Two hours - online hybrid

EXAM PAPER MUST NOT BE REMOVED FROM THE EXAM ROOM

UNIVERSITY OF MANCHESTER
SCHOOL OF COMPUTER SCIENCE

Symbolic AI

Date: Wednesday 15th May 2019
Time: 09:45 - 11:45

This is a hybrid examination with Sections to be answered online and Sections to be answered on paper

Please answer ALL Questions in Section A and Section C online
Please answer ALL Questions in Section B and Section D in separate answerbooks

© The University of Manchester, 2019

This is a CLOSED book examination

The use of electronic calculators is NOT permitted

[PTO]
Section A contains multiple choice questions and is, therefore, restricted
Answer ALL questions. These questions should be answered ON PAPER

11. [Modelling and Representation]

For this question you should select one of the representation formalisms covered in the course e.g. either Datalog, Prolog, or first-order logic.

You have been given the job of designing a new intelligent system for course options and have chosen to make use of what you have learned in Symbolic AI.

a) You have been given a relational database consisting of five tables containing existing data. These tables contain student information, course information, student course choices, course prerequisites (other courses) and the courses belonging to each theme.

<table>
<thead>
<tr>
<th>Id</th>
<th>Name</th>
<th>Start year</th>
<th>Id</th>
<th>Name</th>
<th>Year</th>
<th>Unit Leader</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>John Smith</td>
<td>2017</td>
<td>1</td>
<td>Algorithms</td>
<td>2018</td>
<td>Giles Reger</td>
</tr>
<tr>
<td>2</td>
<td>Amelia Khan</td>
<td>2018</td>
<td>2</td>
<td>Symbolic AI</td>
<td>2018</td>
<td>Giles Reger</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Student</th>
<th>Course</th>
<th>Course</th>
<th>Prerequisite</th>
<th>Theme</th>
<th>Course</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
<td></td>
<td>SE</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td>AI</td>
<td>2</td>
</tr>
</tbody>
</table>

Describe how to represent these tables as a set of extensional relations. Make sure you consider whether your formalism assumes an open or closed world.

(2 marks)

b) For the following relations either specify it in your chosen formalism or state that it cannot be expressed.

- on_theme(S, T) holds iff student S takes a course in theme T
- all_pre(X, Y) as the transitive closure of the prerequisite table e.g. it holds iff course Y is a direct prerequisite of course X or there is a course Z such that Y is a prerequisite of Z and Z is a prerequisite of X.
- no_students(T) holds iff there are no students taking any course associated with theme T.

(2 marks)
c) For each of the following questions write a query to answer it or describe why we cannot express the query in your chosen formalism.

- Which students take two different courses in the AI theme?
- Which students do not satisfy the necessary prerequisites?

(2 marks)

12. [Reasoning]

To make office allocation in the School of Computer Science clearer, the Head of School (robert) has written some rules about which members of staff can sit in which offices. The rules are expressed in first-order logic as follows:

(a) Nobody is both a professor and lecturer

\[ \forall x. \neg (prof(x) \leftrightarrow lecturer(x)) \]

(b) A professor does not share their office with a non-professor

\[ \forall x. (prof(x) \rightarrow \neg \exists y. (\neg prof(y) \land office(x) = office(y))) \]

(c) Robert is a professor

\[ prof(\text{robert}) \]

(d) Giles is a lecturer

\[ lecturer(\text{giles}) \]

Giles is worried that the rules may mean that he has to share an office with Robert. It is your job to check that this cannot possibly happen given the above rules.

a) Transform the above formulas into clausal form. The clausification rules are given at the end of this section. (2 marks)

b) Apply the given clause algorithm with the rules (ordered) resolution, paramodulation, and equality resolution to show that the above set of formulas entails the formula

\[ office(\text{robert}) \neq office(\text{giles}) \]

These rules are given at the end of this section. You may select your own clause ordering and literal ordering. (7 marks)
Classification Rules for First-Order Logic

**Negation Normal Form**

\[-(F_1 \land \ldots \land F_n) \Rightarrow \lnot F_1 \lor \ldots \lor \lnot F_n\]

\[-(F_1 \lor \ldots \lor F_n) \Rightarrow \lnot F_1 \land \ldots \land \lnot F_n\]

\[F_1 \rightarrow F_2 \Rightarrow \lnot F_1 \lor F_2\]

\[-\lnot F \Rightarrow F\]

\[-\forall x_1, \ldots, x_n F \Rightarrow \exists x_1, \ldots, x_n \lnot F\]

\[-\exists x_1, \ldots, x_n F \Rightarrow \forall x_1, \ldots, x_n \lnot F\]

\[-(F_1 \leftrightarrow F_2) \Rightarrow F_1 \not\leftrightarrow F_2\]

\[F_1 \not\leftrightarrow F_2 \Rightarrow (F_1 \rightarrow F_2) \land (F_2 \rightarrow F_1);\]

\[F_1 \not\leftrightarrow F_2 \Rightarrow (F_1 \lor F_2) \land (\lnot F_1 \lor \lnot F_2).\]

**Skolemization**

\[\forall x_1, \ldots, x_n F \Rightarrow F\]

\[\exists x_1, \ldots, x_n F \Rightarrow F\{x_1 \mapsto f_1(y_1, \ldots, y_m), \ldots, x_n \mapsto f_n(y_1, \ldots, y_m)\},\]

**Clausal Normal Form**

\[(A_1 \land \ldots \land A_m) \lor B_1 \lor \ldots \lor B_n \Rightarrow (A_1 \lor B_1 \lor \ldots \lor B_n) \land \ldots \land (A_m \lor B_1 \lor \ldots \lor B_n).\]

**Reasoning Rules for First-Order Logic**

**Ordered Resolution**

\[
\frac{l_1 \lor C}{(C \lor D)\theta} \quad \frac{-l_2 \lor D}{\theta = \mgu(l_1, l_2)}
\]

where \(l_1\) and \(-l_2\) are selected by a well-behaved selection function.

**Paramodulation and Equality Resolution**

\[
\frac{C \lor s = t}{(l[t] \lor C \lor D)\theta} \quad \frac{l[u] \lor D}{\theta = \mgu(s, u)} \quad \frac{s \not= t \lor C}{C\theta} \quad \frac{\theta = \mgu(s, t)}
\]
13. **[Knowledge Representation]**
Convert the following natural language statements into structured KB facts according to RDF-NL.

“Although the Treasury will announce details of the November refunding on Monday, the funding will be delayed if Congress and President Bush fail to increase the Treasury’s borrowing capacity.”

Please follow this procedure: Start by identifying the nuclei and the satellites, identify the rhetorical relations, identify the clause-level predicate-argument structure (subject, predicate, object and context), representing it using the RDF-NL format.

(5 marks)

14. **[Natural Language Inference]**
List 3 basic entailment relations in the context of Natural Language Inference (NLI). Provide an example for each entailment relation. (3 marks)

15. **[Inductive Logic Programming]**
Write whether the following facts about the **FOIL Algorithm** are true or false and correct them when they are false.

(a) It is useful to learn the target function as a set of if-then-rules.
(b) Decision trees are not an expressive and interpretable representation.
(c) FOIL learns sets of first-order rules directly.
(d) Simultaneous covering algorithms learn just one rule at a time and perform many search steps.
(e) FOIL is a sequential covering algorithm.
(f) In FOIL, a general-to-specific search is performed to form the result set.
(g) In FOIL, a general-to-specific search is performed to form each new rule.
(h) Induction can be viewed as the inverse of deduction. However, an inverse resolution operator cannot be defined.

(4 marks)

16. **[Neuro-Symbolic Approaches]**
Please describe in 4 sentences the diff-ILP (differentiable ILP) approach explaining how the differentiable parts of the model (i.e. the neural components) interact with the symbolic-level (logical clauses) part of the model. (4 marks)
17. **[Semantic Parsing]**
Write the parsing tree of the command *mountains in Colorado* using a combinatorial categorical grammar (CCG) that interprets the sentence according to the KB below.
Please make sure that you:

- Identify the POS tags of the sentence.
- Define the lambda expressions.
- Define the derived parsing.

(6 marks)

18. **[End-to-End Neuro-Symbolic Systems]** Draw a diagram of a multi-component Question Answering (QA) end-to-end system obeying the following set of requirements:

- Data sources are unstructured facts.
- Extractors: NER, OpenIE (RDF-NL).
- Semantic Parser using a distributional CCG.
- Indicate the reference distributional reference corpora in your diagram.
- Indicate the requirements for OpenIE (F1: 0.85) and for the Semantic Parser (F1: 0.81).

(8 marks)