Course Guide

PHIL 10134/ PHIL 11115
The Computational Mind

Course Organiser: Dr. Mazviita Chirimuuta (m.chirimuuta@ed.ac.uk)

- Office hours: by appointment

Course Secretaries

- UG: Ann-Marie Cowe (Annmarie.cowe@ed.ac.uk)
- PG: Becky Verdon (Rebecca.Verdon@ed.ac.uk)

Department of Philosophy
School of Philosophy, Psychology and Language Sciences
University of Edinburgh

1. Course Topic and Objectives

Computation is the dominant approach to explaining how the mind works within psychology and neuroscience. Artificial intelligence also now holds out the promise of recreating human-like mental capacities in computing machines. To many, this suggests that cognition (thought, perception, even emotion) *is* a kind of computation. In this course we will explore philosophical background to the computational approach to the mind, centering our study around two questions:

C*omputers compute, but do they think?* (weeks 1-6)
and *Is your brain a computer?* (weeks 7-11)

By the end of the course you should be able to:

1. Understand the basic concepts at play in the computational theory of mind.
2. Give informed assessments of claims made about the thought-like nature of computational processes.
3. Explicate and develop well reasoned arguments for and against the claim that the brain is a kind of computer.

2.1 Assessment Structure (Undergraduate)

- Midterm 1500 Words (40%)
- Final 2500 Words (55%)
- Participation (5%)

Deadlines

- Midterm: 21st October
- Final: 9th December

2.2 Assessment Structure (Postgraduate)

- Essay Plan 1-2 Pages (15%)
- Final 3000 Words (80%)
- Participation (5%)

Deadlines

- Essay Plan: 11th November
- Final: 9th December

Midterm 'formative' paper is optional

- Word limit: 1500 words
- Due date: 21st October

— Further details on course delivery and assessment will be provided on the Learn pages for this course —

3. Content and Readings

For each week, readings are listed below. Readings include *core* and *secondary* readings. The core readings for each week are starred (*).

Core readings are the material that it is your responsibility to read before each class.

Core readings are also the material on which our weekly seminar discussions will be based. Please do delve into the further reading too; these should be your first port of call when writing your essay. The core readings and as many as possible of the secondary readings are accessible via the Library Resources page on Learn.

Some hints: Read the core readings *carefully*. You may find an article challenging or difficult—persist! If you do not understand something, read it again, think about it, try to make sense of it in your own words. If after multiple attempts to make sense of a passage, you still cannot, then there is a good chance that you have identified a real problem in the article—a perfect point to raise in your discussion forum, in the class, or to form the basis of an excellent essay! Jim Pryor has some wonderful tips for reading philosophy (as he says, '*you should expect to read a philosophy article more than once*')

Background reading
The more background you know the better. A good starting point is to read one of the books listed below. Even if you already have a strong background in this topic already, I would encourage you to read one of these books during the semester to consolidate your knowledge.

[1] A. Clark. *Mindware: An Introduction to Cognitive Science*. Oxford University Press, Oxford, 2 edition, 2013.
[2] T. Crane. *The Mechanical Mind*. Routledge, London, 3rd edition, 2015.

And for more advanced readings:
[3] J. Haugeland, editor. Mind Design II. MIT Press, Cambridge, MA, 1999.
[4] M. Sprevak and M. Colombo (eds.) *The Routledge Handbook of the Computational Mind* Routledge, London, 2018.

Part I: Computers Compute, but do they Think?

Week 1: Turing Machines and Defining Computation

Commented [CMD1]: 20/9

* Michie, D. (2008) "Alan Turing's Mind Machines" chap 4 in Husbands and Holland (eds.) *The Mechanical Mind in History.* MIT Press

*Pick at least one of the following:

- Crane, T. "Computers and Thought" chapter 3, *The Mechanical Mind*.
- Haugeland, J. "Automatic Formal Systems chapter 2, *Artificial Intelligence: The Very Idea*

- Daston, L. (2018) "Calculation and the Division of Labor, 1750-1950," *Bulletin of the German Historical Institute*, no. 62 pp.9-30

[1] Crane, T. "Computers and Thought" chap 3, *The Mechanical Mind*.
[2] Proudfoot and Copeland (2018) "Turing and the first electronic brains: What the papers said" in Sprevak and Colombo (eds.) *The Routledge Handbook of the Computational Mind*
[3] Husbands and Holland (2008) "The Ratio Club: a hub of British cybernetics" chap 6 in *The Mechanical Mind in History.* MIT Press
[4] Abraham, T. (2002) "Physio(logical) circuits: the intellectual origins of the McCulloch-Pitts Neural Network". *Journal of the History of the Behavioral Sciences,* Vol. 38(1), 3–25

Week 2: Behaviourism and the Turing Test

Commented [CMD2]: 27/9

*Turing, A. 1950 "Computing Machinery and Intelligence" *Mind*, 59 (236) pp. 433-460
* N. Block, N. 1981 "Psychologism & Behaviorism". *The Philosophical Review.* 90: 5-43

Various essays in:
[1] Robert Epstein, Gary Robert and Grace Beber (2009) *Parsing the Turing Test Philosophical and Methodological Issues in the Quest for the Thinking Computer*. Springer

Week 3: The Functionalist Theory of Mind

Commented [CMD3]: 4/10

*Block, N. 1980 "What is Functionalism?" in *Readings in Philosophy and Psychology* vol. 1. Harvard University Press.
*Putnam, H. (1967) "The nature of mental states." In Mind, Language and Reality, Philosophical Papers, vol. 2, pages 429–440. Cambridge University Press, Cambridge, 1975; also reprinted 1992 *The Philosophy of Mind,* Beakley and Ludlow (eds.) MIT Press.

[1] J. J. C. Smart. Sensations and brain processes. Philosophical Review, 68:141–156, 1959.
[2] N. Block 1978 "Troubles with Functionalism" in W. Savage (ed.), *Perception and Cognition*. University of Minnesota Press. pp. 9—261
[3] O. Shagrir. 2005. The rise and fall of computational functionalism. In Y. Ben-Menahem, editor, *Contemporary Philosophy in Focus: Hilary Putnam*, pages 220–250.
[4] P. M. Churchland 2005 "Functionalism at Forty: A Critical Retrospective". *The Journal of Philosophy* Vol. 102, No. 1, pp. 33-50

Week 4: The Problem of Implementation

Commented [CMD4]: 11/10

* J. R. Searle 1992 "The Critique of Cognitive Reason" Chapter 9 in *The Rediscovery of the Mind*. MIT Press, Cambridge, MA.
*M. Sprevak 2018 "Triviality arguments" in Sprevak and Colombo (eds.) *The Routledge Handbook of the Computational Mind*

[1] 'Appendix' in H. Putnam. Representation and Reality. MIT Press, Cambridge, MA, 1988.
[2] Chalmers, D. 2011. "A Computational Foundation for the Study of Cognition." *The Journal of Cognitive Science* 12: 323-357.
[3] F. Egan. "Metaphysics and computational cognitive science: Let's not let the tail wag the dog." *Journal of Cognitive Science*, 13:39–49, 2012.
[4] P. Godfrey-Smith. "Triviality arguments against functionalism." *Philosophical Studies*, 145:273–295, 2009.
[5] O. Shagrir 2020 "In defense of the semantic view of computation." *Synthese* 197:4083–4108

Week 5: Representations

Commented [CMD5]: 18/10 – school half term
■Mid term due Thursday 21st

* Egan, F. 2018. "The Nature and Function of Content in Computational Models". In M. Sprevak and M. Colombo (eds.), *The Routledge Handbook of the Computational Mind*
* Brooks, R. 1991 "Intelligence without representation" *Artificial Intelligence* 47:139-159

[1] M. Rescorla. "How to integrate representation into computational modelling, and why we should". Journal of Cognitive Science, 13:1–38, 2012
[2] Shea, N. (2018) Chapter 1 "Introduction" *Representation in Cognitive Science*. Oxford University Press.
[3] Ramsey, W. 2017 "Must cognition be representational?" *Synthese* 194:4197–4214
[4] Haugeland, J. 1998 "The Intentionality All-Stars" in *Having Thought* Harvard University Press, pp. 127-170.

Week 6: The Only Game in Town?

Commented [CMD6]: 25/10

* T. van Gelder. 1995. "What might cognition be, if not computation?" *The Journal of Philosophy*, 91:345–381.
*A. Clark and J. Toribo 1994 "Doing without Representing?" *Synthese*, Vol. 101, No. 3 pp. 401-431

[1] Haugeland, J. 1998 "Mind Embodied and Embedded" in *Having Thought* Harvard University Press, pp. 207-237.
[2] R. Beer 2000 "Dynamical Approaches to Cognitive Science" Trends in Cognitive Sciences Vol. 4, No. 3 pp.91-99
[3] F. Faries and A. Chemero 2018. "Dynamic information processing" in Sprevak and Colombo (eds.) *The Routledge Handbook of the Computational Mind*
[4] D. Hutto, E. Myin, A. Peeters and F. Zahnoun "The cognitive basis of computation: Putting computation in its place" RHCM
[5] L. Barsalou, C. Breazeal and L. Smith 2007 "Cognition as coordinated non-cognition" *Cognitive Process* 8:79–91
[6] A. Clark. 2013 'Dynamics' Chapter 7 in *Mindware: An Introduction to Cognitive Science*. Oxford University Press.

Part II: Is Your Brain a Computer?

Week 7: The Machine-Organism Comparison, Historically Considered

Commented [CM7]: 1/11

* Canguilhem, Georges. [1965] 2008b. Machine and Organism. Pp. 75–97 *in Knowledge of Life*. Edited by Paola Marrati and Todd Meyers. New York: Fordham University Press.
* Hesse, M. 1961 "Models" in Chapter 1 *Forces and Fields* pp. 21-28

[1] Chirimuuta, M. (2020) The Reflex Machine and the Cybernetic Brain: The Critique of Abstraction and its Application to Computationalism. *Perspectives on Science* 2020, 28(3):421-457
[2] Levy, A. 2018 "Idealization and abstraction: refining the distinction" *Synthese*
[3] Potochnik, A. 2017 *Idealization and the Aims of Science*. Chicago University Press
[4]Canguilhem, G. 1963 "The Role of Analogies and Models in Biological Discovery" in A. C. Crombie (ed.) *Scientific Change*. New York: Basic Books.
[5] Pickering, A. (2010) chap 3 "Grey Walter" *The Cybernetic Brain*. Chicago University Press
[6] Riskin, J. (2016) chap 9 "Outside In", *The restless clock: A history of the centuries-long argument over what makes living things tick*. Chicago University Press
[7] von Neumann, J. (1958/2000) *The Computer and the Brain*. Yale University Press

Week 8: Marr's levels of explanation

Commented [CM8]: 8/11

* Chapter 1 of D. Marr. Vision. W. H. Freeman, San Francisco, CA, 1982.
* L. Elber-Dorozko and O. Shagrir 2018. "Computation and levels in the cognitive and neural sciences" in Sprevak and Colombo (eds.) *The Routledge Handbook of the Computational Mind.*

[1] O. Shagrir and W. Bechtel. 2017 "Marr's computational level and delineating phenomena."
[2] M. Chirimuuta. 2018 "Marr, Mayr, and MR: What functionalism should now be about," *Philosophical Psychology,* 31(3):403-418
[3] D. C. Dennett. 1994 "Cognitive science as reverse engineering: Several meanings of 'top-down' and 'bottom- up'". In D. Prawitz, B. Skyrms, and D. Westerstahl, editors, Proceedings of the 9th International Congress of Logic, Methodology and Philosophy of Science.
[4] A. Newell. 1982 "The knowledge level." *Artificial Intelligence*, 18:87–127.
[5] Egan, F. (2010) 'Computational Models: A Modest Role for Content', *Studies in History and Philosophy of Science*, 41, pp. 253–259.

Week 9: Computational Explanation in Neuroscience

Commented [CM9]: 15/11

* D. Kaplan. 2011 "Explanation and Description in Computational Neuroscience." *Synthese* 183: 339-73.
* M. Chirimuuta. 2014 "Minimal models and canonical neural computations." *Synthese* 191: 127-53.

[1] F. Egan 2017 "Function-Theoretic Explanation and the Search for Neural Mechanisms" In D. Kaplan, editor, *Explanation and Integration in Mind and Brain Science*. Oxford University Press, Oxford
[2] P. Machamer, L. Darden, & C. Craver. 2000 "Thinking about Mechanisms." Philosophy of Science 67: 1-25,
[3] D. Coelho Mollo. 2018 "Functional individuation, mechanistic implementation." *Synthese* 195:3477–3497
[4] R. Cao 2018 "Computational explanations and neural coding" in Sprevak and Colombo (eds.) *The Routledge Handbook of the Computational Mind*
[5] M. Colombo 2014 "Deep and beautiful. The reward prediction error hypothesis of dopamine." *Studies in History and Philosophy of Biological and Biomedical Sciences* 45: 57–67

Week 10: Brain-Like Computation: Connectionism and Deep Learning

Commented [CM10]: 22/11

* A. Clark. 2013 'Connectionism' Chapter 4 in *Mindware: An Introduction to Cognitive Science*. Oxford University Press.
* C. Buckner 2019 "Deep learning: A philosophical introduction." *Philosophy Compass*. 14:e12625.

[1] G. E. Hinton. How neural networks learn from experience. Scientific American, 267:145–151, 1992.
[2] C. Buckner and J. Garson 2018 "Connectionism and post-connectionist models" in Sprevak and Colombo (eds.) *The Routledge Handbook of the Computational Mind*

Week 11: Deep Learning, Neuroscience and Psychology

Commented [CM11]: 29/11

* C. Stinson 2020 "From Implausible Artificial Neurons to Idealized Cognitive Models: Rebooting Philosophy of Artificial Intelligence." *Philosophy of Science*
* M. Mitchell 2020 "On Crashing the Barrier of Meaning in Artificial Intelligence" *AI Magazine*, *41*(2), 86-92.

[1] R. Cao and D. Yamins 2021. "Explanatory models in neuroscience: Part 1 – taking mechanistic abstraction seriously." *https://arxiv.org/abs/2104.01490*
[2] R. Cao and D. Yamins 2021. "Explanatory models in neuroscience: Part 2 – constraint-based intelligibility." *https://arxiv.org/abs/2104.01489*.
[3] Buckner, C. (2018). Empiricism without magic: Transformational abstraction in deep convolutional neural networks. *Synthese*, *195*(12), 5339–5372