima) in  $Comp(\Sigma)$ ;
  - the morphisms of  $Int(\Sigma)$  that are minima in  $Comp(\Sigma)$  are called agreements; they must form a subcategory  $Int_0(\Sigma)$ , that is, the composition of agreements must be an agreement;
  - the institution where  $Int$  is replaced by  $Int_0$  is weakly abstract;

- 0-symmetry: each agreement  $h : M \rightarrow N$  has a reverse agreement  $h^R : N \rightarrow M$  such that  $(h^R)^R \mathfrak{3} Dh$ ;
- 0-equivalence: for any morphism  $h : B \rightarrow C$  and agreements  $a : A \rightarrow B, c : C \rightarrow D, a; h \equiv h \equiv c$ .

Let us explain some of the motivations for this definition. If  $m, n, m',$  and  $n'$  are interpretations in  $Int(\Sigma)$  for a given signature  $\Sigma$ , and  $h : m \rightarrow n$  and  $h' : m' \rightarrow n'$  are morphisms in the same category  $Int(\Sigma)$ ,  $h : m \rightarrow n \leq_{\Sigma} h' : m' \rightarrow n'$  represents that  $m$  is ‘closer’ to  $n$  (according to  $h$ ) than  $m'$  to  $n'$  (according to  $h'$ );  $\leq_{\Sigma}$  being the pre-order between morphisms for a given signature  $\Sigma$ . The minimal morphisms for each of these orderings are called *agreements*. We want agreements to behave similarly: the fact that a morphism  $h : m \rightarrow n$  is minimal represents that  $m$  is as close to  $n$  as possible, so it is natural to expect that they should be similar. To guarantee that interpretations linked by a minimal morphism (an agreement) behave similarly we impose that the institution  $Int_0$  (where we only consider morphisms between interpretations that are minimal) is weakly abstract. Intuitively, an institution is weakly abstract if our logic does not allow us to look at more details of the interpretations than the morphisms do. Formally, we say that an institution is *weakly abstract* iff given two interpretations  $m$  and  $n$  in  $Int(\Sigma)$ ,  $Cl(\Sigma, m) \subseteq Cl(\Sigma, n)$  whenever there is a morphism  $h : m \rightarrow n$  in  $Int(\Sigma)$ , where  $Cl(\Sigma, m)$  is the set of its properties, i.e.  $\{\varphi \in Sen(\Sigma) | m \Vdash \varphi\}$ . By the condition of 0-symmetry, we have that if there is an agreement from  $m$  to  $n$ , then there is also one from  $n$  to  $m$ . Therefore, these two conditions together (0-symmetry and weak abstractness) imply that two models linked by an agreement satisfy exactly the same formulae.

The 0-equivalence is the requirement that agreements should be transparent with respect to comparisons; here  $h \equiv h'$  means  $(h \leq h' \text{ and } h' \leq h)$ .

The semantics of a default  $D$  and an exception  $E$ , written  $D \mathbf{but} E$ , are defined using the pre-order of the comparison category  $Comp(\Sigma)$ . Let us introduce some notation that is used in the formal definition:  $Mor(E, D)$  is the class of morphisms whose domain satisfy  $E$  and whose codomain satisfy  $D$ ;  $Min(E, D)$  is the class of minimal morphisms of  $Mor(E, D)$ . An interpretation  $m$  is a model of  $D \mathbf{but} E$  if  $m$  is the domain of a minimal morphism between the morphisms whose domain satisfies  $E$  and whose codomain satisfies  $D$ . This means that  $m$  is a model of the exceptions  $E$  that is as close as possible to a model of the defaults  $D$ . Note that a model of  $D \mathbf{but} E$  always satisfies the exception  $E$ .

**2.2 Definition (Semantics of but)**  $m \Vdash D \mathbf{but} E$  iff there is a morphism  $h \in Min(E, D)$  such that  $dom(h) \mathfrak{3} Dm$ .

### 3 Temporal default institution

The temporal default institution is an instantiation of the definition of default institution with propositional linear temporal logic. So, in the temporal case the default institution consists of:

- **Signatures** (Category  $Sign$ ): Sets of propositional symbols.
- **Language** (Category  $Sen(\Sigma)$ ): Formulae constructed in the usual way using the boolean connectives and the temporal operators  $U$  (until) and  $X$  (next); from these temporal operators it is possible to define  $G$  (always) and  $F$  (eventually).
- **Interpretations** (Category  $Int(\Sigma)$ ): The interpretations are triples of the form  $\mathcal{M}3D(\mathbb{N}, <, V)$ , where  $<$  is the usual ordering on the natural numbers and  $V$  is a valuation. The valuation assigns to every timepoint  $t \in \mathbb{N}$  a set of symbols  $V(t) \subseteq \Sigma$ , namely the set of proposition letters that are true at the instant  $t$ . Note that in these interpretations, time is discrete, has an initial moment with no predecessors, and is infinite into the future.
- **Satisfaction Relation** (Relations  $\Vdash_{\Sigma}$ ): The usual for linear temporal logic. We write  $m \Vdash_t A$  if the interpretation  $m$  satisfies the formula  $A$  at the timepoint  $t$ .

We use the anchored version of temporal logic, where a formula  $A$  is said to be *valid* on the interpretation  $m$  iff  $m \Vdash_0 A$ , and we consider the reflexive versions of the modalities. These interpretations are the objects of the category  $Int(\Sigma)$ .

The problem now is how to define morphisms between interpretations and a pre-order on the morphisms in such a way that the defaults in specifications of dynamic systems behave in the expected way. To motivate these definitions let us look at an example.

#### 3.1 An example: a Ferris Wheel

Consider a Ferris Wheel where each chair can be in one of six positions, as illustrated in figure 9.1. We choose one of the chairs and from now on we always talk about the same chair. The chair can be in one of three levels 1,2 or 3, and it moves always in the same direction (clockwise). The name of each position indicates the level of the chair and whether it is going up or down: eg. **3up** indicates that it is at level 3 and it is going up; **2down** indicates that it is at level 2 and it is going down. Following this convention, there are

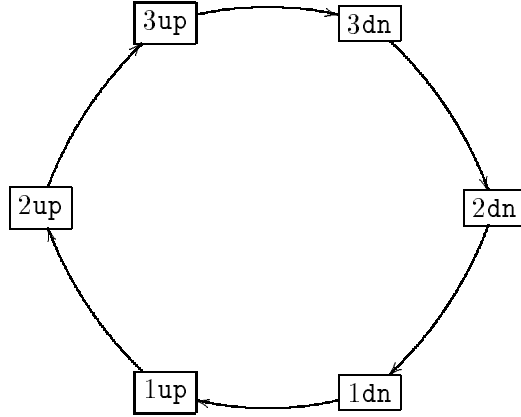


Figure 9.1: A Ferris Wheel.

six different positions where the chair can be: **1up**, **2up**, **3up**, **3down**, **2down** and **1down**. The chair goes from **1up** to **2up**, and from this to **3up**, then it goes to **3down** and it starts going down to **2down** and **1down** and then to **1up** again, and it never stops. A specification of this Ferris Wheel can be seen in figure 9.2. The first formula states the fact that the chair is always in one of the six positions and it is not in more than one at each moment. The other formulae describe the movement of the chair.

$$\begin{array}{ll}
 3D \ G((1up \vee 2up \vee 3up \vee 3down \vee 2down \vee 1down) \wedge & \\
 \neg(1up \wedge 2up) \wedge \neg(1up \wedge 3up) \wedge \neg(1up \wedge 3down) \wedge \dots) & \\
 G(1up \rightarrow X2up) & 3D \ G(2up \rightarrow X3up) \\
 G(3up \rightarrow X3down) & G(3down \rightarrow X2down) \\
 G(2down \rightarrow X1down) & G(1down \rightarrow X1up)
 \end{array}$$

Figure 9.2: A specification of the Ferris Wheel.

Suppose now that we want to add an emergency **lever** that makes the chair reach level 1 as soon as possible. If the chair was in one of the **down** positions, then the faster way of getting to level 1 is by keep going clockwise. However, if the chair is in one of the **up** positions than the best thing is to reverse the direction, and move counter clockwise. Suppose that the **lever** stays pushed until the chair gets to level 1, it stops there and then is released. The action of pushing the lever is call **push** and **lever** represents that the lever is pushed. Figure 9.3 represents what we want to add to the specification.

Instead of rewriting the specification from the beginning, and considering

```

3D G(push→lever)
G((lever ∧ 2up)→X(1up ∧ lever))
G(((lever ∧ 3up))→X(2up ∧ lever))
G(((lever ∧ 1up))→X(1up ∧ ¬lever))
G(((lever ∧ 1down))→ = X(1down ∧ ¬lever))
G(((3down ∨ 2down) ∧ lever)→Xlever)

```

Figure 9.3: Adding an emergency lever.

explicitly whether the `lever` has been pushed or not, we want to add these formulae as exceptions to the specification. If  $D$  is the conjunction of the formulae in the specification of the Ferris Wheel and  $E$  the conjunction of the formulae that describe the `lever`, then the models of the Ferris Wheel with the emergency lever would be the models of the state constraint that satisfy  $D$  **but**  $E$ . However, in order to give semantics to **but** we have to know what are the morphisms between interpretations and how they are ordered, i.e. the comparison category  $Comp(\Sigma)$ . This is our aim now.

### 3.2 The comparison category

When defining the comparison category, the first thing to note is that the state constraints have to be rigid to avoid the satisfaction of the exception by allowing the chair to be in more than one position at the same time. These formulae should not be overridden and we see them as axioms that cannot be overridden. Alternatively, we could see them as defaults with the highest priority. The details of how to do that are not presented here. We can suppose that we only consider for minimisation the models of these axioms. In our example, the first formula of figure 9.2.

If we think about how to order the morphism, a first possibility that follows the ideas of the instantiations defined in [Schobbens,1993] would be the following: if  $h : m \rightarrow n$  and  $h' : m' \rightarrow n'$  are morphisms between temporal interpretations in  $Int(\Sigma)$  for a signature  $\Sigma$ ,  $h : m \rightarrow n \leq_{\Sigma} h' : m' \rightarrow n'$  iff for all  $t \in \mathbb{N}$

$$m \text{ disagrees with } n \text{ at } t \quad \Rightarrow \quad m' \text{ disagrees with } n' \text{ at } t,$$

where by *disagree* we mean that they do not satisfy the same propositional symbols at that instant. This definition has two immediate problems. Firstly, if  $m$  disagrees with  $n$  at some point  $t$  then what happens afterwards

might be meaningless. As an example consider the Ferris Wheel; let  $m$  be a model of the exceptions (the formulae that describe the **lever**) and  $n$  a model of the initial specification. Suppose that at some timepoint  $t$ ,  $m$  was at **2up**, the **lever** was not pushed and, as a pass of magic, in  $t + 1$  it was at **2down**. If  $n$  agreed with  $m$  at  $t$  (i.e.  $m$  was at **2up** as well) then obviously they disagree at  $t + 1$ . Moreover, even if  $m$  behaves as expected after  $t$  it will disagree with  $n$ , and these disagreements are meaningless because they are the result of what happened at  $t$ . This suggests that what we want to do is to see if at some point  $t$ ,  $m$  were the model  $n$ , whether it would behave in the same way. Following this idea, morphisms between interpretations are monotonic functions  $h : \mathbb{N} \rightarrow \mathbb{N}$  such that

$$\forall t \in \mathbb{N}. \forall p \in \Sigma. m \Vdash_t p \Leftrightarrow n \Vdash_{h(t)} p,$$

and we check whether  $m \Vdash_{t+1} p \Leftrightarrow n \Vdash_{h(t)+1} p$ .

Another problem with this definition of the pre-order is that it blocks the occurrence of the actions that change the behaviour, like the action of pushing the **lever**. With this definition, the minimal models of the Ferris Wheel with the **lever** would be the ones where the **lever** is never pushed. This is highly undesirable, since we want to be able to push the **lever** whenever we feel like (whenever there is an emergency), and we do not want to restrict the times we push the **lever** in the minimisation process. This suggests that we only want to compare interpretations that have the same occurrences of actions at the same timepoints. In this way we have to treat actions differently from the attributes in the minimisation process. Hence, we split the signature in two disjoint sets: one for the actions  $Act$ , that cannot be minimised, and another  $Att$  for the attributes. A signature  $\Sigma$  is then a pair of disjoint sets  $(Act, Att)$ , where  $Act$  is the set of actions and  $Att$  is the set of attributes. We also impose that comparable interpretations agree at the first instant, i.e. we suppose that they satisfy the same symbols at the instant 0.

We can now say what are the morphisms between interpretations. Firstly we introduce some notation:

Let  $h : \mathbb{N} \dashrightarrow \mathbb{N}$  be a partial function and  $t \in \mathbb{N}$  be a natural number:

- $h(t) \uparrow$  means that  $h$  is undefined at  $t$ ;
- $h(t) \downarrow$  means that  $h$  is defined at  $t$ ;
- $\bar{h}(t) = \max\{t' \in \mathbb{N} : t' \leq t \text{ and } h(t') \downarrow\}$  is the greatest timepoint less or equal than  $t$  in which the function is defined.

The idea of the morphisms is to relate two temporal interpretations  $m$  and  $n$  that have the same initial state (they satisfy exactly the same atomic

formulae at 0), in a way that, for each instant  $t_1$ , a timepoint  $t_2$  is chosen in a monotonic way, such that the interpretation domain  $m$  at  $t_1$  satisfies exactly the same atomic symbols as the interpretation codomain  $n$  at  $t_2$ . This choice is partial to allow cases where such  $t_2$  does not exist. These conditions are formally expressed by the following definition:

**3.3 Definition** Let  $\Sigma \in \text{Sign}$  be a signature and  $m$  and  $n$  be two temporal interpretations in  $|\text{Int}(\Sigma)|$ . A morphism  $h : m \rightarrow n \in \text{Mor}(\text{Int}(\Sigma))$  is a partial function  $h : \mathbb{N} \dashrightarrow \mathbb{N}$  such that:

- $h(0) \geq 0$ ;
- for all  $t_1, t_2 \in \mathbb{N}$ , if  $t_1 < t_2$ ,  $h(t_1) \downarrow$  and  $h(t_2) \downarrow$  then  $h(t_1) < h(t_2)$ ;
- for all  $t \in \mathbb{N}$ , if  $h(t) \downarrow$  then for all  $p \in \Sigma$ ,  $m \Vdash_t p$  iff  $n \Vdash_{h(t)} p$ .

Now that we have defined morphisms between temporal interpretations, we can go back to think about the notion of ‘closeness’ in order to get the category  $\text{Comp}(\Sigma)$ . Although we cannot define the pre-order yet, taking into account the previous considerations we now know that  $h : m \rightarrow n \leq_{\Sigma} h' : m' \rightarrow n'$  implies:

1.  $\forall p \in \Sigma$ .  $m \Vdash_0 p$  iff  $m' \Vdash_0 p$ ; 09
2.  $\forall a \in \text{Act}$ .  $\forall t \in \mathbb{N}$ .  $m \Vdash_t a$  iff  $m' \Vdash_t a$ ;

Suppose  $h : m \rightarrow n$  is a morphism between two temporal interpretations. If  $m$  was ‘near’  $n$ , it would behave, at every timepoint, in the same way as  $n$  if  $n$  was at a similar state. Discrepancies are the cases when this does not happen.

**3.4 Definition** Let  $\Sigma \in \text{Sign}$  be a signature,  $h : m \rightarrow n$  be a morphism in  $\text{Int}(\Sigma)$  and  $t \in \mathbb{N}$  a timepoint. We say that there is a discrepancy at time  $t$  between  $m$  and  $n$  according to  $h$ , written  $m \overset{h}{\not\leftrightarrow}_t n$ , if  $h(t)$  is defined and there is an attribute  $o \in \text{Att}$  such that

$$(m \Vdash_{t+1} o \text{ and } n \not\Vdash_{h(t)+1} o) \quad \text{or} \quad (m \not\Vdash_{t+1} o \text{ and } n \Vdash_{h(t)+1} o)$$

This notion of discrepancy expresses the fact that, if the model  $n$  was at the state that  $m$  is at some instant  $t$ , it would behave in a different way. This depends on the morphisms  $h$ , that chooses from  $n$  an instant to compare with  $m$  at the instant  $t$ : by definition of morphism, if  $h : m \rightarrow n$  is a morphism, then  $m$  satisfies at  $t$  exactly the same symbols as  $n$  at  $h(t)$ . There is a discrepancy at  $t$  if they progress in different ways: at the following instants,  $t+1$  for  $m$  and  $h(t)+1$  for  $n$ , they satisfy different attributes. This notion of

discrepancy is going to be used to define the pre-order: two interpretations are ‘nearer’ if they have less discrepancies.

If  $h : m \rightarrow n$  and  $h' : m' \rightarrow n'$  are morphisms, we want to know if  $m$  is closer to  $n$  (according to  $h$ ) than  $m'$  to  $n'$  (according to  $h'$ ). The main idea is to minimise the discrepancies. We could say that if  $h : m \rightarrow n \leq_{\Sigma} h' : m' \rightarrow n'$  and there is a discrepancy at  $t$  between  $m$  and  $n$  then there is also one between  $m'$  and  $n'$  at the same instant:

$$\forall t \in \mathbb{N}. m \overset{h}{\leftrightarrow}_t n \Rightarrow m' \overset{h'}{\leftrightarrow}_t n'.$$

The problem with this definition is that if  $m'$  did something ‘*strange*’ before an instant  $t$ , but  $m$  did not disagree with  $n$  at that point, then we could not conclude the relation we wanted. Thinking again about our example, suppose  $h : m \rightarrow n$  and  $h' : m' \rightarrow n'$  are morphisms,  $m$  and  $m'$  satisfy **lever** at the same instants and they agree at 0. Suppose  $t$  is a timepoint where  $m$  is at **2up** and the **lever** is pushed at  $t$ , and this is the first time the **lever** is pushed. We can also suppose that before  $t$ ,  $m$  satisfies all the formulae of the specification of the Ferris Wheel without the **lever**. In these conditions there is a discrepancy between  $m$  and  $n$  at  $t$ . However, if  $m'$  is at a **down** point at  $t$  (because it did not follow the default formulae at some previous instant) then there is no discrepancy  $t$  between  $m'$  and  $n'$ . This is obviously undesirable. To solve this we impose then that if  $h : m \rightarrow n \leq_{\Sigma} h' : m' \rightarrow n'$  and if there is a discrepancy between  $m$  and  $n$  at  $t$  that does not exist between  $m'$  and  $n'$  then there must be a reason for this: either because  $h'$  at  $t$  is undefined or because at a previous instant there was a discrepancy in  $h'$  that did not exist in  $h$ . This gives rise to the following condition that we add to the previous two we already had.  $h : m \rightarrow n \leq_{\Sigma} h' : m' \rightarrow n'$  implies:

1.  $\forall p \in \Sigma. m \Vdash_0 p$  iff  $m' \Vdash_0 p$ ; 09
2.  $\forall a \in Act. \forall t \in \mathbb{N}. m \Vdash_t a$  iff  $m' \Vdash_t a$ ;
3.  $\forall t_1 \in \mathbb{N}. ((m \overset{h}{\leftrightarrow}_{t_1} n \wedge m' \not\overset{h'}{\leftrightarrow}_{t_1} n' \wedge h'(t_1) \downarrow) \Rightarrow (\exists t_2 < t_1. (m \not\overset{h}{\leftrightarrow}_{t_2} n \wedge h(t_2) \downarrow \wedge (m' \overset{h'}{\leftrightarrow}_{t_2} n' \vee h'(t_2) \uparrow))))$

The minimisation of the discrepancies is expressed by condition 3 of the definition 3.5: if, at some instant, there is a discrepancy in a morphism  $h$ , and there is not a discrepancy at  $h'$ , then  $h \leq h'$  only if there was a previous discrepancy at  $h'$  that did not exist at  $h$ . This is like chronological minimisation, where defaults in earlier instants have higher priority than the ones that occur later. It also ensures that if  $h \leq h'$  and there is a

discrepancy in  $h$  that does not exist in  $h'$ , the absence of the discrepancy at  $h'$  is possible only because there was an earlier one that allows the domain of  $h'$  to have a behaviour closer to a model of the defaults.

To define the pre-order there are further two conditions that we want to impose. We want minimal morphisms as defined as possible, which is expressed by condition 4 of definition 3.5: a completely undefined function does not have any discrepancy.

The last condition ensures that the morphisms are as surjective as possible; this has as a result that identities are minimal, and that any minimal morphism is surjective, satisfying one of the conditions of the definition of default institution, namely that the institution where the category  $Int$  is replaced by the category  $Int_0$  is weakly abstract. If we allow non surjective morphisms to be minimal, then we could have two interpretations linked by an agreement that would not satisfy the same formulae. In the points of the codomain interpretation that were not mapped by any point of the domain, the interpretation could satisfy different propositional symbols that would result in the two interpretations not satisfying the same formulae. The formal definition of the pre-order is the following:

**3.5 Definition** *Let  $\Sigma 3D(Act, Att) \in Sign$  be a signature and let  $h : m \rightarrow n$  and  $h' : m' \rightarrow n'$  be morphisms in  $Int(\Sigma)$ . We say that  $h : m \rightarrow n \leq h' : m' \rightarrow n'$  iff:*

1.  $\forall p \in \Sigma. m \Vdash_0 p$  iff  $m' \Vdash_0 p$ ; 09
  2.  $\forall a \in Act. \forall t \in \mathbb{N}. m \Vdash_t a$  iff  $m' \Vdash_t a$ ;
  3.  $\forall t_1 \in \mathbb{N}. ((m \xleftrightarrow{h}_{t_1} n \wedge m' \not\xleftrightarrow{h'}_{t_1} n' \wedge h'(t_1) \downarrow) \Rightarrow (\exists t_2 < t_1. (m \not\xleftrightarrow{h}_{t_2} n \wedge h(t_2) \downarrow \wedge (m' \xleftrightarrow{h'}_{t_2} n' \vee h'(t_2) \uparrow))))$
- 09
4.  $\{t : h(t) \uparrow\} \subseteq \{t : h'(t) \uparrow\}$ ; 09
  5.  $(\forall t \in \mathbb{N}. m \xleftrightarrow{h}_t n$  iff  $m' \xleftrightarrow{h'}_t n')$   $\Rightarrow$   
 $\forall t \in \mathbb{N}. 3D((\exists t_1 \in \mathbb{N}. \bar{h}(t \Leftrightarrow 1) < t_1 < h(t) \wedge \forall p \in \Sigma. (m \Vdash_t p$  iff  $n \Vdash_{t_1} p))$   
 $\Rightarrow (\exists t_2 \in \mathbb{N}. \bar{h}'(t \Leftrightarrow 1) < t_2 < h'(t) \wedge \forall p \in \Sigma. (m' \Vdash_t p$  iff  $n' \Vdash_{t_1} p)))$ .

To sum up, we only compare morphisms if the domains have the same initial state and they satisfy the same actions at the same instants (conditions 1 and 2 of the definition 3.5). The condition 2 makes possible comparisons of interpretations only if the same actions occur at the same time points, so

that we choose, from the models that have the same occurrence of actions, the ones that are as close as possible to the models of the defaults. It is a way of avoiding blockage of the occurrence of actions: if the effect of an action contradicts the defaults, then a model where that action would not occur would be better than the ones where that action occurs. Moreover, we minimise discrepancies (condition 3) and we prefer morphisms ‘more defined’ and ‘more surjective’ (conditions 4 and 5 respectively).

This relation is in fact a pre-order and it defines the category  $Comp(\Sigma)$ : the objects of  $Comp(\Sigma)$  are the morphisms of  $Int(\Sigma)$  and there is a morphism between  $h : m \rightarrow n$  and  $h' : m' \rightarrow n'$  iff  $h : m \rightarrow n \leq_{\Sigma} h' : m' \rightarrow n'$ . This pre-order is used to give semantics of specifications with exceptions: if  $D$  is a specification and  $E$  an exception (formulae over a signature  $\Sigma$ ),  $m$  is a model of  $D$  but  $E$  if  $m$  is the domain of a  $\leq_{\Sigma}$ -minimal morphism of  $Mor(E, D)$ .

Going back to our example, let  $\Sigma 3D\{Act, Att\}$  be the signature with  $Act 3D\{\text{push}\}$  and  $Att 3D\{\text{3down}, \text{2down}, \text{1down}, \text{1up}, \text{2up}, \text{3up}, \text{lever}\}$ . We have that, if  $m$  is a model of the state constraint, and  $m \Vdash D$  but  $E$ , then  $m$  is one of the desired models: if the chair is at **3down**, **2down**, or at **3up** and **2up** but the emergency **lever** has not been pushed, then it behaves as before adding the **lever**; if the chair is at **1down** or **1up** and the **lever** has been pushed it stops; if it is at **2up** or **3up** and the **lever** has been pushed, then it changes direction, as specified in the formulae we added.

Although this notions of morphisms between temporal interpretations and of pre-order on these morphisms give the expected results when dealing with specifications of dynamic systems with defaults, it does not define a default institution. In the definition 2.1 of default institution given in [Schobbens,1993] the author impose that all the minimal morphisms are equivalent, i.e. all the minimal morphisms are minima. Following our discussion, in order to get intuitive results in our framework, we do not want to compare interpretation that differ on the occurrences of actions. In this way, if  $m$  and  $m'$  are temporal interpretations over the same signature  $\Sigma 3D(Act, Att)$ , if they do not agree on the actions  $Act$ , the identities on  $m$  and  $m'$  are incomparable. With this motivation we define a weaker default institution, where we impose that all identities are *minimal* but not necessarily *minima*. This modification of the definition is not discussed in this paper, but most of the results that hold for default institutions also hold in the weaker version.

## 4 Concluding remarks

Starting from the default institution in [Schobbens,1993], we define a weaker version of it where not all minimal morphisms are equivalent. A temporal

instantiation of this weaker version is defined: a propositional linear temporal default institution. This consists of the usual temporal institution, but where the notion of morphism between interpretations is changed, and it is extended with a pre-order between these morphisms. This notion of ordering between morphisms of interpretations is used in order to deal with non-monotonicity in specifications. In particular, an example where a specification is reused but where exceptions to the previous behaviour are added is described. In this example the first specification is seen as a default and a new module is considered, avoiding the necessity of writing the whole specification from the beginning.

## Acknowledgements

Thanks to Mark Ryan and to Pierre-Yves Schobbens for useful discussions, and to Maria João Coutinho and Darryl Davis for reading previous drafts of this paper. The author acknowledges financial support from the European Union through Esprit WG FIREworks (23531) and from PRAXIS XXI in Portugal.

## Bibliography

- [GoguenBurstall,1992] Joseph A. Goguen and Rod M. Burstall. Institutions: Abstract model theory for specification and programming. *Journal of the Association for Computing Machinery*, 39(1):95–146, January 1992.
- [Kraus et al.,1990] S. Kraus, D. Lehmann, and M. Magidor. Non-monotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence*, 44:167–207, 1990.
- [Makinson,1994] D. Makinson. General patterns in non-monotonic reasoning. In Dov Gabbay, C. J. Hogger, and J. A. Robinson, editors, *Nonmonotonic Reasoning and Uncertain Reasoning*, volume 3 of *Handbook of Logic in Artificial Intelligence and Logic Programming*, chapter 2, pages 35–110. Clarendon Press, Oxford, 1994.
- [Schobbens,1993] Pierre-Yves Schobbens, On the meaning of ‘but’. *Science of Computer Programming*, 20(1-2),1993