Abductive Strategies in Human Cognition and in Deep Learning Machines

Lorenzo Magnani
University of Pavia, Italy

INTRODUCTION

On a full-size 19x19 board, in 2015, Google DeepMind's program AlphaGo (capable of playing the famous Go game) defeated Fan Hui, the European Go champion and a 2 dan (out of 9 dan) professional, five times out of five with no handicap. In March 2016, Google also defeated Lee Sedol, a 9 dan player widely regarded as the world's finest champion. In four of the five games, the DeepMind software beat Lee. The program was extremely innovative, with many novel moves and a good capacity to replicate human behavior. It learned from a variety of human-played games using a technique known as "reinforcement learning". The program's more current version, in turn, plays against itself to improve the efficiency of its deep neural networks: AlphaZero is a more modern richer version of AlphaGo, as detailed fully and suitably in https://en.wikipedia.org/wiki/AlphaZero.

We can say that heuristics are used to arrive at a specific target thanks to their organization in strategies. We can consider computational strategies as a compound of heuristic processes, that is, processes composed of good choices of the subsequent state of a cognitive routine – according to some opportunely chosen criteria. In game theory, however, the definition of strategy is broader, encompassing everything from how agents interact with one another to a wide range of interwoven or collective cognitive operations. In turn, ecological thinking (or ecological rationality) (Gigerenzer and Selten, 2002; Raab and Gigerenzer, 2005; Gigerenzer and Brighton, 2009) sees strategies as processes that exploit a large amount of data and knowledge, requiring a lot of computational effort; heuristics, on the other hand, perform simple and effective moves, even if they are less rigorous. Cognitive heuristics are simply described as "rules of thumb" in various parts of computer and cognitive science studies. I'll take the widely held AI viewpoint that sees strategies as a collection of successively selected acceptable heuristics.

I believe that we can evaluate what I just called strategic cognition in the context of Magnani's (2009) studies on abductive cognition, emphasizing the contrast be-tween what I called locked and unlocked strategies (see also Magnani (2019, 2020). I also used these ideas as components of the framework of an altogether new dynamic approach to the nature of computation in my recent book (Magnani, 2022b). In this theory, I emphasize the importance of unconventional computing as a continuous and fantastic process of cognitive domestication of ignorant entities.

Deep learning machines (and thus AlphaGo/AlphaZero programs) are assumed to work with locked strategies, a fact that has a big impact on the type of creativity they may produce. I have extensively explained in my studies that the various sorts of human, animal, and computational hypothetical cognition can be accounted for using the crucial idea of abduction. Selective abduction (Magnani, 2001) – for example, in medical diagnosis (where we must "choose" from a "repository" of al-ready accessible hypotheses) – and creative abduction (Magnani, 2001) – were introduced (abduction that provides new

DOI: 10.4018/978-1-7998-9220-5.ch046

hypotheses).[1] Not only, I have always emphasized that abduction is not only sentential, in the sense that it is carried out using human language resources (oral or written, or artificially constructed using symbols, as in mathematics and logic), but also "model-based" and "manipulative". Model-based abduction is concerned with the exploitation of internal cognitive acts that use models such as simulations, visualizations, and images; manipulative abduction is concerned with the exploitation of the so-called external character of human cognition, in which what I have dubbed the "eco-cognitive" character of cognition is central, because we must consider all those cognitive processes (embodied, embedded, situated, and enacted) in which the function of external models (for example, artifacts) is important. In this situation, manipulative action can provide new data – previously unavailable – and new heuristics capable of improving the methods agents use to solve challenges that necessitate the creative production (or just the selection) of relevant hypotheses. I argued that manipulative abduction is a type of "thinking via doing" and not only about doing in a pragmatic sense (cf. Magnani 2009). It is patent that we face cases of manipulative abduction in the case of deep learning machines (and in the case of the games we are considering in this article): the cognitive processes are intrinsically linked to the manipulation of the stones, and several cognitive embodied moments are at stake, along with the required visualization of the entire external context, the competitor, and so on.

BACKGROUND

It is commonly known that abduction research has advanced our understanding of creative cognition, even in the simple situation of a play in a Go game. The two main concepts I recently introduced during my investigations on abduction, knowledge-enhancing abduction and eco-cognitive openness, explained in (Magnani, 2017), are an excellent resource for delving into the locked and unlocked abductive strategies I mentioned earlier. Locked and unlocked strategies are important conceptual tools for analyzing the central cognitive features of deep learning machines. These strategies are present in human cognition, but they are currently impossible to locate in machines: yes, they yield creative outputs, but they are sadly endowed with varying degrees of creativity, and the presence of locked strategies in computing machines jeopardizes high level creativity. These considerations become obvious in the case of deep learning programs like AlphaGo, which try to automate various sorts of abductive reasoning.

I do believe that these programs exhibit what I call locked abductive strategies, which exhibit weak (even if amazing) types of hypothetical creative cognition because they are constrained in what I call eco-cognitive openness, which is more typical of human cognitive processes dealing with abductive creative reasoning: cognitive strategies are unlocked in this last human case. The fact that these programs are not based on logic, and that the main intellectual tradition associated to a formalization of abduction was in fact based on logic, does not pose an issue for the arguments I have illustrated in (Magnani, 2019). Indeed, abduction can occur at a sub-symbolic level in both humans and machines, so it is no surprise that deep learning machines can produce abductive results in this way. Not only, we have to remember that abduction is also characterized by multimodality (for example, it can be performed exploiting diagrams), as I have described in (Magnani, 2009 and 2017). Indeed, we must emphasize that humans frequently guess abductive hypotheses not only as a result of manipulating the external world, which is appropriately filled with cognitive representations and appropriate artifacts but also as a result of embodied and unconscious capacities (which also characterize some aspects of cognition in higher mammals, that surely do not take advantage of symbolic syntactic language). AI has always presented a variety of