# Augmenting Subsumption Propagation in Distributed Description Logics⋆

Martin Homola[1,2] and Luciano Serafini[2]

[1] Comenius University, Faculty of Mathematics, Physics and Informatics
Mlynská dolina, 84248 Bratislava, Slovakia
`homola@fmph.uniba.sk`
[2] Fondazione Bruno Kessler
Via Sommarive 18, 38050 Povo, Trento, Italy
`serafini@fbk.eu`

**Abstract.** Distributed Description Logics (DDL) enable reasoning with multiple ontologies interconnected by directional semantic mapping, called bridge rules. Bridge rules map concepts of a source ontology into concepts of a target ontology. Concept subsumptions of the source ontology can be propagated according to a propagation pattern expressed by bridge rules into concept subsumptions of the target ontology. In the basic formulation of DDL such a propagation is mostly limited to cases when pairs of ontologies are *directly linked* by means of bridge rules. However, when more than two ontologies are involved, one would expect that subsumption propagates along chains of ontologies linked by bridge rules, but the semantics of DDL is too weak to support this behaviour. In a recent study, an adjusted semantics for DDL that supports subsumption propagation through chains of ontologies has been introduced. This study makes use of so called compositional consistency requirement that has been employed before in Package-based description logics. While the results concerning subsumption propagation under the adjusted semantics are encouraging, there are important drawbacks. In this paper we take a wider perspective, and propose a study of several different alternative extensions of the DDL semantics. For each of them we study the formal properties, and we select the one that, according to our analysis, constitutes a good compromise, and for this case we provide a sound and complete tableaux decision procedure.

## 1 Introduction

Distributed Description Logics (DDL) have been introduced by Borgida & Serafini (2003) and consecutively developed by further research (Serafini et al. 2005, Serafini & Tamilin 2006, Ghidini et al. 2007). DDL are intended especially to

---

enable reasoning over systems of multiple ontologies connected by directional semantic mapping, built upon the formal, logical and well established framework of Description Logics (DL) (Baader et al. 2003) and local models semantics (Ghidini & Giunchiglia 2001). DDL capture the idea of importing and reusing concepts between several heterogeneous ontologies. This idea combines well with the basic assumption of the Semantic Web that no central ontology but rather many ontologies with redundant knowledge will exist (Berners-Lee et al. 2001).

In DL, a knowledge base, also called an ontology, contains a set of terminological axioms that assert relation upon concepts. For instance, with *subsumption axioms* we assert that some particular concept is a subconcept of another concept. While the ontology contains a set of axioms that record selected relations between concepts directly, using the logical semantics of DL some more relations are possibly implied. These relations may then be *queried* by the user. Say, if we assert directly that $C$ is a subconcept of $D$ and $D$ in turn is a subconcept of $E$, this implies that also $C$ is a subconcept of $E$. This relation is called subsumption: if $C$ is a subconcept of $D$, we say that $C$ is subsumed by $D$.

DDL enriches DL with new kind of axioms, called bridge rules. Bridge rules represent a semantic mapping between ontologies in a formal fashion. With bridge rules one is able to assert that some concept $C$, local to ontology $\mathcal{T}_1$, is related to some other concept $D$, local to $\mathcal{T}_2$.[3] While technically this relation is not exactly equivalent to the subsumption present inside the local ontologies, with bridge rules we either indicate a subconcept or a superconcept relation. Moreover, bridge rules are directed, and hence if there is a bridge rule with direction from $\mathcal{T}_1$ to $\mathcal{T}_2$, then $\mathcal{T}_2$ reuses knowledge from $\mathcal{T}_1$ but not necessarily the other way around. Semantic mapping encoded in bridge rules enables for knowledge reuse between local ontologies in a DDL knowledge base via the effect called subsumption propagation.

Consider the example depicted in Fig. 1. Let the ontology on the right be our local backyard ontology. In this ontology we maintain knowledge about all animals in our backyard. Instead of complex modeling of all relations between our animals we reuse a far more complex ontology that contains classification of all animal species. We map MyCat to genus Felis, and we also create the concept DangerousAnimal and map it to Felidae, the family of all cats, since we also keep a hamster and we know that all cats hunt hamsters for food. Thanks to the semantics of DDL we are now able to derive that MyCat is a subconcept of DangerousAnimal, even if this relation is not derived in the original backyard ontology locally. We say that the subsumption between Felis and Felidae has propagated to the backyard ontology thanks to the mapping. Please note also that the direction of the mapping is from the species ontology to the backyard ontology, as we only intend to reuse the knowledge of the species ontology within

---

[3] Even if the DDL framework has been extended for representing mappings between roles and also so called heterogeneous mappings (i.e., mappings between concepts and roles) (Ghidini et al. 2007), in this paper we restrict to the case of mapping between concepts.
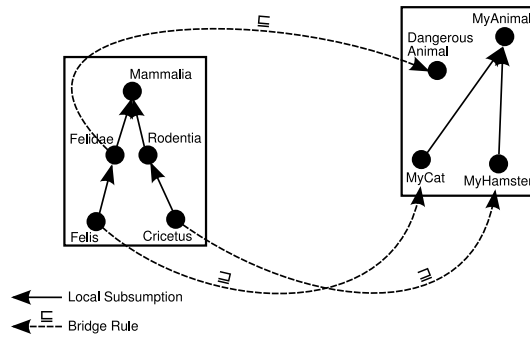
**Fig. 1.** A simple DDL knowledge base with bridge rules. In this knowledge base it always holds that MyCat is a subconcept of DangerousAnimal, even if this fact is not explicitly recorded.

the backyard ontology and not the other way around. The mapping depicted on Fig. 1 does not affect the knowledge within the species ontology in any way.

Subsumption propagation, as seen above, has been described as desired and one of the main features of DDL and it has been studied (Borgida & Serafini 2003, Serafini & Tamilin 2004, Homola 2007). Homola (2008) argues that the original semantics of DDL behaves counterintuitively in certain cases and proposes an adjustment to the semantics in order to cope with this issue.
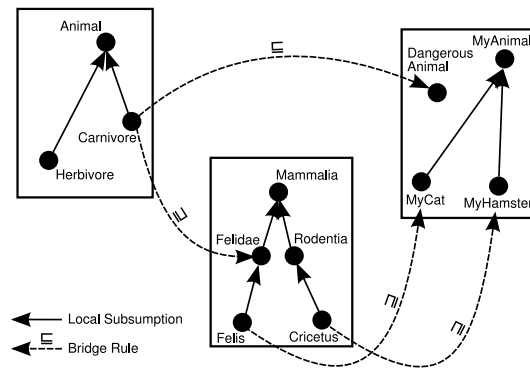


**Fig. 2.** An example of complex concept mapping between three ontologies. In this case the query whether MyCat is a subconcept of DangerousAnimal is not true under the original DDL semantics.

For illustration consider Fig. 2 which extends the previously discussed example depicted in Fig. 1. In this case, the situation is more complex. We no longer

map between concepts Felidae and DangerousAnimal, but instead these two concepts are connected indirectly via the concept Carnivore from yet another ontology. This third ontology deals with animal behaviour in contrast with the species ontology that merely provides classification. In case that there are such ontologies available, we suggest that this way of modelling is even more natural. As in the above example, again we would expect to derive that MyCat is a subconcept of DangerousAnimal within the backyard ontology. This relation is not derived under the original DDL semantics however. In the example from Fig. 2 there is a so called chain of bridge rules. The onto-bridge rule between concepts Carnivore and Felidae and the other one between Felis and MyCat form a chain, as Felis is a subconcept of Felidae in the classification ontology. Similar chains are also possible with into-bridge rules. The adjusted semantics of Homola (2008) allows subsumption propagation along such chains of bridge rules and hence the subsumption between MyCat and DangerousAnimal is entailed under this semantics.

In this paper we take steps forward in order to develop a distributed tableaux reasoning algorithm that would decide satisfiability of concepts and subsumption with respect to the adjusted semantics. The algorithm that is introduced in this paper handles chaining onto-bridge rules correctly, but it is unable to cope with chaining into-bridge rules. We provide precise characterization of the semantics that the algorithm actually implements. The algorithm works with acyclic DDL knowledge bases build on top of $\mathcal{ALC}$ as the local language. As in the case of the tableaux algorithm that is known for the original semantics (Serafini et al. 2005), the algorithm is truly distributed and it permits the scenario where every local ontology is governed by an autonomous reasoning service and these services communicate by passing queries. The local reasoner that has started the computation collects all the answers and makes the decision at the end. Besides the fact that each algorithm works under a different semantics, the main distinguishing feature of the newly introduced algorithm is that communication between local reasoners is divided into multiple messages. We believe that such behaviour may serve as a base for more fine-grained optimization in the future.

## 2   Distributed Description Logics

DDL rely on Description Logics (Baader et al. 2003) as for the representation language of the local knowledge bases. DL provide a well established ontological representation language with formal semantics and with reasoning tasks and associated computational issues well understood. In the following we briefly introduce the $\mathcal{ALC}$ DL in order to build on top of it later in the paper.

A DL knowledge base typically consists of a TBox and an ABox. The TBox contains terminological knowledge, that is, definitions of concepts and their relations (i.e., subsumption axioms). The ABox contains knowledge about individuals, their relations and their membership in concepts. In this work we concentrate above all on the terminological part of the knowledge and hence we will only work with TBoxes below. We will denote TBoxes with $\mathcal{T}_i$ because we deal with several ontologies. Assume that $i$ is an index from some finite index set $I$. Each

TBox is a set of subsumption axioms called general concept inclusions (GCI), each of the form $i : C \sqsubseteq D$, where $C$ and $D$ are concepts. Concepts are either atomic or complex. Atomic concepts have no further structure. Complex concepts are composed from atomic concepts and roles using some of the DL concept constructors. For list of constructors and their meaning see Fig. 3.

The semantics of DL is model-theoretic. Each ontology $\mathcal{T}_i$ is assigned an interpretation $\mathcal{I}_i$ consisting of a domain $\Delta^{\mathcal{I}_i}$ and an interpretation function $\cdot^{\mathcal{I}_i}$ which interprets each concept $C$ by $C^{\mathcal{I}_i} \subseteq \Delta^{\mathcal{I}_i}$. Each role $R$ is interpreted by $R^{\mathcal{I}_i} \subseteq \Delta^{\mathcal{I}_i} \times \Delta^{\mathcal{I}_i}$. Complex concepts must satisfy further semantic constraints, see Fig. 3. There are also two special concepts $\bot$ and $\top$. The bottom concept ($\bot$) is always interpreted by $\bot^{\mathcal{I}_i} = \emptyset$, the top concept ($\top$) is always interpreted by $\top^{\mathcal{I}_i} = \Delta^{\mathcal{I}_i}$, but they are formally defined as syntactic sugar: $\bot \equiv E \sqcap \neg E$ and $\top \equiv E \sqcup \neg E$, for some new concept name $E$.

A GCI axiom $i : C \sqsubseteq D$ , is satisfied by $\mathcal{I}_i$ whenever $C^{\mathcal{I}_i} \subseteq D^{\mathcal{I}_i}$. An interpretation $\mathcal{I}_i$ is a model of $\mathcal{T}_i$ if it satisfies all GCI in $\mathcal{T}_i$. A concept $C$ is satisfiable, if there exists an interpretation $\mathcal{I}_i$ such that $C^{\mathcal{I}_i}$ is nonempty. In addition, $C$ is satisfiable with respect to a TBox $\mathcal{T}_i$ if there exists an interpretation $\mathcal{I}_i$ with $C^{\mathcal{I}_i} \neq \emptyset$ that is a model of $\mathcal{T}_i$. And, subsumption $i : C \sqsubseteq D$ is entailed by $\mathcal{T}_i$ if $C^{\mathcal{I}_i} \subseteq D^{\mathcal{I}_i}$ holds in all models of $\mathcal{T}_i$. Satisfiability of concepts and subsumption entailment are standard decision problems that are investigated in each DL.

Two concepts $C$ and $D$ are equivalent if in each interpretation $\mathcal{I}_i$ we have $C^{\mathcal{I}_i} = D^{\mathcal{I}_i}$. A concept $C$ is in negation normal form (NNF) if negation ($\neg$) only appears in $C$ directly in front of atomic concepts. For each concept, an equivalent concept in NNF exists (Baader et al. 2003), we denote the NNF of $C$ by $\mathrm{nnf}(C)$. For more detailed introduction to DL please refer to the work of Baader et al. (2003) and Horrocks et al. (1999).

| $E$ | $E^{\mathcal{I}_i}$ |
|---|---|
| $\neg C$ | $\Delta^{\mathcal{I}_i} \setminus C^{\mathcal{I}_i}$ |
| $C \sqcap D$ | $C^{\mathcal{I}_i} \cap D^{\mathcal{I}_i}$ |
| $C \sqcup D$ | $C^{\mathcal{I}_i} \cup D^{\mathcal{I}_i}$ |
| $\forall R.C$ | $\{i \in \Delta^{\mathcal{I}_i} \mid (\forall j \in \Delta^{\mathcal{I}_i})\ (i,j) \in R^{\mathcal{I}_i} \implies j \in C^{\mathcal{I}_i}\}$ |
| $\exists R.C$ | $\{i \in \Delta^{\mathcal{I}_i} \mid (\exists j \in \Delta^{\mathcal{I}_i})\ (i,j) \in R^{\mathcal{I}_i} \wedge j \in C^{\mathcal{I}_i}\}$ |

**Fig. 3.** Complex concepts in $\mathcal{ALC}$ and their semantics.

A DDL knowledge base, commonly dubbed distributed TBox $\mathfrak{T}$, includes a set of local TBoxes each over its own DL language and a set of bridge rules $\mathfrak{B}$ that provides mappings between these local TBoxes. The local languages are commonly required to be under $\mathcal{SHIQ}$ (Horrocks et al. 1999) as it is not trivial to extend DDL with nominals (Serafini et al. 2005). In this work we will stay under $\mathcal{ALC}$, which is a sub-language of $\mathcal{SHIQ}$. The formal definition is as follows.

**Definition 1.** *Assume a non-empty index set $I$, a set of concept names $\mathrm{N_C} = \bigcup_{i \in I} \mathrm{N}_{Ci}$ and a set of role names $\mathrm{N_R} = \bigcup_{i \in I} \mathrm{N}_{Ri}$. A Distributed TBox is a pair $\mathfrak{T} = \langle \{\mathcal{T}_i\}_{i \in I}, \mathfrak{B} \rangle$ such that:*

1. *Each local TBox $\mathcal{T}_i$ is a collection of $\mathcal{ALC}$ GCI axioms over $\mathrm{N}_{\mathrm{C}i}$ and $\mathrm{N}_{\mathrm{R}i}$. Each GCI is of the form:*

$$i : C \sqsubseteq D \ .$$

2. *The set of bridge rules $\mathfrak{B}$ divides into sets of bridge rules $\mathfrak{B} = \bigcup_{i,j \in I, i \neq j} \mathfrak{B}_{ij}$. Each $\mathfrak{B}_{ij}$ is a collection of bridge rules in direction from $\mathcal{T}_i$ to $\mathcal{T}_j$ which are of two forms,* into-*bridge rules and* onto-*bridge rules (in the respective order):*

$$i : A \overset{\sqsubseteq}{\to} j : G \ , \qquad\qquad i : B \overset{\sqsupseteq}{\to} j : H \ .$$

While, as we will see bellow, bridge rules are not semantically equivalent to subsumption axioms (GCI), they also work with the notions of subconcept and superconcept. Intuitively speaking, into-bridge rules cause the mapped image of the source concept to always stay inside of the target concept, so hence their name. Conversely, onto-bridge rules cause that the target concept is always inside of the mapped image of the source concept. As already mentioned, the direction of bridge rules matters and hence $\mathfrak{B}_{ij}$ and $\mathfrak{B}_{ji}$ are possibly and expectedly distinct. The bridge graph $\mathrm{G}_{\mathfrak{T}} = \langle V, E \rangle$ of a distributed TBox $\mathfrak{T}$ is defined as follows: $V = I$ and $\langle i, j \rangle \in E$ if $\mathfrak{B}_{ij} \neq \emptyset$. We say that $\mathfrak{T}$ is acyclic if $\mathrm{G}_{\mathfrak{T}}$ is acyclic.

Given a TBox $\mathcal{T}$, a hole is an interpretation $\mathcal{I}^\epsilon = \langle \emptyset, \cdot^\epsilon \rangle$ with empty domain. Holes are used for fighting propagation of inconsistency. We use the most recent definition, introduced by Serafini et al. (2005). A distributed interpretation $\mathfrak{I} = \langle \{\mathcal{I}_i\}_{i \in I}, \{r_{ij}\}_{i \in I, i \neq j} \rangle$ of a distributed TBox $\mathfrak{T}$ consists of a set of local interpretations $\{\mathcal{I}_i\}_{i \in I}$ and a set of domain relations $\{r_{ij}\}_{i \in I, i \neq j}$. For each $i \in I$, either $\mathcal{I}_i = (\Delta^{\mathcal{I}_i}, \cdot^{\mathcal{I}_i})$ is an interpretation of the local TBox $\mathcal{T}_i$ or $\mathcal{I}_i = \mathcal{I}^\epsilon$ is a hole. Each domain relation $r_{ij}$ is a subset of $\Delta^{\mathcal{I}_i} \times \Delta^{\mathcal{I}_j}$. We denote by $r_{ij}(d)$ the set $\{d' \mid \langle d, d' \rangle \in r_{ij}\}$ and by $r_{ij}(D)$ the set $\bigcup_{d \in D} r_{ij}(d)$.

**Definition 2.** *For every $i$ and $j$, a distributed interpretation $\mathfrak{I}$ satisfies the elements of a distributed TBox $\mathfrak{T}$ (denoted by $\mathfrak{I} \models_\epsilon \cdot$) according to the following clauses:*

1. *$\mathfrak{I} \models_\epsilon i : C \sqsubseteq D$ if $\mathcal{I}_i \models C \sqsubseteq D$;*
2. *$\mathfrak{I} \models_\epsilon \mathcal{T}_i$ if $\mathfrak{I} \models_\epsilon i : C \sqsubseteq D$ for each $C \sqsubseteq D \in \mathcal{T}_i$;*
3. *$\mathfrak{I} \models_\epsilon i : C \overset{\sqsubseteq}{\to} j : G$ if $r_{ij}(C^{\mathcal{I}_i}) \subseteq G^{\mathcal{I}_j}$;*
4. *$\mathfrak{I} \models_\epsilon i : C \overset{\sqsupseteq}{\to} j : G$ if $r_{ij}(C^{\mathcal{I}_i}) \supseteq G^{\mathcal{I}_j}$;*
5. *$\mathfrak{I} \models_\epsilon \mathfrak{B}$ if $\mathfrak{I}$ satisfies all bridge rules in $\mathfrak{B}$;*
6. *$\mathfrak{I} \models_\epsilon \mathfrak{T}$ if $\mathfrak{I} \models_\epsilon \mathfrak{B}$ and $\mathfrak{I} \models_\epsilon \mathcal{T}_i$ for each $i$.*

*If $\mathfrak{I} \models_\epsilon \mathfrak{T}$ then we also say that $\mathfrak{I}$ is a (distributed) model of $\mathfrak{T}$.*

The two standard decision problems in DL, satisfiability of concepts and entailment of subsumption, play prominent rôle also in the context of DDL. Formally, the decision problems are defined as follows.

**Definition 3.** *Given a distributed TBox $\mathfrak{T}$, an $i$-local concept $C$ is satisfiable with respect to $\mathfrak{T}$ if there exists a distributed model $\mathfrak{I}$ of $\mathfrak{T}$ such that $C^{\mathcal{I}_i} \neq \emptyset$.*

**Definition 4.** *Given a distributed TBox $\mathfrak{T}$ and two $i$-local concepts $C$ and $D$, it is said that $C$ is subsumed by $D$ with respect to $\mathfrak{T}$ if $C^{\mathcal{I}_i} \subseteq D^{\mathcal{I}_i}$ in every distributed model $\mathfrak{I}$ of $\mathfrak{T}$. We also sometimes say that the subsumption formula $i : C \sqsubseteq D$ is entailed by $\mathfrak{T}$ and denote this by $\mathfrak{T} \models_\epsilon i : C \sqsubseteq D$.*

It is a well known result that in most DL subsumption and (un)satisfiability are inter-reducible (Baader et al. 2003). It follows rather straight forward, that this reduction is also valid for DDL, as follows in Theorem 1.

**Theorem 1.** *Assume a distributed TBox $\mathfrak{T}$ and two $i$-local concepts $C$ and $D$. $\mathfrak{T} \models_\epsilon i : C \sqsubseteq D$ if and only if the concept $\neg C \sqcap D$ is unsatisfiable with respect to $\mathfrak{T}$. Also, $C$ is satisfiable with respect to $\mathfrak{T}$ if and only if the subsumption formula $i : C \sqsubseteq \bot$ is not entailed by $\mathfrak{T}$.*

In the literature several properties have been identified as desired for a logic of modular and distributed ontologies (Loebe 2006). In context of DDL, particularly three properties are considered as desiderata. These are (intuitively explained):

**monotonicity:** addition of new knowledge does not affect the previously entailed knowledge;

**directionality:** some local ontology $\mathcal{T}_i$ possibly affects another local ontology $\mathcal{T}_j$, only if they are connected by a directed path of bridge rules starting from $T_i$ and eventually reaching $T_j$;

**local inconsistency:** if one local ontology is inconsistent this does not necessarily cause the inconsistence of the whole distributed ontology, instead the local inconsistency should be dealt with allowing only minimal necessary damage.

These properties have been formalized in the context of DDL in the literature and it was proved that they hold for the original DDL semantics (Borgida & Serafini 2003, Serafini & Tamilin 2004, Serafini et al. 2005). In any extension of the DDL semantic it is desirable to maintain these three properties, if possible. The properties are formalized as follows.

**Definition 5 (Monotonicity).** *The monotonicity property is satisfied whenever in every distributed TBox $\mathfrak{T}$ it holds that*

$$\mathcal{T}_i \models C \sqsubseteq D \implies \mathfrak{T} \models_\epsilon i : C \sqsubseteq D \ .$$

Directionality states that there should also be no back-flow of information against the direction of bridge rules. That is, if $\mathcal{T}_i$ and $\mathcal{T}_j$ are not connected by any directed path of bridge rules (in direction from $i$ to $j$) then $\mathcal{T}_i$ does not affect $\mathcal{T}_j$; which means that removal of $\mathcal{T}_i$ should have no effect in $\mathcal{T}_j$.

**Definition 6 (Directionality).** *The directionality property is satisfied whenever in every distributed TBox $\mathfrak{T}$ with index set $I$ and bridge graph $G_\mathfrak{T}$ for every $i, j \in I$, $i \neq j$ it holds that if there is no directed path between $i$ to $j$ in $G_\mathfrak{T}$, then*

$$\mathfrak{T} \models_\epsilon j : C \sqsubseteq D \iff \mathfrak{T}^{-i} \models_\epsilon j : C \sqsubseteq D$$

*where $\mathfrak{T}^{-i}$ is obtained by removing $\mathcal{T}_i$, $\mathfrak{B}_{ik}$ and $\mathfrak{B}_{li}$ from $\mathfrak{T}$ for each $k, l \in I$.*

The local inconsistency property states that even if some of the local TBoxes happens to be inconsistent, the inconsistency does not necessarily spread and pollute the whole distributed ontology. There are more options how to formalize such a property and several formalizations occur in the literature (Borgida & Serafini 2003, Serafini & Tamilin 2004, Serafini et al. 2005). We use the most recent definition (Serafini et al. 2005), which is rather technical but in fact describes very well the exact extent of inconsistency propagation that occurs under the original DDL semantics.

The limited amount of inconsistency propagation that occurs in DDL is justified by the following observation. If some local TBox $\mathcal{T}_i$ is inconsistent, then the only admissible local interpretation is a hole. This affects onto-bridge rules outgoing from $\mathcal{T}_i$ – the only admissible local interpretation for every $j$-concept $D$ that is a target of an onto-bridge rule outgoing from $\mathcal{T}_i$ is the empty set (i.e., $D^{\mathcal{I}_j} = \emptyset$ in each distributed model $\mathfrak{I}$ of $\mathfrak{T}$).

We will use some auxiliary notation. Let $\models_d$ be a kind of entailment just as $\models_\epsilon$ only that it does not allow holes as local interpretations and let also each $\mathfrak{I} \models_d \mathfrak{T}$ be called a d-model of $\mathfrak{T}$. Given $J \subseteq I$, $\mathfrak{T}(\epsilon_J)$ is obtained from $\mathfrak{T}$ by removing each $\mathcal{T}_j$ such that $j \in J$, and adding $\{D \sqsubseteq \bot \mid j : C \overset{\sqsupseteq}{\Rightarrow} i : D \in \mathfrak{B} \wedge j \in J\}$ to each $\mathcal{T}_i$, $i \in I \setminus J$.

**Definition 7 (Local inconsistency).** *The local inconsistency property is satisfied whenever in each distributed TBox $\mathfrak{T}$ the following holds: $\mathfrak{T} \models_\epsilon i : C \sqsubseteq D$ if and only if for any $J \subseteq I$, not containing $i$, $\mathfrak{T}(\epsilon_J) \models_d i : C \sqsubseteq D$.*

The mechanism of knowledge reuse within DDL ontologies is formally characterized by the notion called *subsumption propagation*. Intuitively, subsumption propagates from one local ontology where it follows from the local knowledge to some another ontology that is connected by bridge rules. For an example, recall the distributed ontology depicted in Fig. 1. Here, the biological classification ontology on the left locally entails Felis $\sqsubseteq$ Felidae. Thanks a pair of bridge rules this subsumption is propagated into the backyard ontology on the right and MyCat $\sqsubseteq$ DangerousAnimal is distributively entailed.

This subsumption propagation scenario is only a very basic one. The *subsumption propagation property* has been formulated in a more general setting. One such a formulation appears in the work of Serafini & Tamilin (2004). Here subsumption is propagated between $G$ and the union $H_1 \sqcup \cdots \sqcup H_n$, given a locally entailed subsumption $A \sqsubseteq B_1 \sqcup \cdots \sqcup B_n$ in a neighboring local ontology and a sufficient number of bridge rules. This exact form of the property, formulated in form of an inference rule in Theorem 2, is somewhat important for the original DDL framework: it provides the grounds for the distributed reasoning algorithm introduced by Serafini & Tamilin (2004).

**Theorem 2 (Subsumption propagation).** *Given a distributed TBox $\mathfrak{T}$ with an index set $\mathcal{I}$, two local TBoxes $\mathcal{T}_i$ and $\mathcal{T}_j$, $i, j \in I$, $i$-local concepts $A, B_1, \ldots, B_n$ and $j$-local concepts $G, H_1, \ldots, H_n$, for some $n > 0$, the following inference rule*

*is valid:*

$$\mathfrak{T} \models_\epsilon i : A \sqsubseteq B_1 \sqcup \cdots \sqcup B_n$$
$$i : A \overset{\sqsupseteq}{\Rightarrow} j : G \in \mathfrak{B}_{ij}$$
$$\frac{i : B_k \overset{\sqsubseteq}{\Rightarrow} j : H_k \in \mathfrak{B}_{ij}, \, for \, 1 \leq k \leq n}{\mathfrak{T} \models_\epsilon j : G \sqsubseteq H_1 \sqcup \cdots \sqcup H_n}$$

Study of subsumption propagation in more complex cases also appears in the literature (Serafini et al. 2005, Homola 2007). These cases are not as interesting for this study, since they do not extend the case in that particular aspect that only two local ontologies which are directly connected with bridge rules are studied. In this work we concentrate on indirectly connected ontologies, a case with which, as we will see below, the original semantics is not completely able to cope with.

## 3  Improving Subsumption Propagation

As outlined in the introduction with the example of Figure 2, the subsumption propagation pattern in the original DDL is, in some cases, not as strong as one would desire. In a recent paper, Homola (2008) proposes a way to adjust the semantics of DDL in order to improve subsumption propagation in the cases when local ontologies are only connected indirectly. Here we provide a more systematic analysis of the possible improvements, and for one of them we extend the distributed tableaux algorithm for DDL described by Serafini et al. (2005).

This work is concerned with *remote ontologies*. The basic intuition behind this notion is of two local ontologies within a distributed system, that are not connected directly by a single bridge rule. Since we are interested in subsumption propagation between remote ontologies we further refine this notion considering the following two aspects. First, we are only interested in cases when these ontologies are in fact connected by some directed path of bridge rules; this is due to the directionality desideratum which allows no knowledge reuse if this is not the case. Second, even if two ontologies are connected directly, we must still consider also the remote connections between them, that is, paths of bridge rules spanning across multiple ontologies.

The characteristic pattern of interest, when dealing with remote ontologies is characterized by the notion of *chain of bridge rules*. Chain of bridge rules is a directed path of bridge rules which combine together with respect to subconcept-superconcept relations.

**Definition 8.** *Let $\mathfrak{T}$ be distributed TBox with an index set $I$. A chain of bridge rules is either a directed path of into-bridge rules*

$$\langle 1 : C_1 \overset{\sqsubseteq}{\Rightarrow} 2 : D_2, \, 2 : C_2 \overset{\sqsubseteq}{\Rightarrow} 3 : D_3, \, \ldots, \, n-1 : C_{n-1} \overset{\sqsubseteq}{\Rightarrow} n : D_n \rangle$$

*of length $n > 0$ such that for each $1 < i < n$ we have $\mathfrak{T} \models_\epsilon i : D_i \sqsubseteq C_i$ (into-chain); or a directed path of onto-bridge rules*

$$\langle 1 : C_1 \overset{\sqsupseteq}{\Rightarrow} 2 : D_2, \, 2 : C_2 \overset{\sqsupseteq}{\Rightarrow} 3 : D_3, \, \ldots, \, n-1 : C_{n-1} \overset{\sqsupseteq}{\Rightarrow} n : D_n \rangle$$

*of length $n > 0$ such that for each $1 < i < n$ we have $\mathfrak{T} \models_\epsilon i : C_i \sqsubseteq D_i$ (onto-chain).*

While there is no notion of bridge rules composition or bridge rules inference in the literature, and also in this work we do not intend to introduce this kind of reasoning with bridge rules formally, we will use the notion of bridge rules composition as a metaphor in what follows. The point is that the desired extent of subsumption propagation according to our intuitions is characterized by a hypothetical semantics in which each chain of bridge rules between $C$ of $\mathcal{T}_i$ and $D$ of $\mathcal{T}_j$ has the power equal to a single bridge rule between these two concepts. In the basic case this propagation pattern is characterized by the following composition of bridge rules:

$$i : A \overset{\sqsubseteq}{\Rightarrow} j : B \text{ and } j : B \overset{\sqsubseteq}{\Rightarrow} k : C \text{ implies } i : A \overset{\sqsubseteq}{\Rightarrow} k : C \tag{1}$$

$$i : A \overset{\sqsupseteq}{\Rightarrow} j : B \text{ and } j : B \overset{\sqsupseteq}{\Rightarrow} k : C \text{ implies } i : A \overset{\sqsupseteq}{\Rightarrow} k : C \tag{2}$$

$$\tag{3}$$

Semantically these two composition rules correspond to the following two conditions (in the respective order):

$$r_{ij}(A^{\mathcal{I}_i}) \subseteq B^{\mathcal{I}_j} \wedge r_{jk}(B^{\mathcal{I}_j}) \subseteq C^{\mathcal{I}_k} \implies r_{ik}(A^{\mathcal{I}_i}) \subseteq C^{\mathcal{I}_k} \tag{4}$$

$$r_{ij}(A^{\mathcal{I}_i}) \supseteq B^{\mathcal{I}_j} \wedge r_{jk}(B^{\mathcal{I}_j}) \supseteq C^{\mathcal{I}_k} \implies r_{ik}(A^{\mathcal{I}_i}) \supseteq C^{\mathcal{I}_k} \tag{5}$$

$$\tag{6}$$

Since the above two conditions must hold for any interpretation of $A$, $B$ and $C$, then they correspond to the following inclusions between domain relations:

$$r_{ij} \circ r_{jk} \supseteq r_{ik} \tag{7}$$

$$r_{ij} \circ r_{jk} \subseteq r_{ik} \tag{8}$$

Notice that when any two of $i$, $j$ and $k$ are equal, the above properties include $r_{ii}$ for some $i \in I$ which is not defined in a DDL. So we suppose that the above properties hold only in the case when $i, j$ and $k$ are three mutually distinct indices. Our investigation proceeds with studying the properties of an amended semantics of DDL which obeys restrictions (7) and (8).

**Definition 9.** *Given a distributed interpretation $\mathfrak{I}$ with domain relation $r$, we say that $r$ (and also $\mathfrak{I}$) satisfies compositional consistency if for any three mutually distinct indices $i, j, k \in I$ we have $r_{ij} \circ r_{jk} = r_{ik}$.*

Now the adjusted semantics is simply obtained from the original DDL semantics by allowing only distributed interpretations that satisfy compositional consistency. In accordance we often use the wording "DDL under compositional consistency" when referring to this semantics.

The adjusted semantics actually extends the original one, in the sense that if some subsumption formula $\Phi$ is entailed by a distributed TBox $\mathfrak{T}$ in the original

semantics, then it is also entailed by $\mathfrak{T}$ under compositional consistency. The only difference is that in the adjusted semantics possibly some more subsumption formulae are entailed in addition. This is formally stated in the following theorem (for a proof please refer to Homola (2008)).

**Theorem 3.** *Given a distributed TBox $\mathfrak{T}$ and a subsumption formula $\Phi$, if $\mathfrak{T} \models_\epsilon \Phi$ according to the original semantics, then $\mathfrak{T} \models_\epsilon \Phi$ also holds in DDL under compositional consistency.*

To demonstrate the mechanism of improved subsumption propagation within the adjusted semantics, let us revisit the example from Fig. 2 formally.
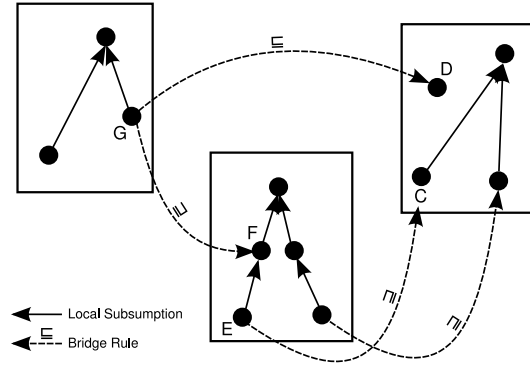


**Fig. 4.** Renaming of concepts from Fig. 2 employed in Example 1. The three local ontologies are referred to as $\mathcal{T}_b$ – behaviour ontology, $\mathcal{T}_c$ – classification ontology, and $\mathcal{T}_y$ – backyard ontology.

*Example 1.* Recall the example depicted in Fig. 2. Let us simplify the notation by renaming the concepts as follows: $C :=$ MyCat, $D :=$ DangerousAnimal, $E :=$ Felis, $F :=$ Felidae, $G :=$ Carnivore. Remaining concepts are of no interest (see Fig. 4). Assume the index set $I = \{b, c, y\}$, where $b$ represents the behaviour ontology (on the left), $c$ represents the classification ontology (in the middle) and $y$ represents the backyard ontology (on the right). There are various axioms in $\mathfrak{T} = \langle\{\mathcal{T}_b, \mathcal{T}_c, \mathcal{T}_y, \}, \mathfrak{B}\rangle$ but of the GCI only $c : E \sqsubseteq F$, actually matters to us, and there are three bridge rules in $\mathfrak{B}$:

$$b : G \overset{\sqsupseteq}{\rightarrow} c : F \ , \qquad\qquad c : E \overset{\sqsupseteq}{\rightarrow} y : C \ ,$$
$$b : G \overset{\sqsubseteq}{\rightarrow} y : D \ .$$

We query whether it holds that $\mathfrak{T} \models_\epsilon y : C \sqsubseteq D$. Assume a distributed interpretation $\mathfrak{I}$. Let us first assume that $\mathfrak{I}$ contains no hole. From Definition 2 we get

$r_{cy}(E^{\mathcal{I}_c}) \supseteq C^{\mathcal{I}_y}$. We also have $r_{bc}(G^{\mathcal{I}_b}) \supseteq F^{\mathcal{I}_c}$, but since $r_{cy}(\cdot)$ is a mapping we get that $r_{cy}(r_{bc}(G^{\mathcal{I}_b})) \supseteq r_{cy}(F^{\mathcal{I}_c})$. And from compositional consistency we get that $r_{cy}(r_{bc}(G^{\mathcal{I}_b})) = r_{by}(G^{\mathcal{I}_b})$ implies $r_{by}(G^{\mathcal{I}_b}) \supseteq r_{cy}(F^{\mathcal{I}_c})$. From the GCI $c : E \sqsubseteq F$ we get $F^{\mathcal{I}_c} \supseteq E^{\mathcal{I}_c}$ and from properties of mapping we again get $r_{cy}(F^{\mathcal{I}_c}) \supseteq r_{cy}(E^{\mathcal{I}_c})$. Putting this all together we derive:

$$r_{by}(G^{\mathcal{I}_b}) \supseteq r_{cy}(F^{\mathcal{I}_c}) \supseteq r_{cy}(E^{\mathcal{I}_c}) \supseteq C^{\mathcal{I}_y} \ ,$$

and that amounts to $r_{by}(G^{\mathcal{I}_b}) \supseteq C^{\mathcal{I}_y}$. On the other hand, from the into-bridge rule between $\mathcal{T}_b$ and $\mathcal{T}_y$ we derive $r_{by}(G^{\mathcal{I}_b}) \subseteq D^{\mathcal{I}_y}$. And so we finally get $C^{\mathcal{I}_y} \subseteq D^{\mathcal{I}_y}$.

For the case with holes assume for instance that $\mathcal{I}_b = \mathcal{I}^\epsilon$. In that case $G^{\mathcal{I}_b} = \emptyset$. Thanks to the constraint generated by bridge rules and the GCI we easily derive that also $F^{\mathcal{I}_c} = \emptyset$, $E^{\mathcal{I}_c} = \emptyset$, and also $C^{\mathcal{I}_y} = \emptyset$. In such a case, however, $C^{\mathcal{I}_y} \subseteq D^{\mathcal{I}_y}$ holds trivially. If we substitute other local interpretations for holes, we get the very same result analogously.

Summing up, in every model of $\mathfrak{T}$ we have $C^{\mathcal{I}_y} \subseteq D^{\mathcal{I}_y}$ and hence $\mathfrak{T} \models_\epsilon y : C \sqsubseteq D$. Recall that we have actually used the compositional consistency requirement in our argumentation. Without it we would not be able to establish the result: we would not be able to prove that $r_{by}(G^{\mathcal{I}_b}) \supseteq r_{cy}(F^{\mathcal{I}_c})$. In fact, the original DDL semantics allows models that violate this inclusion and hence $\mathfrak{T} \models_\epsilon y : C \sqsubseteq D$ does not hold under the original semantics.

Theorem 4 below provides a more general characterization of cases when subsumption propagates to remote ontologies. This characterization generalizes the setting from Figs. 2–4. For a proof, refer again to the work of Homola (2008).
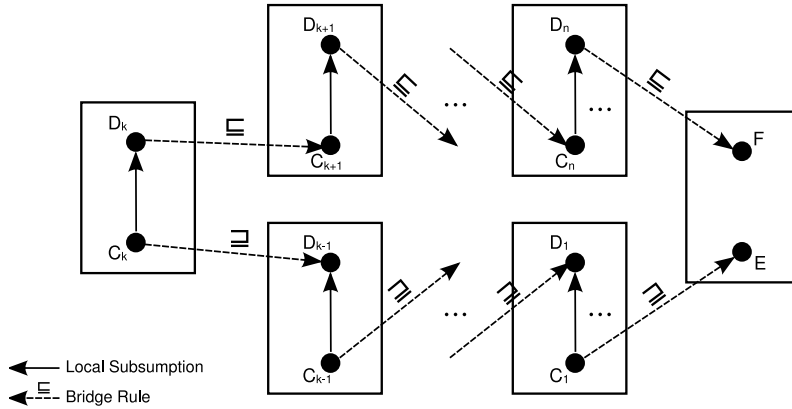


**Fig. 5.** Depiction of the distributed TBox from Theorem 4.

**Theorem 4.** *Given a distributed TBox, as illustrated in Fig. 5, with an index set $I$ and a set of bridge rules $\mathfrak{B}$, that features $n+1$ local TBoxes $\mathcal{T}_0, \mathcal{T}_1, \ldots, \mathcal{T}_n$ with concepts $E, F \in \mathcal{T}_0$, and $C_i, D_i \in \mathcal{T}_i$, for $1 \leq i \leq n$, and $k$ with $1 \leq k \leq n$ such that:*

*1. $\mathfrak{T} \models_\epsilon i : C_i \sqsubseteq D_i$, for $1 \leq i \leq n$,*
*2. $i+1 : C_{i+1} \stackrel{\sqsupseteq}{\Rightarrow} i : D_i \in \mathfrak{B}$, for $1 \leq i < k$,*
*3. $i : D_i \stackrel{\sqsubseteq}{\Rightarrow} i+1 : C_{i+1} \in \mathfrak{B}$, for $k \leq i < n$,*
*4. $1 : C_1 \stackrel{\sqsupseteq}{\Rightarrow} 0 : E \in \mathfrak{B}$ and $n : D_n \stackrel{\sqsubseteq}{\Rightarrow} 0 : F \in \mathfrak{B}$.*

*In DDL under compositional consistency it follows that $\mathfrak{T} \models_\epsilon 0 : E \sqsubseteq F$.*

In a nutshell, Theorem 4 basically says that the effect of bridge rules is now transitive, and hence subsumption propagates even between remote ontologies within the system. Since we deliberately did not formally introduce the notion bridge rules inference as it is not the primary subject of this paper, relying only on the current DDL notation this must be formulated in a rather complex fashion. If such a notion is introduced by future research, the theorem may be reformulated in a more elegant way.

As a counterpart of the positive effects of the newly introduced semantics, we also observe some undesired side effects. In particular, the full compositionality of domain relation spoils the directionality and the local inconsistency properties.
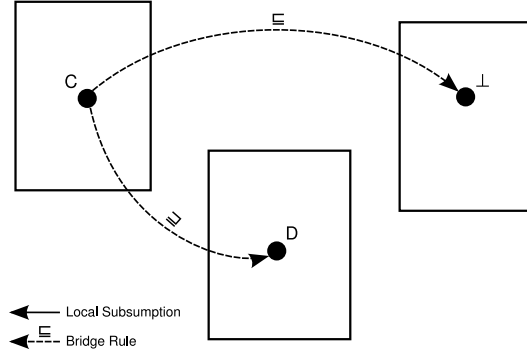


**Fig. 6.** Depiction of the distributed TBox from Example 2.

*Example 2.* Consider the following distributed TBox $\mathfrak{T}$ with the index set $I = \{1, 2, 3\}$, such that $\mathcal{T}_1 = \emptyset$, $\mathcal{T}_2 = \emptyset$, $\mathcal{T}_3 = \emptyset$ and two bridge rules $\mathfrak{B} = \{2 : C \stackrel{\sqsubseteq}{\Rightarrow} 1 : \bot, 2 : C \stackrel{\sqsupseteq}{\Rightarrow} 3 : D\}$.

Under the semantics with compositional consistency, it follows that $\mathfrak{T} \models_\epsilon 3 : D \sqsubseteq \bot$. There is no directed path from $\mathcal{T}_1$ to $\mathcal{T}_3$. If $\mathfrak{T}$ satisfies the directionality property, it should be possible to remove $\mathcal{T}_1$ form $\mathfrak{T}$ without loosing any of the entailed subsumptions. But this is not the case since $\mathfrak{T}^{-1} \not\models_\epsilon 3 : D \sqsubseteq \bot$. The proof that $\mathfrak{T} \models_\epsilon 3 : D \sqsubseteq \bot$ is as follows:

1. $\mathfrak{T} \models_\epsilon 3 : D \sqsubseteq \bot$ because by the second bridge rule, $r_{23}(C^{\mathcal{I}_2}) \supseteq D^{\mathcal{I}_3}$ and $C^{\mathcal{I}_2} = \emptyset$, as we will prove.
2. Assume $r_{23}(C^{\mathcal{I}_2}) \neq \emptyset$.
3. Hence there are $x \in C^{\mathcal{I}_2}$ and $y \in \Delta^{\mathcal{I}_3}$ such that $y \in r_{23}(x)$.
4. Now from compositional consistency there must be $z \in \Delta^{\mathcal{I}_1}$ such that $z \in r_{21}(x)$ and $y \in r_{13}(z)$.
5. but from the first bridge rule we now get $z \in \bot^{\mathcal{I}_1}$ which is a contradiction, as $x \in C^{\mathcal{I}_2}$, $z \in r_{21}(x)$ and $r_{21}(C^{\mathcal{I}_2}) \subseteq \bot^{\mathcal{I}_1}$.

*Example 3.* Consider distributed TBox $\mathfrak{T}$ with index set $I = \{1, 2, 3\}$ and a set of bridge rules $\mathfrak{B}$, such that $\mathcal{T}_1 = \{\top \sqsubseteq \bot\}$, $\mathcal{T}_2 = \emptyset$, $\mathcal{T}_3 = \emptyset$ which only contains one bridge rule

$$2 : C \overset{\sqsupseteq}{\to} 3 : D \ .$$

Since $\mathcal{T}_1$ is inconsistent, the only admissible local interpretation is a hole. Similarly as in the example above, $r_{23} = \emptyset$, because otherwise $\Delta^{\mathcal{I}_1}$ would be non-empty. Hence $\mathfrak{T} \models_\epsilon 3 : D \sqsubseteq \bot$. If local inconsistency would hold, then for each $J \subseteq I$, $3 \notin J$ we must have $\mathfrak{T}(\epsilon_J) \models_d 3 : D \sqsubseteq \bot$. For $J = \{1\}$ this does not hold, however, as we will show. In this case $\mathfrak{T}(\epsilon_J) = \langle \{\mathcal{T}_2, \mathcal{T}_3\}, \mathfrak{B} \rangle$. Consider the d-interpretation $\mathfrak{I}$ with $\Delta^{\mathcal{I}_2} = C^{\mathcal{I}_2} = \{x\}$, $\Delta^{\mathcal{I}_3} = D^{\mathcal{I}_3} = \{y\}$, $r_{23} = \{\langle x, y \rangle\}$ and $r_{32} = \emptyset$. Clearly, $\mathfrak{I}$ is a d-model of $\mathfrak{T}(\epsilon_J)$, computational consistency is satisfied by $r$, and indeed $\mathfrak{I} \not\models_d 3 : D \sqsubseteq \bot$ since $D^{\mathcal{I}_3} \neq \emptyset$.

As showed by Examples 2–3, important properties of the original DDL semantics are violated under the compositional consistency requirement. Willing to preserve these properties, most notably directionality, we have two possible directions: first options is to drop condition (7) completely; the other option is to restrict the applicability of property (7) to the case in which $(i, j)$ and $(j, k)$ are links of the bridge graph. In the following two sections we analyze both these options.

### 3.1 DDL under the Transitivity Condition

From Examples 2–3 we conclude that the compositional consistency requirement is perhaps too strong. It implies all the constraints that are necessary in order to increase subsumption propagation in the amount we have described as desired but it also implies further consequences that are possibly destructive as we have seen in the examples. In the following we study a relaxed version of the compositional consistency condition, which basically is just transitivity of the domain relation in distributed models. We will show that subsumption propagation is more limited under this weaker condition, but on the other hand this semantics satisfies the desiderata postulated for DDL including directionality and the guarded inconsistency propagation. In addition we will be able to introduce a distributed tableaux reasoning algorithm that decides satisfiability for this semantics later in Sect. 4. Formally, the weaker condition is defined as follows.

**Definition 10.** *Given a distributed interpretation $\mathfrak{I}$ with domain relation $r$, we say that $r$ (and also $\mathfrak{I}$) satisfies the transitivity condition if for any three mutually distinct indices $i, j, k \in I$ we have $r_{ij} \circ r_{jk} \subseteq r_{ik}$.*

Apparently, a distributed interpretation satisfies the transitivity condition if and only if its domain relation is transitive. The semantics obtained by allowing only distributed interpretations that satisfy the transitivity condition is hence called DDL under transitivity or DDL with transitive domain relation. The first observation for the new semantics is that the proposition of Theorem 4 does not hold any more if we relax from compositional consistency to transitivity. We demonstrate this problem by an example.

*Example 4.* Consider a distributed TBox $\mathfrak{T}$ with three local TBoxes $\mathcal{T}_1$, $\mathcal{T}_2$ and $\mathcal{T}_3$, with local concepts $C_1, D_1$ in $\mathcal{T}_1$, $E_2$ in $\mathcal{T}_2$ and $F_3$ in $\mathcal{T}_3$, and with bridge rules:

$$2 : E_2 \stackrel{\sqsubseteq}{\rightarrow} 3 : F_3 \ , \qquad\qquad 3 : F_3 \stackrel{\sqsubseteq}{\rightarrow} 1 : D_1 \ ,$$

$$2 : E_2 \stackrel{\sqsupseteq}{\rightarrow} 1 : C_1 \ .$$

This distributed TBox is clearly a very simple instance of the setting assumed by Theorem 4. According to this theorem, we should be able to derive $\mathfrak{T} \models_d 1 : C_1 \sqsubseteq D_1$. This is not true, however, as a distributed model of $\mathfrak{T}$ exists in which $C_1^{\mathcal{I}_1} \not\sqsubseteq D_1^{\mathcal{I}_1}$. This distributed model $\mathfrak{I}$ has three local domains $\Delta^{\mathcal{I}_1} = \{c_1, d_1\}$, $\Delta^{\mathcal{I}_2} = \{e_2\}$, and $\Delta^{\mathcal{I}_3} = \{f_3\}$. Local concepts are interpreted as follows: $C_1^{\mathcal{I}_1} = \{c_1, d_1\}$, $D_1^{\mathcal{I}_1} = \{d_1\}$, $E_2^{\mathcal{I}_2} = \{e_2\}$, and $F_3^{\mathcal{I}_3} = \{f_3\}$. Finally, the domain relation is the following: $r_{21} = \{\langle e_2, c_1 \rangle, \langle e_2, d_1 \rangle\}$, $r_{23} = \{\langle e_2, f_3 \rangle\}$, $r_{31} = \{\langle f_3, d_1 \rangle\}$, $r_{12} = r_{32} = r_{13} = \emptyset$. It is easily verified that $r$ is transitive, hence the transitivity condition is satisfied, and hence Theorem 4 does not stand if compositional consistency is relaxed to transitivity. At the same time, this example does not invalidate Theorem 4 under compositional consistency, as the stronger condition is not satisfied by $\mathfrak{I}$ because $r_{21}(e_2) = \{c_1, d_1\} \not\subseteq \{d_1\} = r_{31}(r_{23}(e_2))$.

So we see that DDL with transitive domain relation has problems with "chaining" into-bridge rules. Luckily enough the problem does not affect chaining onto-bridge rules – if we analyze Example 1 we will see that transitivity is enough to guarantee the two chaining onto-bridge rules in this example to propagate the subsumption relationship as expected. This observation holds in general and is formally established by Theorem 5 below. This theorem is a weaker version of Theorem 4 and it provides a characterization of the semantics of DDL under the transitivity condition.

**Theorem 5.** *Given a distributed TBox, as illustrated in Fig. 7, with an index set $I$ and a set of bridge rules $\mathfrak{B}$, that features $k + 1$ local TBoxes $\mathcal{T}_0, \mathcal{T}_1, \ldots, \mathcal{T}_k$ with concepts $E, F \in \mathcal{T}_0$, and $C_i, D_i \in \mathcal{T}_i$, for $1 \leq i \leq k$, such that:*

*1. $\mathfrak{T} \models_\epsilon i : C_i \sqsubseteq D_i$, for $1 \leq i \leq k$,*
*2. $i + 1 : C_{i+1} \stackrel{\sqsupseteq}{\rightarrow} i : D_i \in \mathfrak{B}$, for $1 \leq i < k$,*
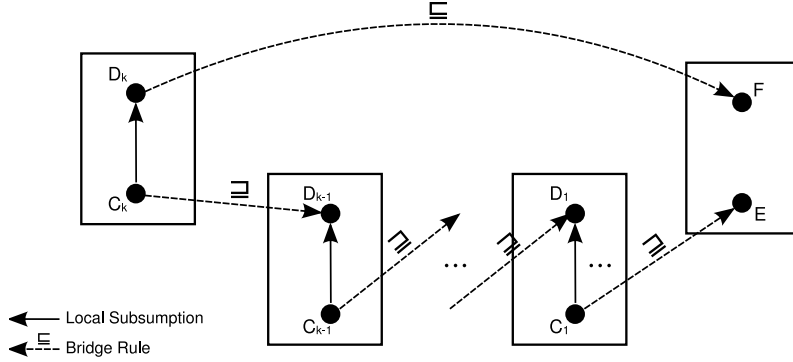
**Fig. 7.** Depiction of the distributed TBox from Theorem 5.

3. $1 : C_1 \overset{\sqsupseteq}{\Rightarrow} 0 : E \in \mathfrak{B}$ and $k : D_k \overset{\sqsubseteq}{\Rightarrow} 0 : F \in \mathfrak{B}$.

In DDL under the transitivity condition it follows that $\mathfrak{T} \models_\epsilon 0 : E \sqsubseteq F$.

*Proof.* The theorem is proved by the following chain of inclusions:

$$E^{\mathcal{I}_0} \overset{1}{\subseteq} r_{k0}\big({C_k}^{\mathcal{I}_k}\big) \overset{2}{\subseteq} r_{k0}\big({D_k}^{\mathcal{I}_k}\big) \overset{3}{\subseteq} F^{\mathcal{I}_0} \quad.$$

Inclusion 3 is a direct consequence of the into-bridge rule $k : D_k \overset{\sqsubseteq}{\Rightarrow} 0 : F \in \mathfrak{B}$. As for Inclusion 2, $\mathfrak{T} \models_d k : C_k \sqsubseteq D_k$ implies ${C_k}^{\mathcal{I}_k} \subseteq {D_k}^{\mathcal{I}_k}$ and since $r_{k0}(\cdot)$ we also have $r_{k0}\big({C_k}^{\mathcal{I}_k}\big) \subseteq r_{k0}\big({D_k}^{\mathcal{I}_k}\big)$. Inclusion 1 divides into:

$$E^{\mathcal{I}_0} \subseteq r_{10}\big(r_{21}\big(\cdots r_{kk-1}\big({C_k}^{\mathcal{I}_k}\big)\cdots\big)\big) \subseteq r_{k0}\big({C_k}^{\mathcal{I}_k}\big) \quad.$$

Both these inclusions ought to be proved by mathematical induction. The first inclusion is proved by induction on $k$, base case is $E^{\mathcal{I}_0} \subseteq r_{10}\big({C_1}^{\mathcal{I}_1}\big)$. This is a direct consequence of the onto-bridge rule $1 : C_1 \overset{\sqsupseteq}{\Rightarrow} 0 : E \in \mathfrak{B}$. In the induction step, by the induction hypothesis we have $E^{\mathcal{I}_0} \subseteq r_{10}\big(r_{21}\big(\cdots r_{k-1k-2}\big({C_{k-1}}^{\mathcal{I}_{k-1}}\big)\cdots\big)\big)$. From the assumptions of the theorem we derive:

$${C_{k-1}}^{\mathcal{I}_{k-1}} \subseteq {D_{k-1}}^{\mathcal{I}_{k-1}} \subseteq r_{kk-1}\big({C_k}^{\mathcal{I}_k}\big) \quad.$$

The composition of mappings $r_{10}(r_{21}(\cdots r_{k-1k-2}(\cdot)\cdots))$ is a mapping, and so $r_{10}\big(r_{21}\big(\cdots r_{k-1k-2}\big({C_{k-1}}^{\mathcal{I}_{k-1}}\big)\cdots\big)\big) \subseteq r_{10}\big(r_{21}\big(\cdots r_{k-1k-2}\big(r_{kk-1}\big({C_k}^{\mathcal{I}_k}\big)\big)\cdots\big)\big)$. Together with the induction hypothesis this amounts to the inclusion we wanted to prove.

As for the second inclusion, we shall prove a slightly more general proposition:

$$r_{n+1n}\big(r_{n+2n+1}\big(\cdots r_{kk-1}\big({C_k}^{\mathcal{I}_k}\big)\cdots\big)\big) \subseteq r_{kn}\big({C_k}^{\mathcal{I}_k}\big) \quad,$$

where $0 \leq n \leq k-2$. The inclusion we ought to prove is then derived by setting $n = 0$. The proposition is proved by mathematical induction on $k - n$. As for

the base case consider $k - n = 2$, and so $n = k - 2$. That means we have to show $r_{k-1k-2}\big(r_{kk-1}\big(C_k{}^{\mathcal{I}_k}\big)\big) \subseteq r_{kk-2}\big(C_k{}^{\mathcal{I}_k}\big)$, which holds because $\mathfrak{I}$ satisfies transitivity. The induction step is for $k - n = j > 2$, and so $n = k - j$. From the induction hypothesis we get $r_{n+2n+1}\big(r_{n+3n+2}\big(\cdots r_{kk-1}\big(C_k{}^{\mathcal{I}_k}\big)\cdots\big)\big) \subseteq r_{kn+1}\big(C_k{}^{\mathcal{I}_k}\big)$. And since $r_{n+1n}(\cdot)$ is a mapping, we also get $r_{n+1n}\big(r_{n+2n+1}\big(\cdots r_{kk-1}\big(C_k{}^{\mathcal{I}_k}\big)\cdots\big)\big) \subseteq r_{n+1n}\big(r_{kn+1}\big(C_k{}^{\mathcal{I}_k}\big)\big)$. Once again, since $\mathfrak{I}$ satisfies the transitivity condition we have $r_{n+1n}\big(r_{kn+1}\big(C_k{}^{\mathcal{I}_k}\big)\big) \subseteq r_{kn}\big(C_k{}^{\mathcal{I}_k}\big)$. By combining the last two inclusions we directly get:

$$r_{n+1n}\big(r_{n+2n+1}\big(\cdots r_{kk-1}\big(C_k{}^{\mathcal{I}_k}\big)\cdots\big)\big) \subseteq r_{kn}\big(C_k{}^{\mathcal{I}_k}\big) \quad .$$

By setting $n = 0$ we derive the inclusion we wanted to prove, and hence the theorem. $\square$

And so we see that this weaker semantics only partially solves the problem of lacking subsumption propagation between remote ontologies, as we have identified it in the beginning of Sect. 3. Subsumption propagates along chains of onto-bridge rules, but it fails to propagate along chains of into-bridge rules. On the other hand, compared to the stronger semantics, it satisfies the requirements that have been placed on DDL (Borgida & Serafini 2003, Serafini et al. 2005), as we formally demonstrate below.

**Theorem 6.** *The monotonicity property is satisfied in DDL under transitivity.*

*Proof.* Given $\mathcal{T}_i \in \mathfrak{T}$ and $C, D \in \mathcal{T}_i$ such that $\mathcal{T}_i \models_\epsilon C \sqsubseteq D$, only models of $\mathcal{T}_i$ and $\mathcal{I}^\epsilon$ are allowed in the $\mathcal{I}_i$ slot of any distributed model $\mathfrak{I}$ of $\mathfrak{T}$. Since in each of these $C \sqsubseteq D$ holds, we get $\mathfrak{T} \models_\epsilon i : C \sqsubseteq D$ directly from the definition. $\square$

While monotonicity is a rather trivial property that expectedly holds also for this semantics, recall that the stronger semantics we have presented in the previous section does not satisfy directionality, as we have showed. Fortunately the semantics of DDL under transitivity behaves more reasonably – directionality is satisfied – as we formally prove below.

**Theorem 7.** *The directionality property is satisfied in DDL under transitivity.*

*Proof.* Given a distributed TBox $\mathfrak{T}$ we shall prove that if there is no directed path from $i$ to $j$ in $G_\mathfrak{T}$, then

$$\mathfrak{T} \models_\epsilon j : C \sqsubseteq D \iff \mathfrak{T}^{-i} \models_\epsilon j : C \sqsubseteq D$$

where $\mathfrak{T}^{-1}$ is obtained by removing $\mathcal{T}_i$, $\mathfrak{B}_{ik}$ and $\mathfrak{B}_{li}$ from $\mathfrak{T}$ for each $k, l \in I$, $k, l \neq i$.

*If part.* By contradiction. Assume that $\mathfrak{T}^{-i} \models_\epsilon j : C \sqsubseteq D$ and $\mathfrak{T} \not\models_\epsilon j : C \sqsubseteq D$. If $\mathfrak{T} \not\models_\epsilon j : C \sqsubseteq D$ then there is a distributed model $\mathfrak{I}$ of $\mathfrak{T}$ in which $C^{\mathcal{I}_j} \not\subseteq D^{\mathcal{I}_j}$. Let $\mathfrak{I}^{-i}$ be obtained from $\mathfrak{I}$ by removing $\mathcal{I}_i$ and $r_{ik}$ and $r_{ki}$ for all $k \neq i$. Since $\mathfrak{I}$ is a model of $\mathfrak{T}$, $\mathfrak{I}^{-i}$ is a model of $\mathfrak{T}^{-i}$ (this follows directly from the

construction). However $C^{\mathcal{I}_j^{-i}} = C^{\mathcal{I}_j}$ and $D^{\mathcal{I}_j^{-i}} = D^{\mathcal{I}_j}$ and hence $C^{\mathcal{I}_j^{-i}} \not\sqsubseteq D^{\mathcal{I}_j^{-i}}$. This implies that $\mathfrak{T}^{-i} \not\models_\epsilon j : C \sqsubseteq D$ but we have assumed the contrary.

*Only-if part.* By contradiction. Assume that $\mathfrak{T} \models_\epsilon j : C \sqsubseteq D$ and $\mathfrak{T}^{-i} \not\models_\epsilon j : C \sqsubseteq D$. If $\mathfrak{T}^{-i} \not\models_\epsilon j : C \sqsubseteq D$, there is a distributed model $\mathfrak{I}^{-i}$ of $\mathfrak{T}^{-i}$ in which $C^{\mathcal{I}_j^{-i}} \not\sqsubseteq D^{\mathcal{I}_j^{-i}}$. Let $J \subset I$ be the set of nodes reachable from $i$ by some directed path in $G_{\mathfrak{T}}$. Let $\mathfrak{I}$ be the following distributed interpretation of $\mathfrak{T}$:

1. $\mathcal{I}_m = \mathcal{I}_m^{-i}$ for each $m \in I \setminus J$;
2. $\mathcal{I}_m = \mathcal{I}^\epsilon$ for each $m \in J$;
3. $r_{mn} = r_{mn}^{-i}$ if $m, n \in I \setminus J$;
4. $r_{mn} = \emptyset$ if either $m \in J$ or $n \in J$.

We now show that $\mathfrak{I}$ is a distributed model of $\mathfrak{T}$. In order to demonstrate this, we need to show that local axioms and bridge rules of $\mathfrak{T}$ are satisfied by $\mathfrak{I}$, and that the domain relation of $\mathfrak{I}$ is transitive.

**Local axioms.** For any $m \in I$, each local GCI axiom $\phi = m : G \sqsubseteq H$ must be satisfied by $\mathcal{I}_m$. If $m \in J$ then $\mathcal{I}_m$ is a hole, which satisfies any local CGI axiom, and hence it also satisfies $\phi$. If $m \in I \setminus J$ then $\mathcal{I}_m = \mathcal{I}_m^{-i}$, which satisfies $\phi$ because $\mathfrak{I}^{-i}$ is a distributed model of $\mathfrak{T}^{-i}$ and $\phi$ occurs in $\mathcal{T}_m^{-i}$ since $\mathcal{T}_m = \mathcal{T}_m^{-i}$.

**Bridge rules.** Given $m, n \in I$ and a bridge rule $b$ that maps from $m : G$ to $n : H$, we must show that $b$ is satisfied by $\mathfrak{I}$. We distinguish between several cases. First, if $m, n \notin J$, then $b$ is satisfied by $\mathfrak{I}$ because $r_{mn}(G^{\mathcal{I}_m}) = r_{mn}^{-i}(G^{\mathcal{I}_m^{-i}})$, $H^{\mathcal{I}_n} = H^{\mathcal{I}_n^{-i}}$, and $b$ is satisfied by $\mathfrak{I}^{-i}$. The case when $m \in J$ and $n \in I \setminus J$ is impossible, because there is a bridge rule from $m$ to $n$ and hence if $m$ belongs to $J$, $n$ must also belong to $J$. In all remaining cases $n \in J$. In this case it follows that $r_{mn}(G^{\mathcal{I}_m}) = \emptyset$ and $H^{\mathcal{I}_n} = \emptyset$, because $r_{mn} = \emptyset$ and $\mathcal{I}_n$ is a hole. In such a case however $\mathfrak{I}$ satisfies $\phi$.

**Transitivity.** We have to show for any three mutually distinct indices $k, m, n \in I$ that $r_{km} \circ r_{mn} \subseteq r_{kn}$. If none of $k, m, n$ belongs to $J$, then the fact that $r_{km} = r_{km}^{-i}$, $r_{mn} = r_{mn}^{-i}$, and $r_{kn} = r_{kn}^{-i}$ implies $r_{km} \circ r_{mn} \subseteq r_{kn}$. If at least one of $k, m$ and $n$ belongs to $J$, then either $r_{km}$ or $r_{mn}$ is equal to $\emptyset$, which implies that $r_{km} \circ r_{mn} = \emptyset \subseteq r_{kn}$.

And so we have showed that $\mathfrak{I}$ is a distributed model of $\mathfrak{T}$. However, since there is no directed path from $i$ to $j$ in $G_{\mathfrak{T}}$, it follows that $j \notin J$ and hence $\mathcal{I}_j = \mathcal{I}_j^{-i}$ and hence $C^{\mathcal{I}_j} \not\sqsubseteq D^{\mathcal{I}_j}$. This in turn implies that $\mathfrak{T} \not\models_\epsilon j : C \sqsubseteq D$, but we have assumed the contrary. $\qquad\square$

Another property which is broken by the stronger semantics is the so called local inconsistency property. It turns out that this problem is indeed caused by the intricate implications of computational consistency and when we only require transitivity, the property holds.

**Theorem 8.** *The local inconsistency property is satisfied in DDL under transitivity.*

*Proof.* Given a distributed TBox $\mathfrak{T}$ over an index set $I$, and given some $i \in I$, a subsumption formula $\phi = i : C \sqsubseteq D$, we ought to prove the equivalence:

$$\mathfrak{T} \models_\epsilon \phi \iff ((\forall J)\ J \subseteq I \wedge i \notin J \implies \mathfrak{T}(\epsilon_J) \models_{\mathrm{d}} \phi)$$

*If part.* Suppose by contradiction that for all subsets $J$ of $I$ such that $i \notin J$, $\mathfrak{T}(\epsilon_J) \models_{\mathrm{d}} \phi$ and $\mathfrak{T} \not\models_\epsilon \phi$. In that case there is an $\epsilon$-model $\mathfrak{I}^1$ of $\mathfrak{T}$ in which $C^{\mathcal{I}_i^1} \not\subseteq D^{\mathcal{I}_i^1}$. Since this is an $\epsilon$-model, some of its local interpretations are possibly holes. Let $J$ be the set of all indices in $I$ that hold a hole as local interpretation, i.e., $J = \{j \in I \mid \mathcal{I}_j^1 = \mathcal{I}^\epsilon\}$. As $\mathfrak{I}^1 \not\models_\epsilon \phi$ we have $i \notin J$, because otherwise $\mathcal{I}_i = \mathcal{I}^\epsilon$ and hole entails any formula including $\phi$. Let $\mathfrak{I}^2$ be the distributed interpretation obtained by cutting the holes off from $\mathfrak{I}^1$, i.e., $\mathfrak{I}^2$ has $I \setminus J$ as its index set, $\mathcal{I}_j^2 = \mathcal{I}_j^1$ for each $j \in I \setminus J$ and $r_{jk}^2 = r_{jk}^1$ for any $j, k \in I \setminus J$. We will now show that $\mathfrak{I}^2$ is a d-model of $\mathfrak{T}(\epsilon_J)$.

1. Indeed $\mathfrak{I}^2$ is a d-interpretation, because it does not contain holes any more.
2. Given any $j \in I \setminus J$ and any $j$-local GCI $\psi$, either $\psi$ also belongs to $\mathcal{T}_j$ of $\mathfrak{T}$ and hence it is satisfied by $\mathcal{I}_j^2 = \mathcal{I}_j^1$ since $\mathfrak{I}^1$ satisfies $\mathcal{T}_j$, or by construction of $\mathfrak{T}(\epsilon_J)$ $\psi = j : D \sqsubseteq \bot$ and there is some bridge rule $b = k : C \overset{\sqsupseteq}{\to} j : D$ in $\mathfrak{T}$ with $k \in J$. Since $b$ is satisfied by $\mathfrak{I}^1$ and $\mathcal{I}_k^1 = \mathcal{I}^\epsilon$ we have $D^{\mathcal{I}_j^2} = D^{\mathcal{I}_j^1} = \emptyset$ and hence $\psi$ is satisfied.
3. Also each bridge of $\mathfrak{T}(\epsilon_J)$ rule, say $b$ between $\mathcal{T}_j$ and $\mathcal{T}_k$, $j, k \in I \setminus J$, is satisfied by $\mathfrak{I}^2$ since $b$ also appears in $\mathfrak{T}$ and $\mathcal{I}_j^2 = \mathcal{I}_j^1$, $\mathcal{I}_k^2 = \mathcal{I}_k^1$, $r_{jk}^2 = r_{jk}^1$ and $\mathfrak{I}^1$ satisfies $b$.
4. Moreover $\mathfrak{I}^2$ satisfies transitivity because $\mathfrak{I}^1$ does and $r_{jk}^2 = r_{jk}^1$ for any $j, k \in I \setminus J$.

Hence we have found a contradiction because $\mathfrak{T}(\epsilon_J)$ ought to entail $\phi$ but it does not, since $\mathfrak{I}^2$ is a model of $\mathfrak{T}(\epsilon_J)$ in which $C^{\mathcal{I}_i^2} \not\subseteq D^{\mathcal{I}_i^2}$ (because $C^{\mathcal{I}_i^1} \not\subseteq D^{\mathcal{I}_i^1}$, $C^{\mathcal{I}_i^2} = C^{\mathcal{I}_i^1}$ and $D^{\mathcal{I}_i^2} = D^{\mathcal{I}_i^1}$).

*Only-if part.* Suppose by contradiction that $\mathfrak{T} \models_\epsilon \phi$ and there is $J \subseteq I$, not containing $i$, such that $\mathfrak{T}(\epsilon_J) \not\models_{\mathrm{d}} \phi$. In this case there must be a $d$-model of $\mathfrak{T}(\epsilon_J)$, say $\mathfrak{I}^1$ such that $C^{\mathcal{I}_i^1} \not\subseteq D^{\mathcal{I}_i^1}$. Let us construct a distributed interpretation $\mathfrak{I}^2$, starting from $\mathfrak{I}^1$ and adding hole into each slot $j \in J$, i.e., $\mathcal{I}_j^2 = \mathcal{I}_j^1$ for each $j \in I \setminus J$ and $\mathcal{I}_j^2 = \mathcal{I}^\epsilon$ for each $j \in J$. Moreover, for any pair of indices $j, k \in I$ we set $r_{jk}^2 = r_{jk}^1$ if both $i, j \in I \setminus J$ and $r_{jk}^2 = \emptyset$ otherwise. We now prove that $\mathfrak{I}^2$ is an $\epsilon$-model of $\mathfrak{T}$.

1. Obviously $\mathfrak{I}^2$ is an $\epsilon$-interpretation of $\mathfrak{T}$.
2. Each local GCI $\psi$ in $\mathfrak{T}$ is satisfied, as if $\psi \in \mathcal{T}_j$, $j \in I \setminus J$ then $\psi$ also appears in $\mathfrak{T}(\epsilon_J)$ and since $\mathcal{I}_j^2 = \mathcal{I}_j^1$ we have $\mathcal{I}_j^2$ entails $\psi$. If $\psi \in \mathcal{T}_k$, $k \in J$ then $\psi$ is satisfied since $\mathcal{I}_k^2$ is a hole.
3. Let $b$ be a bridge rule directed from $\mathcal{T}_j$ and $\mathcal{T}_k$, $j, k \in I$ that maps between $j : C$ and $k : D$. If both $j, k \in I \setminus J$ then $b$ is satisfied because it is satisfied by $\mathfrak{I}^1$ and we have $\mathcal{I}_j^2 = \mathcal{I}_j^1$, $\mathcal{I}_k^2 = \mathcal{I}_k^1$, $r_{jk}^2 = r_{jk}^1$. If both $j, k \in J$ than $b$ is satisfied because $\mathcal{I}_j^2 = \mathcal{I}^\epsilon$, $\mathcal{I}_k^2 = \mathcal{I}^\epsilon$ and $r_{jk}^2 = \emptyset$. If $j \in I \setminus J$ and $k \in J$

than $b$ is satisfied because $\mathcal{I}_k^2 = \mathcal{I}^\epsilon$ and $r_{jk}^2 = \emptyset$. If $j \in J$ and $k \in I \setminus J$ and $b = j : C \overset{\sqsubseteq}{\Rightarrow} k : D$ then $b$ is satisfied because $r_{jk}^2(C^{\mathcal{I}_j^2}) = \emptyset$ which surely is a subset of $D^{\mathcal{I}_k^2}$. The last case occurs when $j \in J$, $k \in I \setminus J$ and $b = j : C \overset{\sqsupseteq}{\Rightarrow} k : D$. In this case $b$ is also satisfied because by construction the local GCI $k : D \sqsubseteq \bot$ belongs to $\mathfrak{T}(\epsilon_J)$ and hence $D^{\mathcal{I}_k^2} = D^{\mathcal{I}_k^1} = \emptyset$ which surely is a subset of $r_{jk}^2(C^{\mathcal{I}_j^2})$.

4. Finally, we show that the domain relation $r^2$ is transitive, that is, for any three mutually distinct $k, m, n \in I$ it holds that $r_{km}^2 \circ r_{mn}^2 \subseteq r_{kn}^2$. If none of $k, m, n$ belongs to $J$ that this holds because $r_{km}^2 = r_{km}^1$, $r_{mn}^2 = r_{mn}^1$, $r_{kn}^2 = r_{kn}^1$, and $r^1$ is transitive. Thanks to the construction of $\mathfrak{J}^2$ however, in any other case either $r_{km}^2 = \emptyset$ (if $k \in J$ or $m \in J$) or $r_{mn}^2 = \emptyset$ (if $m \in J$ or $n \in J$). Hence in all these cases transitivity trivially holds, since $r_{km}^2 \circ r_{mn}^2 = \emptyset \subseteq r_{kn}^2$.

And so we have just showed that $\mathfrak{J}^2$ is an $\epsilon$-model of $\mathfrak{T}$. Now we again reach contradiction, as we have assumed that $\mathfrak{T}$ entails $\phi$ but $C^{\mathcal{I}_i^2} \not\subseteq D^{\mathcal{I}_i^2}$ (because $C^{\mathcal{I}_i^1} \not\subseteq D^{\mathcal{I}_i^1}$, $C^{\mathcal{I}_i^2} = C^{\mathcal{I}_i^1}$ and $D^{\mathcal{I}_i^2} = D^{\mathcal{I}_i^1}$). $\qquad\square$

And so we have showed that in the semantics of DDL under transitivity the level of subsumption propagation between remote ontologies is limited to propagation along chaining onto-bridge rules in scenarios that instantiate Theorem 5. Propagation of subsumption along chaining into-bridge rules is not granted, but all desiderata postulated for DDL, including transitivity and local inconsistency are satisfied.

## 3.2 DDL with restricted compositionality

We will now take the other approach that we have outlined. In this approach we keep both restrictions (7) and (8), but we try to apply them more cautiously. This approach is motivated by the following example.

*Example 5.* Consider a distributed TBox $\mathfrak{T}$ which contains local ontologies about cats $\mathcal{T}_c$, music $\mathcal{T}_m$, botany $\mathcal{T}_b$ and another one about fruit $\mathcal{T}_f$. Suppose that there is some knowledge encoded within each of them and they are all consistent. There are also some bridge rules between $\mathcal{T}_b$ and $\mathcal{T}_f$ for instance:

$$b : \mathsf{Fruit} \overset{\sqsubseteq}{\Rightarrow} f : \mathsf{Fruit} \ , \qquad b : \mathsf{Berry} \overset{\sqsupseteq}{\Rightarrow} f : \mathsf{Berry} \ .$$

In contrast the two local ontologies $\mathcal{T}_c$ and $\mathcal{T}_t$ are completely isolated. Let $\mathfrak{J}$ be a model of $\mathfrak{T}$ in which $\mathsf{Berry}^{\mathcal{I}_f}$ is nonempty and hence there are some $x \in \mathsf{Berry}^{\mathcal{I}_b}$, $y \in \mathsf{Berry}^{\mathcal{I}_f}$ such that $(x, y) \in r_{bf}$. So far this situation is perfectly natural. However, compositional consistency requires existence of $z_1 \in \Delta^{\mathcal{I}_c}$ and $z_2 \in \Delta^{\mathcal{I}_m}$ such that $(x, z_1) \in r_{bc}$, $(z_1, y) \in r_{cf}$, $(x, z_2) \in r_{bm}$ and $(z_2, y) \in r_{mf}$. In turn, $(x, z_1) \in r_{bc}$ implies existence of $(x, w_1) \in r_{bm}$, $(w_1, z_1) \in r_{mc}$, $(x, w_2) \in r_{bf}$ and $(w_2, z_1) \in r_{fc}$, etc. As we can see, even if local TBoxes $\mathcal{T}_c$ and $\mathcal{T}_m$ are totally unrelated, compositional consistency implies existence of elements in $\mathcal{I}_c$ and $\mathcal{I}_m$

and generates many $r$-connections between these elements and elements of the other local TBoxes. Such behaviour is indeed rather strange and as we have seen it even causes trouble in the sense that some important DDL properties are violated.

From the example above we conjecture that compositional consistency is perhaps a stronger condition than we really need in order to support subsumption propagation. It does the job, as established by Theorem 4 but in addition it has undesirable side effects. We try to weaken the condition as follows. Given a distributed TBox and a distributed interpretation over some index set $I$, we do not require compositional consistency to hold for every $i, j, k \in I$, but only when local TBoxes $\mathcal{T}_i$, $\mathcal{T}_j$ and $\mathcal{T}_k$ are actually connected with bridge rules. The relaxed version of the restriction is as follows.

**Definition 11.** *Given a distributed TBox $\mathfrak{T}$ with an index set $I$ and bridge graph $G_{\mathfrak{T}}$, let $\mathfrak{I}$ be a distributed interpretation of $\mathfrak{T}$ with domain relation $r$, we say that $r$ (and also $\mathfrak{I}$) satisfies restricted compositional consistency (also restricted compositionality) if for any three mutually distinct indices $i, j, k \in I$ such that there is a directed path from $i$ to $j$ and another one from $j$ to $k$ in $G_{\mathfrak{T}}$ we have $r_{ij} \circ r_{jk} = r_{ik}$.*

The semantics of DDL under this condition is called DDL under restricted compositionality. This restriction is reasonable in the sense that the proposition of Theorem 4 holds even in this relaxed case. In other words, subsumption does propagate along both: chains of into- and chains of onto-bridge rules. This is verified by inspection of the proof given by Homola (2008): compositional consistency is only applied in this proof in cases which are covered also by the weaker version. Also directionality is satisfied under this semantics.

**Theorem 9.** *The directionality property is satisfied in DDL under restricted compositionality.*

*Proof.* Given a distributed TBox $\mathfrak{T}$ we shall prove that if there is no directed path from $i$ to $j$ in $G_{\mathfrak{T}}$, then

$$\mathfrak{T} \models_{\epsilon} j : C \sqsubseteq D \iff \mathfrak{T}^{-i} \models_{\epsilon} j : C \sqsubseteq D$$

where $\mathfrak{T}^{-1}$ is obtained by removing $\mathcal{T}_i$, $\mathfrak{B}_{ik}$ and $\mathfrak{B}_{li}$ from $\mathfrak{T}$ for each $k, l \in I$, $k, l \neq i$.

*If part.* By contradiction. Assume that $\mathfrak{T}^{-i} \models_{\epsilon} j : C \sqsubseteq D$ and $\mathfrak{T} \not\models_{\epsilon} j : C \sqsubseteq D$. If $\mathfrak{T} \not\models_{\epsilon} j : C \sqsubseteq D$ then there is a distributed model $\mathfrak{I}$ of $\mathfrak{T}$ in which $C^{\mathcal{I}_j} \not\sqsubseteq D^{\mathcal{I}_j}$. Let $\mathfrak{I}^{-i}$ be obtained from $\mathfrak{I}$ by removing $\mathcal{I}_i$ and $r_{ik}$ and $r_{ki}$ for all $k \neq i$. Since $\mathfrak{I}$ is a model of $\mathfrak{T}$, $\mathfrak{I}^{-i}$ is a model of $\mathfrak{T}^{-i}$ (this follows directly from the construction). However $C^{\mathcal{I}_j^{-i}} = C^{\mathcal{I}_j}$ and $D^{\mathcal{I}_j^{-i}} = D^{\mathcal{I}_j}$ and hence $C^{\mathcal{I}_j^{-i}} \not\sqsubseteq D^{\mathcal{I}_j^{-i}}$. This implies that $\mathfrak{T}^{-i} \not\models_{\epsilon} j : C \sqsubseteq D$ but we have assumed the contrary.

*Only-if part.* By contradiction. Assume that $\mathfrak{T} \models_{\epsilon} j : C \sqsubseteq D$ and $\mathfrak{T}^{-i} \not\models_{\epsilon} j : C \sqsubseteq D$. If $\mathfrak{T}^{-i} \not\models_{\epsilon} j : C \sqsubseteq D$, there is a distributed model $\mathfrak{I}^{-i}$ of $\mathfrak{T}^{-i}$ in which $C^{\mathcal{I}_j^{-i}} \not\sqsubseteq D^{\mathcal{I}_j^{-i}}$. Let $J \subset I$ be the set of nodes reachable from $i$ by some directed path in $G_{\mathfrak{T}}$. Let $\mathfrak{I}$ be the following distributed interpretation of $\mathfrak{T}$:

1. $\mathcal{I}_m = \mathcal{I}_m^{-i}$ for each $m \in I \setminus J$;
2. $\mathcal{I}_m = \mathcal{I}^\epsilon$ for each $m \in J$;
3. $r_{mn} = r_{mn}^{-i}$ if $m, n \in I \setminus J$;
4. $r_{mn} = \emptyset$ if either $m \in J$ or $n \in J$.

We now show that $\mathfrak{I}$ is a distributed model of $\mathfrak{T}$. In order to demonstrate this, we need to show that local axioms and bridge rules of $\mathfrak{T}$ are satisfied by $\mathfrak{I}$, and that the domain relation of $\mathfrak{I}$ satisfies the restricted compositionality condition.

**Local axioms.** For any $m \in I$, each local GCI axiom $\phi = m : G \sqsubseteq H$ must be satisfied by $\mathcal{I}_m$. If $m \in J$ then $\mathcal{I}_m$ is a hole, which satisfies any local CGI axiom, and hence it also satisfies $\phi$. If $m \in I \setminus J$ then $\mathcal{I}_m = \mathcal{I}_m^{-i}$, which satisfies $\phi$ because $\mathfrak{I}^{-i}$ is a distributed model of $\mathfrak{T}^{-i}$ and $\phi$ occurs in $\mathcal{T}_m^{-i}$ since $\mathcal{T}_m = \mathcal{T}_m^{-i}$.

**Bridge rules.** Given $m, n \in I$ and a bridge rule $b$ that maps from $m : G$ to $n : H$, we must show that $b$ is satisfied by $\mathfrak{I}$. We distinguish between several cases. First, if $m, n \notin J$, then $b$ is satisfied by $\mathfrak{I}$ because $r_{mn}(G^{\mathcal{I}_m}) = r_{mn}^{-i}(G^{\mathcal{I}_m^{-i}})$, $H^{\mathcal{I}_n} = H^{\mathcal{I}_n^{-i}}$, and $b$ is satisfied by $\mathfrak{I}^{-i}$. The case when $m \in J$ and $n \in I \setminus J$ is impossible, because there is a bridge rule from $m$ to $n$ and hence if $m$ belongs to $J$, $n$ must also belong to $J$. In all remaining cases $n \in J$. In this case it follows that $r_{mn}(G^{\mathcal{I}_m}) = \emptyset$ and $H^{\mathcal{I}_n} = \emptyset$, because $r_{mn} = \emptyset$ and $\mathcal{I}_n$ is a hole. In such a case however $\mathfrak{I}$ satisfies $\phi$.

**Restricted compositionality.** Let $k, m, n \in I$ be three mutually distinct indices such that there exists a directed path from $k$ to $m$ in $G_{\mathfrak{T}}$ and another one from $m$ and $n$. We must now show that $r_{km} \circ r_{mn} = r_{kn}$. If none of $k, m, n$ belongs to $J$, then the fact that $r_{km} = r_{km}^{-i}$, $r_{mn} = r_{mn}^{-i}$, and $r_{kn} = r_{kn}^{-i}$ implies that $r_{km} \circ r_{mn} = r_{kn}$. If $k \in J$, then the fact that $r_{km} = \emptyset$ and $r_{kn} = \emptyset$ implies that $r_{km} \circ r_{mn} = r_{kn}$. If $m \in J$ then necessarily also $n \in J$ since there is a directed path between these two nodes in $G_{\mathfrak{T}}$. But $n \in J$ implies that $r_{mn} = r_{kn} = \emptyset$, thus implying $r_{km} \circ r_{mn} = r_{kn}$. Finally if $n \in J$ then again $r_{mn} = r_{kn} = \emptyset$, and so $r_{km} \circ r_{mn} = r_{kn}$.

And so we have showed that $\mathfrak{I}$ is a distributed model of $\mathfrak{T}$. However, since there is no directed path from $i$ to $j$ in $G_{\mathfrak{T}}$, it follows that $j \notin J$ and hence $\mathcal{I}_j = \mathcal{I}_j^{-i}$ and hence $C^{\mathcal{I}_j} \not\sqsubseteq D^{\mathcal{I}_j}$. This in turn implies that $\mathfrak{T} \not\models_\epsilon j : C \sqsubseteq D$, but we have assumed the contrary. $\square$

And so we have showed that this restricted form of compositional consistency does not break directionality. Unfortunately, still there are some undesired effects that this restriction poses on the semantics as demonstrated by the example below.

*Example 6.* Consider distributed a TBox $\mathfrak{T}$ with index set $I = \{1, 2, 3\}$ and a set of bridge rules $\mathfrak{B}$, such that $\mathcal{T}_1 = \{\top \sqsubseteq \bot\}$, $\mathcal{T}_2 = \emptyset$, $\mathcal{T}_3 = \emptyset$ which contains bridge rules:

$$2 : E \overset{\sqsupseteq}{\Rightarrow} 1 : F \ , \qquad\qquad 1 : G \overset{\sqsupseteq}{\Rightarrow} 3 : H \ ,$$

$$2 : C \overset{\sqsupseteq}{\Rightarrow} 3 : D \ .$$

Since $\mathcal{T}_1$ is inconsistent, the only admissible interpretation is a hole, and so due to compositional consistency, even in the restricted case, $r_{23} = \emptyset$. Hence $\mathfrak{T} \models_\epsilon 3 : D \sqsubseteq \bot$. If local inconsistency would hold, then for each $J \subseteq I$, $3 \notin J$ we must have $\mathfrak{T}(\epsilon_J) \models_d 3 : D \sqsubseteq \bot$. Similarly as in Example 3, for $J = \{1\}$ this does not hold, as $\mathfrak{T}(\epsilon_J)$ admits a d-model $\mathfrak{I}$ with $\Delta^{\mathcal{I}_2} = C^{\mathcal{I}_2} = \{x\}$, $\Delta^{\mathcal{I}_3} = D^{\mathcal{I}_3} = \{y\}$, $r_{23} = \{\langle x, y \rangle\}$ and $r_{32} = \emptyset$, for which $\mathfrak{I} \not\models_d 3 : D \sqsubseteq \bot$ since $D^{\mathcal{I}_3} \neq \emptyset$.

In this section we have examined and evaluated the approach of applying the compositional consistency restriction in DDL in order to improve subsumption propagation (Homola 2008). We have showed that while subsumption propagation is indeed improved, there are also undesirable consequences of such an approach, the directionality and the local inconsistency properties are violated. In order to retain directionality we have verified two options. First, *falling back from computational consistency to transitivity* which amounts to less subsumption propagation but all desired properties of DDL are retained. And second, *applying compositional consistency more cautiously*, only if the local ontologies that are involved are connected by some directed path of bridge rules. The second approach provides just the same level of subsumption propagation as the semantics under unrestricted compositionality and also retains directionality but as we have just seen it still violates the local inconsistency property.

Summing up, with respect to the current state of the art, the semantics of DDL with transitive domain relation provides a reasonable compromise. It ensures subsumption propagation along chaining onto-bridge rules in scenarios that instantiate Theorem 5. Propagation of subsumption under chaining into-bridge rules is not granted, but all desiderata postulated for DDL, including transitivity and localized inconsistency are satisfied. In the next section we look at reasoning with this semantics.

## 4 Distributed Tableaux Algorithm for transitive DDL

In this section, we introduce a distributed tableaux algorithm for deciding satisfiability of concepts with respect to an acyclic distributed TBox for DDL over $\mathcal{ALC}$ under the transitivity requirement. As much as in the original algorithm (Serafini et al. 2005), also in our approach, local reasoners are run independently. We keep precise track of the domain relation $r$ however, and the communication between local reasoners is divided into multiple queries. Hence while the original algorithm can be seen as working with autonomous local tableaux by passing queries, our algorithm can be seen as working with a truly distributed tableau.

**Definition 12.** *Assume a distributed TBox $\mathfrak{T} = \langle \{\mathcal{T}_i\}_{i \in I}, \mathfrak{B} \rangle$ over $\mathcal{ALC}$ with index set $I$, concept names $N_C = \bigcup_{i \in I} N_{Ci}$ and role names $N_R = \bigcup_{i \in I} N_{Ri}$. Let $CC_i$ be the set of all (atomic and complex) concepts over $N_{Ci}$ and $N_{Ri}$ in*

*negation normal form.* A distributed completion tree[4] $T = \{T_i\}_{i \in I}$ *is a set of labeled trees* $T_i = \langle V_i, E_i, \mathcal{L}_i, r_i \rangle$, *such that for each* $i \in I$:

1. *the trees* $\{\langle V_i, E_i \rangle\}_{i \in I}$ *are mutually disjoint;*
2. *the labeling function* $\mathcal{L}_i$ *labels each node* $x \in V_i$ *with* $\mathcal{L}(x) \subseteq 2^{\mathrm{CC}_i}$ *and each edge* $\langle x, y \rangle \in E_i$ *with* $\mathcal{L}(\langle x, y \rangle) \in \mathrm{N}_{\mathrm{R}i}$;
3. *the labeling function* $r_i$ *labels each node* $x \in V_i$ *with a set of references to its r-images* $r_i(x) \subseteq \{j : y \mid j \in I \wedge y \in V_j\}$.

During the run of the tableaux algorithm, tableaux expansion rules are applied on the completion tree and the tree is expanded by each rule application. If no more rules are applicable any more, we say that the completion tree is *complete.* There is a *clash* in the completion tree $T$ if for some $x \in V_i$ and for some $C \in \mathrm{N}_{\mathrm{C}i}$ we have $\{C, \neg C\} \subseteq \mathcal{L}_i(x)$. If there is no clash in $T$ then we say that $T$ is *clash-free.* In order to assure termination we use the standard subset blocking technique that is common for $\mathcal{ALC}$. Given a distributed completion tree $T = \{T_i\}_{i \in I}$, a node $x \in V_i$ is blocked, if it has an ancestor $y \in V_i$ such that $\mathcal{L}_i(x) \subseteq \mathcal{L}_i(y)$. In such a case we also say that $x$ is blocked by $y$. A node $y \in V_i$ is said to be an *R-successor* of $x \in V_i$, if $\langle x, y \rangle \in E_i$ and $\mathcal{L}_i(\langle x, y \rangle) = R$.

The distributed tableaux algorithm for deciding satisfiability of concepts with respect to a distributed TBox takes a distributed TBox $\mathfrak{T}$, a concept $C$ in NNF and $i \in I$ as its inputs. The algorithm then continues in three steps:

1. *Initialization.* Create a new completion tree $T = \{T_j\}_{j \in I}$ such that $T_j = \langle \{s_0\}, \emptyset, \{s_0 \mapsto \{C\}\}, \emptyset \rangle$ for $j = i$ and $T_j = \langle \emptyset, \emptyset, \emptyset, \emptyset \rangle$ for $j \neq i$.
2. *Tableau expansion.* Apply the tableaux expansion rules of Fig. 8 exhaustively.
3. *Answer.* If none of the tableaux expansion rules in Fig. 8 is applicable any more (i.e., the completion tree is now complete), answer "$C$ is satisfiable" if a clash-free completion tree has been constructed. Answer "$C$ is unsatisfiable" otherwise.

Below we present a formal correctness proof for the newly introduced algorithm. The proof is based on the classic proof for $\mathcal{ALC}$ as in fact the only new thing to prove here is that the algorithm uses the $\overset{\exists}{\Rightarrow}$-rule and the $\overset{\sqsubseteq}{\Rightarrow}$-rule to combine several autonomous local $\mathcal{ALC}$ reasoners correctly. The proof is done in three parts. We first prove termination: on every input the algorithm always terminates and never ends up in an infinite loop; then soundness: if the algorithm answers that $C$ is satisfiable with respect to $\mathfrak{T}$ for some $i$-local concept $C$ and a distributed TBox $\mathfrak{T}$ then there actually exists some model of $\mathfrak{T}$ that supports this; and finally we prove completeness: for every concept $C$ that is satisfiable with respect to $\mathfrak{T}$ the algorithm indeed gives a correct answer.

**Theorem 10.** *Given a distributed TBox $\mathfrak{T}$ over $\mathcal{ALC}$ with acyclic bridge graph $\mathrm{G}_{\mathfrak{T}}$ and an $i$-local concept $C$ on the input, the distributed tableaux algorithm for*

---

[4] We prefer the usual naming, and we use the name distributed completion *tree* even if technically it is no longer a tree but rather a forest, that is, a graph composed of several isolated independently rooted trees.

$\sqcap$-**rule:**
If $C_1 \sqcap C_2 \in \mathcal{L}_i(x)$ for some $x \in V_i$ and $\{C_1, C_2\} \nsubseteq \mathcal{L}_i(x)$, and $x$ is not blocked,
then set $\mathcal{L}_i(x) = \mathcal{L}_i(x) \cup \{C_1, C_2\}$.

$\sqcup$-**rule:**
If $C_1 \sqcup C_2 \in \mathcal{L}_i(x)$ for some $x \in V_i$ and $\{C_1, C_2\} \cap \mathcal{L}_i(X) = \emptyset$, and $x$ is not blocked,
then either set $\mathcal{L}_i(x) = \mathcal{L}_i(x) \cup \{C_1\}$ or set $\mathcal{L}_i(x) = \mathcal{L}_i(x) \cup \{C_2\}$.

$\forall$-**rule:**
If $\forall R.C \in \mathcal{L}_i(x)$ for some $x \in V_i$, and there is $R$-successor $y$ of $x$ s.t. $C \notin \mathcal{L}_i(y)$, and $x$ is not blocked,
then set $\mathcal{L}_i(y) = \mathcal{L}_i(y) \cup \{C\}$.

$\exists$-**rule:**
If $\exists R.C \in \mathcal{L}_i(x)$, for some $x \in V_i$ with no $R$-successor $y$ s.t. $C \in \mathcal{L}_i(y)$, and $x$ is not blocked,
then add new node $z$ to $V_i$, add the edge $\langle x, z \rangle$ to $E_i$, and set $\mathcal{L}_i(z) = \{C\}$ and $\mathcal{L}_i(\langle x, z \rangle) = \{R\}$.

$\mathcal{T}$-**rule:**
If $C \sqsubseteq D \in \mathcal{T}$ and for some $x \in V_i$ $\mathrm{nnf}(\neg C \sqcup D) \notin \mathcal{L}_i(x)$, and $x$ is not blocked,
then set $\mathcal{L}_i(x) = \mathcal{L}_i(x) \cup \{\mathrm{nnf}(\neg C \sqcup D)\}$.

$\overset{\sqsupseteq}{\Rightarrow}$-**rule:**
If $G \in \mathcal{L}_j(x)$ for some $x \in V_j$, $i : C \overset{\sqsupseteq}{\Rightarrow} j : G \in \mathfrak{B}$, and there is no $y \in V_i$ s.t. $C \in \mathcal{L}_i(y)$ and $j : x \in r_i(y)$, and $x$ is not blocked,
then add new node $y$ to $V_i$ and set $\mathcal{L}_i(y) = \{C\}$, and set $r_i(y) = \{j : x\} \cup r_j(x)$.

$\overset{\sqsubseteq}{\Rightarrow}$-**rule:**
If $D \in \mathcal{L}_i(x)$ for some $x \in V_i$, $i : D \overset{\sqsubseteq}{\Rightarrow} j : H \in \mathfrak{B}$ and there is $y \in V_j$ s.t. $j : y \in r_i(x)$ and $H \notin \mathcal{L}_j(y)$
then set $\mathcal{L}_j(y) = \mathcal{L}_j(y) \cup \{H\}$.

**Fig. 8.** Tableaux expansion rules for DDL over $\mathcal{ALC}$ under the transitivity requirement. First five rules are standard $\mathcal{ALC}$ tableaux rules. Note that the $\sqcup$-rule is non-deterministic. The final two rules are new and are triggered by bridge rules.

*deciding satisfiability of concepts with respect to a distributed TBox over $\mathcal{ALC}$ for DDL with transitive domain relation always terminates and it is sound and complete.*

*Proof. Termination.* Given a distributed TBox $\mathfrak{T}$ with local TBox $\mathcal{T}_i$ and an $i$-local concept $C$, we ought to prove that the algorithm, once started with $\mathfrak{T}$, $C$ and $i$ on input, eventually terminates. The algorithm initializes the completion tree to $T = \{T_j\}_{j \in I}$ such that $T_j = \langle \{s_0\}, \emptyset, \{s_0 \mapsto \{C\}\}, \emptyset \rangle$ for $j = i$ and $T_j = \langle \emptyset, \emptyset, \emptyset, \emptyset \rangle$ for $j \neq i$. The computation then continues by expanding $\mathcal{T}_i$. If there are no onto-bridge rules ingoing into $\mathcal{T}_i$ the algorithm eventually terminates thanks to subset blocking, this result is known for $\mathcal{ALC}$. If there are some ingoing onto-bridge rules then possibly in some $x \in V_i$ such that $C \in \mathcal{L}_i(x)$ and $C$ appears on a right hand side of an onto-bridge rule, say $k : D \overset{\sqsupseteq}{\Rightarrow} i : C$, the $\overset{\sqsupseteq}{\Rightarrow}$-rule is applied and computation is triggered in $T_k$. By structural subsumption we assume that this computation eventually terminates (the length of the longest incoming path of onto-bridge rules decreased for $\mathcal{T}_k$ compared to $\mathcal{T}_i$, and all such paths are finite because of acyclicity). During this process we possibly get some new concepts in $\mathcal{L}_i(x)$ that are introduced thanks to incoming into-bridge rules that trigger the $\overset{\sqsubseteq}{\Rightarrow}$-rule – but only finitely many. The computation now continues in $x$ and its descendants and possibly the $\overset{\sqsupseteq}{\Rightarrow}$-rule is triggered again in some $y \in V_i$, a descendant of $x$. But thanks to subset blocking, this happens only finitely many times. Hence the algorithm eventually terminates.

*Soundness.* Now that we know that the algorithm terminates, we shall prove that if it answers "$C$ is satisfiable" on input $\mathfrak{T}$, $C$ and $i$, then it also holds that $C$ is satisfiable in $\mathcal{T}_i$ with respect to $\mathfrak{T}$, that is we must show that in such a case there exists a distributed model $\mathfrak{I}$ of $\mathfrak{T}$ such that $C^{\mathcal{I}_i} \neq \emptyset$. Given $\mathfrak{T}$, $C$ and $i$ and let $T$ be the complete and clash-free completion tree that the algorithm has constructed to support the decision. Let us construct the distributed interpretation $\mathfrak{I}$ as follows:

1. Let $\Delta^{\mathcal{I}_i} = V_i$, for each $i \in I$.
2. Let $x \in A^{\mathcal{I}_i}$, for each atomic concept $A \in \mathcal{L}_i(x)$, for each $x \in V_i$ and each $i \in I$.
3. Let $\langle x, y \rangle \in R^{\mathcal{I}_i}$ for each $\langle x, y \rangle \in E_i$ such that $\mathcal{L}_i(\langle x, y \rangle) = R$, for each $i \in I$, if $y$ is not blocked.
4. Let $\langle x, z \rangle \in R^{\mathcal{I}_i}$ for each $\langle x, y \rangle \in E_i$ such that $\mathcal{L}_i(\langle x, y \rangle) = R$, for each $i \in I$, in case that $y$ is blocked by $z$.
5. Let $r_{ij}(x) = y$ for each $x \in V_i$ and for each $j : y \in r_i(x)$.

Please observe that if computation was never triggered within some $T_j$ of $T$ during the run of the algorithm, then $\mathcal{I}_j = \mathcal{I}^\epsilon$. It remains to show that $\mathfrak{I}$ is in fact a model of $\mathfrak{T}$ and $C^{\mathcal{I}_i} \neq \emptyset$. We will first prove the following proposition:

> Given any $i \in I$, for each $E \in \mathcal{L}_i(x)$ (i.e., also for complex concepts) we have $x \in E^{\mathcal{I}_i}$.

This is proved by induction on the structure of $E$. We need to consider the following cases:

1. $E$ is atomic. We know that $x \in E^{\mathcal{I}_i}$ from the construction.
2. $E = \neg E_1$, $E_1$ atomic. Since $T$ is clash-free, $E_1 \notin \mathcal{L}_i(x)$ and by construction $x \notin E_1^{\mathcal{I}_i}$. In that case however $x \in E^{\mathcal{I}_i} = \Delta^{\mathcal{I}_i} \setminus E_1^{\mathcal{I}_i}$.
3. $E = E_1 \sqcap E_2$. Since $T$ is complete, the $\sqcap$-rule is not applicable and hence also $E_1 \in \mathcal{L}_i(x)$ and $E_2 \in \mathcal{L}_i(x)$. By induction we now have that $x \in E_1^{\mathcal{I}_i}$ and $x \in E_2^{\mathcal{I}_i}$ and hence also $x \in E^{\mathcal{I}_i}$.
4. $E = E_1 \sqcup E_2$. Since $T$ is complete, the $\sqcup$-rule is not applicable and hence either $E_1 \in \mathcal{L}_i(x)$ or $E_2 \in \mathcal{L}_i(x)$. By induction we now either have $x \in E_1^{\mathcal{I}_i}$ or we have $x \in E_2^{\mathcal{I}_i}$. In either case however also $x \in E^{\mathcal{I}_i}$.
5. $E = \exists R.E_1$. Since $T$ is complete, the $\exists$-rule is not applicable and hence there must be $y \in V_i$, an R-successor of $x$ such that $E_1 \in \mathcal{L}_i(y)$. By induction $y \in E_1^{\mathcal{I}_i}$ and by construction of $\mathfrak{I}$ $\langle x, y \rangle \in R^{\mathcal{I}_i}$. But that means that $x \in E^{\mathcal{I}_i}$.
6. $E = \forall R.E_1$. Since $T$ is complete, the $\forall$-rule is not applicable and hence for every $y \in V_i$, an R-successor of $x$, we have $E_1 \in \mathcal{L}_i(y)$. By induction $y \in E_1^{\mathcal{I}_i}$ and by construction of $\mathfrak{I}$ $\langle x, y \rangle \in R^{\mathcal{I}_i}$. But that means that $x \in E^{\mathcal{I}_i}$.

Thus we have verified that the local interpretation $\mathcal{I}_i$ is indeed an $\mathcal{ALC}$ interpretation and that $C$ has an instance in this interpretation (since $C \in \mathcal{L}_i(s_0)$). In addition we have in fact also proved that each $i$-local GCI axiom $E_1 \sqsubseteq E_2$ is satisfied by $\mathcal{I}_i$ – from the completeness of $T$, $\mathrm{nnf}(\neg E_1 \sqcup E_2) \in x$, for each $x \in V_i$, and hence $x \in (\mathrm{nnf}(\neg E_1))^{\mathcal{I}_i} \cup E_2^{\mathcal{I}_i}$. It follows that $x \in E_1^{\mathcal{I}_i}$ implies $x \in E_2^{\mathcal{I}_i}$.

It remains to show that $\mathfrak{I}$ is indeed a distributed model of $\mathfrak{T}$ in the adjusted semantics. First, all the bridge rules must be satisfied. Given an onto-bridge rule $k : E_1 \stackrel{\sqsupseteq}{\Rightarrow} l : E_2 \in \mathfrak{B}$, we ought to show that $r_{kl}\big(E_1^{\mathcal{I}_k}\big) \supseteq E_2^{\mathcal{I}_l}$. So let $x \in E_2^{\mathcal{I}_l}$, that is $x \in V_l$ and $E_2 \in \mathcal{L}_l(x)$. But $T$ is complete, $\stackrel{\sqsupseteq}{\Rightarrow}$-rule is not applicable, and hence there must be $y \in V_k$ with $E_1 \in \mathcal{L}_k(y)$ and $l : x \in r_k(y)$. That means however that $y \in E_1^{\mathcal{I}_k}$ and $\langle x, y \rangle \in r_{kl}$, and so $x \in r_{kl}\big(E_1^{\mathcal{I}_k}\big)$. Therefore the bridge rule is satisfied by $\mathfrak{I}$.

Given an into-bridge rule $k : E_1 \stackrel{\sqsubseteq}{\Rightarrow} l : E_2 \in \mathfrak{B}$, we ought to show that $r_{kl}\big(E_1^{\mathcal{I}_k}\big) \subseteq E_2^{\mathcal{I}_l}$. Let $x \in r_{kl}\big(E_1^{\mathcal{I}_k}\big)$. Then there is $y \in V_k$ such that $l : x \in r_l(y)$. But since $T$ is complete, it must be the case that $E_2 \in \mathcal{L}_l(x)$ and hence $x \in E_2^{\mathcal{I}_l}$. Therefore the bridge rule is satisfied by $\mathfrak{I}$.

The last thing in order to verify that $\mathfrak{I}$ is indeed a model of $\mathfrak{T}$ is to show that $\mathfrak{I}$ satisfies the transitivity requirement. We ought to show that for each $k, l, m \in I$, if $y \in r_{kl}(x)$ and $z \in r_{lm}(y)$ then also $z \in r_{km}(x)$. Given the two assumptions we know by construction that $l : y \in r_k(x)$ and $m : z \in r_l(y)$. That means that $x \in V_k$ was initialized after several "chained" applications of the $\stackrel{\sqsupseteq}{\Rightarrow}$-rule starting from $y \in V_l$ and $y$ in turn after several "chained" $\stackrel{\sqsupseteq}{\Rightarrow}$-rule applications starting from $z \in V_m$. In that case however $r_l(y) \subseteq r_k(x)$. Hence $m : z \in r_k(x)$ and so $z \in r_{km}(x)$.

And thus $\mathfrak{I}$ is a distributed model of $\mathfrak{T}$ that satisfies the transitivity condition and $C^{\mathcal{I}_i} \neq \emptyset$ – in other words, $C$ is satisfiable in $\mathcal{T}_i$ with respect to $\mathfrak{T}$.

*Completeness.* Given $\mathfrak{T}$ with index set $I$, a concept $C$ and $i \in I$ such that $C$ is satisfiable in $\mathcal{T}_i$ with respect to $\mathfrak{T}$, we shall prove that the algorithm answers "$C$ is satisfiable", if run on input $\mathfrak{T}$, $C$ and $i$. Let $\mathfrak{I}$ be a distributed model of

$\mathfrak{T}$ with $C^{\mathcal{I}_i} \neq \emptyset$, we know that one must exist. We will simulate the run of the algorithm. During the initialization $T_i$ is set to $\langle \{s_0\}, \emptyset, \{s_0 \mapsto \{C\}\}, \emptyset \rangle$ and all the other $T_j$, $i \neq j$, are set to $\langle \emptyset, \emptyset, \emptyset, \emptyset \rangle$. There is no clash in $T$. We will show by induction (on the number of tableau expansion steps) that after each tableau expansion step the completion tree $T$ is expanded in such a way that no clash is introduced. In order to demonstrate this, we inductively construct an auxiliary mapping $\pi : \bigcup_{i \in I} V_i \to \bigcup_{i \in I} \Delta^{\mathcal{I}_i}$ to track the relation between $\mathfrak{I}$ and $T$. This mapping will keep the property $(*)$:

For each node $x \in V_i$, for each $i \in I$: if $C \in \mathcal{L}_i(x)$ then $\pi(x) \in C^{\mathcal{I}_i}$.

Let $x$ be an arbitrary member of $C^{\mathcal{I}_i}$, place $\pi(s_0) = x$. So, if a tableaux expansion rule is triggered in some $x \in V_i$ of the clash-free completion tree $T$ (from induction hypothesis), we consider several cases, based on the kind of rule that was triggered:

1. $\sqcap$-rule is triggered in $x$ because $E_1 \sqcap E_2 \in \mathcal{L}_i(x)$. Then by induction hypothesis $\pi(x) \in E_1 \sqcap E_2^{\mathcal{I}_i}$. In that case also $\pi(x) \in E_1^{\mathcal{I}_i}$ and $\pi(x) \in E_2^{\mathcal{I}_i}$, and hence the property $(*)$ is maintained after the algorithm adds $E_1$ and $E_2$ to $\mathcal{L}_i(x)$.
2. $\sqcup$-rule is triggered in $x$ because $E_1 \sqcup E_2 \in \mathcal{L}_i(x)$. Then by induction hypothesis $\pi(x) \in E_1 \sqcup E_2^{\mathcal{I}_i}$. In that case however either $\pi(x) \in E_1^{\mathcal{I}_i}$ or $\pi(x) \in E_2^{\mathcal{I}_i}$. Without loss of generality let it be the case that $\pi(x) \in E_1^{\mathcal{I}_i}$. Without loss of generality we assume that the algorithm adds $E_1$ to $\mathcal{L}_i(x)$, as a non-deterministic decision takes place. Hence the property $(*)$ is maintained after this step.
3. $\exists$-rule is triggered in $x$ because it has an $R$-successor $y$, and $\exists R.E_1 \in \mathcal{L}_i(x)$. Then $\pi(x) \in \exists R.E_1^{\mathcal{I}_i}$ and so there is some $y' \in \Delta^{\mathcal{I}_i}$ such that $y' \in E_1^{\mathcal{I}_i}$ and $\langle \pi(x), y' \rangle \in R^{\mathcal{I}_i}$. When the algorithm generates a new node $y$ in this step we set $\pi(y) = y'$ and the property $(*)$ is maintained.
4. $\forall$-rule is triggered in $x$ because it has an $R$-successor $y$, and $\forall R.E_1 \in \mathcal{L}_i(x)$, as a consequence $E_1$ is added to $\mathcal{L}_i(y)$. We know that $y$ was created by an application of the $\exists$-rule and hence $\pi(y)$ is an $R$-successor of $x$ in $\mathcal{I}_i$. But $\mathcal{I}_i$ is a model of $\mathcal{T}_i$ and hence $\pi(y) \in y^{\mathcal{I}_i}$ since from induction hypothesis we know that $\pi(x) \in \forall R.E_1^{\mathcal{I}_i}$. Hence $(*)$ is maintained after this step.
5. $\mathcal{T}$-rule is triggered in $x$, resulting to adding $\mathrm{nnf}(\neg E_1 \sqcup E_2)$ into $\mathcal{L}_i(x)$. Then $(*)$ is maintained as $\mathcal{I}_i$ is a model of $\mathcal{T}_i$ and so it holds that $\pi(x) \in (\mathrm{nnf}(\neg E_1 \sqcup E_2))^{\mathcal{I}_i}$.
6. $\overset{\exists}{\Rightarrow}$-rule is triggered in $x$ because $E_2 \in \mathcal{L}_i(x)$ and $j : E_1 \overset{\exists}{\Rightarrow} i : E_2 \in \mathfrak{B}$. From induction hypothesis, $\pi(x) \in E_2^{\mathcal{I}_i}$, and hence $\pi(x) \in r_{ji}(y')$ for some $y' \in E_1^{\mathcal{I}_j}$. Set $\pi(y) = y'$ for the node $y$ that is newly created in $V_i$ as the result of this expansion step. The label $\mathcal{L}_j(y)$ has been set to $E_1$ but since $y' \in E_1^{\mathcal{I}_j}$ then $(*)$ is maintained.
7. $\overset{\sqsubseteq}{\Rightarrow}$-rule is triggered in $x$ because of $E_1 \in \mathcal{L}_i(x)$, $j : y \in r_i(x)$ and $i : E_1 \overset{\sqsubseteq}{\Rightarrow} j : E_2 \in \mathfrak{B}$, resulting to adding $E_2$ into $\mathcal{L}_j(y)$. From induction hypothesis have $\pi(x) \in E_1^{\mathcal{I}_i}$. Observe that the definition of $\overset{\exists}{\Rightarrow}$-rule and the inductive construction of $\pi(\cdot)$ in previous steps also assure that $j : y \in r_i(x)$ implies

$\pi(y) \in r_{ij}(\pi(x))$. This is because initialization of new $y \in V_j$ always follows incoming $r$-edge and $r$ is transitive because of computational consistency. It follows that $\pi(y) \in E_2^{\mathcal{I}_j}$, since $\mathfrak{I}$ is a distributed model of $\mathfrak{T}$ and the bridge rule assures this. Hence $(*)$ is maintained even after this step.

We already know that the algorithm always terminates. Once this happens, it follows that $T$ is now complete, because no rule is applicable, and clash-free, because the property $(*)$ is maintained all the way up to this point. Hence the algorithm answers "$C$ is satisfiable" and hence the theorem. $\qquad\square$

The algorithm for the original DDL semantics of (Serafini et al. 2005) is obviously truly distributed in the sense that it supports a scenario in which several autonomous reasoning services run independently, one for each local ontology, and communicate by passing queries. While this is also the case for the newly introduced algorithm, it is not necessarily that obvious. In order to clarify this, we provide a message protocol that handles $\overset{\exists}{\Rightarrow}$-rule and $\overset{\sqsubseteq}{\Rightarrow}$-rule execution in a truly distributed fashion, and it also collects the information that the completion tree is complete within the reasoner that has initialized the computation. The algorithm employs three kinds of messages $\mathsf{querySat}(x, r, C)$, $\mathsf{answerSat}(x, C, \mathit{answer})$, and $\mathsf{pushConcept}(x, C)$. In the message based version, the algorithm starts by initializing the local completion tree $T_j$ when asked for satisfiability of some $j$-local concept $C$ with respect to $\mathfrak{T}$. More local completion trees are initialized during the runtime by passing a $\mathsf{querySat}(x, r, C)$ message every time the $\overset{\exists}{\Rightarrow}$-rule is fired. When $\overset{\sqsubseteq}{\Rightarrow}$-rule is fired, the consequences of the into-bridge rule are propagated by passing a $\mathsf{pushConcept}(x, C)$ message. Once a local completion tree $T_k$ is complete this is announced by passing a $\mathsf{answerSat}(x, C, \mathit{answer})$ message to the local reasoner that has triggered the computation in $\mathcal{T}_k$. Detailed specification of the protocol messages is given in Fig. 9.

## 5 Related Work

A distributed tableaux reasoning algorithm for the original semantics of DDL has been introduced by Serafini & Tamilin (2005), Serafini et al. (2005) and implemented in the system DRAGO[5]. This algorithm is based on a fix-point characterization of the original DDL semantics. Given a set of bridge rules $\mathfrak{B}_{ij}$ between $\mathcal{T}_i$ and $\mathcal{T}_j$, define the operator $\mathfrak{B}_{ij}(\cdot)$ as follows:

$$
\mathfrak{B}_{ij}(\mathcal{T}_i) = \left\{ G \sqsubseteq \bigsqcup_{k=1}^{n} H_k \;\middle|\; \begin{array}{l} \mathcal{T}_i \models A \sqsubseteq \bigsqcup_{k=1}^{n} B_k, \\ i : A \overset{\exists}{\Rightarrow} j : G \in \mathfrak{B}_{ij}, \\ i : B_k \overset{\sqsubseteq}{\Rightarrow} j : H_k \in \mathfrak{B}_{ij}, \\ 1 \leq k \leq n \end{array} \right\} \;.
$$

---

[5] DRAGO Reasoner homepage: `http://drago.itc.it/`.

```
querySat(x, r, C):
Send: if $G \in \mathcal{L}_j(x)$ for some $x \in V_j$, $i : C \overset{\exists}{\Rightarrow} j : G \in \mathfrak{B}$, and $x$ is
not blocked, pass the message querySat$(x, r_j(x), C)$ to $T_i$.
Receive: upon receipt of a message querySat$(x, r, C)$ from $T_i$, create
new completion tree $\langle \{s_0\}, \emptyset, \{s_0 \mapsto \{C\}\}, \{s_0 \mapsto \{i : x\} \cup r\} \rangle$
within $T_j$ and start the computation. Record the sent message in
$S_j$.

answerSat(x, C, answer):
Send: if the local completion tree for concept $C$ within $T_j$, that
has been initialized due the message querySat$(x, r, C)$ previously re-
ceived from $T_i$, is now complete, and there are no outgoing messages
recorded in $S_j$, send the message answerSat$(x, C, answer)$ to $\mathcal{T}_i$
with $answer$ set to true if $T_j$ is clash-free and to false otherwise.
Receive: upon receipt of a message answerSat$(x, C, answer)$ from
$\mathcal{T}_i$, remove the message querySat$(x, r, C)$ from $S_j$, that has been
previously send to $T_i$. If $answer$ is false then add $E \sqcap \neg E$ to $\mathcal{L}_j(x)$
for some new concept name $E$.

pushConcept(x, C):
Send: if $D$ has been added to $\mathcal{L}_j(x)$ in the previous step, for some
$x \in V_j$, then for each $j : D \overset{\sqsubseteq}{\Rightarrow} i : H \in \mathfrak{B}$ s.t. $i : y \in r_j(x)$ send
the message pushConcept$(y, H)$ to $T_i$.
Receive: upon receipt of a message pushConcept$(x, C)$ from $\mathcal{T}_i$, add
$C$ to $\mathcal{L}_j(x)$.
```

**Fig. 9.** The message protocol for the newly introduced distributed tableaux algorithm. For each kind of message we specify when the message is sent from some local reasoning service $T_j$ and also what happens when the message is received in some local reasoning service $T_j$ ($T_j$ is always the local reasoning service). An auxiliary data structure $S_j$ is introduced in each $T_j$ in order to track messages that have been sent in order to assure termination.

Consequently another operator, denoted by $\mathfrak{B}$, is defined:

$$\mathfrak{B}\left(\{\mathcal{T}_i\}_{i \in I}\right) = \left\{\mathcal{T}_i \cup \bigcup_{j \neq i} \mathfrak{B}_{ji}\left(\mathcal{T}_j\right)\right\} \ .$$

As showed by (Serafini et al. 2005), the $\mathfrak{B}$ operator always has a fix-point $\mathfrak{B}^*(\mathfrak{T})$ when repeatedly applied on a distributed TBox $\mathfrak{T}$. Moreover, $\mathfrak{T} \models_\epsilon i : \phi$ if and only if the $i$-th component TBox of $\mathfrak{B}^*(\mathfrak{T})$ locally entails $\phi$. Consequently, the standard $\mathcal{SHIQ}$ tableaux reasoning algorithm (Horrocks et al. 1999) is used with one additional bridge rule (see Fig. 10). The unsatisfiability check called by $\mathfrak{B}_{ij}$-rule is done by running the same algorithm again for the $i$-concept that is being checked. Compared to the algorithm introduced in this paper, this algorithm does not keep track of the domain relation $r$ and does not in fact construct a true distributed tableau that would correspond to a particular distributed model of the input concept. Instead it cleverly uses the fix-point characterization and constructs multiple local tableaux in order to guarantee the existence of such a model. Most prominently, all consequences that are added to the triggering node if the unsatisfiability check is successful are estimated prior to the unsatisfiability check is executed. In contrast, in our approach computation in remote ontology is triggered by one message and consequences are announced back to the triggering node once they are computed, possibly by several independent messages. We believe that keeping precise track of all model structures, including the domain relation and dividing the communication into more fine-grained messages may provide better grounds for future optimization. In a sense we may conclude, that in the way how calls are made to the remote tableaux, the original algorithm by Serafini et al. (2005) resembles more a black box approach, whereas the newly introduced algorithm is more like a glass box approach. Another significant difference between the two algorithms is unlike the newly introduced algorithm, the original one is able to handle cyclic bridge rules; a blocking strategy that is employed to achieve this was described by Tamilin (2007, Sect. 5.3.4, p. 91). Finally, it is worth noting that the newly introduced algorithm is easily adjusted to correspond to the original DDL semantics (by setting $r_i(y)$ to $\{j : x\}$ and not to $\{j : x\} \cup r_j(x)$ in the $\stackrel{\exists}{\Rightarrow}$-rule).

> $\mathfrak{B}_{ij}$**-rule:**
> If $G \in \mathcal{L}_j(x)$, $i : A \stackrel{\exists}{\Rightarrow} j : H \in \mathfrak{B}_{ij}$,
> $BH \subseteq \{\langle B_k, H_k \rangle \mid i : B_k \stackrel{\sqsubseteq}{\Rightarrow} j : H_k \in \mathfrak{B}_{ij}\}$,
> $B = \{B \mid \langle B, X \rangle \in BH\}$, $H = \{H \mid \langle Y, H \rangle \in BH\}$, $H \not\subseteq \mathcal{L}_j(x)$
> and $A \sqcap \neg \bigsqcup B$ is unsatisfiable w.r.t $\mathfrak{T}$ in $\mathcal{T}_i$
> then set $\mathcal{L}_j(x) = \mathcal{L}_j(x) \cup \{\bigsqcup H\}$.

**Fig. 10.** The tableaux expansion rule used in the original DDL algorithm (Serafini & Tamilin 2005).

Another distributed ontology framework is called $\mathcal{E}$-connections (Cuenca Grau et al. 2004). Here, inter-ontology roles (called links) are employed instead of concept mapping. A dedicated set of symbols $\epsilon_{ij}$ is used for (directed) links between $\mathcal{T}_i$ and $\mathcal{T}_j$. Links are then used in existential and value restrictions, and complex concepts involving links are formed (for instance the $i$-concept $\exists E.C$ is composed using the link $E \in \epsilon_{ij}$ and the $j$-concept $C$). Semantically, a combined interpretation consists of local interpretations $\mathcal{I}_i$ with non-empty domains $\Delta^{\mathcal{I}_i}$ and each link $E \in \epsilon_{ij}$ is interpreted by $E^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}_i} \times \Delta^{\mathcal{I}_j}$. Reasoning support for $\mathcal{E}$-connections is provided by extending the tableaux reasoning algorithm for $\mathcal{SHIF}(D)$. The two tableaux rules that are essential to handle links are depicted in Fig. 11. There is a notable correspondence between the $\exists_{link}$-rule and our $\overset{\exists}{\Rightarrow}$-rule and also between the $\forall_{link}$-rule and our $\overset{\sqsubseteq}{\Rightarrow}$-rule. This is not that surprising, given the known correspondence between DDL and $\mathcal{E}$-connections (Kutz et al. 2004). Given the nature of $\mathcal{E}$-connections the $\forall_{link}$-rule works exactly the other way around, compared to our $\overset{\sqsubseteq}{\Rightarrow}$-rule, and in this respect the instantiation of a $j$-tree by one application of the $\exists_{link}$-rule and possibly multiple applications of the $\forall_{link}$-rule resembles an unsatisfiability check call from the $\mathfrak{B}_{ij}$-rule of the original DDL algorithm. Remarkably, the exact mechanism how this happens within the $\mathcal{E}$-connections algorithm seems to be more transparent and prone to optimization. The algorithm introduced in this paper in addition keeps track on the domain relation, which is needed to handle subsumption propagation along chains of bridge rules. There is no domain relation within $\mathcal{E}$-connections, and hence no need for such a feature.



---

$\exists_{link}$-**rule:**
**If** $\exists E.C \in \mathcal{L}_i(x)$, $E \in \epsilon_{ij}$, $x$ is not blocked and $x$ has no $E$-successor $y$ with $\mathcal{L}(\langle x, y, \rangle) = \{E\}$,
**then** create a new $E$-successor (a $j$-node) $y$ of $x$ with $\mathcal{L}_j(y) = \{C\}$.
(The new $j$-tree rooted in $y$ will not be expanded until no more rules apply to the $i$-tree.)

$\forall_{link}$-**rule:**
**If** $\forall E.C \in \mathcal{L}_i(x)$, $E \in \epsilon_{ij}$, $x$ is not blocked and there is an $E$-successor $y$ of $x$ such that $C \notin \mathcal{L}_j(y)$
**then** set $\mathcal{L}_j(y)$ to $\mathcal{L}_j(y) \cup \{C\}$.

---

**Fig. 11.** The tableaux expansion rules of the algorithm for $\mathcal{E}$-connections (Cuenca Grau et al. 2004).

Yet another point of view on distributed ontologies is offered by Package-based description logics, or P-DL (Bao et al. 2006). In P-DL, the intuition of importing is pursued: every concept name belongs to some local ontology, but can be also imported to and used by other ontologies in the system. The P-DL semantics uses domain relations. In contrast to other approaches, only one-to-one domain relations are allowed and also compositional consistency (that we have borrowed and applied on DDL) is required. Local domains in P-DL semantics are viewed as partially overlapping, and not disjoint. As a result, some of the

problems known for DDL, such as restricted knowledge propagation between remote ontologies, that we also address in this paper, are not an issue for P-DL according to Bao et al. (2006). On the other hand, in P-DL reasoning over a distributed ontology is always equivalent to reasoning over the union of local ontologies, which is not a goal of DDL. The distributed tableaux algorithm for P-DL is introduced by Bao et al. (2006). It uses $\mathcal{ALC}$ as local language and requires acyclic importing relation. It uses message-based protocol, similar to the ours, and keeps track of the domain relation. Since imported concepts may appear anywhere in the importing ontology, messages are invoked directly from the adjusted $\mathcal{ALC}$-tableaux rules.

A completely different approach to distributed ontology reasoning is taken by Schlicht & Stuckenschmidt (2008), where distributed reasoning algorithm based on resolution techniques is introduced. Yet another distributed ontology framework called Integrated Distributed Description Logics (IDDL) is introduced by Zimmermann (2007), a reasoning algortihm based on reduction to multiple DL knowledge base that are checked for consistency independently is described by Zimmermann & Le Duc (2008).

## 6 Conclusion

In this paper we studied possible adjustment of the standard semantics for DDL in order to improve subsumption propagation between remote ontologies. This can be seen as a completion of the earlier work of Homola (2008). After an analysis of the properties that are (not) enjoyed by the different alternatives, we found that the most appropriate extension is the one that satisfies the so called *transitivity condition*. For this extension we develop a distributed tableaux reasoning algorithm that would decide satisfiability of concepts and subsumption with respect to the adjusted semantics. The algorithm that is introduced in this paper handles chaining onto-bridge rules correctly, but it is unable to cope with chaining into-bridge rules.

As in the case of the tableaux algorithm that is known for the original semantics (Serafini et al. 2005), the newly introduced algorithm is also fully distributed and it permits the scenario where every local ontology is governed by an autonomous reasoning service and these services communicate by passing queries. The local reasoner that has started the computation collects all the answers and makes the final decision at the end. To demonstrate this fact more clearly, we have provided a message-based protocol for the algorithm. Besides the fact that each of these algorithms works under a different semantics, the main distinguishing feature of the newly introduced algorithm is that communication between local reasoners is divided into multiple messages. Computation is sparked in a remote reasoner at some point and both reasoners continue to run independently. If subsumption propagation is proved by the remote reasoner, the local reasoner is acknowledged by a message. We believe that such behaviour may serve as a base for more fine-grained optimization in the future.

The two most prominent open issues put forward by this paper are: extending the algorithm so that it would handle the DDL under compositional consistency semantics in full extent, that is, dealing also with chaining into-bridge rules; and, extending the algorithm towards more expressive DL, since the current version only supports $\mathcal{ALC}$ as the local representation language and requires acyclic concept mapping. Computational handling of distributed knowledge bases in which the concept mapping is not necessarily acyclic is an interesting problem, which is to our best knowledge still unresolved also for the original DDL framework. Remaining research problems that are closely related to this work include practical evaluation of the algorithm by implementation, complexity analysis for the decision problems, and combining, within a unified framework, the current approach with that of Homola (2007), which has addressed the problem of interaction between bridge-rules in DDL.

## Acknowledgement

## Bibliography

Baader, F., Calvanese, D., McGuinness, D., Nardi, D. & Patel-Schneider, P., eds (2003), *The Description Logic Handbook*, Cambridge University Press.

Bao, J., Caragea, D. & Honavar, V. G. (2006), A distributed tableau algorithm for package-based description logics, *in* 'Procs. of CRR 2006', Riva del Garda, Italy.

Berners-Lee, T., Hendler, J. & Lassila, O. (2001), 'The semantic web', *Scientific American* **284**(5), 34–43.

Borgida, A. & Serafini, L. (2003), 'Distributed description logics: Assimilating information from peer sources', *Journal of Data Semantics* **1**, 153–184.

Cuenca Grau, B., Parsia, B. & Sirin, E. (2004), Working with multiple ontologies on the semantic web., *in* 'Procs. of the Third International Semantic Web Conference (ISWC 2004)', Vol. 3298 of *LNCS*, Springer.

Ghidini, C. & Giunchiglia, F. (2001), 'Local models semantics, or contextual reasoning=locality+compatibility', *Artificial Intelligence* **127**(2), 221–259.

Ghidini, C., Serafini, L. & Tessaris, S. (2007), On relating heterogeneous elements from different ontologies, *in* 'Procs. of the Sixth International and Interdisciplinary Conference on Modeling and Using Context (CONTEXT'07)', Vol. 4635 of *LNCS*, Springer.

Homola, M. (2007), Distributed description logics revisited, *in* 'Procs. of 20th International Workshop on Description Logics (DL-2007)', Vol. 250 of *CEUR-WS*.

Homola, M. (2008), Subsumption propagation between remote ontologies in distributed description logic, *in* 'Procs. of the 21st International Workshop on Description Logics (DL2008)', Vol. 353 of *CEUR-WS*.

Horrocks, I., Sattler, U. & Tobies, S. (1999), Practical reasoning for expressive description logics, *in* 'Procs. of the 6th International Conference on Logic for Programming and Automated Reasoning (LPAR'99)', number 1705 *in* 'LNAI', Springer, pp. 161–180.

Kutz, O., Lutz, C., Wolter, F. & Zakharyaschev, M. (2004), '$\mathcal{E}$-connections of abstract description systems', *Artificial Intelligence* **156**(1), 1–73.

Loebe, F. (2006), Requirements for logical modules, *in* 'Procs. of the 1st International Workshop on Modular Ontologies (WoMO'06)', Vol. 232 of *CEUR-WS*.

Schlicht, A. & Stuckenschmidt, H. (2008), Distributed resolution for alc, *in* 'Procs. of the 21st International Workshop on Description Logics (DL2008)', Vol. 353 of *CEUR-WS*.

Serafini, L., Borgida, A. & Tamilin, A. (2005), Aspects of distributed and modular ontology reasoning, *in* 'Procs. of the Nineteenth International Joint Conference on Artificial Intelligence (IJCAI-05)', pp. 570–575.

Serafini, L. & Tamilin, A. (2004), Local tableaux for reasoning in distributed description logics, *in* 'Procs. of the 2004 International Workshop on Description Logics (DL2004)', Vol. 104 of *CEUR-WS*.

Serafini, L. & Tamilin, A. (2005), DRAGO: Distributed reasoning architecture for the semantic web, *in* 'Proc. of the Second European Semantic Web Conference (ESWC'05)', Springer.

Serafini, L. & Tamilin, A. (2006), Distributed instance retrieval in heterogeneous ontologies, *in* 'Proc. of the 2nd Italian Semantic Web Workshop Semantic Web Applications and Perspectives (SWAP'05)', Trento, Italy.

Tamilin, A. (2007), Distributed Ontological Reasoning: Theory, Algorithms, and Applications, PhD thesis, University of Trento.

Zimmermann, A. (2007), Integrated distributed description logics, *in* 'Procs. of 20th International Workshop on Description Logics (DL-2007)', Vol. 250 of *CEUR WS*.

Zimmermann, A. & Le Duc, C. (2008), Reasoning on a Network of Aligned Ontologies, *in* D. Calvanese & G. Lausen, eds, 'Procs. of Web Reasoning and Rule Systems, Second International Conference, RR 2008, Karlsruhe, Germany', Vol. 5341 of *LNCS*, Springer, pp. 43–57.