

# Change Management for Heterogeneous Development Graphs <sup>\*</sup>

Serge Autexier, Dieter Hutter, and Till Mossakowski

German Research Center for Artificial Intelligence (DFKI GmbH)  
Bremen, Germany

**Abstract.** The error-prone process of formal specification and verification of large systems requires an efficient, evolutionary formal development approach. Development graphs have been designed to support such an approach. They can formally represent the actual state of a software development comprising specification and verification work in a structured way and assist the user in her evolutionary development by the incorporated change management support. In this paper we extend this work with respect to heterogeneous development graphs allowing one to make use of different institutions, i.e. logics, for specifying and verifying large developments. We also push forward the idea of stringent locality of definitions by introducing pre-signatures and pre-signature morphisms, which allow us to build up signatures in an incremental and parametric way.

## 1 Introduction

Industrial applications of Formal Methods revealed that an efficient, evolutionary formal development approach is absolutely indispensable as it was hardly ever the case that the development steps were correctly designed in the first attempt. The search for formally correct software and the corresponding proofs is more like a formal reflection of partial developments rather than just a way to assure and prove more or less evident facts. Hence there is a need for computer-aided management of the many types of documents involved in the development of highly dependable systems and thus a need for an approach that keeps track of the various dependencies in and between such documents to provide efficient change management.

Development graphs are designed to encode an actual state of a software development in terms of a specification (represented by the nodes and the definition links of the graph) and of its verification (represented by theorem links and the attached proofs). Proof obligations, such as the requirement that a specification satisfies a security policy or that an implementation satisfies a requirement specification, can be represented as (global) theorem links connecting the corresponding theories.

Since software development is an incremental process, development graphs evolve over time. Changes in the specification are reflected by changes of the

---

<sup>\*</sup> Research supported by BMBF under research grant 01 IW 07002 *FormalSafe*, and DFG under research grant Hu737/3-1 *OMoC*.

nodes and the definition links while improvements in the verification work are reflected by additional theorem links and attached proofs to justify proposed requirements of the system. Specification and verification activities are intertwined resulting in the need for re-justification of existing proofs once the specification has changed.

In previous work we developed a management of change for formal specifications and proofs to save as much verification work as possible upon changing the specification [9, 2, 1]. That work exploited the hierarchical structure of logical dependencies between theories and established theory inclusion relationships, to devise a procedure computing non-interference properties of changes in logical theories and established relationships. In this paper we extend this work with respect to heterogeneous development graphs [13] allowing one to make use of different institutions, i.e. logics, for specifying and verifying large developments. In Section 2 we start with a brief overview on the notions of institutions and their (co-)morphisms. A guiding principle of designing development graphs is the stringent locality of specifications. In Section 3 we introduce pre-signatures and pre-signature morphisms allowing us to build up signatures in an incremental and parametric way and in Section 4 we extend this notion to (co-)morphisms between institutions. Based on this foundation we redefine the framework of heterogeneous development graphs in terms of pre-signatures in Section 5 and provide the proof rules for the verification in the large in Section 6. Section 7 is devoted to an analysis of how changes in the specification will affect the applicability of proof rules as well as their results in order to implement a smart replay of proofs in a changed environment.

## 2 Institutions and Their (Co)Morphisms

Institutions [7] formally capture the notion of logical systems, abstracting away from the details of vocabularies (signatures), sentences, models, and satisfaction (of sentences in models). When working with heterogeneous formal specifications, also mappings between institutions are important.

**Definition 1.** *An institution  $\mathcal{I}$  consists of:*

- a category  $\mathbf{Sign}_{\mathcal{I}}$  of signatures;
- a functor  $\mathbf{Sen}_{\mathcal{I}}: \mathbf{Sign}_{\mathcal{I}} \rightarrow \mathbf{Set}$ , giving a set  $\mathbf{Sen}(\Sigma)$  of  $\Sigma$ -sentences for each signature  $\Sigma \in |\mathbf{Sign}_{\mathcal{I}}|$ , and a function  $\mathbf{Sen}(\sigma): \mathbf{Sen}(\Sigma) \rightarrow \mathbf{Sen}(\Sigma')$  that yields  $\sigma$ -translation of  $\Sigma$ -sentences to  $\Sigma'$ -sentences for each signature morphism  $\sigma: \Sigma \rightarrow \Sigma'$ ;
- a functor  $\mathbf{Mod}_{\mathcal{I}}: \mathbf{Sign}_{\mathcal{I}}^{op} \rightarrow \mathbf{Set}$ , giving a set  $\mathbf{Mod}(\Sigma)$  of  $\Sigma$ -models for each signature  $\Sigma \in |\mathbf{Sign}_{\mathcal{I}}|$ , and a functor  $\mathbf{Mod}(\sigma): \mathbf{Mod}(\Sigma') \rightarrow \mathbf{Mod}(\Sigma)$ , denoted by  $\_|\sigma$ , that yields  $\sigma$ -reducts of  $\Sigma'$ -models for each signature morphism  $\sigma: \Sigma \rightarrow \Sigma'$ ; and
- for each  $\Sigma \in |\mathbf{Sign}_{\mathcal{I}}|$ , a satisfaction relation  $\models_{\mathcal{I}, \Sigma} \subseteq \mathbf{Mod}_{\mathcal{I}}(\Sigma) \times \mathbf{Sen}_{\mathcal{I}}(\Sigma)$

such that for any signature morphism  $\sigma: \Sigma \rightarrow \Sigma'$ ,  $\Sigma$ -sentence  $\varphi \in \mathbf{Sen}_{\mathcal{I}}(\Sigma)$  and  $\Sigma'$ -model  $M' \in \mathbf{Mod}_{\mathcal{I}}(\Sigma')$ :

$$M' \models_{\mathcal{I}, \Sigma'} \sigma(\varphi) \iff M' \upharpoonright_{\sigma} \models_{\mathcal{I}, \Sigma} \varphi \quad [\textit{Satisfaction condition}]$$

The next concept we need is a mapping between institutions. We concentrate here on *institution morphisms* [7] and on *institution comorphisms* (named so in [6]; see “plain maps of institutions” in [11] and “institution representations” in [20, 21]).

**Definition 2.** Let  $\mathcal{I}$  and  $\mathcal{I}'$  be institutions. An institution morphism  $\mu: \mathcal{I} \rightarrow \mathcal{I}'$  consists of:

- a functor  $\mu^{\text{Sign}}: \mathbf{Sign} \rightarrow \mathbf{Sign}'$ ;
- a natural transformation  $\mu^{\text{Sen}}: \mathbf{Sen}' \circ \mu^{\text{Sign}} \rightarrow \mathbf{Sen}$ , that is, a family of functions  $\mu_{\Sigma}^{\text{Sen}}: \mathbf{Sen}'(\mu^{\text{Sign}}(\Sigma)) \rightarrow \mathbf{Sen}(\Sigma)$ , natural in  $\Sigma \in |\mathbf{Sign}|$ ; and
- a natural transformation  $\mu^{\text{Mod}}: \mathbf{Mod} \rightarrow \mathbf{Mod}' \circ (\mu^{\text{Sign}})^{\text{op}}$ , that is, a family of functions  $\mu_{\Sigma}^{\text{Mod}}: \mathbf{Mod}(\Sigma) \rightarrow \mathbf{Mod}'(\mu^{\text{Sign}}(\Sigma))$ , natural in  $\Sigma \in |\mathbf{Sign}|$ ,

such that for any signature  $\Sigma \in |\mathbf{Sign}|$ , the translations  $\mu_{\Sigma}^{\text{Sen}}: \mathbf{Sen}'(\rho^{\text{Sign}}(\Sigma)) \rightarrow \mathbf{Sen}(\Sigma)$  of sentences and  $\mu_{\Sigma}^{\text{Mod}}: \mathbf{Mod}(\Sigma) \rightarrow \mathbf{Mod}'(\rho^{\text{Sign}}(\Sigma))$  of models preserve the satisfaction relation, i.e., for any  $\varphi' \in \mathbf{Sen}'(\mu^{\text{Sign}}(\Sigma))$  and  $M \in \mathbf{Mod}(\Sigma)$ :

$$M \models_{\Sigma} \mu_{\Sigma}^{\text{Sen}}(\varphi') \iff \mu_{\Sigma}^{\text{Mod}}(M) \models'_{\mu^{\text{Sign}}(\Sigma)} \varphi' \quad [\textit{Satisfaction condition}]$$

*Institution morphisms compose in the obvious, component-wise manner.*

An institution comorphism  $\rho: \mathcal{I} \rightarrow \mathcal{I}'$  consists of:

- a functor  $\rho^{\text{Sign}}: \mathbf{Sign} \rightarrow \mathbf{Sign}'$ ;
- a natural transformation  $\rho^{\text{Sen}}: \mathbf{Sen} \rightarrow \mathbf{Sen}' \circ \rho^{\text{Sign}}$ , that is, a family of functions  $\rho_{\Sigma}^{\text{Sen}}: \mathbf{Sen}(\Sigma) \rightarrow \mathbf{Sen}'(\rho^{\text{Sign}}(\Sigma))$ , natural in  $\Sigma \in |\mathbf{Sign}|$ ; and
- a natural transformation  $\rho^{\text{Mod}}: \mathbf{Mod}' \circ (\rho^{\text{Sign}})^{\text{op}} \rightarrow \mathbf{Mod}$ , that is, a family of functions  $\rho_{\Sigma}^{\text{Mod}}: \mathbf{Mod}'(\rho^{\text{Sign}}(\Sigma)) \rightarrow \mathbf{Mod}(\Sigma)$ , natural in  $\Sigma \in |\mathbf{Sign}|$ ,

such that for any  $\Sigma \in |\mathbf{Sign}|$ , the translations  $\rho_{\Sigma}^{\text{Sen}}: \mathbf{Sen}(\Sigma) \rightarrow \mathbf{Sen}'(\rho^{\text{Sign}}(\Sigma))$  of sentences and  $\rho_{\Sigma}^{\text{Mod}}: \mathbf{Mod}'(\rho^{\text{Sign}}(\Sigma)) \rightarrow \mathbf{Mod}(\Sigma)$  of models preserve the satisfaction relation, i.e., for any  $\varphi \in \mathbf{Sen}(\Sigma)$  and  $M' \in \mathbf{Mod}'(\rho^{\text{Sign}}(\Sigma))$ :

$$M' \models'_{\rho^{\text{Sign}}(\Sigma)} \rho_{\Sigma}^{\text{Sen}}(\varphi) \iff \rho_{\Sigma}^{\text{Mod}}(M') \models_{\Sigma} \varphi \quad [\textit{Satisfaction condition}]$$

*Institution comorphisms compose in the obvious, component-wise manner.*

Even though the only essential difference between institution morphisms and comorphisms is in the direction of sentence and model translations wrt. signature translation, the intuition they capture is quite different. Very informally, an institution morphism  $\mu: \mathcal{I} \rightarrow \mathcal{I}'$  shows how a “richer” institution  $\mathcal{I}$  is “projected” onto a “poorer” institution  $\mathcal{I}'$  (by removing some parts of signatures and models of  $\mathcal{I}$  to obtain the simpler signatures and models of  $\mathcal{I}'$ , and by embedding simpler  $\mathcal{I}'$ -sentences into more powerful  $\mathcal{I}$ -sentences). Then, an institution comorphism  $\mathcal{I} \rightarrow \mathcal{I}'$  shows how a “simpler” institution  $\mathcal{I}$  is represented in a “more complex” institution  $\mathcal{I}'$  (by representing the simpler signatures and sentences of

$\mathcal{I}$  as signatures and sentences of  $\mathcal{I}'$ , and extracting simpler  $\mathcal{I}$ -models from more complex  $\mathcal{I}'$ -models).<sup>1</sup>

**Definition 3.** *Institution comorphisms can induce institution morphisms via natural transformations: let  $\rho: \mathcal{I} \rightarrow \mathcal{I}'$  be an institution comorphism, let  $\mu^{\text{Sign}}: \mathbf{Sign}' \rightarrow \mathbf{Sign}$  be a functor and  $\varepsilon: \rho^{\text{Sign}} \circ \mu^{\text{Sign}} \rightarrow \text{id}_{\mathbf{Sign}'}$  a natural transformation. Then  $\rho$   $\varepsilon$ -induces the institution morphism  $\mu = \langle \mu^{\text{Sign}}, \mu^{\text{Sen}}, \mu^{\text{Mod}} \rangle: \mathcal{I}' \rightarrow \mathcal{I}$ , where for  $\Sigma' \in |\mathbf{Sign}'|$ ,  $\mu_{\Sigma'}^{\text{Sen}} = \mathbf{Sen}'(\varepsilon_{\Sigma'}) \circ \rho_{\mu^{\text{Sign}}(\Sigma')}^{\text{Sen}}$  and  $\mu_{\Sigma'}^{\text{Mod}} = \rho_{\mu^{\text{Sign}}(\Sigma')}^{\text{Mod}} \circ \mathbf{Mod}'(\varepsilon_{\Sigma'})$ . Given such an institution morphism  $\mu$ , we denote the corresponding institution comorphism  $\rho$  by  $\text{CoM}(\mu)$ .*

### 3 Institutions with Pre-Signatures

The abstractness of the signatures in the definition of institution (they are just a category) makes it difficult to develop a change management, which is based on manipulation of individual symbols that constitute a signature. Hence, we need to equip institutions with additional structure such that signatures behave more set-like. There are several approaches in this vein in the literature [19, 8, 12]. Most of these approaches formalize the fact that signatures are sets with some structure. Here we follow a radical approach by requiring that signatures in a sense *are* sets (more precisely, can be embedded into sets), such that from a set, we can go back to the corresponding signature in a unique way. This ensures that we can unite signatures, intersect them, etc. Of course, only *some* sets are signatures, while in general, such operations only deliver *pre-signatures*.

**Definition 4.** *An institution with pre-signatures is an institution equipped with an embedding  $|\cdot|: \mathbf{Sign} \rightarrow \mathbf{Set}$ , the symbol functor, and a map  $\text{sym}: \bigcup_{\Sigma \in |\mathbf{Sign}|} \mathbf{Sen}(\Sigma) \rightarrow |\mathbf{Set}|$ , such that*

$$\varphi \in \mathbf{Sen}(\Sigma) \text{ iff } \text{sym}(\varphi) \subseteq |\Sigma|$$

for all  $\varphi \in \bigcup_{\Sigma \in |\mathbf{Sign}|} \mathbf{Sen}(\Sigma)$ . The map  $\text{sym}$  gives the set of symbols used in a sentence, and sentences are uniform in the sense that a well-formed sentence is well-formed over a certain signature iff its symbols belong to that signature. Moreover, we require that any inclusion  $\iota: |\Sigma_1| \hookrightarrow |\Sigma_2|$  is a signature morphism (i.e., is in the image of  $|\cdot|$ ).

In the sequel, we fix an arbitrary institution with pre-signatures. A *pre-signature*  $\Sigma$  is a set, and a *pre-signature morphism*  $\bar{\sigma}$  consists of a right-unique set of pairs  $\text{graph}(\bar{\sigma})$  and a set  $\text{dom}(\bar{\sigma})$ , subject to the requirement that  $\text{dom}(\bar{\sigma}) \subseteq \text{def}(\bar{\sigma})$ , where  $\text{def}(\bar{\sigma}) = \{x \mid \exists y. (x, y) \in \text{graph}(\bar{\sigma})\}$ . We also define  $\text{codef}(\bar{\sigma}) = \{y \mid \exists x. (x, y) \in \text{graph}(\bar{\sigma})\}$ . We write  $\bar{\sigma}(x) = y$  iff  $(x, y) \in \text{graph}(\bar{\sigma})$ , and  $\bar{\sigma}(x) =$

<sup>1</sup> Variants of comorphisms are also used to encode “more complex” institutions into “simpler” ones: e.g. in [6], a so-called simple theoroidal comorphism is used to code first-order logic with equality in first-order logic without equality. See also [16] for discussion of relative strength of logical systems in a similar context.

$\perp$  iff  $x \notin \text{def}(\bar{\sigma})$ . Pre-signature morphisms will be later related to signature morphisms, and the intuition is that the symbols in  $\text{dom}(\bar{\sigma})$  must be mapped by the signature morphism, while the symbols in  $\text{def}(\bar{\sigma}) \setminus \text{dom}(\bar{\sigma})$  may be mapped. Given a pre-signature morphism  $\bar{\sigma}$  and a pre-signature  $\Sigma$ , define the induced function as

$$\text{fun}_{\Sigma}(\bar{\sigma}) = \text{graph}(\bar{\sigma})|_{|\Sigma|} \cup \text{Id}_{|\Sigma| \setminus \text{def}(\bar{\sigma}|_{|\Sigma|})},$$

where  $\text{graph}(\bar{\sigma})$  is construed as a function and  $\bar{\sigma}|_X$  denotes the restriction of  $\bar{\sigma}$  to  $X$ .

A pre-signature morphism  $\bar{\sigma}$  is *well-formed* wrt. a source signature  $\Sigma_1$  and a target signature  $\Sigma_2$ , if  $\text{dom}(\bar{\sigma}) \subseteq |\Sigma_1|$  and there exists a signature morphism  $\sigma: \Sigma_1 \rightarrow \Sigma_2$  with  $|\sigma| = \text{fun}_{|\Sigma_1|}(\bar{\sigma})$ . In this case,  $\sigma$  is unique and is called *the signature morphism from  $\Sigma_1$  to  $\Sigma_2$  induced by  $\bar{\sigma}$*  and we will not distinguish between the  $\sigma$  and  $\bar{\sigma}$  if  $\Sigma_1$  and  $\Sigma_2$  are clear from the context.

Composition of pre-signature morphisms is defined by

$$\bar{\sigma}_2 \circ \bar{\sigma}_1(x) := \begin{cases} \bar{\sigma}_2(\bar{\sigma}_1(x)) & \text{if } x \in \text{def}(\bar{\sigma}_1) \text{ and } \bar{\sigma}_1(x) \in \text{def}(\bar{\sigma}_2) \\ \bar{\sigma}_1(x) & \text{if } x \in \text{def}(\bar{\sigma}_1) \text{ and } \bar{\sigma}_1(x) \notin \text{def}(\bar{\sigma}_2) \\ \bar{\sigma}_2(x) & \text{if } x \notin \text{def}(\bar{\sigma}_1) \text{ and } x \in \text{def}(\bar{\sigma}_2) \\ \perp & \text{otherwise} \end{cases}$$

$$\text{dom}(\bar{\sigma}_2 \circ \bar{\sigma}_1) := \text{dom}(\bar{\sigma}_1)$$

The following is straightforward:

**Proposition 5.** *Composition of pre-signature morphisms is associative.*

The definition of composition ensures the following property:

**Proposition 6.** *If  $\text{codef}(\text{fun}_{\Sigma_1}(\bar{\sigma}_1)) \subseteq |\Sigma_2|$ , then*

$$\text{fun}_{\Sigma_2}(\bar{\sigma}_2) \circ \text{fun}_{\Sigma_1}(\bar{\sigma}_1) = \text{fun}_{\Sigma_1}(\bar{\sigma}_2 \circ \bar{\sigma}_1)$$

*Proof.* Both sides of the equation, when applied to  $x \in |\Sigma_1|$ , yield

$$\bar{\sigma}_2 \circ \bar{\sigma}_1(x) := \begin{cases} \bar{\sigma}_2(\bar{\sigma}_1(x)) & \text{if } x \in \text{def}(\bar{\sigma}_1) \text{ and } \bar{\sigma}_1(x) \in \text{def}(\bar{\sigma}_2) \\ \bar{\sigma}_1(x) & \text{if } x \in \text{def}(\bar{\sigma}_1) \text{ and } \bar{\sigma}_1(x) \notin \text{def}(\bar{\sigma}_2) \\ \bar{\sigma}_2(x) & \text{if } x \notin \text{def}(\bar{\sigma}_1) \text{ and } x \in \text{def}(\bar{\sigma}_2) \\ x & \text{otherwise} \end{cases}$$

□

A pre-signature  $\bar{\Sigma}$  is *well-formed*, if there exists a signature  $\Sigma$  with  $|\Sigma| = \bar{\Sigma}$ . Since  $|\cdot|$  is an embedding, if  $\Sigma$  exists, it is uniquely determined by  $\bar{\Sigma}$ . Hence, in the sequel, we often will not distinguish between (a well-formed)  $\bar{\Sigma}$  and  $\Sigma$ . In particular, the usual set-theoretic operations become available on signatures, the result being in some cases a (well-formed) signature and in some cases only a non-well-formed pre-signature. (Similar for signature morphisms: e.g. a signature morphism  $\sigma$  is an inclusion if  $|\sigma|$  is.) Moreover, we will use pre-signatures also in the role of partial signatures, that is, we will work with pre-signatures that are well-formed only after union with a given (base) pre-signature. A similar remark holds for pre-signature morphisms.

**Proposition 7.** *Given a pre-signature morphism  $\bar{\sigma}$  and three signatures  $\Sigma, \Sigma', \Sigma''$  with  $\Sigma \subseteq \Sigma'$  and let  $\sigma : \Sigma \rightarrow \Sigma''$  (resp.  $\sigma' : \Sigma' \rightarrow \Sigma''$ ) be the signature morphism induced by  $\bar{\sigma}$  from  $\Sigma$  to  $\Sigma''$  (resp.  $\Sigma'$  to  $\Sigma''$ ). Then  $\sigma' \circ \iota = \sigma$ , where  $\iota : \Sigma \rightarrow \Sigma'$  is the inclusion.*

Every signature morphism  $\sigma : \Sigma \rightarrow \Sigma'$  induces a pre-signature morphism  $\bar{\sigma}$  defined by

$$\text{dom}(\bar{\sigma}) := \{x \in \Sigma \mid \sigma(x) \neq x\} \text{ and } \bar{\sigma}(x) := \begin{cases} |\sigma|(x) & \text{if } x \in \text{dom}(\bar{\sigma}) \\ \perp & \text{otherwise} \end{cases}$$

**Definition 8.** *Given two pre-signature morphisms  $\bar{\sigma}$  and  $\bar{\sigma}'$ , we say that  $\bar{\sigma} \prec \bar{\sigma}'$  iff  $\text{def}(\bar{\sigma}) \subseteq \text{def}(\bar{\sigma}')$ ,  $\text{dom}(\bar{\sigma}) \subseteq \text{dom}(\bar{\sigma}')$ , and for all  $x \in \text{dom}(\bar{\sigma})$   $\bar{\sigma}(x) = \bar{\sigma}'(x)$ .*

**Proposition 9.**  *$\prec$  is a well-founded partial order on pre-signature morphisms.*

**Proposition 10.** *Let  $\bar{\sigma}$  be a pre-signature morphism and  $\sigma : \Sigma \rightarrow \Sigma'$  be its induced signature morphism between  $\Sigma$  and  $\Sigma'$ . Then for the pre-signature morphism  $\bar{\sigma}'$  induced by  $\sigma$ , it holds that  $\bar{\sigma}' \prec \bar{\sigma}$  and the signature morphism from  $\Sigma$  to  $\Sigma'$  induced by  $\bar{\sigma}'$  is  $\sigma$ .*

**Definition 11.** *An institution comorphism  $\rho : \mathcal{I} \rightarrow \mathcal{I}'$  is modular if there is*

- $\rho^{\text{PreSign}}$  mapping pre-signatures to pre-signatures and pre-signature morphisms to pre-signature morphisms<sup>2</sup> and
- $\rho^{\text{PreSen}} : \bigcup_{\Sigma \in |\mathbf{Sign}|} \rightarrow \bigcup_{\Sigma \in |\mathbf{Sign}'|}$

satisfying the following conditions:

1.  $|\rho^{\text{Sign}}(\Sigma)| = \rho^{\text{PreSign}}(|\Sigma|)$ ,
2.  $|\rho^{\text{Sign}}(\sigma)| = \rho^{\text{PreSign}}(|\sigma|)$ ,
3.  $\rho^{\text{PreSign}}(\text{fun}_{\Sigma}(\bar{\sigma})) = \text{fun}_{\rho(\Sigma)}(\rho^{\text{PreSign}}(\bar{\sigma}))$ ,
4. for each signature morphism  $\sigma : \Sigma_1 \rightarrow \Sigma_2$ ,

$$|\rho^{\text{Sign}}(\Sigma_2)| = |\rho^{\text{Sign}}(\sigma)|(|\rho^{\text{Sign}}(\Sigma_1)|) \cup \rho^{\text{PreSign}}(|\Sigma_2| \setminus |\sigma|(|\Sigma_1|))$$

5.  $\rho_{\Sigma}^{\text{Sen}}(\varphi) = \rho^{\text{PreSen}}(\varphi)$ , if  $\varphi \in \mathbf{Sen}(\Sigma)$ .

The first three conditions express that the signature translation functor can somehow be extended to pre-signatures (even if we only get something roughly behaving like a functor there), while the fourth condition expresses modularity: we can translate the symbols of  $\Sigma_1$  and those added in  $\Sigma_2$  separately. The fifth condition is a uniformity requirement for the sentence translations: they need to be combinable into a global sentence translation.

**Proposition 12.** *If  $\rho : \mathcal{I} \rightarrow \mathcal{I}'$  is modular, then for any signatures  $\Sigma_1, \Sigma_2$  in  $\mathcal{I}$  such that  $\Sigma_1 \cup \Sigma_2$  is well-formed,*

$$|\rho^{\text{Sign}}(\Sigma_1 \cup \Sigma_2)| = |\rho^{\text{Sign}}(\Sigma_1)| \cup |\rho^{\text{Sign}}(\Sigma_2)|$$

<sup>2</sup> Note that this is not the same as a functor  $\mathbf{Set} \rightarrow \mathbf{Set}$ , since pre-signature morphisms are not equipped with codomains, and domains also differ from their standard meaning.

Indeed, many institutions described in the literature, ranging from propositional over equational, first-order, modal and temporal logics to higher-order logic, can be formalized as institutions with pre-signatures, and many (co)morphisms can be made modular.

## 4 Heterogeneous Logical Environments

Given the two possible ways to link institutions with each other, a notion of a *heterogeneous logical environment* may be formalized as a collection of institutions linked by institution morphisms and comorphisms.

**Definition 13.** A modular heterogeneous logical environment  $\mathcal{HLE}$  is a collection of institutions with pre-signatures and of modular institution morphisms and comorphisms between them, that is, a pair of diagrams  $\langle \mathcal{HLE}^\mu: \mathcal{G}^\mu \rightarrow \mathcal{INS}, \mathcal{HLE}^\rho: \mathcal{G}^\rho \rightarrow \text{coINS} \rangle^3$  in the category  $\mathcal{INS}$  of institutions and their morphisms and  $\text{coINS}$  of institutions and their comorphisms, respectively, such that the two underlying graphs have no common edges and diagrams coincide on common nodes, i.e., for all nodes  $n \in |\mathcal{G}^\mu| \cap |\mathcal{G}^\rho|$ ,  $\mathcal{HLE}^\mu(n) = \mathcal{HLE}^\rho(n)$ .

We write  $\mathcal{G}$  for the union of  $\mathcal{G}^\mu$  and  $\mathcal{G}^\rho$ , and w.l.o.g. assume that all the nodes of the underlying graphs are common,  $|\mathcal{G}| = |\mathcal{G}^\mu| = |\mathcal{G}^\rho|$ .

For simplicity, we assume that each institution morphisms in  $\mathcal{HLE}$  is induced by some institution comorphism in  $\mathcal{HLE}$  via some natural transformation (see Def. 3) which is a pointwise inclusion. Most practical examples obey this additional assumption.

**Definition 14.** Consider institutions  $\mathcal{I}$  and  $\mathcal{I}'$  and signatures  $\Sigma \in |\mathbf{Sign}|$  and  $\Sigma' \in |\mathbf{Sign}'|$ .

A heterogeneous signature morphism is a pair  $\langle \mu, \sigma \rangle: \Sigma \rightarrow \Sigma'$  that consists of an institution morphism  $\mu: \mathcal{I}' \rightarrow \mathcal{I}$  and a signature morphism  $\sigma: \Sigma \rightarrow \mu^{\text{Sign}}(\Sigma')$  in  $\mathbf{Sign}$ . It induces the heterogeneous reduct  $-\downarrow_{\langle \mu, \sigma \rangle}: \mathbf{Mod}'(\Sigma') \rightarrow \mathbf{Mod}(\Sigma)$  defined as the composition  $\mathbf{Mod}(\sigma) \circ \mu_{\Sigma'}^{\text{Mod}}$ , i.e.,  $M' \downarrow_{\langle \mu, \sigma \rangle} = \mu_{\Sigma'}^{\text{Mod}}(M') \downarrow_{\sigma}$ , for all  $M' \in \mathbf{Mod}'(\Sigma')$ .

A heterogeneous pre-signature morphism is a pair  $\langle \mu, \Delta \rangle$  that consists of an institution morphism  $\mu: \mathcal{I}' \rightarrow \mathcal{I}$  and a pre-signature  $\Delta$ . It is well-formed wrt. a source signature  $\Sigma$  and target signature  $\Sigma'$  if there is some heterogeneous signature morphism  $\langle \mu, \sigma \rangle: \Sigma \rightarrow \Sigma'$  such that  $|\sigma|$  is an inclusion and  $\rho^{\text{Sign}}(\Sigma') \setminus \Sigma = \Delta$ . In this case,  $\langle \mu, \sigma \rangle$  is called the heterogeneous signature morphism from  $\Sigma$  to  $\Sigma'$  induced by  $\langle \mu, \Delta \rangle$ .

A heterogeneous signature comorphism is a pair  $\langle \rho, \sigma' \rangle: \Sigma \rightarrow \Sigma'$  that consists of an institution comorphism  $\rho: \mathcal{I} \rightarrow \mathcal{I}'$  and a signature morphism  $\sigma': \rho^{\text{Sign}}(\Sigma) \rightarrow \Sigma'$  in  $\mathbf{Sign}'$ . It induces the heterogeneous reduct  $-\downarrow_{\langle \rho, \sigma' \rangle}: \mathbf{Mod}'(\Sigma') \rightarrow \mathbf{Mod}(\Sigma)$

<sup>3</sup> We assume that  $\mathcal{G}^\mu$  is a graph that gives the shape of the diagram; its nodes  $n \in |\mathcal{G}^\mu|$  carry institutions  $\mathcal{HLE}^\mu(n)$  linked by institution morphisms  $\mathcal{HLE}^\mu(e): \mathcal{HLE}^\mu(n) \rightarrow \mathcal{HLE}^\mu(m)$  for each edge  $e: n \rightarrow m$  in  $\mathcal{G}$ . Similar notation is used for diagrams in other categories.

defined as the composition  $\rho_{\Sigma}^{Mod} \circ \mathbf{Mod}'(\sigma')$ , i.e.,  $M'|_{\langle \rho, \sigma' \rangle} = \rho_{\Sigma}^{Mod}(M'|_{\sigma'})$ , for all  $M' \in \mathbf{Mod}'(\Sigma')$ .

A heterogeneous pre-signature comorphism is a pair  $\langle \rho, \bar{\sigma} \rangle$  that consists of an institution comorphism  $\rho: \mathcal{I} \rightarrow \mathcal{I}'$  and a pre-signature morphism  $\bar{\sigma}$ . It is well-formed if there is some heterogeneous signature comorphism  $\langle \rho, \sigma \rangle: \Sigma \rightarrow \Sigma'$  such that  $\sigma$  is the signature morphism (in  $\mathcal{I}'$ ) from  $\rho^{Sign}(\Sigma)$  to  $\Sigma'$  induced by  $\bar{\sigma}$ . In this case,  $\langle \rho, \sigma \rangle$  is called the heterogeneous signature comorphism from  $\Sigma$  to  $\Sigma'$  induced by  $\langle \rho, \bar{\sigma} \rangle$ .

**Definition 15.** Given institutions  $\mathcal{I}, \mathcal{I}'$  and an  $\mathcal{I}$ -signature  $\Sigma$ . Two heterogeneous pre-signature comorphisms  $\langle \rho_1, \bar{\sigma}_1 \rangle$  and  $\langle \rho_2, \bar{\sigma}_2 \rangle$  with  $\rho_1, \rho_2: \mathcal{I} \rightarrow \mathcal{I}'$  are equivalent on  $\Sigma$ , written  $\langle \rho_1, \bar{\sigma}_1 \rangle \equiv_{\Sigma} \langle \rho_2, \bar{\sigma}_2 \rangle$ , if  $\rho_1 = \rho_2$  and  $\text{fun}_{\rho_1^{PreSign}(\Sigma)}(\bar{\sigma}_1) = \text{fun}_{\rho_1^{PreSign}(\Sigma)}(\bar{\sigma}_2)$ .

**Proposition 16.** Two heterogeneous pre-signature comorphisms  $\langle \rho_1, \bar{\sigma}_1 \rangle$  and  $\langle \rho_2, \bar{\sigma}_2 \rangle$  are equivalent on  $\Sigma$  if and only if  $\rho_1 = \rho_2$  and  $\bar{\sigma}_1(x) = \bar{\sigma}_2(x)$  for any  $x \in \rho_1^{PreSign}(\Sigma)$ .

**Definition 17.** Let  $\langle \rho, \sigma \rangle: \Sigma \rightarrow \Sigma'$  be a heterogeneous signature comorphism. It induces the heterogeneous pre-signature comorphism  $\langle \rho, \bar{\sigma} \rangle$  where  $\bar{\sigma}$  is the pre-signature morphism induced by the signature morphism  $\sigma: \rho^{PreSign}(\Sigma) \rightarrow \Sigma'$ .

**Proposition 18.** The heterogeneous pre-signature comorphism induced by a heterogeneous signature co-morphism  $\langle \rho, \sigma \rangle: \Sigma \rightarrow \Sigma'$  is well-formed and induces the same signature co-morphism  $\langle \rho, \sigma \rangle$  between  $\Sigma$  and  $\Sigma'$ .

Given two heterogeneous signature comorphisms  $\langle \rho_1, \sigma_1 \rangle: \Sigma_1 \rightarrow \Sigma_2$  and  $\langle \rho_2, \sigma_2 \rangle: \Sigma_2 \rightarrow \Sigma_3$ , their composition is defined as

$$\langle \rho_2, \sigma_2 \rangle \circ \langle \rho_1, \sigma_1 \rangle := \langle \rho_2 \circ \rho_1, \sigma_2 \circ \rho_2^{Sign}(\sigma_1) \rangle: \Sigma_1 \rightarrow \Sigma_3.$$

The problem of composing heterogeneous signature morphisms with heterogeneous signature comorphisms is solved by  $\varepsilon$ -inducibility:

**Definition 19.** Given a heterogeneous signature morphism  $\langle \mu, \sigma \rangle: \Sigma \rightarrow \Sigma'$  such that  $\mu$  is  $\varepsilon$ -induced by the institution comorphism  $\rho$  (see Def. 3), the  $\varepsilon$ -translation of  $\langle \mu, \sigma \rangle$  is the heterogeneous signature comorphism  $\langle \rho, \varepsilon_{\Sigma'} \circ \rho^{Sign}(\sigma) \rangle: \Sigma \rightarrow \Sigma'$ .

Now compositions involving heterogeneous signature morphisms are computed by first moving to their  $\varepsilon$ -translations, and then performing composition of heterogeneous signature comorphisms. Note that the resulting heterogeneous signature comorphism in general cannot be translated back to a heterogeneous signature morphism; however, in the light of the following proposition, we do not care about this.

**Proposition 20.** Sentence translation and model reduction for a heterogeneous signature morphism coincide with those of its  $\varepsilon$ -translation.

Given two heterogeneous pre-signature comorphisms  $\langle \rho_1, \bar{\sigma}_1 \rangle$  and  $\langle \rho_2, \bar{\sigma}_2 \rangle$  their *composition* is defined as

$$\langle \rho_2, \bar{\sigma}_2 \rangle \circ \langle \rho_1, \bar{\sigma}_1 \rangle := \langle \rho_2 \circ \rho_1, \bar{\sigma}_2 \circ \rho_2^{PreSign}(\bar{\sigma}_1) \rangle.$$

**Theorem 21 (Compatibility of Compositions).** *Given heterogeneous pre-signature comorphisms  $\langle \rho_1, \bar{\sigma}_1 \rangle$  and  $\langle \rho_2, \bar{\sigma}_2 \rangle$ , such that there are heterogeneous signature comorphisms  $\langle \rho_1, \sigma_1 \rangle: \Sigma_1 \rightarrow \Sigma_2$  and  $\langle \rho_2, \sigma_2 \rangle: \Sigma_2 \rightarrow \Sigma_3$  induced by  $\langle \rho_1, \bar{\sigma}_1 \rangle$  and  $\langle \rho_2, \bar{\sigma}_2 \rangle$ , respectively, then*

$$\langle \rho_2, \sigma_2 \rangle \circ \langle \rho_1, \sigma_1 \rangle: \Sigma_1 \rightarrow \Sigma_3 \text{ is induced by } \langle \rho_2, \bar{\sigma}_2 \rangle \circ \langle \rho_1, \bar{\sigma}_1 \rangle$$

*Proof.* Assume  $|\sigma_1| = fun_{|\rho_1(\Sigma_1)|}(\bar{\sigma}_1)$  and  $|\sigma_2| = fun_{|\rho_2(\Sigma_2)|}(\bar{\sigma}_2)$ . We need to show  $|\sigma_2 \circ \rho_2(\sigma_1)| = fun_{|\rho_2(\rho_1(\Sigma_1))|}(\bar{\sigma}_2 \circ \rho_2^{PreSign}(\bar{\sigma}_1))$ . Now

$$\begin{aligned} |\sigma_2 \circ \rho_2(\sigma_1)| &= |\sigma_2| \circ |\rho_2(\sigma_1)| \\ &= |\sigma_2| \circ \rho_2^{PreSign}(|\sigma_1|) \\ &= fun_{|\rho_2(\Sigma_2)|}(\bar{\sigma}_2) \circ \rho_2^{PreSign}(fun_{|\rho_1(\Sigma_1)|}(\bar{\sigma}_1)) \\ &= fun_{|\rho_2(\Sigma_2)|}(\bar{\sigma}_2) \circ fun_{|\rho_2(\rho_1(\Sigma_1))|}(\rho_2^{PreSign}(\bar{\sigma}_1)) \\ &= fun_{|\rho_2(\rho_1(\Sigma_1))|}(\bar{\sigma}_2 \circ \rho_2^{PreSign}(\bar{\sigma}_1)). \end{aligned}$$

The last equation follows from Prop. 6 by noticing that  $codef(fun_{|\rho_2(\rho_1(\Sigma_1))|}(\rho_2^{PreSign}(\bar{\sigma}_1))) \subseteq |\rho_2(\Sigma_2)|$  because  $\rho_2(\sigma_1): \rho_2(\rho_1(\Sigma_1)) \rightarrow \rho_2(\Sigma_2)$ .  $\square$

**Definition 22.** *Given a heterogeneous pre-signature morphism  $\langle \mu, \Delta \rangle$  such that  $\mu$  is  $\varepsilon$ -induced by the institution comorphism  $\rho$  (see Def. 3), the  $\varepsilon$ -translation of  $\langle \mu, \Delta \rangle$  is the heterogeneous pre-signature comorphism  $\langle \rho, \emptyset \rangle$ . The latter will induce a heterogeneous signature comorphism with a signature morphism component being an inclusion. Note that this is general enough because both  $\varepsilon$  and hiding wrt.  $\Delta$  give inclusion signature morphisms.*

Again, compositions involving heterogeneous pre-signature morphisms are computed by first moving to their  $\varepsilon$ -translations, and then performing composition of heterogeneous pre-signature comorphisms.

## 5 Heterogeneous Development Graphs

We now define development graphs over modular heterogeneous logical environments. In the definition we carefully make explicit how the signature of a theory is built from the pre-signatures of imported nodes. Furthermore, we make explicit how the set of axioms of a theory is obtained from the imported set of axioms, where we also allow for the import of axioms via hiding links if they do not contain hidden symbols.

**Definition 23.** Let  $\mathcal{HLE}$  be a modular heterogeneous logical environment. A heterogeneous development graph over  $\mathcal{HLE}$  is an acyclic, directed graph  $\mathcal{S} = \langle \mathcal{N}, \mathcal{L} \rangle$ .

$\mathcal{N}$  is a set of nodes. Each node  $N \in \mathcal{N}$  is a tuple  $(\mathcal{I}^N, \Sigma^N, \Gamma^N)$  such that  $\mathcal{I}^N$  is an institution from  $\mathcal{HLE}$ ,  $\Sigma^N$  is a  $\mathcal{I}^N$ -pre-signature called the **local signature** of  $N$ , and  $\Gamma^N$  a set of  $\mathcal{I}$ -sentences called the **local axioms** of  $N$ .

$\mathcal{L}$  is a set of directed links, so-called **definition links**, between elements of  $\mathcal{N}$ . Global definition links import the whole subgraph below a node, while local definition links import only its local signature and local axioms. Hiding definition links are like global definition links, with the possibility to hide some symbols of the signature. Free definition links are like global definition links with the possibility to declare which symbols are freely generated in the target node. Formally, each definition link from a node  $M$  to a node  $N$  is either

- **global** (denoted  $M \xrightarrow{\langle \rho, \bar{\sigma} \rangle} N$ ), annotated with a heterogeneous pre-signature comorphism  $\langle \rho, \bar{\sigma} \rangle$  such that  $\rho$  is a comorphism from  $\mathcal{I}^M$  to  $\mathcal{I}^N$ , or
- **local** (denoted  $M \xrightarrow{\langle \rho, \bar{\sigma} \rangle} N$ ), again annotated with a heterogeneous pre-signature comorphism  $\langle \rho, \bar{\sigma} \rangle$  such that  $\rho$  is a comorphism from  $\mathcal{I}^M$  to  $\mathcal{I}^N$ , or
- **hiding** (denoted  $M \xrightarrow[\text{hide}]{\langle \mu, \Delta \rangle} N$ ), annotated with a heterogeneous pre-signature morphism  $\langle \mu, \Delta \rangle$  where  $\Delta$  is a  $\mathcal{I}^M$ -pre-signature of symbols to hide, or
- **free** (denoted  $M \xrightarrow[\text{free}]{\Sigma_F} N$ ), annotated with a pre-signature of symbols over which  $N$  is freely generated.

The global pre-signature  $\text{Sig}_{\mathcal{S}}(N)$  of some node  $N$  wrt.  $\mathcal{S}$  is defined inductively over the definition links:

$$\begin{aligned}
\text{Sig}_{\mathcal{S}}(N) = & \Sigma^N \cup \bigcup_{M \xrightarrow{\langle \rho, \bar{\sigma} \rangle} N \in \mathcal{S}} \bar{\sigma}(\rho^{\text{PreSign}}(\text{Sig}_{\mathcal{S}}(M))) \\
& \cup \bigcup_{M \xrightarrow{\langle \rho, \bar{\sigma} \rangle} N \in \mathcal{S}} \bar{\sigma}(\rho^{\text{PreSign}}(\Sigma^M \cup \text{sym}(\Gamma^M))) \\
& \cup \bigcup_{M \xrightarrow[\text{hide}]{\langle \mu, \Delta \rangle} N \in \mathcal{S}} \mu^{\text{PreSign}}(\text{Sig}_{\mathcal{S}}(M)) \setminus \Delta \\
& \cup \bigcup_{M \xrightarrow[\text{free}]{\Sigma_F} N \in \mathcal{S}} \text{Sig}_{\mathcal{S}}(M)
\end{aligned}$$

A node  $N$  has a well-formed signature iff  $Sig_S(N)$  is a valid  $\mathcal{I}^N$ -signature. A development graph has a well-formed signature iff all its nodes have well-formed signatures.

Let  $M$  be a node with well-founded signature: we call the signature  $Sig_S^{loc}(M) := \langle \Sigma^M \cup sym(\Gamma^M) \rangle_{Sig_S(M)}$  the local signature of  $M$ .

Given two nodes  $M$  and  $N$  with well-formed signatures, then

- $M \xrightarrow{\langle \rho, \bar{\sigma} \rangle} N$  induces a heterogeneous signature comorphism  $\langle \rho, \sigma \rangle$  from  $Sig_S(M) \rightarrow Sig_S(N)$ ;
- $M \xrightarrow{\langle \rho, \bar{\sigma} \rangle} N$  induces a heterogeneous signature comorphism  $\langle \rho, \sigma \rangle$  from  $Sig_S^{loc}(M) \rightarrow Sig_S(N)$ ;
- $M \xrightarrow[\text{hide}]{\langle \mu, \Delta \rangle} N$  induces a heterogeneous signature morphism  $\langle \mu, \iota \rangle$  where  $\iota : Sig_S(N) \rightarrow \mu^{PreSign}(Sig_S(M))$  is the identity inclusion;
- $M \xrightarrow[\text{free}]{\Sigma_F} N$  induces the trivial heterogeneous signature morphism  $\langle Id, \iota \rangle$ .

We agree to use the above relation between heterogeneous pre-signature comorphisms and morphisms on links as notation throughout the rest of this paper.

The set of **global axioms** of some  $N$  with well-formed signature is also defined inductively over the definition link structure:

$$\begin{aligned}
Ax_S(N) = \Gamma^N \cup & \bigcup_{M \xrightarrow{\langle \rho, \bar{\sigma} \rangle} N \in \mathcal{S}} \sigma(\rho^{PreSen}(Ax_S(M))) \\
& \cup \bigcup_{M \xrightarrow{\langle \rho, \bar{\sigma} \rangle} N \in \mathcal{S}} \sigma(\rho^{PreSen}(\Gamma^M)) \\
& \cup \bigcup_{M \xrightarrow[\text{hide}]{\langle \mu, \Delta \rangle} N \in \mathcal{S}} \{\varphi \in \mu^{PreSen}(Ax_S(M)) \mid sym(\varphi) \cap \Delta = \emptyset\} \\
& \cup \bigcup_{M \xrightarrow[\text{free}]{\Sigma_F} N \in \mathcal{S}} Ax_S(M)
\end{aligned}$$

A node  $N$  is well-formed iff it has a well-formed signature  $Sig_S(N)$  and  $Ax_S(N) \subseteq \mathbf{Sen}_{\mathcal{I}^N}(Sig_S(N))$ . A development graph is well-formed, if all its nodes are well-formed.

To simplify matters, we write  $M \xrightarrow{\sigma} N \in \mathcal{S}$  instead of  $M \xrightarrow{\sigma} N \in \mathcal{L}$  when  $\mathcal{L}$  are the links of  $\mathcal{S}$ .

Since development graphs are acyclic, we can use induction principles in definitions and proofs concerning development graphs.

The next definition captures the existence of a path of local and global definition links between two nodes. Notice that such a path must not contain any hiding links.

**Definition 24.** *Let  $\mathcal{S}$  be a development graph. The notion of global reachability is defined inductively: a node  $N$  is globally reachable from a node  $M$  via a heterogeneous pre-signature comorphism  $\langle \rho, \bar{\sigma} \rangle$ ,  $M \xrightarrow{\langle \rho, \bar{\sigma} \rangle} N$  for short, iff*

- either  $M = N$  and  $\rho = id$ ,  $\bar{\sigma} = id$ , or
- $M \xrightarrow{\langle \rho', \bar{\sigma}' \rangle} K \in \mathcal{S}$ , and  $K \xrightarrow{\langle \rho'', \bar{\sigma}'' \rangle} N$ , with  $\langle \rho, \bar{\sigma} \rangle = \langle \rho'', \bar{\sigma}'' \rangle \circ \langle \rho', \bar{\sigma}' \rangle$ .

A node  $N$  is **locally reachable** from a node  $M$  via a heterogeneous pre-signature comorphism  $\langle \rho, \bar{\sigma} \rangle$ ,  $M \xrightarrow{\langle \rho, \bar{\sigma} \rangle} N$  for short, iff  $M \xrightarrow{\langle \rho, \bar{\sigma} \rangle} N$  or there is a node  $K$  with  $M \xrightarrow{\langle \rho', \bar{\sigma}' \rangle} K \in \mathcal{S}$  and  $K \xrightarrow{\langle \rho'', \bar{\sigma}'' \rangle} N$ , such that  $\langle \rho, \bar{\sigma} \rangle = \langle \rho'', \bar{\sigma}'' \rangle \circ \langle \rho', \bar{\sigma}' \rangle$ .

Obviously global reachability implies local reachability.

**Definition 25.** *Let  $\mathcal{S} = \langle \mathcal{N}, \mathcal{L} \rangle$  be a development graph. A node  $N \in \mathcal{N}$  is flattenable iff for all nodes  $M \in \mathcal{N}$  with incoming hiding or free definition links, it holds that  $N$  is not globally reachable from  $M$ .*

The models of flattenable nodes do not depend on existing hiding or free links. For flattenable nodes  $N$ ,  $Ax_{\mathcal{S}}(N)$  captures  $N$  completely. However, this is not the case for nodes that are not flattenable. Therefore, we cannot define a theory semantics of development graphs and use a model-theoretic semantics, which is compatible with the theory semantics (see Prop. 28 below).

**Definition 26.** *Given a node  $N \in \mathcal{N}$  with well-formed signature, its associated class  $\mathbf{Mod}^{\mathcal{S}}(N)$  of models (or  $N$ -models for short) consists of those  $\text{Sig}_{\mathcal{S}}(N)$ -models  $n$  for which*

- (i)  $n$  satisfies the local axioms  $\Gamma^N$ ,
- (ii) for each  $K \xrightarrow{\langle \rho, \bar{\sigma} \rangle} N \in \mathcal{S}$ ,  $n|_{\langle \rho, \bar{\sigma} \rangle}$  is a  $K$ -model,
- (iii) for each  $K \xrightarrow{\langle \rho, \bar{\sigma} \rangle} N \in \mathcal{S}$ ,  $n|_{\langle \rho, \bar{\sigma} \rangle}$  is a  $\text{Sig}_{\mathcal{S}}^{\text{loc}}(K)$ -model which satisfies the local axioms  $\Gamma^K$ , and
- (iv) for each  $K \xrightarrow[\text{hide}]{\langle \mu, \Delta \rangle} N \in \mathcal{S}$  with  $\iota : \mu^{\text{PreSign}}(\text{Sig}_{\mathcal{S}}(K)) \setminus \Delta \rightarrow \text{Sig}_{\mathcal{S}}(N)$  the corresponding inclusion mapping,  $n|_{\langle id, \iota \rangle}$  has a  $\langle \mu, \theta \rangle$ -expansion  $k$  (i.e.  $k|_{\langle \mu, \theta \rangle} = n|_{\langle id, \iota \rangle}$ ) that is a  $K$ -model where  $\langle \mu, \theta \rangle$  is the heterogeneous signature morphism from  $\mu^{\text{PreSign}}(\text{Sig}_{\mathcal{S}}(K)) \setminus \Delta$  to  $\text{Sig}_{\mathcal{S}}(K)$  induced by  $\langle \mu, \Delta \rangle$ ;
- (v) for each  $K \xrightarrow[\text{free}]{\langle Id, \Sigma_F \rangle} N \in \mathcal{S}$ ,  $n$  is a  $K$ -model which is free (in the class of  $K$ -models) over its own  $\iota$ -reduct, where  $\iota : \langle \Sigma_F \rangle_{\text{Sig}_{\mathcal{S}}(K)} \rightarrow \text{Sig}_{\mathcal{S}}(K)$  is the inclusion.

This definition of model classes nicely interacts with reachability:

**Proposition 27.** *Let  $\mathcal{S}$  be a heterogeneous development graph. Then:*

1. *if  $M \xrightarrow{\langle \rho, \bar{\sigma} \rangle} N$  and  $n \in \mathbf{Mod}^{\mathcal{S}}(N)$ , then  $n|_{\langle \rho, \sigma \rangle} \in \mathbf{Mod}^{\mathcal{S}}(M)$  where  $\langle \rho, \sigma \rangle : \text{Sig}_{\mathcal{S}}(M) \rightarrow \text{Sig}_{\mathcal{S}}(N)$  is induced by  $\langle \rho, \bar{\sigma} \rangle$ .*
2. *if  $M \xrightarrow{\langle \rho, \bar{\sigma} \rangle} N$  and  $n \in \mathbf{Mod}^{\mathcal{S}}(N)$ , then  $n|_{\langle \rho, \sigma \rangle} \models \Gamma^M$  where  $\langle \rho, \sigma \rangle : \text{Sig}_{\mathcal{S}}^{\text{loc}}(M) \rightarrow \text{Sig}_{\mathcal{S}}(N)$  is induced by  $\langle \rho, \bar{\sigma} \rangle$ .*

*Proof.* We prove (1) by easy induction over the definition of global reachability, and (2) by (1) and Definition 26.  $\square$

In the following we denote the class of  $\Sigma$ -models that fulfill the  $\Sigma$ -sentences  $\Psi$  by  $\mathbf{Mod}_{\Sigma}(\Psi)$ .

**Proposition 28.** (1)  $\mathbf{Mod}^{\mathcal{S}}(N) \subseteq \mathbf{Mod}_{\text{Sig}_{\mathcal{S}}(N)}(\text{Ax}_{\mathcal{S}}(N))$ .

(2) *If  $N$  is flattenable, then  $\mathbf{Mod}^{\mathcal{S}}(N) = \mathbf{Mod}_{\text{Sig}_{\mathcal{S}}(N)}(\text{Ax}_{\mathcal{S}}(N))$ .*

*Proof.* Let the height of some node  $N$  be the longest sequence of definition links  $N_1 \rightarrow N_2 \dots N_k \rightarrow N$  such that each transition  $N_i \rightarrow N_{i+1}$  corresponds to some global, hiding, or free definition link and the first one  $N_1 \rightarrow N_2$  may also be a local definition link.

- (1) We proceed by induction over the height of  $N$  and assume  $n \in \mathbf{Mod}^{\mathcal{S}}(N)$  and let  $\varphi \in \text{Ax}_{\mathcal{S}}(N)$ . We consider the cases of the definition of  $\text{Ax}_{\mathcal{S}}(-)$ :
  - $\varphi \in \Gamma^N$ : From Def. 26(i) it follows  $n \models \varphi$ .
  - $\varphi \in \langle \rho, \sigma \rangle(\text{Ax}_{\mathcal{S}}(K))$  for some  $K \xrightarrow{\langle \rho, \bar{\sigma} \rangle} N$ : by Def. 26(iii)  $n|_{\langle \rho, \sigma \rangle}$  is in  $\mathbf{Mod}^{\mathcal{S}}(K)$  and by induction hypothesis  $n|_{\langle \rho, \sigma \rangle} \models \text{Ax}_{\mathcal{S}}(K)$ ; it follows  $n \models \langle \rho, \sigma \rangle(\text{Ax}_{\mathcal{S}}(K))$  by the satisfaction condition and hence  $n \models \varphi$ .
  - $\varphi \in \langle \rho, \sigma \rangle(\Gamma^K)$  for some  $K \xrightarrow{\langle \rho, \bar{\sigma} \rangle} N$ : by Def. 26(ii)  $n|_{\langle \rho, \sigma \rangle}$  is a  $\text{Sig}_{\mathcal{S}}^{\text{loc}}$ -model of  $\Gamma^K$ , i.e.  $n \in \mathbf{Mod}_{\text{Sig}_{\mathcal{S}}^{\text{loc}}}(\Gamma^K)$ ; it follows  $n \models \langle \rho, \sigma \rangle(\Gamma^K)$  by the satisfaction condition and hence  $n \models \varphi$ .
  - $\varphi \in \mu^{\text{PreSen}}(\text{Ax}_{\mathcal{S}}(K))$ ,  $\text{sym}(\varphi) \cap \Delta = \emptyset$  for some  $K \xrightarrow[\text{hide}]{\langle \mu, \Delta \rangle} N$ : let  $\iota : \mu^{\text{PreSign}}(\text{Sig}_{\mathcal{S}}(K)) \setminus \Delta \rightarrow \text{Sig}_{\mathcal{S}}(N)$ , be the inclusion, then by Def. 26(iv)  $n|_{\langle id, \iota \rangle}$  has  $\langle \mu, \theta \rangle$ -expansion  $k$  such that  $k \in \mathbf{Mod}^{\mathcal{S}}(K)$ . By induction hypothesis  $k \models \text{Ax}_{\mathcal{S}}(K)$  and hence  $k|_{\langle \mu, \theta \rangle} \models \{\psi \in \mu^{\text{Sen}}(\text{Ax}_{\mathcal{S}}(K)) \mid \text{sym}(\psi) \cap \Delta = \emptyset\}$ ; Since  $n|_{\langle id, \iota \rangle} = k|_{\langle \mu, \theta \rangle}$  (see Def. 26(iv)), it follows that  $n|_{\langle id, \iota \rangle} \models \varphi$  and thus by satisfaction condition and  $\iota(\varphi) = \varphi$  it follows  $n \models \varphi$ .
  - $\varphi \in \langle id, \iota \rangle(\text{Ax}_{\mathcal{S}}(K))$  for some  $K \xrightarrow[\text{free}]{\langle \rho, \Sigma_F \rangle} N$ : by Def. 26(v)  $n \in \mathbf{Mod}^{\mathcal{S}}(K)$  and by induction hypothesis  $n \models \text{Ax}_{\mathcal{S}}(K)$  and thus  $n \models \varphi$ .
- (2) By (1), it suffices to prove the “ $\supseteq$ ” direction. We proceed again by induction of the height of  $N$  and assume  $n$  is an  $\text{Ax}_{\mathcal{S}}(N)$ -model. Since  $N$  is flattenable, we only have to show the clauses (i) to (iii) of Definition 26:

- (i) since  $\Gamma^N \subseteq Ax_S(N)$  it holds that  $n \models \Gamma^N$ .
- (ii) let  $K \xrightarrow{\langle \rho, \bar{\sigma} \rangle} N$ : since  $\langle \rho, \sigma \rangle(Ax_S(K)) \subseteq Ax_S(N)$ , it holds  $n \models \langle \rho, \sigma \rangle(Ax_S(K))$ . By satisfaction condition it follows  $n|_{\langle \rho, \sigma \rangle} \models Ax_S(K)$  and thus  $n \in \mathbf{Mod}^S(K)$  by induction hypothesis.
- (iii) let  $K \xrightarrow{\langle \rho, \bar{\sigma} \rangle} N$ : since  $\langle \rho, \sigma \rangle(\Gamma^K) \subseteq Ax_S(N)$ , it holds  $n \models \langle \rho, \sigma \rangle(\Gamma^K)$ , where  $\langle \rho, \sigma \rangle : Sig_S^{loc}(K) \rightarrow Sig_S(N)$ . By the satisfaction condition  $n|_{\langle \rho, \sigma \rangle} \models \Gamma^K$  and is a  $Sig_S^{loc}(K)$ -model.  $\square$

Complementary to definition links, which *define* the theories of related nodes, we introduce the notion of a *theorem link* with the help of which we are able to *postulate* relations between different theories. Theorem links are the central data structure to represent proof obligations arising in formal developments. Again we distinguish between local and global theorem links (denoted by  $N \xrightarrow{\langle \rho, \bar{\sigma} \rangle} M$  and  $N \xrightarrow{\langle \rho, \bar{\sigma} \rangle} M$  respectively). Moreover, we introduce *local implications* of form  $N \Rightarrow \Gamma$ , where  $\Gamma$  is a set of  $Sig_S(N)$ -sentences.  $N \Rightarrow \{\varphi\}$  also is written  $N \Rightarrow \varphi$ . Finally, we also need theorem links  $N \xrightarrow[\text{hide } \langle \mu, \Delta \rangle]{\langle \rho, \bar{\sigma} \rangle} M$  (where for  $\Sigma_H := \mu^{PreSign}(Sig_S(N)) \setminus \Delta$ ,  $\langle \mu, \Delta \rangle : Sig_S(N) \rightarrow \Sigma_H$  and  $\langle \rho, \sigma \rangle : \Sigma_H \rightarrow Sig_S(M)$ ) involving hiding, as well as  $N \xrightarrow[\text{free } \langle Id, \Sigma_F \rangle]{\langle \rho, \bar{\sigma} \rangle} M$  involving freeness.

The semantics of theorem links is given by the next definition.

**Definition 29.** *Let  $\mathcal{S}$  be a development graph and  $N, M$  nodes in  $\mathcal{S}$ .*

$\mathcal{S}$  **satisfies** a global theorem link  $N \xrightarrow{\langle \rho, \bar{\sigma} \rangle} M$  (denoted  $\mathcal{S} \models N \xrightarrow{\langle \rho, \bar{\sigma} \rangle} M$ ) iff for all  $m \in \mathbf{Mod}^S(M)$ ,  $m|_{\langle \rho, \sigma \rangle} \in \mathbf{Mod}^S(N)$  where  $\langle \rho, \sigma \rangle$  is the heterogeneous signature comorphism from  $Sig_S(N)$  to  $Sig_S(M)$  induced by  $\langle \rho, \bar{\sigma} \rangle$ .

$\mathcal{S}$  **satisfies** a local theorem link  $N \xrightarrow{\langle \rho, \bar{\sigma} \rangle} M$  (denoted  $\mathcal{S} \models N \xrightarrow{\langle \rho, \bar{\sigma} \rangle} M$ ) iff for all  $m \in \mathbf{Mod}^S(M)$ ,  $m|_{\langle \rho, \sigma \rangle} \in \mathbf{Mod}_{Sig_S^{loc}(N)}(\Gamma^N)$  where  $\langle \rho, \sigma \rangle$  is the heterogeneous signature comorphism from  $Sig_S^{loc}(N)$  to  $Sig_S(M)$  induced by  $\langle \rho, \bar{\sigma} \rangle$ .

$\mathcal{S}$  **satisfies** a local implication  $N \Rightarrow \Gamma$ , written  $\mathcal{S} \models N \Rightarrow \Gamma$ , if for all  $n \in \mathbf{Mod}_S(N)$ ,  $n \models \Gamma$ .

$\mathcal{S}$  **satisfies** a hiding theorem link  $N \xrightarrow[\text{hide } \langle \mu, \Delta \rangle]{\langle \rho, \bar{\sigma} \rangle} M$  (denoted  $\mathcal{S} \models N \xrightarrow[\text{hide } \langle \mu, \Delta \rangle]{\langle \rho, \bar{\sigma} \rangle} M$ ) iff for all  $m \in \mathbf{Mod}^S(M)$ ,  $m|_{\langle \rho, \sigma \rangle \circ \langle id, \iota \rangle}$  has a  $\langle \mu, \theta \rangle$ -expansion to some  $N$ -model where  $\langle \mu, \theta \rangle$  is the heterogeneous signature morphism from  $\mu^{PreSign}(Sig_S(N)) \setminus \Delta \rightarrow Sig_S(N)$  induced by  $\langle \mu, \Delta \rangle$ ,  $\langle id, \iota \rangle : \mu^{PreSign}(Sig_S(N)) \setminus \Delta \rightarrow Sig_S(M)$  is the identity inclusion, and  $\langle \rho, \sigma \rangle$  is the heterogeneous signature comorphism from  $\mu^{PreSign}(Sig_S(N)) \setminus \Delta \rightarrow Sig_S(M)$  induced by  $\langle \rho, \bar{\sigma} \rangle$ .

$\mathcal{S}$  **satisfies** a free theorem link  $N \xrightarrow[\text{free } \langle Id, \Sigma_F \rangle]{\langle \rho, \bar{\sigma} \rangle} M$  if for all  $m \in \mathbf{Mod}^S(M)$  it holds that  $m|_{\langle \rho, \sigma \rangle}$  is an  $N$ -model which is free (in the class of  $N$ -models) over its

own  $\iota$ -reduct, where  $\iota: \langle \Sigma_F \rangle_{\text{Sig}_S(N)} \rightarrow \text{Sig}_S(N)$  is the inclusion and  $\langle \rho, \sigma \rangle$  is the heterogeneous signature comorphism from  $\text{Sig}_S(N)$  to  $\text{Sig}_S(M)$  induced by  $\langle \rho, \bar{\sigma} \rangle$ .

Common proof obligations in a formal development can be encoded into properties such that specific global theorem links are implied by the actual development graph.

A global definition link  $M \xrightarrow{\sigma} N$  in a development graph is a *conservative extension*, if every  $M$ -model can be expanded along  $\sigma$  to an  $N$ -model. We will allow to annotate a global definition link as  $M \xrightarrow[\text{cons}]{\sigma} N$ , which shall express that it is a conservative extension. Such annotations can be seen as another kind of proof obligation.

## 6 Proof Rules for Development Graphs

The development graph calculus consists of eight basic rules: global theorem links can either be decomposed into a set of local theorem links, global theorem links, hiding theorem links and free theorem links following the structure of the source node using **(Global-Decomposition)**. Or else they can be subsumed by existing paths in the graph with an equivalent heterogeneous signature comorphism using **(Global-Subsumption)**. Local theorem links can either be decomposed into local implications by **(Local-Inference)** or, like global theorem links, be subsumed by existing paths with equivalent heterogeneous signature comorphism using **(Local-Subsumption)**. Local implications can be discharged by proving the individual conjectures in a sound calculus of the underlying institution by **(Basic-Inference)**.

In order to get rid of hiding links going into the *source* of a global theorem link, one first applies **(Global-Decomposition)**, ending up with some local and hiding theorem links. The rule **(Hide-Theorem-Shift)** allows to prove the latter, using conservativity of definition links. **(Borrowing)** can be used for shifting a proof goal along a conservative extension; hence, it also exploits conservativity of theorem links. Another rule of the proof system is the rule **(Theorem-Hide-Shift)**. It is used to get rid of hiding definition links going into the *target* of a global theorem link.

**Global-Decomposition Rule:**

$$\begin{array}{c}
 N \xrightarrow{\langle \rho, \bar{\sigma} \rangle} K \\
 \hline
 P \xrightarrow{\langle \rho, \bar{\sigma} \rangle \circ \langle \rho', \bar{\tau} \rangle} K \text{ for each } P \xrightarrow{\langle \rho', \bar{\tau} \rangle} N \\
 P \xrightarrow{\langle \rho, \bar{\sigma} \rangle \circ \langle \rho', \bar{\tau} \rangle} K \text{ for each } P \xrightarrow{\langle \rho', \bar{\tau} \rangle} N \\
 P \xrightarrow[\text{hide}]{\langle \rho, \bar{\sigma} \rangle} K \text{ for each } P \xrightarrow[\text{hide}]{\langle \mu, \Delta \rangle} N \\
 P \xrightarrow[\text{free}]{\langle \rho, \bar{\sigma} \rangle} K \text{ for each } P \xrightarrow[\text{free}]{\Sigma_F} N \\
 \hline
 N \xrightarrow{\langle \rho, \bar{\sigma} \rangle} K
 \end{array}$$

*Soundness:* assume that

$$\mathcal{S} \models N \xrightarrow{\langle \rho, \bar{\sigma} \rangle} K \quad (1)$$

$$\mathcal{S} \models P \xrightarrow{\langle \rho, \bar{\sigma} \rangle \circ \langle \rho', \bar{\tau} \rangle} K \text{ for each } P \xrightarrow{\langle \rho', \bar{\tau} \rangle} N \quad (2)$$

$$\mathcal{S} \models P \xrightarrow{\langle \rho, \bar{\sigma} \rangle \circ \langle \rho', \bar{\tau} \rangle} K \text{ for each } P \xrightarrow{\langle \rho', \bar{\tau} \rangle} N \quad (3)$$

$$\mathcal{S} \models P \xrightarrow[\text{hide } \langle \mu, \Delta \rangle]{\langle \rho, \bar{\sigma} \rangle} K \text{ for each } P \xrightarrow[\text{hide } \langle \mu, \Delta \rangle]{\langle \mu, \Delta \rangle} N \quad (4)$$

$$\mathcal{S} \models P \xrightarrow[\text{free } \Sigma_F]{\langle \rho, \bar{\sigma} \rangle} K \text{ for each } P \xrightarrow[\text{free } \Sigma_F]{\Sigma_F} N \quad (5)$$

In order to show  $\mathcal{S} \models N \xrightarrow{\langle \rho, \bar{\sigma} \rangle} K$ , let  $m$  be a  $K$ -model and we show  $m|_{\langle \rho, \sigma \rangle}$  is an  $N$ -model where  $\sigma : \rho^{\text{PreSign}}(\text{Sig}_{\mathcal{S}}(N)) \rightarrow \text{Sig}_{\mathcal{S}}(K)$ . We then consider the cases (i) to (v) of Definition 26:

1.  $\Gamma^N$ : by (1) we know  $m|_{\langle \rho, \sigma' \rangle} \in \mathbf{Mod}_{\text{Sig}_{\mathcal{S}}^{\text{loc}}(N)}(\Gamma^N)$ , where  $\sigma' : \rho^{\text{PreSign}}(\text{Sig}_{\mathcal{S}}^{\text{loc}}(N)) \rightarrow \text{Sig}_{\mathcal{S}}(K)$ . Since  $\text{Sig}_{\mathcal{S}}^{\text{loc}}(N) \subseteq \text{Sig}_{\mathcal{S}}(N)$  it holds  $\sigma' = \sigma \circ \iota$  by Proposition 7 where  $\iota : \text{Sig}_{\mathcal{S}}^{\text{loc}}(N) \rightarrow \text{Sig}_{\mathcal{S}}(N)$  is the identity inclusion. Hence, it holds

$$\begin{aligned} m|_{\langle \rho, \sigma' \rangle} \in \mathbf{Mod}_{\text{Sig}_{\mathcal{S}}^{\text{loc}}(N)}(\Gamma^N) &\Rightarrow m|_{\langle \rho, \sigma \circ \iota \rangle} \in \mathbf{Mod}_{\text{Sig}_{\mathcal{S}}^{\text{loc}}(N)}(\Gamma^N) \\ &\Rightarrow m|_{\langle \rho, \sigma \rangle} \in \mathbf{Mod}_{\text{Sig}_{\mathcal{S}}(N)}(\iota(\Gamma^N)) \\ &\Rightarrow m|_{\langle \rho, \sigma \rangle} \in \mathbf{Mod}_{\text{Sig}_{\mathcal{S}}(N)}(\Gamma^N). \end{aligned}$$

2.  $P \xrightarrow{\langle \rho', \bar{\tau} \rangle} N$  where  $\tau : \rho^{\text{PreSign}}(\text{Sig}_{\mathcal{S}}^{\text{loc}}(P)) \rightarrow \text{Sig}_{\mathcal{S}}(N)$ : from (2) and Theorem 21 it follows  $(m|_{\langle \rho, \sigma \rangle \circ \langle \rho', \tau \rangle}) \in \mathbf{Mod}^{\mathcal{S}}(P)$ .
3.  $P \xrightarrow{\langle \rho', \bar{\tau} \rangle} N$  where  $\tau : \rho^{\text{PreSign}}(\text{Sig}_{\mathcal{S}}(P)) \rightarrow \text{Sig}_{\mathcal{S}}(N)$ : from (3) and Theorem 21 it follows  $(m|_{\langle \rho, \sigma \rangle \circ \langle \rho', \tau \rangle}) \in \mathbf{Mod}^{\mathcal{S}}(P)$ .
4.  $P \xrightarrow[\text{hide } \langle \mu, \Delta \rangle]{\langle \mu, \Delta \rangle} N$  where  $\iota : \mu^{\text{PreSign}}(\text{Sig}_{\mathcal{S}}(P)) \setminus \Delta \rightarrow \text{Sig}_{\mathcal{S}}(N)$  is the identity inclusion morphism and  $\langle \mu, \theta \rangle : \mu^{\text{PreSign}}(\text{Sig}_{\mathcal{S}}(P)) \setminus \Delta \rightarrow \text{Sig}_{\mathcal{S}}(P)$ : from (4) and Theorem 21 it follows  $m|_{\langle \rho, \sigma \rangle \circ \langle \text{id}, \iota \rangle}$  has a  $\langle \mu, \theta \rangle$ -expansion in  $\mathbf{Mod}^{\mathcal{S}}(P)$ .
5.  $P \xrightarrow[\text{free } \Sigma_F]{\langle \text{id}, \Sigma_F \rangle} N$ : from (5) and Theorem 21 it follows  $m|_{\langle \rho, \sigma \rangle \circ \langle \text{id}, \iota \rangle} \in \mathbf{Mod}^{\mathcal{S}}(P)$ .

□

*Remark 30.* Note the use of Theorem 21 in the soundness proof above to lift the composition of heterogeneous pre-signature comorphisms to the induced heterogeneous signature comorphisms. The compatibility of compositions occurs in the soundness proofs of most other proof rules as well and we will make use of it without mentioning it explicitly.

**Global-Subsumption Rule:**

$$\frac{K \xrightarrow{\langle \rho', \bar{\sigma}' \rangle} N}{K \xrightarrow{\langle \rho, \bar{\sigma} \rangle} N} \text{ if } \langle \rho, \bar{\sigma} \rangle \equiv \langle \rho', \bar{\sigma}' \rangle \text{ wrt. } Sig_S(K)$$

*Soundness:* Since the heterogeneous pre-signature comorphisms are equivalent wrt.  $Sig_S(K)$  the induced heterogeneous signature comorphisms are equal, from which the soundness follows trivially.  $\square$

**Local-Subsumption Rule:**

$$\frac{K \xrightarrow{\langle \rho', \bar{\theta} \rangle} N}{K \xrightarrow{\langle \rho, \bar{\sigma} \rangle} N} \text{ if } \langle \rho, \bar{\sigma} \rangle \equiv \langle \rho', \bar{\theta} \rangle \text{ wrt. } Sig_S^{loc}(K)$$

*Soundness:* assume that  $\mathcal{S} \models K \xrightarrow{\langle \rho', \bar{\theta} \rangle} N$  and  $\langle \rho, \bar{\sigma} \rangle \equiv \langle \rho', \bar{\theta} \rangle$  wrt.  $Sig_S^{loc}(K)$ . Then the respective induced heterogeneous signature comorphisms  $\langle \rho, \sigma \rangle, \langle \rho', \theta \rangle : Sig_S^{loc}(K) \rightarrow Sig_S(N)$  are equal (\*). Let  $m$  be an  $N$ -model. By Proposition 27,  $m|_{\langle \rho', \theta \rangle} \models \Gamma^K$ . From (\*) follows  $m|_{\langle \rho, \sigma \rangle} \models \Gamma^K$ , from which follows  $\mathcal{S} \models K \xrightarrow{\langle \rho, \bar{\sigma} \rangle} N$ .  $\square$

**Local-Inference Rule:**

$$\frac{N \Rightarrow \langle \rho, \bar{\sigma} \rangle(\Gamma^K)}{K \xrightarrow{\langle \rho, \bar{\sigma} \rangle} N}$$

*Soundness:* assume that  $m \models \langle \rho, \bar{\sigma} \rangle(\Gamma^K)$  for each  $N$ -model  $m$  and let  $\langle \rho, \sigma \rangle : Sig_S^{loc}(K) \rightarrow Sig_S(N)$  be the heterogeneous signature comorphism induced by  $\langle \rho, \bar{\sigma} \rangle$ . In order to show  $\mathcal{S} \models K \xrightarrow{\langle \rho, \bar{\sigma} \rangle} N$ , let  $m$  be an  $N$ -model. By assumption,  $m \models \langle \rho, \bar{\sigma} \rangle(\Gamma^K)$ . By the satisfaction condition for institutions,  $m|_{\langle \rho, \sigma \rangle} \models \Gamma^K$ . By definition, it follows that  $\mathcal{S} \models K \xrightarrow{\langle \rho, \bar{\sigma} \rangle} N$ .  $\square$

**Basic-Inference Rule:**

$$\frac{P : Ax_S(N) \vdash_{\Sigma^N} \varphi \text{ for each } \varphi \in \Psi}{N \Rightarrow \Psi}$$

where  $P : Ax_S(N) \vdash_{\Sigma^N} \varphi$  denotes that  $P$  is a proof for  $Ax_S(N) \vdash_{\Sigma^N} \varphi$ .

*Soundness:* assume that  $Ax_S(N) \vdash_{Sig_S(N)} \varphi$  for each  $\varphi \in \Psi$ . By soundness of  $\vdash_{Sig_S(N)}$ , we get  $Ax_S(N) \models_{Sig_S(N)} \Psi$ . In order to show  $\mathcal{S} \models N \Rightarrow \Psi$ , let  $m$  be an  $N$ -model. By Proposition 28,  $m \models Ax_S(N)$ . Since  $Ax_S(N) \models_{Sig_S(N)} \Psi$ , also  $m \models \Psi$ .  $\square$

**Hide-Theorem-Shift Rule:**

$$\frac{
\begin{array}{c}
\langle \rho', \bar{\sigma}' \rangle \xrightarrow{\text{cons}} N' \\
\langle \rho', \bar{\sigma}' \rangle \xrightarrow{\text{cons}} N' \\
K \xrightarrow{\text{hide}} \langle \mu, \Delta \rangle
\end{array}
}{
\begin{array}{c}
N' \\
N' \\
\langle \rho, \bar{\sigma} \rangle \xrightarrow{\text{cons}} N \\
K \xrightarrow{\text{hide}} \langle \mu, \Delta \rangle
\end{array}
}
\text{ if } \langle \rho' \circ \text{CoM}(\mu), \bar{\sigma}' \circ \text{CoM}(\mu)^{\text{PreSign}}(\emptyset) \rangle \equiv \langle \rho'' \circ \rho, \bar{\theta} \circ \rho^{\text{PreSign}}(\bar{\sigma}) \rangle$$

wrt. the signature  $\mu^{\text{PreSign}}(\text{Sig}_{\mathcal{S}}(K)) \setminus \Delta$ .

The proof rules are written in a concise notation as above. We will spell out in detail what this notation means for the rule (**Hide-Theorem-Shift**):

$$\frac{
\begin{array}{c}
\langle \rho' \circ \text{CoM}(\mu), \bar{\sigma}' \circ \text{CoM}(\mu)^{\text{PreSign}}(\emptyset) \rangle \equiv \langle \rho'' \circ \rho, \bar{\theta} \circ \rho^{\text{PreSign}}(\bar{\sigma}) \rangle \\
\text{wrt. } \mu^{\text{PreSign}}(\text{Sig}_{\mathcal{S}}(K)) \setminus \Delta \\
N \xrightarrow{\text{cons}} \langle \rho'', \bar{\theta} \rangle \in \mathcal{S} \\
\mathcal{S} \vdash K' \xrightarrow{\text{hide}} \langle \rho', \bar{\sigma}' \rangle \Rightarrow N' \\
\mathcal{S} \vdash K' \xrightarrow{\text{hide}} \langle \mu, \Delta \rangle \Rightarrow N
\end{array}
}{
}$$

*Soundness:* assume that  $\mathcal{S} \models K' \xrightarrow{\text{hide}} \langle \rho', \bar{\sigma}' \rangle \Rightarrow N'$  and  $N = \xrightarrow{\text{cons}} \langle \rho'', \bar{\theta} \rangle \in \mathcal{S}$  is conservative. We have to show that  $\mathcal{S} \models K' \xrightarrow{\text{hide}} \langle \mu, \Delta \rangle \Rightarrow N$ . Let  $m$  be an  $N$ -model. Since  $N = \xrightarrow{\text{cons}} \langle \rho'', \bar{\theta} \rangle \in \mathcal{S}$  is conservative,  $m$  can be expanded to an  $N'$ -model  $m'$  with  $m'|_{\theta'} = m$ . By the assumption,  $m'|_{\sigma'}$  is a  $K'$ -model. Thus,  $m'|_{\sigma' \circ \theta} = m'|_{\theta' \circ \sigma} = m|_{\sigma}$  has a  $\theta$ -expansion to an  $K'$ -model.  $\square$

**Theorem-Hide-Shift Rule:**

$$\frac{
\begin{array}{ccc}
G_N(i) & & L_N(i) \\
\downarrow & & \downarrow \\
\langle \rho_i, \bar{\mu}_i \rangle & & \langle \rho_i, \bar{\mu}_i \rangle \quad (i \in |J| \cap \text{dom}(L)) \\
\downarrow & & \downarrow \\
K = \xrightarrow{\text{hide}} \langle \rho_{(N)}, \bar{\mu}_{(N)} \rangle \circ \langle \rho, \bar{\sigma} \rangle \Rightarrow P
\end{array}
}{
\text{Diag}_N
}$$

$$\begin{array}{c}
\langle \rho, \bar{\sigma} \rangle \Rightarrow N
\end{array}$$

with  $P$  isolated and  $\langle \rho_i, \mu_i \rangle$  induced by  $\langle \rho_i, \bar{\mu}_i \rangle$  a weakly amalgamable co-cone for the diagram  $\text{Diag}_N$  of nodes going into  $N$  (see explanation below)

Since this rule is quite powerful, we need some preliminary notions. Given a node  $N$  in a development graph  $\mathcal{S} = \langle \mathcal{N}, \mathcal{L} \rangle$ , the idea is that we unfold the subgraph below  $N$  into a tree and form a diagram with this tree. More formally, define the *diagram*  $Diag_N: J \rightarrow \mathbf{Sig}$  associated with  $N$  together with partial maps  $G_N: |J| \rightarrow \mathcal{N}$  and  $L_N: |J| \rightarrow \mathcal{L}$  inductively as follows:

- $\langle N \rangle$  is an object in  $J$ , with  $Diag_N(\langle N \rangle) = Sig_{\mathcal{S}}(N)$ . Let  $G_N(\langle N \rangle)$  be just  $N$ .
- if  $i = \langle K \xrightarrow{L_1} \dots \xrightarrow{L_n} N \rangle$  is an object in  $J$  with  $L_1, \dots, L_n$  non-local definition links in  $\mathcal{L}$ , and  $L = M \xrightarrow{\langle \rho, \bar{\sigma} \rangle} K$  a global definition link in  $\mathcal{L}$ , then

$$j = \langle M \xrightarrow{L} K \xrightarrow{L_1} \dots \xrightarrow{L_n} N \rangle$$

is an object in  $J$  with  $Diag_N(j) = Sig_{\mathcal{S}}(M)$ , and  $L$  is a morphism from  $j$  to  $i$  in  $J$  with  $Diag_N(L) = \langle \mu, \sigma \rangle$  where  $\langle \mu, \sigma \rangle: Sig_{\mathcal{S}}(M) \rightarrow Sig_{\mathcal{S}}(K)$  is the heterogeneous signature comorphism induced by  $L$ . If there is no incoming free definition link for  $M$  in  $\mathcal{S}$ , we set  $L_N(j) = M$  and otherwise  $G_N(j) = M$ .

- if  $i = \langle K \xrightarrow{L_1} \dots \xrightarrow{L_n} N \rangle$  is an object in  $J$  with  $L_1, \dots, L_n$  non-local definition links in  $\mathcal{L}$ , and  $L = M \xrightarrow{\langle \rho, \bar{\sigma} \rangle} K$  a local definition link in  $\mathcal{L}$ , then

$$j = \langle M \xrightarrow{L} K \xrightarrow{L_1} \dots \xrightarrow{L_n} N \rangle$$

is an object in  $J$  with  $Diag_N(j) = Sig_{\mathcal{S}}(M)$ , and  $L$  is a morphism from  $j$  to  $i$  in  $J$  with  $Diag_N(L) = \langle \mu, \sigma \rangle$  where  $\langle \mu, \sigma \rangle: Sig_{\mathcal{S}}^{loc}(M) \rightarrow Sig_{\mathcal{S}}(K)$  is the heterogeneous signature comorphism induced by  $L$ . We set  $L_N(j) = M$ .

- if  $i = \langle K \xrightarrow{L_1} \dots \xrightarrow{L_n} N \rangle$  is an object in  $J$  with  $L_1, \dots, L_n$  non-local definition links in  $\mathcal{L}$ , and  $L = M \xrightarrow[\text{hide}]{\langle \mu, \Delta \rangle} K$  a hiding definition link in  $\mathcal{L}$ , then

$$j = \langle M \xrightarrow{L} K \xrightarrow{L_1} \dots \xrightarrow{L_n} N \rangle$$

is an object in  $J$  with  $Diag_N(j) = Sig_{\mathcal{S}}(M)$ , and  $L$  is a morphism from  $i$  to  $j$  in  $J$  with  $Diag_N(L) = \langle \mu, \theta \rangle$  where  $\langle \mu, \theta \rangle: Sig_{\mathcal{S}}(K) \rightarrow Sig_{\mathcal{S}}(M)$  is the heterogeneous signature morphism induced by  $L$ . If there is no incoming free definition link for  $M$  in  $\mathcal{S}$ , we set  $L_N(j) = M$  and otherwise  $G_N(j) = M$ .

Now in order to apply (**Theorem-Hide-Shift**), take a weakly amalgamable cocone  $(\Sigma, (\langle \rho_i, \mu_i \rangle: Diag_N(i) \rightarrow \Sigma)_{i \in |J|})$  for  $Diag_N$  (in general, we know that such a cocone exists only if the institution is quasi-semi-exact), and let  $P$  be a new isolated node with signature  $\Sigma$  and with ingoing global definition links

$G_N(i) \xrightarrow{\langle \rho_i, \bar{\mu}_i \rangle} P$  for  $i \in |J| \cap dom(G)$  where  $\langle \rho_i, \bar{\mu}_i \rangle$  are the heterogeneous pre-signature comorphisms induced by  $\langle \rho_i, \mu_i \rangle$  and with ingoing local definition links

$G_N(i) \xrightarrow{\langle \rho_i, \bar{\mu}_i \rangle} P$  for  $i \in |J| \cap dom(L)$ . Here, an isolated node is one with no local axioms and no ingoing definition links other than those shown in the rule.

*Soundness:* assume that  $\mathcal{S} \models \langle \rho_{\langle N \rangle}, \bar{\mu}_{\langle N \rangle} \rangle \circ \langle \rho, \bar{\sigma} \rangle$ . Let  $m$  be an  $N$ -model. We have to show  $m|_{\langle \rho_{\langle N \rangle}, \mu_{\langle N \rangle} \rangle \circ \langle \rho, \sigma \rangle}$  to be a  $K$ -model in order to establish the holding of  $\langle \rho_{\langle N \rangle}, \bar{\mu}_{\langle N \rangle} \rangle \circ \langle \rho, \bar{\sigma} \rangle$ . We inductively define a family  $(m_i)_{i \in |J|}$  of models  $m_i \in \mathbf{Mod}(G_N(i))$  if  $i \in \text{dom}(G_N)$  or  $m_i \in \mathbf{Mod}^{\mathcal{S}}(L_N(i))$  if  $i \in \text{dom}(L_N)$  by putting

- $m_{\langle N \rangle} = m$ ;
- $m_{\langle M \xrightarrow{L} Q \xrightarrow{L_1} \dots \xrightarrow{L_n} N \rangle} = m'|_{\langle \rho, \sigma \rangle}$ , where  $L = M \xrightarrow{\langle \rho, \bar{\sigma} \rangle} Q$  (note:  $\langle \rho, \sigma \rangle : \text{Sig}_{\mathcal{S}}(M) \rightarrow \text{Sig}_{\mathcal{S}}(Q)$ ) and  $m' = m_{\langle Q \xrightarrow{L_1} \dots \xrightarrow{L_n} N \rangle}$ ;
- $m_{\langle M \xrightarrow{L} Q \xrightarrow{L_1} \dots \xrightarrow{L_n} N \rangle} = m'|_{\langle \rho, \sigma \rangle}$ , where  $L = M \xrightarrow{\langle \rho, \bar{\sigma} \rangle} Q$  (note:  $\langle \rho, \sigma \rangle : \text{Sig}_{\mathcal{S}}^{\text{loc}}(M) \rightarrow \text{Sig}_{\mathcal{S}}(Q)$ ) and  $m' = m_{\langle Q \xrightarrow{L_1} \dots \xrightarrow{L_n} N \rangle}$ ; and
- $m_{\langle M \xrightarrow{L} Q \xrightarrow{L_1} \dots \xrightarrow{L_n} N \rangle}$  is a  $\langle \mu, \theta \rangle$ -expansion of  $m'|_{\langle Id, \iota \rangle}$  to an  $M$ -model (existing since  $m'$  is a  $Q$ -model), where  $L = M \xrightarrow[\text{hide}]{\langle \mu, \Delta \rangle} Q$ ,  $\langle Id, \iota \rangle : \mu^{\text{PreSign}}(\text{Sig}_{\mathcal{S}}(M)) \setminus \Delta \rightarrow \text{Sig}_{\mathcal{S}}(Q)$  the heterogeneous signature identity inclusion comorphism, and  $m' = m_{\langle Q \xrightarrow{L_1} \dots \xrightarrow{L_n} N \rangle}$ .

It is easy to show that this family is consistent with  $\text{Diag}_N$ . Since by the side condition of the rule,  $(\Sigma, (\langle \rho_i, \mu_i \rangle : \text{Diag}_N(i) \rightarrow \Sigma)_{i \in |J|})$  is a weakly amalgamable cocone, there is a  $\Sigma^P$ -model  $m_P$  with  $m_P|_{\langle \rho_i, \mu_i \rangle} = m_i$ ,  $i \in |J|$ . The latter implies that  $m_P$  is a  $P$ -model. By the assumption,  $m_P|_{\langle \rho_{\langle N \rangle}, \mu_{\langle N \rangle} \rangle \circ \langle \rho, \sigma \rangle} = m_{\langle N \rangle}|_{\sigma} = m|_{\sigma}$  is a  $K$ -model.

It remains to show that it is safe to annotate the new definition links in the graph with the induced heterogeneous pre-signature comorphisms  $\langle \rho_i, \bar{\mu}_i \rangle$  rather than with the heterogeneous signature comorphisms  $\langle \rho_i, \mu_i \rangle$ . This is ensured by the fact that the induced heterogeneous pre-signature comorphism induces the same heterogeneous signature comorphism between the same source and target signatures (Proposition 18).  $\square$

### **Borrowing Rule:**

$$\begin{array}{c}
 \begin{array}{ccc}
 K & & N \\
 \parallel & & \parallel \\
 \langle \rho, \bar{\theta} \rangle \parallel & & \langle \rho', \bar{\theta}' \rangle \parallel \text{cons} \\
 \downarrow & & \downarrow \\
 K' \equiv \equiv \Rightarrow N' & & \\
 \langle \rho_{\bar{\sigma}'}, \bar{\sigma}' \rangle & & \\
 \hline
 K \equiv \equiv \Rightarrow N & & \text{if } \langle \rho_{\bar{\sigma}'}, \bar{\sigma}' \rangle \circ \langle \rho, \bar{\theta} \rangle \equiv \langle \rho', \bar{\theta}' \rangle \circ \langle \rho_{\bar{\sigma}}, \bar{\sigma} \rangle \\
 \parallel & & \parallel \\
 \langle \rho, \bar{\theta} \rangle \parallel & & \langle \rho', \bar{\theta}' \rangle \parallel \text{cons} \\
 \downarrow & & \downarrow \\
 K' & & N'
 \end{array}
 \end{array}
 \text{ wrt. } \text{Sig}_{\mathcal{S}}(K)$$

*Soundness:* Assume that (1)  $\mathcal{S} \models K \stackrel{\langle \rho, \bar{\theta} \rangle}{=} \Rightarrow K'$ , (2)  $\mathcal{S} \models N \stackrel{\langle \rho', \bar{\theta}' \rangle}{\underset{\text{cons}}{=} \Rightarrow} N'$ , and that (3)  $\mathcal{S} \models K' \stackrel{\langle \rho_{\sigma'}, \bar{\sigma}' \rangle}{=} \Rightarrow N'$ . Let  $m$  be an  $N$ -model. By (2),  $m$  has an  $\langle \rho', \theta' \rangle$ -expansion to an  $N'$ -model  $m'$  with  $m'|_{\langle \rho', \theta' \rangle} = m$ . By (3),  $m'|_{\langle \rho_{\sigma'}, \sigma' \rangle}$  is an  $K'$ -model, and hence, by (1) and Proposition 16  $m'|_{\langle \rho_{\sigma'}, \sigma' \rangle \circ \langle \rho, \theta \rangle} = m'|_{\langle \rho', \theta' \rangle \circ \langle \rho_{\sigma}, \sigma \rangle} = m|_{\langle \rho_{\sigma}, \sigma \rangle}$  is a  $K$ -model.  $\square$

**Theorem 31.** *The proof rules are sound. The rules are incomplete, but complete if the underlying institutions have complete calculi and relative to a given oracle deciding conservativity and freeness of global definition links.*

*Proof.* For each rule a soundness proof has been given. For incompleteness, see the counter-example published in [15]. For the relative completeness results see [14].  $\square$

## 7 Change Impact Analysis

In this section we are concerned with the question of transferring proof work done in one particular development graph to another graph that differs from the first one only in some locally constricted areas such that we are able to relate most of the nodes and (definition) links of the first to nodes and links of the latter. Given this mapping we are interested in the problem of mapping the proof work encoded in theorem links, their decompositions, and proofs of theorems to the new development graph. Proof work is done by applying the development graph rules presented in Section 6 to the actual development graph.

The transfer of proof work from the original development graph to the changed one can be easily done by reapplying the “proof scripts”, i.e. the rules used in the first are reused to verify theorem links at corresponding positions in the latter. However, in general the application of the development graph rules has a high price because it involves, for instance, the use of theorem proving or the calculation of compositions of (co-)morphisms. Hence, a major design goal is to define the scope of the rules (in terms of the affected subgraph) as small as possible in order to minimize the test for applicability to a limited area within the development graph but also to keep the rules as generic as possible. As a consequence we may be able to transfer a proof (script) from the original development graph to the new one although the semantics of the represented proof obligation has changed.

We introduced the notions of pre-signatures and local axioms to specify the signature and sentences of a theory in an incremental way. While a node (as the root of its subgraph) represents semantically a (full-fledged) theory, its syntactical representation is restricted to those bits that are not imported from other theories via some incoming definition link. From this point of view, a node considered in its own right denotes a function that maps its imported theories to the theory represented by the subgraph of the node. Several development graph rules (e.g. the global decomposition rule) exploit the constructive nature of these functions rather than inspecting the individual theory represented by

the node. In this sense, the rules are rather generic with respect to the semantics of the imported theories. This is important for change management since we are interested in a smart replay of the development graph proofs in a changed environment. The same development graph rules may still be applicable although the semantics of the involved nodes have been significantly affected by a change of the specification. Smart replay means that we want to anticipate the result of applying a rule in a changed setting by adaptation of the result of application in the original setting. Doing this we will know the differences between the two rule applications allowing us to transfer also proof steps incorporating the results of original rule applications to the new setting. In the following we will analyze each individual development graph rule according to how changes in the area in which the rule is applied will affect the result of the rule application. Therefore, we distinguish between the *domain*, which is the subgraph of that graph with all elements that contribute to the semantics of the involved entities, and the *pre-domain* of a proof rule that consists only of those parts that actually (syntactically) matter for the rule application.

**Global-Decomposition Rule:** Applied on  $N \stackrel{\langle \rho, \bar{\sigma} \rangle}{=} \Rightarrow K$  the pre-domain consists exclusively of  $\langle \rho, \bar{\sigma} \rangle$ , the node  $N$  and all direct incoming definition links (local, global, hiding, free) along with their heterogeneous pre-signature morphisms and comorphisms and their source nodes. The domain contains the theorem link and the subgraphs imported into  $N$  and  $K$  including all signature elements and axioms.

*Impact analysis:* if some definition link from the pre-domain is deleted, the corresponding (local/global) theorem link needs to be deleted as well. If some definition link has been added to the pre-domain, a new (local/global) theorem link needs to be added. If  $\langle \rho, \bar{\sigma} \rangle$  has changed, the heterogeneous pre-signature comorphisms and heterogeneous pre-signature morphisms of the introduced theorem links are affected and must be recomputed. Analogously, if the heterogeneous pre-signature comorphism of some incoming definition link has been changed, the heterogeneous pre-signature comorphism of the corresponding theorem link is affected and must be recomputed. If the heterogeneous pre-signature morphism of some incoming hiding definition link has been changed, the heterogeneous pre-signature morphism of the corresponding hiding theorem link is affected and must be recomputed.

**Global-Subsumption Rule:** Applied on  $K \stackrel{\langle \rho, \bar{\sigma} \rangle}{=} \Rightarrow N$  to subsume it by the path computed by  $K \stackrel{\langle \rho', \bar{\sigma}' \rangle}{\succ} \Rightarrow N$  the pre-domain comprises  $K$ ,  $N$ ,  $\langle \rho, \bar{\sigma} \rangle$ , and all nodes, links and their respective heterogeneous pre-signature comorphisms on the path  $K \stackrel{\langle \rho', \bar{\sigma}' \rangle}{\succ} \Rightarrow N$ . The domain includes in addition the transitive closure of incoming definition links, nodes, local signatures and local axioms for  $K$ ,  $N$  and all nodes on the path.

*Impact analysis:* If some node or link on the path is deleted, the rule is invalid. Otherwise, if  $\langle \rho, \bar{\sigma} \rangle$  or some of the heterogeneous pre-signature comorphisms on

the path has been changed then the rule must check again if they are still equal wrt. the (new) global signature of  $K$ .

**Local-Inference Rule:** For this proof rule applied on  $K \xrightarrow{\langle \rho, \bar{\sigma} \rangle} N$  the pre-domain and the domain coincide and are exclusively the node  $K$ , its local signature  $Sig_S^{loc}(N)$  and the local axioms  $\Gamma^K$  as well as the heterogeneous pre-signature comorphism. Any changes outside of these are irrelevant for the proof rule application.

*Impact Analysis:* any change in the local signature  $Sig_S^{loc}(K)$  has no effect on the proof rule as long as the heterogeneous pre-signature comorphism  $\langle \rho, \bar{\sigma} \rangle$  is unchanged. If  $\langle \rho, \bar{\sigma} \rangle$  changed, it must be checked whether the mappings of the local axioms from  $\Gamma^K$  are still the same. If a local axiom has been deleted from  $\Gamma^K$ , then the corresponding conjecture in the local implication is superfluous and can be removed. If a local axiom has been added, then a corresponding conjecture has to be added to the local implication.

**Local-Subsumption Rule:** Applied on  $K \xrightarrow{\langle \rho, \bar{\sigma} \rangle} N$  to subsume it by the path computed by  $K \xrightarrow{\langle \rho', \bar{\sigma}' \rangle} N$  the pre-domain comprises  $P$  and its local signature and axioms  $N$ ,  $\langle \rho, \bar{\sigma} \rangle$ , and all nodes, links and their respective heterogeneous pre-signature comorphisms on the path  $K \xrightarrow{\langle \rho', \bar{\sigma}' \rangle} N$ . The domain includes in addition  $K$  and its local signature and local axioms, as well as the transitive closure of incoming definition links, nodes, local signatures and local axioms of  $N$  and all nodes on the path.

*Impact analysis:* If some node or link on the path is deleted, the rule is invalid. Otherwise, any change of  $\langle \rho, \bar{\sigma} \rangle$  or some of the heterogeneous pre-signature comorphisms on the path requires to check again if they are still equal wrt. the (new) local signature of  $K$ .

**Basic-Inference Rule:** Applied to reduce  $N \Rightarrow \varphi_1 \dots \varphi_n$  to  $P_1 : Ax_S(N) \vdash_{\Sigma^N} \varphi_1 \dots P_n : Ax_S(N) \vdash_{\Sigma^N} \varphi_n$  the domain consist of the  $\varphi_i$  and the subgraph imported into  $N$  including all the axioms and signature elements. The pre-domain, though, only consists of the  $\varphi_i$  and the global set of axioms of  $N$  determined by  $Ax_S(N)$ .

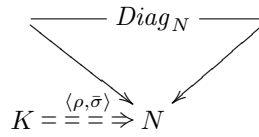
*Impact analysis:* If some  $\varphi_i$  is deleted, then the rule application is still sound, but over-complete and the superfluous premise  $P_i : Ax_S(N) \vdash_{\Sigma^N} \varphi_i$  must be deleted. If there is an additional  $\varphi$ , then the rule application is incomplete and the respective additional premise must be added. If some axiom has been deleted from  $Ax_S(N)$ , which was used in the proof  $P_i$  of some  $\varphi_i$ , then that proof is affected and needs to be redone. If some axiom has been added to  $Ax_S(N)$ , nothing is affected.

**Hide-Theorem-Shift Rule:** Applied to prove the hiding theorem link in  $K \xrightarrow[\text{hide } \langle \mu, \Delta \rangle]{\langle \rho, \bar{\sigma} \rangle} N \xrightarrow[\text{cons } \langle \rho', \bar{\theta} \rangle]{\langle \rho', \bar{\theta} \rangle} N'$  the domain comprises the two links with their heterogeneous pre-signature morphisms and comorphisms, the subgraph imported in the

source node  $K$  including all signature elements and axioms, as well as the subgraph imported in  $N'$  also including all signature elements and axioms. The pre-domain, however, are only the two links with  $\langle \rho, \bar{\sigma} \rangle$ ,  $\langle \mu, \Delta \rangle$ , and  $\langle \rho'', \bar{\theta} \rangle$ , as well as the nodes  $K$ ,  $N$ , and  $N'$ .

*Impact analysis:* If  $\langle \rho, \bar{\sigma} \rangle$ ,  $\langle \rho'', \bar{\theta} \rangle$  or  $\mu$  change, then the heterogeneous pre-signature comorphism  $\langle \rho', \bar{\sigma}' \rangle$  is also affected and must be recomputed. If the global signature  $Sig_S(K)$  or  $\Delta$  change, the equivalence condition needs to be rechecked. Finally, if the conservativity status of the definition link changes, the rule is incorrect and must be deleted.

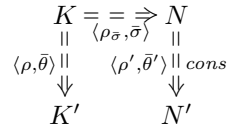
**Theorem-Hide-Shift Rule:** Applying the rule to the global theorem link in



the domain and pre-domain coincide and consist of  $K$ , the theorem link with  $\langle \rho, \bar{\sigma} \rangle$ , as well as  $N$  and the whole subgraph imported in  $N$  including all signature elements and axioms.

*Impact analysis:* Any change in the diagram affecting the signatures of the nodes and the heterogeneous pre-signature comorphisms requires to recompute the (heterogeneous) colimit, as there is no means to analyze the impact of changes here to the content of  $P$ . To optimize this and extend the impact analysis to the colimit computation would require to transfer the colimit computation from heterogeneous signature comorphisms to heterogeneous pre-signature comorphisms, which is an open task.

**Borrowing Rule:** For application of the rule to shift the theorem link between  $K$  and  $N$  in



to a theorem link between  $K'$  and  $N'$ , the domain comprises the imported subgraphs with all signature elements and axioms of  $K$ ,  $N$ ,  $K'$  and  $N'$ , as well as all three theorem links with heterogeneous pre-signature comorphisms. The pre-domain, however, only consists of these three theorem links and the global signature of  $K$ .

*Impact Analysis:* if one of the involved heterogeneous pre-signature comorphisms is affected, the heterogeneous pre-signature comorphism of the new theorem link between  $K'$  and  $N'$  needs to be recomputed and the side condition rechecked. If the global signature of  $K$  has changed, then the side-condition needs to be rechecked.

*Remark 32.* Note the benefits for change impact analysis of using pre-signature morphisms: first, pre-signature morphisms change less frequently in general than full signature morphisms, which are affected by each change in the global signature of the source node. Thus, changes occur less often using that representation. Second, whether a new heterogeneous pre-signature comorphism maps a given axiom differently than the precedent heterogeneous pre-signature comorphism can be checked locally and without computing the full heterogeneous signature comorphism: take the symbols of an axiom and compare the intersection with the domain of the old and the new heterogeneous pre-signature comorphism: if the intersection is equal and the heterogeneous pre-signature comorphisms are equivalent on this intersection, then the axiom is mapped in the same way and in the rule (**Local-Inference**) the old conjectures in the local implication can remain in place.

*Realization.* The change impact analysis has been realized for the HETS-tool which is an implementation of heterogeneous development graphs, but with signatures and signature morphisms rather than pre-signatures and pre-signature morphisms. Adapting HETS would have been a major enterprise and we thus implemented it in the *GMoc*-tool for generic change impact analysis [3], which will also allow us to combine change impact analysis for development graphs with change impact analysis of other documents occurring in a software development process, such as source code, requirements documents and general documentation. *GMoc* performs the impact analysis as described above based on a serialized representation of heterogeneous development graphs and provides the results of the analysis in form of impact annotations to the serialized representation. In addition to the information about affected theorem links it also provides information about those development graph nodes and links for which the signature and respectively the signature morphisms need to be recomputed by HETS. The change impact analysis in *GMoc* works on typed graphs obtained from the serialized representation and the impact analysis is formalized as a set of graph rewriting rules (see [3] for details).

## 8 Conclusion

In this paper we provided a thorough definition of heterogeneous development graphs that allow for an efficient management of change. We introduced pre-signatures and pre-signature morphisms as elementary building blocks to specify complex signatures and signature morphisms, which allows us to specify theories in a completely modular way. The proof rules make use of this modularity in order to restrict the focus of testing the applicability of rules to some few nodes and their relations in the development graph. The locality of rule applications supports the reuse of proofs in case of changing the development graph. We use a smart replay mechanism which anticipates the result of applying a rule in a changed setting by adapting the result of its application in the previous setting.

The idea of development graphs turned out to be very fruitful in the past and various research groups adopted this approach. The proof semantics of

CASL [17], the theory structuring mechanisms in OMDoc [10], the recent module system for Twelf [18] and also locales in Isabelle [4] were influenced by the notion of development graphs. The presented approach allows for an efficient support of specification and verification with the help of heterogeneous development graphs. The next step will be to extend the framework of change management to the use of generalized theoroidal institution comorphisms (cf. [5]) generalizing the way how signatures of a source logic can be translated to entire theories of the target one and broadening the class of logic encodings that can be formalized as comorphisms. We already have defined the notion of modular comorphism in a way that anticipates this extension.

## References

1. S. Autexier and D. Hutter. Maintenance of formal software developments by stratified verification. In *Proceedings 9th International Conference on Logic for Programming Artificial Intelligence and Reasoning*. Springer-Verlag, LNAI, 2002.
2. S. Autexier and D. Hutter. Mind the gap - maintaining formal developments in MAYA. In *Festschrift in Honor of J.H. Siekmann*. Springer-Verlag, LNCS 2605, 2005.
3. S. Autexier and N. Müller. Semantics-based change impact analysis for heterogeneous collections of documents. In M. Gormish and R. Ingold, editors, *Proceedings of 10th ACM Symposium on Document Engineering (DocEng2010)*, Manchester, UK, september 2010.
4. C. Ballarin. Interpretation of locales in Isabelle: Theories and proof contexts. In J. M. Borwein and W. M. Farmer, editors, *Mathematical Knowledge Management, 5th International Conference, MKM 2006, Wokingham, UK, August 11-12, 2006, Proceedings*, volume 4108 of *Lecture Notes in Computer Science*, pages 31–43. Springer, 2006.
5. M. Codrescu. Generalized theoroidal institution comorphisms. In *Recent Trends in Algebraic Development Techniques: 19th International Workshop, WADT 2008, Pisa, Italy, June 13-16, 2008, Revised Selected Papers*, pages 88–101, Berlin, Heidelberg, 2009. Springer-Verlag.
6. J. Goguen and G. Roşu. Institution morphisms. *Formal Aspects of Computing*, 13:274–307, 2002. 10.1007/s001650200013.
7. J. A. Goguen and R. M. Burstall. Institutions: Abstract model theory for specification and programming. *Journal of the Association for Computing Machinery*, 39:95–146, 1992. Predecessor in: LNCS 164, 221–256, 1984.
8. J. A. Goguen and G. Rosu. Composing hidden information modules over inclusive institutions. In O. Owe, S. Krogdahl, and T. Lyche, editors, *Essays in Memory of Ole-Johan Dahl*, volume 2635 of *Lecture Notes in Computer Science*, pages 96–123. Springer, 2004.
9. D. Hutter. Management of change in verification systems. In *Proceedings 15th IEEE International Conference on Automated Software Engineering, ASE-2000*, pages 23–34. IEEE Computer Society, 2000.
10. M. Kohlhase. *OMDOC - An Open Markup Format for Mathematical Documents [Version 1.2]*, volume 4180 of *LNAI*. Springer, August 2006.
11. J. Meseguer. General logics. In *Logic Colloquium 87*, pages 275–329. North Holland, 1989.

12. T. Mossakowski. Specifications in an arbitrary institution with symbols. In D. Bert, C. Choppy, and P. D. Mosses, editors, *WADT*, volume 1827 of *Lecture Notes in Computer Science*, pages 252–270. Springer, 1999.
13. T. Mossakowski. Heterogeneous development graphs and heterogeneous borrowing. In M. Nielsen and U. Engberg, editors, *Foundations of Software Science and Computation Structures*, volume 2303 of *Lecture Notes in Computer Science*, pages 326–341. Springer-Verlag, 2002.
14. T. Mossakowski. Heterogeneous specification and the heterogeneous tool set. Habilitation thesis, University of Bremen, 2005.
15. T. Mossakowski, S. Autexier, and D. Hutter. Development graphs - proof management for structured specifications. *Journal of Logic and Algebraic Programming, special issue on Algebraic Specification and Development Techniques*, 67(1-2):114–145, april 2006.
16. T. Mossakowski, R. Diaconescu, and A. Tarlecki. What is a logic translation? *Logica Universalis*, 3(1):95–124, 2009.
17. P. D. Mosses, editor. *CASL Reference Manual*. Number 2960 in LNCS. Springer, 2004.
18. F. Rabe and C. Schürmann. A practical module system for lf. In *LFMTP '09: Proceedings of the Fourth International Workshop on Logical Frameworks and Meta-Languages*, pages 40–48, New York, NY, USA, 2009. ACM.
19. D. T. Sannella and A. Tarlecki. Extended ML: an institution-independent framework for formal program development. In *Proc. Workshop on Category Theory and Computer Programming*, volume 240 of *Lecture Notes in Computer Science*, pages 364–389. Springer, 1986.
20. A. Tarlecki. Institution representation. Unpublished note, Dept. of Computer Science, University of Edinburgh, 1987.
21. A. Tarlecki. Moving between logical systems. In *Recent Trends in Data Type Specification*, pages 478–502. Springer, 1998.