

A Modal Specification Theory for Components with Data

Bauer, Sebastian S.; Larsen, Kim Guldstrand; Legay, Axel; Nyman, Ulrik; Wasowski, Andrzej

Published in:
Formal Aspects of Component Software

DOI (link to publication from Publisher):
[10.1007/978-3-642-35743-5_5](https://doi.org/10.1007/978-3-642-35743-5_5)

Publication date:
2012

Document Version
Accepted author manuscript, peer reviewed version

[Link to publication from Aalborg University](#)

Citation for published version (APA):
Bauer, S. S., Larsen, K. G., Legay, A., Nyman, U., & Wasowski, A. (2012). A Modal Specification Theory for Components with Data. In *Formal Aspects of Component Software: 8th International Symposium, FACS 2011, Oslo, Norway, September 14-16, 2011, Revised Selected Papers* (Vol. 7253). Springer Publishing Company. https://doi.org/10.1007/978-3-642-35743-5_5

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

A Modal Specification Theory for Components with Data^{*}

Sebastian S. Bauer^{1,2}, Kim G. Larsen², Axel Legay³,
Ulrik Nyman², and Andrzej Wąsowski⁴

¹ Institut für Informatik, Ludwig-Maximilians-Universität München, Germany

² Department of Computer Science, Aalborg University, Denmark

³ INRIA/IRISA, Rennes, France

⁴ IT University of Copenhagen, Denmark

Abstract. Modal specification is a well-known and widely used formalism used as an abstraction theory for transition systems. Modal specifications are transition systems equipped with two types of transitions: *must*-transitions that are mandatory to any implementation, and *may*-transitions that are optional. The duality of transitions allows to develop a unique approach for both logical and structural compositions, and eases the step-wise refinement process for building implementations.

We propose Modal Specifications with Data (MSD), the first *modal* specification theory with explicit representation of data. Our new theory includes all the essential ingredients of a specification theory. As MSD are potentially infinite-state systems, we propose symbolic representations based on effective predicates. Our theory serves as a new abstraction-based formalism for transition systems with data.

1 Introduction

Modern IT systems are often large and consist of complex assemblies of numerous reactive and interacting components. The components are often designed by independent teams, working under a common agreement on what the interface of each component should be. Consequently, the search for mathematical foundations which support *compositional reasoning* on interfaces is a major research goal. A framework should support inferring properties of the global implementation, designing and advisedly reusing components.

Interfaces are specifications and components that implement an interface are understood as models/implementations. Specification theories should support various features including (1) *refinement*, which allows to compare specifications as well as to replace a specification by another one in a larger design, (2) *structural composition*, which allows to combine specifications of different components, (3) *logical conjunction*, expressing the intersection of the set of requirements expressed by two or more specifications, and last (4) a *quotient operator* that is dual to structural composition and allows synthesizing a component from a set of assumptions.

^{*} Work supported by the German Academic Exchange Service (DAAD), grant D/10/46169, by an “Action de Recherche Collaborative” ARC (TP)I, by MT-LAB, VKR Centre of Excellence, and by the EU project ASCENS, 257414.

Among existing specification theories, one finds modal specifications [1], which are labeled transition systems equipped with two types of transitions: *must*-transitions that are mandatory for any implementation, and *may*-transitions which are optional for an implementation. Modal specifications are known to achieve a more flexible and easy-to-use compositional development methodology for CCS [2], which includes a considerable simplification of the step-wise refinement process proposed by Milner and Larsen. While being very close to logics (conjunction), the formalism takes advantage of a behavioral semantics allowing for easy composition with respect to process construction (structural composition) and synthesis (quotient). However, despite the many advantages, only a few implementations have been considered so far. One major problem is that contrary to other formalisms based on transition systems, there exists no theory of modal specification equipped with rich information such as data variables.

In this paper, we add a new stone to the cathedral of results on modal specifications [3, 4], that is we propose the first such theory equipped with rich data values. Our first contribution is to design a semantical version of modal specifications whose states are split into locations and valuations for possibly infinite-domain variables. For every component, we distinguish between local variables, that are locally controlled by the component, and uncontrolled variables that are controlled by other components and can be accessed, but not modified. Combining variables with sets of actions labeling transitions offers a powerful set of communication primitives that cannot be captured by most existing specification theories. We also propose a symbolic predicate-based representation of our formalism. We consider effective predicates that are closed under conjunction, union, and membership—classical assumptions in existing symbolic theories (e.g. [5]). While the semantic level is possibly infinite-state, the syntactical level permits us to reason on specifications just like one would with the original modal specifications, but with the additional power of rich data.

Continuing our quest, we study modal refinement between specifications. Refinement, which resembles simulation between transition systems, permits to compare sets of implementations in a syntactic manner. Modal refinement is defined at the semantic level, but can also be checked at the symbolic level. We propose a predicate abstraction approach that simplifies the practical complexity of the operation by reducing the number of states and simplifying the predicates. This approach is in line with the work of Godefroid et al. [6], but is applied to specification-based verification rather than to model checking.

We then propose definitions for both logical and structural composition, on the level of symbolic representations of specifications. These definitions are clearly not direct extensions of the ones defined on modal specifications as behaviors of both controlled and uncontrolled variables have to be taken into account. As usual, structural composition offers the property of independent implementability, hence allowing for elegant step-wise refinement. In logical composition, two specifications which disagree on their requirements can be reconciled by synthesizing a new component where conflicts have been removed. This can be done with a symbolic pruning of bad states, which terminates if the system is finite-state, or if the structure of the transition system induced by the specification relies, for instance, on a well-quasi order [7]. Finally, we also propose a quotient operation, that is the dual operation of structural composition, which works for

a subclass of systems, and we discuss its limitation. This operator, absent from most existing behavioral and logical specification theories, allows synthesizing a component from a set of assumptions.

In Sect. 2 we introduce modal specifications with data and their finite symbolic representations, refinement, an implementation relation and consistency. In Sect. 3 we define the essential operators of every specification theory, that is parallel composition, conjunction and quotient. For verification of refinement between infinite-state specifications we propose in Sect. 4 an approach based on predicate abstraction techniques. We summarize related works in Sect. 5 and conclude in Sect. 6.

Acknowledgment. We would like to thank Rolf Hennicker for valuable comments on a draft of the paper.

2 Modal Specifications with Data

We will first introduce specifications which are finite symbolic representations of modal specifications with data. We will then propose modal refinement and derive an implementation relation and a consistency notion.

In the following, $\mathcal{P}(M)$ denotes the powerset of M , $\mathcal{P}_{\geq 1}(M) = \mathcal{P}(M) \setminus \{\emptyset\}$, and the union of two disjoint sets is denoted by $M \uplus N$, which is $M \cup N$ with $M \cap N = \emptyset$.

Let \mathbb{V} be a fixed set of variables, each variable ranging over a fixed domain \mathbb{D} . For a given subset $V \subseteq \mathbb{V}$, a *data state* s over V is a mapping $s : V \rightarrow \mathbb{D}$. If $V = \{x_1, x_2, \dots, x_n\}$ and $d_1, d_2, \dots, d_n \in \mathbb{D}$, we write $[x_1 \mapsto d_1, x_2 \mapsto d_2, \dots, x_n \mapsto d_n]$ for the data state s which maps every x_i to d_i , for $1 \leq i \leq n$. We write $\llbracket V \rrbracket$ for the set of all possible data states over V . For disjoint sets of variables V_1 and V_2 and data states $s_1 \in \llbracket V_1 \rrbracket$ and $s_2 \in \llbracket V_2 \rrbracket$, the operation $(s_1 \cdot s_2)$ composes the data states resulting in a new state $s = (s_1 \cdot s_2) \in \llbracket V_1 \uplus V_2 \rrbracket$, such that $s(x) = s_1(x)$ for all $x \in V_1$ and $s(x) = s_2(x)$ for all $x \in V_2$. This is naturally lifted to sets of states: if $S_1 \subseteq \llbracket V_1 \rrbracket$ and $S_2 \subseteq \llbracket V_2 \rrbracket$ then $(S_1 \cdot S_2) = \{(s_1 \cdot s_2) \mid s_1 \in S_1, s_2 \in S_2\} \subseteq \llbracket V_1 \uplus V_2 \rrbracket$.

Like in the work of de Alfaro et al. [8] we define specifications with respect to an assertion language allowing suitable predicate representation. Given a set V of variables, we denote by $Pred(V)$ the set of first-order predicates with free variables in V ; we assume that these predicates are written in some specified first-order language with existential (\exists) and universal (\forall) quantifiers and with interpreted function symbols and predicates; in our examples, the language contains the usual arithmetic operators and boolean connectives ($\vee, \wedge, \neg, \Rightarrow$). Syntactic equality of predicates is written with the symbol \equiv . Given a set of variables V we denote by $(V)'$ an isomorphic set of 'primed' variables from V : so if $x \in V$ then $(x)' \in (V)'$. We use this construction to represent pre- and post-values of variables. A variable $(x)' \in (V)'$ represents the next state value of the variable $x \in V$. Given a formula $\varphi \in Pred(V)$ and a data state $s \in \llbracket V \rrbracket$, we write $\varphi(s)$ if the predicate formula φ is true when its free variables are interpreted as specified by s . Given a formula $\psi \in Pred(V_1 \uplus (V_2)')$ and states $s_1 \in \llbracket V_1 \rrbracket$, $s_2 \in \llbracket V_2 \rrbracket$, we often write $\psi(s_1, s_2)$ for $\psi(s_1 \cdot t_2)$ where $t_2 \in \llbracket (V_2)' \rrbracket$ such that $t_2((x)') = s_2(x)$ for all $x \in V_2$. Given a predicate $\varphi \in Pred(V)$, we write $(\varphi)' \in Pred((V)')$ for the predicate obtained by substituting x with $(x)'$ in φ , for all $x \in V$. We write $\llbracket \varphi \rrbracket$ for the

set $\{s \in \llbracket V \rrbracket \mid \varphi(s)\}$ which consists of all states satisfying $\varphi \in \text{Pred}(V)$ (for predicates with primed and unprimed variables), and φ is *consistent* if $\llbracket \varphi \rrbracket \neq \emptyset$. We write $\exists V \varphi$ meaning existential quantification of φ over all variables in the set V , and similar for universal quantification. Finally, for a predicate $\psi \in \text{Pred}(V_1 \uplus (V_2)')$, we write ${}^\circ\psi$ for $\exists(V_2)'\psi$, and ψ° for $\exists V_1 \psi$.

Our theory enriches modal automata with variables. Specifications not only express constraints on the allowed sequences of actions, but also their dependence and effect on the values of variables. Like in the loose approach of modal specifications [1] which allows under-specification using *may* and *must* modalities on transitions, we allow loose specification of the effects of actions on the data state. From a given location and a given data state, a transition to another location is allowed to lead to several next data states. Unlike in modal specifications, variables are observable in our framework, allowing for modeling shared variable communication.

A *signature* $\text{Sig} = (\Sigma, V^L, V^G)$ determines the alphabet of actions Σ and the set of variables $V = V^L \uplus V^G$ of an interface. The variables in V^L are *local (controlled) variables*, owned by the interface and visible to any other component. V^G contains the *uncontrolled variables* owned by the environment, which are read-only for the interface.

Specifications are finite modal transition systems where transitions are equipped with predicates. A transition predicate $\psi \in \text{Pred}(V \uplus (V^L)')$ relates a previous state, determined by all controlled and uncontrolled data states, with the next possible controlled data state.

Definition 1. A specification is a tuple $\mathbf{A} = (\text{Sig}, \text{Loc}, \ell^0, \varphi^0, E_\Diamond, E_\Box)$ where $\text{Sig} = (\Sigma, V^L, V^G)$ is a signature, Loc is a finite set of locations, $\ell^0 \in \text{Loc}$ is the initial location, $\varphi^0 \in \text{Pred}(V^L)$ is a predicate on the initial local state, and E_\Diamond, E_\Box are finite *may- and must-transition relations* respectively:

$$E_\Diamond, E_\Box \subseteq \text{Loc} \times \Sigma \times \text{Pred}(V \uplus (V^L)') \times \text{Loc}.$$

Given a specification \mathbf{A} , locations $\ell, \ell' \in \text{Loc}$, and action $a \in \Sigma$, we refer to the set of transition predicates on may-transitions by $\text{May}^a(\ell, \ell') = \{\psi \mid (\ell, a, \psi, \ell') \in E_\Diamond\}$ and on must-transitions by $\text{Must}^a(\ell, \ell') = \{\psi \mid (\ell, a, \psi, \ell') \in E_\Box\}$.

Example 1. Consider a specification of a print server, shown in Fig. 1. Must-transitions are drawn with solid arrows and may-transitions with dashed ones. Every solid arrow representing a must-transition has an implicit may-transition shadowing it which is not shown. Every transition is equipped with a transition predicate over unprimed variables, referring to the pre-state, and primed variables, referring to the poststate. The print server receives new print jobs (**newPrintJob**), stores them and assigns them either a low or high priority; the numbers of low and high priority jobs are modeled by controlled variables l and h , respectively; l and h are natural numbers. A job with low priority can also be reclassified to high priority (**incPriority**). The printer server can send (**send**) a job to a printer, and then wait for the acknowledgment (**ack**). In state ℓ_1 , if there is a job with high priority and the uncontrolled boolean variable *priorityMode* is true, then there must be a send transition. The specification is loose in the sense that if a second print job is received in state ℓ_1 , then the behavior is left unspecified.

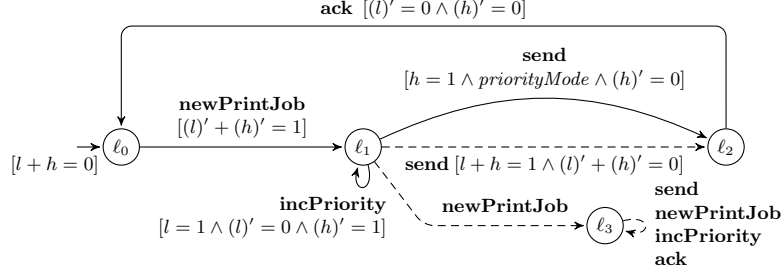


Fig. 1. Abstract specification **P** of a print server.

We now define the kind of transition systems which will be used for formalizing the semantics of specifications. A specification is interpreted as a variant of modal transition systems where the *state space* is formed by the cartesian product $Loc \times \llbracket V^L \rrbracket$, i.e. a *state* is a pair (ℓ, s) where $\ell \in Loc$ is a location and $s \in \llbracket V^L \rrbracket$ is a valuation of the controlled variables. To motivate the choice of the transition relations in the semantics of specifications, we first describe the intended meaning of may- and must-transitions.

A may-transition $(\ell, a, \psi, \ell') \in E_\diamond$ in the specification expresses that in any implementation, in any state (ℓ, s) and for any guard $g \in \llbracket V^G \rrbracket$ (that is a valuation of uncontrolled variables V^G) the implementation is *allowed* to have a transition with guard g and action a to a next state (ℓ', s') such that $\psi(s \cdot g, s')$. The interpretation of a must-transition $(\ell, a, \psi, \ell') \in E_\square$ is a bit more involved: Any implementation, in state (ℓ, s) , and for any guard $g \in \llbracket V^G \rrbracket$, if there is a valuation $s' \in \llbracket V^L \rrbracket$ such that $\psi(s \cdot g, s')$, then the implementation is *required* to have a transition from state (ℓ, s) with guard g and action a to *at least some* state t' such that $\psi(s \cdot g, t')$. The requirement expressed by must-transitions cannot be formalized by standard modal transition systems, but fortunately, a generalization called disjunctive modal transition systems introduced in [9] can precisely capture these requirements. May-transitions target (as usual) only one state, but must-transitions branch to several possible next states (thus must-transitions are hypertransitions), with an existential interpretation: there must exist at least one transition with some target state which is an element from the set of target states of the hypertransition.

Definition 2. A modal specification with data (MSD) is a tuple

$$\mathbf{S} = (Sig, Loc, \ell^0, S^0, \longrightarrow_\diamond, \longrightarrow_\square)$$

where Sig, Loc, ℓ^0 are like in Def. 1, $S^0 \subseteq \llbracket V^L \rrbracket$ is a set of initial data states, and $\longrightarrow_\diamond, \longrightarrow_\square \subseteq Loc \times \llbracket V^L \rrbracket \times \llbracket V^G \rrbracket \times \Sigma \times (Loc \times \mathcal{P}_{\geq 1}(\llbracket V^L \rrbracket))$ are the may- (\diamond) and must- (\square) transition relations such that every may-transition targets a single state: if $(\ell, s, g, a, (\ell', S')) \in \longrightarrow_\diamond$ then $|S'| = 1$.

A state $(\ell, s) \in Loc \times \llbracket V^L \rrbracket$ is called *syntactically consistent* iff targets reachable by must-transitions are also reachable by may-transitions: if $(\ell, s, g, a, (\ell', S')) \in \longrightarrow_\square$ then $(\ell, s, g, a, (\ell', \{s'\})) \in \longrightarrow_\diamond$ for all $s' \in S'$. \mathbf{S} is syntactically consistent iff all states are syntactically consistent, and the set of initial data states is nonempty, i.e. $S^0 \neq \emptyset$.

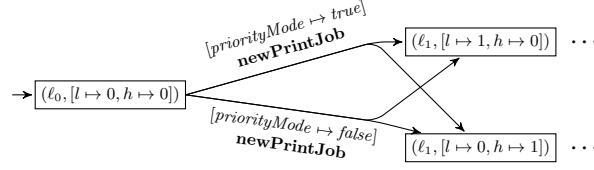


Fig. 2. Excerpt of the semantics of the abstract print server specification.

May-transitions $(\ell, s, g, a, (\ell', S')) \in \longrightarrow_{\diamond}$ are often written $(\ell, s) \xrightarrow{g, a}_{\diamond} (\ell', S')$, and similarly for must-transitions.

We can now define formally how a specification translates to its semantics in terms of an MSD. As already described above, the semantics of a may-transition of the specification is given by the set of may-transitions pointing to single admissible target states, and a must-transition gives rise to (must-)hypertransitions targeting all the admissible poststates.

Definition 3. The semantics of a specification $\mathbf{A} = (Sig, Loc, \ell^0, \varphi^0, E_{\diamond}, E_{\square})$ is given by the MSD $\langle \mathbf{A} \rangle_{\text{sem}} = (Sig, Loc, \ell^0, S^0, \longrightarrow_{\diamond}, \longrightarrow_{\square})$ where $S^0 = \llbracket \varphi^0 \rrbracket$ and the transition relations are defined as follows. For each $\ell, \ell' \in Loc$, $s, s' \in \llbracket V^L \rrbracket$, $g \in \llbracket V^G \rrbracket$, and $a \in \Sigma$:

- i. If $(\ell, a, \psi, \ell') \in E_{\diamond}$ and $\psi(s \cdot g, s')$ then $(\ell, s) \xrightarrow{g, a}_{\diamond} (\ell', \{s'\})$.
- ii. If $(\ell, a, \psi, \ell') \in E_{\square}$ and $\psi(s \cdot g, s')$ then $(\ell, s) \xrightarrow{g, a}_{\square} (\ell', \{t' \in \llbracket V^L \rrbracket \mid \psi(s \cdot g, t')\})$.

A specification \mathbf{A} is called *syntactically consistent* iff its semantics $\langle \mathbf{A} \rangle_{\text{sem}}$ is syntactically consistent. In the following we will always assume that specifications and MSD are syntactically consistent.

Example 2. An excerpt of the semantics of our abstract specification of the print server (see Fig. 1) can be seen Fig. 2. As before, we draw must-transitions with a solid arrow, and has an implicit set of may-transitions shadowing it which are not shown, i.e. for each target (ℓ, S') of a must-transition and each $s \in S'$ there is a may-transition with the same source state and with target state $(\ell, \{s\})$.

The first must-transition $(\ell_0, \text{newPrintJob}, (l)' + (h)' = 1, \ell_1) \in E_{\square}$ of the print server specification gives rise to the transitions shown in Fig. 2. Any new print job must be stored in either l or h but which one is not yet fixed by the specification. Thus in the semantics this is expressed as a disjunctive must-transition to the unique location ℓ_1 and the next possible data states $[l \mapsto 1, h \mapsto 0]$ and $[l \mapsto 0, h \mapsto 1]$.

A *refinement relation* allows to relate a concrete specification with an abstract specification. Refinement should satisfy the following substitutability property: If \mathbf{A} refines \mathbf{B} then replacing \mathbf{B} with \mathbf{A} in a context $C[\cdot]$ gives a specification $C[\mathbf{A}]$ refining $C[\mathbf{B}]$. Refinement will be a precongruence, i.e. it is compatible with the structural and logical operators on specifications in the above sense.

Our definition of refinement is based on modal refinement [10, 9] for (disjunctive) modal transition systems, where the may-transitions determine which actions are permitted in a refinement while the must-transitions specify which actions must be present in a refinement and hence in any implementation. We adapt it with respect to data states.

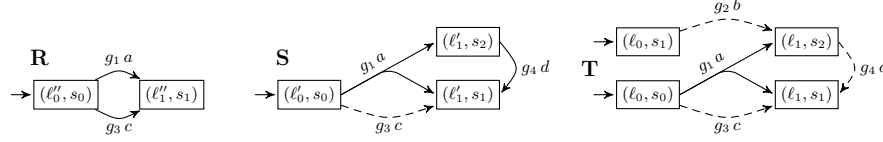


Fig. 3. Successive refinement of an MSD \mathbf{T} .

Example 3. We motivate our adaption of modal refinement to take into account data states with the help of a small example shown in Fig. 3. We draw may-transitions with a dashed arrow, and must-transitions with a solid arrow. Every must-transition has an implicit set of may-transitions shadowing it which are not shown. The MSD \mathbf{T} (to the right) has two initial states, both having ℓ_0 as the initial location. The must-transition starting from (ℓ_0, s_0) expresses that in any implementation there must be a transition leading to at least one of the states (ℓ_1, s_1) and (ℓ_1, s_2) . The MSD \mathbf{T} can be refined to the MSD \mathbf{S} (by dropping one may-transition and turning one may-transition to a must-transition), and then \mathbf{S} is refined by the MSD \mathbf{R} , by refining the must-transition $(\ell'_0, s_0, g_1, a, (\ell'_1, \{s_1, s_2\}))$ in \mathbf{S} to the must-transition $(\ell''_0, s_0, g_1, a, (\ell''_1, \{s_1\}))$ in \mathbf{R} , and by strengthening the transition with guard g_3 and action c to a must-transition.

Definition 4. Let $\mathbf{T}_1 = (Sig, Loc_1, \ell_1^0, S_1^0, \longrightarrow_{\diamond,1}, \longrightarrow_{\square,1})$ and $\mathbf{T}_2 = (Sig, Loc_2, \ell_2^0, S_2^0, \longrightarrow_{\diamond,2}, \longrightarrow_{\square,2})$ be MSD over the same signature $Sig = (\Sigma, V^L, V^G)$. A relation $R \subseteq Loc_1 \times Loc_2 \times \llbracket V^L \rrbracket$ is a refinement relation iff for all $(\ell_1, \ell_2, s) \in R$:

- i. Whenever $(\ell_1, s) \xrightarrow{g a}_{\diamond,1} (\ell'_1, \{s'\})$ then there exists $(\ell_2, s) \xrightarrow{g a}_{\diamond,2} (\ell'_2, \{t'\})$ such that $s' = t'$ and $(\ell'_1, \ell'_2, s') \in R$.
- ii. Whenever $(\ell_2, s) \xrightarrow{g a}_{\square,2} (\ell'_2, S'_2)$ then there exists $(\ell_1, s) \xrightarrow{g a}_{\square,1} (\ell'_1, S'_1)$ such that $S'_1 \subseteq S'_2$ and $(\ell'_1, \ell'_2, s') \in R$ for all $s' \in S'_1$.

We say that \mathbf{T}_1 refines \mathbf{T}_2 , written $\mathbf{T}_1 \leq_{\text{sem}} \mathbf{T}_2$, iff $S_1^0 \subseteq S_2^0$ and there exists a refinement relation R such that for any $s \in S_1^0$ also $(\ell_1^0, \ell_2^0, s) \in R$. A specification \mathbf{A}_1 refines another specification \mathbf{A}_2 , written $\mathbf{A}_1 \leq \mathbf{A}_2$, iff $\langle \mathbf{A}_1 \rangle_{\text{sem}} \leq_{\text{sem}} \langle \mathbf{A}_2 \rangle_{\text{sem}}$.

The refinement relation is a preorder on the class of all specifications. Refinement can be checked in polynomial time in the size of the state space of the MSD (for variables with finite domains). In general the domain may be infinite, or prohibitively large, so in Sect. 4 we revisit the question of refinement checking using abstraction techniques.

Example 4. The semantics of our abstract print server specification, shown in Fig. 2, can be refined as shown in Fig. 4. Now, both must-transitions point to the location ℓ_1 with the data state $[l \mapsto 1, h \mapsto 0]$ which means that any new incoming print job is assigned a low priority, independent of the uncontrolled variable *priorityMode*.

An MSD for which the conditions (1) $\longrightarrow_{\diamond} = \longrightarrow_{\square}$ and (2) $|S^0| = 1$ are satisfied, can be interpreted as (an abstraction of) an *implementation*: there are no design choices left open as (1) all may-transitions are covered by must-transitions and (2) there is only one initial data state possible. Any MSD for which the conditions (1) and (2) are satisfied, is called *transition system with data (TSD)* in the following. Note that TSD cannot be

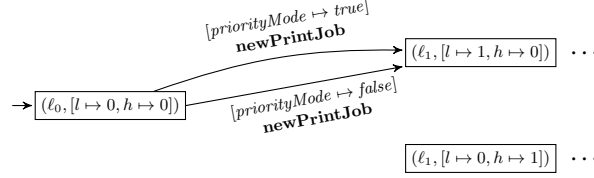


Fig. 4. Refinement of the MSD shown in Fig. 2.

strictly refined, i.e. for any TSD \mathbf{I} and any MSD \mathbf{S} with the same signature, $\mathbf{S} \leq_{\text{sem}} \mathbf{I}$ implies $\mathbf{I} \leq_{\text{sem}} \mathbf{S}$.

An implementation relation connects specifications to implementations (given as TSD) satisfying them. We can simply use refinement as the implementation relation. Given a specification \mathbf{A} and some TSD \mathbf{I} , we write $\mathbf{I} \models \mathbf{A}$ for $\mathbf{I} \leq_{\text{sem}} \langle \mathbf{A} \rangle_{\text{sem}}$, so our implementation \mathbf{I} is seen as the model which satisfies the property expressed by the specification \mathbf{A} . Now the set of implementations of a specification is the set of all its refining TSD: given a specification \mathbf{A} , we define $\text{Impl}(\mathbf{A}) = \{\mathbf{I} \mid \mathbf{I} \models \mathbf{A}\}$.

Our implementation relation \models immediately leads to the classical notion of consistency as existence of models. A specification \mathbf{A} is *consistent* iff $\text{Impl}(\mathbf{A})$ is non-empty. Consequently, as modal refinement is reflexive, any specification \mathbf{A} for which $\langle \mathbf{A} \rangle_{\text{sem}}$ is a TSD, is consistent.

By transitivity, modal refinement entails implementation set inclusion: for specifications \mathbf{A} and \mathbf{B} , if $\mathbf{A} \leq \mathbf{B}$ then $\text{Impl}(\mathbf{A}) \subseteq \text{Impl}(\mathbf{B})$. The relation $\text{Impl}(\mathbf{A}) \subseteq \text{Impl}(\mathbf{B})$ is sometimes called *thorough refinement* [11]. Just like for modal transition systems, thorough refinement does not imply modal refinement in general [12]. To establish equivalence we follow [13] by imposing a restriction on \mathbf{B} , namely that it is deterministic. An MSD is *deterministic* if

- (1) if $(\ell, s, g, a, (\ell', S')), (\ell, s, g, a, (\ell'', S'')) \in \longrightarrow_{\square}$ then $(\ell', S') = (\ell'', S'')$,
- (2) if $(\ell, s, g, a, (\ell', S')), (\ell, s, g, a, (\ell'', S'')) \in \longrightarrow_{\diamond} \cup \longrightarrow_{\square}$ then $\ell' = \ell''$.

A specification \mathbf{B} is *deterministic*, if the MSD $\langle \mathbf{B} \rangle_{\text{sem}}$ is deterministic. Note that for may-transitions, determinism only requires that for the same source state, guard and action, the transition leads to a unique next location. The reason why this is sufficient is that modal refinement explicitly distinguishes states by their data state part: two states (ℓ, s) and (ℓ', s') can only be related if their data state parts s, s' coincide.

Now, turning back to the relationship of modal refinement and inclusion of implementation sets (thorough refinement), we can prove the following theorem. Under the restriction of determinism of the refined (abstract) specification we can prove completeness of refinement. This theorem effectively means that modal refinement, as defined for MSD, is characterized by set inclusion of admitted implementations.

Theorem 1. *Let \mathbf{A} and \mathbf{B} be two specifications with the same signature such that \mathbf{B} is deterministic. Then $\mathbf{A} \leq \mathbf{B}$ if and only if $\text{Impl}(\mathbf{A}) \subseteq \text{Impl}(\mathbf{B})$.*

3 Compositional Reasoning

In this section we propose all the essential operators on specifications a good specification theory should provide. We will distinguish between structural and logical composition.

Structural composition mimics the classical composition of transition systems at the specification level. Logical composition allows to compute the intersection of sets of models and hence can be used to represent the conjunction of requirements made on an implementation. Furthermore we will introduce a quotient operator which is the dual operator to structural composition.

From now on, we assume that for any two specifications with the signatures $Sig_1 = (\Sigma_1, V_1^L, V_1^G)$ and $Sig_2 = (\Sigma_2, V_2^L, V_2^G)$, respectively, we can assume that $\Sigma_1 = \Sigma_2$ and $V_1^L \uplus V_1^G = V_2^L \uplus V_2^G$. This is not a limitation, as one can apply the constructions of [4] to equalize alphabets of actions and sets of variables.

Parallel composition. Two specifications \mathbf{A}_1 and \mathbf{A}_2 with $Sig_1 = (\Sigma_1, V_1^L, V_1^G)$, $Sig_2 = (\Sigma_2, V_2^L, V_2^G)$, respectively, are *composable* iff $V_1^L \cap V_2^L = \emptyset$. Then their signatures can be composed in a straightforward manner to the signature

$$Sig_1 \parallel Sig_2 = (\Sigma_1, V_1^L \cup V_2^L, (V_1^G \cup V_2^G) \setminus (V_1^L \cup V_2^L))$$

in which the set of controlled variables is the union of the sets of controlled variables of \mathbf{A}_1 and \mathbf{A}_2 , and the set of uncontrolled variables consists of all those uncontrolled variables of \mathbf{A}_1 and \mathbf{A}_2 which are controlled neither by \mathbf{A}_1 nor by \mathbf{A}_2 .

Definition 5. Let \mathbf{A}_1 and \mathbf{A}_2 be two composable specifications. The parallel composition of \mathbf{A}_1 and \mathbf{A}_2 is defined as the specification

$$\mathbf{A}_1 \parallel \mathbf{A}_2 = (Sig_1 \parallel Sig_2, Loc_1 \times Loc_2, (\ell_1^0, \ell_2^0), \varphi_1^0 \wedge \varphi_2^0, E_\diamond, E_\square)$$

where the transition relations E_\diamond and E_\square are the smallest relations satisfying the rules:

1. if $(\ell_1, a, \psi_1, \ell'_1) \in E_{\diamond,1}$ and $(\ell_2, a, \psi_2, \ell'_2) \in E_{\diamond,2}$ then
 $((\ell_1, \ell_2), a, \psi_1 \wedge \psi_2, (\ell'_1, \ell'_2)) \in E_\diamond$,
2. if $(\ell_1, a, \psi_1, \ell'_1) \in E_{\square,1}$ and $(\ell_2, a, \psi_2, \ell'_2) \in E_{\square,2}$ then
 $((\ell_1, \ell_2), a, \psi_1 \wedge \psi_2, (\ell'_1, \ell'_2)) \in E_\square$.

Composition of specifications, similar to the classical notion of modal composition for modal transition systems [10], synchronizes on matching shared actions and only yields a must-transition if there exist corresponding matching must-transitions in the original specifications. Composition is commutative (up to isomorphism) and associative. Our theory supports independent implementability of specifications, which is a crucial requirement for any compositional specification framework [14].

Theorem 2. Let $\mathbf{A}_1, \mathbf{A}_2, \mathbf{B}_1, \mathbf{B}_2$ be specifications such that \mathbf{A}_1 and \mathbf{B}_1 are composable. If $\mathbf{A}_1 \leq \mathbf{A}_2$ and $\mathbf{B}_1 \leq \mathbf{B}_2$, then $\mathbf{A}_1 \parallel \mathbf{B}_1 \leq \mathbf{A}_2 \parallel \mathbf{B}_2$.

The analog of parallel composition on the level of specifications is parallel composition \parallel_{sem} on the level of MSD which is a straightforward translation of the above symbolic rules. In fact one can prove that both parallel compositions \parallel and \parallel_{sem} are equivalent, i.e. that $\langle \mathbf{A}_1 \parallel \mathbf{A}_2 \rangle_{\text{sem}} = \langle \mathbf{A}_1 \rangle_{\text{sem}} \parallel_{\text{sem}} \langle \mathbf{A}_2 \rangle_{\text{sem}}$ for any two composable specifications $\mathbf{A}_1, \mathbf{A}_2$.

Remark 1. Interface theories based on transition systems labeled with input/output actions usually involve a notion of compatibility, which is a relation between interfaces determining whether two components can work properly together. Since the present theory does not have a notion of input/output it is enough to require that two components are composable, i.e. that their local variables do not overlap. A pessimistic input/output compatibility notion has been proposed in our previous work [15]. Optimistic input/output compatibility based on a game semantics allows computing all the environments in which two components can work together. Following our recent works in [16, 4], one can enrich labels of transitions in the present theory with input and output and apply the same game-based semantics in order to achieve an optimistic composition.

Syntactical consistency. Our next two specification operators, conjunction and quotient, may yield specifications which are *syntactically inconsistent*, i.e. either there is no legal initial data state or there are states with a must-transition but without corresponding may-transition.

In general, given a specification \mathbf{A} , syntactic consistency implies consistency, i.e. $\text{Impl}(\mathbf{A}) \neq \emptyset$, but in general, the reverse does not hold. However, every consistent specification can be “pruned” to a syntactically consistent one, by pruning backwards from all syntactically inconsistent states, removing states which have to reach some of the “bad” states. Pruning will be shown to preserve the set of implementations.

For a specification $\mathbf{A} = (\text{Sig}, \text{Loc}, \ell^0, \varphi^0, E_\diamond, E_\square)$, the pruning (or reduction) of \mathbf{A} , denoted by $\rho(\mathbf{A})$, is done as follows. Let $B : \text{Loc} \rightarrow \text{Pred}(V^L)$ be a mapping of locations to predicates over the local variables. We define a predecessor operation, iteratively computing all states that are forced to reach a “bad” state. Define a weakest precondition predicate, for $\psi \in \text{Pred}(V \uplus (V^L)'), \varphi \in \text{Pred}(V^L)$, by

$$\text{wp}_\psi[\varphi] \equiv \exists V^G. \circ\psi \wedge (\forall (V^L)'. \psi \Rightarrow (\varphi)') \quad (1)$$

which computes the largest set of local states such that there exists an uncontrolled state $g \in \llbracket V^G \rrbracket$ such that ψ maps to at least one next state, and all next states satisfy φ . Then

$$\text{predec}(B)(\ell) \equiv B(\ell) \vee \bigvee_{a \in \Sigma, \ell' \in \text{Loc}, \psi \in \text{Must}^a(\ell, \ell')} \text{wp}_\psi[B(\ell')]$$

and $\text{predec}^0(B) \equiv B$, $\text{predec}^{j+1}(B) \equiv \text{predec}(\text{predec}^j(B))$ for $j \geq 0$, and then finally $\text{predec}^*(B) \equiv \bigcup_{j \geq 0} \text{predec}^j(B)$. Define $\text{bad} : \text{Loc} \rightarrow \text{Pred}(V^L)$, for any $\ell \in \text{Loc}$, by

$$\text{bad}(\ell) \equiv \bigvee_{a \in \Sigma, \ell' \in \text{Loc}, \psi \in \text{Must}^a(\ell, \ell')} \exists V^G. \circ\psi \wedge \left(\forall (V^L)'. \psi \Rightarrow \bigwedge_{\psi' \in \text{May}^a(\ell, \ell')} \neg \psi' \right)$$

and thus $\text{bad}(\ell)$ is satisfied by a valuation $s \in \llbracket V^L \rrbracket$ iff there is a must-transition for which no choice of the next data state is permitted by the may-transitions.

In general, for infinite-domain variables, the computation of $\text{predec}^*(\text{bad})$ may not terminate. In [7], it was shown that reachability and related properties in well-structured transition systems with data values, that are monotonic transition systems with a well-quasi ordering on the set of data values, is decidable. This result can be used for specifications with infinite-domain variables to show that under these assumptions, there

is some $j \geq 0$ such that for all $\ell \in Loc$, $\llbracket \text{predec}^j(\text{bad})(\ell) \rrbracket = \llbracket \text{predec}^{j+1}(\text{bad})(\ell) \rrbracket$. In the following, for the specification operators conjunction and quotient (which may result in a syntactically inconsistent specification and hence need to be pruned) we assume that such a $j \geq 0$ exists.

The *pruning* $\rho(\mathbf{A})$ of \mathbf{A} is defined if $\varphi^0 \wedge \neg \text{predec}^j(\text{bad})(\ell^0)$ is consistent; and in this case, $\rho(\mathbf{A})$ is the specification $(\text{Sig}, Loc, \ell^0, \varphi^0 \wedge \neg \text{predec}^j(\text{bad})(\ell^0), E_\Diamond^\rho, E_\Box^\rho)$ where, for $\chi_{\text{good}} = \neg \text{predec}^j(\text{bad})$,

$$\begin{aligned} E_\Diamond^\rho &= \{ (\ell_1, a, \chi_{\text{good}}(\ell_1) \wedge \psi \wedge (\chi_{\text{good}}(\ell_2))', \ell_2) \mid (\ell_1, a, \psi, \ell_2) \in E_\Diamond \}, \\ E_\Box^\rho &= \{ (\ell_1, a, \chi_{\text{good}}(\ell_1) \wedge \psi \wedge (\chi_{\text{good}}(\ell_2))', \ell_2) \mid (\ell_1, a, \psi, \ell_2) \in E_\Box \}. \end{aligned}$$

Crucially the pruning operator has the expected properties:

Theorem 3. *Let \mathbf{A} be a deterministic, possibly syntactically inconsistent specification. Then $\rho(\mathbf{A})$ is defined if and only if \mathbf{A} is consistent. And if $\rho(\mathbf{A})$ is defined, then*

1. $\rho(\mathbf{A})$ is a specification (hence syntactically consistent),
2. $\rho(\mathbf{A}) \leq \mathbf{A}$,
3. $\text{Impl}(\mathbf{A}) = \text{Impl}(\rho(\mathbf{A}))$, and
4. for any specification \mathbf{B} , if $\mathbf{B} \leq \mathbf{A}$, then $\mathbf{B} \leq \rho(\mathbf{A})$.

Logical composition. Conjunction of two specifications yields the greatest lower bound with respect to modal refinement. Syntactic inconsistencies arise if one specification requires a behavior disallowed by the other.

Definition 6. *Let \mathbf{A}_1 and \mathbf{A}_2 be two specifications with the same signature Sig . The conjunction of \mathbf{A}_1 and \mathbf{A}_2 is defined as the possibly syntactically inconsistent specification*

$$\mathbf{A}_1 \wedge \mathbf{A}_2 = (\text{Sig}, Loc_1 \times Loc_2, (\ell_1^0, \ell_2^0), \varphi_1^0 \wedge \varphi_2^0, E_\Diamond, E_\Box)$$

where the transition relations E_\Diamond, E_\Box are the smallest relations satisfying the rules, for any $\ell_1 \in Loc_1, \ell_2 \in Loc_2, a \in \Sigma$,

1. If $(\ell_1, a, \psi_1, \ell'_1) \in E_{\Diamond,1}, (\ell_2, a, \psi_2, \ell'_2) \in E_{\Diamond,2}$, then $((\ell_1, \ell_2), a, \psi_1 \wedge \psi_2, (\ell'_1, \ell'_2)) \in E_\Diamond$,
2. If $(\ell_1, a, \psi_1, \ell'_1) \in E_{\Box,1}$, then $((\ell_1, \ell_2), a, \psi_1 \wedge (\bigvee_{\psi_2 \in \text{May}_2^a(\ell_2, \ell'_2)} \psi_2), (\ell'_1, \ell'_2)) \in E_\Box$,
3. If $(\ell_2, a, \psi_2, \ell'_2) \in E_{\Box,2}$, then $((\ell_1, \ell_2), a, \psi_2 \wedge (\bigvee_{\psi_1 \in \text{May}_1^a(\ell_1, \ell'_1)} \psi_1), (\ell'_1, \ell'_2)) \in E_\Box$,
4. If $(\ell_1, a, \psi_1, \ell'_1) \in E_{\Box,1}$ then $((\ell_1, \ell_2), a, \circ \psi_1 \wedge (\bigvee_{\psi_2 \in M} \neg \psi_2), (\ell'_1, \ell'_2)) \in E_\Box$,
where $M = \bigcup_{\ell'_2 \in Loc_2} \text{May}_2^a(\ell_2, \ell'_2)$,
5. If $(\ell_2, a, \psi_2, \ell'_2) \in E_{\Box,2}$ then $((\ell_1, \ell_2), a, \circ \psi_2 \wedge (\bigvee_{\psi_1 \in M} \neg \psi_1), (\ell_1, \ell'_2)) \in E_\Box$,
where $M = \bigcup_{\ell'_1 \in Loc_1} \text{May}_1^a(\ell_1, \ell'_1)$.

The first rule composes may-transitions (with the same action) by conjoining their predicates. Rule (2) and (3) express that any required behavior of \mathbf{A}_1 (\mathbf{A}_2 resp.), as long as it is allowed by \mathbf{A}_2 (\mathbf{A}_1 resp.), is also a required behavior in $\mathbf{A}_1 \wedge \mathbf{A}_2$. Rules (4) and (5) capture the case when a required behavior of \mathbf{A}_1 is not allowed by \mathbf{A}_2 . Conjunction is commutative and associative.

Refinement is a precongruence with respect to conjunction for deterministic specifications. Moreover, under the assumption of determinism, the conjunction construction yields the greatest lower bound with respect to modal refinement:

Theorem 4. *Let \mathbf{A} , \mathbf{B} , \mathbf{C} be specifications with the same signature and let \mathbf{A} and \mathbf{B} be deterministic. If $\mathbf{A} \wedge \mathbf{B}$ is consistent then*

1. $\rho(\mathbf{A} \wedge \mathbf{B}) \leq \mathbf{A}$ and $\rho(\mathbf{A} \wedge \mathbf{B}) \leq \mathbf{B}$,
2. $\mathbf{C} \leq \mathbf{A}$ and $\mathbf{C} \leq \mathbf{B}$ implies $\mathbf{C} \leq \rho(\mathbf{A} \wedge \mathbf{B})$,
3. $\text{Impl}(\rho(\mathbf{A} \wedge \mathbf{B})) = \text{Impl}(\mathbf{A}) \cap \text{Impl}(\mathbf{B})$.

Quotient as the dual operator to structural composition. The quotient operator allows factoring out behaviors from larger specifications. Given two specifications \mathbf{A} and \mathbf{B} the quotient of \mathbf{B} by \mathbf{A} , in the following denoted $\mathbf{B} \parallel \mathbf{A}$, is the most general specification that can be composed with \mathbf{A} and still refines \mathbf{B} .

In the following, we assume for the signatures $\text{Sig}_{\mathbf{A}} = (\Sigma, V_{\mathbf{A}}^L, V_{\mathbf{A}}^G)$ and $\text{Sig}_{\mathbf{B}} = (\Sigma, V_{\mathbf{B}}^L, V_{\mathbf{B}}^G)$ that $V_{\mathbf{A}}^L \subseteq V_{\mathbf{B}}^L$. The signature of the quotient $\mathbf{B} \parallel \mathbf{A}$ is then $\text{Sig}_{\mathbf{B} \parallel \mathbf{A}} = (\Sigma, V_{\mathbf{B} \parallel \mathbf{A}}^L, V_{\mathbf{B} \parallel \mathbf{A}}^G)$ with $V_{\mathbf{B} \parallel \mathbf{A}}^L = V_{\mathbf{B}}^L \setminus V_{\mathbf{A}}^L$ and $V_{\mathbf{B} \parallel \mathbf{A}}^G = V_{\mathbf{B}}^G \cup V_{\mathbf{A}}^L$. Note that, as said before, we restrict ourselves to the case where $V_{\mathbf{A}}^L \uplus V_{\mathbf{A}}^G = V_{\mathbf{B}}^L \uplus V_{\mathbf{B}}^G$.

It is unknown if in our general model of specifications a finite quotient exists. For specifications involving variables with finite domains only, a semantic quotient operation can be defined, which works on the (finite) semantics of \mathbf{A} and \mathbf{B} . As already noticed in previous works, e.g. [17], non-determinism is problematic for quotienting, and thus specifications are assumed to be deterministic. In our case, even when assuming deterministic specifications, the non-determinism with respect to the next local data state is still there: thus the quotient $\mathbf{B} \parallel \mathbf{A}$, when performing a transition, does not know the next data state of \mathbf{A} . However, due to our semantics, in which transitions are guarded by uncontrolled states, the quotient can always observe the current data state of \mathbf{A} . This extension of the usual quotient can be shown that it satisfies the following soundness and maximality property: Given MSD \mathbf{S} and \mathbf{T} such that \mathbf{S} is deterministic and $\mathbf{T} \parallel_{\text{sem}} \mathbf{S}$ is consistent, and assume a semantic pruning operator ρ_{sem} which is the straightforward translation of pruning ρ to the semantic level. Then $\mathbf{X} \leq_{\text{sem}} \rho_{\text{sem}}(\mathbf{T} \parallel_{\text{sem}} \mathbf{S})$ if and only if $\mathbf{S} \parallel_{\text{sem}} \mathbf{X} \leq_{\text{sem}} \mathbf{T}$ for any MSD \mathbf{X} .

Now our goal is to compute the quotient at the symbolic level of specifications. We do this for a restricted subclass of specifications in which each occurring transition predicate ψ is *separable*, meaning that ψ is equivalent to ${}^\circ\psi \wedge \psi^\circ$. Although this might seem as a serious restriction, we can often transform the transition systems with transition predicates of the form $(x)' = x + 1$ to a transition system with transition predicates which are separable and keep the same set of implementations. For instance, if we know that there are only finitely many possible values v_1, \dots, v_n for x in the current state, we can “unfold” the specification and replace the transition predicates $(x)' = x + 1$ by $(x)' = v_i$, for $1 \leq i \leq n$.

The symbolic quotient introduces two new locations, the universal state (univ) and an error state (\perp). In the universal state the quotient can show arbitrary behavior and is needed to obtain maximality, and the error state is a syntactically inconsistent state used to encode conflicting requirements. The state space of the quotient is given by $\text{Loc}_B \times \text{Loc}_A \times \text{Pred}(V_A^L)$, so every state stores not only the current location of B and A (like in [17]) but includes a predicate about the current possible data states of A . For notational convenience, for $\varphi \in \text{Pred}(V_1 \uplus V_2)$ and $\varphi_1 \in \text{Pred}(V_1)$, we write $\varphi \parallel \varphi_1$ for $(\forall V_1. \varphi_1 \Rightarrow \varphi) \in \text{Pred}(V_2)$.

Definition 7. Let A and B be two specifications such that $V_A^L \subseteq V_B^L$. The quotient of B by A is defined as the possibly syntactically inconsistent specification $B \parallel A = (\text{Sig}_{B \parallel A}, (\text{Loc}_B \times \text{Loc}_A \times \text{Pred}(V_A^L)) \cup \{\text{univ}, \perp\}, (\ell_B^0, \ell_A^0, \varphi_A^0), \varphi_B^0 \parallel \varphi_A^0, E_\diamond, E_\square)$ where the transition relations are given by, for all $a \in \Sigma$ and all $\xi_A \in \text{Pred}(V_A^L)$,

1. if $(\ell_B, a, \psi_B, \ell'_B) \in E_{\diamond, B}$ and $(\ell_A, a, \psi_A, \ell'_A) \in E_{\diamond, A}$, then
 $((\ell_B, \ell_A, \xi_A), a, \xi_A \wedge \circ \psi_B \wedge \circ \psi_A \wedge (\psi_B^\circ \parallel \psi_A^\circ), (\ell'_B, \ell'_A, \psi_A^\circ)) \in E_\diamond$,
2. if $(\ell_B, a, \psi_B, \ell'_B) \in E_{\square, B}$ and $(\ell_A, a, \psi_A, \ell'_A) \in E_{\square, A}$, then
 $((\ell_B, \ell_A, \xi_A), a, \xi_A \wedge \circ \psi_B \wedge \circ \psi_A \wedge (\psi_B^\circ \parallel \psi_A^\circ), (\ell'_B, \ell'_A, \psi_A^\circ)) \in E_\square$,
3. if $(\ell_B, a, \psi_B, \ell'_B) \in E_{\square, B}$ and $(\ell_A, a, \psi_A, \ell'_A) \in E_{\square, A}$, then
 $((\ell_B, \ell_A, \xi_A), a, \xi_A \wedge \circ \psi_B \wedge \circ \psi_A \wedge \neg(\psi_B^\circ \parallel \psi_A^\circ), \perp) \in E_\square$,
4. if $(\ell_B, a, \psi_B, \ell'_B) \in E_{\square, B}$, then
 $((\ell_B, \ell_A, \xi_A), a, \xi_A \wedge \circ \psi_B \wedge \neg(\bigvee_{\psi_A \in M} \circ \psi_A), \perp) \in E_\square$
where $M = \bigcup_{\ell'_A \in \text{Loc}_A} \text{Must}_A^a(\ell_A, \ell'_A)$,
5. $((\ell_B, \ell_A, \xi_A), a, \neg \xi_A, \text{univ}) \in E_\diamond$,
6. $((\ell_B, \ell_A, \xi_A), a, \xi_A \wedge \neg(\bigvee_{\psi_A \in M} \circ \psi_A), \text{univ}) \in E_\diamond$
where $M = \bigcup_{\ell'_A \in \text{Loc}_A} \text{May}_A^a(\ell_A, \ell'_A)$,
7. $(\text{univ}, a, \text{true}, \text{univ}) \in E_\diamond$,
8. $(\perp, a, \text{true}, \perp) \in E_\square$.

Rules (1) and (2) capture the cases when both A and B can perform a may- and must-transition, respectively. Rules (3) and (4) capture any inconsistencies which can arise if for a must-transition in B there is no way to obtain a must-transition by composition of the quotient with A . In order to obtain maximality, we add a universal state univ in which the behavior of the quotient is not restricted (rules (5)–(7)). Finally, the rule (8) makes the error state syntactically inconsistent.

Since we only have finitely many transition predicates ψ_A in A , and they are all separable, the set of locations $(\text{Loc}_B \times \text{Loc}_A \times (\{\psi_A^\circ \mid \psi_A \text{ occurring in } A\} \cup \{\varphi_A^0\})) \cup \{\text{univ}, \perp\}$ of $B \parallel A$ is also finite. Thus we can construct the symbolic quotient in a finite number of steps, starting in the initial state $(\ell_B^0, \ell_A^0, \varphi_A^0)$, and iteratively constructing the transitions. Soundness and maximality of the quotient follows from the following theorem.

Theorem 5. Let A and B be specifications such that $V_A^L \subseteq V_B^L$, all transition predicates of A and B are separable, A is deterministic and $B \parallel A$ is consistent. Then for any specification C such that $\text{Sig}_C = \text{Sig}_{B \parallel A}$, $C \leq \rho(B \parallel A)$ if and only if $A \parallel C \leq B$.

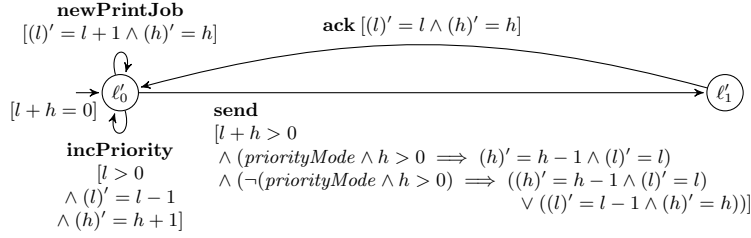


Fig. 5. Refined print server specification **Q**.

4 Predicate Abstraction for Verification of Refinement

We now switch our focus to the problem of deciding whether a specification **A** refines another specification **B** (which reduces to checking $\langle \mathbf{A} \rangle_{\text{sem}} \leq_{\text{sem}} \langle \mathbf{B} \rangle_{\text{sem}}$). As soon as domains of variables are infinite, $\langle \mathbf{A} \rangle_{\text{sem}}$ and $\langle \mathbf{B} \rangle_{\text{sem}}$ may be MSD with infinitely many states and transitions. In this case, this problem is known to be undecidable in general. Thus we propose to resort to predicate abstraction techniques [18]. Given two specifications **A** and **B** we derive over- and under-approximations \mathbf{A}^o and \mathbf{B}^u which are guaranteed to be *finite* MSD. Then, we show that $\mathbf{A}^o \leq_{\text{sem}} \mathbf{B}^u$ implies $\mathbf{A} \leq \mathbf{B}$.

Example 5. Fig. 5 shows a print server specification **Q** which we will show is a refinement of the abstract specification **P** in Fig. 1. The behavior of the print server is now fixed for any number of print jobs. Moreover, the send transition has been refined such that depending on the priority mode (provided by the environment of the print server) a job with high priority (in case *priorityMode* is true) or a job with low priority (otherwise) is chosen next.

Given a specification $\mathbf{A} = (Sig, Loc, \ell^0, \varphi^0, \longrightarrow_\diamond, \longrightarrow_\square)$ with $Sig = (\Sigma, V^L, V^G)$, we partition the local state space and the uncontrolled state space using finitely many predicates $\phi_1, \phi_2, \dots, \phi_N \in Pred(V^L)$ and $\chi_1, \chi_2, \dots, \chi_M \in Pred(V^G)$. We fix these predicates in the following to simplify the presentation. The signature of the abstraction is then given by $Sig_{abstr} = (\Sigma, V_{abstr}^L, V_{abstr}^G)$, where $V_{abstr}^L = \{x_1, x_2, \dots, x_N\}$ and $V_{abstr}^G = \{y_1, y_2, \dots, y_M\}$. All variables x_i, y_j have Boolean domain. A variable x_i (y_j) encodes whether the predicate ϕ_i (χ_j) holds or not.

Any abstract state $\nu \in \llbracket V_{abstr}^L \rrbracket$ is a conjunction of predicates $\bigwedge_{i=1}^N \phi_i^{\nu(x_i)}$, where $\phi_i^{\nu(x_i)} = \phi_i$ if $\nu(x_i) = 1$, else $\phi_i^{\nu(x_i)} = \neg \phi_i$. Further, a set of abstract states $N \subseteq \llbracket V_{abstr}^L \rrbracket$ corresponds to $\bigvee_{\nu \in N} \nu$. Similarly for any $\omega \in \llbracket V_{abstr}^G \rrbracket$ and for $M \subseteq \llbracket V_{abstr}^G \rrbracket$.

The transition relation of the over-approximation expands the allowed behaviors and limits the required behaviors. Dually, the under-approximation will further restrict the allowed behavior and add more required transitions. In other words, over-approximation is an *existential* abstraction on may-transitions and *universal* abstraction on must-transitions; dually for the under-approximation.

Formally, the *over-approximation* \mathbf{A}^o of **A** is defined by the finite TSD $(Sig_{abstr}, Loc, \ell^0, S_{abstr}^0, \longrightarrow_{\diamond, abstr}, \longrightarrow_{\square, abstr})$, where the initial abstract state contains all partitions overlapping with concrete initial states $S_{abstr}^0 = \{\nu \in \llbracket V_{abstr}^L \rrbracket \mid \exists V^L. \nu \wedge \varphi^0\}$,

and the abstract transition relations are derived as follows. For all $\ell, \ell' \in Loc$, $a \in Act$, $\nu, \dot{\nu} \in \llbracket V_{abstr}^L \rrbracket$, $\omega \in \llbracket V_{abstr}^G \rrbracket$,

- i. If $\exists V. \exists (V^L)'. \nu \wedge \omega \wedge (\bigvee_{\psi \in May^a(\ell, \ell')} \psi) \wedge (\dot{\nu})'$, then $(\ell, \nu) \xrightarrow{\omega^a}_{\diamond, abstr} (\ell', \{\dot{\nu}\})$, so there is a may-transition between partitions in the abstraction if there was a may-transition between any states in these partitions in the concrete system.
- ii. Whenever, for some $N \subseteq \llbracket V_{abstr}^L \rrbracket$, the predicate

$$\forall V. \nu \wedge \omega \Rightarrow \bigvee_{\psi \in Must^a(\ell, \ell')} \psi \wedge (\forall (V^L)'. \psi \Rightarrow (N)') \quad (2)$$

is true and N is minimal with respect to this property, then $(\ell, \nu) \xrightarrow{\omega^a}_{\square, abstr} (\ell', N)$.

For the *under-approximation* \mathbf{B}^u of \mathbf{B} , we assume that every transition predicate ψ on a must-transition must be separable (see page 12). Moreover, in order to soundly capture must-transitions, we must be able to exactly describe the target set of (concrete) local states by a union of abstract states; so for any $(\ell, a, \psi, \ell') \in E_{\square, \mathbf{B}}$, there exists a set $N \subseteq \llbracket V_{abstr}^L \rrbracket$ such that $\forall (V^L)'. \psi^\circ \Leftrightarrow (N)'$. The under-approximation \mathbf{B}^u is the finite TSD $(Sig_{abstr}^0, Loc, \ell^0, S_{abstr}^0, \xrightarrow{\cdot}_{\diamond, abstr}, \xrightarrow{\cdot}_{\square, abstr})$, where $S_{abstr}^0 = \{\nu \in \llbracket V_{abstr}^L \rrbracket \mid \forall V^L. \nu \Rightarrow \varphi^0\}$, and for all $\ell, \ell' \in Loc$, $a \in Act$, $\nu, \dot{\nu} \in \llbracket V_{abstr}^L \rrbracket$, $\omega \in \llbracket V_{abstr}^G \rrbracket$,

- i. If $\forall V. \forall (V^L)'. \nu \wedge \omega \wedge (\dot{\nu})' \Rightarrow \bigvee_{\psi \in May^a(\ell, \ell')} \psi$ then $(\ell, \nu) \xrightarrow{\omega^a}_{\diamond, abstr} (\ell', \{\dot{\nu}\})$,
- ii. For every $(\ell, a, \psi, \ell') \in E_{\square, \cdot}$, if $\exists V. \nu \wedge \omega \wedge \psi^\circ$, then $(\ell, \nu) \xrightarrow{\omega^a}_{\square, abstr} (\ell', N)$ where $N \subseteq \llbracket V_{abstr}^L \rrbracket$ such that $\forall (V^L)'. \psi^\circ \Leftrightarrow (N)'$.

Correctness of the abstraction follows from the following theorem.

Theorem 6. $\mathbf{A}^\circ \leq_{\text{sem}} \mathbf{B}^u$ implies $\mathbf{A} \leq \mathbf{B}$.

Example 6. Fig. 6 and Fig. 7 are over- and under-approximations of \mathbf{Q} and \mathbf{P} , respectively. The MSD represent abstractions w.r.t. the predicates $\phi_{0,0} \equiv h = l = 0$, $\phi_{0,1} \equiv l = 0 \wedge h = 1$, $\phi_{1,0} \equiv l = 1 \wedge h = 0$, and $\phi_{>1} \equiv h + l > 1$ for the controlled variables l and h , and $\omega_1 \equiv \text{priorityMode}$, $\omega_2 \equiv \neg \text{priorityMode}$ for the uncontrolled variable *priorityMode*. Note that all transition predicates in \mathbf{P} are separable, and all possible (concrete) poststates can be precisely captured by the predicates $\phi_{0,0}, \phi_{0,1}, \phi_{1,0}, \phi_{>1}$. For better readability we have omitted most of the guards ω_1, ω_2 , i.e. every transition without guard stands for two transitions with the same action, source and target state(s), and with ω_1 and ω_2 as guard, respectively. Moreover, the state $(\ell_3, \phi_{0,0} \vee \phi_{0,1} \vee \phi_{1,0} \vee \phi_{>1})$ is a simplified notation which represents all the states (ℓ_3, ϕ) with $\phi \in \{\phi_{0,0}, \phi_{0,1}, \phi_{1,0}, \phi_{>1}\}$ and all may-transitions leading to it lead to each of the states, and the may-loop stands for all the transitions between each of the states. Obviously, $\mathbf{Q}^\circ \leq_{\text{sem}} \mathbf{P}^u$, and from Thm. 6 it follows that $\mathbf{Q} \leq \mathbf{P}$.

Even though this abstraction technique requires separability of predicates, it is applicable to a larger set of specifications. Sometimes, as already described in the previous section, transitions with non-separable predicates can be replaced by finite sets of transitions to achieve separability, without changing the semantics of the specification. Automatic procedures for generation of predicates are subject of future work. Finally, our abstraction also supports compositional reasoning about parallel composition in the following sense:

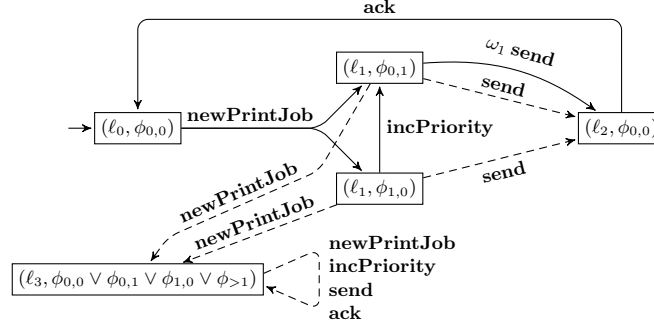


Fig. 6. Under-approximation P^u .

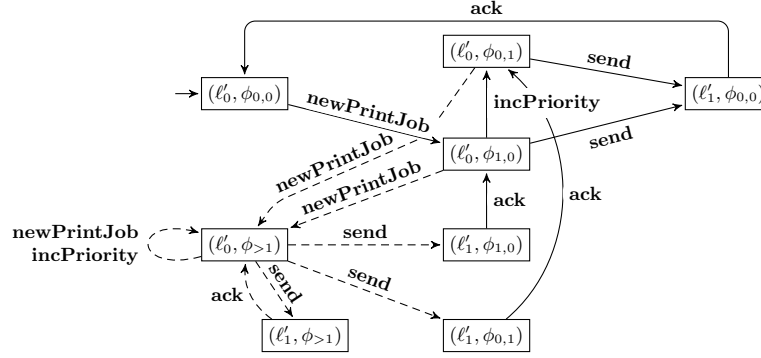


Fig. 7. Over-approximation Q^o .

Theorem 7. Let \mathbf{A} and \mathbf{B} be two composable specifications, and $V_{\mathbf{A} \parallel \mathbf{B}}^G = (V_{\mathbf{A}}^G \cup V_{\mathbf{B}}^G) \setminus (V_{\mathbf{A}}^L \uplus V_{\mathbf{B}}^L)$. Let $E_{\mathbf{A}} \subseteq \text{Pred}(V_{\mathbf{A}}^L)$, $E_{\mathbf{B}} \subseteq \text{Pred}(V_{\mathbf{B}}^L)$, and $F \subseteq \text{Pred}(V_{\mathbf{A} \parallel \mathbf{B}}^G)$ be sets of predicates partitioning the respective data states.

\mathbf{A} is approximated w.r.t. $E_{\mathbf{A}}$ for $V_{\mathbf{A}}^L$, and $E_{\mathbf{B}} \cup F$ for $V_{\mathbf{A}}^G = V_{\mathbf{A} \parallel \mathbf{B}}^G \uplus V_{\mathbf{B}}^L$ and similarly, \mathbf{B} is approximated w.r.t. $E_{\mathbf{B}}$ and $E_{\mathbf{A}} \cup F$. Finally, $\mathbf{A} \parallel \mathbf{B}$ is approximated w.r.t. $E_{\mathbf{A}} \cup E_{\mathbf{B}}$ for $V_{\mathbf{A} \parallel \mathbf{B}}^L = V_{\mathbf{A}}^L \uplus V_{\mathbf{B}}^L$, and F for $V_{\mathbf{A} \parallel \mathbf{B}}^G$. We assume that each predicate, in any abstraction of \mathbf{A} , \mathbf{B} , or $\mathbf{A} \parallel \mathbf{B}$, are encoded with the same variable.

Then $(\mathbf{A} \parallel \mathbf{B})^o \leq_{\text{sem}} \mathbf{A}^o \parallel_{\text{sem}} \mathbf{B}^o$, and $\mathbf{A}^u \parallel_{\text{sem}} \mathbf{B}^u \leq_{\text{sem}} (\mathbf{A} \parallel \mathbf{B})^u$.

This result allows reusing abstractions of individual components in a continued development and verification process. For instance, if we want to verify $\mathbf{A} \parallel \mathbf{B} \leq \mathbf{C}$ then we can compute (or reuse) the less complex abstractions \mathbf{A}^o and \mathbf{B}^o . Thm. 7 implies then that from $\mathbf{A}^o \parallel_{\text{sem}} \mathbf{B}^o \leq_{\text{sem}} \mathbf{C}^u$ we can infer $\mathbf{A} \parallel \mathbf{B} \leq \mathbf{C}$.

5 Related work

The main difference to related approaches based on modal process algebra taking data states into account, e.g. [19] is that they cannot naturally express logical and structural composition in the same formalism. A comparison between modal specifications and

other theories such as interface automata [20] and process algebra [2] can be found in [3]. In [8], the authors introduced sociable interfaces, that is a model of I/O automata [21] equipped with a data and a game-based semantics. While their communication primitives are richer, sociable interfaces do not encompass any notion of logical composition and quotient, and their refinement is based on an alternating simulation.

Transition systems enriched with predicates are used, for instance, in the approach of [22, 23] where they use symbolic transition systems (STS), but STS do not support modalities and loose data specifications as they focus more on model checking than on the (top down) development of concurrent systems by refinement.

In [15] modal I/O automata has been extended by pre- and postconditions viewed as contracts, however, only semantics in terms of sets of implementations have been defined (implementations with only input actions correspond to our TSD). Modal refinement as defined in [15] is coarser than in this paper, and moreover, neither conjunction nor a quotient operation are defined.

6 Conclusion

We have proposed a specification theory for reasoning about components with rich data state. Our formalism, based on modal transition systems, supports: refinement checking, consistency checking with pruning of inconsistent states, structural and logical composition, and a quotient operator. We have defined symbolic representations of the operators and have shown that they are equivalent to the semantic definitions—this allows for automatic analysis of specifications. We have also presented a predicate abstraction technique for modal specifications with data. We believe that this work is a significant step towards practical use of specification theories based on modal transition systems. The ability to reason about data domains permits the modeling of industrial case studies.

In the future, we intend to develop larger case studies. Furthermore, we would like to extend the formalism with more complex communication patterns and to investigate in which cases we can still obtain all the operators on specifications, in particular the quotient operator. We are also planning to implement the theory in the MIO Workbench [24, 25, 26], a verification tool for modal input/output interfaces.

References

- [1] Larsen, K.G.: Modal specifications. In Sifakis, J., ed.: *Automatic Verification Methods for Finite State Systems*. Volume 407 of *Lecture Notes in Computer Science*, Springer (1989)
- [2] Milner, R.: *A Calculus of Communicating Systems*. Volume 92 of *Lecture Notes in Computer Science*. Springer (1980)
- [3] Nyman, U.: *Modal Transition Systems as the Basis for Interface Theories and Product Lines*. PhD thesis, Department of Computer Science, Aalborg University (October 2008)
- [4] Raclet, J.B., Badouel, E., Benveniste, A., Caillaud, B., Legay, A., Passerone, R.: Modal interfaces: unifying interface automata and modal specifications. In Chakraborty, S., Halbwachs, N., eds.: *EMSOFT*, ACM (2009)
- [5] Abdulla, P.A., Bouajjani, A., d’Orso, J.: Monotonic and downward closed games. *J. Log. Comput.* **18**(1) (2008) 153–169

- [6] Godefroid, P., Huth, M., Jagadeesan, R.: Abstraction-based model checking using modal transition systems. In: CONCUR. (2001) 426–440
- [7] Abdulla, P.A., Cerans, K., Jonsson, B., Tsay, Y.K.: Algorithmic analysis of programs with well quasi-ordered domains. *Inf. Comput.* **160**(1-2) (2000) 109–127
- [8] de Alfaro, L., da Silva, L.D., Faella, M., Legay, A., Roy, P., Sorea, M.: Sociable interfaces. In Gramlich, B., ed.: *FroCos*. Volume 3717 of *Lecture Notes in Computer Science.*, Springer (2005) 81–105
- [9] Larsen, K.G., Xinxin, L.: Equation solving using modal transition systems. In: *LICS*, IEEE Computer Society (1990) 108–117
- [10] Larsen, K.G., Thomsen, B.: A modal process logic. In: *LICS*, IEEE Computer Society (1988) 203–210
- [11] Antonik, A., Huth, M., Larsen, K.G., Nyman, U., Wasowski, A.: Complexity of decision problems for mixed and modal specifications. In: *FoSSaCS 2008*. Volume 4962 of *Lecture Notes in Computer Science.*, Springer (2008) 112–126
- [12] Larsen, K.G., Nyman, U., Wasowski, A.: On modal refinement and consistency. In Caires, L., Vasconcelos, V.T., eds.: *CONCUR*. Volume 4703 of *Lecture Notes in Computer Science.*, Springer (2007) 105–119
- [13] Benes, N., Kretínský, J., Larsen, K.G., Srba, J.: On determinism in modal transition systems. *Theor. Comput. Sci.* **410**(41) (2009) 4026–4043
- [14] de Alfaro, L., Henzinger, T.A.: Interface theories for component-based design. In: *Embedded Software, First International Workshop, EMSOFT 2001, Tahoe City, CA, USA, October, 8-10, 2001, Proceedings*. Volume 2211 of *Lecture Notes in Computer Science.*, Springer (2001) 148–165
- [15] Bauer, S.S., Hennicker, R., Wirsing, M.: Interface theories for concurrency and data. *Theor. Comput. Sci.* (2011) To appear.
- [16] Larsen, K.G., Nyman, U., Wasowski, A.: Modal I/O automata for interface and product line theories. In Nicola, R.D., ed.: *ESOP*. Volume 4421 of *Lecture Notes in Computer Science.*, Springer (2007) 64–79
- [17] Raclet, J.B.: Residual for component specifications. *Electr. Notes Theor. Comput. Sci.* **215** (2008) 93–110
- [18] Graf, S., Saïdi, H.: Construction of abstract state graphs with pvs. In Grumberg, O., ed.: *CAV*. Volume 1254 of *Lecture Notes in Computer Science.*, Springer (1997) 72–83
- [19] van de Pol, J., Espada, M.V.: Modal Abstractions in μ CRL. In Rattray, C., Maharaj, S., Shankland, C., eds.: *AMAST*. Volume 3116 of *Lecture Notes in Computer Science.*, Springer (2004) 409–425
- [20] de Alfaro, L., Henzinger, T.A.: Interface automata. *SIGSOFT Softw. Eng. Notes* **26** (September 2001) 109–120
- [21] Lynch, N., Tuttle, M.R.: An introduction to Input/Output automata. *CWI-quarterly* **2**(3) (1989)
- [22] Fernandes, F., Royer, J.C.: The STSLib project: Towards a formal component model based on STS. *Electr. Notes Theor. Comput. Sci.* **215** (2008) 131–149
- [23] Barros, T., Ameur-Boulifa, R., Cansado, A., Henrio, L., Madelaine, E.: Behavioural models for distributed fractal components. *Annales des Télécommunications* **64**(1-2) (2009) 25–43
- [24] Bauer, S.S., Mayer, P., Schroeder, A., Hennicker, R.: On weak modal compatibility, refinement, and the mio workbench. In Esparza, J., Majumdar, R., eds.: *TACAS*. Volume 6015 of *Lecture Notes in Computer Science.*, Springer (2010) 175–189
- [25] Bauer, S.S., Mayer, P., Legay, A.: MIO Workbench: A Tool for Compositional Design with Modal Input/Output Interfaces. In: *Automated Technology for Verification and Analysis (ATVA 2011)*. (2011) Accepted for publication.
- [26] MIO Workbench: <http://www.miowb.net/>