Designing & Understanding Cognitive Systems

Peter Danielson
UBC COGS 300.002 2014-5
6 Jan2015
Outline

• Course
  – Some social contracts

• Subject
Course Team

• Instructors
  – Peter Danielson peter.danielson@gmail.com
  • 225 LSK

• TAs email: cogs300@gmail.com
  – Mario Cimet Lab 03 (Mon 3 - 5)
  – Katya Feffer Lab 04 (Wed 3 - 5)

• http://www.cogsys.ubc.ca/300/
• http://piazza.com/ubc.ca/winterterm22014/30002/home
Social Contract 1: Quizzes

A. Midterm
B. Daily Quizzes
C. B + some in class exam questions

• Social Contract
  – Veil of Ignorance
  – Anonymity
  – Social Choice
  – Default/baseline?
Social Contract 2

Presentations

When?

A. End of Half Term Blocks
B. Daily (in a few weeks)
C. C?

How Schedule:

A. Randomly assigned slots
B. A + OK to trade
C. ?
Course Evaluation

- No Midterm: 0%
- Daily quizzes & exercises: 20%
- Final exam: 30%
- Lab Projects [group]: 30%
- Assignments: 20%
  - Portfolio Blog: 5%
  - Presentation: 5%
  - Portfolio Wiki: 5%
  - Quiz/Discussion Questions: 5%
  - Games: 5% extra
Readings & Quizzes

• No text (but wiki)
• Various Perspectives & types of reading
  – Nobel/Turing Prize lectures
  – Engineering briefing
  – Journal articles
• Read in advance of lecture
  – Quiz ques due by 10 PM
  – Good questions (and review)
  – Attend & participate (drop 2 worst classes)
• Generate & Rank via Piazza
  – Subject to instructor modification
• Answer with i>clickers
  – Also for games & experiments
• Author answer
• Class discuss/evaluate
Labs

• Lego Mindstorms NXT
  – BYO Computers

• Problems
  – Lab 0: A basic mobile robot
  – Lab 1: Competitive game: X Wrestling
  – Lab 2: internal representation (maze)
  – Lab 3: Team cooperation/competition: Capture the X
Subject

• COGS emphasizing
  – Social science
    • Interacting Agents
  – Human – Computer/robot comparisons
  – Ethics of cognitive systems
1. A Moral Dilemma

A runaway trolley is about to run over and kill five people, but a bystander can throw a switch that will turn the trolley onto a side track, where it will kill only one person.

Is it permissible to throw the switch?

A. Yes
B. No
Standard Trolley Problem 2

A runaway trolley is about to run over and kill five people, but a bystander who is standing on a footbridge can shove a man in front of the train, saving the five people but killing the man. Is it permissible to shove the man?

A. Yes
B. No
Ethics for Cognitive Systems

• Controversy – yet some universals
  – Death is bad, more is worse
  – Footbridge “harder” than Switch
    • Yes Footbridge but No Switch rare
    • Timing data (for anonymous polling?)

• Social contract: Understanding Cog Sys:
A. Anonymous Polling
B. De-identified open data
C. ?
Ethics for Cognitive Systems: Design

• How should a robot
  – Drive the trolley? (prescriptive)
  – Will this make any difference to how people choose/answer? (descriptive)
  – How simple a situation will reveal these differences (methodological)
    • E.g. games

• New kinds of agent make moralized life more complex
Thanks for rescuing me from the streets of LA.
2. A Social Dilemma

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>3,3</td>
<td>0,5</td>
</tr>
<tr>
<td>D</td>
<td>5,0</td>
<td>1,1</td>
</tr>
</tbody>
</table>

• Past Example
  – C = Drive Civic
  – D = Drive Escalade

• Future Example: Set your AutoAudi to
  – C = Drive civilly
  – D = Drive aggressively

• Issues:
  – Descriptive:
    • What would you choose?
  – Prescriptive:
    • What should a rational agent do? ....moral
Motivating A Social Dilemma

- What would you do if the payoffs were grades for the course!
- Just kidding!
  - Real payoffs vs. Research ethics
  - A: don’t show or count
  - B: Show only
  - C: Extra grade points
  - D: Real grade points
Modeled as a Single-Play
N-player Prisoner’s Dilemma

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>C</strong></td>
<td>R = 3</td>
<td>S = 0</td>
</tr>
<tr>
<td><strong>D</strong></td>
<td>T = 5</td>
<td>P = 1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cooperators:</th>
<th>0</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>100%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>C</strong></td>
<td>(0)</td>
<td>.75</td>
<td>1.5</td>
<td>2.25</td>
<td>3</td>
</tr>
<tr>
<td><strong>D</strong></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>(5)</td>
</tr>
</tbody>
</table>
Next

• Newell & Simon on symbols
• Bring Clicker
• Try out Quiz/Discussion generator