COGS300: UNDERSTANDING & DESIGNING INTELLIGENT SYSTEMS

COURSE INFORMATION

Class Meeting Time: Tuesdays and Thursdays, 11-12:30 pm
Location: Geography, Room 147
  ▪ This is the 2nd edition of Mindware. Additional readings and resources will be provided throughout the course.

Lab Meeting Time: Mondays, 3-5pm; Tuesdays, 5-7pm; Wednesdays, 3-5pm
Location: FRDM Basement
  ▪ The labs will begin during the second week of the course, that is, from Monday, Sep 10.

COURSE TEAM

Instructor: Sina Fazelpour
Website: https://sinafazelpour.com/
Email Address: sina.fazelpour@alumni.ubc.ca
Office Location: Room 228, The W. Maurice Young Centre for Applied Ethics, Located in Leonard S. Klinck Building, 6356 Agricultural Road
Office Hours: Mondays and Wednesdays, 2-3pm
I very much encourage you to visit me during my office hours! If necessary, I would be more than happy to meet by appointment outside the regular hours as well.

Teaching Assistants
Ignacio Iturralde
Theresa Kikuchi
Eric Patterson
Email Address: cogs300@gmail.com
  ▪ Please see the guideline about emails in the Course Policy section.
COURSE DESCRIPTION

The aim of this course is to expose students to some of the key models and frameworks for understanding and designing intelligent systems. The first half of the course provides a close examination of computational models of cognition: we familiarize ourselves with the theoretical underpinnings of computational models, develop a concrete sense of how they work often with the aid of hands-on exercises, and then critically assess the adequacy of these models from empirical, computational, and conceptual perspectives. The second half of the course provides students with the opportunity to choose how they want to build on the background gained in the first half: some possibilities include considering challenges and alternatives to the computational framework, discussing novel and exciting extensions to the computational framework, and examining some of the social and ethical implications of the rapid growth in the design of artificial intelligence.

The lecture discussions are complemented with a lab component, where using Arduino microprocessors students build robots that sense and control their environment. The lab component provides students with hands-on experience about how the concepts discussed in lectures may work in practice. The practice of design offers a setting for thinking clearly about how a given task can be specified such that one can formulate different solutions on its basis and for considering how, given the peculiarities of a task environment, one can go about selecting solutions that yield effective physical implementations of that solution. What is more, simple as they may be, the designed systems provide a valuable opportunity for reflecting on how the practice of design, viewed from the “internal” perspective of the designer, may challenge or support “external” theoretical and methodological perspectives on how intelligent behavior should be modelled and understood.

COURSE EVALUATION

Schematic Overview

- Best 4 of 5 bi-weekly responses (5% each, total of 20%)
- Midterm take-home test (10%)
- Term project (20%)
- Lab work (25%)
- Final exam (25%)

*Make sure* to see the separate handouts for each of these components for deadlines, details, guidance, and specific assessment criteria.
WEEK 1

Lecture 1, Thursday, September 6. Introduction
The goal of this session is to provide a general introduction to the course and its different components. We will discuss different aspects of course evaluation and the motivation behind them. And, more generally still, we will use this session to get to know each other a bit.

▪ Andy Clark, Introduction
▪ If you are reading this, send me an email with a picture of your favorite artist (musician, actor, painter, writer, …)

Other resources:
▪ Andy Clark, Appendix I

WEEK 2

Lecture 2, Tuesday, September 11. Cognition as Computation
The goal of this session is to refresh our minds about some core concepts of logic and computability with the help of an in-class exercise.

▪ Andy Clark, Chapter 1, pp. 7-14 should provide a refresher; pp. 15-22 consider some initial motivations and worries)
▪ Alan Turing (1950) “Computing Machinery & Intelligence”
▪ (In-class exercise) Interactive Module: The Mindproject, Turing Machine

Other resources:
▪ A brief post on Persian mathematician Muhammad Al-Khwarizmi from whose (Latinized) name the term “algorithm” is derived
▪ Podcast: Konnie Huq on Ada Lovelace, acknowledged as the world’s first computer programmer

Lecture 3, Thursday, September 13. Cognition as Computation & Physical Computing
We start this session is to critically examine the Turing Test as a way of diagnosing intelligent behavior.

▪ Andy Clark, Chapter 1, pp. 15-22
▪ Interactive Module: The Mindproject, The Turing Test
▪ Douglas Hofstadter & Daniel Dennett, “The Turing Test: A Coffeehouse Conversation”
Other resources:

- **Interactive Module: Chat with a (Celebrity) Chatterbot**: you can try talking to Mitsuku, which holds the record of winning the Loebner Prize three times in 2013, 2016, and 2017, or Rose, which won the prize in 2014 and 2015.
  - [Homepage](#) of the Loebner Prize
- Tom Igoe, “What is Physical Computing”

**WEEK 3**

**Lecture 4, Tuesday, September 18. Classical Symbol Systems: Introduction**

The goal of this session is to introduce the concept of physical symbol system and the idea of problem solving as heuristic search.

- Andy Clark, Chapter 2, pp. 30-36 (ignore Box 2.2)

Other resources:

- **Video**: Herbert Simon on the Early Days of AI (1992)


The goal of this session is to continue our discussion of the physical symbol system hypothesis. We consider whether problem solving, understood as heuristic search, can be “creative” by examining some early heuristics for automated discovery.


Other resources:


**WEEK 4**

**Lecture 6, Tuesday, September 25. Cognitive Architectures**

The goal of this session is to introduce the concept of a cognitive architecture and critically exam the adequacy of the cognitive architecture approach in its aim of unify intelligence.

- **Online**: Newell & Simon’s [General Problem Solver](#)
- Andy Clark, Chapter 2, Box 2.2 on Soar architecture.
- Andy Clark, Chapter 2, Section C, pp. 42-45.
Lecture 7, Thursday, September 27. Meaning and Relevance
The goal of this session is to introduce and critically evaluate two of the main challenges to the symbol system hypothesis: Searle’s Chinese Room argument and the Frame problem. In terms of readings, if you do not have enough time, you can choose to focus on one of the two following options:

Option A: Problem of Meaning
- Andy Clark, Chapter 2, Section A, pp. 36-40 (excluding Box 2.4)
- John Searle (1980) Minds, Brains, and Programs
- Interactive Module: The Mindproject, Searle and the Chinese Room Argument

Option B: Problem of Relevance
- Andy Clark, Chapter 2, Section B, pp. 40-42

We end Thursday’s session with a class vote about the topics to be covered during weeks 11-12. So make sure you have looked at some of the sample sources before then.

Lecture 8, Tuesday, October 2. Intentionality
Intelligence, Turing argued, is a matter of how a system behaves. But what is it about (some) behavior that makes it intelligent? A promising answer that seems to fit nicely with the symbolic approach is based on postulating internal intentional states such as beliefs and desires. But do such internal states exist? The goal of this session is to critically examine the debate surrounding the reality of these internal states.

- Andy Clark, Chapter 3, pp. 47-57
Other sources:
- Video: Daniel Dennett, “A Difference That Makes a Difference” from Edge.

Lecture 9, Thursday, October 4. Intentionality
The goal of this session is to continue our discussion of intentionality and its relation to a system’s behavior by examining in further detail Dennett’s three stances.

- Interactive Module: John Conway’s Game of Life
- Daniel Dennett (1991) “Real Patterns”.
- Andy Clark, Chapter 3, Box 3.4; pp. 62-67 (Only Sections on “Stances” and “Upgrading the Basic Believer”)

Other resources:
- Joshua Rothman, “Daniel Dennett’s Science of the Soul”. The New Yorker, March 27, 2017 is a nice profile of Daniel Dennett and his general approach to understanding the mind.
- Video: Daniel Dennett’s talk at Google (February 8, 2017), talking about his last book From Bacteria to Bach and Back: The Evolution of Minds, which may be of interest to those who want to see how Dennett’s ideas have evolved
- Daniel Dennett (1995) “Back from the Drawing Board” in Dennett and His Critics (in Section 6, Dennett replies to Millikan’s criticism)

WEEK 6
Lecture 10, Tuesday, October 9. Connectionism: How to Build a Perceptron?
The goal of this session is to introduce the connectionist models of cognition by examining one of the earliest (and simplest) neural networks: the linear classifier Perceptron. We then examine some limits to the “expressive reach” of the Perceptron classifier. Depending on class interest and demand, we may begin the session by introducing some basic notions of linear algebra, relevant to understanding the workings of neural nets. If there is time, we will introduce some core concepts from statistical learning theory to better understand what it means for a class of learning rules to converge on the best solution for certain types of data, and what it means for certain other types of data to go beyond their “expressive reach”.

- Genevieve Orr’s Lecture on Perceptron
- Interactive Module: play with a perceptron

Other resources:
- Frank Rosenblatt (1957) “The Perceptron - a perceiving and recognizing automaton”
- Mikel Olazaran (1993) “A sociological history of the neural network controversy” is an excellent sociological study of the controversy surrounding the Perceptron project (and neural nets in general) during the 60s, fueled by (perhaps a misunderstanding of) Minsky’s results about the Perceptron. The controversy resulted in the stagnation of research on neural networks for decades.
Lecture 11, Thursday, October 11. Connectionism – Guest Lecture – Mark Schmidt
The goal of this session is to introduce further features of connectionist networks that allow them to overcome the difficulties faced by the perceptron. But, as we also learn, there is no such a thing as a free lunch!

- Andy Clark, Chapter 4, pp. 69-81
- Interactive Module: The Mindproject, Connectionism: An Introduction
- Interactive Module: TensorFlow is an accessible and elegant tool for playing with neural networks with different properties and learning rules

Other resources:
- Marvin Minsky (1990) “Logical vs. Analogical or Symbolic vs. Connectionist or Neat vs. Scruffy”
- Interactive Module: Neuroduino, a two-layer perceptron network for Arduinos
- Interactive Module: Simbrain is a free and accessible tool for building, running, and analyzing neural-networks. It also comes with a useful YouTube manual

WEEK 7
Lecture 12, Tuesday, October 16. Connectionism: From a Psychological Perspective
The goal of this session is to introduce some important challenges to the biological and psychological plausibility of early connectionist models of cognition and the developments inspired by these challenges. We finish by thinking about how certain cognitive phenomena such as mind-wandering, daydreaming and creativity could be viewed as features of neural networks.

- Andy Clark, Chapter 4, pp. 87-89

Other resources:
- Online: Google’s Inceptionism

Lecture 13, Thursday, October 18. Connectionism: Systematicity and Productivity
The goal of this session is to introduce and critically evaluate Fodor & Pylyshyn’s “systematicity” and “productivity” challenges to connectionist models of cognition (we will use Pinker’s more succinct and accessible formulation of these challenges. But feel free to read Fodor & Pylyshyn’s original article).

- Andy Clark, Chapter 4, pp. 84-86
- Steven Pinker, How the Mind Works, pp. 112-131.
Other resources:

- Nick Chater & Mike Oaksford (1990) “Autonomy, implementation and cognitive architecture: A reply to Fodor and Pylyshyn”

**Week 8**

**Lecture 14, Tuesday, October 23. Review**

The goal of this session is to provide a review of our progress so far for the midterm.

**Lecture 15, Thursday, October 25. Professional Development**

The goal of this session is to review the material covered from another perspective by examining how the tools gained in the course so far can be used for purposes of professional development.

**Week 9**

**Lecture 16, Tuesday, October 30. Marr’s Levels of Analysis and Perception**

The goal of this session is to critically examine the explanatory perspective we have thus far assumed in the course. To this end, we introduce David Marr’s highly influential distinction between three “levels of analysis”.

- Andy Clark, Chapter 5, pp. 93-94
- Jose Luis Bermudez, Selections.

Other resources:


**Lecture 17, Thursday, November 1. Levels of Analysis**

The goal of this session is to continue our discussion of Marr’s levels of analysis by considering its adequacy with respect to psychological and neurophysiological findings about perception, cognition, and action.

- Andy Clark, Chapter 5, pp. 95-111 (Discussion sections A and C only)

Other resources:

- Eric Jonas & Konrad Kording (2017) “Could a Neuroscientist Understand a Microprocessor?”

**Week 10**

**Lecture 18, Tuesday, November 6. Robotics**

The goal of this session is to use some of the problems with Marr’s approach to motivate an alternative approach to understanding intelligence based on robotics and design. We examine in
some detail the subsumption architecture, which embodies some of the core ideas of the
robotics approach, with the help of an in-class exercise.

- Andy Clark, Chapter 6, pp. 118-128
- Rodney Brooks (1991) “Intelligence without representation”

Other resources:
- **Video**: Andy Clark and Barbara Webb on the robotics approach part 1 and 2
- **Podcast**: Interview with Rolf Pfeifer on Talking Robots podcast.
- *Excerpts* from Valentino Braitenberg’s *Vehicles*

**Lecture 19, Thursday, November 8. Robotics and Artificial Life**

The goal of this session is to continue our discussion of robotics by considering evolutionary-
based robotics and artificial life. We end by assessing the explanatory scope of the non-
representation-based robotics approach as well as to consider the questions raised by the
robotics approach about the relation between mind and life.

- Andy Clark, Chapter 5, pp. 95-111 (Discussion sections A and C only)

Other resources:

For weeks 11 and 12, 2 of the following 5 options will be chosen by class vote. I have
provided some sample sources for each of these topics so that you can form a general idea about
them. Of course, given time limitations, not all these sample sources will be assigned as required
reading. Moreover, if a topic were chosen by popular vote, it is possible that similar, but different
readings could be assigned.

**WEEKS 11-12**

1. **Embodied, Embedded, and Enacted Cognition**

- **Sample sources:**
  - Andy Clark, Chapters 9 and 10
  - Diego Cosmelli & Evan Thompson (2010) “Embodiment or Envatment?: Reflections
    on the Bodily Basis of Consciousness”.
  - **Blog**: Evan Thompson’s blog posts (1, 2, 3, and 4) about the revised edition of his book
    (co-authored with Francisco Varela and Eleanor Rosch) *The Embodied Mind* on the Brains Blog
  - **Online**: Daniel Dennett chooses Affordances in response to *Edge’s* 2017 question
    “What Scientific Term or Concept Ought to be More Widely Known?”
2. Hierarchical Learning and Predictive Processing
   - Sample sources:
     - Andy Clark, Chapter 11.
     - Blog: Jakob Hohwy’s blog posts (1, 2, 3, and 4) about his book *The Predictive Mind* on the Brains Blog.
     - Podcast: Andy Clark on the *Brain Science* Podcast
     - Video: Karl Friston on the *Free Energy Principle*

3. Causal Graphical Models
   - Sample sources:
     - Eugene Charniak (1991) “Bayesian Networks without Tears”
     - Alison Gopnik & Laura Schulz (2007) “Causal Bayes Nets for Dummies, The Psychology of Causal Inference for Nerds – A Correspondence” (this is the Introduction to Alison Gopnik & Laura Schulz’s *Causal Learning: Psychology, Philosophy, and Computation*).
     - Podcast: Interview with David Danks about his 2014 book *Unifying the Mind* (New Books Network podcasts)
     - Video: Chris Hitchcock reviewing Causal Graphical Models at Judea Pearl Symposium
     - Online Module: DAGitty is an online environment for creating and analyzing causal graphical models (also known as directed acyclic graphs, DAGs).

4. Creativity & Artificial Intelligence
   - Sample sources:
     - Margaret Boden (1998) “Creativity and Artificial Intelligence”
     - Online: Google’s Inceptionism
     - Video: Margaret Boden, Creativity and AI: Asking the Right Questions

5. Ethics of Artificial Intelligence, Future of Humans
   - Sample sources:
     - Blog post: Rodney Brooks, “The Seven Deadly Sins of Predicting the Future of AI”
     - Video: Some of the videos from TEDTalk’s Artificial Intelligence playlist
     - Video: David Autor, “Will Automation Take Away All Our Jobs?”, TEDTalk.
**WEEK 13**

**Lecture 24, Tuesday, November 27.**
The goal of this session is to provide a synthetic overview of the course.

**Lecture 25, Thursday, November 29.**
The goal of this session is to go through some of the outstanding questions from the covered material.


**Course Policy**

**Plagiarism**
Here is UBC’s policy on plagiarism: Plagiarism, which is intellectual theft, occurs where an individual submits or presents the oral or written work of another person as his or her own. Scholarship quite properly rests upon examining and referring to the thoughts and writings of others. However, when another person’s words (i.e. phrases, sentences, or paragraphs), ideas, or entire works are used, the author must be acknowledged in the text, in footnotes, in endnotes, or in another accepted form of academic citation. Where direct quotations are made, they must be clearly delineated (for example, within quotation marks or separately indented). Failure to provide proper attribution is plagiarism because it represents someone else’s work as one’s own. Plagiarism should not occur in submitted drafts or final works. A student who seeks assistance from a tutor or other scholastic aids must ensure that the work submitted is the student’s own. Students are responsible for ensuring that any work submitted does not constitute plagiarism. Students who are in any doubt as to what constitutes plagiarism should consult their instructor before handing in any assignments. A link about Academic misconduct is below:

http://www.calendar.ubc.ca/Vancouver/index.cfm?tree=3,54,111,959

You are professionals and there is no excuse for plagiarism in this course.

**Missed Work**
If you miss a deadline for graded work due to illness or some other legitimate reason, you should immediately contact me, so that alternative arrangements can be made. For papers, this will typically mean a modest extension. Please note that the paper assignments will be given well in advance of the deadlines and that a personal problem the day before a deadline is not an adequate excuse. In addition, you may be required to provide documentation that would warrant granting an extension. In general, contact me as soon as you realize there may be some complications. Please see UBC’s general policy on missed work.

**Late Project Penalty**
If you don’t have an approved excuse for submitting a later project, you may still do so, but the project will be penalized—up to one grade per day (i.e., 10 points). To submit a late project, you must notify me, and provide me with a copy of your project.

**Email**
I will not be able to engage in e-mail discussions about course content. For example, please do not ask me to e-mail you an explanation of material that is discussed in lecture or the texts, or to e-mail you an explanation of one of the assignments. This is what the office hours are for, and you are most welcome to discuss all of these during the office hours. Still, e-mail is fine if you need to notify me of an illness or if you need to arrange an appointment. Be sure to include your full name in your email return address and to use a subject line that includes “COGS300”, otherwise your message may be treated as spam (by some algorithm unknown to me). If you are unsure about whether your email return address does include your full name, try emailing yourself. You should see your name in the from column in your inbox.
If you are emailing your TAs, please use the following format as part of your email’s subject line: “COGS300-Your TA's name”. This would make it easier for the specific TA to locate your email and respond swiftly.

Finally, whether emailing your TAs or me, please allow up to 24 hours (48 hours over the weekends or holidays) for our reply, before sending another email.

**Office Hours**
If you have questions about course content, you are very much encouraged to visit me during my office hours. You don’t need an appointment to attend scheduled office hours. If I am with another student or my door is closed, just knock on the door to let me know you are waiting.

**Announcements**
You are responsible for all announcements made in lecture and via the class e-mail list. The class e-mail list is automatically generated by the university. You should make sure that UBC has your current e-mail address so that you don’t miss any important messages.

**Lectures**
You are responsible for all of the material covered in lecture. Please bear in mind that the lectures and the readings are meant to complement one another. The lectures don’t simply rehash the readings. That would be pretty boring.

**Students with Disabilities**
If you have a documented disability and need special accommodations (e.g., extra time on the exam, a computer for notes), you should let me know within the first two weeks of classes. Please note that accommodations can be made only when students are registered with UBC’s Access & Diversity office.

**Changes to the Requirements**
I may see the need to change the basic requirements of the course as things progress. Such changes will be announced in lecture and will be posted on this website.

**Distributing Handouts, Course Material, and Passwords**
The handouts and study materials on this course website are for your personal educational use and are not to be posted on other websites, social media, or distributed via electronic means. Access to the course website is restricted to registered students in this course.