Subsumption of concepts in \mathcal{FL}_0 for (cyclic) terminologies with respect to descriptive semantics is PSPACE-complete.

Yevgeny Kazakov joint work with Hans de Nivelle

Max-Plank Institut für Informatik, Saarbrücken, Germany

Description Logics

The building blocks in description logics are:

- \mathcal{A} atomic concepts (unary relations)
- \mathcal{R} atomic roles (binary relations)

Description Logics The building blocks in description logics are: • \mathcal{A} – atomic concepts (unary relations) • \mathcal{R} – atomic roles (binary relations) The basic description logic: ALC :=A $C_1 \sqcap C_2$ $C_1 \sqcup C_2$ $\neg C$ $\forall R.C$ $\exists R.C$.

Description Logics The building blocks in description logics are: • \mathcal{A} – atomic concepts (unary relations) • \mathcal{R} – atomic roles (binary relations) The basic description logic: ALC ::=A(x)A $C_1(x) \wedge \overline{C_2(x)}$ $C_1 \sqcap C_2$ $C_1(x) \vee C_2(x)$ $C_1 \sqcup C_2$ $\neg C(x)$ $\neg C$ $\forall R.C$ $\forall y(R(x,y) \to C(y))$ $\exists y (R(x, y) \land C(y))$ $\exists R.C$.

The reasoning tasks in DL (nowledge base (or terminology) \mathcal{T} : $Human \doteq Mammal \sqcap \forall p^{arent} \cdot Human$ $Elephant \doteq Mammal \sqcap \forall p^{arent} \cdot Elephant$ $Adam \doteq Mammal \sqcap \forall p^{arent} \cdot L$

The reasoning tasks in DL (nowledge base (or terminology) \mathcal{T} : $Human = Mammal \sqcap \forall p^{arent}.Human$ $Elephant = Mammal \sqcap \forall p^{arent}. Elephant$ $Adam = Mammal \sqcap \forall p^{arent} \perp$ Basic reasoning task – subsumption checking: $Adam \Box_{\mathcal{T}} Human ?$ $Adam \sqsubseteq_{\tau} Elephant?$ Human $\Box_{\mathcal{T}}$ Elephant?

The reasoning tasks in DL (nowledge base (or terminology) \mathcal{T} : $Human = Mammal \sqcap \forall p^{arent}.Human \leftarrow Cyclic!$ $Elephant = Mammal \sqcap \forall p^{arent} \cdot Elephant \leftarrow Cyclic$ $Adam = Mammal \sqcap \forall parent. \bot$ Basic reasoning task – subsumption checking: $Adam \Box_{\mathcal{T}} Human ?$ $Adam \Box_{\mathcal{T}} Elephant?$ Human $\Box_{\mathcal{T}}$ Elephant?

The reasoning tasks in DL (nowledge base (or terminology) \mathcal{T} : $Human = Mammal \sqcap \forall p^{arent}. Human \leftarrow Cyclic!$ $Elephant = Mammal \sqcap \forall p^{arent} \cdot Elephant \leftarrow Cyclic$ $Adam = Mammal \sqcap \forall p^{arent} \perp$ Basic reasoning task – subsumption checking: parent $Adam \Box_{\mathcal{T}} Human ?$ $Adam \sqsubseteq_{\mathcal{T}} Elephant?$ Human Human $\Box_{\mathcal{T}}$ Elephant? Mamma

Should really all models satisfying \mathcal{T} be considered?

Three types of semantics for terminological cycles (B.Nebel):

Three types of semantics for terminological cycles (B.Nebel):

Descriptive semantics: all interpretations satisfying definitions are considered (*definitions-constrains*):
 BlueNode ÷ Node □ ∀arc.RedNode RedNode ÷ Node □ ∀arc.BlueNode

Three types of semantics for terminological cycles (B.Nebel):

- Descriptive semantics: all interpretations satisfying definitions are considered (*definitions-constrains*):
 BlueNode ÷ Node □ ∀arc.RedNode RedNode ÷ Node □ ∀arc.BlueNode
- If p-semantics: cyclic concepts are interpreted by minimal possible sets (recursive definitions):
 DAGnode = Node □ ∀arc.DAGnode

Three types of semantics for terminological cycles (B.Nebel):

- Descriptive semantics: all interpretations satisfying definitions are considered (*definitions-constrains*):
 BlueNode ÷ Node □ ∀arc.RedNode RedNode ÷ Node □ ∀arc.BlueNode
- If p-semantics: cyclic concepts are interpreted by minimal possible sets (recursive definitions):
 DAGnode = Node □ ∀arc.DAGnode

 g∫p-semantics: cyclic definitions are evaluated in maximal possible way ("all"-definitions):
 MOMO = Man □ ∀child.MOMO

The small terminological language

 $\mathcal{FL}_0 ::= A |$ $C_1 \sqcap C_2 |$ $C_1 \sqcup C_2 |$ $\neg C |$ $\forall R.C |$ $\exists R.C .$

A(x) $C_1(x) \land C_2(x)$ $C_1(x) \lor C_2(x)$ $\neg C(x)$ $\forall y(R(x,y) \rightarrow C(y))$ $\exists y(R(x,y) \land C(y))$

The small terminological language

$\mathcal{FL}_0 ::=$	$A \mid$	A(x)
	$C_1 \sqcap C_2$	$C_1(x) \wedge C_2(x)$
		$C_1(x) \lor C_2(x)$
	$\neg C \mid$	$\neg C(x)$
	$\forall R.C \mid$	$\forall y (R(x, y) \to C(y))$
	$\exists R.C$.	$\exists y (R(x,y) \land C(y))$

Subsumption in \mathcal{FL}_0	Cyclic T-Boxes	Acyclic T-Boxes
descriptive semantics	in PSPACE, PSPACE-hard	
Ifp-semantics	PSPACE-complete	co-NP-complete
gfp-semantics	PSPACE-complete	

The description graph

We focus our attention on terminologies \mathcal{T} of the form:

$$A_i \doteq \forall R_{i,1}.B_{i,1} \sqcap \cdots \sqcap \forall R_{i,k_i}.B_{i,k_i}$$

(1)

with definitions for every atomic concept in \mathcal{T} .

The description graph

We focus our attention on terminologies \mathcal{T} of the form:

$$A_i \doteq \forall R_{i,1}.B_{i,1} \sqcap \cdots \sqcap \forall R_{i,k_i}.B_{i,k_i}$$

(1)

with definitions for every atomic concept in \mathcal{T} . The description graph $\mathcal{G}_{\mathcal{T}}$ is a graph, where:

Nodes are labelled by concept names;

The description graph

We focus our attention on terminologies \mathcal{T} of the form:

 $A_i = \forall R_{i,1}.B_{i,1} \sqcap \cdots \sqcap \forall R_{i,k_i}.B_{i,k_i}$ (1)

with definitions for every atomic concept in \mathcal{T} . The description graph $\mathcal{G}_{\mathcal{T}}$ is a graph, where:

Nodes are labelled by concept names;

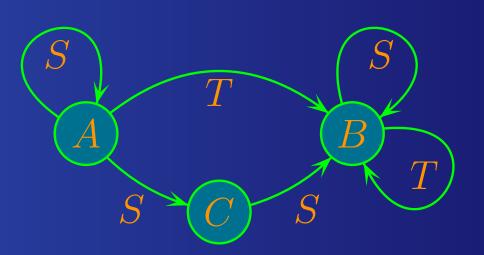
• Oriented edges are labelled by role names such that: the edge *e* comes from the node n_1 to the node n_2 iff

• n_1 is labelled by A, n_2 is labelled by B,

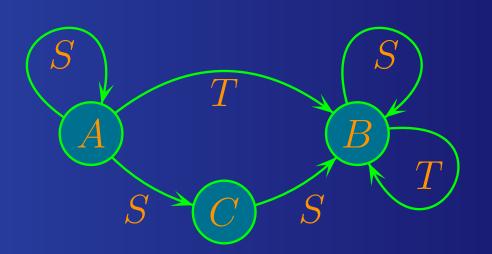
e is labelled by R and

• $A \stackrel{\cdot}{=} \cdots \sqcap \forall R.B \sqcap \cdots \in \mathcal{T}.$

Example Consider the terminology \mathcal{T} : $A \doteq \forall S.A \sqcap \forall T.B \sqcap \forall S.C$ $B \doteq \forall S.B \sqcap \forall T.B$ $C \doteq \forall S.B$ Example Consider the terminology \mathcal{T} : $A \doteq \forall S.A \sqcap \forall T.B \sqcap \forall S.C$ $B \doteq \forall S.B \sqcap \forall T.B$ $C \doteq \forall S.B$

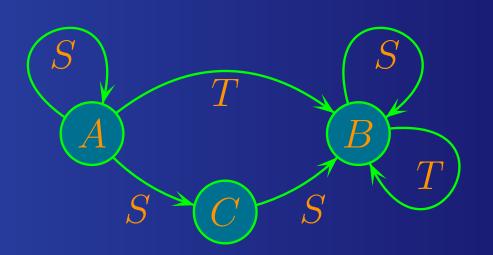


Consider the terminology \mathcal{T} : $A \doteq \forall S.A \sqcap \forall T.B \sqcap \forall S.C$ $B \doteq \forall S.B \sqcap \forall T.B$ $C \doteq \forall S.B$



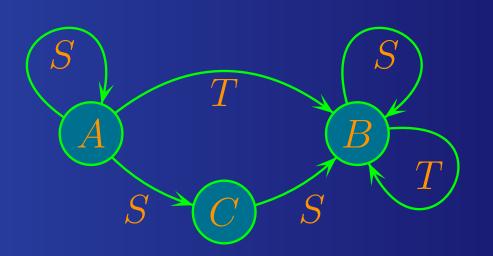
• To check the subsumption $A \sqsubseteq_{\mathcal{T}} B$ assume, there is a model \mathcal{M} with some $a_0 \in A^{\mathcal{M}} \setminus B^{\mathcal{M}}$;

Consider the terminology \mathcal{T} : $A \doteq \forall S.A \sqcap \forall T.B \sqcap \forall S.C$ $B \doteq \forall S.B \sqcap \forall T.B$ $C \doteq \forall S.B$



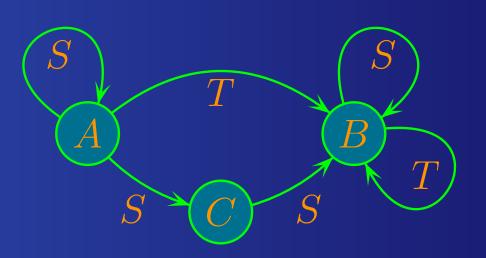
- To check the subsumption $A \sqsubseteq_{\mathcal{T}} B$ assume, there is a model \mathcal{M} with some $a_0 \in A^{\mathcal{M}} \setminus B^{\mathcal{M}}$;
- \mathcal{T} implies: $\neg B \doteq \exists S. \neg B \sqcup \exists T. \neg B$, so, there exists some a_1 with $(a_0, a_1) \in S^{\mathcal{M}}$ or $\in T^{\mathcal{M}}$ and $a_1 \in (\neg B)^{\mathcal{M}}$

Consider the terminology \mathcal{T} : $A \doteq \forall S.A \sqcap \forall T.B \sqcap \forall S.C$ $B \doteq \forall S.B \sqcap \forall T.B$ $C \doteq \forall S.B$



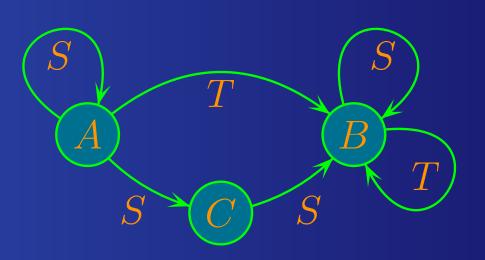
- To check the subsumption $A \sqsubseteq_{\mathcal{T}} B$ assume, there is a model \mathcal{M} with some $a_0 \in A^{\mathcal{M}} \setminus B^{\mathcal{M}}$;
- \mathcal{T} implies: $\neg B \doteq \exists S. \neg B \sqcup \exists T. \neg B$, so, there exists some a_1 with $(a_0, a_1) \in S^{\mathcal{M}}$ or $\in T^{\mathcal{M}}$ and $a_1 \in (\neg B)^{\mathcal{M}}$
- Repeating, we construct the infinite sequence of $a_i \in (\neg B)^{\mathcal{M}}$ with $(a_i, a_{i+1}) \in S^{\mathcal{M}}$ or $\in T^{\mathcal{M}}$.

Consider the terminology \mathcal{T} : $A \doteq \forall S.A \sqcap \forall T.B \sqcap \forall S.C$ $B \doteq \forall S.B \sqcap \forall T.B$ $C \doteq \forall S.B; \quad a_0 \in A^{\mathcal{M}} \setminus B^{\mathcal{M}}$



The sequence of a_i can be one of the following:

Consider the terminology \mathcal{T} : $A \doteq \forall S.A \sqcap \forall T.B \sqcap \forall S.C$ $B \doteq \forall S.B \sqcap \forall T.B$ $C \doteq \forall S.B; \quad a_0 \in A^{\mathcal{M}} \setminus B^{\mathcal{M}}$



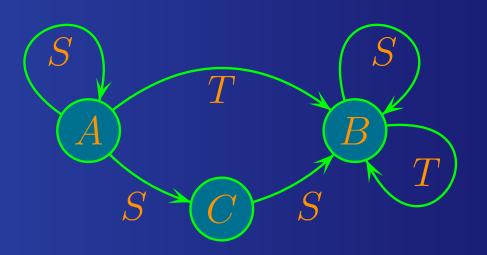
The sequence of *a_i* can be one of the following: *a*₀ *T^M a*₁ ... then *a*₁ ∈ *B^M*;

Consider the terminology \mathcal{T} :

 $A \stackrel{\cdot}{=} \forall S.A \sqcap \forall T.B \sqcap \forall S.C$

 $B \stackrel{\cdot}{=} \forall S.B \sqcap \forall T.B$

 $C \doteq \forall S.B; \qquad a_0 \in A^{\mathcal{M}} \setminus B^{\mathcal{M}}$

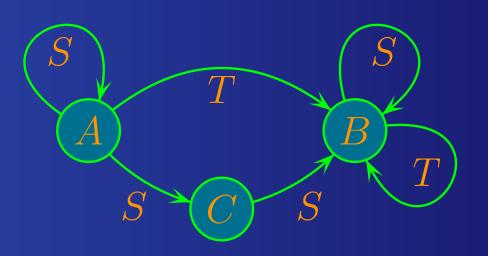


The sequence of a_i can be one of the following:
a₀ T^M a₁ ... then a₁ ∈ B^M;
a₀ S^M a₁ T^M a₂ ... then a₁ ∈ A^M and a₂ ∈ B^M;

Consider the terminology \mathcal{T} :

 $A \stackrel{\cdot}{=} \forall S.A \sqcap \forall T.B \sqcap \forall S.C$ $B \stackrel{\cdot}{=} \forall S.B \sqcap \forall T.B$

 $C \stackrel{\cdot}{=} \forall S.B; \qquad a_0 \in A^{\mathcal{M}} \setminus B^{\mathcal{M}}$

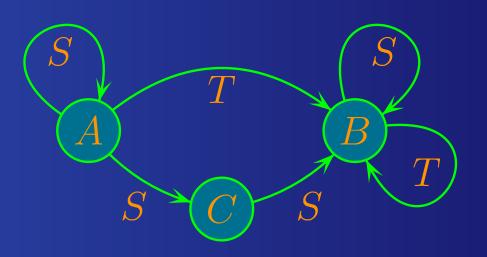


The sequence of a_i can be one of the following:
a₀ T^M a₁ ... then a₁ ∈ B^M;
a₀ S^M a₁ T^M a₂ ... then a₁ ∈ A^M and a₂ ∈ B^M;
a₀ S^M a₁ S^M a₂ ... then a₁ ∈ C^M and a₂ ∈ B^M;

Consider the terminology \mathcal{T} :

 $A \stackrel{\cdot}{=} \forall S.A \sqcap \forall T.B \sqcap \forall S.C$ $B \stackrel{\cdot}{=} \forall S.B \sqcap \forall T.B$

 $C \stackrel{\cdot}{=} \forall S.B; \qquad a_0 \in A^{\mathcal{M}} \setminus B^{\mathcal{M}}$



The sequence of a_i can be one of the following:
a₀ T^M a₁ ... then a₁ ∈ B^M;
a₀ S^M a₁ T^M a₂ ... then a₁ ∈ A^M and a₂ ∈ B^M;
a₀ S^M a₁ S^M a₂ ... then a₁ ∈ C^M and a₂ ∈ B^M;
Since a_i ∈ (¬B)^M, no such M exists, thus A ⊑_T B.

Characterization of subsumption

The following can be shown using the similar arguments: Lemma. (Characterization of concept subsumption)

 $A \sqsubseteq_{\mathcal{T}} B$ iff in the description graph $\mathcal{G}_{\mathcal{T}}$ for every infinite path $B = B_0, \ldots, B_i, \ldots$ there exists an infinite path $A = A_0, \ldots, A_i, \ldots$ with the correspondent labels on the edges such that $A_k = B_k$ for some $k \ge 0$.





Characterization of subsumption

The following can be shown using the similar arguments: Lemma. (Characterization of concept subsumption)

 $A \sqsubseteq_{\mathcal{T}} B$ iff in the description graph $\mathcal{G}_{\mathcal{T}}$ for every infinite path $B = B_0, \ldots, B_i, \ldots$ there exists an infinite path $A = A_0, \ldots, A_i, \ldots$ with the correspondent labels on the edges such that $A_k = B_k$ for some $k \ge 0$.



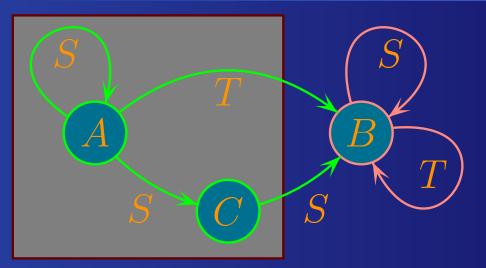
Characterization of subsumption

The following can be shown using the similar arguments: Lemma. (Characterization of concept subsumption)

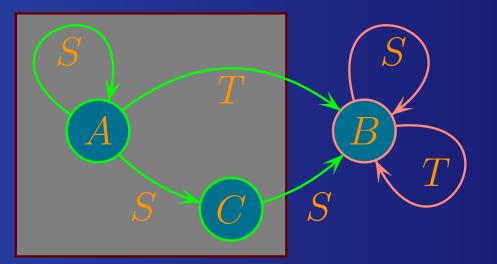
 $A \sqsubseteq_{\mathcal{T}} B$ iff in the description graph $\mathcal{G}_{\mathcal{T}}$ for every infinite path $B = B_0, \ldots, B_i, \ldots$ there exists an infinite path $A = A_0, \ldots, A_i, \ldots$ with the correspondent labels on the edges such that $A_k = B_k$ for some $k \ge 0$.



Consider the terminology \mathcal{T} : $A \doteq \forall S.A \sqcap \forall T.B \sqcap \forall S.C$ $B \doteq \forall S.B \sqcap \forall T.B$ $C \doteq \forall S.B$

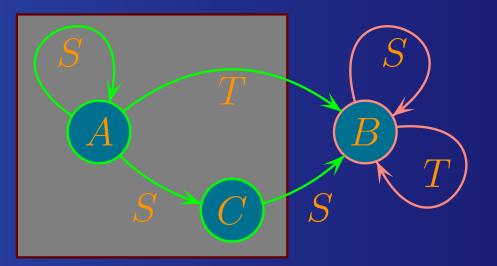


Consider the terminology \mathcal{T} : $A \doteq \forall S.A \sqcap \forall T.B \sqcap \forall S.C$ $B \doteq \forall S.B \sqcap \forall T.B$ $C \doteq \forall S.B$



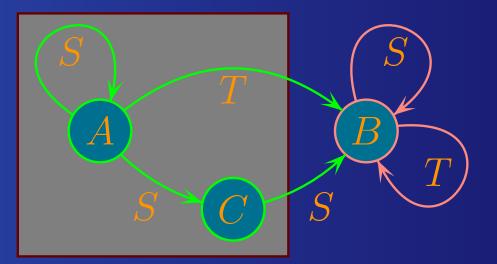
All paths from *B* are passing the node *B* only;

Consider the terminology \mathcal{T} : $A \doteq \forall S.A \sqcap \forall T.B \sqcap \forall S.C$ $B \doteq \forall S.B \sqcap \forall T.B$ $C \doteq \forall S.B$



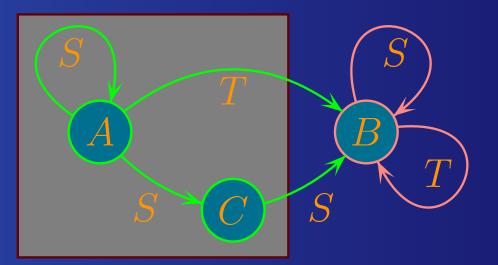
All paths from *B* are passing the node *B* only;
One can find a path for any sequence in {*S*, *T*}*;

Consider the terminology \mathcal{T} : $A \doteq \forall S.A \sqcap \forall T.B \sqcap \forall S.C$ $B \doteq \forall S.B \sqcap \forall T.B$ $C \doteq \forall S.B$



- All paths from B are passing the node B only;
- One can find a path for any sequence in $\{S, T\}^*$;
- Lemma \implies a concept subsumes *B* iff for any infinite sequence in $\{S, T\}^*$ there is a path leading to *B*;

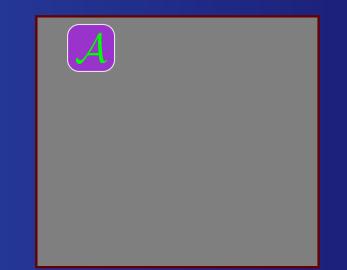
Consider the terminology \mathcal{T} : $A \doteq \forall S.A \sqcap \forall T.B \sqcap \forall S.C$ $B \doteq \forall S.B \sqcap \forall T.B$ $C \doteq \forall S.B$



- All paths from *B* are passing the node *B* only;
- One can find a path for any sequence in $\{S, T\}^*$;
- Lemma \implies a concept subsumes *B* iff for any infinite sequence in $\{S, T\}^*$ there is a path leading to *B*;
- Thus $A \sqsubseteq_{\mathcal{T}} B$ and $C \not\sqsubseteq_{\mathcal{T}} B$

The "hard" instance

Take any NFA \mathcal{A} over Σ :



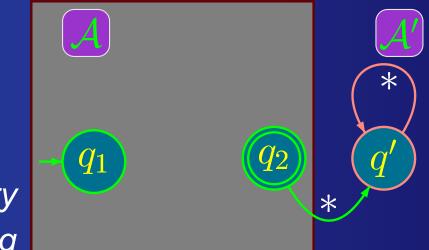
Take any NFA \mathcal{A} over Σ :

- with one initial and one accepting state;
- without blocking states: (every state, except, perhaps, the accepting state has an outcoming transition)

A	q 1	<u>(92</u>)

Take any NFA \mathcal{A} over Σ :

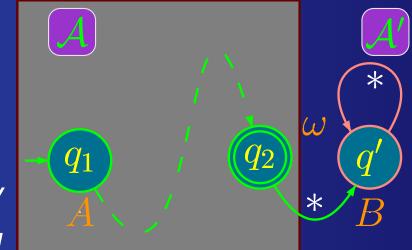
- with one initial and one accepting state;
- without blocking states: (every state, except, perhaps, the accepting state has an outcoming transition)



Add a new state and make the transition to it from the accepting state and itself for any letter in ∑.

Take any NFA \mathcal{A} over Σ :

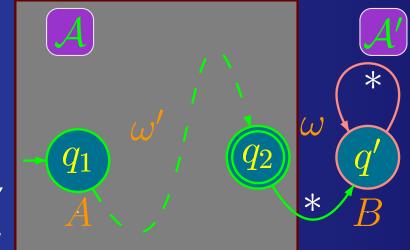
- with one initial and one accepting state;
- without blocking states: (every state, except, perhaps, the accepting state has an outcoming transition)



- Add a new state and make the transition to it from the accepting state and itself for any letter in Σ .
- Consider the correspondent terminology \mathcal{T} : $A \sqsubseteq_{\mathcal{T}} B$ iff there is a path from q_1 through q_2 for any word in Σ^{ω} .

Take any NFA \mathcal{A} over Σ :

- with one initial and one accepting state;
- without blocking states: (every state, except, perhaps, the accepting state has an outcoming transition)



- Add a new state and make the transition to it from the accepting state and itself for any letter in Σ .
- Consider the correspondent terminology \mathcal{T} : $A \sqsubseteq_{\mathcal{T}} B$ iff there is a path from q_1 through q_2 for any word in Σ^{ω} .
- $A \sqsubseteq_{\mathcal{T}} B$ if for any word $\omega \in \Sigma^{\omega}$ there is a finite prefix ω' which is accepted by \mathcal{A} .

We introduce a prefix acceptance condition for NFA: "An infinite word is accepted iff NFA accepts it's finite prefix."

- We introduce a prefix acceptance condition for NFA: "An infinite word is accepted iff NFA accepts it's finite prefix."
- The problem: "Given NFA_p^{ω} check whether it accepts every infinite word". How hard is it?

- We introduce a prefix acceptance condition for NFA: "An infinite word is accepted iff NFA accepts it's finite prefix."
- The problem: "Given NFA^{\u0356} check whether it accepts every infinite word". How hard is it?
- Similar problem for finite automata and buchi automata is known as the universality problem.

- We introduce a prefix acceptance condition for NFA: "An infinite word is accepted iff NFA accepts it's finite prefix."
- The problem: "Given NFA^{\u0356} check whether it accepts every infinite word". How hard is it?
- Similar problem for finite automata and buchi automata is known as the universality problem.
- It is PSPACE-complete

- We introduce a prefix acceptance condition for NFA: "An infinite word is accepted iff NFA accepts it's finite prefix."
- The problem: "Given NFA^{\u0356} check whether it accepts every infinite word". How hard is it?
- Similar problem for finite automata and buchi automata is known as the universality problem.
- It is PSPACE-complete
- It is reasonable to view our problem as the universality problem for automata with prefix acceptance condition.

- We introduce a prefix acceptance condition for NFA: "An infinite word is accepted iff NFA accepts it's finite prefix."
- The problem: "Given NFA_p^{ω} check whether it accepts every infinite word". How hard is it?
- Similar problem for finite automata and buchi automata is known as the universality problem.
- It is PSPACE-complete
- It is reasonable to view our problem as the universality problem for automata with prefix acceptance condition.
- The alternative formulation for the problem:

$$L \cdot \Sigma^{\omega} = \Sigma^{\omega} ?$$

The main result

_emma.(The reduction lemma)

Concept subsumption for descriptive semantics is not easy than the universality problem for automata with prefix acceptance condition.

The main result

_emma.(The reduction lemma)

Concept subsumption for descriptive semantics is not easy than the universality problem for automata with prefix acceptance condition.

Theorem.

The universality problem for $\mathbf{NFA}_{p}^{\omega}$ is PSPACE-complete.

The main result

_emma.(The reduction lemma)

Concept subsumption for descriptive semantics is not easy than the universality problem for automata with prefix acceptance condition.

Theorem.

The universality problem for $\mathbf{NFA}_{p}^{\omega}$ is PSPACE-complete.

Corollary

Subsumption of concepts in \mathcal{FL}_0 for (cyclic) terminologies with respect to descriptive semantics is **PSPACE**-complete.

Conclusions and related work

 We have confirmed the relationship between subsumption problem and automata-theoretic problems.

Conclusions and related work

 We have confirmed the relationship between subsumption problem and automata-theoretic problems.

 New interest in subboolean description logics: *EL* ::= A | C₁ ⊓ C₂ | ∃R.C.

 Franz Baader (2002): Subsumption in *EL* is polynomial.

Conclusions and related work

 We have confirmed the relationship between subsumption problem and automata-theoretic problems.

 New interest in subboolean description logics: *EL* ::= A | C₁ ⊓ C₂ | ∃R.C.

 Franz Baader (2002): Subsumption in *EL* is polynomial.

 Description logics with mixed semantics?
 T.Henzinger, O. Kupferman, R.Majumdar (2003): satisfiability of ∀MC is PSPACE-complete, satisfiability of ∃MC is NP-complete.
 However, implication problem (~ subsumption) is still EXPTIME.

Thank you!

Thank You!