Computational creativity beyond machine learning

Physics of Life Reviews, https://doi.org/10.1016/j.plrev.2020.06.007
(This is author's version, please see the DOI above for the authoritative version).

Hannu Toivonen
Department of Computer Science and HIIT, University of Helsinki, Finland
hannu.toivonen@helsinki.fi, http://www.cs.helsinki.fi/u/htoivone/

"*The difference between greater and lesser creativity lies not in how you solve problems, but rather in what problems you choose to solve"* (Getzels and Csikszentmihalyi, 1976).

IDyOT, Information Dynamics of Thinking, is a minimal cognitive architecture motivated by human cognition and creativity (Wiggins 2018). From an artificial intelligence point of view, where a lot of research is invested into automating problem solving, it is interesting how IDyOT is *not* goal-oriented but rather driven by its cognitive cycle and information theory.

IDyOT is based on the mathematical principle of information efficiency: information is used to learn representations from the data, to assist perception via prediction, and to direct attention. Learning is at the core of IDyOT, but its use is subtly different from mainstream machine learning.

Machine learning can be characterized as search for generalisations from data. These generalisations then afford production of appropriate response in novel situations, as long as they are sufficiently similar to situations encountered before. All this also holds for IDyOT. It is an emergent cognitive architecture in the sense that it learns its internal representations using the principle of information efficiency. Inspired by Gärdenfors' (2000) theory of conceptual spaces, however, IDyOT aims to bridge non-conscious cognitive processing with symbolic, conscious awareness, so the emerging representations are symbolic.

In many machine learning settings, such as classification or regression, the response of the system cannot be novel; it is, by construction, a familiar one even if the situation is novel. Recently, generative machine learning models have been used increasingly to produce artefacts in much more open classes, e.g. music, text or images. In these cases, the response often is novel. The same goal still holds: identifying or generating an appropriate response to given data.

The operation of IDyOT is based on prediction and expectation as tools toward information efficiency. In contrast to plain machine learning, the anticipatory cognitive cycle of IDyOT is active also in the absence of input (i.e. sensory stimuli). This allows concepts to develop in the non-conscious mind as response to the "noise of the resting brain" and then possibly to appear in the conscious mind as creative ideas.

The representations learned by IDyOT have several layers of abstraction, affording imagination—or hallucination?—on different levels of abstraction. Deep models used in machine learning have similar

structure, and future work on the IDyOT theory hopefully will shed light on its relationship to other learning architectures.

IDyOT is an interesting model for creativity, affording emergence of both conceptual representations and concepts (cf. Xiao et al. 2019). Furthermore, IDyOT is not necessarily externally triggered nor goal-oriented. Unlike regular intelligent systems, IDyOT promises to be able to set its own goals: goals presumably emerge from IDyOT's current state and its current focus of attention. Non-conscious development of goals thus seems to be intertwined with the development of the respective ideas or solutions, eventually giving an "experience" of goal-oriented operation.

The ability to set one's own creative goals is a major step towards greater creativity. A computational system, such as an implementation of IDyOT, cannot be considered a creator on its own right unless it has an ability to pursue a course independent of its developer's intentions (Jennings 2010). Given that IDyOT is described as symbolic, it should be feasible—and most interesting—to examine goals that emerge in actual systems and how they relate to the data observed by the system. Even more interesting is if IDyOT's imagination, together with its different levels of abstraction, produce novel concepts that could be meaningfully communicated to outsiders.

Creativity is often defined as the ability to produce novel and valuable ideas or artefacts (Boden 2004). A main contribution of IDyOT is to propose a minimal cognitive architecture that affords creativity of different kinds, including transformational creativity where the creator modifies its creative spaces or rules, even its goals. Given the view that creativity is not the product of a single mind but rather of cultural, social and personal forces (Csikszentmihalyi 1988), an interesting topic for future research would be social IDyOTs developing traditions and conventions between themselves as well as with humans.

References

Boden MA. *The creative mind: myths and mechanisms*. 2nd edition. London, UK: Routledge. ISBN 0-349-10469-7, 2004.

Csikszentmihalyi, M. Society, culture and person: a systems view of creativity. In R. J. Sternberg (Ed.), *The nature of creativity: Contemporary psychological perspectives* (pp. 325–339). Cambridge, UK: Cambridge University Press. 1988.

Gärdenfors Peter. *Conceptual spaces: the geometry of thought.* Cambridge, MA: MIT Press; 2000.

Getzels, J.W., Csikszentmihalyi, M.: *The Creative Vision: A Longitudinal Study of Problem Finding in Art.* John Wiley & Sons, New York, 1976.

Jennings KE. Developing creativity: artificial barriers in artificial intelligence. *Minds and Machines* 2010, **20**: 489– 501.

Wiggins, GA. Creativity, information, and consciousness: The information dynamics of thinking. *Phys Life Rev* (2018).

Ping Xiao, Toivonen Hannu, Gross Oskar, Cardoso Amílcar, Correia João, Machado Penousal, et al. Conceptual representations for computational concept generation. *ACM Comput Surv* **52**, 1, Article 9, 33 pages. 2019. DOI:https://doi.org/10.1145/3186729