Turing, Ashby, and “the Action of the Brain”*

dr hab. Hajo Greif

NCN Grant Proposal, OPUS 19 Call, June 6, 2020

1 Research Objectives

Not very much has been written to date on the relation between Alan M. Turing and W. Ross Ashby, both members of the “Ratio Club” (1949–1958). Not much of the communication between the two seems to have been preserved or discovered either, besides citing and briefly discussing a letter from Turing to Ashby in which he suggested using an early digital computer for “producing models of the action of the brain” (Turing 1946). Despite the personal acquaintance between them, and despite the partial proximity of their research fields, the two are often cited as respective figureheads of the competing research programmes of Artificial Intelligence (AI) and cybernetics.

In objection to this dichotomy, this inquiry in history and philosophy of science pursues two complementary aims: First, it is a historical investigation into the interactions between their related-but-distinct views. Second, it will help to answer two closely related systematic questions:

Q1 What are the epistemically relevant formal and material properties required of the “models of the action of the brain”?

Q2 What are the relevant formal and material properties ascribed to “the action of the brain”?

There are various key motives shared between Turing’s and Ashby’s work, and there are elements to be discovered in their writings and their modelling endeavours that would later figure both in AI and in cybernetics. Both authors believed that “the action of the brain” can be subject to a method of modelling that casts it in a mathematical description and breaks it down into elementary operations in such a way that the model could be implemented in some kind of machine.

However, on a preliminary take on Q1, Turing and Ashby differed, first, in their formal versus material approaches to modelling (see Black 1962; Hesse 1966) and, consequently, in choosing digital versus analog machine models. Second, they diverged in their takes on the mathematical methods involved. Third, their interpretations of the models’ target systems were at variance: Ashby (1947, 1960, 1962) was concerned with adaptive behaviours of brains and other systems, their functions and their relationships to their environments, all understood in explicitly Darwinian terms. However, he restricted his focus to the origins of adaptive behaviour by learning, leaving aside “genic” adaptation, and therefore the organic basis of that behaviour. Conversely, Turing developed a notion of idealised theoretical machines, known as “logical computing machines”, which originally served metamathematical purposes but informed the concrete design of the digital computer. He used his theoretical machines for inquiries into a varied set of phenomena, from proto-connectionist models of the brain (1948) via simulation of conversational behaviour (1950) to pattern development in organisms (1952). Notably, in the latter he relied on the non-Darwinian account of morphogenesis in Sir D’Arcy Thompson’s *On Growth and Form* (1942).

* An earlier version of this proposal was admitted to the second stage of merit-based evaluation in OPUS 17 but ultimately not selected for funding by Panel II. The new version has been revised in light of the earlier reviews.

On a preliminary take on Q2, Turing's and Ashby's approaches to modelling cognitive phenomena each in their own way display a notable indifference towards those capacities which are considered the key constituents of human thought in classical cognitive science and AI: semantics and symbolic representation. Although these concepts became central only in the course of the establishment of those research programmes, one might expect that two of their founding figures either anticipated them to some degree or could be retrospectively aligned with them. Instead, Turing and Ashby each in their own way partly prefigure what would become alternative approaches to classical AI and cognitive science: views of cognition as being and "situated" or "embodied", "embedded", "extended" and "enactive" (known as "4E cognition"; see, for example, Chemero 2009; Clark 1997; Hutto and Myin 2013; Noë 2004; Varela et al. 1991). Turing was concerned with questions of symbolic representation only to the extent that they can be subject to computational modelling and only as one among other phenomena of interest. Ashby, in turn, entirely disregarded those questions in favour of non-symbolically modelling the foundational mechanisms of adaptive behaviour in expressly mechanistic but non-computational fashion.

1.1 Research Questions

On the background of the preliminary observations, three research questions that specify the general questions Q1 and Q2 and two sets of working hypotheses can be articulated:

RQ 1 What influence did the choice of modelling methods, especially the role of formal and material models, have on Turing's and Ashby's theorising about cognitive phenomena?

RQ 2 What influence did the biological sciences of their time, in particular their views of Darwinian evolution, have on Turing's and Ashby's theorising about cognitive phenomena?

RQ 3 What bearing do Turing's and Ashby's partly contrasting views on modelling (RQ 1) and biology (RQ 2) have on "situated" or "4E" approaches in contemporary cognitive inquiries?

A key aspect of the proposed inquiry will be to elucidate the implications of the answers to RQ 1 and RQ 2 for the theme of RQ 3, namely an increased focus on cognition as being of the "4E" kind, at the expense of the classical questions of semantics and symbolic representation. This shift of perspective occurs in numerous contemporary cognitive inquiries. The apparent parallels between these approaches and Turing's and Ashby's views may have either of two possible sources: First, the parallel may be coincidental while pointing towards the same set of fundamental problems of representationalist accounts in AI and cognitive science. Second, contemporary 4E approaches may be directly or indirectly informed by Turing's and Ashby's views – which seems straightforward to establish for the the cybernetic lines of reasoning that directly build on Ashby, but which is less obvious for Turing, who is usually taken to have inaugurated the research programme of AI.

1.2 Working Hypotheses

The proposed research will be organised around two working hypotheses, one historical, one systematic.

The historical working hypothesis

H There are two seeming lacunae in either author's argument:

- *h* 1 The question of the evolutionary origins of the action of the brain, and thus the organic level of adaptation, is largely left out of the picture even where a behavioural level of adaptation is addressed.
- *h* 2 Turing and Ashby, when modelling "the action of the brain", do not systematically concern themselves with its capacities of symbolic representation.

> These seeming lacunae are not omissions, but characteristic of Turing's and Ashby's systematic theorising, and their similarities and differences.

Part of the purpose of this project is to elucidate the reasons for the former putative lacuna (h 1) and its possible bearing on the latter (h 2). In a first approximation, leaving the mechanisms of biological adaptation out of the picture may, in Turing's reliance on D'Arcy Thompson's view of biological form as well as in Ashby's focus on behavioural adaptation, partly owe to the fact that, at the time of their writing, the authors were not in a position to rely on what would become known as the "modern synthesis" in evolutionary biology (Depew and Weber 1995; Huxley 1942). The objects of that synthesis were the Darwinian mechanisms of evolution by random variation and natural selection and the statistical laws of mathematical population genetics. For the first time, a comprehensive, largely consensual and empirically grounded paradigm of Darwinian biology emerged that sent to the sidelines D'Arcy Thompson's laws of form and other non-Darwinian theories. Hence, Ashby's and Turing's partial omission of the biological roots of cognition might not be an omission after all, but owes to the state of the sciences of life in their time.

Two of the key insights for a systematic inquiry to be derived from this comparative discussion are, first, the general importance of biological and in particular evolutionary models of human cognitive abilities, and the focus on embodiment and environment implied by this perspective. Unless representational capacities are considered an epiphenomenon of evolved traits or entirely outside the scope of biological naturalism, they will be subsumable under the same kind of evolutionary model, too. Their structure and functions should be expected to build upon, and be derived from lower-level adaptive traits. Second, the relation between the formal or material nature of the concrete models used and the materiality and embodiment of its target systems will be of methodological importance in defining the scope of the models in question.

The systematic working hypothesis

S There is a core set of assumptions that appear in Turing and Ashby and that are shared by various 4E approaches to cognition:

- s 1 An empirically adequate model of cognitive abilities will trace, by formal or material means, an organism's embodied interactions with the environment.
- s 2 An empirically adequate model according to s 1 will address the biological functions of cognitive traits both on an evolutionary and on a behavioural level.
- s 3 An empirically adequate model according to s 1 and s 2 will trace back any representational capacities to embodied interactions with the environment.

The extent to which and the ways in which Turing and Ashby actually informed 4E approaches are a matter of empirical investigation.

There are two levels of biological function that can be identified for cognitive traits: first, the mechanisms for producing outward behaviours or inward representations of world affairs are a product of natural selection, in terms of having been relevant to the biological reproduction of an organism's ancestors, with representational capacities building upon more direct traffic with the environment in bottom-up fashion. Second, the reference of the behaviours and representations to certain world affairs is fixed in the course of the reproduction of those behaviours and representations by the former mechanisms in processes of learning and cultural transmission.

A model of cognitive mechanisms and the behaviours and any representations they produce will be most far-reaching if it succeeds in not only formally describing the processes of variation and selection in question, but also in simulating them on the various levels identified. If arguments from embodiment and enaction are defensible, such a model will not merely benefit from being embodied in some form, but will have to interact with a natural environment in order to be instructive in the first place. Only then will such a system be able to approximate the mechanisms present in, and the environmental conditions present to, a natural cognitive system.

2 Significance of the Project

Turing and Ashby have been granted preference over other early proponents of cybernetics and computer science on the grounds, first, of preliminary evidence of the specific relationship between their views (Asaro 2011; Greif 2018) and, second, because of their concrete relationship as members of a group whose ideas were formative to the very notion of machine models of cognition (Dewhurst 2018; Husbands and Holland 2008).

By tracing some of the lesser-known interrelations between Turing's and Ashby's work, their influences and their import, the proposed research will help to resolve the seemingly strict, and intellectually sterile, symbolic / embodied dichotomy that has dominated the debates in and around AI for many years: The primary targets of modelling the "action of the brain" envisioned by the two authors might not be higher-order cognitive functions and symbolic representation but the basic forms of neuronal processes and the adaptive functions of the brain as a biological organ respectively. There is no straightforward mapping from Turing's work onto the tradition of symbol-based AI on the one hand and from Ashby's cybernetics onto embodiment-centred views on the other. Instead, the relations between their views are uniquely disposed to explain, and possibly overcome, the seeming dichotomy between these two research programmes.

This latter aim in particular will move the project beyond a purely historical domain and establish its relevance to the debates in 4E cognition and Nouvelle AI, where sensing and interacting with real environments have become much more central to cognitive modelling than classical AI's chess playing or theorem proving. The systematic part of the inquiry will be innovative in showing a path towards a more finely grained, biologically grounded view than hitherto available of the action of the brain as a necessarily embodied, but not necessarily representational phenomenon.

3 General Concept and Plan of Research

The project will be organised into three parts. Parts I and II will address the historical (H) and systematic (S) working hypotheses respectively, whereas Part III is designed to forge a synthesis between them. The three parts will be subdivided into seven work packages (WPs) that are identified in relation to the research questions (RQ 2–RQ 3):

Part I: The Historical Perspective

- WP 1: Turing's and Ashby's views of modelling and simulation – philological work
- WP 2: Identifying Turing's and Ashby's biological reference points – philological work
- WP 3: Turing and Ashby on embodiment, environment and representation – philological and conceptual work

Part II: The Systematic Perspective

- WP 4: Contemporary views on embodiment, environment and representation – conceptual and philological work
- WP 5: Identifying Turing's and Ashby's influences on contemporary views – philological and conceptual work
- WP 6: Embodiment, environment and representation in contemporary cognitive science and AI – conceptual and empirical work

Part III: Synthesis

- WP 7: Synthesis of results of previous research tasks – conceptual work

These research tasks will be completed by a team of three researchers: the PI (who works on the project 0.5 full-time equivalent/FTE; specialisation in philosophy of AI and philosophy of technology), a co-investigator (0.25 FTE; philosophy of computer science) and a full-time postdoc (history and philosophy of science). The project is planned for a duration of 36 months.

4 Research Methodology

Five kinds of methods will be applied in this research project.

First, a significant part of the work will be philological in nature: surveying and analysing primary and secondary sources, retracing the key concepts, arguments and references used, considering their context, comparing and evaluating them.

Second, to the extent that preliminary inquiries indicate that there is material that has not been sighted and published to date, some archival work will be done in order to trace documents that will feed into philological analysis. Candidate places include the Alan M. Turing archive at King's College, Cambridge, the National Archive for the History of Computing, Manchester, the personal archive of W. Ross Ashby at the British Library and the John Bates Archive at the Wellcome Library for the History of Medicine, London.

Third, a significant part of the work will be argumentative, in terms of developing hypotheses on the grounds of the results of the philological and empirical work.

Fourth, some expert interviews with researchers in the fields of Nouvelle AI and situated approaches in cognitive science will complement the prior, more literature-based approaches, so as to empirically ground and provide a corrective to the claims that the conceptual work will give rise to (for an introduction to the method, see Bogner et al. 2009).

Fifth, first-hand expertise in computer science and AI will provide a second gauge of, and corrective to, claims on computer models developed on the philological and conceptual levels.

References

Asaro, P. M. (2011). Computers as Models of the Mind: On Simulations, Brains, and the Design of Computers. In: *The Search for a Theory of Cognition. Early Mechanisms and New Ideas.* Ed. by S. Franchi and F. Bianchini. Amsterdam/New York: Rodopi, 89–114.

Ashby, W. R. (1947). The Nervous System as Physical Machine: With Special Reference to the Origin of Adaptive Behaviour. *Mind* 56.221, 44–59.

— (1960). *Design for a Brain. The Origin of Adaptive Behaviour.* 2nd ed. New York/London: John Wiley & Sons.

— (1962). Simulation of a Brain. In: *Computer Applications in the Behavioral Sciences.* Ed. by H. Borko. New York: Plenum Press, 452–466.

Black, M. (1962). *Models and Metaphors.* Ithaca: Cornell University Press.

Bogner, A., B. Littig, and W. Menz, eds. (2009). *Interviewing Experts.* Research Methods. Basingstoke: Palgrave Macmillan.

Chemero, A. (2009). *Radical Embodied Cognitive Science.* Cambridge/London: MIT Press.

Clark, A. (1997). *Being There: Putting Brain, Body and World Together Again.* Cambridge: MIT Press.

Depew, D. J. and B. H. Weber (1995). *Darwinism Evolving. Systems Dynamics and the Genealogy of Natural Selection.* Cambridge: MIT Press.

Dewhurst, J. (2018). British Cybernetics. In: *The Routledge Handbook of the Computational Mind.* Ed. by M. Sprevak and M. Colombo. London: Routledge, 272–282.

Greif, H. (2018). 'The Action of the Brain'. Machine Models and Adaptive Functions in Turing and Ashby. In: *Philosophy and Theory of Artificial Intelligence 2017.* Ed. by V. C. Müller. Cham: Springer, 24–35.

Hesse, M. B. (1966). *Models and Analogies in Science.* Notre Dame: University of Notre Dame Press.

Husbands, P. and O. Holland (2008). The Ratio Club: A Hub of British Cybernetics. In: *The Mechanical Mind in History.* Ed. by P. Husbands, O. Holland, and M. Wheeler. Cambridge/London: MIT Press, 91–148.

Hutto, D. D. and E. Myin (2013). *Radicalizing Enactivism: Minds without Content.* Cambridge: MIT Press.

Huxley, J. (1942). *Evolution. The Modern Synthesis.* London: George Allen & Unwin.

Noë, A. (2004). *Action in Perception.* Cambridge: MIT Press.

Thompson, D. W. (1942). *On Growth and Form.* 2nd ed. Cambridge: Cambridge University Press.

Turing, A. M. (1946). *Letter to W. Ross Ashby of 19 November 1946 (approx.)* The W. Ross Ashby Digital Archive.

— (1948). *Intelligent Machinery: A Report by A.M. Turing.* Tech. rep. London: National Physical Laboratory.

— (1950). Computing Machinery and Intelligence. *Mind* 59, 433–460.

— (1952). The Chemical Basis of Morphogenesis. *Philosophical Transactions of the Royal Society, B* 237, 37–72.

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