Devoir Maison : Description Logics

Date limite : vendredi 2 décembre, début du cours (10h15)

Question 1 : Basic notions

Consider the following interpretation \mathcal{I} :

$$\Delta^{\mathcal{I}} = \{e_1, e_2, e_3, e_4\} \quad A^{\mathcal{I}} = \{e_1, e_3\} \quad B^{\mathcal{I}} = \{e_4\}$$
$$G^{\mathcal{I}} = \{e_2, e_3\} \quad s^{\mathcal{I}} = \{(e_1, e_2), (e_1, e_4), (e_2, e_3), (e_3, e_4), (e_3, e_1)\}$$

(a) For each of the following concepts C, give the corresponding set $C^{\mathcal{I}}$:

$$A \sqcap \neg B \qquad A \sqcup \neg G \qquad \forall s.G \qquad \exists s^-.(A \sqcap \exists s. \neg G) \qquad \exists s.\forall s. \bot$$

(b) State which of the following inclusions is satisfied in \mathcal{I} .

$$G \sqsubseteq \neg B \qquad \qquad A \sqcap \neg B \sqsubseteq \exists s^-.(A \sqcap \exists s. \neg G) \qquad \exists s. \forall s. \bot \sqsubseteq \forall s. G$$

(c) For each of the following Boolean CQs, state whether it is satisfied in the interpretation \mathcal{I} :

$$\exists x, y, z.G(x) \land s(x, y) \land s(z, y) \land G(z) \qquad \exists x, y, z.G(x) \land s(x, y) \land s(z, y) \land B(z) \\ \exists x, y.A(x) \land s(x, y) \land A(y) \land s(y, x) \qquad \exists x, y, z.s(x, z) \land s(z, y) \land A(x) \land A(y)$$

Briefly justify your answers.

Question 2 : Acyclic terminologies and tableaux algorithm

A terminology is a special kind of TBox consisting only of concept definitions

$$A \equiv C$$

where A is a concept name and C a (possibly complex) concept, and no two concept definitions have the same atomic concept on the left. A concept definition $A \equiv C$ corresponds to the pair of inclusions $A \sqsubseteq C$ and $C \sqsubseteq A$, and so it is satisfied by \mathcal{I} if $A^{\mathcal{I}} = C^{\mathcal{I}}$.

Given a terminology \mathcal{T} and concept C, the unfolding of C w.r.t. \mathcal{T} is obtained by applying the following operation as long as possible:

• replace any concept name A such that $A \equiv D \in \mathcal{T}$ by D

For example, the unfolding of $A \sqcap \exists r.B$ w.r.t. the terminology

$$\{A \equiv \forall s. \forall r. B \quad B \equiv E \sqcap F \quad F \equiv \neg G\}$$

is the concept $\forall s. \forall r. (E \sqcap \neg G) \sqcap \exists r. (E \sqcap \neg G)$. Note that for some terminologies and concepts, the unfolding is not well defined as the replacement process continues forever, e.g. take C = A and $\mathcal{T} = \{A \equiv \exists R.A\}$.

Given a terminology \mathcal{T} , we say that A directly uses B if \mathcal{T} contains a concept definition $A \equiv C$ such that B occurs in C. We say that A uses B (w.r.t. \mathcal{T}) if A directly uses B or if there is a concept name B' such that A uses B' and B' directly uses B. We call a terminology \mathcal{T} acyclic if there is no concept name that uses itself.

We consider the following procedure for testing satisfiability of an \mathcal{ALC} concept C w.r.t. acyclic \mathcal{ALC} :

Step 1 Compute the unfolding U of C w.r.t. \mathcal{T} .

Step 2 Put U into negation normal form.

Step 3 Run CSat on U, and output the same answer as CSat.

(a) Use the above procedure to test satisfiability of the concepts:

1.
$$D \sqcap H$$

2. $\neg D \sqcap E$

with respect to the following acyclic terminology:

$$\{B \equiv A \sqcup \forall r. \neg A \quad D \equiv \exists r. B \sqcap \exists r. E \quad E \equiv \exists r. A \quad H \equiv \forall r. \neg A\}$$

Show the steps, indicate which tableau rules were used, and state how you reached the final result (satisfiable or not).

- (b) Prove that if \mathcal{T} is an acyclic terminology and C is any concept, then the unfolding of C w.r.t. \mathcal{T} is well defined (i.e. the unfolding process terminates).
- (c) What can be said about the size of the unfolding of a concept w.r.t. an acyclic terminology? Either argue that the unfolding can be computed in polynomial time, or explain why this is not always possible.
- (d) Briefly explain why the above procedure is a decision procedure (i.e., sound, complete, and terminating) for \mathcal{ALC} concept satisfiability w.r.t. acyclic terminologies. You may assume the properties of the NNF transformation and CSat algorithm seen in class.

Question 3: Conservative extensions

We call a TBox \mathcal{T}_2 a conservative extension of \mathcal{T}_1 if

- $sig(\mathcal{T}_1) \subseteq sig(\mathcal{T}_2)$
- every model of \mathcal{T}_2 is a model of \mathcal{T}_1
- for every model \mathcal{I}_1 of \mathcal{T}_1 , there exists a model \mathcal{I}_2 of \mathcal{T}_2 such that \mathcal{I}_1 and \mathcal{I}_2 interpret all concept and role names in $sig(\mathcal{T}_1)$ identically.

where $sig(\mathcal{T})$ denotes the signature of \mathcal{T} , i.e. the set of concept and role names occurring in \mathcal{T} .

- (a) Let $\mathcal{T}_1 = \{A \sqsubseteq B, E \sqsubseteq \exists s.D\}, \mathcal{T}_2 = \mathcal{T}_1 \cup \{A \sqsubseteq G\}, \mathcal{T}_3 = \mathcal{T}_1 \cup \{A \sqsubseteq E\}, and \mathcal{T}_4 = \{A \sqsubseteq B, D \sqsubseteq E\}.$ Determine which of the TBoxes $\mathcal{T}_2, \mathcal{T}_3, and \mathcal{T}_4$ are conservative extensions of \mathcal{T}_1 . Explain your answer.
- (b) Prove the following statement: if \mathcal{T}_1 and \mathcal{T}_2 are \mathcal{EL} -TBoxes, \mathcal{T}_2 is a conservative extension of \mathcal{T}_1 , and C and D are \mathcal{EL} -concepts containing only concept and role names from $sig(\mathcal{T}_1)$, then

$$\mathcal{T}_1 \models C \sqsubseteq D \quad \text{iff } \mathcal{T}_2 \models C \sqsubseteq D$$

(c) Consider the analogous statement for conjunctive queries: if \mathcal{T}_1 and \mathcal{T}_2 are \mathcal{EL} -TBoxes, \mathcal{T}_2 is a conservative extension of \mathcal{T}_1 , and $q(\vec{x})$ is a CQ containing only concept and role names from $sig(\mathcal{T}_1)$, then

$$\operatorname{cert}(q, (\mathcal{T}_1, \mathcal{A})) = \operatorname{cert}(q, (\mathcal{T}_2, \mathcal{A}))$$

for every ABox using only concept and role names from $sig(\mathcal{T})$. Does this statement hold? Justify your answer.

(d) What if we consider \mathcal{ALC} TBoxes rather than \mathcal{EL} TBoxes? Explain whether this changes your answers to parts (b) and (c).