COGS 300: Understanding and Designing Cognitive Systems

1. Teaching team

Instructor
Márton Sóskuthy
email: marton.soskuthy@ubc.ca
office hour: 10:00–11:00 Tues,
location TBA

Teaching Assistants
Daichi Furukawa
Victoria Lim
Amy Wang
email: cogs.300@ubc.ca
technical office hour: TBA

2. Course description

COGS 300 is an exploration of foundational ideas and topics in cognitive science and related areas. This course serves two main goals: (i) to give you a taste of some of the main problems, debates and arguments in areas that lie at the intersection of the four COGS subjects (and maybe whet your appetite for more!); and (ii) to introduce you to key methodologies and technical ideas related to these problems/debates/arguments. This second point is especially important: we will be discussing big ideas and their broad implications, but I also want you to properly ground yourselves in the main areas of COGS by developing a feel for their 'tools of the trade'. Without this grounding, it's hard to appreciate what the big debates are really about!

Here is how the course is structured. In the first eight weeks (W1–W8), we are going to take a deep dive into the debate about symbolic *versus* connectionist models of cognition, while also taking the opportunity to introduce some key concepts such as self-organisation and emergence, which will become especially important later on in the course. The symbolic *versus* connectionist debate will give you a sense of some of the basic issues in cognitive science, and will also show you two radically different ways of approaching them. In the following two weeks (W9–W10), we will consider another important question about cognition: whether it is limited to the brain or if it extends

beyond it. Looking at the two closely related areas of embodied cognition and situated cognition will also reveal some embarrassing cracks in the edifice of cognitive science, which are now widely recognised as part of a 'replication crisis' in the social sciences. The final weeks of the course build further on the ideas of embodied/situated cognition and introduce yet another influential approach to cognition that relies on the idea of evolution and, in particular, cultural evolution.

This course is supported by labs built around the idea that you can better understand cognitive systems by implementing them in practice. The labs are all centred around Unity, which is a 3D software environment primarily used for developing computer games. You will explore various concepts introduced in the course through practical simulations, e.g. teaching a robot to navigate an obstacle course, coding a digital robot arm, simulating collective behaviour in animals etc. The labs culminate in a final project that allows you to show off your newly acquired skills in Unity and cognitive science: a robot tournament!

3. Course delivery

Below is a summary of the teaching schedule for COGS 300.

Lectures

Tuesday 11:00–12:30, MacMillan 160
Thursday 11:00–12:30, MacMillan 160

Labs

Monday 15:00–17:00, LIFE 2532
Wednesday 13:00–15:00, LIFE 2532
Friday 10:00–12:00, LIFE 2532

The *lectures* follow the progression of topics outlined in the 'course content' section. We will reflect on some of the key points of the readings, dissect difficult or technical arguments, go through a range of demonstrations and exercises related to the course content, and engage in conversation. To encourage discussion in the classroom, I will solicit conversation starters (questions or controversial statements) in advance of the Thursday lectures from a selection of students, and we'll choose a few of these for in-class discussion (see the 'assessment' section for more detail).

The *labs* focus on coding simulations in Unity. While they function as a more or less standalone component of this course, they are structured around topics covered in class such as physical symbol systems, simple neurally inspired models, self-organisation, etc. The model systems that you will code are relatively simple. Nonetheless, they provide

plenty of opportunities to think about key issues in cognitive systems in a slightly different context. Lab materials will be shared via the following website: `https://blogs.ubc.ca/cogs300/`

4. Assessment

Conversation starters (3 of them $\times$ 2%, 6% total). The Thursday lectures will include some structured in-class discussion organised around *conversation starters* provided by you. These conversation starters should be no longer than three lines, and may come in the form of questions or controversial statements related to one of the week's readings (your choice!). While they should have a clear connection to the reading, they should also go beyond it in some way: for instance, you could make a proposal that follows from an argument in the reading but isn't explicitly mentioned; you could present a mini-argument (remember: no longer than three lines!) for or against a point in the reading; or you could ask how a given theory would extend to a phenomenon not discussed in the reading. A few examples:

- Models of self-organisation present elegant explanations for simple behaviours such as the coordinated flashing of fireflies, but how could they possibly account for the emergence of complex behaviours such as language?
- The replication crisis in cognitive science suggests that many or perhaps even most empirical findings are false. Are there good heuristics for identifying results that are less likely to be false without digging into statistical details?
- The readings on deep learning show that learning algorithms are becoming incredibly powerful, but they don't even mention any related ethical problems. Are the killer robots coming for us?

Each week, students from one of the lab groups (i.e. around 25 students) will be asked to come up with a single conversation starter for one of the readings for that week. These will have to be submitted via Canvas no later than **midnight on Tuesday**. The anonymised conversation starters will then be made available to the class on Thursday in class, and you will be invited to vote for the ones you like best at the beginning of class. You can vote for any number of conversation starters and everyone gets to vote. We'll use the most popular ones in class. Your participation is absolutely essential for this to work: this is your chance to shape the topics that get discussed in class!

The conversation starters are graded based on participation: as long as you put one in before midnight on Tuesday and it's based on that week's readings, you will get the full 2%. **If it's off-topic (i.e. not connected to that week's readings) or you don't submit one, you will get 0%.**

The schedule of conversation starters is shown in the 'Course content, readings & assessment schedule' section below. Basically, we will follow an alternating schedule, with conversation starters solicited from students in the Monday lab group in week 2; from the Wednesday lab group in week 3; from the Friday lab group in week 4… Hopefully, you get the idea!

Online quizzes (4 of them × 4%, 16% total). There will be four half-hour online quizzes that test your knowledge and understanding of technical concepts and foundational ideas. The format is straightforward: multiple choice, matching related concepts, entering a number, etc. The quizzes are calibrated so that you need thorough understanding of the course materials to get an A (85%). The deadlines for the quizzes will be published on Canvas and also in the assessment schedule below, but they will be spread out relatively evenly throughout the course, and each of them will test one of the four main modules covered in the course (see the course outline below).

Reading responses (5 of them × 5% minus 5% for lowest grade, 20% total). This is a biweekly assessment that requires you to provide a short (no more than 400 word) response to one of the readings. The response should reference at least one other source (does not need to be a reading from this class) and – similarly to the conversation starters – it should go beyond the reading in some way. You can discuss implications of the reading, present applications or examples not covered in the reading, challenge or contest the reading in some way. It's up to you how you respond. The responses should be submitted online via Canvas. The deadlines will be announced in class and shown on Canvas as well.

Please use the first sentence or two of your reading response to highlight what you believe is the main message of the reading. Don't summarise the entire paper (400 words doesn't give you the space to do that). You can respond to any point in the reading (i.e. it doesn't have to be the main message), and you can bring in external ideas from other related fields. But please make sure that you pitch your essay at a level where a non-specialist (i.e. someone with a generic COGS background but no specialisation in linguistics / philosophy / etc.) can follow along. Note that the word limit is a hard one: we will take off marks for going over the limit of 400 words. Shorter essays are OK in principle, but, realistically, anything shorter than 300 words won't give you enough space to articulate a well-reasoned argument.

The grading rubric for the reading responses is available online in the modules section of Canvas. A few important pointers:

- Make sure that you include appropriately formatted references to both the original paper that you are responding to and the external paper(s) that you bring in (e.g. APA style).

- You can include references to websites and popular articles, but make sure that these are intellectually serious and from reputable sources.
- Use paragraphs to structure your reading response – if it's a single paragraph, you'll be marked down!
- Make sure that your reading response has a coherent structure, and indicate this structure explicitly through sign-posting, such as:referring to earlier points in the reading response and using phrases such as 'however', 'for example', 'moreover', etc. to indicate the logical structure of your argument.
- Choose a topic that goes beyond our class discussions; if you just rehash a point that we've already covered, you'll lose marks. But try not to stray too far either. A bit of wackiness is appreciated, but there should be a clear connection to the readings and course topics.
- Read the grading rubric carefully – this is what we'll use to evaluate your response!

We appreciate that there's a bit of a learning curve in figuring out how to write a good reading response, so your reading response with the lowest grade will be excluded from the calculation of your final grade.

Labs (10 of them $\times$ 2%, 20% total). You will be responsible to attend labs and work together in groups of 3-5 on the lab assignments. More information on marking will be provided in the labs.

Final lab project (11% total). The labs will culminate in a virtual robot tournament. Each lab group of 3-5 people will submit their own virtual robot, and these robots will compete in pairs in an arena. The last three labs of the course will help you design your robot and tackle various subparts of the final challenge, but it will be up to you to decide how you want to make your robot competitor stand out: perhaps by incorporating a simple neural network under the hood? Or by writing a complex symbolic representation of the task? Maybe by using an evolutionary algorithm to train your robot?

As part of this exercise, you will be asked to make a one-minute video that introduces your robot and sketches out its unique features; and a poster that gives a little more detail about how it is implemented. We will also ask you to make it clear how your implementation relates to concepts introduced in the course.

Your projects will be presented and marked in a tournament-style session in the last lecture of the term. Marking will be based on originality, relevance and execution, and will be **independent of how your robot performs in the tournament**. The grading

rubric for the projects will be made available on Canvas in the module section. Each team member will receive the same mark.

And now for the magnificent final prize: the winners (i.e. all team members involved in the construction of the winning robot) will receive an extra 2% on their final course grade (up to a maximum of 100%, of course!).

Final take-home exam (25% total). There will be a final take-home exam which will test knowledge of material from the readings and lectures, as well as your ability to integrate and communicate material. The topics covered in the final take-home exam are available on Canvas in the Modules section. The timing of this take-home exam will be shared later in the term.

Chat with Márton (2% total). This one is simple: come and see me in my office hour for a five minute chat or arrange another time via email. COGS 300 is a large class and the lecture format (plus the online teaching!) makes it difficult to interact like normal humans, so this meeting is supposed to help us have a real conversation. To avoid everyone coming to see me right at the beginning of term or right before the end of term, each one of you will be assigned a specific week to come and see me in. If you absolutely cannot make that week, or if you have another class during my office hour, email me and we'll come up with a solution. It will not be possible to arrange a chat after the end of term.

In summary, here's how you will be graded.

- 6%: conversation starters (3 per student)
- 16%: online quizzes (4 of them)
- 20%: reading responses (5 of them, lowest grade excluded)
- 20%: labs (10 of them)
- 11%: final lab project (+2% for the winners!)
- 25%: final take-home exam
- 2%: chat with Márton

5. Course content, readings and assessment schedule

The following table presents a detailed outline of the contents of the lectures and labs, and specifies the readings for each lecture. You are expected to read the assigned paper(s) *before* each lecture. One of the main goals of the classes is precisely to make some of the more technical arguments in the readings clearer, which assumes that you've actually done the reading!

Some of the readings are *challenging*. We'll try to tackle this in various ways, but to give you a sense of what to expect each week, I'll use a set of symbols to mark the level of difficulty of the assigned readings:

🍰: a piece of cake
🔗: follow link, read/explore online
⌛: a long one!
⚙: fairly technical paper
💣: this one's a real bastard

If there's no symbol, the reading is not particularly easy, nor particularly challenging, technical or long. Readings between square brackets are optional (though highly recommended!). They can be chosen as targets for your reading responses and we may mention them in class, but you will not be tested on them.

All readings will be provided for free on Canvas as PDFs.

WEEK	TOPIC & READINGS	LABS	ASSESSMENT
W1 Jan 8	TUES: *Introduction* no reading! THURS: *Big ideas in Cognitive Systems* `https://www.youtube.com/watch?v=PBdZi_JtV4c`	no lab!	no assessment!
Module 1: Physical Symbol Systems			
W2 Jan 15	TUES: *The Turing Machine* Clark (2013, ch. 1, pp. 7–14) https://turingmachine.io (to be explored in class) THURS: *Symbol systems* Clark (2013, ch. 2, pp. 30–36) Newell and Simon (1976, 113–120)	Introducing Unity	C. Starter: Mon group, due midnight Tues
W3 Jan 22	TUES: *Language as a symbol system* Pinker (2007, pp. 74–105) THURS: *Language is not an instinct* Tomasello (1995, pp. 131–146)	Obstacle course	C. Starter: Weds group, due midnight Tues Reading R. 1: due midnight Friday
Module 2: Emergence & Connectionism			
W4 Jan 29	TUES: *Emergence & Self-organisation* Kelso (1995, pp. 1–15) THURS: *Synchronized flashing of fireflies* Camazine et al. (2001, pp. 143–159)	Self-driving car	C. Starter: Fri group, due midnight Tues Quiz 1: released on Mon

WEEK	TOPIC & READINGS	LABS	ASSESSMENT
W5 Feb 5	TUES: *Perceptron* goo.gl/2QAC96 (What the Hell is a Perceptron?) Gurney (1997, ch. 1, 2.1–2.3, 3.1, 4.1–4.4) THURS: *Introducing connectionism* Clark (2013, ch. 4, pp. 69–81)	Fireflies Lab	C. Starter: Mon group, due midnight Tues Quiz 1: due midnight Mon Reading R. 2: due midnight Friday
W6 Feb 12	TUES: *Deep learning* LeCun et al. (2015) www.youtube.com/watch?v=FmpDIaiMIeA THURS: *Deep learning II, AI ethics* Vallor and Bekey (2017)	Perceptron Lab	C. Starter: Weds group, due midnight Tues Quiz 2: released on Mon
W7 19 Feb	Midterm break – no classes!		
W8 Feb 26	TUES: *Large Generative Language Models* Mahowald et al. (2023) THURS: *What is understanding?* Mitchell and Krakauer (2023)	Connectionism Lab	C. Starter: Fri group, due midnight Tues Quiz 2: due midnight Mon Readin R. 3: due midnight Fri
Module 3: Embodied & Situated Cognition			
W9 Mar 4	TUES: *Embodied cognition* Wilson and Golonka (2013, pp. 1–8) THURS: *Situated cognition* Robbins and Aydede (2009)	Embodied Cognition Lab	C. Starter: Mon group, due midnight Tues

WEEK	TOPIC & READINGS	LABS	ASSESSMENT
W10 Mar 11	TUES: *Power posing…* Carney et al. (2010) Ranehill et al. (2015) 🎂 THURS: *Of dead salmon and chronological rejuvenation* Simmons et al. (2011) [Bennett et al. (2010) 🎂]	Robot Tournament 1	C. Starter: Weds group, due midnight Tues Readin R. 4: due midnight Fri
Module 4: Cognitive & Cultural Evolution			
W11 Mar 18	TUES: *The evolution of cognition* Heyes (2012) THURS: *Cognitive technology* Clark (2013, ch. 8, pp. 166–179) 🎂	Robot Tournament 2	Quiz 3: released on Mon C. Starter: Fri group, due midnight Tues
W12 Mar 25	TUES: *Cultural evolution* Henrich and McElreath (2003) THURS: *Cultural evolution cont'd* no additional reading	no lab (Good Friday)	Quiz 3: due midnight Mon Readin R. 5: due midnight Fri
W13 Apr 1	TUES: *Iterated learning* Kirby et al. (2008) THURS: *Project troubleshooting* no reading	Robot Tournament 3	Quiz 4: released on Mon
Last week of term			
W14 Apr 8	TUES: *Revision* THURS: *Robot Tournament* no reading!	no lab!	Final project: due midnight Mon Quiz 4: due midnight Tues

6. Course policy

Attendance. Regular attendance of lecture and lab are expected.

Academic misconduct. Cheating and plagiarism are wrong; you all know that. I am obligated to report any kind of academic misconduct, so just don't do it. Please review the UBC Calendar "Academic regulations" for the university policy on cheating, plagiarism, and other forms of academic dishonesty. When you use text or ideas from another author, they 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.

Accommodations. The University accommodates students with disabilities who have registered with the Centre for Accessibility. The University accommodates students whose religious obligations conflict with attendance, submitting assignments, or completing scheduled tests and examinations. Please let your instructor know in advance, preferably in the first week of class, if you will require any accommodation on these grounds. Students who plan to be absent for varsity athletics, family obligations, or other similar commitments cannot assume they will be accommodated, and should discuss their commitments with the instructor before the drop date.

Missed and late work. We will only provide opportunities to submit missed work (reading responses, conversation starters or lab projects) when there is a serious and legitimate reason for missing the original assignment. If you anticipate that you will have trouble finishing a specific assignment, you will need to miss a lab, or you are struggling to keep up in general, please contact Márton as soon as possible. If you fall ill, have to care for someone who has fallen ill or are facing other serious difficulties, please try to get in touch as soon as you can so we can make arrangements for you.

Late work will be penalised at a rate of 10% per day of delay (unless you have an approved excuse for late submission). The conversation starters are an exception: they must be submitted on Tuesday before midnight; **late conversation starters receive a mark of 0**. If you submit something after the deadline, you are responsible for letting us know as we may already have started marking and may not be aware of your submission. In general, if you are anticipating issues with submitting coursework in time, please let us know as soon as you can.

Announcements. Important announcements will be made via Canvas. You are responsible

for reading these and following whatever instructions they contain. You should make sure to monitor your e-mail so that you don't miss any important messages.

UBC Values & Policies. UBC provides resources to support student learning and to maintain healthy lifestyles but recognizes that sometimes crises arise and so there are additional resources to access including those for survivors of sexual violence. UBC values respect for the person and ideas of all members of the academic community. Harassment and discrimination are not tolerated nor is suppression of academic freedom. UBC provides appropriate accommodation for students with disabilities and for religious and cultural observances. UBC values academic honesty and students are expected to acknowledge the ideas generated by others and to uphold the highest academic standards in all of their actions. Details of the policies and how to access support are available here: `https://senate.ubc.ca/policies-resources-support-student-success`

COVID-19 Safety. Please respect everyone else's decisions around COVID-19. For our meetings in this class, it is important that all of us feel as comfortable as possible engaging in class activities while sharing an indoor space.

If you are sick, and you suspect it may be COVID, I encourage you to take a rapid test to be on the safe side. If you are negative, it is up to you to decide whether you are well enough to come to class. I personally would encourage you to err on the side of caution, and stay home until you are mostly recovered.

In this course, class materials will be available on Canvas and all assignments are to be submitted online, so that you can prioritise your health and still succeed.

If you miss class because of illness:

- Consult the class resources on Canvas.
- Use the Piazza discussion forum for help.
- Make a connection early in the term to another student or a group of students in the class. You can help each other by sharing notes. If you don't yet know anyone in the class, post on the discussion forum to connect with other students.
- If you are concerned that you will miss an assignment deadline due to illness, contact the instructor to discuss.

If I (the instructor) am feeling ill: If I am unwell, I will test for COVID, and will only come to class if I am negative and well enough to lecture. I will also make sure to wear a mask if I still have any symptoms. I will make every reasonable attempt to communicate plans for class as soon as possible (by email, on Canvas, etc.).

References

Bennett, C. M., Baird, A. A., Miller, M. B., and Wolford, G. L. (2010). Neural correlates of interspecies perspective taking in the post-mortem atlantic salmon: An argument for proper multiple comparisons correction. *Journal of Serendipitous and Unexpected Results*, 1:1–5.

Camazine, S., Franks, N. R., Sneyd, J., Bonabeau, E., Deneubourg, J.-L., and Theraula, G. (2001). *Self-Organization in Biological Systems*. Princeton University Press.

Carney, D. R., Cuddy, A. J., and Yap, A. J. (2010). Power posing: Brief nonverbal displays affect neuroendocrine levels and risk tolerance. *Psychological science*, 21(10):1363–1368.

Clark, A. (2013). *Mindware: An Introduction to the Philosophy of Cognitive Science*. Oxford University Press, Oxford.

Gurney, K. (1997). *An Introduction to Neural Networks*. CRC Press.

Henrich, J. and McElreath, R. (2003). The evolution of cultural evolution. *Evolutionary Anthropology: Issues, News, and Reviews*, 12(3):123–135.

Heyes, C. (2012). New thinking: the evolution of human cognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1599):2091–2096.

Kelso, J. S. (1995). *Dynamic Patterns: The Self-Organization of Brain and Behavior*. The MIT Press.

Kirby, S., Cornish, H., and Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*.

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, 521(7553):436.

Mahowald, K., Ivanova, A. A., Blank, I. A., Kanwisher, N., Tenenbaum, J. B., and Fedorenko, E. (2023). Dissociating language and thought in large language models: a cognitive perspective. Preprint: *arXiv:2301.06627*.

Mitchell, M., and Krakauer, D. (2023). The debate over understanding in AI's large language models. *Proceedings of the National Academy of Sciences*, 120(13):e2215907120.

Newell, A. and Simon, H. A. (1976). Computer science as empirical inquiry: Symbols and search. *Communications of the Association for Computing Machinery*, 19(3):113–126.

Pinker, S. (2007). *The Language Instinct*. Harper Perennial Modern Classics, New York, NY.

Pinker, S. and Ullman, M. T. (2002). The past and future of the past tense. *Trends in cognitive sciences*, 6(11):456–463.

Ranehill, E., Dreber, A., Johannesson, M., Leiberg, S., Sul, S., and Weber, R. A. (2015). Assessing the robustness of power posing: No effect on hormones and risk tolerance in a large sample of men and women. *Psychological science*, 26(5):653–656.

Robbins, P. and Aydede, M. (2009). A short primer on situated cognition. In Robbins, P. and Aydede, M., editors, *The Cambridge Handbook of Situated Cognition*, pages 3–10. Cambridge University Press, Cambridge.

Rumelhart, D. and McClelland, J. (1986). On learning the past tenses of English verbs. In Rumelhart, D. and McClelland, J., editors, *Parallel distributed processing: explorations in the microstructure of cognition, vol. 2: psychological and biological models*, pages 216–271. MIT Press, Cambridge, MA.

Simmons, J. P., Nelson, L. D., and Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22(11):1359–1366.

Tomasello, M. (1995). Language is not an instinct. *Cognitive Development*, 10:131–156.

Vallor, S. and Bekey, G. A. (2017). Artificial intelligence and the ethics of self-learning robots. In P. Lin, K. Abney, R. J., editor, *Robot ethics 2.0*, pages 338–353. Oxford University Press, Oxford.

Wilson, A. and Golonka, S. (2013). Embodied cognition is not what you think it is. *Frontiers in Psychology*, 4:58.