

# SHOIQ with transitive closure of roles is decidable

Chan Le Duc, Myriam Lamolle, Olivier Curé

► **To cite this version:**

Chan Le Duc, Myriam Lamolle, Olivier Curé. SHOIQ with transitive closure of roles is decidable. Description Logics, 2013, Ulm, Germany. hal-01740557

**HAL Id: hal-01740557**

**<https://hal-upec-upem.archives-ouvertes.fr/hal-01740557>**

Submitted on 22 Mar 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# *SHOIQ* with transitive closure of roles is decidable

Chan Le Duc<sup>1</sup>, Myriam Lamolle<sup>1</sup>, and Olivier Curé<sup>2</sup>

<sup>1</sup> LIASD Université Paris 8 - IUT de Montreuil, France  
{chan.leduc, myriam.lamolle}@iut.univ-paris8.fr  
<sup>2</sup> LIGM Université Paris-Est, France  
ocure@univ-mlv.fr

**Abstract.** The Semantic Web makes an extensive use of the OWL DL ontology language, underlied by the *SHOIQ* description logic, to formalize its resources. In this paper, we propose a decision procedure for this logic extended with the transitive closure of roles in concept axioms, a feature needed in several application domains. To address the problem of consistency in this logic, we introduce a new structure for characterizing models which may have an infinite non-tree-like part.

## 1 Introduction

The ontology language OWL-DL [1] is widely used to formalize data resources on the Semantic Web. This language is mainly based on the description logic *SHOIN* which is known to be decidable [2]. Although *SHOIN* provides *transitive roles* to model transitivity of relations, we can find several applications in which the *transitive closure of roles*, that is more expressive than transitive roles, is needed. For instance, if we denote by  $R^-$  and  $R^+$  the inverse and transitive closure of a role  $R$  respectively then it is obvious that the concept  $\exists R^+.\forall R^-. \perp$  is unsatisfiable w.r.t. an empty TBox. If we now substitute  $R^+$  for a transitive role  $R_t$  such that  $R \sqsubseteq R_t$  (i.e. we substitute each occurrence of  $R^+$  in axioms for  $R_t$ ) then the concept  $\exists R_t.\forall R^-. \perp$  is satisfiable. The point is that an instance of  $R^+$  represents a sequence of instances of  $R$  but an instance of  $R_t$  corresponds to a sequence of instances of *itself*.

In this paper, we consider an extension of *SHOIQ* by enabling transitive closure of roles in concept axioms. In the general case, transitive closure is not expressible in the first order logic [3], the logic from which DL is a sublanguage, while the second order logic is sufficiently expressive to do so.

In the DL literature ([4]; [5]), there have been works dealing with transitive closure of roles. Recently, Ortiz [5] has proposed an algorithm for deciding consistency in the logic  $ALCQIb_{reg}^+$  which allows for transitive closure of roles. However, nominals are disallowed in this logic. It is known that reasoning with a DL including number restrictions, inverse roles, nominals and transitive closure of roles is hard. The reason for this is that there exists an ontology in that DL whose models have an *infinite* non-tree-like part. Calvanese *et al.* [6] have presented an automata-based technique for dealing with the logic *ZOIQ* that includes transitive closure of roles, and showed that the sublogics *ZIQ*, *ZOQ* and *ZOI* are decidable. To obtain this result, the authors have introduced the *quasi-forest model property* to characterize models of ontologies in these sublogics.

Although they are very expressive, none of these sublogics includes  $\mathcal{SHOIQ}$  with transitive closure of roles, namely  $\mathcal{SHOIQ}_{(+)}$ . The following example<sup>3</sup>, noted  $\mathcal{K}_1$ , shows that there is an ontology in  $\mathcal{SHOIQ}_{(+)}$  which does not enjoy the quasi-forest model property. We consider the following axioms:

- (1)  $\{o\} \sqsubseteq A$ ;  $A \sqcap B \sqsubseteq \perp$ ;  $A \sqsubseteq \exists R.A \sqcap \exists R'.B$ ;  $B \sqsubseteq \exists S^+.\{o\}$
- (2)  $\{o\} \sqsubseteq \forall X^-. \perp$ ;  $\top \sqsubseteq \leq 1 X.\top$ ;  $\top \sqsubseteq \leq 1 X^-. \top$  where  $X \in \{R, R', S\}$

Figure 1 shows an infinite non-tree-like model of  $\mathcal{K}_1$ . In fact, each individual  $x$  that satisfies  $\exists S^+.\{o\}$  must have two distinct paths from  $x$  to the individual satisfying nominal  $o$ . Intuitively, we can see that (i) such a  $x$  must satisfy  $\exists S^+.\{o\}$  and B, (ii) an individual satisfying B must connect to another individual satisfying A which must have a  $R$ -path to nominal  $o$ , and (iii) two concepts A and B are disjoint.

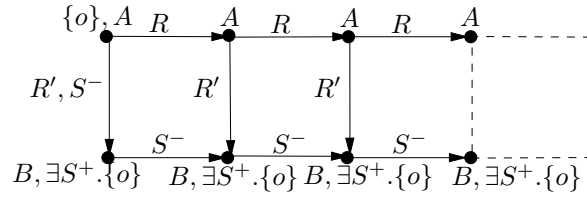


Fig. 1. An infinite non tree-like model of  $\mathcal{K}_1$

This example shows that methods ([7], [8], [6]) based on the hypothesis which says that if an ontology is consistent it has a *quasi-forest model*, could fail to address the problem of consistency in a DL including simultaneously  $\mathcal{O}$  (nominals),  $\mathcal{I}$  (inverse roles),  $\mathcal{Q}$  (number restrictions) and transitive closure of roles.

In this paper, we propose a decision procedure for the problem of consistency in  $\mathcal{SHOIQ}$  with transitive closure of roles in concept axioms. The underlying idea of our algorithm is founded on the *star-type* and *frame* notions introduced by Pratt-Hartmann [9]. This technique uses star-types to represent individuals and “tiles” them together to form a frame for representing a model. For each star-type  $\sigma$ , we maintain a function  $\delta(\sigma)$  which stores the number of individuals satisfying this star-type. To obtain termination condition, we introduce two additional structures into a frame : (i) the first one, namely *cycles*, describes duplicate parts of a model resulting from interactions of logic constructors in  $\mathcal{SHOIQ}$ , (ii) the second one, namely *blocking-blocked cycles*, describes parts of a model bordered by cycles which allow a frame to satisfy transitive closure of roles occurring in concepts of the form  $\exists R^+.C$ .

## 2 The Description Logic $\mathcal{SHOIQ}_{(+)}$

In this section, we present the syntax, the semantics and main inference problems of  $\mathcal{SHOIQ}_{(+)}$ . In addition, we introduce a tableau structure for  $\mathcal{SHOIQ}_{(+)}$ , which allows us to represent a model of a  $\mathcal{SHOIQ}_{(+)}$  knowledge base.

<sup>3</sup> This example is initially proposed by Sebastian Rudolph from an informal discussion

**Definition 1.** Let  $\mathbf{R}$  be a non-empty set of role names and  $\mathbf{R}_+ \subseteq \mathbf{R}$  be a set of transitive role names. We use  $\mathbf{R}_1 = \{P^- \mid P \in \mathbf{R}\}$  to denote a set of inverse roles, and  $\mathbf{R}_\oplus = \{Q^+ \mid Q \in \mathbf{R} \cup \mathbf{R}_1\}$  to denote a set of transitive closure of roles. Each element of  $\mathbf{R} \cup \mathbf{R}_1 \cup \mathbf{R}_\oplus$  is called a  $\text{SHOIQ}_{(+)}$ -role. A role inclusion axiom is of the form  $R \sqsubseteq S$  for two  $\text{SHOIQ}_{(+)}$ -roles  $R$  and  $S$  such that  $R \notin \mathbf{R}_\oplus$  and  $S \notin \mathbf{R}_\oplus$ . A role hierarchy  $\mathcal{R}$  is a finite set of role inclusion axioms. An interpretation  $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$  consists of a non-empty set  $\Delta^{\mathcal{I}}$  (domain) and a function  $\cdot^{\mathcal{I}}$  which maps each role name to a subset of  $\Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$  such that

$$\begin{aligned} R^{-\mathcal{I}} &= \{\langle x, y \rangle \in \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}} \mid \langle y, x \rangle \in R^{\mathcal{I}}\} \text{ for all } R \in \mathbf{R}, \\ \langle x, z \rangle \in S^{\mathcal{I}}, \langle z, y \rangle \in S^{\mathcal{I}} &\text{ implies } \langle x, y \rangle \in S^{\mathcal{I}} \text{ for each } S \in \mathbf{R}_+, \text{ and} \\ (Q^+)^{\mathcal{I}} &= \bigcup_{n>0} (Q^n)^{\mathcal{I}} \text{ with } (Q^1)^{\mathcal{I}} = Q^{\mathcal{I}}, \end{aligned}$$

$$(Q^n)^{\mathcal{I}} = \{\langle x, y \rangle \in (\Delta^{\mathcal{I}})^2 \mid \exists z \in \Delta^{\mathcal{I}}, \langle x, z \rangle \in (Q^{n-1})^{\mathcal{I}}, \langle z, y \rangle \in Q^{\mathcal{I}}\} \text{ for } Q^+ \in \mathbf{R}_\oplus$$

\* An interpretation  $\mathcal{I}$  satisfies a role hierarchy  $\mathcal{R}$  if  $R^{\mathcal{I}} \subseteq S^{\mathcal{I}}$  for each  $R \sqsubseteq S \in \mathcal{R}$ . Such an interpretation is called a model of  $\mathcal{R}$ , denoted by  $\mathcal{I} \models \mathcal{R}$ . To simplify notations for nested inverse roles and transitive closures of roles, we define two functions  $\cdot^\ominus$  and  $\cdot^\oplus$  as follows:

$$R^\ominus = \begin{cases} R^- & \text{if } R \in \mathbf{R}; \\ S & \text{if } R = S^- \text{ and } S \in \mathbf{R}; \\ (S^-)^+ & \text{if } R = S^+, S \in \mathbf{R}, \\ S^+ & \text{if } R = (S^-)^+, S \in \mathbf{R} \end{cases} \quad R^\oplus = \begin{cases} R^+ & \text{if } R \in \mathbf{R}; \\ S^+ & \text{if } R = (S^+)^+ \text{ and } S \in \mathbf{R}; \\ (S^-)^+ & \text{if } R = S^- \text{ and } S \in \mathbf{R}; \\ (S^-)^+ & \text{if } R = (S^+)^- \text{ and } S \in \mathbf{R} \end{cases}$$

\* A relation  $\boxsubseteq$  is defined as the transitive-reflexive closure  $\mathcal{R}^+$  of  $\sqsubseteq$  on  $\mathbf{R} \cup \{R^\ominus \sqsubseteq S^\ominus \mid R \sqsubseteq S \in \mathcal{R}\} \cup \{R^\oplus \sqsubseteq S^\oplus \mid R \sqsubseteq S \in \mathcal{R}\} \cup \{Q \sqsubseteq Q^\oplus \mid Q \in \mathbf{R} \cup \mathbf{R}_1\}$ . We define a function  $\text{Trans}(R)$  which returns true iff there is some  $Q \in \mathbf{R}_+ \cup \{P^\ominus \mid P \in \mathbf{R}_+\} \cup \{P^\oplus \mid P \in \mathbf{R} \cup \mathbf{R}_1\}$  such that  $Q \boxsubseteq R \in \mathcal{R}^+$ . A role  $R$  is called simple w.r.t.  $\mathcal{R}$  if  $\text{Trans}(R) = \text{false}$ .

The reason for the introduction of two functions  $\cdot^\ominus$  and  $\cdot^\oplus$  in Definition 1 is that they avoid using  $R^{-}$  and  $R^{++}$ . Moreover, it remains a unique nested case  $(R^-)^+$ . According to Definition 1, axiom  $R \sqsubseteq Q^\oplus$  is not allowed in a role hierarchy  $\mathcal{R}$  since this may lead to undecidability [10]. Notice that the closure  $\mathcal{R}^+$  may contain  $R \sqsubseteq Q^\oplus$  if  $R \sqsubseteq Q$  belongs to  $\mathcal{R}^+$ .

**Definition 2 (terminology).** Let  $\mathbf{C}$  be a non-empty set of concept names with a non-empty subset  $\mathbf{C}_o \subseteq \mathbf{C}$  of nominals. The set of  $\text{SHOIQ}_{(+)}$ -concepts is inductively defined as the smallest set containing all  $C$  in  $\mathbf{C}$ ,  $\top$ ,  $C \sqcap D$ ,  $C \sqcup D$ ,  $\neg C$ ,  $\exists R.C$ ,  $\forall R.C$ ,  $(\leq n S.C)$  and  $(\geq n S.C)$  where  $n$  is a positive integer,  $C$  and  $D$  are  $\text{SHOIQ}_{(+)}$ -concepts,  $R$  is an  $\text{SHOIQ}_{(+)}$ -role and  $S$  is a simple role w.r.t. a role hierarchy. We denote  $\perp$  for  $\neg \top$ . The interpretation function  $\cdot^{\mathcal{I}}$  of an interpretation  $\mathcal{I} = (\Delta^{\mathcal{I}}, \cdot^{\mathcal{I}})$  maps each concept name to a subset of  $\Delta^{\mathcal{I}}$  such that  $\top^{\mathcal{I}} = \Delta^{\mathcal{I}}$ ,  $(C \sqcap D)^{\mathcal{I}} = C^{\mathcal{I}} \cap D^{\mathcal{I}}$ ,  $(C \sqcup D)^{\mathcal{I}} = C^{\mathcal{I}} \cup D^{\mathcal{I}}$ ,  $(\neg C)^{\mathcal{I}} = \Delta^{\mathcal{I}} \setminus C^{\mathcal{I}}$ ,  $|\{o^{\mathcal{I}}\}| = 1$  for all  $o \in \mathbf{C}_o$ ,  $(\exists R.C)^{\mathcal{I}} = \{x \in \Delta^{\mathcal{I}} \mid \exists y \in \Delta^{\mathcal{I}}, \langle x, y \rangle \in R^{\mathcal{I}} \wedge y \in C^{\mathcal{I}}\}$ ,  $(\forall R.C)^{\mathcal{I}} = \{x \in \Delta^{\mathcal{I}} \mid \forall y \in \Delta^{\mathcal{I}}, \langle x, y \rangle \in R^{\mathcal{I}} \Rightarrow y \in C^{\mathcal{I}}\}$ ,  $(\geq n S.C)^{\mathcal{I}} = \{x \in \Delta^{\mathcal{I}} \mid |\{y \in C^{\mathcal{I}} \mid \langle x, y \rangle \in S^{\mathcal{I}}\}| \geq n\}$ ,  $(\leq n S.C)^{\mathcal{I}} = \{x \in \Delta^{\mathcal{I}} \mid |\{y \in C^{\mathcal{I}} \mid \langle x, y \rangle \in S^{\mathcal{I}}\}| \leq n\}$  where  $|S|$  is denoted for the cardinality of a set  $S$ . An axiom  $C \sqsubseteq D$  is called a general concept inclusion (GCI)

where  $C, D$  are  $\mathcal{SHOIQ}_{(+)}$ -concepts (possibly complex), and a finite set of GCIs is called a terminology  $\mathcal{T}$ . An interpretation  $\mathcal{I}$  satisfies a GCI  $C \sqsubseteq D$  if  $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$  and  $\mathcal{I}$  satisfies a terminology  $\mathcal{T}$  if  $\mathcal{I}$  satisfies each GCI in  $\mathcal{T}$ . Such an interpretation is called a model of  $\mathcal{T}$ , denoted by  $\mathcal{I} \models \mathcal{T}$ . A pair  $(\mathcal{T}, \mathcal{R})$  is called a  $\mathcal{SHOIQ}_{(+)}$  knowledge base where  $\mathcal{R}$  is a  $\mathcal{SHOIQ}_{(+)}$  role hierarchy and  $\mathcal{T}$  is a  $\mathcal{SHOIQ}_{(+)}$  terminology. A knowledge base  $(\mathcal{T}, \mathcal{R})$  is said to be consistent if there is a model  $\mathcal{I}$  of both  $\mathcal{T}$  and  $\mathcal{R}$ , i.e.,  $\mathcal{I} \models \mathcal{T}$  and  $\mathcal{I} \models \mathcal{R}$ . A concept  $C$  is called satisfiable w.r.t.  $(\mathcal{T}, \mathcal{R})$  iff there is some interpretation  $\mathcal{I}$  such that  $\mathcal{I} \models \mathcal{R}$ ,  $\mathcal{I} \models \mathcal{T}$  and  $C^{\mathcal{I}} \neq \emptyset$ . Such an interpretation is called a model of  $C$  w.r.t.  $(\mathcal{T}, \mathcal{R})$ . A concept  $D$  subsumes a concept  $C$  w.r.t.  $(\mathcal{T}, \mathcal{R})$ , denoted by  $C \sqsubseteq D$ , if  $C^{\mathcal{I}} \subseteq D^{\mathcal{I}}$  holds in each model  $\mathcal{I}$  of  $(\mathcal{T}, \mathcal{R})$ .  $\triangleleft$

Since unsatisfiability, subsumption and consistency w.r.t. a  $\mathcal{SHOIQ}_{(+)}$  knowledge base can be reduced to each other, it suffices to study knowledge base consistency. For the ease of construction, we assume all concepts to be in *negation normal form* (NNF), i.e., negation occurs only in front of concept names. Any  $\mathcal{SHOIQ}_{(+)}$ -concept can be transformed to an equivalent one in NNF by using DeMorgan's laws and some equivalences as presented in [11]. According to [12],  $\text{nnf}(C)$  can be computed in polynomial time in the size of  $C$ . For a concept  $C$ , we denote the nnf of  $C$  by  $\text{nnf}(C)$  and the nnf of  $\neg C$  by  $\dot{C}$ . Let  $D$  be a  $\mathcal{SHOIQ}_{(+)}$ -concept in NNF. We define  $\text{cl}(D)$  to be the smallest set that contains all sub-concepts of  $D$  including  $D$ . For a knowledge base  $(\mathcal{T}, \mathcal{R})$ , we reuse  $\text{cl}(\mathcal{T}, \mathcal{R})$ , which was introduced by Horrocks *et al.* [7], to denote all sub-concepts occurring in the axioms of  $(\mathcal{T}, \mathcal{R})$ . We have  $\text{cl}(\mathcal{T}, \mathcal{R})$  is bounded by  $\mathcal{O}(|(\mathcal{T}, \mathcal{R})|)$  [7]. To translate *star-type* and *frame* structures presented by Pratt-Hartmann (2005) for  $\mathcal{C}^2$  into those for  $\mathcal{SHOIQ}$ , we need to add new sets of concepts, denoted  $\text{cl}_1(\mathcal{T}, \mathcal{R})$  and  $\text{cl}_2(\mathcal{T}, \mathcal{R})$ , to the signature of a  $\mathcal{SHOIQ}_{(+)}$  knowledge base  $(\mathcal{T}, \mathcal{R})$ .

$$\text{cl}_1(\mathcal{T}, \mathcal{R}) = \{ \leq mS.C \mid \{(\leq nS.C), (\geq nS.C)\} \cap \text{cl}(\mathcal{T}, \mathcal{R}) \neq \emptyset, 1 \leq m \leq n \} \cup \{ \geq mS.C \mid \{(\leq nS.C), (\geq nS.C)\} \cap \text{cl}(\mathcal{T}, \mathcal{R}) \neq \emptyset, 1 \leq m \leq n \}$$

For a generating concept  $(\geq nS.C)$  and a set  $I \subseteq \{0, \dots, \log n + 1\}$ , we denote  $\mathcal{C}_{(\geq nS.C)}^I = \prod_{i \in I} C_{(\geq nS.C)}^i \sqcap \prod_{j \notin I} \dot{C}_{(\geq nS.C)}^j$  where  $C_{(\geq nS.C)}^i$  are new concept names

for  $0 \leq i \leq \log n + 1$ . We define  $\text{cl}_2(\mathcal{T}, \mathcal{R})$  as follows:

$$\text{cl}_2(\mathcal{T}, \mathcal{R}) = \{ C_{(\geq nS.C)}^i \mid (\geq nS.C) \in \text{cl}(\mathcal{T}, \mathcal{R}) \cup \text{cl}_1(\mathcal{T}, \mathcal{R}), 0 \leq i \leq \log n + 1 \} \cup \{ \mathcal{C}_{(\geq nS.C)}^I \mid (\geq nS.C) \in \text{cl}(\mathcal{T}, \mathcal{R}) \cup \text{cl}_1(\mathcal{T}, \mathcal{R}), I \subseteq \{0, \dots, \log n + 1\} \}$$

*Remark 1.* If numbers are encoded in binary then the number of new concept names  $C_{(\geq nS.C)}^i$  for  $0 \leq i \leq \log n + 1$ , is bounded by  $\mathcal{O}(|(\mathcal{T}, \mathcal{R})|)$  since  $n$  is bounded by  $\mathcal{O}(2^{|\mathcal{T}, \mathcal{R}|})$ . This implies that  $|\text{cl}_2(\mathcal{T}, \mathcal{R})|$  is bounded by  $\mathcal{O}(2^{|\mathcal{T}, \mathcal{R}|})$ . Note that two concepts  $\mathcal{C}_{(\geq nS.C)}^I$  and  $\mathcal{C}_{(\geq nS.C)}^J$  are disjoint for all  $I, J \subseteq \{0, \dots, \log n + 1\}$ ,  $I \neq J$ . The concepts  $\mathcal{C}_{(\exists S.C)}$  and  $\mathcal{C}_{(\geq nS.C)}^I$  will be used for building chromatic star-types. This notion will be clarified after introducing the frame structure (Definition 5).

Finally, we denote  $\mathbf{CL}(\mathcal{T}, \mathcal{R}) = \text{cl}(\mathcal{T}, \mathcal{R}) \cup \text{cl}_1(\mathcal{T}, \mathcal{R}) \cup \text{cl}_2(\mathcal{T}, \mathcal{R})$ , and use  $\mathbf{R}(\mathcal{T}, \mathcal{R})$  to denote the set of all role names occurring in  $\mathcal{T}, \mathcal{R}$  with their inverse. The definition of  $\mathbf{CL}(\mathcal{T}, \mathcal{R})$  is inspired from the Fischer-Ladner closure that was introduced in [13]. The closure  $\mathbf{CL}(\mathcal{T}, \mathcal{R})$  contains not only sub-concepts syntactically obtained from  $\mathcal{T}$

but also sub-concepts that are semantically derived from  $\mathcal{T}$  w.r.t.  $\mathcal{R}$ . For instance, if  $\forall S.C$  is a sub-concept from  $\mathcal{T}$  and  $R \sqsubseteq S \in \mathcal{R}$  then  $\forall R.C \in \mathbf{CL}(\mathcal{T}, \mathcal{R})$ .

To describe a model of a  $\mathcal{SHOIQ}_{(+)}$  knowledge base in a more intuitive way, we use a tableau structure that expresses semantic constraints resulting directly from the logic constructors in  $\mathcal{SHOIQ}_{(+)}$ . A tableau definition for  $\mathcal{SHOIQ}_{(+)}$  can be found in [14].

### 3 A Decision Procedure For $\mathcal{SHOIQ}_{(+)}$

This section starts by translating *star-type* and *frame* structures presented by Pratt-Hartmann (2005) for  $\mathcal{C}^2$  into those for  $\mathcal{SHOIQ}_{(+)}$ .

**Definition 3 (star-type).** Let  $(\mathcal{T}, \mathcal{R})$  be a  $\mathcal{SHOIQ}_{(+)}$  knowledge base. A star-type is a pair  $\sigma = \langle \lambda(\sigma), \xi(\sigma) \rangle$ , where  $\lambda(\sigma) \in 2^{\mathbf{CL}(\mathcal{T}, \mathcal{R})}$  is called core label,  $\xi(\sigma) = \langle \langle r_1, l_1 \rangle, \dots, \langle r_d, l_d \rangle \rangle$  is a  $d$ -tuple over  $2^{\mathbf{R}(\mathcal{T}, \mathcal{R})} \times 2^{\mathbf{CL}(\mathcal{T}, \mathcal{R})}$ . A pair  $\langle r, l \rangle$  is a ray of  $\sigma$  if  $\langle r, l \rangle = \langle r_i, l_i \rangle$  for some  $1 \leq i \leq d$ . We use  $\langle r(\rho), l(\rho) \rangle$  to denote a ray  $\rho = \langle r, l \rangle$  where  $r(\rho) = r$  and  $l(\rho) = l$ .

- A star-type  $\sigma$  is nominal if  $o \in \lambda(\sigma)$  for some  $o \in \mathbf{C}_o$ .
- A star-type  $\sigma$  is chromatic if  $\rho \neq \rho'$  implies  $l(\rho) \neq l(\rho')$  for two rays  $\rho, \rho'$  of  $\sigma$ . When a star-type  $\sigma$  is chromatic,  $\xi(\sigma)$  can be considered as a set of rays.
- Two star-types  $\sigma, \sigma'$  are equivalent if  $\lambda(\sigma) = \lambda(\sigma')$ , and there is a bijection  $\pi$  between  $\xi(\sigma)$  and  $\xi(\sigma')$  such that  $\pi(\rho) = \rho'$  implies  $r(\rho') = r(\rho)$  and  $l(\rho') = l(\rho)$ .

We denote  $\Sigma$  for the set of all star-types for  $(\mathcal{T}, \mathcal{R})$ . ◁

Note that for a chromatic star-type  $\sigma$ ,  $\xi(\sigma)$  can be considered as a set of rays since rays are distinct and not ordered. We can think of a star-type  $\sigma$  as the set of individuals  $x$  satisfying all concepts in  $\lambda(\sigma)$ , and each ray  $\rho$  of  $\sigma$  corresponds to a “neighbor” individual  $x_i$  of  $x$  such that  $r(\rho)$  is the label of the link between  $x$  and  $x_i$ ; and  $x_i$  satisfies all concepts in  $l(\rho)$ . In this case, we say that  $x$  satisfies  $\sigma$ .

**Definition 4 (valid star-type).** Let  $(\mathcal{T}, \mathcal{R})$  be a  $\mathcal{SHOIQ}_{(+)}$  knowledge base. Let  $\sigma$  be a star-type for  $(\mathcal{T}, \mathcal{R})$  where  $\sigma = \langle \lambda(\sigma), \xi(\sigma) \rangle$ . The star-type  $\sigma$  is valid if  $\sigma$  is chromatic and the following conditions are satisfied:

1. If  $C \sqsubseteq D \in \mathcal{T}$  then  $\text{nnf}(\neg C \sqcup D) \in \lambda(\sigma)$ ;
2.  $\{A, \neg A\} \not\subseteq \lambda$  for every concept name  $A$  where  $\lambda = \lambda(\sigma)$  or  $\lambda = l(\rho)$  for each  $\rho \in \xi(\sigma)$ ;
3. If  $C_1 \sqcap C_2 \in \lambda(\sigma)$  then  $\{C_1, C_2\} \subseteq \lambda(\sigma)$ ;
4. If  $C_1 \sqcup C_2 \in \lambda(\sigma)$  then  $\{C_1, C_2\} \cap \lambda(\sigma) \neq \emptyset$ ;
5. If  $\exists R.C \in \lambda(\sigma)$  then there is some ray  $\rho \in \xi(\sigma)$  such that  $C \in l(\rho)$  and  $R \in r(\rho)$ ;
6. If  $(\leq nS.C) \in \lambda(\sigma)$  and there is some ray  $\rho \in \xi(\sigma)$  such that  $S \in r(\rho)$  then  $C \in l(\rho)$  or  $\dot{\neg}C \in l(\rho)$ ;
7. If  $(\leq nS.C) \in \lambda(\sigma)$  and there is some ray  $\rho \in \xi(\sigma)$  such that  $C \in l(\rho)$  and  $S \in r(\rho)$  then there is some  $1 \leq m \leq n$  such that  $\{(\leq mS.C), (\geq mS.C)\} \subseteq \lambda(\sigma)$ ;
8. For each ray  $\rho \in \xi(\sigma)$ , if  $R \in r(\rho)$  and  $R \sqsubseteq S$  then  $S \in r(\rho)$ ;

9. If  $\forall R.C \in \lambda(\sigma)$  and  $R \in r(\rho)$  for some ray  $\rho \in \xi(\sigma)$  then  $C \in l(\rho)$ ;
10. If  $\forall R.D \in \lambda(\sigma)$ ,  $S \sqsubseteq R$ ,  $\text{Trans}(S)$  and  $R \in r(\rho)$  for some ray  $\rho \in \xi(\sigma)$  then  $\forall S.D \in l(\rho)$ ;
11. If  $\forall Q^\oplus.C \in \lambda(\sigma)$ ,  $R \sqsubseteq Q$  and  $R \in r(\rho)$  for some ray  $\rho \in \xi(\sigma)$  then  $\forall Q^\oplus.C \in l(\rho)$ ;
12. If  $\exists Q^\oplus.C \in \lambda(\sigma)$  then  $(\exists Q.C \sqcup \exists Q^\oplus.C) \in \lambda(\sigma)$ ;
13. If  $(\geq nS.C) \in \lambda(\sigma)$  then there are  $n$  distinct rays  $\rho_1, \dots, \rho_n \in \xi(\sigma)$  such that  $\{C, \mathcal{C}_{(\geq nS.C)}^{I_i}\} \subseteq l(\rho_i)$ ,  $S \in r(\rho_i)$  for all  $1 \leq i \leq n$ ; and  $I_j, I_k \subseteq \{0, \dots, \log n + 1\}$ ,  $I_j \neq I_k$  for all  $1 \leq j < k \leq n$ .
14. If  $(\leq nS.C) \in \lambda(\sigma)$  and there do not exist  $n + 1$  rays  $\rho_0, \dots, \rho_n \in \xi(\sigma)$  such that  $C \in l(\rho_i)$  and  $S \in r(\rho_i)$  for all  $0 \leq i \leq n$ .  $\triangleleft$

Roughly speaking, a star-type  $\sigma$  is valid if each individual  $x$  satisfies *semantically* all concepts in  $\lambda(\sigma)$ . In fact, each condition in Definition 4 represents the semantics of a constructor in  $\mathcal{SHOIQ}_{(+)}$  except for transitive closure of roles. From valid star-types, we can “tile” a model instead of using expansion rules for generating nodes as described in tableau algorithms. Before presenting how to “tile” a model from star-types, we need some notation that will be used in the remainder of the paper.

**Notation 1** We call  $\mathcal{P} = \langle (\sigma_1, \rho_1, d_1), \dots, (\sigma_k, \rho_k, d_k) \rangle$  a sequence where  $\sigma_i \in \Sigma$ ,  $\rho_i \in \xi(\sigma_i)$  and  $d_i \in \mathbb{N}$  for  $1 \leq i \leq k$ .

- $\text{tail}(\mathcal{P}) = (\sigma_k, \rho_k, d_k)$ ,  $\text{tail}_\sigma(\mathcal{P}) = \sigma_k$ ,  $\text{tail}_\rho(\mathcal{P}) = \rho_k$ ,  $\text{tail}_\delta(\mathcal{P}) = d_k$  and  $|\mathcal{P}| = k$ . We denote  $\mathcal{L}(\mathcal{P}) = \lambda(\text{tail}_\sigma(\mathcal{P}))$ .
- $\mathbf{p}^i(\mathcal{P}) = (\sigma_i, \rho_i, d_i)$ ,  $\mathbf{p}_\sigma^i(\mathcal{P}) = \sigma_i$ ,  $\mathbf{p}_\rho^i(\mathcal{P}) = \rho_i$  and  $\mathbf{p}_\delta^i(\mathcal{P}) = d_i$  for each  $1 \leq i \leq k$ .
- an operation  $\text{add}(\mathcal{P}, (\sigma, \rho, d))$  extends  $\mathcal{P}$  to a new sequence with  $\text{add}(\mathcal{P}, (\sigma, \rho, d)) = \langle \mathcal{P}, (\sigma, \rho, d) \rangle$ .

**Definition 5 (frame).** Let  $(\mathcal{T}, \mathcal{R})$  be a  $\mathcal{SHOIQ}_{(+)}$  knowledge base. A frame for  $(\mathcal{T}, \mathcal{R})$  is a tuple  $\mathcal{F} = \langle \mathcal{N}, \mathcal{N}_o, \Omega, \delta \rangle$ , where

1.  $\mathcal{N}$  is a set of valid star-types such that  $\sigma$  is not equivalent to  $\sigma'$  for all  $\sigma, \sigma' \in \mathcal{N}$ ;
2.  $\mathcal{N}_o \subseteq \mathcal{N}$  is a set of nominal star-types;
3.  $\Omega$  is a function that maps each pair  $(\sigma, \rho)$  with  $\sigma \in \mathcal{N}$  and  $\rho \in \xi(\sigma)$  to a sequence  $\Omega(\sigma, \rho) = \langle (\sigma_1, \rho_1, d_1), \dots, (\sigma_m, \rho_m, d_m) \rangle$  with  $\sigma_i \in \mathcal{N}$ ,  $\rho_i \in \xi(\sigma_i)$ ,  $d_i \in \mathbb{N}$  for  $1 \leq i \leq m$  such that for each  $\sigma_i$  with  $1 \leq i \leq m$ , it holds that  $l(\rho) = \lambda(\sigma_i)$ ,  $l(\rho_i) = \lambda(\sigma)$  and  $r(\rho_i) = r^-(\rho)$  where  $r^-(\rho) = \{R^\ominus \mid R \in r(\rho)\}$ .
4.  $\delta$  is a function  $\delta : \mathcal{N} \rightarrow \mathbb{N}$ . By abuse of notation, we also use  $\delta$  to denote a function which maps each pair  $(\sigma, \rho)$  with  $\sigma \in \mathcal{N}$  and  $\rho \in \xi(\sigma)$  into a number in  $\mathbb{N}$ , i.e.,  $\delta(\sigma, \rho) \in \mathbb{N}$ .  $\triangleleft$

The frame structure, as introduced in Definition 5, allows us to compress individuals of a model into star-types. For each star-type  $\sigma$  and each ray  $\rho \in \xi(\sigma)$ , a list  $\Omega(\sigma, \rho)$  of triples  $(\sigma_i, \rho_i, d_i)$  with  $\rho_i \in \xi(\sigma_i)$  is maintained where  $\sigma_i$  is a “neighbor” star-type of  $\sigma$  via  $\rho \in \xi(\sigma)$ , and  $d_i$  indicates the  $d_i$ -th “layer” of rays of  $\sigma_i$ . We can think a layer of rays of  $\sigma_i$  as an individual that connects to its neighbor individuals via the rays of  $\sigma_i$ . The following definition presents how to connect such layers to form paths in a frame.

**Definition 6 (path).** Let  $\mathcal{F} = \langle \mathcal{N}, \mathcal{N}_o, \Omega, \delta \rangle$  be a frame for a  $\mathcal{SHOIQ}_{(+)}$  knowledge base  $(\mathcal{T}, \mathcal{R})$ . A path is inductively defined as follows:

1. A sequence  $\langle \emptyset, (\sigma, \rho, 1) \rangle$  is a path if  $\sigma \in \mathcal{N}_o$  and  $\rho \in \xi(\sigma)$ ;
2. A sequence  $\langle \mathcal{P}, (\sigma, \rho, d) \rangle$  with  $\mathcal{P} \neq \emptyset$  and  $\text{tail}(\mathcal{P}) = (\sigma_0, \rho_0, d_0)$ , is a path if  $(\sigma, \rho, d) = \mathbf{p}^{d_0}(\Omega(\sigma_0, \rho'))$  for each  $\rho' \neq \rho_0$ . In this case, we say that  $\langle \mathcal{P}, (\sigma, \rho, d) \rangle$  is the  $\rho'$ -neighbor of  $\mathcal{P}$ , and two paths  $\mathcal{P}, \langle \mathcal{P}, (\sigma, \rho, d) \rangle$  are neighbors. Additionally, if  $\langle \mathcal{P}, (\sigma, \rho, d) \rangle$  is a  $\rho'$ -neighbor of  $\mathcal{P}$  and  $Q \in r(\rho')$  then  $\langle \mathcal{P}, (\sigma, \rho, d) \rangle$  is a  $Q$ -neighbor of  $\mathcal{P}$ . In this case, we say that  $\langle \mathcal{P}, (\sigma, \rho, d) \rangle$  is a  $Q$ -neighbor of  $\mathcal{P}$ , or  $\mathcal{P}$  is a  $Q^\ominus$ -neighbor of  $\langle \mathcal{P}, (\sigma, \rho, d) \rangle$ .

We define  $\mathcal{P} \sim \mathcal{P}'$  if  $\text{tail}_\sigma(\mathcal{P}) = \text{tail}_\sigma(\mathcal{P}')$  and  $\text{tail}_\delta(\mathcal{P}) = \text{tail}_\delta(\mathcal{P}')$ . Since  $\sim$  is an equivalence relation over the set of all paths, we use  $\mathcal{P}$  to denote the set of all equivalence classes  $[\mathcal{P}]$  of paths in  $\mathcal{F}$ . For  $[\mathcal{P}], [\mathcal{Q}] \in \mathcal{P}$ , we define:

1.  $[\mathcal{P}]$  is a neighbor ( $\rho'$ -neighbor) of  $[\mathcal{Q}]$  if there are  $\mathcal{P}' \in [\mathcal{P}]$  and  $\mathcal{Q}' \in [\mathcal{Q}]$  such that  $\mathcal{Q}'$  is a neighbor ( $\rho'$ -neighbor) of  $\mathcal{P}'$ ;
2.  $[\mathcal{Q}]$  is a reachable path of  $[\mathcal{P}]$  if there are  $[\mathcal{P}_1], \dots, [\mathcal{P}_n] \in \mathcal{P}$  such that  $[\mathcal{P}_{i+1}]$  is a neighbor of  $[\mathcal{P}_i]$  for all  $1 \leq i < n$  where  $[\mathcal{P}_1] = [\mathcal{P}]$  and  $[\mathcal{Q}] = [\mathcal{P}_n]$ .
3.  $[\mathcal{Q}]$  is a  $Q$ -neighbor of  $[\mathcal{P}]$  if there are  $\mathcal{P}' \in [\mathcal{P}]$  and  $\mathcal{Q}' \in [\mathcal{Q}]$  such that  $\mathcal{Q}'$  is a  $Q$ -neighbor of  $\mathcal{P}'$ , or  $\mathcal{P}'$  is a  $Q^\ominus$ -neighbor of  $\mathcal{Q}'$ ;
4.  $[\mathcal{Q}]$  is a  $Q$ -reachable path of  $[\mathcal{P}]$  if there are  $[\mathcal{P}_1], \dots, [\mathcal{P}_n] \in \mathcal{P}$  such that  $[\mathcal{P}_{i+1}]$  is a  $Q$ -neighbor of  $[\mathcal{P}_i]$  for all  $1 \leq i < n$  where  $[\mathcal{P}_1] = [\mathcal{P}]$  and  $[\mathcal{Q}] = [\mathcal{P}_n]$ .  $\triangleleft$

Note that for two paths  $\mathcal{P}, \mathcal{P}'$  with  $\text{tail}_\rho(\mathcal{P}) \neq \text{tail}_\rho(\mathcal{P}')$ , we have  $\mathcal{P} \sim \mathcal{P}'$  if  $\text{tail}_\sigma(\mathcal{P}) = \text{tail}_\sigma(\mathcal{P}')$  and  $\text{tail}_\delta(\mathcal{P}) = \text{tail}_\delta(\mathcal{P}')$ . This does not allow for extending  $\text{tail}_\rho(\mathcal{P})$  to  $\text{tail}_\rho([\mathcal{P}])$ . As a consequence, there may be several “predecessors” of an equivalence class  $[\mathcal{P}]$ . However, we can define  $\text{tail}_\sigma([\mathcal{P}]) = \text{tail}_\sigma(\mathcal{P})$ ,  $\text{tail}_\delta([\mathcal{P}]) = \text{tail}_\delta(\mathcal{P})$  and  $\mathcal{L}([\mathcal{P}]) = \mathcal{L}(\mathcal{P})$ . In the sequel, we use  $\mathcal{P}$  instead of  $[\mathcal{P}]$  whenever it is clear from the context.

**Definition 7 (cycle).** Let  $\mathcal{F} = \langle \mathcal{N}, \mathcal{N}_o, \Omega, \delta \rangle$  be a frame for a  $\mathcal{SHOI}\mathcal{Q}_{(+)}$  knowledge base  $(\mathcal{T}, \mathcal{R})$  with a set  $\mathcal{P}$  of paths in  $\mathcal{F}$ .

1. A cycle is a set  $\Theta$  of triples  $(\mathcal{P}, \rho, \mathcal{Q})$  with  $\mathcal{P}, \mathcal{Q} \in \mathcal{P}$  and  $\rho \in \xi(\text{tail}_\sigma(\mathcal{P}))$  such that for each  $(\mathcal{P}, \rho, \mathcal{Q}) \in \Theta$  the following conditions are satisfied:
  - (a)  $\text{tail}_\delta(\mathcal{P}) > 1$  and  $\text{tail}_\delta(\mathcal{Q}) > 1$ ;
  - (b) If  $\mathcal{P}'$  is the  $\rho$ -neighbor of  $\mathcal{P}$  then  $\text{tail}_\sigma(\mathcal{P}') = \text{tail}_\sigma(\mathcal{Q})$ ;
  - (c) for each sequence  $\mathcal{P}_1, \dots, \mathcal{P}_n \in \mathcal{P}$  such that  $\mathcal{P}_1 = \mathcal{P}$ ,  $\mathcal{P}_2$  is not the  $\rho$ -neighbor of  $\mathcal{P}$ , and  $\mathcal{P}_{i+1}$  is a neighbor of  $\mathcal{P}_i$  for  $1 \leq i < n$ , there is some  $(\mathcal{P}'', \rho'', \mathcal{Q}'') \in \Theta$  such that
    - i. either  $\mathcal{Q}'' = \mathcal{P}_j$ ,  $\text{tail}_\sigma(\mathcal{P}_{j+1}) = \text{tail}_\sigma(\mathcal{P}'')$ ,  $\text{tail}_\delta(\mathcal{P}_{j+1}) \geq \text{tail}_\delta(\mathcal{P}'')$  and  $\mathcal{P}_j$  is the  $\rho$ -neighbor of  $\mathcal{P}_{j+1}$  for some  $1 < j < n$ ,
    - ii. or there are  $\mathcal{P}_{n+1}, \dots, \mathcal{P}_{n+m} \in \mathcal{P}$  with  $\mathcal{Q}'' = \mathcal{P}_{n+m-1}$ ,  $\text{tail}_\sigma(\mathcal{P}_{n+m}) = \text{tail}_\sigma(\mathcal{P}'')$ ,  $\text{tail}_\delta(\mathcal{P}_{n+m}) \geq \text{tail}_\delta(\mathcal{P}'')$  and  $\mathcal{P}_{i+1}$  is a neighbor of  $\mathcal{P}_i$  for  $n \leq i < n+m$ .

In this case, we say that  $\mathcal{Q}$  is cycled by  $\mathcal{P}$  via  $\rho$ .

2. A cycle  $\Theta'$  is a reachable cycle of  $\Theta$  if for each  $(\mathcal{P}, \rho, \mathcal{Q}) \in \Theta$  and for each sequence  $\mathcal{P}_1, \dots, \mathcal{P}_n \in \mathcal{P}$  such that  $\mathcal{P}_1 = \mathcal{P}$ ,  $\mathcal{P}_2$  is not a  $\rho$ -neighbor of  $\mathcal{P}$ , and  $\mathcal{P}_{i+1}$  is a neighbor of  $\mathcal{P}_i$  for  $1 \leq i < n$ , there is some  $(\mathcal{P}'', \rho'', \mathcal{Q}'') \in \Theta'$  such that



- (a) either  $\mathcal{Q}'' = \mathcal{P}_j$ ,  $\text{tail}_\sigma(\mathcal{P}_{j+1}) = \text{tail}_\sigma(\mathcal{P}'')$ ,  $\text{tail}_\delta(\mathcal{P}_{j+1}) \geq \text{tail}_\delta(\mathcal{P}'')$  and  $\mathcal{P}_j$  is a  $\rho$ -neighbor of  $\mathcal{P}_{j+1}$  for some  $1 < j < n$ ,
- (b) or there are  $\mathcal{P}_{n+1}, \dots, \mathcal{P}_{n+m} \in \mathcal{P}$  with  $\mathcal{Q}'' = \mathcal{P}_{n+m-1}$ ,  $\text{tail}_\sigma(\mathcal{P}_{n+m}) = \text{tail}_\sigma(\mathcal{P}'')$ ,  $\text{tail}_\delta(\mathcal{P}_{n+m}) \geq \text{tail}_\delta(\mathcal{P}'')$  and  $\mathcal{P}_{i+1}$  is a neighbor of  $\mathcal{P}_i$  for  $n \leq i < n+m$ .  $\triangleleft$

Note that cycles may encapsulate loops if  $\text{tail}_\delta(\mathcal{P}_{j+1}) = \text{tail}_\delta(\mathcal{P}'')$  holds in Conditions 1(c)i and 1(c)ii, Definition 7. Let  $\Theta$  be a cycle in a frame. Definition 7 ensures that each reachable path of some path  $\mathcal{P}$  with  $(\mathcal{P}, \rho, \mathcal{Q}) \in \Theta$  goes through a star-type  $\sigma = \text{tail}_\sigma(\mathcal{Q})$  with some  $(\mathcal{P}', \rho', \mathcal{Q}') \in \Theta$ . Such a cycle, which is similar to blocking-blocked nodes in completion graphs for *SHOIQ* [7], allows for “unravelling” infinitely the frame to obtain a model of a KB in *SHOIQ* (without transitive closure of roles). This means that we can extend the set  $\mathcal{P}$  of paths by adding infinitely paths which lengthen  $\mathcal{Q}$  such that  $(\mathcal{P}, \rho, \mathcal{Q}) \in \Theta$  and  $\mathcal{P}$  is not a neighbor of  $\mathcal{Q}$ . However, such a cycle structure is not sufficient to represent models of a KB with transitive closure of roles since a concept such as  $\exists Q^\oplus.D \in \mathcal{L}(\mathcal{P})$  can be satisfied by a  $Q$ -reachable path  $\mathcal{P}'$  of  $\mathcal{P}$  which is arbitrarily far from  $\mathcal{P}$ . There are the following possibilities for an algorithm which builds a frame: (i) the algorithm stops building the frame as soon as a cycle  $\Theta$  is detected such that each concept of the form  $\exists Q^\oplus.D$  occurring in  $\mathcal{L}(\mathcal{P})$  for each cycling path  $\mathcal{P}$  of  $\Theta$  is satisfied, i.e.,  $\mathcal{P}$  has a  $Q$ -reachable path  $\mathcal{P}'$  with  $\exists Q.D \in \mathcal{L}(\mathcal{P})$ , (ii) despite of several detected cycles, the algorithm continues building the frame until each concept of the form  $\exists Q^\oplus.D$  occurring in  $\mathcal{L}(\mathcal{P})$  is satisfied for each cycling path  $\mathcal{P}$  of  $\Theta$ . If we adopt the first possibility, the completeness of such an algorithm cannot be established since there are models in which paths satisfying concepts of the form  $\exists Q^\oplus.D$  can spread over several “iterative structures” such as cycles. For this reason, we adopt the second possibility by introducing into frames an additional structure, namely *blocking-blocked cycles*, which determines a sequence of cycles  $\Theta_1, \dots, \Theta_k$  such that  $\Theta_{i+1}$  is a reachable cycle of  $\Theta_i$ . Reachability of cycles allows for “unravelling” the frame between cycled paths  $\mathcal{Q}'$  with  $(\mathcal{P}', \rho', \mathcal{Q}') \in \Theta_k$  and cycling paths  $\mathcal{P}$  with  $(\mathcal{P}, \rho, \mathcal{Q}) \in \Theta_1$ .

**Definition 8 (blocking).** Let  $\mathcal{F} = \langle \mathcal{N}, \mathcal{N}_o, \Omega, \delta \rangle$  be a frame for a *SHOIQ*<sub>(+)</sub> knowledge base  $(\mathcal{T}, \mathcal{R})$  with a set  $\mathcal{P}$  of all paths in  $\mathcal{F}$ .

1. A cycle  $\Theta'$  is blockable by a cycle  $\Theta$  if  $\Theta'$  is a reachable cycle of  $\Theta$ , and for each  $(\mathcal{P}', \rho', \mathcal{Q}') \in \Theta'$  there is some  $(\mathcal{P}, \rho, \mathcal{Q}) \in \Theta$  such that  $\mathcal{L}(\mathcal{P}) = \mathcal{L}(\mathcal{P}')$ ,  $\mathcal{L}(\mathcal{Q}) = \mathcal{L}(\mathcal{Q}')$  and  $r(\rho) = r(\rho')$ . In this case, we say that  $\mathcal{Q}'$  is blockable by  $\mathcal{P}$  via  $\rho$ .
2. A cycle  $\Theta'$  is blocked by a cycle  $\Theta$  if there are  $\Theta_1, \dots, \Theta_k$  with  $\Theta = \Theta_1$ ,  $\Theta' = \Theta_k$  such that  $\Theta_{i+1}$  is blockable by  $\Theta_i$  for  $1 \leq i < k$ , and for each  $(\lambda, s, \lambda') \in 2^{\text{CL}(\mathcal{T}, \mathcal{R})} \times 2^{\mathbf{R}(\mathcal{T}, \mathcal{R})} \times 2^{\text{CL}(\mathcal{T}, \mathcal{R})}$  with  $\exists Q^\oplus.D \in \lambda$ ,
  - if there is some path  $\mathcal{P}_k$  with  $(\mathcal{P}_k, \rho_k, \mathcal{Q}_k) \in \Theta_k$ ,  $\mathcal{L}(\mathcal{P}_k) = \lambda$ ,  $\mathcal{L}(\mathcal{Q}_k) = \lambda'$ ,  $s = r(\rho_k)$  such that  $\mathcal{P}_k$  has no  $Q$ -reachable path  $\mathcal{P}'_k$  with  $\exists Q.D \in \mathcal{L}(\mathcal{P}'_k)$ ,
  - then there is some  $\mathcal{P}_1$  with  $(\mathcal{P}_1, \rho_1, \mathcal{Q}_1) \in \Theta_1$ ,  $\mathcal{L}(\mathcal{P}_1) = \lambda$ ,  $\mathcal{L}(\mathcal{Q}_1) = \lambda'$ ,  $r(\rho_1) = s$  such that for each  $\exists P^\oplus.C \in \mathcal{L}(\mathcal{P}_1)$ ,
    - there is a  $P$ -reachable path  $\mathcal{Q}''$  of  $\mathcal{P}_1$  with  $\exists P.C \in \mathcal{L}(\mathcal{Q}'')$ , and

- there are two triples  $(\mathcal{P}_j, \rho_j, \mathcal{Q}_j) \in \Theta_j$  and  $(\mathcal{P}_{j+1}, \rho_{j+1}, \mathcal{Q}_{j+1}) \in \Theta_{j+1}$  for some  $1 \leq j < k$ , which satisfy that  $\mathcal{Q}''$  is a reachable path of  $\mathcal{Q}_j$  and  $\mathcal{Q}_{j+1}$  is a reachable path of  $\mathcal{Q}''$ .

In this case, we say that  $\mathcal{P}_1$  blocks  $\mathcal{P}_k$  via  $\rho_k$ .  $\triangleleft$

According to Definition 8, there are a sequentially reachable cycles between a blocking cycle  $\Theta_1$  and a blocked cycle  $\Theta_k$ , which allows for unravelling the frame between  $\Theta_k$  and  $\Theta_1$ . Condition 2 Definition 8 says that if  $(\mathcal{P}, \rho, \mathcal{Q}) \in \Theta_k$  and  $\mathcal{L}(\mathcal{P})$  contains concepts of the form  $\exists Q^\oplus.D$  which are not satisfied by reachable paths of  $\mathcal{P}$  then there exists some  $\mathcal{P}'$  with  $(\mathcal{P}', \rho', \mathcal{Q}') \in \Theta_1$  which allows for satisfying these concepts  $\exists Q^\oplus.D$  in  $\mathcal{L}(\mathcal{P})$  by unravelling. We would like to note that a path  $\mathcal{P}$  is blocked if there is some blocked cycle  $\Theta_k$  such that  $(\mathcal{P}, \rho, \mathcal{Q}) \in \Theta_k$ .

**Definition 9 (valid frame).** Let  $(\mathcal{T}, \mathcal{R})$  be a SHOIQ knowledge base. A frame  $\mathcal{F} = \langle \mathcal{N}, \mathcal{N}_o, \Omega, \delta \rangle$  is valid if the following conditions are satisfied:

1. For each  $o \in \mathbf{C}_o$  there is a unique  $\sigma_o \in \mathcal{N}_o$  such that  $o \in \lambda(\sigma_o)$  and  $\delta(\sigma_o) = 1$ ;
2. For each  $\sigma \in \mathcal{N}$ ,  $\sigma$  is valid;
3. If  $\exists Q^\oplus.C \in \lambda(\text{tail}_\sigma(\mathcal{P}_0))$  for some  $\mathcal{P}_0 \in \mathcal{P}$  then there are  $\mathcal{P}, \mathcal{P}' \in \mathcal{P}$  such that one of the following conditions is satisfied:
  - (a)  $\mathcal{P}_0 = \mathcal{P} = \mathcal{P}'$  and  $\exists Q.C \in \mathcal{L}(\mathcal{P}_0)$ ;
  - (b)  $\mathcal{P}'$  is a  $Q$ -reachable of  $\mathcal{P}$ , and  $\exists Q.C \in \mathcal{L}(\mathcal{P}')$  where  $\mathcal{P} = \mathcal{P}_0$  or  $\mathcal{P}$  blocks  $\mathcal{P}_0$ ;
  - (c)  $\mathcal{P}$  is a  $Q^\ominus$ -reachable of  $\mathcal{P}'$ , and  $\exists Q.C \in \mathcal{L}(\mathcal{P}')$  where  $\mathcal{P} = \mathcal{P}_0$  or  $\mathcal{P}$  blocks  $\mathcal{P}_0$ .  $\triangleleft$

Conditions 1-3 in Definition 9 ensure satisfaction of tableau properties [14]. In particular, Condition 3 takes into account satisfaction of transitive closure of roles. In fact, for each blocking path  $\mathcal{P}$  with  $\exists Q^\oplus.D \in \mathcal{L}(\mathcal{P})$ , this condition says that  $\mathcal{P}$  must be satisfied by a  $Q$ -reachable path  $\mathcal{P}'$  of  $\mathcal{P}$  between a blocking cycle  $\Theta_1$  and a blocked cycle  $\Theta_k$ . If there is some path  $\mathcal{Q}$  between  $\Theta_1$  and  $\Theta_k$  with  $\exists S^\oplus.C \in \mathcal{L}(\mathcal{Q})$  that is not satisfied, due to the construction of star-types and frames, a concept  $\exists S^\oplus.C$  is propagated along  $S$ -reachable paths of  $\mathcal{Q}$ . This implies that  $\mathcal{Q}$  has a  $S$ -reachable path  $\mathcal{Q}'$  such that  $\mathcal{Q}'$  is blocked and  $\exists S^\oplus.C \in \mathcal{L}(\mathcal{Q}')$ . By unravelling (details will be given in soundness proof), we can build an (extended)  $S$ -reachable path  $\mathcal{Q}''$  of  $\mathcal{Q}'$  such that  $\exists S.C \in \mathcal{L}(\mathcal{Q}'')$ .

We now present Algorithm 1 for building a frame which is valid if the conditions in Definition 9 are satisfied. This algorithm starts by adding nominal star-types to the frame. For each non blocked path  $\mathcal{P}$  with a ray  $\rho \in \xi(\text{tail}_\sigma(\mathcal{P}))$  such that  $\delta(\text{tail}_\sigma(\mathcal{P}))$  is minimal and there is a difference between  $\delta(\text{tail}_\sigma(\mathcal{P}))$  and  $\delta(\text{tail}_\sigma(\mathcal{P}), \rho)$ , the algorithm picks in a nondeterministic way a valid star-type  $\omega$  that matches  $\text{tail}_\sigma(\mathcal{P})$  via  $\rho$ , and updates  $\Omega(\text{tail}_\sigma(\mathcal{P}), \rho)$ ,  $\Omega(\omega, \rho')$ ,  $\delta(\text{tail}_\sigma(\mathcal{P}), \rho)$ ,  $\delta(\omega, \rho')$ , eventually,  $\delta(\text{tail}_\sigma(\mathcal{P}))$  and  $\delta(\omega)$  by calling `updateFrame( $\dots$ )` [14]. The algorithm terminates when a blocked cycle is detected.

Figure 2 depicts a frame when executing Algorithm 1 for  $\mathcal{K}_1$  in the example presented in Section 1. The algorithm builds a frame  $\mathcal{F} = \langle \mathcal{N}, \mathcal{N}_o, \Omega, \delta \rangle$  where  $\mathcal{N} = \{\sigma_0, \sigma_1, \sigma_2, \sigma_3, \sigma_4\}$  and  $\mathcal{N}_o = \{\sigma_0\}$ . The dashed arrows indicate how the function

**Require:** A  $\mathcal{SHOIQ}_{(+)}$  knowledge base  $(\mathcal{T}, \mathcal{R})$   
**Ensure:** A frame  $\langle \mathcal{N}, \mathcal{N}_o, \Omega, \delta \rangle$  for  $(\mathcal{T}, \mathcal{R})$

- 1: Let  $\Sigma$  be the set of all star-types for  $(\mathcal{T}, \mathcal{R})$
- 2: **for all**  $o \in \mathbf{C}_o$  **do**
- 3:   **if** there is no  $\sigma \in \mathcal{N}$  such that  $o \in \lambda(\sigma)$  **then**
- 4:     Choose a star-type  $\sigma_o \in \Sigma$  such that  $o \in \lambda(\sigma_o)$
- 5:     Set  $\delta(\sigma_o) = 1$ ,  $\mathcal{N} = \mathcal{N} \cup \{\sigma_o\}$  and  $\mathcal{N}_o = \mathcal{N}_o \cup \{\sigma_o\}$
- 6:     Set  $\delta(\sigma_o, \rho) = 0$ ,  $\Omega(\sigma_o, \rho) = \emptyset$  for all  $\rho \in \xi(\sigma_o)$
- 7:   **end if**
- 8: **end for**
- 9: **while** there is a path  $\mathcal{P}$  that is not blocked and a ray  $\rho \in \xi(\text{tail}_\sigma(\mathcal{P}))$  such that  
 $\text{tail}_\delta(\mathcal{P}) = \delta(\text{tail}_\sigma(\mathcal{P}), \rho) + 1$  and  $\delta(\text{tail}_\sigma(\mathcal{P})) \leq \delta(\omega)$  for all  $\omega \in \mathcal{N}$  **do**
- 10:   Choose a star-type  $\sigma' \in \Sigma$  such that there is a ray  $\rho' \in \xi(\sigma')$  satisfying  
 $l(\rho) = \lambda(\sigma')$ ,  $l(\rho') = \lambda(\sigma)$ ,  $r(\rho') = r^-(\rho)$ , and  
 $\sigma' \in \mathcal{N}$  implies  $\delta(\sigma') = \delta(\sigma', \rho') + 1$
- 11:   updateFrame( $\sigma, \rho, \sigma', \rho'$ )
- 12: **end while**

**Algorithm 1:** An algorithm for building a frame

$\Omega(\sigma, \rho)$  can be built. For example,  $\Omega(\sigma_0, \rho_0) = \{(\sigma_1, \nu_0, 1)\}$ ,  $\Omega(\sigma_0, \rho_1) = \{(\sigma_2, \rho'_0, 1)\}$  where  $\rho_0$  and  $\rho_1$  are the respective horizontal and vertical rays of  $\sigma_0$ ;  $\nu_0$  is the left ray of  $\sigma_1$ ;  $\rho'_0$  is the vertical ray of  $\sigma_2$ . Moreover, the directed dashed arrow from  $\sigma_0$  to  $\sigma_1$  indicates that the ray  $\rho_0$  of  $\sigma_0$  can match the ray  $\nu_0$  on the left ray of  $\sigma_1$  since  $l(\rho_0) = \lambda(\sigma_1)$ ,  $r(\nu_0) = \lambda(\sigma_0)$ ,  $r(\nu_0) = r^-(\rho_0)$ .

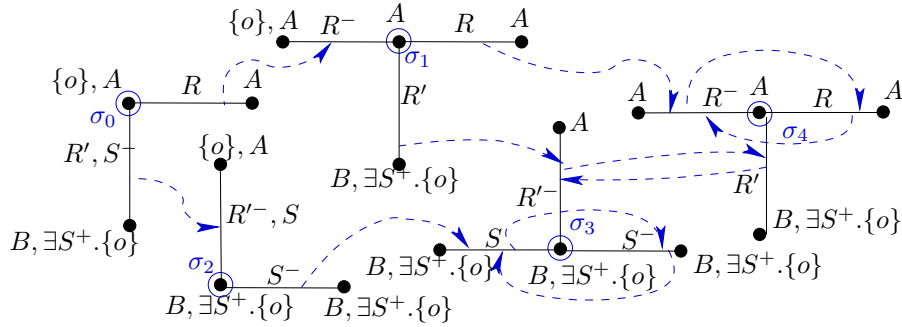
Then, the algorithm generates  $\delta(\sigma_0) = 1$ ,  $\delta(\sigma_1) = 1$ ,  $\delta(\sigma_2) = 1$  and forms a cycle  $\Theta$  consisting of the following triples :  $((\sigma_3, 2), \rho_1, (\sigma_3, 3))$  ( $\rho_1$  is the left ray of  $\sigma_3$ ),  $((\sigma_3, 2), \rho_3, (\sigma_4, 1))$  ( $\rho_3$  is the vertical ray of  $\sigma_3$ ),  $((\sigma_4, 1), \rho_4, (\sigma_4, 2))$  ( $\rho_4$  is the left ray of  $\sigma_4$ ) and  $((\sigma_4, 1), \rho_5, (\sigma_3, 2))$  ( $\rho_5$  is the vertical ray of  $\sigma_4$ ). Note that for the sake of brevity, we use just  $\text{tail}_\sigma(\mathcal{P})$  and  $\text{tail}_\delta(\mathcal{P})$  to denote a path in the triples. We can check that any path that is an extension of a path  $\mathcal{P}$  gets through a path  $\mathcal{Q}$  where  $\mathcal{P}$  is the first component of a triple and  $\mathcal{Q}$  is the third component of a triple.

The algorithm may add some more paths that go through  $\sigma_3$  and  $\sigma_4$  to form a blocked cycle. A model of the ontology can be built by starting from  $\sigma_0$  and getting (i)  $\sigma_4$  via  $\sigma_1$ , (ii)  $\sigma_3$  via  $\sigma_1$ , and (iii)  $\sigma_3$  via  $\sigma_2$ . From  $\sigma_3$  and  $\sigma_4$ , the model goes through  $\sigma_3$  and  $\sigma_4$  infinitely. Note that from any individual  $x$  satisfying  $\sigma_3$  (or  $\sigma_4$ ), i.e. the “label” of  $x$  contains  $\exists Q^+.\{o\}$ , there is a path containing  $S$  which goes back the individual satisfying  $\sigma_0$ . Thus, the concept  $\exists Q^+.\{o\}$  is satisfied for each individual whose label contains  $\exists Q^+.\{o\}$ .

**Lemma 1.** Let  $(\mathcal{T}, \mathcal{R})$  be a  $\mathcal{SHOIQ}_{(+)}$  knowledge base.

1. Algorithm 1 terminates.
2. If Algorithm 1 can build a valid frame for  $(\mathcal{T}, \mathcal{R})$  then  $(\mathcal{T}, \mathcal{R})$  is consistent.
3. If  $(\mathcal{T}, \mathcal{R})$  is consistent then Algorithm 1 can build a valid frame  $\mathcal{F}$  for  $(\mathcal{T}, \mathcal{R})$ .

*Proof (sketch).* Since the functions  $\delta(\sigma)$  and  $\delta(\sigma, \langle r, l \rangle)$  is increased monotonously by Algorithm 1, termination of the algorithm can be proved if we can show that : (i) the



**Fig. 2.** A frame obtained by Algorithm 1 for  $\mathcal{K}_1$  in the example in Section 1

number of different star-types is bounded; (ii) the detection of a blocked cycle according to Definition 8 terminates. For the soundness of Algorithm 1, we can extend the set  $\mathcal{P}$  of paths to a set  $\widehat{\mathcal{P}}$  of extended paths by “unravelling” the frame between blocking-blocked cycles. A tableau [14] can be built from  $\widehat{\mathcal{P}}$ . The main argument is that when extending a path  $\mathcal{P}$  if  $\text{tail}_\sigma(\mathcal{P}) \neq \text{tail}_\sigma(\mathcal{P}')$  for all blocked path  $\mathcal{P}'$  then this extension process can be continued up to a star-type  $\sigma' = \text{tail}_\sigma(\mathcal{P}'')$  for some blocked path  $\mathcal{P}''$ . This holds due to the definition of cycles and blockable cycles. Otherwise, i.e.,  $\text{tail}_\sigma(\mathcal{P}) = \text{tail}_\sigma(\mathcal{P}')$  for some blocked path  $\mathcal{P}'$  by  $Q$  then  $\mathcal{P}$  can be extended by getting through  $\text{tail}_\sigma(Q)$ .

Regarding completeness, a tableau can guide the algorithm (i) to choose valid star-types, (ii) to ensure that  $\delta(\sigma) = 1$  for each nominal star-type  $\sigma$ , and (iii) to detect a pair  $(\Theta_1, \Theta_k)$  of blocking and blocked cycles as soon as each concept of the form  $\exists Q^\oplus.D$  in  $\Theta_1$  is satisfied. We refer the readers to [14] for a complete proof of Lemma 1.

The following theorem is a consequence of Lemma 1.

**Theorem 1.**  $SHOIQ_{(+)}$  is decidable.

## 4 Conclusion

In this paper, we have presented a decision procedure for the description logic  $SHOIQ$  with transitive closure of roles in concept axioms, whose decidability was not known. The most significant feature of our contribution is to introduce a structure for characterizing models which have an infinite non-tree-like part. This structure would provide an insight into regularity of such models which would be enjoyed by a more expressive DL, such as  $ZOIQ$  [6], whose decidability remains open. In future work, we aim to improve the algorithm by making it more goal-directed and aim to investigate another open question about the hardness of  $SHOIQ_{(+)}$ .

## References

1. Patel-Schneider, P., Hayes, P., Horrocks, I.: OWL web ontology language semantics and abstract syntax. In: W3C Recommendation. (2004)
2. Tobies, S.: The complexity of reasoning with cardinality restrictions and nominals in expressive description logics. *Journal of Artificial Intelligence Research* **12** (2000) 199–217
3. Aho, A.V., Ullman, J.D.: Universality of data retrieval languages. In: Proceedings of the 6th of ACM Symposium on Principles of Programming Language. (1979)
4. Baader, F.: Augmenting concept languages by transitive closure of roles: An alternative to terminological cycles. In: Proceedings of the Twelfth International Joint Conference on Artificial Intelligence. (1991)
5. Ortiz, M.: An automata-based algorithm for description logics around *SRIQ*. In: Proceedings of the fourth Latin American Workshop on Non-Monotonic Reasoning 2008, CEUR-WS.org (2008)
6. Calvanese, D., Eiter, T., Ortiz, M.: Regular path queries in expressive description logics with nominals. In: IJCAI. (2009) 714–720
7. Horrocks, I., Sattler, U.: A tableau decision procedure for *SHOIQ*. *Journal Of Automated Reasoning* **39**(3) (2007) 249–276
8. Motik, B., Shearer, R., Horrocks, I.: Hypertableau reasoning for description logics. *J. of Artificial Intelligence Research* **36** (2009) 165–228
9. Pratt-Hartmann, I.: Complexity of the two-variable fragment with counting quantifiers. *Journal of Logic, Language and Information* **14**(3) (2005) 369–395
10. Le Duc, C., Lamolle, M.: Decidability of description logics with transitive closure of roles. In: Proceedings of the 23rd International Workshop on Description Logics (DL 2010), CEUR-WS.org (2010)
11. Horrocks, I., Sattler, U., Tobies, S.: Practical reasoning for expressive description logics. In: Proceedings of the International Conference on Logic for Programming, Artificial Intelligence and Reasoning (LPAR 1999), Springer (1999)
12. Baader, F., Nutt, W.: Basic description logics. In: *The Description Logic Handbook: Theory, Implementation and Applications* (2nd edition), Cambridge University Press (2007) 47–104
13. Fischer, M.J., Ladner, R.I.: Propositional dynamic logic of regular programs. *Journal of Computer and System Sciences* **18**(18) (1979) 174–211
14. Le Duc, C., Lamolle, M., Curé, O.: A decision procedure for *SHOIQ* with transitive closure of roles in concept axioms. In: Technical Report, <http://www.iut.univ-paris8.fr/~leduc/papers/RR-SHOIQTr.pdf> (2013)