

Verification and Synthesis in Description Logic Based Dynamic Systems (Abridged Version)*

Diego Calvanese¹, Giuseppe De Giacomo², Marco Montali¹, and Fabio Patrizi²

¹ Free University of Bozen-Bolzano, Piazza Domenicani 3, 39100 Bolzano, Italy
lastname@inf.unibz.it

² Sapienza Università di Roma, Via Ariosto, 25, 00185 Rome, Italy
lastname@dis.uniroma1.it

Abstract. We devise a general framework for formalizing Description Logic Based Dynamic Systems that is parametric w.r.t. the description logic knowledge base and the progression mechanism. We study verification and adversarial synthesis for specifications expressed in a variant of first-order μ -calculus, with a controlled form of quantification across successive states. We provide key decidability results for both verification and synthesis, under a “bounded-state” assumption. We then study two instantiations of this general framework, where the knowledge base is expressed in *DL-Lite* and *ALCQT*, respectively.

1 Introduction

Integrating semantic mechanisms like Description Logics (DLs), to describe static knowledge, and full-fledged mechanisms like transition systems, to describe dynamics is of great interest, especially in the context of a semantic view of data-aware and artifact-based processes [16,7,13]. The tight combination of these two aspects into a single logical formalism is notoriously difficult [20]. In particular, due to the nature of DL assertions we get one of the most difficult kinds of static constraints for reasoning about actions [18]. This combination gives rise to a semantics based on two-dimensional structures (DL domain + dynamics/time), which easily leads to undecidability [20,14]. To regain decidability, one has to restrict the combination only to concepts (vs. roles) [2]. This limitation is unsatisfactory in all those applications where we want to progress a full-fledged knowledge base (KB) representing a shared conceptualization of a domain of interest, such as an ontology, (a formalization of) a UML Class Diagram, or an Ontology Based Data Access System [9].

Recently, to overcome such difficulties, a looser coupling of the static and dynamic representation mechanisms has been proposed, giving rise to a rich body of research [3,10,5]. Virtually all such work is implicitly or explicitly based on Levesque’s functional approach [15], where a KB is seen as a system that provides the ability of querying its knowledge in terms of logical implication/certain answers (“ask” operation), and the ability of progressing it through forms of updates (“tell” operation). As a consequence,

* This paper is an abridged version of [11], best paper award at RR 2013.

the knowledge description formalism becomes decoupled from the formalism that describes the progression: we can define the dynamics through a transition system, whose states are DL KBs, and transitions are labeled by the action (with object parameters) that causes the transition. The key issue in this context is that such transition systems are *infinite* in general, and hence some form of faithful abstraction is needed. Note that, if for any reason we learn that the number of states is finite, then verifying dynamic properties over such systems amounts to a form a finite-state model checking.

In this paper, we follow this approach and devise a general framework for *DL Based Dynamic Systems*, which is parametric w.r.t. the DL used for the knowledge base and the mechanism used for progressing the state of the system. We study verification and (adversarial) synthesis for specifications expressed in a variant of first-order μ -calculus, with a controlled form of quantification across successive states. We recall that μ -calculus subsumes virtually all logics used in verification, including LTL, CTL and CTL*. Adversarial synthesis for μ -calculus captures a wide variety of synthesis problems, including conditional planning with full-observability, behaviour composition, and several other sophisticated forms of synthesis and planning [12].

We provide key decidability results for verification under a “bounded-state” assumption. Such assumption states that while progressing, the system can change arbitrarily the stored individuals but their total number at each time point cannot exceed a certain bound. Notice that along an infinite run (and hence in the overall transition system), the total number of individuals can still be infinite. We then turn to adversarial synthesis, where we consider the system engaged in a sort of game with an adversarial environment. The two agents (the system and the environment) move in alternation, and the problem is to synthesize a strategy for the system to force the evolution of the game so as to satisfy a given synthesis specification. Such a specification is expressed in the above first-order μ -calculus, using the temporal operators to express that the system is able to force a formula Φ in the next state regardless of the environment moves, as in the strategy logic ATL [1]. We show again decidability under the “bounded-state” assumption.

This paper is an abridged version of [11], best paper award at the 7th International Conference on Web Reasoning and Rule Systems (RR 2013). Proofs are omitted and can be found in [11].

2 Framework

In this paper we follow Leveque’s functional approach [15]. We introduce Description Logic Based Dynamic Systems (DLDSs), which are systems constituted by a DL knowledge base, consisting of a fixed TBox and an ABox that changes as the system evolves, and a set of actions that step-wise progress the knowledge base by changing the ABox. We formalize such systems in a general form, referring neither to any specific DL nor to any specific action representation formalism, and making only minimal assumptions on the various components. In Section 6, we show concrete instantiations of the framework.

Object universe. We fix a countably infinite (object) universe Δ of individuals. These constants act as standard names and allow us to refer to individuals across time.

DL knowledge bases. A (DL) *knowledge base* (KB) (T, A) is constituted by a TBox T representing the intensional knowledge about the domain of interest in terms of concepts

and roles, and an ABox A representing extensional knowledge about individuals. The TBox T is a finite set of universal assertions (we don't allow nominals, to avoid confusion between intensional and extensional levels). The ABox A is a finite set of *assertions* consisting of facts, i.e., atomic formulas of the form $N(d)$ and $P(d, d')$, where N is a concept name, P is a role name (we use binary roles, but all our results can be extended to n -ary relations), and d, d' are individual constants in Δ . We use $\text{ADOM}(A)$ to denote the set of constants actually appearing in A . We adopt the standard DL semantics based on first-order interpretations and on the notion of (first-order) model. Naturally, the only TBoxes of interest for us are those that are *satisfiable*, i.e., admit at least one model. We say that an ABox A is *consistent w.r.t. a TBox T* if (T, A) is satisfiable, and that (T, A) *logically implies* an ABox assertion α , denoted $(T, A) \models \alpha$, if every model of (T, A) is also a model of α . As usual in DLs, we assume that reasoning, i.e., checking KB satisfiability and logical implication are decidable tasks.

Description Logic based Dynamic Systems. A *Description Logic based Dynamic System (DLDS)* is a tuple $\mathcal{S} = (T, A_0, \Gamma)$, where (T, A_0) is a KB and Γ is a finite set of *actions*. The set $\text{ADOM}(A_0)$ of constants in A_0 are called *distinguished* and we denote them by C , i.e., $C = \text{ADOM}(A_0)$. Such constants play a special role because they are the only ones that can be used in verification formulas (see Section 4). Actions are parametrized, and when an action is executed formal parameters are substituted with individual constants from Δ . Formal parameters are taken from a countably infinite set P of *parameter names*. Interestingly, we do not assume that the number of formal parameters of an action is fixed a priori, but we allow it to depend on the KB (actually, the ABox, since the TBox is fixed) on which the action is executed.

Notice that this mechanism is more general than what typically found in procedures/functions of programming languages, where the number of parameters is fixed by the signature of the procedure/function. Our mechanism is directly inspired by web systems, in which input forms are dynamically constructed and customized depending on data already acquired. For example, when inserting author data in a conference submission system, the input fields that are presented to the user depend on the (previously specified) number of authors.

Once formal parameters are substituted by actual ones, executing the action has the effect of generating a new ABox. We use \mathcal{A}_T to denote the set of all ABoxes that can be constructed using concept and role names in T , and individuals in Δ . Formally, each action in Γ has the form (π, τ) , where

- $\pi : \mathcal{A}_T \rightarrow 2^P$ is a *parameter selection function* that, given an ABox A , returns the *finite* set $\pi(A) \subseteq P$ of parameters of interest for τ w.r.t. A (see below);
- $\tau : \mathcal{A}_T \times \Delta^P \mapsto \mathcal{A}_T$, is a (partial) *effect function* that, given an ABox A and a *parameter assignment* $m : \pi(A) \rightarrow \Delta$, returns (if defined) the ABox $A' = \tau(A, m)$, which (i) is consistent wrt T , and (ii) contains only constants in $\text{ADOM}(A) \cup \text{IM}(m)$.³

Observe that since $\pi(A)$ is finite, so is $\text{IM}(m)$, thus only finitely many new individuals, w.r.t. A , can be added to A' .

We focus our attention on DLDS's that are *generic*, which intuitively means that the two functions constituting an action are invariant w.r.t. renaming of individuals⁴.

³ By $\text{IM}(\cdot)$ we denote the *image* of a function.

⁴ This name is due to the notion of genericity in databases, also called *uniformity* in [6].

Genericity is a natural assumption that essentially says that the properties of individuals are only those that can be inferred from the KB. To capture genericity, we first introduce the notion of equivalence of ABoxes modulo renaming. Specifically, given two ABoxes A_1, A_2 , sets $S_1 \supseteq \text{ADOM}(A_1) \cup C$ and $S_2 \supseteq \text{ADOM}(A_2) \cup C$, and a bijection $h : S_1 \rightarrow S_2$ that is the identity on C , we say that A_1 and A_2 are *logically equivalent modulo renaming h w.r.t. a TBox T* , written $A_1 \cong_T^h A_2$, if: (i) for each assertion α_1 in A_1 , $(T, A_2) \models h(\alpha_1)$, (ii) for each assertion α_2 in A_2 , $(T, A_1) \models h^{-1}(\alpha_2)$, where $h(\alpha_1)$ (resp. $h^{-1}(\alpha_2)$) is a new assertion obtained from α_1 (resp., α_2), by replacing each occurrence of an individual $d \in \Delta$ with $h(d)$ (resp., $h^{-1}(d)$). We say that A_1 and A_2 are *logically equivalent modulo renaming w.r.t. T* , written $A_1 \cong_T A_2$, if $A_1 \cong_T^h A_2$ for some h . We omit T when clear from the context.

A DLDS $\mathcal{S} = (T, A_0, \Gamma)$ is *generic*, if for every $(\pi, \tau) \in \Gamma$ and every $A_1, A_2 \in \mathcal{A}_T$ s.t. $A_1 \cong_T A_2$, we have that (i) $\pi(A_1) = \pi(A_2)$, and (ii) for every two parameter assignments $m_1 : \pi(A_1) \rightarrow \Delta$ and $m_2 : \pi(A_2) \rightarrow \Delta$, if there exists a bijection $h : \text{ADOM}(A_1) \cup C \cup \text{IM}(m_1) \rightarrow \text{ADOM}(A_2) \cup C \cup \text{IM}(m_2)$ s.t. $A_1 \cong_T^h A_2$, then, whenever $A'_1 = \tau(A_1, m_1)$ is defined then also $A'_2 = \tau(A_2, m_2)$ is defined and vice-versa, and moreover $A'_1 \cong_T^h A'_2$.

DLDS Transition System. The dynamics of a DLDS \mathcal{S} is characterized by the *transition system* $\mathcal{Y}_{\mathcal{S}}$ it generates. The kind of transition systems we consider here have the general form $\mathcal{Y} = (U, T, \Sigma, s_0, \text{abox}, \Rightarrow)$, where: (i) U is the universe of individual constants, which includes the distinguished constants C ; (ii) T is a TBox; (iii) Σ is a set of states; (iv) $s_0 \in \Sigma$ is the initial state; (v) abox is a function that, given $s \in \Sigma$, returns an ABox that has terms of Δ as individuals, and conforms to T ; (vi) $\Rightarrow \subseteq \Sigma \times \mathcal{L} \times \Sigma$ is a labeled transition relation between states, where \mathcal{L} is the set of labels. We also introduce an unlabeled transition relation $\Rightarrow \subseteq \Sigma \times \Sigma$ obtained from the labeled transition relation by projecting out the labels (we use this notion when dealing with verification properties). Given a DLDS $\mathcal{S} = (T, A_0, \Gamma)$, its (*generated*) *transition system* $\mathcal{Y}_{\mathcal{S}} = (U, T, \Sigma, s_0, \text{abox}, \Rightarrow)$ is defined as: (i) $U = \Delta$; (ii) abox is the identity function (thus $\Sigma \subseteq \mathcal{A}_T$); (iii) $s_0 = A_0$; (iv) $\Rightarrow \subseteq \Sigma \times \mathcal{L} \times \Sigma$ is a labeled transition relation where $\mathcal{L} = \Gamma \times \mathcal{M}$, with \mathcal{M} the domain of action parameter assignments, is a set of labels containing one pair (a, m) for every action $a \in \Gamma$ and corresponding parameter assignment $m \in \mathcal{M}$; (v) Σ and \Rightarrow are defined by mutual induction as the smallest sets satisfying the following property: if $A \in \Sigma$ then for every $(\pi, \tau) \in \Gamma$, $m : \pi(A) \rightarrow \Delta$, and $A' \in \mathcal{A}_T$, s.t. $A' = \tau(A, m)$, we have $A' \in \Sigma$ and $A \xRightarrow{\ell} A'$, s.t. $\ell = ((\pi, \tau), m)$.

3 Specification Logic μDL_p

To specify dynamic properties over DLDSs, we use a first-order variant of μ -calculus [19]. μ -calculus is virtually the most powerful temporal logic used for model checking of finite-state transition systems, and is able to express both linear time logics such as LTL and PSL, and branching time logics such as CTL and CTL*. Technically, μ -calculus separates local properties, asserted on the current state or on states that are immediate successors of the current one, from properties talking about states that are arbitrarily far away from the current one [19]. The latter are expressed through the use of fixpoints.

In our variant of μ -calculus, we allow local properties to be expressed as queries in any language for which query entailment for DL KBs is decidable [8]. Specifically, given a KB (T, A) , a query Q , and an assignment v for the free variables of Q , we say that Qv is *entailed* by (T, A) , if $(T, A) \models Qv$, i.e., (T, A) logically implies the formula Qv obtained from Q by substituting its free variables according to v . Notice that the set of v such that $(T, A) \models Qv$ are the so-called *certain answers*.

Following [4], we allow for a controlled form of first-order quantification across states: it ranges over individual objects across time only as long as they persist in the active domain. Formally, we define the logic μDL_p as:

$$\Phi ::= Q \mid \neg\Phi \mid \Phi_1 \wedge \Phi_2 \mid \exists x. \text{LIVE}(x) \wedge \Phi \mid \text{LIVE}(\mathbf{x}) \wedge \langle \neg \rangle \Phi \mid \text{LIVE}(\mathbf{x}) \wedge [\neg] \Phi \mid Z \mid \mu Z. \Phi$$

where Q is a (possibly open) query as described above, in which the only constants that may appear are those in C , Z is a second order predicate variable (of arity 0), and, with a slight abuse of notation, we write $\text{LIVE}(x_1, \dots, x_n) = \bigwedge_{i \in \{1, \dots, n\}} \text{LIVE}(x_i)$. For μDL_p , the following assumption holds: in $\text{LIVE}(\mathbf{x}) \wedge \langle \neg \rangle \Phi$ and $\text{LIVE}(\mathbf{x}) \wedge [\neg] \Phi$, the variables \mathbf{x} are exactly the free variables of Φ , once we substitute to each bounded predicate variable Z in Φ its bounding formula $\mu Z. \Phi'$. We use the usual abbreviations, including: $\text{LIVE}(\mathbf{x}) \rightarrow \langle \neg \rangle \Phi = \neg(\text{LIVE}(\mathbf{x}) \wedge [\neg] \neg \Phi)$ and $\text{LIVE}(\mathbf{x}) \rightarrow [\neg] \Phi = \neg(\text{LIVE}(\mathbf{x}) \wedge \langle \neg \rangle \neg \Phi)$. Intuitively, the use of $\text{LIVE}(\cdot)$ in μDL_p ensures that individuals are only considered if they persist along the system evolution, while the evaluation of a formula with individuals that are not present in the current database trivially leads to false or true.

The formula $\mu Z. \Phi$ denotes the least fixpoint of the formula Φ (seen as the predicate transformer $\lambda Z. \Phi$). As usual in μ -calculus, formulae of the form $\mu Z. \Phi$ must obey to the *syntactic monotonicity* of Φ w.r.t. Z , which states that every occurrence of the variable Z in Φ must be within the scope of an even number of negation symbols. This ensures that the least fixpoint $\mu Z. \Phi$ always exists.

The semantics of μDL_p formulae is defined over possibly infinite transition systems of the form $(U, T, \Sigma, s_0, \text{abox}, \Rightarrow)$ seen above. Since μDL_p also contains formulae with both individual and predicate free variables, given a transition system \mathcal{T} , we introduce an *individual variable valuation* v , i.e., a mapping from individual variables x to U , and a *predicate variable valuation* V , i.e., a mapping from the predicate variables Z to subsets of Σ . With these three notions in place, we assign meaning to formulae by associating to \mathcal{T} , v , and V an *extension function* $(\cdot)_{v,V}^{\mathcal{T}}$, which maps formulae to subsets of Σ . Formally, the extension function $(\cdot)_{v,V}^{\mathcal{T}}$ is defined inductively as:

$$\begin{aligned} (Q)_{v,V}^{\mathcal{T}} &= \{s \in \Sigma \mid (T, \text{abox}(s)) \models Qv\} \\ (\neg\Phi)_{v,V}^{\mathcal{T}} &= \Sigma \setminus (\Phi)_{v,V}^{\mathcal{T}} \\ (\Phi_1 \wedge \Phi_2)_{v,V}^{\mathcal{T}} &= (\Phi_1)_{v,V}^{\mathcal{T}} \cap (\Phi_2)_{v,V}^{\mathcal{T}} \\ (\exists x. \text{LIVE}(x) \wedge \Phi)_{v,V}^{\mathcal{T}} &= \{s \in \Sigma \mid \exists d \in \text{ADOM}(\text{abox}(s)). s \in (\Phi)_{v, [x/d], V}^{\mathcal{T}}\} \\ (\text{LIVE}(\mathbf{x}) \wedge \langle \neg \rangle \Phi)_{v,V}^{\mathcal{T}} &= \{s \in \Sigma \mid \mathbf{x}/\mathbf{d} \in v \text{ implies } \mathbf{d} \subseteq \text{ADOM}(\text{abox}(s)) \text{ and } \exists s'. s \Rightarrow s' \text{ and } s' \in (\Phi)_{v,V}^{\mathcal{T}}\} \\ (\text{LIVE}(\mathbf{x}) \wedge [\neg] \Phi)_{v,V}^{\mathcal{T}} &= \{s \in \Sigma \mid \mathbf{x}/\mathbf{d} \in v \text{ implies } \mathbf{d} \subseteq \text{ADOM}(\text{abox}(s)) \text{ and } \forall s'. s \Rightarrow s' \text{ implies } s' \in (\Phi)_{v,V}^{\mathcal{T}}\} \\ (Z)_{v,V}^{\mathcal{T}} &= V(Z) \\ (\mu Z. \Phi)_{v,V}^{\mathcal{T}} &= \bigcap \{\mathcal{E} \subseteq \Sigma \mid (\Phi)_{v, V[Z/\mathcal{E}]}^{\mathcal{T}} \subseteq \mathcal{E}\} \end{aligned}$$

Intuitively, $(\cdot)_{v,V}^{\mathcal{T}}$ assigns to such constructs the following meaning. (i) The boolean connectives have the expected meaning. (ii) The quantification of individuals is done over the individuals of the “current” ABox, using the special $\text{LIVE}(\cdot)$ predicate (*active domain*

quantification). Notice that such individuals can be referred in a later state, provided that they persist in between (see below). (iii) The extension of $\langle - \rangle \Phi$ consists of the states s s.t.: for *some* successor state s' of s , Φ holds in s' under v . (iv) The extension of $\mu Z.\Phi$ is the *smallest subset* \mathcal{E}_μ of Σ s.t., assigning to Z the extension \mathcal{E}_μ , the resulting extension of Φ (under v) is contained in \mathcal{E}_μ . That is, the extension of $\mu Z.\Phi$ is the *least fixpoint* of the operator $(\Phi)_{v, V[Z/\mathcal{E}]}$, where $V[Z/\mathcal{E}]$ denotes the predicate valuation obtained from V by forcing the valuation of Z to be \mathcal{E} . When Φ is a closed formula, $(\Phi)_{v, V}$ does not depend on v or V , and we denote the extension of Φ simply by $(\Phi)^\mathcal{Y}$. A closed formula Φ holds in a state $s \in \Sigma$ if $s \in (\Phi)^\mathcal{Y}$. In this case, we write $\mathcal{Y}, s \models \Phi$. A closed formula Φ holds in \mathcal{Y} , denoted by $\mathcal{Y} \models \Phi$, if $\mathcal{Y}, s_0 \models \Phi$. Given a DLDS \mathcal{S} , we say that $\mathcal{S} \models \Phi$ if $\mathcal{Y}_\mathcal{S} \models \Phi$. We call *model checking* the problem of verifying whether $\mathcal{Y} \models \Phi$ holds.

4 Verification

Next we focus on checking a μDL_p formula against a DLDS. It is easy to show, by reduction from the halting problem, that this is in general undecidable, even under the assumption of genericity (see, e.g., [5]). On the other hand, it can be shown that for finite Δ , the problem is decidable: only finitely many ABoxes exist, thus, by quantifier elimination, one can reduce the problem to model checking of propositional μ -calculus. Undecidability calls for the identification of decidable classes of the problem. To this end, we focus our attention on a particular class of DLDS, which we call *state-bounded*. Under the assumption of genericity, we are able to prove decidability of verification for state-bounded DLDS. A *state-bounded* DLDS \mathcal{K} is one for which there exists a finite bound b s.t., for each state s of $\mathcal{Y}_\mathcal{K}$, $|\text{ADOM}(\text{abox}(s))| < b$. Observe that state-bounded DLDS contain in general infinitely many states, and that a DLDS \mathcal{K} can be state-unbounded even if, for every state s of $\mathcal{Y}_\mathcal{K}$, $|\text{ADOM}(\text{abox}(s))|$ is finite (but not bounded). W.l.o.g., for state-bounded DLDS, we assume that the maximum number n of parameters that actions may request is known (nevertheless, it can be obtained can be obtained in PSPACE.) We are able to show the following result.

Theorem 1. *Model checking of μDL_p over a state-bounded, generic DLDS \mathcal{K} is decidable, and can be reduced to model checking of propositional μ -calculus over a finite-state transition system, whose number of state is at most exponential in the size of \mathcal{K} .*

5 Adversarial Synthesis

We turn now to the problem of adversarial synthesis, i.e., we consider a setting in which two agents act in turn as adversaries. The first agent, called *environment*, acts autonomously, whereas we control the second agent, called *system*. The joint behavior of the two agents gives rise to a so-called *two-player game structure* (2GS) [12], which can be seen as the arena of a game. On top of the 2GS we can formulate, using variants of μ -calculus, what the system should obtain in spite of the adversarial moves of the environment. This specification can be considered the goal of the game for the system. The synthesis problem amounts to synthesizing a *strategy*, i.e., a suitable refined behavior

for the system that guarantees to the system the fulfillment of the specification. Many synthesis problems can be rephrased using 2GS. An example is *conditional planning in nondeterministic fully observable domains*, where the system is the action executor, and the environment is the domain that nondeterministically chooses the (fully observable) effect of the action among those possible. Another example is *behavior composition*, which aims at synthesizing an orchestrator that realizes a given target behavior by delegating the execution of actions to available behaviors. Here the system is the orchestrator, and the environment is formed by the target and the available behaviors. The goal of the system is to maintain over time the ability of delegating requested actions. Several other sophisticated forms of synthesis and planning can be captured through 2GS [12].

We can deal with adversarial synthesis in our framework by building a 2GS in which we encode explicitly in the DL KB the alternation of the moves of the two players. Specifically, we introduce a fresh concept name *Turn* and two fresh distinguished constants t_e and t_s , whose (mutual exclusive) presence in *Turn* indicates which of the two players will move next. We denote with \mathcal{A}_T^e the set of ABoxes in \mathcal{A}_T that contain *Turn*(e) (and not *Turn*(s)). Similarly, for \mathcal{A}_T^s .

A *DL based 2GS (DL2GS)* is a DLDS $\mathcal{K} = (T, A_0, \Gamma)$, where $\text{Turn}(e) \in A_0$ and the set of actions Γ is partitioned into a set Γ_e of *environment actions* and a set Γ_s of *system actions*. The effect function of each environment action is defined only for ABoxes in Γ_e and brings about an ABox in Γ_s . Symmetrically, the effect function of each system action is defined only for ABoxes in Γ_s and brings about an ABox in Γ_e . In this way we achieve the desired alternation of environment and system moves: starting from the initial state, the environment moves arbitrarily and the system suitably responds to the environment move, and this is repeated (possibly forever).

Logics of interest for 2GSs are temporal logics in the style of ATL [1], in which “next Φ ” has the meaning of “system can force Φ ” to hold in the next state, by suitably responding to every move done by the environment. Here, we introduce a specialization of the logic μDL_p , called μADL_p , defined as follows:

$$\begin{aligned} \Phi ::= & Q \mid \neg Q \mid \Phi_1 \wedge \Phi_2 \mid \Phi_1 \vee \Phi_2 \mid Z \mid \mu Z.\Phi \mid \nu Z.\Phi \mid \exists x.\text{LIVE}(x) \wedge \Phi \mid \forall x.\text{LIVE}(x) \rightarrow \Phi \mid \\ & \text{LIVE}(\mathbf{x}) \wedge \llbracket - \rrbracket_s \Phi \mid \text{LIVE}(\mathbf{x}) \rightarrow \llbracket - \rrbracket_s \Phi \mid \text{LIVE}(\mathbf{x}) \wedge \llbracket - \rrbracket_w \Phi \mid \text{LIVE}(\mathbf{x}) \rightarrow \llbracket - \rrbracket_w \Phi \end{aligned}$$

where $\llbracket - \rrbracket_s \Phi$ stands for $\llbracket - \rrbracket(\text{LIVE}(\mathbf{x}) \wedge \langle - \rangle \Phi)$, and $\llbracket - \rrbracket_w \Phi$ stands for $\llbracket - \rrbracket(\text{LIVE}(\mathbf{x}) \rightarrow \langle - \rangle \Phi)$. Technically, μADL_p is a fragment of μDL_p , where negation normal form is enforced, i.e., negation is pushed inwards so that it appears only in front of atoms (in our case, queries Q). Intuitively, both $\llbracket - \rrbracket_s \Phi$ and $\llbracket - \rrbracket_w \Phi$ mean that the system can force Φ to hold, i.e., for every (environment) move there is a (system) move leading to Φ . The difference between the two is that $\llbracket - \rrbracket_s \Phi$ precludes the environment from dropping objects existing before its move, while $\llbracket - \rrbracket_w \Phi$ does not.

As an immediate consequence of Theorem 1, we get that checking μADL_p formulas over state-bounded, generic DL2GS’s is decidable. However, we are not only interested in verifying μADL_p formulas, but also in *synthesizing* strategies to actually fulfill them. A *strategy* is a partial function $f : (\mathcal{A}_T^e \times \mathcal{A}_T^s)^* \times \mathcal{A}_T^e \mapsto \Gamma_s \times \mathcal{M}$ where \mathcal{M} is the domain of action parameter assignments, such that for every *history* $\lambda = A_0^e, A_0^s \dots A_n^e, A_n^s, A_{n+1}^e$, it is the case that $A_i^s = \tau_{f(\lambda[i])}(A_i^e, m_{f(\lambda[i])})$ where $\lambda[i]$ is prefix of λ of length i , and $m_{f(\lambda[i])} : \pi_{f(\lambda[i])}(A_i^e) \rightarrow \Delta$, and symmetrically $A_{i+1}^e = \tau(A_i^s, m)$ for some

$m : \pi(A_i^s) \rightarrow \Delta$. We say that a strategy f is *winning* if by resolving the existential choice in evaluating the formulas of the form $\llbracket - \rrbracket_s \Phi$ and $\llbracket - \rrbracket_w \Phi$ according to f , the goal formula is satisfied. Notably, model checking algorithms provide a *witness* of the checked property [17], which, in our case, consists of a *labeling* produced during the model checking process of the abstract DL2GS's. From *labelled* game states, one can read how the controller is meant to react to the environment in order to fulfill the formulas that *label* the state itself, and from this, define a strategy to fulfill the goal formula. It remains to lift such abstract strategy on the finite abstract DL2GS to the actual DL2GS. The abstract strategy \bar{f} models a family of concrete strategies. Thus, in principle, in order to obtain an actual strategy, it would be sufficient *concretizing* \bar{f} by replacing the abstract individuals and parameter assignments with concrete individuals and assignments that satisfy, step-by-step, the same equality commitments. While theoretically correct, this procedure cannot be realized in practice, as the resulting family of strategies is in general infinite. However, we adopt a lazy approach that allows us to generate and follow a concrete strategy, as the game progresses. Formally we have:

Theorem 2. *There exists an algorithm that, given a state-bounded, generic DL2G \mathcal{K} and a μADL_p formula Φ , realizes a concrete strategy to force Φ .*

The algorithm iterates over three steps: (i) matching of the current concrete history λ with an abstract history $\bar{\lambda}$ over which \bar{f} is defined; (ii) extraction of the action and abstract parameter assignment; (iii) concretization of the obtained parameter assignment.

6 Instantiations of the Framework

As a concrete instantiation of the abstract framework introduced in Section 2, we consider a variant of Knowledge and Action Bases (KABs) [5], instantiated on the lightweight DL *DL-Lite* and the expressive DL *ALCQL*. A *KAB* is a tuple $\mathcal{K} = (T, A_0, Act)$ where T and A_0 form the *knowledge base*, and Act is the *action base*. In practice, \mathcal{K} is a stateful device that stores the information of interest into a KB, formed by a fixed TBox T and an initial ABox A_0 , which evolves by executing actions in Act .

An *action* $\alpha \in Act$ modifies the current ABox A by adding or deleting assertions, thus generating a new ABox A' . α consists of a set $\{e_1, \dots, e_n\}$ of effects, that take place simultaneously. An *effect* e_i has the form $Q_i \rightsquigarrow A'_i$, where

- Q_i is an ECQ, i.e., a domain independent first-order query whose atoms represent certain answers of unions of conjunctive queries (UCQs) [8].
- A'_i is a set of ABox assertions that include as terms: individuals in A_0 , free variables of Q_i , and Skolem terms $f(\mathbf{x})$ having as arguments such free variables.

Given $\alpha \in Act$, we can capture its execution by defining an action $(\pi_\alpha, \tau_\alpha)$ as follows:

- $\pi_\alpha(A)$ returns the ground Skolem terms obtained by executing the queries Q_i of the effects over (T, A) , instantiating the facts A'_i using the returned answers, and extracting the (ground) Skolem terms occurring therein.
- $\tau_\alpha(A, m)$ returns the ABox obtained by executing over (T, A) the queries Q_i of the effects, instantiating the facts A'_i using the returned answers, and assigning to the (ground) Skolem terms occurring therein values according to a freely chosen parameter assignment m , as long as such an ABox is consistent with T . If the ABox is not consistent with T , then $\tau_\alpha(A, m)$ is undefined.

In this way, we can define Γ from Act , and hence the DLDS $\mathcal{S} = (T, A_0, \Gamma)$ corresponding to the KAB $\mathcal{K} = (T, A_0, Act)$. Notice that, being the parameter assignment freely chosen, the resulting DLDS is actually generic.

In this context, we use ECQs also as the query language for the local properties in the verification and synthesis formalism. Now observe that if the TBox is expressed in a lightweight DL of the *DL-Lite* family, answering ECQ queries is PSPACE-complete in combined complexity (and in AC^0 in data complexity, i.e., the complexity measured in the size of the ABox only). The same complexity bounds hold for the construction of $\pi_\alpha(A)$ and $\tau_\alpha(A, m)$, and hence, under the assumption of state-boundedness, the abstract transition system generated by \mathcal{S} can be constructed in EXPTIME. It follows that both verification and synthesis can be done in EXPTIME. Instead, if the TBox is expressed in an expressive DL such as *ALCQI*, the cost that dominates the construction of the abstract transition system is the 2EXPTIME cost of answering over (T, A) the UCQs that are the atoms of the EQL queries in the actions. If instead of UCQs we use atomic concepts and roles (i.e., we do instance checking), the cost of query evaluation drops to EXPTIME, and so does the cost of building the abstract transition system.

These results can be extended to other DLs as long as they do not include nominals, given that the presence of nominals blurs the distinction between the extensional and the intensional knowledge, and hence requires more careful handling.

7 Conclusion

This work complements and generalizes two previous papers focussing on forms of verification on DL-based dynamics. One is [5] from which we took the formalism for the instantiations. The crucial difference is that in their framework they use Skolem terms to denote new values, which as a consequence remain unknown during the construction of the transition system, while we substitute these Skolem terms with actual values. Decidability of *weakly acyclic systems* is shown. The other one is [10], where a sort of light-weight DL-based dynamic was proposed. There, a semantic layer in *DL-Lite* is built on top of a data-aware process. The *DL-Lite* ontology plus mapping is our knowledge component, while the dynamic component (the actions) are induced by the process working directly on the data-layer. Exploiting *DL-Lite* first-order rewritability properties of conjunctive queries, the verification can be done directly on the data-aware process. Decidability of checking properties in μ -calculus *without quantification across* is shown for state-bounded data-aware process. In both, synthesis is not considered.

References

1. Alur, R., Henzinger, T.A., Kupferman, O.: Alternating-time temporal logic. J. of the ACM 49(5), 672–713 (2002)
2. Artale, A., Franconi, E.: Temporal description logics. In: Gabbay, D., Fisher, M., Vila, L. (eds.) Handbook of Temporal Reasoning in Artificial Intelligence. Foundations of Artificial Intelligence, Elsevier (2005)
3. Baader, F., Ghilardi, S., Lutz, C.: LTL over description logic axioms. ACM Trans. on Computational Logic 13(3), 21:1–21:32 (2012)

4. Bagheri Hariri, B., Calvanese, D., De Giacomo, G., Deutsch, A., Montali, M.: Verification of relational data-centric dynamic systems with external services. In: Proc. of the 32nd ACM SIGACT SIGMOD SIGART Symp. on Principles of Database Systems (PODS 2013) (2013)
5. Bagheri Hariri, B., Calvanese, D., Montali, M., De Giacomo, G., De Masellis, R., Felli, P.: Description logic Knowledge and Action Bases. *J. of Artificial Intelligence Research* 46, 651–686 (2013)
6. Belardinelli, F., Lomuscio, A., Patrizi, F.: An abstraction technique for the verification of artifact-centric systems. In: Proc. of the 13th Int. Conf. on the Principles of Knowledge Representation and Reasoning (KR 2012). pp. 319–328 (2012)
7. Bhattacharya, K., Gerede, C., Hull, R., Liu, R., Su, J.: Towards formal analysis of artifact-centric business process models. In: Proc. of the 5th Int. Conference on Business Process Management (BPM 2007). Lecture Notes in Computer Science, vol. 4714, pp. 288–234. Springer (2007)
8. Calvanese, D., De Giacomo, G., Lembo, D., Lenzerini, M., Rosati, R.: EQL-Lite: Effective first-order query processing in description logics. In: Proc. of the 20th Int. Joint Conf. on Artificial Intelligence (IJCAI 2007). pp. 274–279 (2007)
9. Calvanese, D., De Giacomo, G., Lembo, D., Lenzerini, M., Rosati, R.: Tractable reasoning and efficient query answering in description logics: The *DL-Lite* family. *J. of Automated Reasoning* 39(3), 385–429 (2007)
10. Calvanese, D., De Giacomo, G., Lembo, D., Montali, M., Santoso, A.: Ontology-based governance of data-aware processes. In: Proc. of the 6th Int. Conf. on Web Reasoning and Rule Systems (RR 2012). Lecture Notes in Computer Science, vol. 7497, pp. 25–41. Springer (2012)
11. Calvanese, D., De Giacomo, G., Montali, M., Patrizi, F.: Verification and synthesis in description logic based dynamic systems. In: Proc. of the 7th Int. Conf. on Web Reasoning and Rule Systems (RR 2013). Lecture Notes in Computer Science, vol. 7994, pp. 50–64. Springer (2013)
12. De Giacomo, G., Felli, P., Patrizi, F., Sardiña, S.: Two-player game structures for generalized planning and agent composition. In: Proc. of the 24th AAAI Conf. on Artificial Intelligence (AAAI 2010). pp. 297–302 (2010)
13. Deutsch, A., Hull, R., Patrizi, F., Vianu, V.: Automatic verification of data-centric business processes. In: Proc. of the 12th Int. Conf. on Database Theory (ICDT 2009). pp. 252–267 (2009)
14. Gabbay, D., Kurusz, A., Wolter, F., Zakharyashev, M.: *Many-dimensional Modal Logics: Theory and Applications*. Elsevier Science Publishers (2003)
15. Levesque, H.J.: Foundations of a functional approach to knowledge representation. *Artificial Intelligence* 23, 155–212 (1984)
16. Martin, D.L., Burstein, M.H., McDermott, D.V., McIlraith, S.A., Paolucci, M., Sycara, K.P., McGuinness, D.L., Sirin, E., Srinivasan, N.: Bringing semantics to web services with OWL-S. In: Proc. of the 16th Int. World Wide Web Conf. (WWW 2007). pp. 243–277 (2007)
17. Piterman, N., Pnueli, A., Sa’ar, Y.: Synthesis of reactive(1) designs. In: Proc. of the 7th Int. Conf. on Verification, Model Checking, and Abstract Interpretation (VMCAI 2006). Lecture Notes in Computer Science, vol. 3855, pp. 364–380. Springer (2006)
18. Reiter, R.: *Knowledge in Action: Logical Foundations for Specifying and Implementing Dynamical Systems*. The MIT Press (2001)
19. Stirling, C.: *Modal and Temporal Properties of Processes*. Springer (2001)
20. Wolter, F., Zakharyashev, M.: Temporalizing description logic. In: Gabbay, D., de Rijke, M. (eds.) *Frontiers of Combining Systems*, pp. 379–402. Studies Press/Wiley (1999)