Contents

Introduction I-III

Chapter 1. Cognitive Science and Artificial Intelligence: Death and Rebirth of a Collaboration

1. When Cognitive Science was AI
1.2. From The General Problem Solver to the Society of Mind: cognitivist insights from the early AI era
1.3. Heuristics and AI Eras
1.4. Modelling Paradigms and AI Eras: Cognitivist and Emergentist Perspective
1.5. Death and Rebirth of a Collaboration

Chapter 2. Cognitive and Machine Oriented Approaches to Intelligence in Artificial Systems

2.1. Nature vs Machine Inspired Approaches to Artificial Systems
2.2. Functionalist vs Structuralist Design Approaches
2.3. Levels of Analysis of Computational Systems
2.4. The Space of Cognitive Systems
2.5. Functional and Structural Neural Systems
2.6. Functional and Structural Symbolic Systems

Chapter 3. Principles of the Cognitive Design Approach

3.1. Classical, Bounded and Bounded-Rational Models of Cognition
3.2. Resource-Rationality Models
3.3 Kinds of Explanations
3.4 Levels of Plausibility and the Minimal Cognitive Grid (MCD)

Chapter 4. Examples of Cognitively Inspired Systems and application of the Minimal Cognitive Grid

4.1 Modern AI Systems: Cognitive Computing?
4.2 Cognitive Architectures

4.3 SOAR

4.4. ACT-R

4.5 Two Problems for the Knowledge Level in Cognitive Architectures

4.6. Knowledge Size and Knowledge Heterogeneity in SOAR and ACT-R

4.7. DUAL PECCS

Chapter 5. Evaluating the Performances of Artificial Systems

5.1. “Thinking” Machines and Turing Test(s)

5.2. The Chinese Room

5.3. The Newell Test for a Theory of Cognition

5.4. The Winograd Schema Challenge

5.5. DARPA Challenges, Robocup and Robocup@Home

5.6. Comparison

Chapter 6. The Next Steps

6.1. The Road Travelled

6.2. The Way Forward

6.3. Towards a Standard Model of Mind/Common Model of Cognition

6.4. Community
List of Figures and Tables

Chapter 1. Cognitive Science and Artificial Intelligence: Death and Rebirth of a Collaboration

Section 1.4. Fig. 1. Overview of the internal dynamics of Physical Symbol Systems
Section 1.4. Fig. 2. Brooks Subsumption Architecture

Chapter 2. Cognitive and Machine Oriented Approaches to Intelligence in Artificial Systems

Section 2.4 Fig. 3. Enriched 2D Space of Cognitive Systems

Chapter 4. Examples of Cognitively Inspired Systems and application of the Minimal Cognitive Grid

Section 4.3 Fig. 4. The SOAR cognitive architecture
Section 4.3 Table 1. Newell’s timescale of human actions
Section 4.4 Fig. 5. The ACT-R cognitive architecture
Section 4.5 Table 2. Prototype models vs Exemplar Models
Section 4.7. Fig. 6. An example of the hybrid conceptual architecture in DUAL PECCS

Chapter 5. Evaluating the Performances of Artificial Systems

Section 5.1 Fig. 7. A pictorial representation of the “Imitation Game”
Section 5.6 Table 3. Comparative table of the different evaluation approaches
This book is about (re)building a bridge between two different “sciences of the artificial”: Artificial Intelligence and Cognitive Science that, nowadays, apart from some notable exceptions, do not talk much to each other as they should. Here, I review some of the main themes that have characterized the historical paths of these two disciplines and argue that the technological maturity reached in several domains now calls for a renewed joint enterprise finalised at addressing more substantial challenges that these two disciplines have to face from a scientific viewpoint.

The book explicitly targets a multidisciplinary audience. As such it is mainly an act of courage (or probably of irresponsibility) since experts in the specific subfields will have for sure much more things to say and would surely be able to communicate their own ideas in a better way than I can do. However, as mentioned, this book privileges the breath of the connections between the disciplines rather than the depths of the exploration within each single discipline. As such, it is not a manual or a handbook since it presupposes the knowledge of same basic elements of each of the disciplines that will be touched by our arguments. Of course, scholars and students of the diverse fields have knowledge of different pieces of the entire puzzle and need to be briefly introduced to the aspects that they do not know. This service is provided in the book that, however, reminds to the specialistic literature for the details.

One of the main goals of this manuscript is to show to the reader that the so called “cognitive design approach” has still an important role to play in the development of intelligent AI technologies as well as in the context of development of plausible computational models of cognition. In other words: the study of the “Cognitive Design” principles for building “Artificial Minds” will be hopefully a useful instrument for the current and future generation of AI and cognitive science scholars and students. In this respect, a first caveat is necessary: in the philosophical literature on the AI there are many different, and well known, positions about whether or not it it is justifiable to use the terms “mind”, “intelligence” or “thinking”, to describe the constitutive or the behavioural elements of a computational system. In this book we will not enter in the details of such a monumental and decades-long debate, that also involve the attribution of such faculties to other “species” (from non-human mammals to bacteria). Given the actual purpose of the book, we will also avoid to roughly summarize it because such an attempt would be necessarily incomplete. Sometimes, however, we will refer to some instances of such a debate. For the moment we just mention here, as a reference for the position about why the term “mind” can be justifiably
associated to the term “artificial”, the book “Artificial Minds” by Stan Franklin (Franklin, 1995). The position defended by Franklin, that sees the possession of a “mind” as a matter of degrees and not as a mere boolean notion and that - as such - foresee the possibility of implementing (to some degree) a “mind” in an artificial system, can be considered our starting working hypothesis.
A Paola e Francesca
Chapter 1 - Cognitive Science and Artificial Intelligence: Death and Rebirth of a collaboration

The first chapter proposes a brief historical overview of some of the main insights developed in 65 years of research in Artificial Intelligence (AI), by introducing the early vision of the discipline (based on a mutual collaboration with the Cognitive Psychology) and its “paradigm shift” started from the mid’80’s of the last century. Starting from that period, Artificial Intelligence and the interdisciplinary enterprise known as Cognitive Science started to produce several sub-fields, each with its own goals, methods and criteria for evaluation. The reasons for the current renewed interest of a cognitively inspired approach in the AI research are discussed.

Chapter 2 - Cognitive and Machine oriented Approaches to Intelligence in Artificial Systems

This chapter presents the different possible routes to build an Artificial Intelligence system. On one hand it presents the design assumptions underlying the cognitive approaches to AI and, on the other hand, it presents the tenets of machine oriented approaches aimed at obtaining AI systems able to exhibit intelligent behaviour without making any assumption about the biological or cognitive plausibility of the implemented mechanisms. It additionally introduces the reader to the main instances about the debate on the levels of analysis of computational systems (being cognitively inspired or not).

Chapter 3 – Principles of the Cognitive Design Approach

This chapter introduces the classical notions of rationality developed in the field of cognitive modelling and presents different types of explanatory accounts available in the literature. Finally it presents the “Minimal Cognitive Grid”, a pragmatic methodological tool proposed to rank the different degrees of structural accuracy of artificial systems in order project and predict their explanatory power.

Chapter 4 – Examples of Cognitively Inspired Systems and application of the MCD

Given the proposal presented in the previous sections of the book, this chapter describes some practical applications of the Minimal Cognitive Grid by showing how it allows to collocate
different types of artificial systems in the landscape formed by the cognitive design approach. Examples of artificial models of cognition and cognitive architectures will be shown and compared with examples of functionalists AI systems that, despite called as instances of “cognitive computing”, cannot be considered realistic models of our cognition.

Chapter 5 - Evaluating the Performances of Artificial Systems

This chapter introduces the main proposals that have been developed in order to evaluate the performance of artificial systems (cognitively inspired or not) and to justify the ascription of faculties coming from the “cognitive” vocabulary (like “intelligence”) to such systems. After introducing the Turing Test, its problematic aspects and some of the main modifications proposed (e.g. the Super Turing Test and other variations), we will analyze other frameworks like the Newell Test for a theory of cognition and other tasks and challenges that have been used - with different purposes - as a testbed for the evaluation of artificial systems. These tasks go from the Robocup World Soccer to the DARPA Challenges for autonomous vehicles to the recently proposed Winograd Schema Challenge and the RoboCup@Home. We will analyse these proposal both in the light of their eventual explanatory role in the context of a computationally-driven science of the mind and with respect to their actual capacity of evaluating the “intelligence” of artificial systems.

Chapter 6 - The Next Steps

This concluding chapter will synthesize the main issues presented along the book and will try to provide a roadmap for the next years in the context of cognitive AI research, by suggesting fields where the cognitive design approach can provide valuable inputs for the realization of better AI systems.
References


Fodor, J.A. (1986), Why Paramecia don't have mental representations, Midwest Studies in Philosophy, 10: 3-23.


Hopfield, J.J. (1982), Neural networks and physical systems with emergent collective computational abilities, Proceedings of the National Academy of Sciences of USA, 79: 2554-2558


LaCurts, K. (2011). Criticisms of the turing test and why you should ignore (most of) them. CSAIL, MIT Philosophy and theoretical computer science.


