On Bisimulations for Description Logics

Ali Rezaei Divroodi and Linh Anh Nguyen

Institute of Informatics, University of Warsaw
Banacha 2, 02-097 Warsaw, Poland
{rezaei,nguyen}@mimuw.edu.pl

Abstract. We formulate bisimulations for useful description logics. The simplest among the considered logics is a variant of PDL (propositional dynamic logic). The others extend that logic with inverse roles, nominals, quantified number restrictions, the universal role, and/or the concept constructor $\exists r.\mathtt{Self}$. They also allow role axioms. We give results about invariance of concepts, TBoxes and ABoxes, preservation of RBoxes and knowledge bases, and the Hennessy-Milner property w.r.t. bisimulations in the considered description logics. We also provide results on the largest auto-bisimulations and quotient interpretations w.r.t. such equivalence relations.

1 Introduction

Bisimulations arose in modal logic [15–17] and state transition systems [14, 7]. They were introduced by van Benthem under the name *p-relation* in [15, 16] and the name *zigzag relation* in [17]. Bisimulations reflect, in a particularly simple and direct way, the locality of the modal satisfaction definition. The famous Van Benthem Characterization Theorem states that modal logic is the bisimulation invariant fragment of first-order logic. Bisimulations have been used to analyze the expressivity of a wide range of extended modal logics (see, e.g., [2] for details). In state transition systems, bisimulation is viewed as a binary relation associating systems which behave in the same way in the sense that one system simulates the other and vice versa. Kripke models in modal logic are a special case of labeled state transition systems. Hennessy and Milner [7] showed that weak modal languages could be used to classify various notions of process invariance. In general, bisimulations are a very natural notion of equivalence for both mathematical and computational investigations.[1]

Bisimilarity between two states is usually defined by three conditions (the states have the same label, each transition from one of the states can be simulated by a similar transition from the other, and vice versa). As shown in [2], the four program constructors of PDL (propositional dynamic logic) are "safe" for these three conditions. That is, we need to specify the mentioned conditions only for atomic programs, and as a consequence, they hold also for complex programs. For bisimulation between two pointed-models, the initial states of the models are

[1] This paragraph is based on [2].

CONCURRENCY, SPECIFICATION AND PROGRAMMING
M. Szczuka et al. (eds.): Proceedings of the international workshop CS&P 2011
September 28-30, Pułtusk, Poland, pp. 99-110

also required to be bisimilar. When converse is allowed (the case of CPDL), two additional conditions are required for bisimulation [2]. Bisimulation conditions for dealing with graded modalities were studied in [4, 3, 9]. In the field of hybrid logic, the bisimulation condition for dealing with nominals is well known (see, e.g., [1]).

Description logics (DLs) are variants of modal logic. They represent the domain of interest in terms of concepts, individuals, and roles. A concept is interpreted as a set of individuals, while a role is interpreted as a binary relation among individuals. A description logic is characterized by a set of concept constructors, a set of role constructors, and a set of allowed forms of role axioms and individual assertions. A knowledge base in a DL usually has three parts: an RBox consisting of axioms about roles, a TBox consisting of terminology axioms, and an ABox consisting of assertions about individuals. Description logics are used, amongst others, as the logical base of the Web Ontology Language OWL.

In this paper we study bisimulations for the family of DLs which extend $\mathcal{ALC}_{reg}$ (the variant of PDL) with an arbitrary combination of inverse roles, quantified number restrictions, nominals, the universal role, and the concept constructor $\exists r.\mathtt{Self}$. Inverse roles are like converse modal operators, quantified number restrictions are like graded modalities, and nominals are as in hybrid logic. We present conditions for bisimulation in a uniform way for the whole considered family of DLs and prove the standard invariance property (Theorem 3.4) and the Hennessy-Milner property (Theorem 4.1). Our condition for quantified number restrictions is simpler than the ones given for graded modalities in [4, 3]. It is weaker than the one given for counting modalities in [9], but is strong enough to guarantee the Hennessy-Milner property. We are not aware of previous work on bisimulation for logics with a constructor like the universal role or the concept constructor $\exists r.\mathtt{Self}$. Another special point of our approach is that (named) individuals are treated as initial states, which requires an appropriate condition for bisimulation.

We also address the following problems:

- When is a TBox invariant for bisimulation? (Corollary 3.5 and Theorem 3.6)
- When is an ABox invariant for bisimulation? (Theorem 3.7)
- What can be said about preservation of RBoxes w.r.t. bisimulation? (Theorem 3.9)
- What can be said about invariance or preservation of knowledge bases w.r.t. bisimulation? (Theorems 3.8 and 3.10)

Furthermore, we give some results (Theorems 5.3, 5.4 and 5.5) on the largest auto-bisimulation of an interpretation in a DL, the quotient interpretation w.r.t. that equivalence relation, and minimality of such a quotient interpretation.

1.1 Related Work

In [10] Kurtonina and de Rijke studied expressiveness of concept expressions in some description logics by using bisimulations. They considered a family of DLs that are sublogics of the DL $\mathcal{ALCNR}$, which extend $\mathcal{ALC}$ with (unquantified) number restrictions and role conjunction. They did not consider individuals, nominals, quantified number restrictions, the concept constructor $\exists r.\mathtt{Self}$, the universal role, and the role constructors like the program constructors of PDL. They did not study invariance or preservation of TBoxes, ABoxes and RBoxes, and did not study minimality of quotient interpretations.

In [11] Lutz et al. characterized the expressive power of TBoxes in the DL $\mathcal{ALCQIO}$ and its sublogics, including the lightweight DLs such as DL-Lite and $\mathcal{EL}$. They used the notion of ω-saturatedness, which is more general than the notion of finite-image (used for the Hennessy-Milner property). Their bisimulation condition for dealing with quantified number restrictions is thus different from ours. They also studied invariance of TBoxes and the problem of TBox rewritability. Note, however, that the work [11] cannot be used to judge novelty or originality of our work (in particular, w.r.t. invariance of TBoxes), as the first version [6] of the current paper appeared to the public a few days earlier than [11]. Also note that $\mathcal{ALCQIO}$ lacks the role constructors of PDL, the concept constructor $\exists r.\mathtt{Self}$ and the universal role, and the work [11] does not deal with invariance or preservation of ABoxes and RBoxes as well as minimality of quotient interpretations.

1.2 The Structure of This Work

In Section 2 we present notation and semantics of the DLs considered in this paper. In Section 3 we define bisimulations in those DLs and give our results on invariance and preservation w.r.t. such bisimulations. In Section 4 we give our results on the Hennessy-Milner property of the considered DLs. Section 5 is devoted to auto-bisimulation and minimization. Section 6 concludes this work. Due to the lack of space, proofs of our results are presented only in the long version [5] of the current paper.

2 Notation and Semantics of Description Logics

Our languages use a finite set Σ_C of *concept names* (atomic concepts), a finite set Σ_R of *role names* (atomic roles), and a finite set Σ_I of *individual names*. We denote concept names by letters like A and B, denote role names by letters like r and s, and denote individual names by letters like a and b.

We consider some (additional) *DL-features* denoted by I (*inverse*), O (*nominal*), Q (*quantified number restriction*), U (*universal role*), $\mathtt{Self}$. A *set of DL-features* is a set consisting of some of these names. We sometimes abbreviate sets of DL-features, writing e.g., OIQ instead of $\{O, I, Q\}$.

Let Φ be any set of DL-features and let $\mathcal{L}$ stand for $\mathcal{ALC}_{reg}$, which is the name of a description logic corresponding to propositional dynamic logic (PDL). The DL language $\mathcal{L}_\Phi$ allows *roles* and *concepts* defined inductively as follows:

- if $r \in \Sigma_R$ then r is a role of $\mathcal{L}_\Phi$

- if $A \in \Sigma_C$ then A is a concept of $\mathcal{L}_\Phi$
- if R and S are roles of $\mathcal{L}_\Phi$ and C is a concept of $\mathcal{L}_\Phi$ then
  - ε, $R \circ S$, $R \sqcup S$, R^* and $C?$ are roles of $\mathcal{L}_\Phi$
  - $\top$, $\bot$, $\neg C$, $C \sqcap D$, $C \sqcup D$, $\forall R.C$ and $\exists R.C$ are concepts of $\mathcal{L}_\Phi$
  - if $I \in \Phi$ then R^- is a role of $\mathcal{L}_\Phi$
  - if $O \in \Phi$ and $a \in \Sigma_I$ then $\{a\}$ is a concept of $\mathcal{L}_\Phi$
  - if $Q \in \Phi$, $r \in \Sigma_R$ and n is a natural number then $\geq n\, r.C$ and $\leq n\, r.C$ are concepts of $\mathcal{L}_\Phi$
  - if $\{Q, I\} \subseteq \Phi$, $r \in \Sigma_R$ and n is a natural number then $\geq n\, r^-.C$ and $\leq n\, r^-.C$ are concepts of $\mathcal{L}_\Phi$
  - if $U \in \Phi$ then $U \in \Sigma_R$
  - if $\mathtt{Self} \in \Phi$ and $r \in \Sigma_R$ then $\exists r.\mathtt{Self}$ is a concept of $\mathcal{L}_\Phi$.

We use letters like R and S to denote arbitrary roles, and use letters like C and D to denote arbitrary concepts. A role stands for a binary relation, while a concept stands for a unary relation.

The intended meaning of the role constructors is the following:

- $R \circ S$ stands for the sequential composition of R and S
- $R \sqcup S$ stands for the set-theoretical union of R and S
- R^* stands for the reflexive and transitive closure of R
- $C?$ stands for the test operator (as of PDL)
- R^- stands for the *inverse* of R.

The concept constructors $\forall R.C$ and $\exists R.C$ correspond respectively to the modal operators $[R]C$ and $\langle R\rangle C$ of PDL. The concept constructors $\geq n\, R.C$ and $\leq n\, R.C$ are called *quantified number restrictions*. They correspond to graded modal operators.

An *interpretation* $\mathcal{I} = \langle \Delta^\mathcal{I}, \cdot^\mathcal{I}\rangle$ consists of a non-empty set $\Delta^\mathcal{I}$, called the *domain* of $\mathcal{I}$, and a function $\cdot^\mathcal{I}$, called the *interpretation function* of $\mathcal{I}$, which maps every concept name A to a subset $A^\mathcal{I}$ of $\Delta^\mathcal{I}$, maps every role name r to a binary relation $r^\mathcal{I}$ on $\Delta^\mathcal{I}$, and maps every individual name a to an element $a^\mathcal{I}$ of $\Delta^\mathcal{I}$. The interpretation function $\cdot^\mathcal{I}$ is extended to complex roles and complex concepts as shown in Figure 1, where $\#\Gamma$ stands for the cardinality of the set Γ. We write $C^\mathcal{I}(x)$ to denote $x \in C^\mathcal{I}$, and write $R^\mathcal{I}(x, y)$ to denote $\langle x, y\rangle \in R^\mathcal{I}$.

We say that a role R is in the *converse normal form* (CNF) if the inverse constructor is applied in R only to role names different from U. Since every role can be translated to an equivalent role in CNF,[2] in this paper we assume that roles are presented in the CNF.

We refer to elements of Σ_R also as *atomic roles*. Let $\Sigma_R^\pm = \Sigma_R \cup \{r^- \mid r \in \Sigma_R\}$. From now on, by *basic roles* we refer to elements of $\Sigma_R^\pm$ if the considered language allows inverse roles, and refer to elements of Σ_R otherwise. In general, the language decides whether inverse roles are allowed in the considered context.

[2] For example, $((r \sqcup s^-) \circ r^*)^- = (r^-)^* \circ (r^- \sqcup s)$.

$$
\begin{array}{rlrl}
(R \circ S)^{\mathcal{I}} &= R^{\mathcal{I}} \circ S^{\mathcal{I}} & \top^{\mathcal{I}} &= \Delta^{\mathcal{I}} \\
(R \sqcup S)^{\mathcal{I}} &= R^{\mathcal{I}} \cup S^{\mathcal{I}} & \bot^{\mathcal{I}} &= \emptyset \\
(R^*)^{\mathcal{I}} &= (R^{\mathcal{I}})^* & (\neg C)^{\mathcal{I}} &= \Delta^{\mathcal{I}} \setminus C^{\mathcal{I}} \\
(C?)^{\mathcal{I}} &= \{\langle x, x\rangle \mid C^{\mathcal{I}}(x)\} & (C \sqcap D)^{\mathcal{I}} &= C^{\mathcal{I}} \cap D^{\mathcal{I}} \\
\varepsilon^{\mathcal{I}} &= \{\langle x, x\rangle \mid x \in \Delta^{\mathcal{I}}\} & (C \sqcup D)^{\mathcal{I}} &= C^{\mathcal{I}} \cup D^{\mathcal{I}} \\
U^{\mathcal{I}} &= \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}} & \{a\}^{\mathcal{I}} &= \{a^{\mathcal{I}}\} \\
(R^-)^{\mathcal{I}} &= (R^{\mathcal{I}})^{-1} & (\exists r.\mathtt{Self})^{\mathcal{I}} &= \{x \in \Delta^{\mathcal{I}} \mid r^{\mathcal{I}}(x, x)\}
\end{array}
$$

$$
\begin{aligned}
(\forall R.C)^{\mathcal{I}} &= \{x \in \Delta^{\mathcal{I}} \mid \forall y\,[R^{\mathcal{I}}(x, y) \text{ implies } C^{\mathcal{I}}(y)]\} \\
(\exists R.C)^{\mathcal{I}} &= \{x \in \Delta^{\mathcal{I}} \mid \exists y\,[R^{\mathcal{I}}(x, y) \text{ and } C^{\mathcal{I}}(y)]\} \\
(\geq n\, R.C)^{\mathcal{I}} &= \{x \in \Delta^{\mathcal{I}} \mid \#\{y \mid R^{\mathcal{I}}(x, y) \text{ and } C^{\mathcal{I}}(y)\} \geq n\} \\
(\leq n\, R.C)^{\mathcal{I}} &= \{x \in \Delta^{\mathcal{I}} \mid \#\{y \mid R^{\mathcal{I}}(x, y) \text{ and } C^{\mathcal{I}}(y)\} \leq n\}
\end{aligned}
$$

Fig. 1. Interpretation of complex roles and complex concepts.

A *role (inclusion) axiom* in $\mathcal{L}_\Phi$ is an expression of the form $\varepsilon \sqsubseteq r$ or $R_1 \circ \ldots \circ R_k \sqsubseteq r$, where $k \geq 1$ and $R_1, \ldots, R_k$ are basic roles of $\mathcal{L}_\Phi$.[3] Given an interpretation $\mathcal{I}$, define that:

$$
\begin{aligned}
\mathcal{I} \models \varepsilon \sqsubseteq r &\text{ if } \varepsilon^{\mathcal{I}} \subseteq r^{\mathcal{I}} \\
\mathcal{I} \models R_1 \circ \ldots \circ R_k \sqsubseteq r &\text{ if } R_1^{\mathcal{I}} \circ \ldots \circ R_k^{\mathcal{I}} \subseteq r^{\mathcal{I}}
\end{aligned}
$$

We say that a role axiom φ is *valid* in $\mathcal{I}$ and $\mathcal{I}$ *validates* φ if $\mathcal{I} \models \varphi$. Note that reflexiveness and transitiveness of atomic roles are expressible by role axioms. When $I \in \Phi$ symmetry of an atomic role can also be expressed by a role axiom.

An *RBox* in $\mathcal{L}_\Phi$ is a finite set of role axioms in $\mathcal{L}_\Phi$. An interpretation $\mathcal{I}$ is a *model* of an RBox $\mathcal{R}$, denoted by $\mathcal{I} \models \mathcal{R}$, if it validates all the role axioms of $\mathcal{R}$.

A *terminological axiom* in $\mathcal{L}_\Phi$, also called a *general concept inclusion* (GCI) in $\mathcal{L}_\Phi$, is an expression of the form $C \sqsubseteq D$, where C and D are concepts in $\mathcal{L}_\Phi$. An interpretation $\mathcal{I}$ *validates* an axiom $C \sqsubseteq D$, denoted by $\mathcal{I} \models C \sqsubseteq D$, if $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$.

A *TBox* in $\mathcal{L}_\Phi$ is a finite set of terminological axioms in $\mathcal{L}_\Phi$. An interpretation $\mathcal{I}$ is a *model* of a TBox $\mathcal{T}$, denoted by $\mathcal{I} \models \mathcal{T}$, if it validates all the axioms of $\mathcal{T}$.

An *individual assertion* in $\mathcal{L}_\Phi$ is an expression of one of the forms $C(a)$ (*concept assertion*), $R(a, b)$ (*positive role assertion*), $\neg R(a, b)$ (*negative role assertion*), $a = b$, and $a \neq b$, where C is a concept and R is a role in $\mathcal{L}_\Phi$.

Given an interpretation $\mathcal{I}$, define that:

[3] This definition depends only on whether $\mathcal{L}_\Phi$ allows inverse roles, i.e., whether $I \in \Phi$.

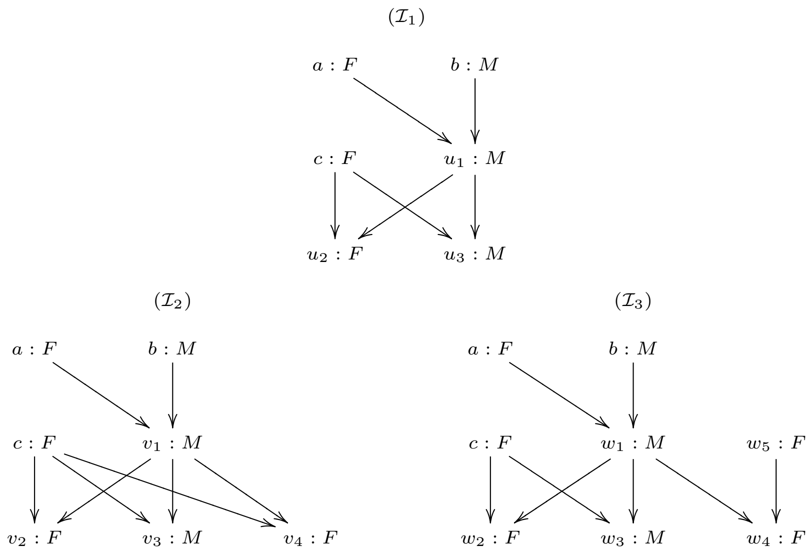

Fig. 2. Exemplary interpretations for Examples 2.1 and 3.2.

$$\begin{array}{ll} \mathcal{I} \models a = b & \text{if } a^{\mathcal{I}} = b^{\mathcal{I}} \\ \mathcal{I} \models a \neq b & \text{if } a^{\mathcal{I}} \neq b^{\mathcal{I}} \\ \mathcal{I} \models C(a) & \text{if } C^{\mathcal{I}}(a^{\mathcal{I}}) \text{ holds} \\ \mathcal{I} \models R(a,b) & \text{if } R^{\mathcal{I}}(a^{\mathcal{I}}, b^{\mathcal{I}}) \text{ holds} \\ \mathcal{I} \models \neg R(a,b) & \text{if } R^{\mathcal{I}}(a^{\mathcal{I}}, b^{\mathcal{I}}) \text{ does not hold.} \end{array}$$

We say that $\mathcal{I}$ *satisfies* an individual assertion φ if $\mathcal{I} \models \varphi$.

An *ABox* in $\mathcal{L}_\Phi$ is a finite set of individual assertions in $\mathcal{L}_\Phi$. An interpretation $\mathcal{I}$ is a *model* of an ABox $\mathcal{A}$, denoted by $\mathcal{I} \models \mathcal{A}$, if it satisfies all the assertions of $\mathcal{A}$.

A *knowledge base* in $\mathcal{L}_\Phi$ is a triple $\langle \mathcal{R}, \mathcal{T}, \mathcal{A} \rangle$, where $\mathcal{R}$ (resp. $\mathcal{T}$, $\mathcal{A}$) is an RBox (resp. a TBox, an ABox) in $\mathcal{L}_\Phi$. An interpretation $\mathcal{I}$ is a *model* of a knowledge base $\langle \mathcal{R}, \mathcal{T}, \mathcal{A} \rangle$ if it is a model of all $\mathcal{R}$, $\mathcal{T}$, and $\mathcal{A}$.

Example 2.1. Let $\Sigma_I = \{a, b, c\}$, $\Sigma_C = \{F, M\}$ and $\Sigma_R = \{r\}$. One can think of these names as *Alice* (a), *Bob* (b), *Claudia* (c), *female* (F), *male* (M), and *has_child* (r). In Figure 2 we give three interpretations $\mathcal{I}_1$, $\mathcal{I}_2$ and $\mathcal{I}_3$. The edges are instances of r. We have, for example, $\Delta^{\mathcal{I}_1} = \{a^{\mathcal{I}_1}, b^{\mathcal{I}_1}, c^{\mathcal{I}_1}, u_1, u_2, u_3\}$, where these six elements are pairwise different, $F^{\mathcal{I}_1} = \{a^{\mathcal{I}_1}, c^{\mathcal{I}_1}, u_2\}$, and $M^{\mathcal{I}_1} = \{b^{\mathcal{I}_1}, u_1, u_3\}$.[4] All of these interpretations are models of the following ABox in $\mathcal{L}_{OIQ}$, where r^- can be read as *has_parent*:

$$\{\ F(a),\ M(b),\ F(c),\ \big(\exists r.(\exists r^-.\{b\} \sqcap \geq 2\, r.\exists r^-.\{c\})\big)(a)\ \}$$

[4] The elements u_i, v_j, w_k are unnamed objects. (The elements of Σ_I can be called *named individuals*, while the elements u_i, v_j, w_k can be called *unnamed individuals*.)

Assuming that r means *has_child*, then the last assertion of the above ABox means "a and b have a child which in turn has at least two children with c".

All the interpretations $\mathcal{I}_1, \mathcal{I}_2$ and $\mathcal{I}_3$ validate the terminological axioms $\neg F \sqsubseteq M$ and $\{a\} \sqsubseteq \forall r^*.(\{a\} \sqcup \geq 2\, r^-.\top)$ of $\mathcal{L}_{OIQ}$. ◁

3 Bisimulations and Invariance Results

Let $\mathcal{I}$ and $\mathcal{I}'$ be interpretations. A binary relation $Z \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}'}$ is called an $\mathcal{L}_\Phi$-*bisimulation* between $\mathcal{I}$ and $\mathcal{I}'$ if the following conditions hold for every $a \in \Sigma_I$, $A \in \Sigma_C$, $r \in \Sigma_R$, $x, y \in \Delta^{\mathcal{I}}$, $x', y' \in \Delta^{\mathcal{I}'}$:

$$Z(a^{\mathcal{I}}, a^{\mathcal{I}'}) \tag{1}$$

$$Z(x, x') \Rightarrow [A^{\mathcal{I}}(x) \Leftrightarrow A^{\mathcal{I}'}(x')] \tag{2}$$

$$[Z(x, x') \wedge r^{\mathcal{I}}(x, y)] \Rightarrow \exists y' \in \Delta^{\mathcal{I}'}[Z(y, y') \wedge r^{\mathcal{I}'}(x', y')] \tag{3}$$

$$[Z(x, x') \wedge r^{\mathcal{I}'}(x', y')] \Rightarrow \exists y \in \Delta^{\mathcal{I}}[Z(y, y') \wedge r^{\mathcal{I}}(x, y)], \tag{4}$$

if $I \in \Phi$ then

$$[Z(x, x') \wedge r^{\mathcal{I}}(y, x)] \Rightarrow \exists y' \in \Delta^{\mathcal{I}'}[Z(y, y') \wedge r^{\mathcal{I}'}(y', x')] \tag{5}$$

$$[Z(x, x') \wedge r^{\mathcal{I}'}(y', x')] \Rightarrow \exists y \in \Delta^{\mathcal{I}}[Z(y, y') \wedge r^{\mathcal{I}}(y, x)], \tag{6}$$

if $O \in \Phi$ then

$$Z(x, x') \Rightarrow [x = a^{\mathcal{I}} \Leftrightarrow x' = a^{\mathcal{I}'}], \tag{7}$$

if $Q \in \Phi$ then

if $Z(x, x')$ holds then, for every role name r, there exists a bijection $h : \{y \mid r^{\mathcal{I}}(x, y)\} \to \{y' \mid r^{\mathcal{I}'}(x', y')\}$ such that $h \subseteq Z$, (8)

if $\{Q, I\} \subseteq \Phi$ then (additionally)

if $Z(x, x')$ holds then, for every role name r, there exists a bijection $h : \{y \mid r^{\mathcal{I}}(y, x)\} \to \{y' \mid r^{\mathcal{I}'}(y', x')\}$ such that $h \subseteq Z$, (9)

if $U \in \Phi$ then

$$\forall x \in \Delta^{\mathcal{I}}\, \exists x' \in \Delta^{\mathcal{I}'}\, Z(x, x') \tag{10}$$

$$\forall x' \in \Delta^{\mathcal{I}'}\, \exists x \in \Delta^{\mathcal{I}}\, Z(x, x'), \tag{11}$$

if `Self` $\in \Phi$ then

$$Z(x, x') \Rightarrow [r^{\mathcal{I}}(x, x) \Leftrightarrow r^{\mathcal{I}'}(x', x')]. \tag{12}$$

For example, if $\Phi = \{Q, I\}$ then only the conditions (1)-(6), (8), (9) (and all of them) are essential.

Lemma 3.1.

1. *The relation $\{\langle x, x\rangle \mid x \in \Delta^{\mathcal{I}}\}$ is an $\mathcal{L}_\Phi$-bisimulation between $\mathcal{I}$ and $\mathcal{I}$.*

2. *If Z is an $\mathcal{L}_\Phi$-bisimulation between $\mathcal{I}$ and $\mathcal{I}'$ then Z^{-1} is an $\mathcal{L}_\Phi$-bisimulation between $\mathcal{I}'$ and $\mathcal{I}$.*
3. *If Z_1 is an $\mathcal{L}_\Phi$-bisimulation between $\mathcal{I}_0$ and $\mathcal{I}_1$, and Z_2 is an $\mathcal{L}_\Phi$-bisimulation between $\mathcal{I}_1$ and $\mathcal{I}_2$, then $Z_1 \circ Z_2$ is an $\mathcal{L}_\Phi$-bisimulation between $\mathcal{I}_0$ and $\mathcal{I}_2$.*
4. *If $\mathcal{Z}$ is a set of $\mathcal{L}_\Phi$-bisimulations between $\mathcal{I}$ and $\mathcal{I}'$ then $\bigcup \mathcal{Z}$ is also an $\mathcal{L}_\Phi$-bisimulation between $\mathcal{I}$ and $\mathcal{I}'$.*

The proof of this lemma is straightforward.

An interpretation $\mathcal{I}$ is *$\mathcal{L}_\Phi$-bisimilar* to $\mathcal{I}'$ if there exists an $\mathcal{L}_\Phi$-bisimulation between them. By Lemma 3.1, this $\mathcal{L}_\Phi$-bisimilarity relation is an equivalence relation between interpretations. We say that $x \in \Delta^\mathcal{I}$ is *$\mathcal{L}_\Phi$-bisimilar* to $x' \in \Delta^{\mathcal{I}'}$ if there exists an $\mathcal{L}_\Phi$-bisimulation Z between $\mathcal{I}$ and $\mathcal{I}'$ such that $Z(x, x')$ holds. This latter $\mathcal{L}_\Phi$-bisimilarity relation is also an equivalence relation (between elements of interpretations' domains).

Example 3.2. Consider the interpretations $\mathcal{I}_1$, $\mathcal{I}_2$ and $\mathcal{I}_3$ given in Figure 2 and described in Example 2.1. All of them are $\mathcal{L}$-bisimilar. The elements u_2 (of $\mathcal{I}_1$) and v_2, v_4 (of $\mathcal{I}_2$) are $\mathcal{L}_\Phi$-bisimilar for $\Phi \subseteq \{I, O\}$. The elements u_1 (of $\mathcal{I}_1$) and v_1 (of $\mathcal{I}_2$) are not $\mathcal{L}_Q$-bisimilar. The interpretations $\mathcal{I}_1$ and $\mathcal{I}_2$ are $\mathcal{L}_\Phi$-bisimilar for $\Phi \subseteq \{I, O\}$, but not $\mathcal{L}_Q$-bisimilar. The interpretation $\mathcal{I}_3$ is not $\mathcal{L}_I$-bisimilar to $\mathcal{I}_1$ and $\mathcal{I}_2$, but it is $\mathcal{L}_Q$-bisimilar to $\mathcal{I}_1$. ◁

Lemma 3.3. *Let $\mathcal{I}$ and $\mathcal{I}'$ be interpretations and Z be an $\mathcal{L}_\Phi$-bisimulation between $\mathcal{I}$ and $\mathcal{I}'$. Then the following properties hold for every concept C in $\mathcal{L}_\Phi$, every role R in $\mathcal{L}_\Phi$, every $x, y \in \Delta^\mathcal{I}$, every $x', y' \in \Delta^{\mathcal{I}'}$, and every $a \in \mathcal{I}$:*

$$Z(x, x') \Rightarrow [C^\mathcal{I}(x) \Leftrightarrow C^{\mathcal{I}'}(x')] \tag{13}$$

$$[Z(x, x') \wedge R^\mathcal{I}(x, y)] \Rightarrow \exists y' \in \Delta^{\mathcal{I}'}[Z(y, y') \wedge R^{\mathcal{I}'}(x', y')] \tag{14}$$

$$[Z(x, x') \wedge R^{\mathcal{I}'}(x', y')] \Rightarrow \exists y \in \Delta^\mathcal{I}[Z(y, y') \wedge R^\mathcal{I}(x, y)] \tag{15}$$

if $O \in \Phi$ then:

$$Z(x, x') \Rightarrow [R^\mathcal{I}(x, a^\mathcal{I}) \Leftrightarrow R^{\mathcal{I}'}(x', a^{\mathcal{I}'})]. \tag{16}$$

A concept C in $\mathcal{L}_\Phi$ is said to be *invariant for $\mathcal{L}_\Phi$-bisimulation* if, for any interpretations $\mathcal{I}$, $\mathcal{I}'$ and any $\mathcal{L}_\Phi$-bisimulation between $\mathcal{I}$ and $\mathcal{I}'$, if $Z(x, x')$ holds then $x \in C^\mathcal{I}$ iff $x' \in C^{\mathcal{I}'}$.

Theorem 3.4. *All concepts in $\mathcal{L}_\Phi$ are invariant for $\mathcal{L}_\Phi$-bisimulation.*

This theorem follows immediately from the assertion (13) of Lemma 3.3.

A TBox $\mathcal{T}$ in $\mathcal{L}_\Phi$ is said to be *invariant for $\mathcal{L}_\Phi$-bisimulation* if, for every interpretations $\mathcal{I}$ and $\mathcal{I}'$, if there exists an $\mathcal{L}_\Phi$-bisimulation between $\mathcal{I}$ and $\mathcal{I}'$ then $\mathcal{I}$ is a model of $\mathcal{T}$ iff $\mathcal{I}'$ is a model of $\mathcal{T}$. The notions of whether an ABox or a knowledge base in $\mathcal{L}_\Phi$ is invariant for $\mathcal{L}_\Phi$-bisimulation are defined similarly.

Corollary 3.5. *If $U \in \Phi$ then all TBoxes in $\mathcal{L}_\Phi$ are invariant for $\mathcal{L}_\Phi$-bisimulation.*

An interpretation $\mathcal{I}$ is said to be *unreachable-objects-free* (w.r.t. the considered language) if every element of $\Delta^{\mathcal{I}}$ is reachable from some $a^{\mathcal{I}}$, where $a \in \Sigma_I$, via a path consisting of edges being instances of basic roles.

Like Corollary 3.5, the following theorem concerns invariance of TBoxes w.r.t. $\mathcal{L}_\Phi$-bisimulation.

Theorem 3.6. *Let $\mathcal{T}$ be a TBox in $\mathcal{L}_\Phi$ and $\mathcal{I}$, $\mathcal{I}'$ be unreachable-objects-free interpretations (w.r.t. $\mathcal{L}_\Phi$) such that there exists an $\mathcal{L}_\Phi$-bisimulation between $\mathcal{I}$ and $\mathcal{I}'$. Then $\mathcal{I}$ is a model of $\mathcal{T}$ iff $\mathcal{I}'$ is a model of $\mathcal{T}$.*

The following theorem concerns invariance of ABoxes w.r.t. $\mathcal{L}_\Phi$-bisimulation.

Theorem 3.7. *Let $\mathcal{A}$ be an ABox in $\mathcal{L}_\Phi$. If $O \in \Phi$ or $\mathcal{A}$ contains only assertions of the form $C(a)$ then $\mathcal{A}$ is invariant for $\mathcal{L}_\Phi$-bisimulation.*

The following theorem concerns invariance of knowledge bases w.r.t. $\mathcal{L}_\Phi$-bisimulation. It follows immediately from Theorems 3.6 and 3.7.

Theorem 3.8. *Let $\langle \mathcal{R}, \mathcal{T}, \mathcal{A} \rangle$ be a knowledge base in $\mathcal{L}_\Phi$ such that $\mathcal{R} = \emptyset$ and either $O \in \Phi$ or $\mathcal{A}$ contains only assertions of the form $C(a)$. Let $\mathcal{I}$ and $\mathcal{I}'$ be unreachable-objects-free interpretations (w.r.t. $\mathcal{L}_\Phi$) such that there exists an $\mathcal{L}_\Phi$-bisimulation between $\mathcal{I}$ and $\mathcal{I}'$. Then $\mathcal{I}$ is a model of $\langle \mathcal{R}, \mathcal{T}, \mathcal{A} \rangle$ iff $\mathcal{I}'$ is a model of $\langle \mathcal{R}, \mathcal{T}, \mathcal{A} \rangle$.*

An interpretation $\mathcal{I}'$ is an r-extension of an interpretation $\mathcal{I}$ if $\Delta^{\mathcal{I}'} = \Delta^{\mathcal{I}}$, $\cdot^{\mathcal{I}'}$ differs from $\cdot^{\mathcal{I}}$ only in interpreting role names, and for all $r \in \Sigma_R$, $r^{\mathcal{I}'} \supseteq r^{\mathcal{I}}$.

Given an interpretation $\mathcal{I}$ and an RBox $\mathcal{R}$, the *least r-extension of $\mathcal{I}$ validating $\mathcal{R}$* is the r-extension $\mathcal{I}'$ of $\mathcal{I}$ such that $\mathcal{I}'$ is a model of $\mathcal{R}$ and, for every r-extension $\mathcal{I}''$ of $\mathcal{I}$, if $\mathcal{I}''$ is a model of $\mathcal{R}$ then $r^{\mathcal{I}'} \subseteq r^{\mathcal{I}''}$ for all $r \in \Sigma_R$. That r-extension exists and is unique because the axioms of $\mathcal{R}$ correspond to non-negative Horn clauses of first-order logic.

In general, RBoxes are not invariant for $\mathcal{L}_\Phi$-bisimulations.[5] The following theorem is a result not directly related to invariance.

Theorem 3.9. *Suppose $\Phi \subseteq \{I, O, U\}$ and let $\mathcal{R}$ be an RBox in $\mathcal{L}_\Phi$. Let $\mathcal{I}_0$ be a model of $\mathcal{R}$, Z be an $\mathcal{L}_\Phi$-bisimulation between $\mathcal{I}_0$ and an interpretation $\mathcal{I}_1$, and $\mathcal{I}_1'$ be the least r-extension of $\mathcal{I}_1$ validating $\mathcal{R}$. Then Z is an $\mathcal{L}_\Phi$-bisimulation between $\mathcal{I}_0$ and $\mathcal{I}_1'$.*

The following theorem concerns preservation of knowledge bases under $\mathcal{L}_\Phi$-bisimulation. It follows immediately from Theorems 3.9, 3.7, and 3.6.

Theorem 3.10. *Suppose $\Phi \subseteq \{I, O, U\}$ and let $\langle \mathcal{R}, \mathcal{T}, \mathcal{A} \rangle$ be a knowledge base in $\mathcal{L}_\Phi$ such that if $O \notin \Phi$ then $\mathcal{A}$ contains only assertions of the form $C(a)$. Let $\mathcal{I}_0$ and $\mathcal{I}_1$ be unreachable-objects-free interpretations (w.r.t. $\mathcal{L}_\Phi$) such that $\mathcal{I}_0$ is a model of $\mathcal{R}$ and there is an $\mathcal{L}_\Phi$-bisimulation Z between $\mathcal{I}_0$ and $\mathcal{I}_1$. Let $\mathcal{I}_1'$ be the least r-extension of $\mathcal{I}_1$ validating $\mathcal{R}$. Then:*

[5] The Van Benthem Characterization Theorem states that a first-order formula is invariant for bisimulations iff it is equivalent to the translation of a modal formula (see, e.g., [2]).

1. $\mathcal{I}'_1$ *is a model of* $\langle \mathcal{R}, \mathcal{T}, \mathcal{A} \rangle$ *iff* $\mathcal{I}_0$ *is a model of* $\langle \mathcal{R}, \mathcal{T}, \mathcal{A} \rangle$
2. Z *is an* $\mathcal{L}_\Phi$*-bisimulation between* $\mathcal{I}_0$ *and* $\mathcal{I}'_1$.

4 The Hennessy-Milner Property

An interpretation $\mathcal{I}$ is *finitely branching* (or *image-finite*) w.r.t. $\mathcal{L}_\Phi$ if, for every $x \in \Delta^\mathcal{I}$ and every basic role R in $\mathcal{L}_\Phi$, the set $\{y \in \Delta^\mathcal{I} \mid R^\mathcal{I}(x, y)\}$ is finite.

Let $\mathcal{I}$ and $\mathcal{I}'$ be interpretations, and let $x \in \Delta^\mathcal{I}$ and $x' \in \Delta^{\mathcal{I}'}$. We say that x is $\mathcal{L}_\Phi$*-equivalent* to x' if, for every concept C in $\mathcal{L}_\Phi$, $x \in C^\mathcal{I}$ iff $x' \in C^{\mathcal{I}'}$.

Theorem 4.1 (The Hennessy-Milner Property). *Let* $\mathcal{I}$ *and* $\mathcal{I}'$ *be finitely branching interpretations (w.r.t.* $\mathcal{L}_\Phi$*) such that, for every* $a \in \Sigma_I$, $a^\mathcal{I}$ *is* $\mathcal{L}_\Phi$*-equivalent to* $a^{\mathcal{I}'}$. *Suppose that if* $U \in \Phi$ *then* $\Sigma_I \neq \emptyset$. *Then* $x \in \Delta^\mathcal{I}$ *is* $\mathcal{L}_\Phi$*-equivalent to* $x' \in \Delta^{\mathcal{I}'}$ *iff there exists an* $\mathcal{L}_\Phi$*-bisimulation* Z *between* $\mathcal{I}$ *and* $\mathcal{I}'$ *such that* $Z(x, x')$ *holds.*

Corollary 4.2. *Let* $\mathcal{I}$ *and* $\mathcal{I}'$ *be finitely branching interpretations (w.r.t.* $\mathcal{L}_\Phi$*). Suppose that* $\Sigma_I \neq \emptyset$ *and, for every* $a \in \Sigma_I$, $a^\mathcal{I}$ *is* $\mathcal{L}_\Phi$*-equivalent to* $a^{\mathcal{I}'}$. *Then the relation* $\{\langle x, x' \rangle \in \Delta^\mathcal{I} \times \Delta^{\mathcal{I}'} \mid x$ *is* $\mathcal{L}_\Phi$*-equivalent to* $x'\}$ *is an* $\mathcal{L}_\Phi$*-bisimulation between* $\mathcal{I}$ *and* $\mathcal{I}'$.

5 Auto-Bisimulation and Minimization

An $\mathcal{L}_\Phi$-bisimulation between $\mathcal{I}$ and itself is called an $\mathcal{L}_\Phi$*-auto-bisimulation of* $\mathcal{I}$. An $\mathcal{L}_\Phi$-auto-bisimulation of $\mathcal{I}$ is said to be the *largest* if it is larger than or equal to ($\supseteq$) any other $\mathcal{L}_\Phi$-auto-bisimulation of $\mathcal{I}$.

Proposition 5.1. *For every interpretation* $\mathcal{I}$, *the largest* $\mathcal{L}_\Phi$*-auto-bisimulation of* $\mathcal{I}$ *exists and is an equivalence relation.*

This proposition follows from Lemma 3.1.

Given an interpretation $\mathcal{I}$, by $\sim_{\Phi,\mathcal{I}}$ we denote the largest $\mathcal{L}_\Phi$-auto-bisimulation of $\mathcal{I}$, and by $\equiv_{\Phi,\mathcal{I}}$ we denote the binary relation on $\Delta^\mathcal{I}$ with the property that $x \equiv_{\Phi,\mathcal{I}} x'$ iff x is $\mathcal{L}_\Phi$-equivalent to x'.

Theorem 5.2. *For every finitely branching interpretation* $\mathcal{I}$, $\equiv_{\Phi,\mathcal{I}}$ *is the largest* $\mathcal{L}_\Phi$*-auto-bisimulation of* $\mathcal{I}$ *(i.e. the relations* $\equiv_{\Phi,\mathcal{I}}$ *and* $\sim_{\Phi,\mathcal{I}}$ *coincide).*

The *quotient interpretation* $\mathcal{I}/_{\sim_{\Phi,\mathcal{I}}}$ of $\mathcal{I}$ w.r.t. $\sim_{\Phi,\mathcal{I}}$ is defined as usual:

- $\Delta^{\mathcal{I}/\sim_{\Phi,\mathcal{I}}} = \{[x]_{\sim_{\Phi,\mathcal{I}}} \mid x \in \Delta^\mathcal{I}\}$, where $[x]_{\sim_{\Phi,\mathcal{I}}}$ is the abstract class of x w.r.t. $\sim_{\Phi,\mathcal{I}}$
- $a^{\mathcal{I}/\sim_{\Phi,\mathcal{I}}} = [a^\mathcal{I}]_{\sim_{\Phi,\mathcal{I}}}$, for $a \in \Sigma_I$
- $A^{\mathcal{I}/\sim_{\Phi,\mathcal{I}}} = \{[x]_{\sim_{\Phi,\mathcal{I}}} \mid x \in A^\mathcal{I}\}$, for $A \in \Sigma_C$
- $r^{\mathcal{I}/\sim_{\Phi,\mathcal{I}}} = \{\langle [x]_{\sim_{\Phi,\mathcal{I}}}, [y]_{\sim_{\Phi,\mathcal{I}}} \rangle \mid \langle x, y \rangle \in r^\mathcal{I}\}$, for $r \in \Sigma_R$.

Theorem 5.3. *If $\Phi \subseteq \{I, O, U\}$ then, for every interpretation $\mathcal{I}$, the relation $Z = \{\langle x, [x]_{\sim_{\Phi,\mathcal{I}}}\rangle \mid x \in \Delta^{\mathcal{I}}\}$ is an $\mathcal{L}_\Phi$-bisimulation between $\mathcal{I}$ and $\mathcal{I}/_{\sim_{\Phi,\mathcal{I}}}$.*

The following theorem concerns invariance of terminological axioms and concept assertions, as well as preservation of role axioms and other individual assertion under the transformation of an interpretation to its quotient using the largest $\mathcal{L}_\Phi$-auto-bisimulation.

Theorem 5.4. *Suppose $\Phi \subseteq \{I, O, U\}$ and let $\mathcal{I}$ be an interpretation. Then:*

1. *For every expression φ which is either a terminological axiom in $\mathcal{L}_\Phi$ or a concept assertion (of the form $C(a)$) in $\mathcal{L}_\Phi$, $\mathcal{I} \models \varphi$ iff $\mathcal{I}/_{\sim_{\Phi,\mathcal{I}}} \models \varphi$.*
2. *For every expression φ which is either a role inclusion axiom or an individual assertion of the form $R(a, b)$ or $a = b$, if $\mathcal{I} \models \varphi$ then $\mathcal{I}/_{\sim_{\Phi,\mathcal{I}}} \models \varphi$.*

An interpretation $\mathcal{I}$ is said to be *minimal* among a class of interpretations if $\mathcal{I}$ belongs to that class and, for every other interpretation $\mathcal{I}'$ of that class, $\#\Delta^{\mathcal{I}} \leq \#\Delta^{\mathcal{I}'}$ (the cardinality of $\Delta^{\mathcal{I}}$ is less than or equal to the cardinality of $\Delta^{\mathcal{I}'}$). The following theorem concerns minimality of quotient interpretations generated by using the largest $\mathcal{L}_\Phi$-auto-bisimulations.

Theorem 5.5. *Suppose $\Phi \subseteq \{I, O, U\}$ and let $\mathcal{I}$ be an unreachable-objects-free interpretation. Then:*

1. *$\mathcal{I}/_{\sim_{\Phi,\mathcal{I}}}$ is a minimal interpretation $\mathcal{L}_\Phi$-bisimilar to $\mathcal{I}$.*
2. *If $\mathcal{I}/_{\sim_{\Phi,\mathcal{I}}}$ is finite then it is a minimal interpretation that validates the same terminological axioms in $\mathcal{L}_\Phi$ as $\mathcal{I}$.*
3. *If $\mathcal{I}/_{\sim_{\Phi,\mathcal{I}}}$ is finitely branching then it is a minimal interpretation that satisfies the same concept assertions in $\mathcal{L}_\Phi$ as $\mathcal{I}$.*

Computing the largest auto-bisimulations is standard like Hopcroft's automaton minimization algorithm [8] and the Paige-Tarjan algorithm [13], and is sketched in the long version [5] of the current paper.

6 Conclusions

We have formulated bisimulations in a uniform way for a large class of expressive DLs and provided results on invariance of concepts, TBoxes and ABoxes, preservation of RBoxes and knowledge bases, and the Hennessy-Milner property w.r.t. those bisimulations. We have also given some results on the largest auto-bisimulations and quotient interpretations w.r.t. such equivalence relations.

This paper is a (reasonably) systematic work on bisimulations for DLs. It differs from the related work [10, 11] on the class of considered DLs and on a considerable number of studied problems.

Our results found the logical base of [12], which is a pioneering work on applying bisimulation to concept learning and approximation. That is, our results, especially the ones on the largest auto-bisimulations, are very useful for machine learning in the context of description logics.

Acknowledgements

This work was supported by the National Centre for Research and Development (NCBiR) under Grant No. SP/I/1/77065/10 by the strategic scientific research and experimental development program: "Interdisciplinary System for Interactive Scientific and Scientific-Technical Information".

We would like to thank the anonymous reviewer for helpful comments.

References

1. C. Areces, P. Blackburn, and M. Marx. Hybrid logics: Characterization, interpolation and complexity. *J. Symb. Log.*, 66(3):977–1010, 2001.
2. P. Blackburn, M. de Rijke, and Y. Venema. *Modal Logic.* Number 53 in Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 2001.
3. W. Conradie. Definability and changing perspectives: The Beth property for three extensions of modal logic. Master's thesis, ILLC, University of Amsterdam, 2002.
4. M. de Rijke. A note on graded modal logic. *Studia Logica*, 64(2):271–283, 2000.
5. A.R. Divroodi and L.A. Nguyen. The long version of the current paper. `http://arxiv.org/abs/1104.1964v3`, 2011.
6. A.R. Divroodi and L.A. Nguyen. The first version of the current paper. `http://arxiv.org/abs/1104.1964v1`, 2011-04-11.
7. M. Hennessy and R. Milner. Algebraic laws for nondeterminism and concurrency. *Journal of the ACM*, 32(1):137–161, 1985.
8. J. Hopcroft. An $n \log n$ algorithm for minimizing states in a finite automaton. Available at `ftp://reports.stanford.edu/pub/cstr/reports/cs/tr/71/190/CS-TR-71-190.pdf`, 1971.
9. D. Janin and G. Lenzi. On the relationship between monadic and weak monadic second order logic on arbitrary trees, with applications to the mu-calculus. *Fundam. Inform.*, 61(3-4):247–265, 2004.
10. N. Kurtonina and M. de Rijke. Expressiveness of concept expressions in first-order description logics. *Artif. Intell.*, 107(2):303–333, 1999.
11. C. Lutz, R. Piro, and F. Wolter. Description logic tboxes: Model-theoretic characterizations and rewritability. *CoRR*, abs/1104.2844, 2011.
12. L.A. Nguyen and A. Szałas. Roughification: Knowledge-based indiscernibility modeling. Submitted.
13. R. Paige and R.E. Tarjan. Three partition refinement algorithms. *SIAM J. Comput.*, 16(6):973–989, 1987.
14. D.M.R. Park. Concurrency and automata on infinite sequences. In Peter Deussen, editor, *Proceedings of the 5th GI-Conference*, volume 104 of *LNCS*, pages 167–183. Springer, 1981.
15. J. van Benthem. *Modal Correspondence Theory.* PhD thesis, Mathematisch Instituut & Instituut voor Grondslagenonderzoek, University of Amsterdam, 1976.
16. J. van Benthem. *Modal Logic and Classical Logic.* Bibliopolis, Naples, 1983.
17. J. van Benthem. Correspondence theory. In D. Gabbay and F. Guenther, editors, *Handbook of Philosophical Logic, Volume II*, pages 167–247. Reidel, Dordrecht, 1984.