Automata-Theoretic Decision Procedures for Information Logics

Stéphane Demri[1]* and Ulrike Sattler[2]

[1] Lab. Spécification et Vérification, ENS de Cachan & CNRS UMR 8643
61, av. Pdt. Wilson, 94235 Cachan Cedex, France
email: demri@lsv.ens-cachan.fr
[2] TU Dresden, Institut für Theoretische Informatik
01062 Dresden, Germany
email: sattler@tcs.inf.tu-dresden.de

Abstract. Automata-theoretic decision procedures for solving model-checking and satisfiability problems for temporal, dynamic, and description logics have flourished during the past decades. In the paper we define an ExpTime decision procedure based on the emptiness problem of Büchi automata on infinite trees for the very expressive information logic SIM designed for reasoning about information systems. This logic involves modal parameters satisfying certain properties to capture the relevant properties of information systems, and provides nominals at the formula level, Boolean expressions and nominals at the modal level, an implicit intersection operation for relations, and a universal modality. The original combination of known techniques allows us to solve the open question related to the ExpTime-completeness of SIM. Furthermore, we discuss how variants of SIM can be treated similarly although the decidability status of some of them is still unknown.

Keywords: computational complexity, Büchi tree automaton, information logic, hybrid logic

1 Introduction

From logic to automata. After the works of Büchi and Rabin [Büc62,Rab69], various classes of automata turned out to be well-suited to solve decision procedures for logical problems, including some for temporal logics (see e.g., [VW94,Var97,KVW00]), for the μ-calculus and its fragments (see e.g., [EJ99,SE89,VW86,EJS01,Var98]), and for description logics (see e.g., [CDGL99,CGL02]) to quote three families of logics. For instance, translating formulae in temporal logics to automata is a standard approach for implementing model checking, see e.g., the model-checking tool SPIN [Hol97]. More recently, such techniques have also been applied successfully in [LS01] to fragments of the Boolean modal logic BML introduced in

* We acknowledge financial support from the RWTH Aachen, Germany for a visit of S. Demri at the RWTH in february 2002.

[GP90] and to hybrid full μ-calculus in [SV01] (see also [KSV02]) opening an avenue to design similar decision procedures for other hybrid logics [Bla00b]. In this paper, we will use automata-theoretic decision procedures to prove complexity results for *information logics*.

Information logics. Such logics were introduced in [Paw81], and we refer the reader to [Orło98,DO02] for a comprehensive survey on information logics. Information logics are designed to model and reason about *information systems*. To this purpose, information logics provide a formal specification language to talk about relations in information systems. An *information system* S is defined as a structure $S = \langle OB, AT \rangle$ such that OB is a non-empty set of *objects*, AT is a non-empty set of *attributes*, and every attribute $a \in AT$ is a mapping $a : OB \to \mathcal{P}(VAL_a) \setminus \{\emptyset\}$, where VAL_a is a non-empty set of *values*. For every object x and for every attribute a, $a(x)$ can be read as the set of possible values of the attribute a for the object x. In that setting, various derived relations between objects can be defined. We recall some of them below (see e.g. [Orło98]). For all $x_1, x_2 \in OB$, for every $A \subseteq AT$,

(I) $\langle x_1, x_2 \rangle \in ind_A$ iff for every $a \in A$, $a(x_1) = a(x_2)$ (indiscernability);
(II) $\langle x_1, x_2 \rangle \in fin_A$ iff for every $a \in A$, $a(x_1) \subseteq a(x_2)$ (forward inclusion);
(III) $\langle x_1, x_2 \rangle \in bin_A$ iff for every $a \in A$, $a(x_2) \subseteq a(x_1)$ (backward inclusion);
(IV) $\langle x_1, x_2 \rangle \in sim_A$ iff for every $a \in A$, $a(x_1) \cap a(x_2) \neq \emptyset$ (similarity).

$\langle x_1, x_2 \rangle \in ind_A$ can be read as follows: the objects x_1 and x_2 cannot be distinguished modulo the set A of attributes. Similarly, $\langle x_1, x_2 \rangle \in sim_A$ iff x_1 and x_2 are similar modulo A. The other relations fin_A and bin_A admit a reading in a similar vein.

Given an information system $S = \langle OB, AT \rangle$, we can define a structure $\langle OB, (\mathcal{R}_A)_{A \subseteq AT} \rangle$, where $(\mathcal{R}_A)_{A \subseteq AT}$ is a family of relations derived from S (see e.g., the above clauses (I)-(IV)). In a more abstract setting, an *information frame* is a pair $\langle W, (\mathcal{R}_P)_{P \subseteq PAR} \rangle$ such that W and PAR are non-empty sets and $(\mathcal{R}_P)_{P \subseteq PAR}$ is a family of binary relations indexed by subsets of PAR. An information logic is defined as a multi-modal logic characterised by a class of information frames. Since the relations derived from information systems satisfy certain properties, the information frames usually satisfy additional conditions. For example, it is not hard to see that, for every $\mathcal{R} \in \{ind, fin, bin, sim\}$, we have

$$\mathcal{R}_{P \cup Q} = \mathcal{R}_P \cap \mathcal{R}_Q \text{ for all } P, Q \subseteq PAR, \text{ and} \tag{1}$$

$$\mathcal{R}_\emptyset \text{ is the cartesian product of the domain.} \tag{2}$$

Moreover, every relation $\mathcal{R}_P$ satisfies certain local conditions: for instance, the indiscernability relations are equivalence relations, and the similarity relations are reflexive and symmetric. The first information logic has been introduced in [OP81] and many others appeared later (see e.g., [Vak91,Bal96,Kon97,Vak98,Ste98]). Most information logics include further expressive means such a nominals whose combination with the intersection of modal parameters and the universal modality are known to make reasoning rather complex. In the following, we will concentrate ourselves on one such logic, SIM, and show how a combination of a suitable normal form for modal expressions, a tree model property, and tree automata can be used to overcome this difficulty and define an optimal decision procedure for SIM satisfiability.

The logic SIM. Among the class of information logics, the logic SIM introduced in [Kon97] plays a special role since it provides various expressive ingredients: an implicit universal modality, nominals at both levels of formulae and modal expressions, and Boolean operators in modal expressions. This highly expressive logic was designed to represent and reason about relevant properties of similarity relations sim_A. Additionally, it turned out that this logic is also well-suited for the internalization of deduction in proof systems [Kon97] (see also [Dem99b,Tza99,Bla00a,Sel01]). A SIM-model is a Kripke structure with reflexive and symmetric accessibility relations $(sim_P)_{P \subseteq PAR}$ as abstract counterparts of the similarity relations derived from information systems. Hence the relations in $(sim_P)_{P \subseteq PAR}$ interact according to the above conditions (1) and (2).

Our contribution. In this paper, we show that the satisfiability problem for the logic SIM is ExpTime-complete. The ExpTime lower bound is a consequence of more general results since SIM contains a universal modal connective with a family of B modal connectives (see e.g., [Spa93,CL94,Hem96]). The ExpTime upper bound is established by an exponential reduction into the emptiness problem for Büchi automata on infinite trees that is known to be in PTime (see e.g., [VW86,EJ88]). As mentioned previously, this technique is nowadays standard for logics of programs, but it has never been applied to information logics. Indeed, relative information logics contain features that are not traditionally present in most logics of programs (e.g., the presence of nominals on the formula and modal level, and Boolean operators in modal expressions). Recently in [LS01], the fragment of Boolean modal logic (BML) with only the complement operator $\neg$ on relations is shown to be in ExpTime by a reduction into the emptiness problem for Büchi automata on infinite trees. In [SV01], such an upper bound is also established for the hybrid full μ-calculus by reduction into the emptiness problem for parity alternating automata on infinite trees. A combination of such recent results and an appropriate extension to handle intersection and nominals at the level of modal expressions allows us to prove the ExpTime lower bound for SIM. Observe that, in [LS01], it is already shown that the fragment of BML with only complement and intersection is NExpTime-hard even if no $\neg$ is in the scope of $\cap$ and no $\cap$ is in the scope of $\neg$ (see e.g., [LS01, Figure 1]). We show that the information logic SIM is of a lower complexity: it is in ExpTime even though it provides full Boolean operators in modal expressions.[1] This lower complexity is due to the restrictions of the semantics of the modal expressions

[1] Please note that this difference is not due to an bounded/unbounded number of atomic modal expressions—this number is unbounded in both logics mentioned.

designed to represent derived relations in information systems. A remarkable side-effect of our result for SIM is the following. The multi-modal logic with a universal modality $[U]$, modal connectives of the form $[c_1 \cap \cdots \cap c_n]$, where $\cap$ is interpreted as intersection on binary relations, each c_i is interpreted as a reflexive and symmetric relation, and the logic contains propositional variables and nominals can be shown to have an ExpTime-complete satisfiability problem.

Figure 1 shows the presence of ingredients in the logics SIM, BML, and in the hybrid μ-calculus. The operators $\neg$ and $\cap$ refers respectively to complementation and intersection operated on binary accessibility relations. SIM also contains Boolean operators but only at the level of parameters (see Sect. 2) and of course at the level of formulae. Moreover, in Figure 1, FO2[=] refers to the existence of a relational translation into FO2[=], the fragment of classical logic with two variables and equality. As an outcome, the logic SIM has features that prevent from having a natural translation into BML or the hybrid μ-calculus

	SIM	BML	hybrid μ-calculus
nominals	X		X
universal modality	X	X	X
$\cap$	X	X	
$\neg$		X	
FO2[=]	X [DK98]	X	
in ExpTime	X, this paper		X [SV01]

Fig. 1. Comparing SIM, BML, and the hybrid μ-calculus

Our technical developments for SIM can be extended to the logics FORIN and IND (see e.g., [Kon98]), where the similarity relations are replaced by forward inclusion relations and indiscernability relations, respectively but without nominals at the object level. This improves significantly the upper bound from [DK98] whereas decidability for FORIN and IND could not be obtained from the reduction into the (undecidable) fragment of first order logic with three variables [Kah62]. The decidability status of full FORIN and IND is still open.

Plan of the paper. The rest of the paper is structured as follows. Sect. 2 presents the logics for which the computational complexity is studied in the paper. Sect. 3 deals with normal forms for SIM formulae whereas Sect. 4 introduces the concept of the global information for SIM-models that will play an important role. In Sects. 5 and 6 we provide a notion of Hintikka trees for SIM-models preparing the automata construction. For the logic SIM and variants of it, the satisfiability problem is reduced to the emptiness problem for Büchi automata on infinite trees in Sect. 7. Finally, we give some concluding remarks in Sect. 8.

The paper has been designed to be self-contained. Standard definitions we use concerning automata on infinite objects can be found in [Tho90], and concerning computational complexity in [Pap94].

2 Information Logics

In this section, we first introduce syntax and semantics of the logic SIM, then describe the closely connected logics FORIN and IND.

The set of primitive symbols of the language for SIM is composed of

- a countably infinite set $\mathrm{PRP} = \{\mathrm{p}_1, \mathrm{p}_2, \ldots\}$ of *propositional variables*,
- a countably infinite set $\mathrm{NOM} = \{\mathrm{x}_1, \mathrm{x}_2, \ldots\}$ of *object nominals*,
- a set P of *parameter expressions*, which is the smallest set containing a countably infinite set $\mathrm{PNOM} = \{\mathrm{E}_1, \mathrm{E}_2, \ldots\}$ of *parameter nominals* and a countably infinite set $\mathrm{PVAR} = \{\mathrm{C}_1, \mathrm{C}_2, \ldots\}$ of *parameter variables*, and that is closed under the Boolean operators $\cap, \cup, -$.

The formation rules of the set FOR(SIM) of SIM-formulae are those of propositional logic, where object nominals can be used in the place of propositional variables, plus the rule: if $\phi \in \mathrm{FOR(SIM)}$ and $\mathrm{A} \in \mathrm{P}$, then $[\mathrm{A}]\phi \in \mathrm{FOR(SIM)}$. The following is an example of a (valid) SIM-formula:

$$[\mathrm{E}_2 \cap -\mathrm{E}_2]\mathrm{x} \Rightarrow [\mathrm{E}_1 \cup \mathrm{C}_1](\mathrm{x} \vee \mathrm{p}).$$

Moreover, for every syntactic object O, we write $|\mathrm{O}|$ to denote its *length* (or *size*), that is the number of symbol occurrences in O viewed as a string. As usual, $\mathrm{sub}(\phi)$ denotes the set of *subformulae* of the formula ϕ (including ϕ itself). For every $\mathrm{X} \in \{\mathrm{NOM}, \mathrm{PNOM}, \mathrm{PVAR}, \mathrm{P}\}$, we write $\mathrm{X}(\phi)$ to denote the elements of X occurring in the formula ϕ. Obviously, $\mathrm{card}(\mathrm{X}(\phi)) < |\phi|$.

Definition 1. Let PAR be a non-empty set. A P-*interpretation* m is a map $m : \mathrm{P} \to \mathcal{P}(PAR)$ such that, for all $\mathrm{A}_1, \mathrm{A}_2 \in \mathrm{P}$,

- if $\mathrm{A}_1, \mathrm{A}_2 \in \mathrm{PNOM}$ and $\mathrm{A}_1 \neq \mathrm{A}_2$, then $m(\mathrm{A}_1) \neq m(\mathrm{A}_2)$;
- if $\mathrm{A}_1 \in \mathrm{PNOM}$, then $m(\mathrm{A}_1)$ is a singleton;
- $m(\mathrm{A}_1 \cap \mathrm{A}_2) = m(\mathrm{A}_1) \cap m(\mathrm{A}_2)$ and $m(\mathrm{A}_1 \cup \mathrm{A}_2) = m(\mathrm{A}_1) \cup m(\mathrm{A}_2)$;
- $m(-\mathrm{A}_1) = PAR \setminus m(\mathrm{A}_1)$.

PAR is referred to as a set of parameters that is the obvious counterpart of the set of attributes in information systems. Given parameter expressions A and B, we write $\mathrm{A} \equiv \mathrm{B}$ iff for every P-interpretation m, we have $m(\mathrm{A}) = m(\mathrm{B})$.

Definition 2. A *SIM-model* $\mathcal{M}$ is a structure $\mathcal{M} = \langle W, (\mathcal{R}_P)_{P \subseteq PAR}, m \rangle$, where W and PAR are non-empty sets and $(\mathcal{R}_P)_{P \subseteq PAR}$ is a family of binary relations on W such that

(uni) $\mathcal{R}_\emptyset$ is the cartesian product $W \times W$;
(refl) $\mathcal{R}_P$ is reflexive for every $P \subseteq PAR$;
(sym) $\mathcal{R}_P$ is symmetric for every $P \subseteq PAR$;
(inter) $\mathcal{R}_{P \cup Q} = \mathcal{R}_P \cap \mathcal{R}_Q$ for all $P, Q \subseteq PAR$.

Moreover, m is a mapping $m : \mathrm{NOM} \cup \mathrm{PRP} \cup \mathrm{P} \to \mathcal{P}(W) \cup \mathcal{P}(PAR)$ such that $m(\mathrm{p}) \subseteq W$ for every $\mathrm{p} \in \mathrm{PRP}$, $m(\mathrm{x}) = \{w\}$, where $w \in W$ for every $\mathrm{x} \in \mathrm{NOM}$, and the restriction of m to P is a P-interpretation.

Consequently, two levels of interpretation are used to define the relations in the SIM-models. On the one hand, the parameter expressions are interpreted within the Boolean algebra

$$\mathcal{B} = \langle \mathcal{P}(PAR), \cup, \cap, -, 1, 0 \rangle$$

for some non-empty set PAR. On the other hand, the conditions on $(\mathcal{R}_P)_{P \subseteq PAR}$ induce a semi-lattice structure of $\mathcal{L} = \langle \{\mathcal{R}_P : P \in \mathcal{B}\}, \cap \rangle$ with zero element $W \times W$.

Condition **(inter)** allows SIM to capture intersection on relations. Indeed, let us write R_{A} for $\mathcal{R}_{m(\mathrm{A})}$. Then, for all parameter expressions A, B, we have $R_{\mathrm{A} \cup \mathrm{B}} = R_{\mathrm{A}} \cap R_{\mathrm{B}}$. By contrast, complementation and union cannot be expressed in a similar fashion (otherwise we would get operators similar to those in BML). Additionally, SIM contains universal modality since $R_{\mathrm{A} \cap -\mathrm{A}}$ is precisely the product $W \times W$.

The object nominals can be viewed as constants for objects and parameter nominals as constants for attributes in information systems. Similarly, $(\mathcal{R}_P)_{P \subseteq PAR}$ is an abstraction of the family $(sim_A)_{A \subseteq PAR}$ derived from information systems. Please note that, for parameter nominals, we assume that different nominals are interpreted as different relations, i.e., we admit the so-called *unique name assumption*. In contrast, object nominals can be interpreted as arbitrary singletons, i.e., we do not admit the unique name assumption. Since the set of parameter nominals is countably infinite, an obvious consequence of the definition of the SIM-models is that every SIM-model has an infinite set of parameters. Let $\mathcal{M} = \langle W, (\mathcal{R}_P)_{P \subseteq PAR}, m \rangle$ be a model. As usual, we say that a formula ϕ is *satisfied by* $w \in W$ *in* $\mathcal{M}$ (written $\mathcal{M}, w \models \phi$) if the following conditions are satisfied.

$\mathcal{M}, w \models \mathrm{p}$ iff $w \in m(\mathrm{p})$ for $\mathrm{p} \in \mathrm{PRP} \cup \mathrm{NOM}$;
$\mathcal{M}, w \models \neg\phi$ iff not $\mathcal{M}, w \models \phi$;
$\mathcal{M}, w \models \phi \wedge \psi$ iff $\mathcal{M}, w \models \phi$ and $\mathcal{M}, w \models \psi$;
$\mathcal{M}, w \models [\mathrm{A}]\phi$ iff for every $w' \in W$, if $\langle w, w' \rangle \in \mathcal{R}_{m(\mathrm{A})}$, then $\mathcal{M}, w' \models \phi$.

A formula ϕ is *true* in a SIM-model $\mathcal{M}$ (written $\mathcal{M} \models \phi$) iff for every $w \in W$, $\mathcal{M}, w \models \phi$. A formula ϕ is said to be *SIM-valid* iff ϕ is true in every SIM-model. A formula ϕ is said to be *SIM-satisfiable* iff $\neg\phi$ is not SIM-valid.

Theorem 3. *[Vak87] The class of information frames $\langle OB, (sim_A)_{A \subseteq AT} \rangle$ derived from information systems is precisely the class of SIM-frames.*

The frames are understood as parts of the models without the meaning function m. Hence, from a SIM-model for a given SIM-formula ϕ, one can extract an information system satisfying the specification ϕ. Theorem 3 guarantees that the SIM-models are the adequate structures to deal with the information frames based on similarity derived from information systems.

The similarity logic with an infinite set of parameters defined in [Kon98] is not strictly the logic SIM defined above but one can show that both logics have the same set of valid formulae [DK98, Proposition 9]. Variants of SIM can be easily designed by considering relations derived from information systems different from similarity (e.g., forward inclusion, indiscernability). Let FORIN [resp. IND] be the relative logic sharing its language with SIM such that a FORIN-model [resp. IND-model] is obtained from Definition 2 by adding the condition **(trans)** and by withdrawing **(sym)** [resp. from Definition 2 by adding the condition **(trans)**]:

(trans) $\mathcal{R}_P$ is transitive for every $P \subseteq PAR$.

Decidability of the satisfiability problem for the logic SIM is shown in [DK98] by translating SIM satisfiability into satisfiability for FO2[=], the fragment of classical logic with two variables and equality. The reduction increases exponentially the size of the formulae and FO2[=] satisfiability is in NEXPTIME [GKV97]. Hence, the best known upper bound for SIM satisfiability is N2EXPTIME. Additionally, the proof in [DK98] cannot be adapted to show the decidability of IND and FORIN since transitivity requires three variables.

More about the logic SIM and analogous information logics can be found in [DO02].

3 Normal Forms for Parameter Expressions

In this section, we recall a notion of normal form for parameter expressions inspired by the canonical disjunctive normal form for propositional logic. Such normal forms play a special role for the relative information logics. Normal forms for Boolean modal expressions with nominals have been introduced in [Kon98] in order to facilitate the design of Rasiowa–Sikorski-style proof systems (dual tableaux) for SIM. Such a technique has been also useful to show decidability of SIM [DK98] and for some fragments of Boolean modal logic BML [LS01, Sect. 5] (see also [DG00]). In this paper, we use a normal form for the Boolean modal expressions with nominals. We recall below some definitions.

For $l \geq 1$ and $n \geq 1$, let $Y = \{\mathrm{E}_1, \ldots, \mathrm{E}_l\}$ be distinct parameter nominals and be $X = \{\mathrm{C}_1, \ldots, \mathrm{C}_n\}$ distinct parameter variables. For every integer $k \in \{0, \ldots, 2^n - 1\}$, we denote by B_k the parameter expression $\mathrm{B}_k \stackrel{\text{def}}{=} \mathrm{A}_1 \cap \ldots \cap \mathrm{A}_n$ where, for every $s \in \{1, \ldots, n\}$, $\mathrm{A}_s = \mathrm{C}_s$ if $bit_s(k) = 0$ and $\mathrm{A}_s = -\mathrm{C}_s$ otherwise, and $bit_s(k)$ denotes the sth bit in the binary representation of k with n bits. Although not essential, the use of binary representation will facilitate the presentation of technical developments. For every integer $k' \in \{0, \ldots, l\}$, we denote by $\mathrm{D}_{k'}$ the parameter expression

$$\mathrm{D}_{k'} \stackrel{\text{def}}{=} \begin{cases} -\mathrm{E}_1 \cap \ldots \cap -\mathrm{E}_l \text{ if } k' = 0; \\ \mathrm{E}_{k'} \text{ otherwise.} \end{cases}$$

For every integer $k \in \{0, \ldots, 2^n - 1\}$ and for every $k' \in \{0, \ldots, l\}$, $\mathrm{A}_{k,k'} \stackrel{\text{def}}{=} \mathrm{B}_k \cap \mathrm{D}_{k'}$. For instance, if $n = l = 2$, then $\mathrm{A}_{3,2} = -\mathrm{C}_1 \cap -\mathrm{C}_2 \cap \mathrm{E}_2$. The set $\mathrm{Comp}(X, Y)$ of $\langle X, Y \rangle$-*components*, is defined as follows:

$$\mathrm{Comp}(X, Y) \stackrel{\text{def}}{=} \{\mathrm{A}_{k,k'} \mid k \in \{0, \ldots, 2^n - 1\}, k' \in \{0, \ldots, l\}\}.$$

The set $\mathrm{Comp}(X,Y)$ of $\langle X,Y\rangle$-components enables us to partition every set of parameters. Indeed, for every P-interpretation $m : \mathrm{P} \rightarrow \mathcal{P}(PAR)$, the family $\{m(\mathrm{A}) \mid \mathrm{A} \in \mathrm{Comp}(X,Y)\}$ is a partition of PAR [Kon98]. As a consequence, we obtain the following property.

Lemma 4. *Let* A *be a parameter expression built from* $X \cup Y$*. Then either* $\mathrm{A} \equiv -\mathrm{A} \cap \mathrm{A}$ *or there is a unique non-empty subset* $\{\mathrm{A}'_1, \ldots, \mathrm{A}'_u\}$ *of* $\mathrm{Comp}(X,Y)$ *such that* $\mathrm{A} \equiv \mathrm{A}'_1 \cup \ldots \cup \mathrm{A}'_u$.

Lemma 4 enables us to define normal forms of parameter expressions. Let A be a parameter expression built from $X \cup Y$. The normal form of A, $\mathrm{N}_{X,Y}(\mathrm{A})$, is defined as follows:

$$\mathrm{N}_{X,Y}(\mathrm{A}) \stackrel{\mathrm{def}}{=} \begin{cases} \emptyset \text{ if } \mathrm{A} \equiv (\mathrm{A} \cap -\mathrm{A}); \\ \{\mathrm{A}_{k_1,k'_1}, \ldots, \mathrm{A}_{k_u,k'_u}\} \text{ if } \mathrm{A} \equiv \mathrm{A}_{k_1,k'_1} \cup \ldots \cup \mathrm{A}_{k_u,k'_u}. \end{cases}$$

Observe that there exists an effective procedure that computes $\mathrm{N}_{X,Y}(\mathrm{A})$ in deterministic time exponential in $|\mathrm{A}|+n+l$. Moreover, it is known that, for all parameter expressions A, B built from $X \cup Y$, we have $\mathrm{A} \equiv \mathrm{B}$ iff $\mathrm{N}_{X,Y}(\mathrm{A}) = \mathrm{N}_{X,Y}(\mathrm{B})$. Please note that this normal form is not thought to be applied to all parameter expressions in a SIM-formula to be tested for satisfiability (since this would obviously yield an exponential blow-up), but it is used in the following section to decide the implication relation between parameter expressions.

4 Global Information for SIM-models

Due to the presence of nominals, SIM does not have the tree model property. Hence, to use automata-based techniques, we will define appropriate tree abstractions of models, so-called Hintikka-trees. However, the expressive power of SIM is such that the Hintikka-trees will be defined w.r.t. "global" information. In this section, we describe this global information in SIM models. Intuitively, global information is true at any point of the model or concerns edges which are omitted when considering tree abstractions of (non-tree) models, i.e., edges relating an individual to the instance of a nominal.

For instance, given an object nominal x occurring in ϕ, the set of subformulae of ϕ that hold true in the unique state satisfying x is a global information. In this section, we generalize the global information about object nominals and the universal modality from [SV01]. Guessing a global information for a given formula ϕ will correspond to the primary non-deterministic choice in the automata built for ϕ (see Sect. 7).

Let ϕ be a SIM-formula, C a parameter constant, E a parameter nominal, and x an object nominal. To avoid considering formulae containing no parameter nominals or no parameter variables, in the remainder, we assume w.l.o.g that (1) each formula contains at least one object nominal, and (2) each formula is of the form $\phi' \wedge \bigwedge_{i=1}^{\gamma} \neg[(\mathrm{C} \cap -\mathrm{C} \cap \mathrm{E})]\neg \mathrm{x}_i$, where $\mathrm{x}_1, \ldots, \mathrm{x}_\gamma$ are all the object nominals occurring in ϕ'. The first assumption is without loss of generality because we can transform each SIM-formula without object nominals into an equi-satisfiable one by conjoining it with x. The second assumption is without loss of generality because each SIM-model interprets both $\mathrm{C} \cap -\mathrm{C}$ and $\mathrm{C} \cap -\mathrm{C} \cap \mathrm{E}$ as the universal relation.

In the remainder of this section, we discuss all aspects of global information which we use to design the Büchi tree automaton accepting all (tree abstractions of) models of a SIM formula ϕ. To do so, we first consider a fixed model $\mathcal{M} = \langle W, (\mathcal{R}_P)_{P \subseteq PAR}, m\rangle$ of ϕ and collect, step by step, all information we keep globally track of when abstracting from this model to the corresponding Hintikka-tree.

4.1 Parameter Nominals

Let $\mathrm{PNOM}(\phi) = \{\mathrm{E}_1, \ldots, \mathrm{E}_l\}$ be the set of parameter nominals and $\mathrm{PVAR}(\phi) = \{\mathrm{C}_1, \ldots, \mathrm{C}_n\}$ be the set of parameter variables occurring in ϕ. Recall that $n, l \geq 1$. Given a P-interpretation m, there is a unique map $f : \{1, \ldots, l\} \rightarrow \{0, \ldots, 2^n - 1\}$ such that, for every $k' \in \{1, \ldots, l\}$, we have

$$\{k \in \{0, \ldots, 2^n - 1\} \mid m(\mathrm{E}_{k'}) \in m(\mathrm{B}_k)\} = \{f(k')\} \qquad \text{(UNI)}$$

since we assume the unique name assumption for parameter nominals, and moreover the set

$$\{m(\mathrm{B}_k) \mid k \in \{0, \ldots, 2^n - 1\}\}$$

is a partition of PAR. Such a map f can be encoded with $\mathcal{O}(n \times l \times log(l))$ bits. Moreover, for every $k' \in \{1, \ldots, l\}$, for every set $X \subseteq \{0, \ldots, 2^n - 1\}$, at most one parameter expression in $\{\mathrm{A}_{k,k'} \mid k \in X\}$ is not interpreted as the universal relation. Hence we have a variety of different parameter expressions that are all interpreted as the universal relation—a situation obviously more complex than the one in which one explicit universal modal connective $[U]$ is part of the language.

Let A, B be parameter expressions built on $\mathrm{PNOM}(\phi) \cup \mathrm{PVAR}(\phi)$. Given the map $f : \{1, \ldots, l\} \rightarrow \{0, \ldots, 2^n - 1\}$, we write $\mathrm{A} \sqsubseteq_f \mathrm{B}$ iff for every P-interpretation m satisfying (UNI), we have $m(\mathrm{A}) \subseteq m(\mathrm{B})$. We have chosen to define $\sqsubseteq_f$ rather than $\sqsubseteq_m$ because there are far less mappings f than there are ms, and this difference will be crucial in the following. The relation $\mathrm{A} \sqsubseteq_f \mathrm{B}$ can be checked in exponential-time in $|\mathrm{A}| + |\mathrm{B}| + n + l$ since $\mathrm{A} \sqsubseteq_f \mathrm{B}$ iff

$$\mathrm{N}_{\mathrm{PVAR}(\phi),\mathrm{PNOM}(\phi)}(\mathrm{A}) \setminus (\bigcup_{k'=1}^{l} \{\mathrm{A}_{k,k'} \in \mathrm{Comp}(\mathrm{PVAR}(\phi), \mathrm{PNOM}(\phi)) \mid k \neq f(k')\})$$
$$\subseteq$$
$$\mathrm{N}_{\mathrm{PVAR}(\phi),\mathrm{PNOM}(\phi)}(\mathrm{B}) \setminus (\bigcup_{k'=1}^{l} \{\mathrm{A}_{k,k'} \in \mathrm{Comp}(\mathrm{PVAR}(\phi), \mathrm{PNOM}(\phi)) \mid k \neq f(k')\}).$$

Indeed, the problem can be shown to be co-NP-complete since it is a slight variant of the validity problem of propositional logic. We write $\mathrm{A} \equiv_f \emptyset$ to denote $\mathrm{A} \sqsubseteq_f \mathrm{A} \cap -\mathrm{A}$. Obviously, $\mathrm{A} \sqsubseteq_f \mathrm{B}$ iff, for every SIM-model $\langle W, (\mathcal{R}_P)_{P \subseteq PAR}, m\rangle$ with m satisfying (UNI), we have $\mathcal{R}_{m(\mathrm{B})} \subseteq \mathcal{R}_{m(\mathrm{A})}$.

4.2 Universal Modalities

Set $UF = \{[\mathrm{A}]\psi \in \mathrm{sub}(\phi) \mid \mathrm{A} \equiv_f \emptyset,\ \mathcal{M} \models \psi\}$ and $EF = \{[\mathrm{A}]\psi \in \mathrm{sub}(\phi) \mid \mathrm{A} \equiv_f \emptyset,\ \mathcal{M} \not\models \psi\}$. Observe that UF and EF depend on the map f but for a given model $\mathcal{M}$, the structure $\langle f, UF, EF\rangle$ is unique. The structure $\langle UF, EF\rangle$ can be also encoded using $\mathcal{O}(|\phi| \times log(|\phi|))$ bits.

4.3 Object Nominals

Let $\mathrm{NOM}(\phi)$ be the set of object nominals and $\mathrm{P}(\phi)$ the parameter expressions occurring in ϕ. We will fix which nominals are interpreted by the same object, what formulae are satisfied by these objects, and how they are inter-related. Let EQ be the unique equivalence relation on $\mathrm{NOM}(\phi)$, NOM be the unique map $NOM : \mathrm{NOM}(\phi) \to \mathcal{P}(\mathrm{sub}(\phi))$ and $\mathrm{R_N}$ be the unique ternary relation in $\mathrm{NOM}(\phi)^2 \times \mathrm{P}(\phi)$ such that

- for all $\mathrm{x}, \mathrm{y} \in \mathrm{NOM}(\phi)$, $\langle \mathrm{x}, \mathrm{y}\rangle \in EQ$ iff $m(\mathrm{x}) = m(\mathrm{y})$;
- for every $\mathrm{x} \in \mathrm{NOM}(\phi)$,
  $NOM(\mathrm{x}) \stackrel{\text{def}}{=} \{\psi \in \mathrm{sub}(\phi) \mid \text{for } m(\mathrm{x}) = \{w\},\ \mathcal{M}, w \models \psi\}$;
- for all $\mathrm{x}, \mathrm{y} \in \mathrm{NOM}(\phi)$, $\mathrm{A} \in \mathrm{P}(\phi)$, if $m(\mathrm{x}) = \{w\}$ and $m(\mathrm{y}) = \{w'\}$, then $\langle \mathrm{x}, \mathrm{y}, \mathrm{A}\rangle \in \mathrm{R_N}$ iff $\langle w, w'\rangle \in \mathcal{R}_{m(\mathrm{A})}$.

The triple $\langle EQ, NOM, \mathrm{R_N}\rangle$ can be encoded using $\mathcal{O}(|\phi|^3)$ bits. Such a global information about the model $\mathcal{M}$ is actually a variant of the global information used in [SV01].

4.4 Abstract Global Information

Next, we summarize the above mentioned aspects of global information and define it independently of a specific model.

A *global information* G for ϕ is a structure $\langle f, UF, EF, EQ, NOM, \mathrm{R_N}\rangle$ such that

1. f is a map $f : \{1, \ldots, l\} \to \{0, \ldots, 2^n - 1\}$ (it describes how parameter nominals are interpreted);
2. UF and EF are subsets of $\{\varphi \in \mathrm{sub}(\phi) : \varphi = [\mathrm{A}]\psi\}$ (UF contains the formulae quantified universally that are true in a model, and EF contains those formulae quantified universally in ϕ that are not true);
3. $EQ \subseteq \mathrm{NOM}(\phi)^2$ (it describes which object nominals are interpreted by the same individual);
4. NOM is a map $NOM : \mathrm{NOM}(\phi) \to \mathcal{P}(\mathrm{sub}(\phi))$ (it describes the formulae satisfied by the interpretations of nominals);
5. $\mathrm{R_N} \subseteq \mathrm{NOM}(\phi)^2 \times \mathrm{P}(\phi)$ (it describes the inter-relationship between nominals).

We write $\mathrm{R_N}(\mathrm{A})$ to denote the binary relation $\{\langle \mathrm{x}, \mathrm{x}' \rangle \mid \langle \mathrm{x}, \mathrm{x}', \mathrm{A} \rangle \in \mathrm{R_N}\}$. A global information G for ϕ can be easily encoded using $\mathcal{O}(|\phi|^3)$ bits.

Next, we define consistency of global informations. So far, a global information G is simply a structure of a certain type, whereas the SIM-consistency of G reflects the semantics of SIM.

A global information $G = \langle f, UF, EF, EQ, NOM, \mathrm{R_N} \rangle$ is said to be *SIM-consistent* iff G satisfies the following conditions:

(G1) EQ is an equivalence relation;
(G2) for every $\mathrm{x} \in \mathrm{NOM}(\phi)$, $NOM(\mathrm{x})$ is locally SIM-consistent (to be defined in Definition 6 below) and $\mathrm{x} \in NOM(\mathrm{x})$;
(G3) $\{UF, EF\}$ is a bipartition of $\{[\mathrm{A}]\psi \in \mathrm{sub}(\phi) \mid \mathrm{A} \equiv_f \emptyset\}$;
(G4) for all $\mathrm{x}, \mathrm{y} \in \mathrm{NOM}(\phi)$, $\langle \mathrm{x}, \mathrm{y} \rangle \in EQ$ iff $NOM(\mathrm{x}) = NOM(\mathrm{y})$;
(G5) for all $\mathrm{A}, \mathrm{B} \in \mathrm{P}(\phi)$, $\mathrm{A} \sqsubseteq_f \mathrm{B}$ implies $\mathrm{R_N}(\mathrm{B}) \subseteq \mathrm{R_N}(\mathrm{A})$;
(G6) for every $\mathrm{A} \in \mathrm{P}(\phi)$, EQ is a congruence for $\mathrm{R_N}(\mathrm{A})$, and the relation $\mathrm{R_N}(\mathrm{A})$ is reflexive and symmetric;
(G7) for all $\mathrm{x}, \mathrm{y} \in \mathrm{NOM}(\phi)$, if $[\mathrm{A}]\psi \in NOM(\mathrm{x})$ and $\langle \mathrm{x}, \mathrm{y}, \mathrm{B} \rangle \in \mathrm{R_N}$ for some $\mathrm{A} \sqsubseteq_f \mathrm{B}$, then $\psi \in NOM(\mathrm{y})$;
(G8) for all $\langle \mathrm{x}, \mathrm{y}, \mathrm{A}_1 \rangle, \ldots, \langle \mathrm{x}, \mathrm{y}, \mathrm{A}_n \rangle \in \mathrm{R_N}$, $n \geq 1$, and $\mathrm{B} \in \mathrm{P}(\phi)$, if $\mathrm{B} \sqsubseteq_f \mathrm{A}_1 \cup \ldots \cup \mathrm{A}_n$, then $\langle \mathrm{x}, \mathrm{y}, \mathrm{B} \rangle \in \mathrm{R_N}$;
(G9) for every $\mathrm{A} \in \mathrm{P}(\phi)$, for all $\mathrm{x}, \mathrm{y} \in \mathrm{NOM}(\phi)$, $\mathrm{A} \equiv_f \emptyset$ implies $\langle \mathrm{x}, \mathrm{y}, \mathrm{A} \rangle \in \mathrm{R_N}$.

Please note that **(G6)** is the place where it is important that we are considering SIM, and which would need to be modified when adapting the approach to FORIN or IND. In order the establish the ExpTime upper bound for SIM, we need the result below.

Lemma 5. *Checking whether a global information for ϕ is SIM-consistent can be done in time in $2^{\mathcal{O}(|\phi|)}$.*

The exponential bound is due to the relation $\sqsubseteq_f$ and to the exponential amount of triple in **(G8)** since in **(G8)**, $1 \leq n \leq \mathrm{card}(\mathrm{P}(\phi))$. We write $\mathrm{GCONS}(\phi)$ to denote the set of SIM-consistent global informations for ϕ.

5 Symbolic States

In this section, we define the notion of symbolic states which represent objects in SIM-models.

Definition 6. Let X be a subset of $\mathrm{sub}(\phi)$ for some formula ϕ. The set X is said to be *locally SIM-consistent* iff each $\psi \in \mathrm{sub}(\phi)$ satisfies the following conditions:

(L1) if $\psi = \neg\varphi$, then $\varphi \in X$ iff $\psi \notin X$;
(L2) if $\psi = \varphi_1 \wedge \varphi_2$, then $\{\varphi_1, \varphi_2\} \subseteq X$ iff $\psi \in X$;
(L3) if $\psi = [\mathrm{A}]\varphi$ and $\psi \in X$, then $\varphi \in X$.

Let G be a SIM-consistent global information. Given two locally SIM-consistent sets X and Y and a parameter expression A occurring in ϕ, we write $X \sim_{G,\mathrm{A}} Y$ to denote that, for every $[\mathrm{B}]\psi \in X$, if $\mathrm{B} \sqsubseteq_f \mathrm{A}$, then $\psi \in Y$ and, for every $[\mathrm{B}]\psi \in Y$, if $\mathrm{B} \sqsubseteq_f \mathrm{A}$, then $\psi \in X$.

Observe that $\sim_{G,\mathrm{A}}$ depends on G by the map f. The relation $\sim_{G,\mathrm{A}}$ is the abstract counterpart of a maximal relation $\mathcal{R}_{m(\mathrm{A})}$ in SIM-models. More precisely, let $\mathcal{M} = \langle W, (\mathcal{R}_P)_{P \subseteq PAR}, m\rangle$ be a SIM-model, let $\langle w, w'\rangle \in \mathcal{R}_{m(\mathrm{A})}$ for some A occurring in ϕ, and let G be a SIM-consistent global information for ϕ built from $\mathcal{M}$ as done in Sect. 4. Then

$$\{\psi \in \mathrm{sub}(\phi) \mid \mathcal{M}, w \models \psi\} \sim_{G,\mathrm{A}} \{\psi \in \mathrm{sub}(\phi) \mid \mathcal{M}, w' \models \psi\}.$$

We are now ready to define symbolic states. Each such state contains information on the relation between the associated node and its (unique) predecessor, the formulae the respective object satisfies, and how it is related to (instances of) object nominals. The latter information is crucial since these edges will be omitted when abstracting/unravelling models to Hintikka trees (if they were not omitted, unravelling would either not yield trees or instances of object nominals would not be unique).

A *symbolic state for* ϕ is either $\perp$ or a triple $q = \langle \mathrm{A}, X, T\rangle$ such that $\mathrm{A} \in \mathrm{P}(\phi)$, $X \in \mathcal{P}(\mathrm{sub}(\phi))$, and $T \subseteq \mathrm{P}(\phi) \times \mathrm{NOM}(\phi)$.

In $q = \langle \mathrm{A}, X, T\rangle$, A refers to the relation $\mathcal{R}_{m(\mathrm{A})}$ which relates q's (unique) predecessor to q, X is the set of formulae satisfied in q, and T is a *table* such that, for every $\langle \mathrm{B}, \mathrm{x}\rangle \in T$, $\langle q, w\rangle \in \mathcal{R}_{m(\mathrm{A})}$ for $m(\mathrm{x}) = \{w\}$. We often use $\langle \mathrm{A}_q, X_q, T_q\rangle$. The "dummy" value $\perp$ is used for those nodes in a tree not representing objects, and we call a symbolic state q *dummy* if $q = \perp$. Similarly, a symbolic state $\langle \mathrm{A}, X, T\rangle$ is a *named state* if $X \cap \mathrm{NOM}(\phi)$ is non-empty. We will also write $\psi \in q = \langle \mathrm{A}, X, T\rangle$ [resp. $\langle \mathrm{A}, \mathrm{x}\rangle \in q$] instead of $\psi \in X$ [resp. $\langle \mathrm{A}, \mathrm{x}\rangle \in T$].

Let G be a (SIM-consistent) global information. A symbolic state $q = \langle \mathrm{A}, X, T\rangle$ is said to be *locally SIM-consistent with respect to* G iff q is dummy or if it satisfies the following conditions:

(SC1) X is locally SIM-consistent;
(SC2) for every $\mathrm{x} \in \mathrm{NOM}(\phi)$, $\mathrm{x} \in q$ implies $X = NOM(\mathrm{x})$ and $T = \{\langle \mathrm{B}, \mathrm{y}\rangle \mid \langle \mathrm{x}, \mathrm{y}, \mathrm{B}\rangle \in \mathrm{R_N}\}$;
(SC3) for every $\langle \mathrm{A}, \mathrm{x}\rangle \in T$, $X \sim_{G,\mathrm{A}} NOM(\mathrm{x})$;
(SC4) for all $\langle \mathrm{A}_1, \mathrm{x}_1\rangle, \ldots, \langle \mathrm{A}_n, \mathrm{x}_n\rangle \in T$ with $n \geq 1$, if $\mathrm{x}_1 = \ldots = \mathrm{x}_n$ then, for every $\mathrm{A} \in \mathrm{P}(\phi)$ with $\mathrm{A} \sqsubseteq_f \mathrm{A}_1 \cup \ldots \cup \mathrm{A}_n$, we have $\langle \mathrm{A}, \mathrm{x}_1\rangle \in T$;
(SC5) for every $\mathrm{B} \in \mathrm{P}(\phi)$ such that $\mathrm{B} \equiv_f \emptyset$, for every $\mathrm{x} \in \mathrm{NOM}(\phi)$, $\langle \mathrm{B}, \mathrm{x}\rangle \in T$;
(SC6) $UF \subseteq X$ and $EF \cap X = \emptyset$.

We use $\mathrm{SYMB}(\phi)$ to denote the set of symbolic states of ϕ, and $\mathrm{SYMB}_G(\phi)$ to denote the set of symbolic states of ϕ that are locally SIM-consistent with respect to a (SIM-consistent) global information G.

(SC3) ensures that the "omitted" edges to instances of nominals are semantically possible. In order to establish the ExpTime upper bound for SIM, we need also the result below.

Lemma 7. *Deciding whether a symbolic state is locally SIM-consistent with respect to a (SIM-consistent) global information can be done in time* $2^{\mathcal{O}(|\phi|)}$.

6 Hintikka Trees

We are now ready to introduce Hintikka trees for SIM with respect to a given global information G. As usual, such trees are abstractions of SIM-models that allow a further treatment with Büchi automata on infinite trees. A nice example of existing such abstractions are those for the μ-calculus (see e.g., the well-founded pre-models in [SE89]). We will show that each SIM-model can be unravelled into a Hintikka tree, and thus prove a tree model property for SIM (such properties are known to be helpful for the decidability of modal logics [Grä99]). This section is the core of the paper since it combines the preliminary results from the previous sections with the ideas underlying the introduction of Hintikka trees.

For ϕ a SIM-formula, a Hintikka-tree for ϕ is labelled with symbolic states, has a dummy root node, and, at its first level, we find a node satisfying ϕ as well as nodes for all nominals occurring in ϕ. Since a negated box formulae can be either witnessed by an "anonymous" successor node in the tree or by a node labelled with named states representing an instance of a nominal, **(H7)** is split into two conditions, one for each case.

We recall that, given $K \geq 1$ and a finite alphabet Σ, an infinite Σ, K-tree $\mathcal{T}$ is a mapping $\mathcal{T} : \{1, \ldots, K\}^* \rightarrow \Sigma$.

Let ϕ be a SIM-formula with $K = |\phi|$, $\mathrm{PNOM}(\phi)$ the set of parameter nominals occurring in ϕ with $l = \mathrm{card}(\mathrm{PNOM}(\phi)) \geq 1$, and $\mathrm{PVAR}(\phi)$ the set of parameter variables occurring in ϕ with $n = \mathrm{card}(\mathrm{PVAR}(\phi)) \geq 1$.

Definition 8. A $\mathrm{SYMB}(\phi), K$-tree $\mathcal{T}$ is a *Hintikka tree for* ϕ iff there exists a SIM-consistent global information $G = \langle f, UF, EF, EQ, NOM, \mathrm{R_N} \rangle \in \mathrm{GCONS}(\phi)$ for ϕ such that

(H1) $\mathcal{T}(\epsilon)$ is dummy;
(H2) there is $i \in \{1, \ldots, K\}$ such that $\phi \in \mathcal{T}(i)$;
(H3) for every $\mathrm{x} \in \mathrm{NOM}(\phi)$, there is a unique $i \in \{1, \ldots, K\}$ such that $\mathrm{x} \in \mathcal{T}(i)$ (this i is then written i_{x});

and each $s \in \{1, \ldots, K\}^+$ satisfies the following conditions:

(H4) $\mathcal{T}(s)$ is locally SIM-consistent with respect to G;
(H5) if $\mathcal{T}(s)$ is dummy, then $\mathcal{T}(s \cdot 1), \ldots, \mathcal{T}(s \cdot K)$ are also dummy;
(H6) if s is of length at least 2, then $\mathcal{T}(s)$ is not a named symbolic state;
(H7) if $\mathcal{T}(s) = \langle \mathrm{A}, X, T \rangle$ is not dummy and $[\mathrm{B}]\psi \in \mathrm{sub}(\phi) \setminus X$, then
1. either there is $i \in \{1, \ldots, K\}$ with $\mathcal{T}(s \cdot i) = \langle \mathrm{B}, X', T' \rangle$, $\mathcal{T}(s \cdot i)$ is not dummy, and $\psi \notin X'$ or
2. there is $\mathrm{x} \in \mathrm{NOM}(\phi)$ such that $\langle \mathrm{B}, \mathrm{x} \rangle \in T$ and $\psi \notin \mathcal{T}(i_{\mathrm{x}})$;

(H8) for every $i \in \{1, \ldots, K\}$, if both $\mathcal{T}(s) = \langle \mathrm{A}, X, T \rangle$ and $\mathcal{T}(s \cdot i) = \langle \mathrm{B}, X', T' \rangle$ are not dummy, then $X \sim_{G,\mathrm{B}} X'$.

Such a Hintikka tree is said *to respect G*.

All the preliminary work done so far yields Lemma 9 below.

Lemma 9. *For every SIM-formula ϕ, (I) ϕ is SIM-satisfiable iff (II) ϕ has a Hintikka tree.*

Proof. (II) $\rightarrow$ (I). Let $\mathcal{T}$ be a Hintikka tree respecting the SIM-consistent global information G.

The construction of $\mathcal{M}$. We construct a SIM-model $\mathcal{M} = \langle W, (\mathcal{R}_P)_{P \subseteq PAR}, m \rangle$ of ϕ as follows:

- $W \stackrel{\text{def}}{=} \{s \in \{1, \ldots, K\}^+ : \mathcal{T}(s) \text{ is not dummy}\}$;
- $PAR \stackrel{\text{def}}{=} \mathbb{N}$;
- for every $i \in \mathbb{N}$, $m(\mathrm{E}_i) \stackrel{\text{def}}{=} \{2^n - 1 + i\}$;
- for every $i \in \{1, \ldots, n\}$,

$$m(\mathrm{C}_i) \stackrel{\text{def}}{=} \{2^n - 1 + j \mid j \in \{1, \ldots, l\},\ bit_i(f(j)) = 0\} \cup \{k \in \{0, \ldots, 2^n - 1\} \mid bit_i(k) = 0\}$$

  (the other parameter variables are interpreted as the empty set);
- for every $s \in W$, for every $\mathrm{p} \in \mathrm{PRP}$, $s \in m(\mathrm{p})$ iff $\mathrm{p} \in \mathcal{T}(s)$;
- for every $\mathrm{A} \in \mathrm{P}(\phi)$, let R_{A} be the binary relation on $W \times W$ defined as the reflexive and symmetric closure of the union of the following three sets

  1. $\mathrm{R_N}(\mathrm{A})$;
  2. $\mathrm{R}(\mathrm{A}) = \{\langle s, s \cdot i \rangle \in W^2 \mid s \in \{1, \ldots, K\}^+,\ i \in \{1, \ldots, K\},\ \mathcal{T}(s \cdot i) = \langle \mathrm{A}, X, T \rangle\}$;
  3. $\mathrm{R}'(\mathrm{A}) = \{\langle s, i_{\mathrm{x}} \rangle \in W^2 \mid s \in \{1, \ldots, K\}^+,\ \langle \mathrm{A}, \mathrm{x} \rangle \in \mathcal{T}(s)\}$;

- for every $i \in \mathbb{N} \setminus \{0, \ldots, 2^n - 1 + l\}$, $\mathcal{R}_{\{i\}} \stackrel{\text{def}}{=} W \times W$;
- for every $i \in \{0, \ldots, 2^n - 1 + l\}$, $\mathcal{R}_{\{i\}} \stackrel{\text{def}}{=} \bigcup \{R_{\mathrm{A}} \mid \mathrm{A} \in \mathrm{P}(\phi),\ i \in m(\mathrm{A})\}$;
- for every $P \subseteq \mathbb{N}$ such that $\mathrm{card}(P) \geq 2$, $\mathcal{R}_P \stackrel{\text{def}}{=} \bigcap_{i \in P} \mathcal{R}_{\{i\}}$;
- for every $\mathrm{x} \in \mathrm{NOM}(\phi)$, $m(\mathrm{x}) = \{i_{\mathrm{x}}\}$ (object nominals not occurring in ϕ are interpreted as arbitrary singletons).

Basic properties of $\mathcal{M}$. By construction, each relation R_{A} is reflexive and symmetric. The same holds for each relation $\mathcal{R}_P$ since symmetry and reflexivity are properties preserved by taking arbitrary intersection.

It is not difficult to check that $\mathcal{M}$ is a SIM-model, that $\mathcal{M}$ respects G, and that, moreover, the following properties are satisfied:

1. for every $k \in \{0, \ldots, 2^n - 1\}$, $m(\mathrm{A}_{k,0}) = \{k\}$;
2. for every $k' \in \{1, \ldots, l\}$, $m(\mathrm{A}_{f(k'),k'}) = \{2^n - 1 + k'\}$;
3. if $\langle s, s' \rangle \in \mathcal{R}_{m(\mathrm{A})}$ and $[\mathrm{A}]\psi \in \mathcal{T}(s)$, then $\psi \in \mathcal{T}(s')$.

A nice consequence of the points (1) and (2) is that reasoning about the normal form of A can be reduced to reasoning on the elements in $m(\mathrm{A})$. By way of example, we show the Property (3).

Assume $\langle s, s'\rangle \in \mathcal{R}_{m(\mathrm{A})}$ and $[\mathrm{A}]\psi \in \mathcal{T}(s)$.

Case 1: $\mathrm{A} \equiv_f \emptyset$.

Hence $m(\mathrm{A}) = \emptyset$. If $[\mathrm{A}]\psi \in \mathcal{T}(s)$, then **(SC6)** implies $[\mathrm{A}]\psi \in UF$. Since s' is not dummy and $\mathcal{T}(s')$ is also locally SIM-consistent with respect to G, we obtain $[\mathrm{A}]\psi \in \mathcal{T}(s')$. By **(L3)**, we thus get $\psi \in \mathcal{T}(s')$.

Case 2: $m(\mathrm{A}) = \{i_1, \ldots, i_k\} \neq \emptyset$.

Then $\langle s, s'\rangle \in \mathcal{R}_{m(\mathrm{A})}$ iff for every $i \in \{i_1, \ldots, i_k\}$, $\langle s, s'\rangle \in \mathcal{R}_{\{i\}}$.

Case 2.0: If $s = s'$ then **(L3)** implies $\psi \in \mathcal{T}(s') = \mathcal{T}(s)$.

Case 2.1: $s \neq s'$ and both $\mathcal{T}(s)$ and $\mathcal{T}(s')$ are named symbolic states.

Let $s = i_{\mathrm{x}}$ and $s' = i_{\mathrm{y}}$ for some $\mathrm{x}, \mathrm{y} \in \mathrm{NOM}(\phi)$. Since $\langle s, s'\rangle \in \mathcal{R}_{\{i\}}$ for every $i \in \{i_1, \ldots, i_k\}$, we have that, for every $i \in \{i_1, \ldots, i_k\}$, there is $\mathrm{A}_i \in \mathrm{P}(\phi)$, such that $i \in m(\mathrm{A}_i)$ and $\langle \mathrm{x}, \mathrm{y}, \mathrm{A}_i\rangle \in \mathrm{R_N}$. Thus **(G8)** together with $\mathrm{A} \sqsubseteq_f \mathrm{A}_{i_1} \cup \ldots \cup \mathrm{A}_{i_k}$ implies that $\langle \mathrm{x}, \mathrm{y}, \mathrm{A}\rangle \in \mathrm{R_N}$, and **(G7)** implies $\psi \in \mathcal{T}(s')$.

Case 2.2: $s \neq s'$ and neither $\mathcal{T}(s) = \langle \mathrm{A}_s, X_s, T_s\rangle$ nor $\mathcal{T}(s') = \langle \mathrm{A}_{s'}, X_{s'}, T_{s'}\rangle$ are named symbolic states.

W.l.o.g. , let $s' = s \cdot i$ for some $i \in \{1, \ldots, K\}$. Then $\langle s, s'\rangle \in R_{\mathrm{A}_{s'}}$ and, for every $i \in \{0, \ldots, 2^n - 1 + l\}$, $\langle s, s'\rangle \in \mathcal{R}_{\{i\}}$ iff $i \in m(\mathrm{A}_{s'})$.

Since $\langle s, s'\rangle \in \mathcal{R}_{m(\mathrm{A})}$, we have $\mathrm{A} \sqsubseteq_f \mathrm{A}_{s'}$. **(H8)** implies that $X_s \sim_{G,\mathrm{A}_{s'}} X_{s'}$, and thus $\psi \in \mathcal{T}(s')$.

Case 2.3: $s \neq s'$, $\mathcal{T}(s) = \langle \mathrm{A}_s, X_s, T_s\rangle$ is not a named symbolic state, and $\mathcal{T}(s') = \langle \mathrm{A}_{s'}, X_{s'}, T_{s'}\rangle$ is a named symbolic state.

Let $i_{\mathrm{x}} = s'$ for some $\mathrm{x} \in \mathrm{NOM}(\phi)$. For every $i \in \{i_1, \ldots, i_k\}$, $\langle s, s'\rangle \in \mathcal{R}_{\{i\}}$ implies that, for every $i \in \{i_1, \ldots, i_k\}$, there is $\mathrm{A}_i \in \mathrm{P}(\phi)$ such that $i \in m(\mathrm{A}_i)$ and $\langle \mathrm{A}_i, \mathrm{x}\rangle \in T_s$. **(SC4)** and $\mathrm{A} \sqsubseteq_f \mathrm{A}_{i_1} \cup \ldots \cup \mathrm{A}_{i_k}$ imply that $\langle \mathrm{A}, \mathrm{x}\rangle \in T_s$. Finally, **(H4)** and **(SC3)** imply that $X_s \sim_{G,\mathrm{A}} X_{s'}$, and thus $\psi \in \mathcal{T}(s')$.

Case 2.4: $s \neq s'$ and $\mathcal{T}(s)$ is a named symbolic state and $\mathcal{T}(s')$ is not a named symbolic state.

Due to the symmetry of $\sim_{G,\mathrm{A}}$, this case is similar to Case 2.3.

The induction. Since $\mathcal{T}$ is a Hintikka tree, there is $i \in \{1, \ldots, K\}$ such that $\phi \in \mathcal{T}(i)$. In order to show that $\mathcal{M}, i \models \phi$ (and therefore $\mathcal{M}$ is a model for ϕ), we prove by induction on the formula structure that, for every $\psi \in \mathrm{sub}(\phi)$, for every $s \in W$, we have $\psi \in \mathcal{T}(s)$ iff $\mathcal{M}, s \models \psi$. The base case (with object nominals and propositional variables) and the induction steps for conjunction and negation are by an easy verification. Let us treat in detail the remaining case. Let $[\mathrm{A}]\psi$ be a subformula of ϕ and assume that $[\mathrm{A}]\psi \in \mathcal{T}(s)$. As we have seen above, this implies that, for every $\langle s, s'\rangle \in \mathcal{R}_{m(\mathrm{A})}$, we have $\psi \in \mathcal{T}(s')$. By the induction hypothesis, we have $\mathcal{M}, s' \models \psi$. So, $\mathcal{M}, s \models [\mathrm{A}]\psi$.

Now let $[\mathrm{A}]\psi$ be a subformula of ϕ and assume that $\mathcal{M}, s \models [\mathrm{A}]\psi$ and that $[\mathrm{A}]\psi \notin \mathcal{T}(s)$. Due to **(H7)**,

- either $\psi \notin \mathcal{T}(s \cdot i)$ for some $i \in \{1, \ldots, K\}$ and $\mathcal{T}(s \cdot i)$ is not dummy
- or for some $\mathrm{x} \in \mathrm{NOM}(\phi)$, $\psi \notin \mathcal{T}(i_{\mathrm{x}})$ and $\langle \mathrm{A}, \mathrm{x} \rangle$ is a pair of the table of $\mathcal{T}(s)$.

By the induction hypothesis, either $\mathcal{M}, s \cdot i \not\models \psi$ or $\mathcal{M}, i_{\mathrm{x}} \not\models \psi$. However, $\langle s, s' \rangle \in \mathcal{R}_{m(\mathrm{A})}$ since $\langle s, s' \rangle \in R_{\mathrm{A}}$ by construction. Consequently, $\mathcal{M}, s \not\models [\mathrm{A}]\psi$ which leads to a contradiction.

(I) $\rightarrow$ (II) Let $\mathcal{M} = \langle W, (\mathcal{R}_P)_{P \subseteq PAR}, m \rangle$ be a SIM-model and $w_0 \in W$ such that $\mathcal{M}, w_0 \models \phi$. Let $G_0 = \langle f, UF, EF, EQ, NOM, \mathrm{R_N} \rangle$ be a SIM-consistent global information for ϕ built from $\mathcal{M}$ as described in Sect. 4. We define a Hintikka tree $\mathcal{T}$ for ϕ respecting G_0. In the construction of $\mathcal{T}$, we use an auxiliary mapping $\tau : \{1, \ldots, K\}^* \rightarrow W \cup \{\perp\}$ which is defined inductively together with $\mathcal{T}$ as follows.

We first need some notation. Let α be the number of equivalence classes of EQ, and let $h : \mathrm{NOM}(\phi) \rightarrow \{1, \ldots, \alpha\}$ be a mapping that associates each nominal $\mathrm{x} \in \mathrm{NOM}(\phi)$ with the unique element $w_i \in W$ with $m(\mathrm{x}) = \{w_{h(\mathrm{x})}\}$. Let $[\mathrm{A}_1]\psi_1, \ldots, [\mathrm{A}_\beta]\psi_\beta$ be all box formulae in $\mathrm{sub}(\phi)$. For every $w \in W$, we write X_w to denote $\{\psi \in \mathrm{sub}(\phi) \mid \mathcal{M}, w \models \psi\}$ and T_w to denote

$$\{\langle \mathrm{A}, \mathrm{x} \rangle \in \mathrm{P}(\phi) \times \mathrm{NOM}(\phi) \mid \langle w, w_{h(\mathrm{x})} \rangle \in \mathcal{R}_{m(\mathrm{A})}\}.$$

As usual, $\mathcal{T}$ is obtained unravelling $\mathcal{M}$, but taking special care with objects $w_1, \ldots, w_\alpha$. We define τ and $\mathcal{T}$ as follows.

- $\tau(\epsilon) \stackrel{\text{def}}{=} \perp$ and $\mathcal{T}(\epsilon) \stackrel{\text{def}}{=} \perp$.
- If $w_0 = w_j$ for some $j \in \{1, \ldots, \alpha\}$, then, for every $i \in \{1, \ldots, \alpha\}$, $\tau(i) \stackrel{\text{def}}{=} w_i$ and $\mathcal{T}(i) \stackrel{\text{def}}{=} \langle \mathrm{A}, X_{w_i}, T_{w_i} \rangle$ for some arbitrary $\mathrm{A} \in \mathrm{P}(\phi)$ and, for every $i \in \{\alpha + 1, \ldots, K\}$, $\tau(i) \stackrel{\text{def}}{=} \perp$ and $\mathcal{T}(i) \stackrel{\text{def}}{=} \perp$.
- Otherwise ($w_0 \neq w_i$ for every $i \in \{1, \ldots, \alpha\}$), for every $i \in \{1, \ldots, \alpha + 1\}$, $\tau(i) \stackrel{\text{def}}{=} w_{i-1}$ and $\mathcal{T}(i) \stackrel{\text{def}}{=} \langle \mathrm{A}, X_{w_{i-1}}, T_{w_{i-1}} \rangle$ for some arbitrary $\mathrm{A} \in \mathrm{P}(\phi)$, and for every $i \in \{\alpha + 2, \ldots, K\}$, $\tau(i) \stackrel{\text{def}}{=} \perp$ and $\mathcal{T}(i) \stackrel{\text{def}}{=} \perp$.
- For every $s \in \{1, \ldots, K\}^+$,
  - for every $i \in \{\beta + 1, \ldots, K\}$, $\tau(s \cdot i) \stackrel{\text{def}}{=} \perp$ and $\mathcal{T}(s \cdot i) \stackrel{\text{def}}{=} \perp$;
  - if $\tau(s) = \perp$ then, for every $i \in \{1, \ldots, \beta\}$, $\tau(s \cdot i) \stackrel{\text{def}}{=} \perp$ and $\mathcal{T}(s \cdot i) \stackrel{\text{def}}{=} \perp$;
  - otherwise, if $[\mathrm{A}_i]\psi_i \notin \mathcal{T}(s)$ for some $i \in \{1, \ldots, \beta\}$, then
    - $*$ either for some $j \in \{1, \ldots, \alpha\}$, $\langle \tau(s), w_j \rangle \in \mathcal{R}_{m(\mathrm{A}_i)}$ and $\mathcal{M}, w_j \not\models \psi_i$; in that case $\tau(s \cdot i) \stackrel{\text{def}}{=} \perp$ and $\mathcal{T}(s \cdot i) \stackrel{\text{def}}{=} \perp$;
    - $*$ or there is $w' \in W \setminus \{w_1, \ldots, w_\alpha\}$ such that $\langle \tau(s), w' \rangle \in \mathcal{R}_{m(\mathrm{A}_i)}$ and $\mathcal{M}, w' \not\models \psi_i$; in that case $\tau(s \cdot i) \stackrel{\text{def}}{=} w'$ and $\mathcal{T}(s \cdot i) \stackrel{\text{def}}{=} \langle \mathrm{A}_i, X_{w'}, T_{w'} \rangle$.

    If $[\mathrm{A}_i]\psi_i \in \mathcal{T}(s)$ for some $i \in \{1, \ldots, \beta\}$, then $\tau(s \cdot i) \stackrel{\text{def}}{=} \perp$ and $\mathcal{T}(s \cdot i) \stackrel{\text{def}}{=} \perp$.

We can easily check that $\mathcal{T}$ is a Hintikka tree for ϕ respecting G_0.

7 Tree Automata for Relative Formulae

In the second part of the proof of Lemma 9, a SIM-model $\mathcal{M}$ is unravelled in an almost standard way to a Hintikka tree. Thus we have proved a tree model property for SIM—which could also be called a forest model property. In this section, we will exploit this forest model property and describe a decision procedure based on automata on infinite trees, so-called *Büchi tree automata*. For a given SIM-formula ϕ, we construct a Büchi tree automaton $\mathcal{A}_\phi$ that accepts exactly all Hintikka trees for ϕ. At first glance, the construction may look intricate but it simply mimicks the local conditions of the Hintikka trees.

We recall that a *Büchi tree automaton* $\mathcal{A} = \langle \Sigma, Q, \delta, I, F \rangle$ for Σ, K-trees is an operational model where Q is a non-empty, finite set of states, Σ is a finite alphabet, $\delta \subseteq Q \times \Sigma \times Q^K$ is a transition relation, I and F are non-empty subsets of Q, the set of initial states and the set of terminal states, respectively. A *run* r on a Σ, K-tree $\mathcal{T}$ is a Q, K-tree such that, for every $s \in \{1, \ldots, K\}^*$, $\langle r(s), \mathcal{T}(s), r(s \cdot 1), \ldots, r(s \cdot K) \rangle \in \delta$. A run r on $\mathcal{T}$ is *accepting* iff for every path in $\mathcal{T}$ there is a state in F that occurs infinitely often. Deciding whether a Büchi tree automaton for Σ, K-trees has an accepting run can be done in polynomial-time [VW86] (see also [Rab70,EJ88]). For SIM, we only need to consider a restricted class of tree automata, namely those automata in which all the states are terminal, often referred to as *safety automata*.

7.1 The Construction

Before giving the formal definition of $\mathcal{A}_\phi$, we give an intuitive description of it and the conditions it imposes on the trees it accepts:

- Each state consists of a symbolic state and a global information G, and $\mathcal{A}_\phi$ ensures that the global information part of all states involved in an accepting run coincide.
- ϕ and each object nominal in ϕ is found in the label of one of the nodes at the first level of the input tree. Moreover, object nominals are found in the label of the same node if they belong to the same equivalence class according to the EQ component of G.
- Nodes at level ≥ 2 do not have object nominals in their labels.
- If a node is labelled with $\perp$, then so are all its descendants.
- Successors of a node s satisfy conditions imposed by the box formulae in s's label.
- Diamond formulae in a node s's label (i.e., box formulae not in s's label) are either witnessed by one of s's successors or by a node on the first level representing an object nominal.

For those familiar with tree automata, it can be easily seen that the above conditions are all local and can thus be "encoded" in the transition function of a tree automaton. Let us now give the formal definition for $\mathcal{A}_\phi$ when ϕ is a SIM-formula satisfying the hypotheses at the beginning of Sect. 6. $\mathcal{A}_\phi$ is the Büchi tree automaton $\langle \Sigma, Q, \delta, I, Q \rangle$ defined as follows.

1. $\Sigma \stackrel{\text{def}}{=} \mathrm{SYMB}(\phi)$.
2. $Q \stackrel{\text{def}}{=} \{\iota\} \cup \{\langle q, G \rangle \mid q \in \mathrm{SYMB}_G(\phi),\ G \in \mathrm{GCONS}(\phi)\}$.

3. $I \stackrel{\text{def}}{=} \{\iota\}$;
4. $\langle q', a, q'_1, \ldots, q'_K \rangle \in \delta$ iff either
   (first) $q' = \iota$, $a = \perp$, and there is $G = \langle f, UF, EF, EQ, NOM, \mathrm{R_N} \rangle \in \mathrm{GCONS}(\phi)$ such that, for every $i \in \{1, \ldots, K\}$, $q'_i = \langle q_i, G \rangle$ for some $q_i \in \mathrm{SYMB}_G(\phi)$,
   (H2′) there is $i \in \{1, \ldots, K\}$, such that $\phi \in q_i$, and
   (H3′) for every $\mathrm{x} \in \mathrm{NOM}(\phi)$, there is a unique $i \in \{1, \ldots, K\}$, such that $\mathrm{x} \in q_i$, or
   (H5′) $q' = \langle \perp, G \rangle$ for some $G \in \mathrm{GCONS}(\phi)$, $a = \perp$ and, for every $i \in \{1, \ldots, K\}$, $q'_i = q'$; or
   (witnesses) $q' = \langle q, G \rangle$ for some non-dummy $q \in \mathrm{SYMB}_G(\phi)$, $G \in \mathrm{GCONS}(\phi)$, $a = q$ and, for every $i \in \{1, \ldots, K\}$, $q'_i = \langle q_i, G \rangle$ and the following conditions are satisfied:
   (H6′) for every $i \in \{1, \ldots, K\}$, q_i is not a named symbolic state;
   (H7′) if $[\mathrm{B}]\psi \in \mathrm{sub}(\phi) \setminus q$, then
   (a) either there is $i \in \{1, \ldots, K\}$ such that $q_i = \langle \mathrm{B}, X', T' \rangle$ is not dummy and $\psi \notin q_i$;
   (b) or there is $\mathrm{x} \in \mathrm{NOM}(\phi)$ such that $\langle \mathrm{B}, \mathrm{x} \rangle \in T_q$ and $\psi \notin NOM(\mathrm{x})$ (where NOM is the fifth component of G);
   (H8′) for every $i \in \{1, \ldots, K\}$, if q_i is not dummy, then $X_q \sim_{G, \mathrm{A}_{q_i}} X_{q_i}$.

The conditions (**H**i') are the obvious counterparts of the conditions (**H**i). It is worth noting that although Hintikka trees for ϕ require the satisfaction of conditions between trees of the forest, this can be handled by a Büchi tree automaton. Indeed, the symbolic links are encoded locally by the table and by the global information G, which is ensured to coincide on all nodes in a tree accepted by the automaton.

Lemma 10. *A* $\mathrm{SYMB}(\phi), K$*-tree* $\mathcal{T}$ *is a Hintikka tree for* ϕ *iff* $\mathcal{A}_\phi$ *has an accepting run on* $\mathcal{T}$*.*

Proof. Let $\mathcal{T}$ be a Hintikka tree for ϕ respecting the SIM-consistent global information G and $r : \{1, \ldots, K\}^* \to Q$ be the Q, K-tree such that $r(\epsilon) = \iota$ and, for every $s \in \{1, \ldots, K\}^+$, $r(s) = \langle \mathcal{T}(s), G \rangle$. One can check easily that r is an accepting run for $\mathcal{T}$.

For the converse, let $\mathcal{T}$ be an infinite tree accepted by $\mathcal{A}_\phi$, and let $r : \{1, \ldots, K\}^* \to Q$ be an accepting run of $\mathcal{A}_\phi$ on $\mathcal{T}$. Then $r(i) = \langle q_0, G_0 \rangle$ with $\phi \in q_0$ for some $i \in \{1, \ldots, K\}$ and, for every $s \in \{1, \ldots, K\}^+$, if $r(s) = \langle q, G \rangle$, then $G = G_0$ and $q = \mathcal{T}(s)$. By construction, $\mathcal{T}$ is a Hintikka tree for ϕ respecting the SIM-consistent global information G_0.

We are now in the position to establish the main result of the paper.

Theorem 11. *The satisfiability problem for the logic* SIM *is* ExpTime*-complete.*

Proof. The lower bound is by an easy verification from the results in [CL94] and [Hem96, Theorem 5.1]. Let us establish the ExpTime upper bound. Lemma 9 together with Lemma 10 implies that every SIM-formula ϕ is SIM-satisfiable iff $\mathcal{A}_\phi$ accepts at least one tree. Since card(SYMB(ϕ)) $\leq |\phi| \times 2^{|\phi|+|\phi|^2}$ and card(GCONS(ϕ)) is in $2^{\mathcal{O}(|\phi|^3)}$, $\mathcal{A}_\phi$ has $2^{\mathcal{O}(|\phi|^3)}$ states. Moreover, card(δ) is in $2^{\mathcal{O}(|\phi|^4)}$ and checking whether $\langle q, a, q_1, \ldots, q_K \rangle \in \delta$ can be done in time $2^{\mathcal{O}(|\phi|)}$ (using Lemmas 5 and 7). Consequently, computing $\mathcal{A}_\phi$ requires time in $2^{\mathcal{O}(|\phi|^4)}$. Since the emptiness problem for Büchi tree automata of the form $\mathcal{A}_\phi$ can be checked in time $\mathcal{O}(|\delta|^2)$, SIM-satisfiability can be checked in time $2^{\mathcal{O}(|\phi|^4)}$.

7.2 Other constraints on nominals

The parameter nominals in SIM are strong in the sense that two distinct parameter nominals are interpreted by different parameters. This is a constraint introduced in [Kon97,Kon98]. Alternatively, it is possible to relax this condition by allowing that two different parameter nominals can be interpreted identically while preserving the ExpTime-completeness of SIM-satisfiability. Indeed, it is sufficient to add, in a global information G, an equivalence relation for the parameter nominals and to slightly modify the definition of the normal forms $\mathrm{N}_{X,Y}(\mathrm{A})$ (see Sect. 3). Two parameter nominals in the same equivalence class are then interpreted identically. Additionally, constraints of the form "M distinct object [resp. parameter] nominals, $M \geq 2$, cannot be interpreted by the same objects [resp. parameters]" can also be handled by the present framework by requiring, in the equivalence relations for nominals, that each equivalence class has less than M elements.

7.3 A standard version of SIM

Let SIM$^{\mathrm{st}}$ be the multi-modal logic with a universal modality $[U]$, modal connectives of the form $[c_1 \cap \cdots \cap c_n]$, where $\cap$ is interpreted as intersection on binary relations, each c_i is interpreted as a reflexive and symmetric relation, and the logic contains propositional variables and nominals. The logic SIM$^{\mathrm{st}}$ can be viewed as the standard (and simplified) version of SIM; more details about the relationsip between SIM and SIM$^{\mathrm{st}}$ can be found in [Dem99a,DG00]. More importantly, by slightly adapting the ExpTime-completeness proof for SIM one can show the following result.

Corollary 12. *The satisfiability problem for the logic SIM$^{\mathrm{st}}$ is* ExpTime*-complete.*

However, it is open whether replacing symmetry by transitivity in SIM$^{\mathrm{st}}$ preserves decidability (see also Sect. 7.4 below).

7.4 Extensions for FORIN and IND?

At a first glance, it seems as if the decision procedure for SIM could be easily adapted to IND and FORIN since transitivity can be handled for the following aspects:

- the constraints on $\mathrm{R_N}$ (by imposing the appropriate frame condition);
- the definition of $\sim_{G,\mathrm{A}}$ (by updating the propagation rules for the [A]-formulae);
- the definition of the family $(\mathcal{R}_P)_{P \subseteq PAR}$ in the proof of Lemma 9((II) $\rightarrow$ (I)) (by considering the appropriate closure operation on relations).

However, such an adaptation does not allow us to prove the point (3) in the proof of Lemma 9((II) $\rightarrow$ (I)). More precisely, Case (2) is problematic.

By way of example, consider the logic FORIN. Suppose that we have updated the conditions for $\mathrm{R_N}$ and $\sim_{G,\mathrm{A}}$ adequately. In the proof of Lemma 9((II) $\rightarrow$ (I)), assume that $\mathrm{R_A}$ was defined as the reflexive and transitive closure of $\mathrm{S(A)} = \mathrm{R_N(A)} \cup \mathrm{R_A} \cup \mathrm{R'_A}$. Let us consider the subcases 2.3 ($\mathcal{T}(s)$ is not and $\mathcal{T}(s')$ is a named symbolic state). Then $\langle s, s' \rangle \in \mathcal{R}_{m(\mathrm{A})}$ entails, for every $i \in \{i_1, \ldots, i_k\}$, $\langle s, s' \rangle \in \mathcal{R}_{\{i\}}$, which implies the existence of the following paths:

$$s = s_1^{i_1} \xrightarrow{\mathrm{S}(\mathrm{A}_1^{i_1})} s_2^{i_1} \xrightarrow{\mathrm{S}(\mathrm{A}_2^{i_1})} \ldots \xrightarrow{\mathrm{S}(\mathrm{A}_{n_{i_1}}^{i_1})} s_{n_{i_1}+1}^{i_1} = s'$$

$$\vdots$$

$$s = s_1^{i_j} \xrightarrow{\mathrm{S}(\mathrm{A}_1^{i_j})} s_2^{i_j} \xrightarrow{\mathrm{S}(\mathrm{A}_2^{i_j})} \ldots \xrightarrow{\mathrm{S}(\mathrm{A}_{n_{i_j}}^{i_j})} s_{n_{i_j}+1}^{i_j} = s'$$

$$\vdots$$

$$s = s_1^{i_k} \xrightarrow{\mathrm{S}(\mathrm{A}_1^{i_k})} s_2^{i_1} \xrightarrow{\mathrm{S}(\mathrm{A}_2^{i_k})} \ldots \xrightarrow{\mathrm{S}(\mathrm{A}_{n_{i_k}}^{i_k})} s_{n_{i_k}+1}^{i_k} = s'$$

where, for every $j \in \{1, \ldots, k\}$, $n_{i_j} \geq 0$ and for every $j' \in \{0, \ldots, n_{i_j}\}$, $i_j \in m(\mathrm{A}_{j'}^{i_j})$. Because of the presence of named symbolic states (and therefore object nominals in the language) and the transitive closure involved, the above paths are of unbounded length and they may be different. By contrast, for SIM, only paths of length one need to be handled simultaneously, which can be done locally. If $[\mathrm{A}]\psi \in \mathcal{T}(s)$, we need to ensure $\psi \in \mathcal{T}(s')$, which would involve a path of the form

$$s = s_1 \xrightarrow{\mathrm{S(B_1)}} s_2 \xrightarrow{\mathrm{S(B_2)}} \ldots \xrightarrow{\mathrm{S(B}_n)} s_{n+1} = s'$$

where, for every $j \in \{1, \ldots, k\}$ and $j' \in \{0, \ldots, n\}$, $i_j \in m(\mathrm{B}_{j'})$. Obviously, this is not possible using a simple "local" propagation.

By contrast, in the absence of object nominals in the language, in the remaining subcase 2.2 (without named symbolic states), the existence of such a path σ is guaranteed. Indeed, if $\langle s, s' \rangle \in \mathcal{R}_{m(\mathrm{A})}$, then there is unique path σ of minimal length of the above form satisfying the required condition. Consequently, by slightly adapting the developments for SIM, we can show:

Theorem 13. *The satisfiability problem for the logics FORIN and IND without object nominals is* EXPTIME*-complete.*

8 Conclusion

On the basis of existing automata-theoretic techniques for logical problems, we have shown that the logic SIM introduced in [Kon98] has an EXPTIME-complete satisfiability problem, improving significantly the best known upper bound from [DK98]. The proof is by a reduction to the emptiness problem for Büchi automata on infinite trees. The most original parts of this reduction rely on the normalisation of parameter expressions for nominals, on the introduction of global information for models (extending what is done in [SV01] for nominals and the universal modality), and on our treatment of intersection for relations. The proof for SIM can be successfully adapted to the logics IND and FORIND [Kon98, Sect. 8] with minor changes only in the case we discard object nominals from the language. By contrast, the decidability status of full IND and full FORIN remains a challenging open question. This highlights the technical difficulty encountered to establish the EXPTIME upper bound for SIM satisfiability.

These new results obtained with automata-theoretic techniques show that such techniques are also powerful for information logics. Indeed, the automata framework can cope uniformely with object and parameter nominals, with Boolean parameter expressions, with the universal modality and with local conditions such as reflexivity and symmetry. For transitivity, the adaptation of the method is still unkown. In this paper, we have actually used a small fragment of the automata machinery, namely the Büchi tree automata in which all the states are terminal, leaving some room for further extensions with richer operational models such as the tree automata with parity acceptance conditions [Var98].

References

[Bal96] Ph. Balbiani. Modal logics with relative accessibility relations. In D. Gabbay and H.J. Ohlbach, editors, *Conference on Formal and Applied Practical Reasoning*, volume 1085 of *Lecture Notes in Computer Science*, pages 29–41. Springer, 1996.

[Bla00a] P. Blackburn. Internalizing labelled deduction. *Journal of Logic and Computation*, 10(1):137–168, 2000.

[Bla00b] P. Blackburn. Representation, reasoning, and relational structures: a hybrid logic manifesto. *Logic Journal of the IGPL*, 8(3):339–365, 2000.

[Büc62] R. Büchi. On a decision method in restricted second-order arithmetic. In *International Congress on Logic, Method and Philosophical Science'60*, pages 1–11, 1962.

[CDGL99] D. Calvanese, G. De Giacomo, and M. Lenzerini. Reasoning in expressive description logics with fixpoints based on automata on infinite trees. In *IJCAI'99*, pages 84–89, 1999.

[CGL02] D. Calvanese, G. De Giacomo, and M. Lenzerini. Description logics: Foundations for class-based knowledge representation. In *LICS'02*, pages 359–370. IEEE Computer Society, 2002.

[CL94] C. Chen and I. Lin. The complexity of propositional modal theories and the complexity of consistency of propositional modal theories. In A. Nerode and Yu. V. Matiyasevich, editors, *LFCS-3, St. Petersburg*, volume 813 of *Lecture Notes in Computer Science*, pages 69–80. Springer, 1994.

[Dem99a] S. Demri. A logic with relative knowledge operators. *Journal of Logic, Language, and Information*, 8(2):167–185, 1999.

[Dem99b] S. Demri. Sequent calculi for nominal tense logics: a step towards mechanization? In N. Murray, editor, *TABLEAUX'99*, volume 1617 of *Lecture Notes in Artificial Intelligence*, pages 140–154. Springer, 1999.

[DG00] S. Demri and D. Gabbay. On modal logics characterized by models with relative accessibility relations: Part II. *Studia Logica*, 66(3):349–384, 2000.

[DK98] S. Demri and B. Konikowska. Relative similarity logics are decidable: reduction to FO^2 with equality. In J. Dix, L. Fariñas del Cerro, and U. Furbach, editors, *JELIA'98*, volume 1489 of *Lecture Notes in Artificial Intelligence*, pages 279–293. Springer, 1998.

[DO02] S. Demri and E. Orlowska. *Incomplete Information: Structure, Inference, Complexity*. EATCS Monographs. Springer, Berlin, 2002.

[EJ88] A. Emerson and C. Jutla. The complexity of tree automata and logics of programs. In *29th Annual Symposium on Foundations of Computer Science*, pages 328–337. IEEE Computer Society Press, 1988.

[EJ99] A. Emerson and C. Jutla. The complexity of tree automata and logics of programs. *SIAM Journal of Computing*, 29:132–158, 1999.

[EJS01] A. Emerson, C. Jutla, and A. Sistla. On model-checking for μ-calculus and its fragments. *Theoretical Computer Science*, 258(1-2):491–522, 2001.

[GKV97] E. Grädel, Ph. Kolaitis, and M. Vardi. On the decision problem for two-variable first-order logic. *The Bulletin of Symbolic Logic*, 3(1):53–69, 1997.

[GP90] G. Gargov and S. Passy. A note on Boolean modal logic. In P. Petkov, editor, *Summer School and Conference on Mathematical Logic '88*, pages 299–309. Plenum, New York, 1990.

[Grä99] E. Grädel. Why are modal logics so robustly decidable? *Bulletin of the EATCS*, 68:90–103, 1999.

[Hem96] E. Hemaspaandra. The price of universality. *Notre Dame Journal of Formal Logic*, 37(2):173–203, 1996.

[Hol97] G. Holzmann. The model checker SPIN. *IEEE Transations on Software Engineering*, 23(5):279–295, 1997.

[Kah62] A. Kahr. Improved reductions of the Entscheidungsproblem to subclasses of $\forall\exists\forall$ formulas. In *Symposium on Mathematical Theory of Automata*, pages 57–70. Brooklyn Polytechnic Institute, New York, 1962.

[Kon97] B. Konikowska. A logic for reasoning about relative similarity. *Studia Logica*, 58(1):185–226, 1997.

[Kon98] B. Konikowska. A logic for reasoning about similarity. In *[Orło98]*, pages 462–491, 1998.

[KSV02] O. Kupferman, U. Sattler, and M. Vardi. The complexity of the graded μ-calculus. In A. Voronkov, editor, *Proc. 18th Conference on Automated Deduction, Copenhagen, Denmark*, volume 2392 of *Lecture Notes in Computer Science*, pages 423–437. Springer, Berlin, 2002.

[KVW00] O. Kupferman, M. Vardi, and P. Wolper. An automata-theoretic approach to branching-time model checking. *Journal of the Association for Computing Machinery*, 47(2):312–360, 2000.

[LS01] C. Lutz and U. Sattler. The complexity of reasoning with Boolean modal logics. In M. de Rijke, H. Wansing, F. Wolter, and M. Zakharyaschev, editors, *Advances in Modal Logics (AIML'2000)*. CSLI, 2001. To appear.

[OP81] E. Orłowska and Z. Pawlak. Expressive power of knowledge representation system. Technical Report 432, Institute of Computer Science, Polish Academy of Sciences, 1981.

[Orło98] E. Orłowska, editor. *Incomplete Information: Rough Set Analysis.* Studies in Fuzziness and Soft Computing. Physica, Heidelberg, 1998.

[Pap94] Ch. Papadimitriou. *Computational Complexity.* Addison-Wesley, Reading, MA, 1994.

[Paw81] Z. Pawlak. Information systems theoretical foundations. *Information Systems*, 6(3):205–218, 1981.

[PS98] L. Polkowski and A. Skowron, editors. *Rough Sets in Knowledge Discovery.* Studies in Fuzziness and Soft Computing. Physica, Heidelberg, 1998.

[Rab69] M. Rabin. Decidability of second-order theories and automata on infinite trees. *Transactions of the American Mathematical Society*, 41:1–35, 1969.

[Rab70] M. Rabin. Weakly definable relations and special automata. In Y. Bar-Hillel, editor, *Symposium on Mathematical Logic and Foundations of Set Theory*, pages 1–23. North-Holland, 1970.

[SE89] R. Streett and A. Emerson. An automata theoretic decision procedure for the propositional μ-calculus. *Information and Computation*, 81:249–264, 1989.

[Sel01] J. Seligman. Internalization: the case of hybrid logics. *Journal of Logic and Computation*, 11:671–689, 2001.

[Spa93] E. Spaan. *Complexity of Modal Logics.* PhD thesis, ILLC, Amsterdam University, 1993.

[Ste98] J. Stepaniuk. Rough relations and logics. In *[PS98]*, pages 248–260, 1998.

[SV01] U. Sattler and M. Vardi. The hybrid mu-calculus. In A. Leitsch, R. Goré, and T. Nipkow, editors, *IJCAR'01*, volume 2083 of *Lecture Notes in Artificial Intelligence*, pages 76–91. Springer, 2001.

[Tho90] W. Thomas. Automata on infinite objects. In J. Van Leeuwen, editor, *Handbook of Theoretical Computer Science, Volume B, Formal models and semantics*, pages 133–191. Elsevier, 1990.

[Tza99] M. Tzakova. Tableau calculi for hybrid logics. In N. Murray, editor, *International Conference on Theorem Proving with Analytic Tableaux and Related Methods, Saratoga Springs, USA*, pages 278–292. Lecture Notes in Artificial Intelligence, Vol. 1617. Springer, Berlin, 1999.

[Vak87] D. Vakarelov. Abstract characterization of some knowledge representation systems and the logic NIL of nondeterministic information. In Ph. Jorrand and V. Sgurev, editors, *Artificial Intelligence: Methodology, Systems, Applications*, pages 255–260. North-Holland, Amsterdam, 1987.

[Vak91] D. Vakarelov. Modal logics for knowledge representation systems. *Theoretical Computer Science*, 90:433–456, 1991.

[Vak98] D. Vakarelov. Information systems, similarity and modal logics. In *[Orło98]*, pages 492–550, 1998.

[Var97] M. Vardi. Alternating automata: unifying truth and validity checking for temporal logics. In *CADE'97*, volume 1249 of *Lecture Notes in Artificial Intelligence*, pages 191–206. Springer, 1997.

[Var98] M. Vardi. Reasoning about past with two-way automata. In *ICALP'98*, volume 1443 of *Lecture Notes in Computer Science*, pages 628–641. Springer, 1998.

[VW86] M. Vardi and P. Wolper. Automata-theoretic techniques for modal logics of programs. *Journal of Computer and System Sciences*, 32:183–221, 1986.

[VW94] M. Vardi and P. Wolper. Reasoning about infinite computations. *Information and Computation*, 115:1–37, 1994.