CS886 - AI and Philosophy
Fall 2017

Prof. Robin Cohen

Time and location: (Monday mornings 9am-11am; DC2568)

This seminar-style course is ideally suited for graduate students who are either writing theses in Artificial Intelligence or who have a good background in artificial intelligence. Other interested students are still welcome: simply ask the instructor. The course will be limited to 12 students. Any student is welcome to Audit the course and can request official Audit status.

The course will be discussion-oriented. The aim is to enable students to explore theories in philosophy at length, to then imagine applying these theories to the design of a particular artificial intelligence model. Each student will be required to present their selected philosophy theory with the class, and to later present their proposed application of that theory to artificial intelligence, ultimately leading to a completed final project which is handed in just after our last class. An additional exercise will involve working in small groups on the specific topic of ethics for AI. Initial classes will serve to introduce the students to the topic area, presenting examples of AI concepts which may benefit from the application of philosophy theories.

Workload (Tentative Plan)

<table>
<thead>
<tr>
<th>Workload</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Presentation of philosophy theory</td>
<td>10</td>
</tr>
<tr>
<td>Handout on philosophy theory</td>
<td>10</td>
</tr>
<tr>
<td>Presentation of proposed solution</td>
<td>10</td>
</tr>
<tr>
<td>Group project on ethics in AI</td>
<td>10</td>
</tr>
<tr>
<td>Group project presentation</td>
<td>5</td>
</tr>
<tr>
<td>Class participation</td>
<td>5</td>
</tr>
<tr>
<td>Final individual project</td>
<td>50</td>
</tr>
</tbody>
</table>

Schedule

This is a tentative outline for the course.

Sep 11    Introduction to course
Sep 18    Exemplar cases of utility functions and norms
Sep 25    Ethics in AI: discussion; time for groups to meet
Oct 2     Presentations of Ethics in AI solutions (projects due Oct 6)
Oct 16    Presentations of philosophy theories (4) (10 mins discussion on each)
Oct 23    Presentations of philosophy theories (4) (10 mins discussion on each)
Oct 30    Presentations of philosophy theories (4) (10 mins discussion on each)
Nov 6     Guest speaker (Paul Thagard)
Nov 13    Presentations of projects (5) (5 mins discussion on each)
Nov 20    Presentations of projects (5) (5 mins discussion on each)
Nov 27    Presentations of projects (2) (5 mins discuss); Course Wrap-up
Dec 4     (NO CLASS) Final Projects due
          Hardcopy to be placed in my CS mailbox by 4pm and softcopy sent

Contact: rcohen@uwaterloo.ca