Universitat Autònoma de Barcelona

2017/2018

Minds, Machines and Cognition

Code: 43842
ECTS Credits: 6

Degree	Type	Year	Semester
4316227 Applied Philosophy	OT	0	2

Contact

Name: Olga Fernández Prat

Email: Olga.Fernandez@uab.cat

Use of languages

Principal working language: catalan (cat)

Teachers

Targeta Provisional

Prerequisites

Competent english reading.

Objectives and Contextualisation

Cognitive Science is the current science of mind. Clearly this is a discipline in interdisciplinary research that converges in philosophy, computer science, linguistics, psychology, neuroscience, sociology and anthropology.

Scientists working in cognitive science have proposed models on a variety of cognitive processes and have offered explanations about the language, vision/perception, memory, concepts, reasoning, problem solving and decision. One of the distinctive features of the first work in cognitive science is the assumption that the mind can be understood as a special type of a computer. In fact, many researchers in artificial intelligence conceived their work as an attempt to implement cognitive processes in machines.

In this module we will explore some of the main works on the philosophy of cognitive science and artificial intelligence and analyze, with a critical and philosophical approach, the influence of the computer analogy in cognitive science. The following is an indicative list of subjects related to the module: the Turing test; the Chinese room argument; the frame problem; connectionism; extended and embodied cognition; artificial conscience; logics to represent knowledge in artificial intelligence (modal logic, temporal logic, fuzzy logic…).

Students are expected to be able to formulate and argue their points of view on the topics covered in the course. In addition, it is intended that they develop the following competences:

- To demonstrate their knowledge and critical understanding of the main arguments about the possibility of the existence of intelligent machines.

- To understand critically some of the central philosophical issues about the nature of thought and consciousness.

- To show critical understanding of some explanations about the mind and cognitive sciences.

- To demonstrate ability to critically understand some of the main texts in the philosophy of cognitive sciences and artificial intelligence.

- To prove the hability to represent arguments and show their correctness in different non-classical logics.

Skills

- Analyze critically and synthesize information obtained from an article or a specialized monograph, and from quality information distributed on the web.
- Apply knowledge of classical authors in the western philosophical tradition to current philosophical questions.
- Continue the learning process, to a large extent autonomously.
- Critically assess the implications on the human condition of new ideological, political, economic and technological forms that impact on the contemporary world.
- Define, design, plan and prepare an original and unpublished work of philosophical research, following established academic-scientific parameters.
- Establish and apply the implications that scientific knowledge and research have for advanced philosophical research.
- Organize ones own time and resources to undertake research: design a plan by prioritizing objectives, schedules and commitments.
- Reconstruct and analyze critically the positions of the main current researchers in the field of philosophy of each of the main subject areas of the masters degree (science, art, politics) using their characteristic categories and lexis.
- Relate the concepts and knowledge of the various areas of current philosophical research in relation to dependencies between science and technology, and the ethical and political implications of such dependencies.
- Search for, select and manage information autonomously, both from structured sources (data bases, bibliographies, specialized journals) and from information distributed on the web.
- Solve problems in new or little-known situations within broader (or multidisciplinary) contexts related to the field of study.
- Use acquired knowledge as a basis for originality in the application of ideas, often in a research context.

Learning outcomes

1. Analyze critically and synthesize information obtained from an article or a specialized monograph, and from quality information distributed on the web.
2. Apply and adapt current philosophical proposals to problems in the cognitive sciences.
3. Apply current philosophical language, content and theories to the problems related to philosophy of the mind.
4. Apply knowledge of philosophy to debates on the regulation of cybernetics.
5. Continue the learning process, to a large extent autonomously.
6. Demonstrate critical understanding of some explanations concerning the mind and the cognitive sciences and the technical applications of these.
7. Draw up a paper in philosophy within the framework of problem areas in the cognitive sciences.
8. Organize ones own time and resources to undertake research: design a plan by prioritizing objectives, schedules and commitments.
9. Relate scientific knowledge to philosophical proposals and their explanatory and critical orientations.
10. Search for, select and manage information autonomously, both from structured sources (data bases, bibliographies, specialized journals) and from information distributed on the web.
11. Solve problems in new or little-known situations within broader (or multidisciplinary) contexts related to the field of study.
12. Use acquired knowledge as a basis for originality in the application of ideas, often in a research context.

Content

1. Turing machines and the problem of consciousness.

References:

- Turing, A. (1947), ¿Puede pensar una máquina? (translations available on-line).

- Searle, J. (1980), "Minds, Brains and Programs", Behavioural and Brain Sciences 3, 417-457.

2. Traditional Cognitive Science and Classical Artificial Intelligence.

References:

- Pylyshyn, Z. (1984), Computation and cognition: toward a foundation for cognitive science, MIT Press, Cambridge. (A selection will be provided).

3. Criticism of traditional Cognitive Science.

References:

- Dreyfus, H. (1972), What Computers Can't Do, New York: Harper & Row. (A selection will be provided).

4. The new Cognitive Science: extended and embodied cognition.

References:

- Haugeland, J. (1998), Mind Embodied and Embedded, in J. Haugeland, Having Thought: Essays in the Metaphysics of Mind, Harvard University Press.

- Prinz, J.J. (2009), "Is Consciousness Embodied?", in P. Robbins and M. Aydede (eds.), The Cambridge Handbook of Situated Cognition, Cambridge: Cambridge University Press, pp. 419- 436.

5. Epistemology of robots.

References:

- Boden, M. (1995), 'Could a Robot Be Creative - And Would We Know?', in Android Epistemology, Cambridge: MIT Press.

- Gomila, A. (1995), 'From Cognitive Systems to Persons', in Android Epistemology, Cambridge: MIT Press.

6. Artificial Intelligence: origins, achievements and challenges.

References:

- COPELAND, J. (1993), Artificial Intelligence: A Philosophical Introduction, Wiley-Blackwell, páginas 20-33.

- RUSELL, S., NORVIG P. (2009), Artificial Intelligence: A Modern Approach, Prentice-Hall, capítols 1, 27.

- LÓPEZ DE MANTARAS, R. (2015), De Turing als robots humanoides: Passat, present i futur de la Intel.ligència Artificial. Available video at www.uab.cat.

7. Turing Machines. Philosophy of Artificial Intelligence.

References:

- COPELAND, J. (1993), Artificial Intelligence: A Philosophical Introduction, Wiley-Blackwell, páginas 72-96, 204-271.

- NICOLELIS, M, CICUREL, R. (2015), The Relativistic Brain: How it works and why it cannot be simulated by a Turing machine, Kios Press, chapters 3, 4.

- RENDELL, P. (2015), Turing Machine Universality of the Game of Life (Emergence, Complexity and Computation), Springer, 1st edition, chapter 1.

- RUSELL, S., NORVIG P. (2009), Artificial Intelligence: A Modern Approach, Prentice-Hall, capítol 26.

- University of Cambridge (2012), Alan Turing- The life of a genius. Available video at iTunes.

- Stanford Encyclopedia of Philosophy, "Turing Machines" (https://plato.stanford.edu/entries/ turing-machine/).

8. The Role of Logic: Representation of Knowledge and Reasoning.

References:

- BADESA, C., JANÉ, I., JANSANA, R. (2000), Elementos de lógica formal, Ariel, chapters 6, 7,11.

- BENTHEM, J, Hans van DITMARCH, H., Jan van EIJCK, J. Eijck, JASPARS, J (2011), Logic in Action, chapter 1.

- RUSELL, S., NORVIG P. (2009), Artificial Intelligence:A Modern Approach, Prentice-Hall, chapter 7.

9. Deontic logics.

References:

- BENTHEM, J, Hans van DITMARCH, H., Jan van EIJCK, J. Eijck, JASPARS, J (2011), Logic in Action. (A selection will be provided) .

- BLACKBURN, P., RIJKE, M. de, VENEMA, Y. (2001): Modal Logic, volume 53 of Cambridge Tracts in Theoretical Computer Science, Cambridge University Press, chapter 1.

- GARSON, J.W. (2013): Modal logic for philosophers, Cambridge University Press, chapters 1,3.

- Notes provided by the teacher.

10. Fuzzy Logics.

References:

- GARRIDO, Ángel (2014). Lógicas de nuestro tiempo, Dykinson, chapter IV.

- SMITH, N.J.J. (2013):Vagueness and degrees of truth, Oxford, chapters 3,5.

- Notes provided by the teacher.

Other references

Block, N., 2005, "Review of Alva Noë Action in Perception," Journal of Philosophy, 102: 259-272.

Calvo, P. y Gomila, A., 2008, Handbook of Cognitive Science. An Embodied Approach, Elsevier.

Clark, A., 2008, Supersizing the Mind: Embodiment, Action, and Cognitive Extension, New York: Oxford University Press.

Dubois, D., Godo, L., Prade, H.(2014): Weighted logics for artificial intelligence : an introductory discussion. International Journal of Approximate Reasoning, Elsevier, vol. 55 (n9), pp. 1819-1829.

Gardner, H. (1985). The mind's new science. Basic Books, New York.

Marr, D., 1982, Vision: A Computational View, San Francisco: Freeman Press.

Moore, C., Mertens, S. (2011),The Nature of Computation, Oxford University Press, Oxford.

Pérez Chico, D., González Bedia, M. (eds.), La Nueva Ciencia Cognitiva, Ed. Plaza y Valdés.

Varela, F., Thompson, E. E. Rosch, 1991, The Embodied Mind: Cognitive Science and Human Experience, Cambridge, MA: MIT Press.

Methodology

There are three elements:

1. Lectures.

2. Seminars: discussions of the required readings.

3. Student's work at home. Students must do some required readings. The work will be supervised by the teachers during seminars.

Activities

Title	Hours	ECTS	Learning outcomes
Type: Directed			
Seminars and lectures.	30	1.2	3, 4, 2, 6, 7, 11, 5, 9, 12
Type: Supervised			
Tutorials.	26	1.04	1, 10, 9
Type: Autonomous			
Compulsory and required readings.	61	2.44	1, 10, 8
Doing exercises.	7	0.28	1, 8, 5
Preparing oral presentation.	7	0.28	1, 3, 4, 2, 10, 6, 7, 8, 5, 9

Evaluation

The assessment of the course has the following elements:

During the classes 1 to 5 it will be assigned to a student one of the compulsory readings and he/she must make an oral presentation of it. After he/she must deliver the written part that he/she had used in order to make the presentation. It is expected that other students have also read the papers in order to get involved in the discussion afterwards. The oral presentation, the written part and the participation in the seminars worth 50% of the final grade.

During the classes 6 to 10 the teacher wil distribute exercises for doing at home. Students will dispose of a week in order to resolve them. The answers must be delivered via e-mail or in hand the following class. The last two exercises can be delivered via e-mail or at the teacher's office. All these exercises worth 50% of the final grade. There are seven exercisews: the first one worth 1 point over the final grade, the second and the tercer 0,25 each, the fourth and fifth 0,75 each, the sixth and the last 1 point each.

Evaluation activities

Title	Weighting	Hours	ECTS	Learning outcomes
Class presentation.	50%	12	0.48	1, 3, 4, 2, 10, 6, 7, 8, 11, 5, 9, 12
Exercises.	50%	7	0.28	11, 5

Bibliography

BADESA, C., JANÉ, I., JANSANA, R. (2000), Elementos de lógica formal, Ariel, chapters 6, 7 and 11.

BENTHEM, J, Hans van DITMARCH, H., Jan van EIJCK, J. Eijck, JASPARS, J (2011), Logic in Action.

BLACKBURN, P., RIJKE, M. de, VENEMA, Y. (2001): Modal Logic, volume 53 of Cambridge Tracts in Theoretical Computer Science, Cambridge University Press, chapter 1.

BLOCK, N., 2005, "Review of Alva Noë Action in Perception," Journal of Philosophy, 102: 259 272.

BODEN, M. (1995), "Could a Robot Be Creative - And Would We Know?" in Android Epistemology, Cambridge: MIT Press.

CALVO, P., GOMILA, A., Handbook of Cognitive Science. An Embodied Approach, Elsevier.

CLARK, A., 2008, Supersizing the Mind: Embodiment, Action, and Cognitive Extension, New York: Oxford University Press.

COPELAND, J. (1993), Artificial Intelligence: A Philosophical Introduction, Wiley-Blackwell, pages 20-33, 72-96, 204-271.

DUBOIS, D., GODO, L., PRADE, H.(2014): Weighted logics for artificial intelligence : an introductory discussion. International Journal of Approximate Reasoning, Elsevier, vol. 55 (9), pp. 1819-1829.

DREYFUS, H. (1972), What Computers Can't Do, New York: Harper & Row.

GARDNER, H. (1985). The mind's new science. Basic Books, New York.

GARRIDO, Ángel (2014). Lógicas de nuestro tiempo, Dykinson, chapter 4.

GARSON, J.W. (2013): Modal logic for philosophers, Cambridge University Press, chapters 1 and 3.

GOMILA, A. (1995), "From Cognitive Systems to Persons" a Android Epistemology, Cambridge: MIT Press.

HAUGELAND, J. (1998). Mind Embodied and Embedded en J. Haugeland, Having Thought: Essays inthe Metaphysics of Mind, Harvard University Press.

HAYES, P. (1995), Android Epistemology, MIT Press.

LÓPEZ DE MANTARAS, R. (2015), De Turing als robots humanoides: Passat, present i futur de la Intel.ligència Artificial. Available video in www.uab.cat.

MARR, D., 1982, Vision: A Computational View, San Francisco: Freeman Press.

MOORE, C., MERTENS, S. (2011),The Nature of Computation. Oxford University Press, Oxford.

NICOLELIS, M, CICUREL, R. (2015), The Relativistic Brain: How it works and why it cannot be simulated by a Turing machine, Kios Press, chapters 3 and 4.

PÉREZ CHICO, D., GONZÁLEZ BEDIA, M. (eds.): La Nueva Ciencia Cognitiva, Ed. Plaza y Valdés.

PRINZ, J.J. (2009), "Is Consciousness Embodied?" en The Cambridge Handbook of Situated Cognition, P. Robbins and M. Aydede (eds.), Cambridge: Cambridge University Press, pp.

419-436.

PYLYSHYN, Z. (1984), Computation and cognition: toward a foundation for cognitive science, MIT Press, Cambridge.

RENDELL, P. (2015), Turing Machine Universality of the Game of Life (Emergence, Complexity and Computation), Springer, 1st edition, chapter 1.

RUSELL, S., NORVIG P. (2009), Artificial Intelligence: A Modern Approach, Prentice-Hall, chapters 1, 7, 26 and 27.

SEARLE, J. (1980), "Minds, Brains and Programs", Behavioural and Brain Sciences 3, 417-457.

SMITH, N.J.J. (2013):Vagueness and degrees of truth, Oxford, chapters 3 and 5.

STANFORD ENCYCLOPEDIA of PHILOSOPHY, "Turing Machines".

TURING, A. (1947), "¿Puede pensar una máquina?" a Mentes y Máquinas, Tecnos, 1985.

UNIVERSITY of CAMBRIDGE (2012), Alan Turing- The life of a genius. Free video available at iTunes.

VARELA, F., THOMPSON, E. E. ROSCH, 1991, The Embodied Mind: Cognitive Science and Human Experience, Cambridge, MA: MIT Press.