Exercises for the Course "Ontology-Mediated Query Answering"

Rough grading scheme: Exercises 1-5 to get 3, also do Exercises 6 (a)-(c) and 7(a)-(b) to get 4, and complete the whole exercise sheet for 5.

Exercise 1 (Basic notions). Consider the following interpretation $\mathcal{I}$:

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

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

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

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

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

(c) For each of the following Boolean[1] CQs, state whether it is satisfied in $\mathcal{I}$:

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

Briefly justify your answers.

Exercise 2 (CQs vs UCQs). Show that for any DL-Lite$_{\mathcal{R}}$ or $\mathcal{EL}$ KB $\mathcal{K}$, and every Boolean UCQ $q_1 \vee \ldots \vee q_n$, we have $\mathcal{K} \models q_1 \vee \ldots \vee q_n$ iff $\mathcal{K} \models q_i$ for some $1 \leq i \leq n$. Does the same hold for $\mathcal{ALC}$ KBs? Justify your answer.

Exercise 3 (Reducing IQs to KB satisfiability). Give a (short) proof of the following statement: $(\mathcal{T}, \mathcal{A}) \models C(b)$ if and only if the KB $(\mathcal{T}, \mathcal{A} \cup \{\neg C(b)\})$ is unsatisfiable.

Exercise 4 (Instance queries in $\mathcal{ALC}$). Consider the following $\mathcal{ALC}$ knowledge base:

$$\mathcal{T} = \{D \sqsubseteq \forall s.E, E \sqsubseteq \forall u.F, \exists r.B \sqsubseteq D, \exists u.\neg B \sqsubseteq E\}$$
$$\mathcal{A} = \{r(a, b), s(a, c), u(c, b)\}$$

Show that $\mathcal{T}, \mathcal{A} \models F(b)$ by using the reduction of IQ answering to satisfiability together with the tableau algorithm KBSat (and optimizations) presented in the course. It may be helpful to first convince yourself why this entailment holds, in order to guide your choices of which tableau rules to apply first (and avoid exploring more branches than needed).

Exercise 5 (Query rewriting in DL-Lite). Consider the DL-Lite$_{\mathcal{R}}$ TBox:

$$\mathcal{T} = \{B \sqsubseteq \exists r, \exists r^- \sqsubseteq \exists p, \exists p^- \sqsubseteq A, A \sqsubseteq \exists u, \exists u^- \sqsubseteq H, \exists s \sqsubseteq D, D \sqsubseteq \neg H, r \sqsubseteq s^-, p \sqsubseteq s\}$$

(a) First list all DL-Lite$_{\mathcal{R}}$ inclusions that are entailed from $\mathcal{T}$.

(b) Use (a) to construct a rewriting of the IQ $D(x)$ w.r.t. $\mathcal{T}$ and consistent ABoxes.

(c) Use (a) to construct a rewriting of unsatisfiability w.r.t. $\mathcal{T}$.

(d) Construct the canonical model of $(\mathcal{T}, \mathcal{A})$ where $\mathcal{A} = \{B(a), p(a, c)\}$.

(e) Give an FO-rewriting of the Boolean CQ

$$q = \exists x, y, z, z' \; r(x, y) \wedge s(y, z) \wedge F(z) \wedge u(z, z')$$

w.r.t. consistent ABoxes. Please explain how you obtained your rewriting.

[1] A Boolean CQ is a CQ whose variables are all existentially quantified.

Exercise 6 (Models and query answering in $\mathcal{EL}$).

(a) Create an $\mathcal{EL}$ TBox $\mathcal{T}_n$ that uses concept names $A_0, \ldots, A_n$ and role names r and s such that the canonical model of the KB $(\mathcal{T}_n, \{A_0(a)\})$ is a binary tree of depth n, where each element at depth k belongs to concept A_k, and every non-leaf element belongs to one r-successor and one s-successor.

(b) Briefly explain how the rewriting approach introduced in the course would handle the input query $\exists x.A_n(x)$ w.r.t. the TBox $\mathcal{T}_n$.

(c) Next give the compact canonical model for $(\mathcal{T}_n, \{A_0(a)\})$ and an example of a Boolean CQ that is satisfied in the compact canonical model but not in the canonical model.

(d) Now consider the following restricted classes of Boolean CQs:

- directed linear Boolean CQs, whose role atoms form a chain $r_1(x_0, x_1), r_2(x_1, x_2), \ldots r_m(x_{m-1}, x_m)$
- acyclic Boolean CQs, whose role atoms induce an acyclic undirected graph (whose edges are the pairs $\{t, t'\}$ such that the query contains an atom of the form $r(t, t')$ or $r(t', t)$)

For each of the preceding classes, decide whether the following statement holds: for every $\mathcal{EL}$ KB and every query q from the class, q is satisfied in the compact canonical model of $\mathcal{K}$ iff it is satisfied in the canonical model of $\mathcal{K}$. Explain your answers.

Exercise 7 (Repairs). We say that an ABox $\mathcal{A}$ is $\mathcal{T}$-consistent if $(\mathcal{T}, \mathcal{A})$ is satisfiable. Recall that a subset $\mathcal{R} \subseteq \mathcal{A}$ is a repair of KB $(\mathcal{T}, \mathcal{A})$ if (i) $\mathcal{R}$ is $\mathcal{T}$-consistent, and (ii) there is no $\mathcal{T}$-consistent subset $\mathcal{B} \subseteq \mathcal{A}$ with $\mathcal{R} \subsetneq \mathcal{B}$.

(a) How many repairs can DL-Lite$_\mathcal{R}$ KBs possess: one, polynomially many, or exponentially many? Justify your answer.

(b) Repeat the previous question for $\mathcal{EL}$ and $\mathcal{ALC}$ KBs.

(c) Focusing on DL-Lite$_\mathcal{R}$ KBs, sketch polynomial-time algorithms (w.r.t. combined complexity) for the following problems:

- deciding whether a given subset of the ABox is a repair,
- computing a single repair,
- deciding whether an assertion belongs to every repair.

(d) Consider now $\mathcal{EL}_\bot$ (which extends $\mathcal{EL}$ with $\bot$) and show that the problem of deciding whether an assertion belongs to every repair of an $\mathcal{EL}_\bot$ KB is coNP-hard in data complexity.