To appear in *Behavioral and Brain Sciences*

Deep-Learning Networks and the Functional Architecture of Executive Control

(Commentary on Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J. (In press). Building machines that learn and think like people. To appear in *Behavioral and Brain Sciences*.)

Richard P. Cooper
R.Cooper@bbk.ac.uk
Centre for Cognition, Computation and Modelling,
Department of Psychological Sciences,
Birkbeck, University of London

Abstract: Lake et al. underrate both the promise and the limitations of contemporary deep learning techniques. The promise lies in combining those techniques with broad multisensory training as experienced by infants and children. The limitations lie in the need for such systems to possess functional subsystems that generate, monitor and switch goals and strategies in the absence of human intervention.

Lake et al. present a credible case for why natural intelligence requires the construction of compositional, causal generative models that incorporate intuitive psychology and physics. Several of their arguments (e.g., for compositionality and theory construction, and for learning from limited experience) echo arguments that have been made throughout the history of cognitive science (e.g., Fodor & Pylyshyn, 1988). Indeed, in the context of Lake et al.'s criticisms the closing remarks of Fodor and Pylyshyn's seminal critique of 1980s-style connectionism make sobering reading: "some learning is a kind of theory construction … We seem to remember having been through this argument before. We find ourselves with a gnawing sense of deja vu" (Fodor & Pylyshyn, 1988, p. 69). It would appear that cognitive science has advanced little in the last 30 years with respect to the underlying debates.

Yet Lake et al. underrate both the promise and the limitations of contemporary deep learning (DL) techniques with respect to natural and artificial intelligence. While contemporary DL approaches to, say, learning and playing Atari games, undoubtedly employ psychologically unrealistic training regimes and are undoubtedly inflexible with respect to changes to the reward/goal structure, to fixate on these limitations overlooks the promise of such approaches. It is clear the DL nets are not normally trained with anything like the experiences had by the developing child, whose learning is based on broad multisensory experience and is cumulative with new motor and cognitive skills building on old (Vygotsky, 1978). Until DL nets are trained in this way it is not reasonable to critique the outcomes of such approaches for unrealistic training regimes, of, for example, "almost 500 times as much experience as the human received" (Lake et al., p. 17), for that 500 times as much experience neglects the prior experience that the human brought to the task. DL networks (as currently organised) require that much experience precisely because they bring nothing but a learning algorithm to the task.

A more critical question is whether contemporary DL approaches, might, with appropriate training, be able to acquire intuitive physics – the kind of thing an infant learns through their earliest interactions with the world (that there are solids and liquids, and that solids can be grasped and that some can be picked up, but that they fall when dropped, etc.). Similarly can DL acquire intuitive psychology through interaction with other agents? And what kind of input representations and motor abilities might allow DL networks to develop representational structures that support reuse across tasks? The promise of DL networks (and at present it remains a promise) is that, with sufficiently broad training, they may support the development of systems that capture intuitive physics and intuitive psychology. To neglect this possibility is to see the glass as half empty, rather than half full.

The suggestion is not simply that training an undifferentiated DL network with the ordered multisensory experiences of a developing child will automatically yield an agent with natural intelligence. As Lake et al. note, gains come from combining DL with reinforcement learning (RL) and Monte Carlo Tree Search to support extended goal-directed activities (such as playing Atari games) and problem solving (as in the game of Go). These extensions are of particular interest because they parallel cognitive psychological accounts of more complex cognition. More specifically, accounts of behaviour generation and regulation have long distinguished between automatic and deliberative behaviour. Thus, the contention scheduling / supervisory system theory of Norman and Shallice (1986) proposes that one system – the contention scheduling system – controls routine, over-learned, or automatic behaviour, while a second system – the supervisory system – may bias or modulate the contention scheduling system in non-routine situations where deliberative control is exercised. Within this account the routine system may plausibly employ a DL-type network combined with (a hierarchical variant of) model-free reinforcement learning, while the non-routine system is more plausibly conceived of in terms of a model-based system (cf. Daw et al., 2005).

Viewing DL-type networks as models of the contention scheduling system suggests that their performance should be compared to those aspects of expert performance that are routinized or over-learned. From this perspective, the limits of DL-type networks are especially informative as they indicate which cognitive functions cannot be routinized and should be properly considered as supervisory. Indeed, classical model-based RL is impoverished compared to natural intelligence. The evidence from patient and imaging studies suggests that the non-routine system is not an undifferentiated whole, as might befit a system that simply performs Monte Carlo Tree Search. The supervisory system appears to perform a variety of functions such as goal generation (to create one's own goals and to function in real domains outside of the laboratory), strategy generation and evaluation (to create and evaluate potential strategies that might achieve goals), monitoring (to detect when one's goals are frustrated, and to thereby trigger generation of new plans/strategies or new goals), switching (to allow changing goals), response inhibition (to prevent selection of prepotent actions which may conflict with one's high-level goals), and perhaps others. (See Shallice & Cooper, 2011, for an extended review of relevant evidence, and Fox et al., 2013, and Cooper, 2016, for detailed suggestions for the potential organisation of higher-level modulatory systems.) These functions must also support creativity and autonomy, as expressed by naturally intelligent systems. Furthermore "exploration" is not unguided as in the classical exploration / exploitation trade-off of RL. Natural intelligence appears to combine the largely reactive *perception-action* cycle of RL with a more active *action-perception* cycle, in which the cognitive system can act and deliberatively explore in order to test hypotheses.

To achieve natural intelligence it is likely that a range of supervisory functions will need to be incorporated into the model-based system or as modulators of a model-free system. Identifying the component functions and their interactions, i.e., identifying the functional architecture (Newell, 1990), will be critical if we are to move beyond Lake et al.'s "character" and "frostbite" challenges, which remain highly circumscribed tasks that draw upon limited world knowledge.

References

Cooper, R. P. (2016). Executive functions and the generation of "random" sequential responses: A computational account. *Journal of Mathematical Psychology*, *73*, 153-168. DOI: 10.1016/j.jmp.2016.06.002

Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature neuroscience*, *8*(12), 1704-1711. doi:10.1038/nn1560

Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, *28*(1-2), 3-71.

Fox, J., Cooper, R. P., & Glasspool, D. W. (2013). A canonical theory of dynamic decision-making. *Frontiers in Psychology, 4*(150), 1-19. DOI: 10.3389/fpsyg.2013.00150

Newell, A. (1990). *Unified Theories of Cognition.* Cambridge, MA: Harvard University Press.

Norman, D. A. & Shallice, T. (1986). Attention to action: willed and automatic control of behaviour. In R. Davidson, G. Schwartz & D. Shapiro (eds.) *Consciousness and Self Regulation, Volume 4*. NY: Plenum.

Shallice, T., & Cooper, R. P. (2011). *The Organisation of Mind*. Oxford University Press.

Vygotsky, L. S. (1978). Interaction between learning and development. In Cole, M., John-Steiner, V., Scribner, S., & Souberman, E. (eds.) *Mind in Society: The Development of Higher Psychological Processes*. (Chapter 6, pp. 79-91.) Cambridge, MA: Harvard University Press.