Knowledge Representation and Processing

Course Lecturer: Yizheng Zhao

June 25, 2022

⋆ **This examination contributes to 50% of the total mark of the course.**

Question 1. Basics of KR and Ontologies

In 1993, Gruber originally defined the notion of an ontology as "explicit specification of a conceptualization". In 1997, Borst defined an ontology as "formal specification of a shared conceptualization". This definition additionally required that the conceptualization should express a shared view between several parties, a consensus rather than an individual view. Also, such conceptualization should be expressed in a (formal) machine readable format. In 1998, Studer et al. merged these two definitions stating that:"An ontology is a formal, explicit specification of a shared conceptualization."

- Do you think this is a good definition for ontology? Why and why not? (3 marks)

Model Solution. Ontologies can be interpreted by computers thanks to a formal (logical) semantics that defines the terms and logical statements using the usual Tarski-style set-theoretic semantics, which enables automated reasoning — the domain is interpreted as a set of elements, an individual as an element in the domain, a class as a subset of the domain, and a relationship as a pair of elements in the domain. This provides human users and computers with a shared understanding of domain knowledge.

Marking Scheme. Award 2 marks when students have, somehow, mentioned the idea of "ontologies are equipped with formal semantics", and unlock the remaining 1 mark when students have (briefly) described the Tarski-style set-theoretic semantics.

Question 2. Description Logics as KR formalism

- If you could have only $\exists$-*restriction* or $\forall$-*restriction* in your ontology, which, from a modelling perspective, would be preferable, and why? (3 marks)

Model Solution. From a modelling perspective, existential restrictions are preferable because they are generally more committal. Universal restrictions can be satisfied in models where the instance of the restriction has no relations (i.e., P only C is true if having no P successors). Such individuals are not what we commonly expect in our models. Such vacuous definitions are hard to detect without existentials (though we could achieve a similar effect with an ABox). The lack of universals means we cannot express closure axioms for our properties thus we cannot prevent extra properties with values outside the permitted range. But, in general, such models are easier to ignore than those with missing values.

Marking Scheme. Award 4 marks when students answered with existential restrictions and unlock the remaining when they also gave a good argument. However, answering with universal restrictions does not mean loss of all marks. Award 4 marks as long as a good argument is found, but in this case the marks are capped at 4.

Question 3. DL Syntax and Semantics
Consider the interpretation $\mathcal{I}$ defined by

$$\begin{aligned} \Delta^{\mathcal{I}} &= \{a, b, c, d\} \\ P^{\mathcal{I}} &= \{a, b, c\} \\ Q^{\mathcal{I}} &= \{c\} \\ r^{\mathcal{I}} &= \{(a, c), (b, c), (b, d)\} \end{aligned}$$

Model Solution.

Determine the following sets (3 marks):

- $(P \sqcap \exists r.Q)^{\mathcal{I}} = \{a, b\}$
- $(\forall r.Q)^{\mathcal{I}} = \{a, c, d\}$
- $(\exists r.Q \sqcap \forall r.\neg Q)^{\mathcal{I}} = \emptyset$
- $(\neg P \sqcup \neg \exists r.Q)^{\mathcal{I}} = \{c, d\}$
- $(\exists r.\forall r.\bot)^{\mathcal{I}} = \{a, b\}$
- $(\forall r.(\forall r.\top \sqcap (P \sqcup \neg P)))^{\mathcal{I}} = \{a, b, c, d\}$

Question 4. Tableau for $\mathcal{ALC}$-Concept
Consider the $\mathcal{ALC}$-concept:

$$C = \neg A \sqcap \exists r.(A \sqcup B) \sqcap \forall r.\neg B$$

- Apply the $\mathcal{ALC}$-tableau algorithm to C to determine whether C is satisfiable or not. In your answer, show how the completion rules are applied step by step to the constraint system $x : C$. If C is satisfiable, construct an interpretation $\mathcal{I}$ satisfying C. (6 marks)

Model Solution.

- As C is already in NNF, we immediately apply the ALC-Tableau algorithm to it.

  The ALC-Tableau algorithm starts with:
  $S_0 = \{x : \neg A \sqcap \exists r.(A \sqcup B) \sqcap \forall r.\neg B\}$

  An application of the $\sqcap$-rule gives:
  $S_1 = S_0 \cup \{x : \neg A, x : \exists r.(A \sqcup B), x : \forall r.\neg B\}$

  An application of the $\exists$-rule gives:
  $S_2 = S_1 \cup \{(x, y) : r, y : A \sqcup B\}$

  An application of the $\sqcup$-rule gives:
  $S_3 = S_2 \cup \{y : A\}$ or $S_3^* = S_2 \cup \{y : B\}$

  An application of the $\forall$-rule gives:
  $S_4^* = S_3^* \cup \{y : \neg B\}$

  Clash obtained, thus we proceed with the other branch:

  An application of the $\forall$-rule gives:
  $S_4 = S_3 \cup \{y : B\}$

No rule is applicable to S_4 and S_4 contains no clash. Thus, C is *satisfiable*. A model $\mathcal{I}$ of C is given by:

$\Delta^{\mathcal{I}} = \{x, y\}$

$A^{\mathcal{I}} = \{y\}$

$B^{\mathcal{I}} = \{y\}$

$r^{\mathcal{I}} = \{(x, y)\}$

Question 5. Tableau for $\mathcal{ALC}$ without TBox

Use the $\mathcal{ALC}$-Tableaux algorithm to determine whether

$$\emptyset \models \forall r.A \sqsubseteq \exists r.A$$

In words: determine whether the GCI $\forall r.A \sqsubseteq \exists r.A$ follows from the empty TBox (without TBox). (5 marks)

Model Solution.

- To show $\emptyset \models \forall r.A \sqsubseteq \exists r.A$ is to show that there is no an interpretation $\mathcal{I}$ such that $\mathcal{I} \not\models \forall r.A \sqsubseteq \exists r.A$. We prove by contradiction. We assume that there exists such an interpretation $\mathcal{I}$, that is, there exists an element $d \in \Delta^{\mathcal{I}}$ such that $d \in (\forall r.A)^{\mathcal{I}}$ but $d \notin (\exists r.A)^{\mathcal{I}}$, that is, $d \in (\forall r.A \sqcap \neg\exists r.A)^{\mathcal{I}}$. Therefore, our task is to check whether $\forall r.A \sqcap \neg\exists r.A$ is satisfiable.

  As $\forall r.A \sqcap \neg\exists r.A$ is not in NNF yet, the first step is to transform it into NNF using the transformation rules. This gives $\forall r.A \sqcap \forall r.\neg A$.

  The ALC-Tableau algorithm starts with:
  $S_0 = \{x : \forall r.A \sqcap \forall r.\neg A\}$

  An application of the $\sqcap$-rule gives:
  $S_1 = S_0 \cup \{x : \forall r.A, x : \forall r.\neg A\}$

  No rule is applicable to S_1 and S_1 contains no clash. Thus, $\forall r.A \sqcap \neg\exists r.A$ is *satisfiable*, which means $\emptyset \not\models \forall r.A \sqsubseteq \exists r.A$.

Question 6. Tableau for $\mathcal{ALC}$ with TBox

Consider the following $\mathcal{ALC}$-TBox $\mathcal{T}$:

$$\mathsf{Vegan} \equiv \mathsf{Person} \sqcap \forall \mathsf{eats}.\mathsf{Plant}$$
$$\mathsf{Vegetarian} \equiv \mathsf{Person} \sqcap \forall \mathsf{eats}.(\mathsf{Plant} \sqcup \mathsf{Dairy})$$

We want to know if $\mathcal{T} \models \mathsf{Vegan} \sqsubseteq \mathsf{Vegetarian}$. (5 marks)

Model Solution.

This amounts to showing $\mathcal{T} \models \mathsf{Person} \sqcap \forall \mathsf{eats}.\mathsf{Plant} \sqsubseteq \mathsf{Person} \sqcap \forall \mathsf{eats}.(\mathsf{Plant} \sqcup \mathsf{Dairy})$ (**unfolding**). As in Question 5, our task is to check whether $\forall \mathsf{eats}.\mathsf{Plant} \sqcap \neg\forall \mathsf{eats}.(\mathsf{Plant} \sqcup \mathsf{Dairy})$ is satisfiable. Since $\forall \mathsf{eats}.\mathsf{Plant} \sqcap \neg\forall \mathsf{eats}.(\mathsf{Plant} \sqcup \mathsf{Dairy})$ is not in NNF yet, the first step is to transform it into NNF using the transformation rules. This gives $\forall \mathsf{eats}.\mathsf{Plant} \sqcap \exists \mathsf{eats}.(\neg\mathsf{Plant} \sqcap \neg\mathsf{Dairy})$.

The ALC-Tableau algorithm starts with:
$S_0 = \{x : \forall \mathsf{eats}.\mathsf{Plant} \sqcap \exists \mathsf{eats}.(\neg\mathsf{Plant} \sqcap \neg\mathsf{Dairy})\}$

An application of the $\sqcap$-rule gives:
$S_1 = S_0 \cup \{x : \forall \text{eats.Plant}, x : \exists \text{eats.}(\neg\text{Plant} \sqcap \neg\text{Dairy})\}$

An application of the $\exists$-rule gives:
$S_2 = S_1 \cup \{(x, y) : \text{eats}, y : \neg\text{Plant} \sqcap \neg\text{Dairy}\}$

An application of the $\sqcap$-rule gives:
$S_3 = S_2 \cup \{y : \neg\text{Plant}, y : \neg\text{Dairy}\}$

An application of the $\forall$-rule gives:
$S_4 = S_3 \cup \{y : \text{Plant}\}$

Clash, thus $\forall\text{eats.Plant} \sqcap \neg\forall\text{eats.}(\text{Plant} \sqcup \text{Dairy})$ is unsatisfiable, and $\mathcal{T} \models \text{Vegan} \sqsubseteq \text{Vegetarian}$.

Question 7. Reasoning in $\mathcal{ALC}$
Consider the following $\mathcal{ALC}$-TBox $\mathcal{T}$:

$$\begin{aligned}
\text{Professor} &\sqsubseteq \exists\text{work_for.School} \\
\text{Professor} &\sqsubseteq \exists\text{supervises.Research_Group} \\
\text{Professor} &\sqsubseteq \forall\text{supervises.Research_Student} \\
\text{School} &\sqsubseteq \exists\text{affiliated_with.University}
\end{aligned}$$

- Define a model $\mathcal{I}$ of $\mathcal{T}$ in which Professor is satisfiable.
- Add $\text{ResearchGroup} \sqcap \text{ResearchStudent} \sqsubseteq \bot$ to $\mathcal{T}$. Is the resulting $\mathcal{T}$ consistent? Why? (5 marks)

Model Solution.

The interpretation $\mathcal{I}$ that is a model of $\mathcal{T}$ with $\text{Professor} \neq \emptyset$ is given by setting

$\Delta^{\mathcal{I}} = \{a, b, c, d\}$,

$\text{Professor}^{\mathcal{I}} = \{a\}$,

$\text{School}^{\mathcal{I}} = \{b\}$,

$\text{Research\_Group}^{\mathcal{I}} = \{c\}$,

$\text{Research\_Student}^{\mathcal{I}} = \{c\}$,

$\text{University}^{\mathcal{I}} = \{d\}$,

$\text{works\_for}^{\mathcal{I}} = \{(a, b)\}$,

$\text{supervises}^{\mathcal{I}} = \{(a, c)\}$,

$\text{affiliated\_with}^{\mathcal{I}} = \{(b, d)\}$.

The resulting TBox is satisfiable. For example, every $\mathcal{I}$ in which $\text{Professor}^{\mathcal{I}} = \emptyset$ and $\text{School}^{\mathcal{I}} = \emptyset$ is a model of the extended TBox. Note, however, that Professor is not satisfiable anymore because Research_Group and Research_Student are disjoint (according to the new inclusion) but every Professor supervises a Research_Group and any such Research_Group is, by the value restriction in the third inclusion, a Research_Student.

Question 8. Ontology-Based Data Access

Consider the following $\mathcal{ALC}$-TBox $\mathcal{T}$:

$$
\begin{aligned}
\mathsf{Cat} \sqcap \mathsf{Dog} \sqcap \mathsf{Bird} &\sqsubseteq \bot \\
\mathsf{Cat} &\sqsubseteq \mathsf{Mammal} \\
\mathsf{Dog} &\sqsubseteq \mathsf{Mammal} \\
\exists \mathsf{breatheWith}.\mathsf{Lung} &\sqsubseteq \mathsf{Mammal} \\
\mathsf{Mammal} &\sqsubseteq \forall \mathsf{breatheWith}.\mathsf{Lung} \\
\top &\sqsubseteq \exists \mathsf{breatheWith}.\mathsf{Lung}
\end{aligned}
$$

Consider the following $\mathcal{ALC}$-ABox $\mathcal{A}$:

$$
\begin{aligned}
&\mathsf{Cat}(\mathsf{paper}) \\
&\neg \mathsf{Dog}(\mathsf{cotton\_candy}) \\
&\mathsf{Bird}(\mathsf{polly}) \\
&\mathsf{breatheWith}(\mathsf{dudu}, \mathsf{krp})
\end{aligned}
$$

Recall that the answers to Boolean queries given by knowledge bases are "yes", "no", or "do not know" (OWA). Find the answers given by the knowledge base $(\mathcal{T}, \mathcal{A})$ to the following Boolean queries: (5 marks)

- Cat(cotton_candy) Don't know
- Dog(paper) Don't know
- Mammal(cotton_candy) Yes
- Mammal(dudu) Yes
- Lung(krp) Yes
- ∀breatheWith.Lung(dudu) Yes
- ∀breatheWith.Lung(polly) Yes
- ¬Dog(polly) Don't know

Question 9. Reasoning in $\mathcal{EL}$

Let $\mathcal{T}$ be an $\mathcal{EL}$-TBox containing:

$$\mathsf{X} \sqsubseteq \exists \mathsf{r}.\mathsf{Y},\ \mathsf{Y} \sqsubseteq \exists \mathsf{r}.\mathsf{Y},$$

and let $\mathcal{T}$ be an $\mathcal{EL}$-ABox containing:

$$\mathsf{r(a,b), X(b)},$$

Compute the interpretation $\mathcal{I}_{\mathcal{T},\mathcal{A}}$ so that all $\mathcal{EL}$-concepts C and d $\in \{a, b\}$:

$$\mathcal{T}, \mathcal{A} \models \mathsf{C(d)} \iff \mathcal{I}_{\mathcal{T},\mathcal{A}} \models \mathsf{C(d)}$$

(5 marks)

Question 10. Ontology-Mediated Query Answering

Consider the following $\mathcal{ALC}$-TBox $\mathcal{T}$:

$$\top \sqsubseteq \mathsf{red} \sqcup \mathsf{green},\ \mathsf{red} \sqcap \exists \mathsf{r}.\mathsf{green} \sqsubseteq \mathsf{clash},\ \mathsf{green} \sqcap \exists \mathsf{r}.\mathsf{red} \sqsubseteq \mathsf{clash}$$

and the following $\mathcal{ALC}$-ABox $\mathcal{A}$:

$$\mathsf{r(0,\ 1),\ r(1,\ 2),\ r(2,\ 3),\ r(3,\ 0)}.$$

What is the answer of $(\mathcal{T}, \mathcal{A})$ to the Boolean query $\exists x\ \mathsf{clash}(x)$? Explain your answer. (5 marks)

Question 11 (BONUS question). Non-standard Reasoning

Consider the following $\mathcal{EL}$-TBox $\mathcal{T}$:

$$\mathsf{A} \sqcap \mathsf{C} \sqsubseteq \mathsf{D}, \mathsf{B} \sqsubseteq \exists \mathsf{r}.\mathsf{A}.$$

For any $\mathcal{EL}$-GCI α with (i) $\mathsf{sig}(\alpha) \subseteq \mathsf{sig}(\mathcal{T})$ and (ii) $\mathsf{A} \notin \mathsf{sig}(\alpha)$, show (on the semantics level) that $\alpha = \mathsf{B} \sqsubseteq \exists \mathsf{r}.\top$ is the strongest logical consequence of $\mathcal{T}$.

(Hint: α is the logical consequence of $\mathcal{T}$ means: $\mathcal{T} \models \mathsf{B} \sqsubseteq \exists \mathsf{r}.\top$, while α is the STRONGEST logical consequence of $\mathcal{T}$ means: there is NO another β with (i) $\mathsf{sig}(\beta) \subseteq \mathsf{sig}(\mathcal{T})$ and (ii) $\mathsf{A} \notin \mathsf{sig}(\beta)$ such that $\mathcal{T} \models \beta \models \alpha$).

(5 marks)