

Lingnan University
Department of Philosophy

Course Title	:	Consciousness, Creativity, and Computation
Course Code	:	PAI4002
Recommended Study Term	:	Term 2, Fourth Year
No. of Credits/Term	:	3
Mode of Tuition	:	Sectional approach
Class Contact Hours	:	3 hours lecture per week
Category	:	Required course
Prerequisite(s)	:	NIL
Co-requisite(s)	:	NIL
Exclusion(s)	:	NIL
Exemption Requirement(s)	:	NIL

Brief Course Description

Artificial intelligence research has produced computers that perform reasoning tasks, learn, formulate plans, classify data, and interact via natural language. Many have argued on this basis that extant computers already count as thinking agents, employing at least rudimentary versions of mental states like belief and intention. In contrast, it remains acutely controversial whether there is any prospect whatsoever for a computational system that exhibits the mental capacities regarded as most characteristic of human mentality. Does the nature of computation permit or forbid a computational system that is conscious, or genuinely creative, or artistic, or capable of guilt or love, or free, or a moral agent on par with a human being? Are questions along this line even susceptible to rational inquiry, or are they unanswerable? These issues exist at the frontier of philosophy and science, interacting with lively debates on how best to characterize, theorize, and study the mental capacities in question. The aim of this course is to familiarize students with such debates and their implications for AI research, with a focus on the philosophical perspective. The course covers three interacting core topics: (i) the fundamental nature of computational systems and their physical implementation; (ii) the nature of consciousness, with a focus on the possibility of conscious AI systems; and (iii) the nature and value of creativity, with a focus on the possibility of significantly creative AI systems. Supplementary possible topics include: AI emotion; AI art; AI free will; AI selfhood; AI rights; AI personhood.

Aims

The course aims to:

- Introduce students to contemporary philosophical theories and debate on the nature of computation, the nature of consciousness, and the nature of creativity.
- Provide students with sufficient conceptual background to understand the central controversies raised by the idea of a conscious or creative AI.
- Enable students to engage with, analyse, and contribute to influential debates on consciousness, creativity, and AI.
- Enable students to draw connections between ongoing debates on the possibility of human-like AI and relevant debates in philosophy and science about characteristic features of the human mind.

Learning Outcomes

On completion of the course, students will be able to:

1. Explain and compare prominent foundational theories of computation, consciousness, and creativity, and their discussion in the context of AI research.
2. Identify, analyse, and critically assess prominent issues and arguments concerning the nature of consciousness, creativity and related mental capacities.
3. Identify, analyse, and critically assess prominent issues and arguments concerning the nature and possibility of human-like AI.
4. Specify connections between (i) ongoing debate on the nature of human capacities like consciousness and creativity and (ii) debates about the prospects for the development of human-like AI.

Indicative Content

1. Fundamentals of computation: algorithms, models of computation (finite automata, pushdown automata, Turing machines), universal machines, the Church-Turing Thesis.
2. Physical computers, the threat of pancomputationalism.
3. Theories of consciousness: higher-order theories, reflexive theories, representationalist theories, cognitive theories, phenomenal experience.
4. Theories of creativity: value, surprise, originality, agency.
5. Are consciousness or creativity necessary for intelligence? The Turing test versus Lovelace test for machine intelligence.
6. The possibility of machine consciousness.

7. The nature and value of AI art and computer art.
8. Possible additional topics: machine emotion, machine will, machine selfhood, machine personhood.

Teaching Method

A sectional approach: the major portion of lecture time will be used to introduce the course material (LO1, 2, 3, 4). The latter part of lectures will provide the opportunity to critically assess the theories introduced in the lectures through discussions and debates in order to specify the various connections between different theories of AI (LO2, 3, 4). Writing term papers and participation in class discussion will help the students develop their own views on the topics of the course (LO2, 3, 4).

Measurement of Learning Outcomes

Learning Outcomes	Assessment Methods		
	Class Participation	Two Research Essays	Midterm and Final Examination
LO1: Explain and compare prominent foundational theories of computation, consciousness, and creativity, and their discussion in the context of AI research.			✓
LO2: Identify, analyse, and critically assess prominent issues and arguments concerning the nature of consciousness, creativity and related mental capacities.	✓	✓	✓
LO3: Identify, analyse, and critically assess prominent issues and arguments concerning the nature and possibility of human-like AI.	✓	✓	✓
LO4: Specify connections		✓	

between (i) ongoing debate on the nature of human capacities like consciousness and creativity and (ii) debates about the prospects for the development of human-like AI.			
---	--	--	--

Assessment

Class Participation	: 10%
Two Individual Research Essays	: 20% each
Midterm Examination	: 20%
Final Examination	: 30%

Required/Essential Readings

Suggested textbooks:

- I. Consciousness
 - A. Uriah Kriegel (ed.), *The Oxford Handbook of the Philosophy of Consciousness*, Oxford University Press 2023
- II. Creativity
 - A. Amy Kind, *Imagination and Creative Thinking*, Cambridge University Press, 2022
 - B. Elliot Samuel Paul and Scott Barry Kaufman (eds.), *The Philosophy of Creativity: New Essays*, Oxford University Press 2017
- III. Computation and Human-level Artificial Intelligence
 - A. Jack Copeland, *Artificial Intelligence: A Philosophical Introduction*, Blackwell, 1993
 - B. Margaret Boden, *Artificial Intelligence: A Very Short Introduction*, Oxford University Press, 2018
 - C. Goran Hermeren, *Art and Artificial Intelligence*, Cambridge University Press, 2024.

Further suggested readings (selections):

- I. Computation
 - A. The nature of computation:

1. Hilary Putnam, 'Minds and Machines', in *Philosophical Papers: Volume 2, Mind, Language and Reality*, Cambridge University Press 1975.
2. W. Daniel Hillis, *The Pattern on the Stone*, Basic Books 2015
3. Tim Crane, *The Mechanical Mind*, 3rd edition, Routledge 2015
4. Richard Zach, 'Turing Machines', in *Sets, Logic, Computation: An Open Introduction to Metalogic*, The Open Logic Project 2021

B. Physical computation and pancomputationalism

1. Hilary Putnam, *Representation and Reality*, MIT Press 1988

II. Consciousness:

A. The nature of consciousness

1. Bernard J. Baars, *A Cognitive Theory of Consciousness*, Cambridge University Press 1988
2. David Chalmers, 'Facing up to the problem of consciousness', *Journal of Consciousness Studies* 2, 1995
3. Patricia Churchland, 'The Hornswoggle Problem', *Journal of Consciousness Studies* 3 (5-6), 1996
4. Susan Blackmore 'There Is No Stream of Consciousness', *Journal of Consciousness Studies* 9 (5-6), 2002
5. Daniel Dennett, 'Illusionism as the Obvious Default Theory of Consciousness', *Journal of Consciousness Studies* 23 (11-12), 2016
6. Hedda Hassel Morch, *Non-physicalist Theories of Consciousness*, Cambridge University Press 2023

B. Knowledge arguments

1. Thomas Nagel, 'What is it like to be a bat?' *Philosophical Review* 83, 1974
2. Frank Jackson, 'What Mary didn't know', *Journal of Philosophy* 83, 1986
3. David Lewis, "What experience teaches." In W. Lycan, ed. *Mind and Cognition: A Reader*. Oxford: Blackwell, 1990
4. Colin McGinn, *The Problem of Consciousness*. Oxford: Blackwell, 1991

III. Creativity

1. John Hospers, 'Artistic Creativity', *The Journal of Aesthetics and Art*

Criticism 43, 2005

2. Beth Hennessey and Teresa Amabile, 2010, 'Creativity', *Annual Review of Psychology* 61, 2010
3. Hills, Alison and Alexander Bird, 2018, "Creativity Without Value", in Gaut and Kieran (eds.), *Philosophy and Creativity*, 2018
4. Lindsay Brainard, 'What is creativity?', *Philosophical Quarterly*, Forthcoming

IV. Artificial Intelligence:

A. Could an AI system be conscious or feel emotion?

1. Alan Turing, 'Computing Machinery and Intelligence', *Mind* (59): pp. 433–60, 1950
2. Aaron Sloman and Ronald L. Chrisley, 'Virtual Machines and Consciousness', *Journal of Consciousness Studies* 10, 2003
3. Matthias Scheutz, 'Artificial emotions and machine consciousness', in *The Cambridge Handbook of Artificial Intelligence*, Frankish & Ramsey eds., Cambridge 2014
4. Susan Schneider, *Artificial You: AI and the Future of Your Mind*, Princeton University Press, 2019
5. Marcus Arvan and Corey Maley, 'Panpsychism and AI Consciousness', *Synthese* 200, 2022

B. Can an AI be creative and artistic?

1. Margaret Boden, 'Creativity and artificial intelligence', *Artificial Intelligence* 103, 1998
2. Brinjsjord et al, 'Creativity, the Turing test, and the (better) Lovelace test', *Minds and Machines* 11, 2001
3. Marta Halina, 'Insightful Artificial Intelligence', *Mind & Language*, 36, 2021
4. Margaret Boden, 'Skills and appreciation of computer art', *Connection Science* 28, 2016
5. Steffen Steinert, 'Art: brought to you by creative machines', *Philosophy and Technology* 30, 2016
6. Ted Chiang, 'Why A.I. isn't going to make art', *The New Yorker*, September 9, 2024

Recommended/Supplementary Readings

- Michael Wooldridge, *A Brief History of Artificial Intelligence*, Flatiron Books, 1st Edition, 2021
- Paul Churchland, *Matter and Consciousness*, MIT Press 2013
- Philip Goff, *Consciousness and Fundamental Reality*, Oxford University Press 2023

Important Notes

1. Students are expected to spend a total of 9 hours (i.e. 3 hours of class contact and 6 hours of personal study) per week to achieve the course learning outcomes.
2. Students shall be aware of the University regulations about dishonest practice in coursework, tests and examinations, and the possible consequences as stipulated in the Regulations Governing University Examinations and Course Work. In particular, plagiarism, being a kind of dishonest practice, is “the presentation of another person’s work without proper acknowledgement of the source, including exact phrases, or summarised ideas, or even footnotes/citations, whether protected by copyright or not, as the student’s own work”. Students are required to strictly follow university regulations governing academic integrity and honesty.
3. Students are required to submit writing assignment(s) using Turnitin.
4. To enhance students’ understanding of plagiarism, a mini-course “Online Tutorial on Plagiarism Awareness” is available on <https://pla.ln.edu.hk/>.
5. Students must adhere to the University’s guidelines and practices when using Generative Artificial Intelligence (AI) tools. The official documents “Guidelines for Using GAI Tools at Lingnan University” and “Best Practices for Ethical and Responsible Use of GAI Tools in Course Assessments” can be found here: <https://www.ln.edu.hk/tlc/generative-artificial-intelligence/gai-guidelines-best-practices>.

Consciousness, Creativity, and Computation

Individual Research Essay (20%) Assessment Rubrics

Assessment Categories	Excellent (A)	Good (B)	Pass (C)	Fail (F)

Knowledge of Course Material PA: 8/6/4/2/0	Knowledge goes substantially beyond what is expected of students in this course.	Essay/answer demonstrates knowledge of most important points; shows understanding of the basic philosophical positions and disagreements.	Some of the basic points are understood, but the student shows substantial misunderstandings.	Student's understanding deviates significantly from the positions discussed in class/ possibly fails to write on the assigned topic.
Argumentation and Originality PA: 8/6/4/2/0	Essay presents compelling arguments, shows understanding of important objections, and offers novel arguments or ideas that show a high degree of understanding.	Essay presents compelling arguments and shows understanding of important objections. Novel portions are less compelling or absent.	There are discernible arguments and engagement with the ideas of the class, but the arguments are often vague and impressionistic, or supported with dubious claims or fallacious modes of reasoning.	No discernible arguments/ no engagement with the ideas of the class.
Organisation PA: 5/4/3/2/0	Ideas are clearly arranged to allow easy understanding of their relations.	As before, but with slightly more mistakes in arrangement of ideas.	Essay seems to switch topics at points.	Total lack of organisational structure/ incomplete.

Writing PA: 4/3/2/1/0	Flawless, engaging writing.	Occasional mistakes in grammar or spelling which do not interfere with comprehension.	Some portions of the paper are clearly comprehensible, but other portions cannot be accurately assessed because of problems with the writing.	The content of the paper cannot be evaluated/ student is making little to no effort to complete the assignment.
------------------------------	--------------------------------	---	--	---

Consciousness, Creativity, and Computation
Midterm/Final Exam (20%/30%) Assessment Rubrics

Assessment Categories	Excellent (A)	Good (B)	Pass (C)	Fail (F)
Knowledge (Weight: 60%) 21/14/7/0	Demonstrates knowledge of all the relevant points, with a clear and accurate exposition.	Demonstrates knowledge of most relevant points, with a largely clear and accurate exposition.	Demonstrates some understanding of the basic points, but only a limited understanding.	Demonstrates very little understanding of even the basic points.

Argumentation (Weight: 30%) 7/4/2/0	Presents novel, compelling arguments with enough details.	Presents compelling arguments, with less novelty or some gaps.	Attempts to express an argument, but the result is unclear or messy.	Makes little or no attempt to present arguments.
Presentation (Weight: 10%) 7/4/2/0	Ideas are clearly arranged and expressed.	Ideas are mostly clearly arranged and expressed, but with elements of unclarity.	Some important ideas are clearly expressed, while others are hard to interpret.	Total lack of clarity.

**Consciousness, Creativity, and Computation
Class Participation (10%) Assessment Rubrics**

Assessment Categories	Excellent (A) 20-16%	Good (B) 15-10%	Pass (C) 9-6%	Fail (F) 5-1%
Attendance (Weight:70%) 20/14/7/0	The student attends class very regularly (about 90% or higher).	Attends class regularly (between 70% and 90%).	Attends class infrequently (between 50% and 70%).	Attends class infrequently (below 50%).

Discussion Participation (Weight: 15%) 20/14/7/0	The student always contributes to the discussion when in attendance.	The student sometimes contributes to the discussion when in attendance.	The student rarely contributes to the discussion when in attendance.	The student never contributes to the discussion when in attendance.
Quality of Contribution (Weight: 15%) 20/14/7/0	When the student contributes to the discussion, they raise thoughtful questions, analyse relevant issues, and demonstrate excellent preparation (e.g., knowledge of the assigned readings).	When the student contributes to the discussion, the level of engagement and preparation is good but not outstanding.	When the student contributes to the discussion, the level of engagement and preparation is sufficient but could be significantly improved.	The student never contributes to the discussion, and, when they do, they are seldom engaged with the course material and the preparation for the lesson is lacking in most cases.