PHL342H Minds and Machines

Fall 2020 Syllabus

Description
An investigation of key topics to do with the philosophical foundations of modern cognitive science and artificial intelligence (AI). Issues to be addressed include (among many others): whether machines can think and whether we might be thinking machines; functionalism, representationalism, and the computational theory of mind; the nature of computation and information processing; cognitive architecture; intentionality and consciousness; the extended mind hypothesis; the "singularity"; the moral, political, social, and existential implications of AI and cognitive science research.

Delivery Method
This course is online-synchronous. Students are required to view short prerecorded video lectures prior to attending a weekly live streamed full-class "town hall" and their weekly live streamed tutorial section. Everyone enrolled in the course will need a computer with a microphone and camera to participate in the town halls and tutorial sections.

- Weekly town hall: Wednesdays 2-3pm
    - The town halls will involve a blend of lecture-type presentation and class discussion. They will take place on Bb Collaborate Ultra via our main course Quercus page. Students should attend having already done the week's reading and having already viewed the week's video lectures. Further details regarding the town halls and how to participate will be furnished in Week 2, the first full week of classes (see the course calendar below).
- Weekly tutorial sections: Fridays 12-1pm, 1-2pm, or 2-3pm
    - The tutorial sections are venues in which students discuss course material with each other and with their tutorial leader. They will take place on Bb Collaborate Ultra via the various tutorial section Quercus pages. Further details regarding the tutorials and how to participate will be furnished in Week 2.
        - IMPORTANT: Please register for a spot in a tutorial section as soon as possible!

Instructor
Prof. James John
Email: jim.john@utoronto.ca

Email policy: I will read messages only if they are from a UTORmail email account and their subject lines include the course identifier and a clear statement of purpose (e.g. "PHL342H: I Have a Conflict with the Final Assessment"). Where a question cannot be answered with a brief reply email, I will indicate that you should speak to me during office hours. I will take up generic questions that could be of interest to all students at the beginning of the next town hall.

Office: UC F209 (but due to COVID-19 restrictions there will be no in-person office hours)
Office hours: Mondays and Thursdays 3-4pm, beginning in Week 3

Office hours policy: You are encouraged to attend the weekly office hours sessions. They will take place on Bb Collaborate Ultra via our main course Quercus page. The office hours will run on a drop-in basis. If you would like to schedule a private discussion, please contact me by email. Further details regarding office hours and how to participate will be furnished in the first full week of classes.

Course Websites, Recordings, and Readings
To access the main course website, login (with your UTORid and password) to Quercus (https://q.utoronto.ca) and select the "LEC0101" page for PHL342H. The websites for the various tutorial sections are in the same place but under "TUT" headings. Please check the main course website and your tutorial website regularly for important announcements.

IMPORTANT: *The weekly town halls will be recorded and made available on Quercus for the duration of the term for the benefit of students in remote time zones. However, the weekly tutorial sections will not be recorded*.

Recording privacy policy: All recorded town halls and course video lectures belong to the instructor, the University, and/or other sources depending on the specific facts of each situation, and they are protected by copyright. Do not download, copy, or share any course or student materials or videos without the explicit permission of the instructor.

- Students who wish to remain fully anonymous in the recorded town halls have the option of joining anonymously and of disengaging their cameras and microphones. Please inform me as soon as possible if you have questions or concerns.

Below, in the course calendar section, is a week-by-week listing of all of the *required* readings for PHL342H. Every reading is available for free online. There are no other required readings, and there are no physical books to purchase. (A list of recommended but not required supplementary books and articles will be provided later.)

Information Security
If you are a citizen of another country, and/or accessing your courses at the University of Toronto from a jurisdiction outside of Canada, please note that you may be subject to the laws of the country in which you are residing, or any country of which you have citizenship. The University of Toronto has a long-established commitment to freedom of expression, with this right enabled by an environment valuing respect, diversity, and inclusion. In this and your other courses, you may be assigned readings or discuss topics that are against the law in other jurisdictions. You are encouraged to become familiar with any local laws that may apply to you and any potential impact on you if course content and information could be considered illegal, controversial, or politically sensitive. If you have any concerns about these issues, please contact me for discussion.

Assignments, Grading, and Late Policy
Your grade will be determined as follows:

Participation in tutorial discussion	15%
Online Quercus quizzes	20%
2 5-6pp. essays	40% (20% each)
Final assessment (cumulative)	25%

On participation: There are ten weeks of tutorial sections for this course (see the course calendar for details). Their purpose is to provide you with an opportunity to discuss the course material. Participation in tutorials is required and attendance will be taken, though every student is allowed two "no questions asked" excused absences. Information regarding tutorials and how to get the most out of them will be distributed later.

- If you need to miss more than two tutorials, you should first report your absence through the online absence declaration available on ACORN (under the "Profile and Settings" menu) and then notify me as well as your TA. These additional absences will be excused only if a brief reading response on the week's readings is submitted to your TA within one week.
    - *If you are in a remote time zone and anticipate difficulty attending the tutorials, please contact me as soon as possible.*

On quizzes: There will be ten weekly Quercus quizzes, beginning in Week 3. Students must complete at least seven of the quizzes to receive a quiz grade for the course. A student's quiz grade will be determined by averaging their grades on the (seven or more) quizzes they completed. More information on the quizzes will be provided later.

On essays: There will be two short critical essays. Detailed instructions on how to write them will be furnished later. Essays are due via Quercus as indicated below on the course calendar. Work will be counted late from the due date on and will be penalized by dropping 1/3 of a letter grade for each day (including weekends and holidays) the work is late.[1] The Faculty deadline for the submission of term work is the last day of classes (12/9). Extensions beyond this deadline may be granted only if they do not interfere with the submission of grades. Otherwise, you must petition your College Registrar.

- All requests for extensions must be directed to me, not to TAs.

On the final assessment: There will be a short answer-based take-home final test due on a date to be determined during the December final assessment period. Detailed instructions on how to complete the test will be furnished before the end of term.

Academic Honesty

The University of Toronto takes academic integrity very seriously, and there are significant consequences if you are caught cheating or engaging in academic misconduct. All academic work in this course must adhere to the Code of Behaviour on Academic Matters (https://governingcouncil.utoronto.ca/secretariat/policies/code-behaviour-academic-matters-july-1-2019). Students are also urged to consult the Student Code of Conduct (https://governingcouncil.utoronto.ca/secretariat/policies/code-student-conduct-december-13-2019).

Some specifics to keep in mind:

- In written work, all sources used must be correctly cited, and if material is copied directly, appropriately cited and placed within quotation marks.

- If you work or study with friends, protect your work by not sharing or emailing your notes or assignments. You can help friends by discussing your ideas together and comparing your notes from lectures, but all written work must be completed on your own.

- You can find help with your assigned work by seeing me in office hours, speaking with your TA, or taking advantage of the many resources available on campus. I urge you to visit, in particular, the Office of Student Academic Integrity's website (http://www.artsci.utoronto.ca/osai/students), where you can find useful information on how to avoid plagiarism, and the Writing Centre's website (http://writing.utoronto.ca), where you can find further tips.

In Week 3 I will post to our course website a handout on how to avoid plagiarism. It is your responsibility to familiarize yourself with this material.

Course Code of Conduct and Accessibility

The University of Toronto is committed to equity, human rights, and respect for diversity. All members of the learning environment in this course should strive to create an atmosphere of mutual respect where all members of our community can express themselves, engage with each other, and respect one another's differences. The University does not condone discrimination or harassment against any persons or communities.

The University of Toronto is committed to accessibility, and students with special learning needs are welcome in this course. Please feel free to approach me or Accessibility Services (http://www.studentlife.utoronto.ca/as) if you

[1] For example, a paper that earns an A and is one day late will receive an A-, a B+ if two days late, a B if three days late, and so on.

have any accessibility concerns about the course, the online delivery procedures, or any course materials. We want to assist you as best we can in achieving academic success.

Course Calendar: Schedule of Topics, Readings, and Assignments

The URLs for each reading are listed below. "Public link" indicates that the reading is publicly available online. "Library link" indicates that the reading is made available by the University of Toronto Library and that to access it you will need to login with your UTORid and password.

Note: there will be tutorials every week of the term with the exceptions of Weeks 1, 2, 10, and 14.

Week 1 — Start Week (no class)

Week 2 9/16 — *Introduction to the Course: Artificial Intelligence and the Computer Model of the Mind*

Kevin Lande, "Do You Compute?" Public link: https://aeon.co/essays/your-brain-probably-is-a-computer-whatever-that-means

Robert Epstein, "The Empty Brain" Public link: https://aeon.co/essays/your-brain-does-not-process-information-and-it-is-not-a-computer

Week 3 9/23 — *The Turing Test*

Alan Turing, "Computing Machinery and Intelligence" Library link: https://www-jstor-org.myaccess.library.utoronto.ca/stable/2251299?pq-origsite=summon&seq=1#metadata_info_tab_contents

Ned Block, "The Mind as the Software of the Brain" (sec. 11.1.1) http://www.nyu.edu/gsas/dept/philo/faculty/block/papers/msb.pdf

- **TUTORIALS BEGIN! OFFICE HOURS BEGIN!**
- **ESSAY 1 ASSIGNMENT POSTED!**

Week 4 9/30 — *Metaphysical Interlude: Souls, Behavior, and Brains: The Rise of Functionalism*

Jerry Fodor, "The Mind-Body Problem" (pp.114-122, stop at "An obvious objection…") Library link: https://www-jstor-org.myaccess.library.utoronto.ca/stable/24964264?seq=1#metadata_info_tab_contents

Janet Levin, "Functionalism" in *Stanford Encyclopedia of Philosophy* (secs. 1-3) Public link: https://plato.stanford.edu/entries/functionalism/

Michael Rescorla, "The Computational Theory of Mind" in *SEP* (sec. 1) Public link: https://plato.stanford.edu/entries/computational-mind/

Week 5 10/7 — *Classical Computationalism*

Michael Rescorla, "The Computational Theory of Mind" in *SEP* (secs. 3 and 7.1)

Ned Block, "The Mind as the Software of the Brain" (secs. 11.1.2, 11.1.3, 11.1.4, 11.1.5, and 11.2)

Joel Walmsley, "Gödel's Incompleteness Theorem" Library link: https://link-springer-com.myaccess.library.utoronto.ca/content/pdf/10.1057%2F9781137283429_4.pdf

Week 6 10/14 — *The Chinese Room Argument*

John Searle, "Minds, Brains, and Programs" Library link: https://www-cambridge-org.myaccess.library.utoronto.ca/core/journals/behavioral-and-brain-sciences/article/minds-brains-and-programs/DC644B47A4299C637C89772FACC2706A

- Selected comments from the "Open Peer Commentary" (which ones, exactly, to be determined): Jerry Fodor, "Searle on What Only Brains Can Do"; John Haugeland, "Programs, Causal Powers, and Intentionality"; Zenon Pylyshyn, "The 'Causal Power' of Machines"; and Searle's "Author's Response"

Michael Rescorla, "The Computational Theory of Mind" in *SEP* (sec. 5.2)

- ****ESSAY 1 DUE VIA QUERCUS BY WEDNESDAY AT 11:59PM!!!****

Week 7 10/21 — *Mental Representation and the Problem of Intentionality*

Alex Byrne, "Intentionality" Public link: https://philpapers.org/rec/BYRI-2

Pierre Jacob, "Intentionality" in *SEP* (sec. 9) Public link: https://plato.stanford.edu/entries/intentionality/

David Pitt, "Mental Representation" in *SEP* (sec. 6) Public link: https://plato.stanford.edu/entries/mental-representation/

Ned Block, "Conceptual Role Semantics" (Public link: https://philpapers.org/rec/BLOCRS-2)

Week 8 10/28 — *Connectionist Computationalism and the Deep Learning Revolution*

Michael Rescorla, "The Computational Theory of Mind" in *SEP* (sec. 4)

Cameron Buckner and James Garson, "Connectionism" in *SEP* (sec. 11) Public link: https://plato.stanford.edu/entries/connectionism/

Cameron Buckner, "Deep Learning: A Philosophical Introduction" Library link: https://onlinelibrary-wiley-com.myaccess.library.utoronto.ca/doi/full/10.1111/phc3.12625

Week 9 11/4 — *Dynamical Systems*

Timothy van Gelder, "Dynamics and Cognition" Library link: http://cognet.mit.edu.myaccess.library.utoronto.ca/pdfviewer/book/9780262275071/chap16

Michael Rescorla, "The Computational Theory of Mind" in *SEP* (sec. 7.4)

- **ESSAY 2 ASSIGNMENT POSTED!**

Week 10 — Reading Week (no class)

Week 11 11/18 — *Embodiment*

Rodney Brooks, “Intelligence without Representation” Library link: https://www-sciencedirect-com.myaccess.library.utoronto.ca/science/article/pii/000437029190053M?via%3Dihub

Week 12 11/25 *The Extended Mind*

Andy Clark and David Chalmers, “The Extended Mind” Library link: https://www-jstor-org.myaccess.library.utoronto.ca/stable/3328150?pq-origsite=summon&seq=1#metadata_info_tab_contents

Joel Walmsley, “The Extended Mind Hypothesis” Library link: https://link-springer-com.myaccess.library.utoronto.ca/content/pdf/10.1057%2F9781137283429_8.pdf

- ****ESSAY 2 DUE VIA QUERCUS BY WEDNESDAY AT 11:59PM!!!****

Week 13 12/2 *The Challenge of Phenomenal Consciousness: Subjective and Qualitative Character*

Thomas Nagel, “What Is It Like to Be a Bat?” (Library link: https://www-jstor-org.myaccess.library.utoronto.ca/stable/2183914?pq-origsite=summon&seq=1#metadata_info_tab_contents

Frank Jackson, “What Mary Didn’t Know” Library link: https://www-jstor-org.myaccess.library.utoronto.ca/stable/2026143?pq-origsite=summon&seq=1#metadata_info_tab_contents

Ned Block, “Comparing the Major Theories of Consciousness” Public link: http://www.nyu.edu/gsas/dept/philo/faculty/block/papers/Theories_of_Consciousness.pdf

Week 14 12/9 *Course Conclusion: A Look Towards the Future: Utopia or Dystopia?*

Raffi Katchadourian, “The Doomsday Invention” Public link: https://www.newyorker.com/magazine/2015/11/23/doomsday-invention-artificial-intelligence-nick-bostrom

Kenneth Taylor, “The Robots Are Coming” Public link: https://bostonreview.net/science-nature-philosophy-religion/kenneth-taylor-robots-are-coming