

Homo Practicus: Practice Theory and Post-Cognitivist AI

Rodger Kibble¹

Abstract. We consider some similar trends in social theory and artificial intelligence, which have both seen a focus on autonomous, rational self-interested agents supplemented or opposed by normative frameworks where agent autonomy is circumscribed by various rules, protocols or implicit conventions. A further development in social theory has been the ‘practice turn’, the emergence of practice theory in the work of Bourdieu and others, which was at least partly stimulated by encounters with Heidegger and Wittgenstein. We conclude by considering whether more recent strands in AI such as ‘Heideggerian’ and enactive approaches can be seen as playing a similar role in AI as practice theory has in sociology.

1 INTRODUCTION

Social theory and artificial intelligence share the goals of characterising, describing and explaining the behaviour of persons/actors/agents. The agent-based tradition within AI [6] treats ‘agents’ as autonomous software entities which operate in some kind of social settings where they interact with other artificial agents and/or human principals, clients or users.

We can see some broad similarities in the way agents have been characterised in the different disciplines. Both have included a notion of the instrumentally rational, self-serving agent or *homo economicus* [15], who acts on the basis of information about the state of the world and a given set of ‘preferences’ in order to maximise some expected reward or ‘utility’. This approach has been formalised in AI as the BDI (belief-desire-intention) model in which agents maintain representations of the state of the world (*beliefs*) and generate plan-based *intentions* according to their encoded preferences or *desires* [6:155] (It is no accident that one of the founding figures in AI, Herbert Simon, was also a Nobel laureate in economics.) In both fields, this model has been opposed or supplemented by a ‘normative’ approach, where individual autonomy is seen to be

circumscribed by constraints of various kinds, whether these take the form of explicit rules or laws, or tacit/emergent conventions that give rise to sets of mutual expectations. (See e.g. [6], ch 10). One motivation for this approach is that agents’ ‘preferences’ are not necessarily pre-given and fixed, but may themselves be shaped by acquired social norms.

Finally, Reckwitz [15] identifies a third strand of ‘cultural theory’ of which practice theory is a subtype, which seeks to overcome a shared ‘blind spot’ in the ‘rational choice’ and ‘normative’ traditions by paying attention to the ‘implicit, tacit or unconscious layer of knowledge which enables a symbolic organisation of reality’. Practice theories see the individual as a ‘bodily and mental agent’ which ‘acts as the carrier [...] of a practice’ [15:246]; ‘A practice can be understood as the regular, skilful “performance” of human bodies’ [15:251]. This regard for the bodily nature of situated action echoes recent developments in artificial intelligence, a field which Bishop [1] sees as having reached a ‘branch point’ between classical, ‘cognitivist’ approaches which would encompass both BDI and norm-oriented models, and what he calls the ‘4Es’ framework of radical ‘Embodied, Embedded, Ecological, Enactivist’ approaches.

In the remainder of this short paper: section 2 makes some broad comparisons between the models of agency offered by social theory and AI, limited to the first two approaches mentioned above. Section 3 briefly introduces practice theory, introduced as a possible way around regress problems which threaten some standard accounts of normativity. These sections draw in particular on Reckwitz’s [15] account of practice theory against the background of classic ‘purpose-oriented’ and ‘norm-oriented’ approaches to ‘explaining action and social order’, on

¹ r.kibble@gold.ac.uk; Department of Computing, Goldsmiths University of London, New Cross, London SE14 6NW.

survey papers on practice theory by Rouse [16] and Schatzki [18], d’Inverno and Luck’s monograph *Understanding Agent Systems* [6] and on Hollander and Wu’s survey of normative agent-based systems [10]. Section 4 argues that related difficulties afflict normative approaches in both social theory and artificial agent-based systems, and tentatively considers whether ‘post-cognitivist’ AI is concerned with similar dissatisfactions with these two ways of thinking about agency. Finally in section 5 we speculate on possibilities for extending insights into both AI and social theory by fostering dialogue among practitioners in the two fields.

2 AGENT MODELS IN SOCIAL THEORY AND AI

Neoclassical economics and contemporary Rational Choice Theory assume that individuals are rational, self-interested decision makers (‘homo economicus’), coldly maximising their own interests even if they come at the expense of others. It is generally assumed that the sum total of egoistic actions tends to result in increased prosperity for the community as a whole – though this is unlikely to be evenly distributed among the populace. This characterisation tends to strike lay readers as a grotesque caricature, and studies have found that ‘homo economicus’ types only make up a small proportion of the population [9].

As noted above, rational agents within AI have been modelled using the BDI framework (i.a.), according to which agents carry out formal (mathematical or logic-based) reasoning in order to compute an effective way to achieve goals based on representations of the state of the world and their encoded preferences. In fact the notion of fully autonomous agents has proved problematic from both theoretical and practical points of view. This is especially borne out when we consider *communication*. Standard specifications for agent communication such as FIPA’s ACL [8] model utterances as actions which are intended to modify or update the belief state of a recipient agents. The specs assume that agents are capable of reasoning about their interlocutors’ prior beliefs. There is a tension here between the requirements of successful communication and the notion of agent autonomy, which would bar any agent from directly modifying another’s internal state. In practice, communicating

agents are likely to employ *protocols* which mandate the form of particular responses and cut down on the chain of reasoning required, at the expense of limiting agent autonomy [11].

We should also observe at this point that the “pure” model of agents as genuinely autonomous, generic reasoners is more of a concern for academic researchers than for developers in the software industry. Poslad [14] notes that

BDI models have incomplete axiomatisations and can be computationally complex or even intractable. The BDI model focuses on private belief and intention transfer between individuals. It doesn’t take into account third-party or societal interaction and associated constraints. BDI models seldom focus on pragmatic issues such as belief and intention management ... These can make the model computationally complex or even intractable.

Thus in the AI field of agent-based systems there are both theoretical and practical problems with the notion of pure autonomy.

An alternative to ‘homo economicus’ is the ‘homo sociologicus’, Reckwitz’s term for the norm-oriented frameworks of e.g. Durkheim and Parsons. These have their counterparts in the emerging discipline of normative multi-agent systems. A recent survey [10] proposes a consensus model of the “norm life cycle” incorporating the processes of *creation, transmission, recognition, enforcement, acceptance, modification, internalisation, emergence, forgetting* and *evolution*. The authors distinguish in terms of *norm origins* between *Type I* norms, which are decreed by an authority, and *Type II* which emerge from interactions between agents, either by spontaneous emergence or negotiation.

It should be stressed that while both disciplines seem to have taken similar turns, each has done so in part for reasons internal to the discipline: the ‘normative turn’ in AI has been driven as much by theoretical and engineering considerations as by taking inspiration from sociological theories.

3 THE 'PRACTICE' TURN IN SOCIAL THEORY

One can perhaps trace the genesis of practice theory to Marx's Theses on Feuerbach [13], with their insistence that 'all social life is essentially *practical*' ('Alles gesellschaftliche Leben ist wesentlich *praktisch*' emphasis in original). (Reckwitz inverts this formula, perhaps unwittingly, when he says 'To say that practices are "social practices" is ... a tautology: A practice is social...' [15:250].) However, the philosophical background of 20th century practice theory tends to be identified with the work of the later Wittgenstein [20] and the early Heidegger – indeed Reckwitz [14] claims that 'we find everything that is original in practice theory already in the work of these authors' (though he also states that they are 'hardly ever systematically scrutinised'). Rouse [16] identifies Wittgenstein's key contribution as his account of rule-following [20], and that of Heidegger as the claim that "all interpretation (including linguistic assertion) draws upon a more basic understanding or competence that is not explicitly articulated".

Brandom [4] claims to be following Wittgenstein in making a distinction between types of rule-following as *regulism* and *regularism*: in the first case, agents are taken to be following explicit rules, while in the latter the rule may be inferred from observations of regularities in behaviour, possibly including sanctions for 'deviancy'. Brandom argues that neither can serve as a foundational account of rule-following: regulism relies not only on the rules that have been explicitly promulgated but on rules about how to follow rules, the normative authority of the rule-giver and so on; while regularism fails because identifying a particular sequence of actions as a norm-conformant performance is inevitably a matter of interpretation, which may be shaped by the analyst's normative presuppositions. His position is that norms are ultimately grounded in 'social practices' which may not be amenable to further reductive analysis – as Wittgenstein [20] has it, one's spade eventually turns on 'bedrock' and all that can be said is '*so handle ich eben*' ('this is just what I do').

Bourdieu cites the same Wittgenstein passage in his highly influential [2]. He contrasts practices with rule-

governed behaviour by observing that the latter "ceases to convince as one considers the practical mastery of the symbolism of social interaction... presupposed by the most everyday games of social interaction", rather we apply a "practical knowledge" which functions like a "self-regulating device" providing for continual "adjustment of practices and expressions to the reactions and expectations of other agents". One is not consciously consulting internalised lists of rules for social behaviour, but carrying out appropriate practices in a constantly-changing social milieu according to the constraints of the *habitus*.

Practice theory is a term that has been applied to a variety of approaches (or practices?) in the social sciences and humanities that seek to study the behaviour of individuals in social contexts by focussing on performances classed as *practices* against a background of other practices, in place of such monolithic categories as culture, class, gender, rules, values, norms and so on. One motivation for this is that analysts can focus on observable events rather than postulating unobservable entities such as beliefs, values or traditions, or speculating about the psychology of the participants' motives. There is quite a range of activities that have been considered under this heading, from eating with specific utensils to gift-giving, playing chess or conducting scientific research. Rouse [16] provides a useful survey of practice theories across a variety of disciplines.

4 PRACTICE THEORY AND 'POST-COGNITIVIST' AI

We noted above that classical economics has assumed that the outcome of a series of spontaneous self-serving actions is an increased level of prosperity for the community, even if no individual has this as an explicit goal. In the field of agent-based modelling (ABM) a number of researchers have sought to demonstrate that social norms can emerge from repeated interactions between instrumentally rational agents, again without any agent deliberately seeking to establish such a norm. We have argued elsewhere [12] that attempts to model normativity as emergent or supervenient on instrumental calculations are subject to the regress problems identified by

Brandom, whether the norms under consideration are Hollander and Wu's Type I ('regulist') or Type II ('regularist'). For example, Savarimuthu et al [17] consider a scenario where a traveller is trying to establish whether a tipping norm is in place in a particular locality by observing whether waiters are disposed to 'sanction' customers who leave a small or no tip. But this assumes that the traveller is also able to identify sanctioning actions, which are themselves norm-governed, may well not be transparent to strangers in the region, and that the traveller can tell whether they are done correctly or incorrectly.

The notion of 'practice' as 'regular, skilful "performance" of human bodies' seems congruent with Bishop's [op cit] statement that 'At a fundamental level, enactivism is anti-dualist: the self arises as part of the process of an embodied entity interacting with its *umwelt* in precise ways determined by its physiology.'

Dreyfus [7] rehearses his well-known objections to 'cognitivist' AI as a recapitulation of the errors of rationalist philosophy, singling out the 'relevance problem' of identifying which of myriad atomised fragments of 'sense data' are significant or relevant to the task at hand. A special case is the well-known 'frame problem': if there is some change in the environment, perhaps as a result of an agent's actions, how should it determine what else changes and what stays the same, and whether it matters? For Dreyfus, treating the computer or brain as "a passive receiver of bits of meaningless data, which then have to have significance attached to them" is a dead end. Rather, the task of developing human-level or even animal-level AI must build on an insight he attributes to Heidegger: that we somehow respond directly to what is significant or relevant in the environment, and that this ability is more fundamental than "*thinking* and solving problems" (emphasis in original). He argues that none of the supposedly post-cognitivist, 'Heideggerian' approaches has succeeded in modelling the ways that humans 'cope with what is significant', and appears pessimistic of any prospect of success any time soon. He envisages a mission for 'Heideggerian AI' as follows:

Showing in detail how the representational un-ready-to-hand in all its forms depends upon a background of holistic,

nonrepresentational coping is exactly the Heideggerian project and would, indeed, be the most important contribution that Heideggerian AI could make to Cognitive Science.

In fact this 'holistic, nonrepresentational coping' sounds somewhat like the 'implicit, tacit or unconscious layer of knowledge which enables a symbolic organisation of reality' with which social practice theory is concerned, according to the passage from Reckwitz quoted in section 1. From this point of view, it may be that success in this venture could potentially make important contributions to both social and cognitive science.

Collins [5] discusses 'tacit knowledge' from the point of view of a practising sociologist, and foregrounds the idea that this kind of knowledge is embodied within learned motor skills: for instance one can learn how to ride a bicycle, dance, or swing a golf club through repeated practice, but it is hardly possible to convey this knowledge to another person simply by giving them a list of instructions, however explicitly formulated. He argues that there is little hope for 'classic' symbolic or connectionist AI systems to replicate the knowledge/skills (what Brandom might call 'practices-or-abilities') which are needed for socially embedded practices such as riding a bicycle in traffic, which involves not only maintaining one's balance but paying constant attention to the actions and likely behaviour of fellow road-users.

5 CONCLUSION/DISCUSSION

A minimal aim of this paper has been to point out that social practice theories and post-cognitivist schools in AI share some areas of concern and display a similar stance towards the way humans engage with themselves and the world, and that in both fields this stance has in part arisen in opposition to 'classic' individualist or normative models of agency. I have found little evidence of exchanges of ideas between practice theoreticians and AI researchers such as 'radical enactivists', though it is worth noting that Bourdieu's [3] is cited in [19] even if his name does not appear in the index. We have sought to show in this short paper that there are researchers in both fields

who have followed a broadly similar path, with comparable decisions being taken for different but hopefully mutually reinforcing reasons. I believe there is further scope for two-way dialogue between these enterprises: for example, social theory can provide a broader range of exemplars for post-cognitive AI than the usual (Heideggerian) business with hammering, opening doors and picking things off tables. And distinctions such as those Reckwitz [15] proposes between practice theories and the ‘intellectualist’ cultural theories of mentalism, textualism and intersubjectivity may provide AI researchers with a more fine-grained framework for agent-based simulations. On the other hand, AI and Cognitive Science can hopefully assist social theorists toward a more concrete understanding of what it means for agents to operate ‘skilfully’ with a layer of ‘tacit or unconscious knowledge’.

REFERENCES

- [1] Bishop, JM, ‘Artificial Intelligence at a New Branch Point’, introduction to Kibble et al (eds), *Selected Papers from AISB50*, 2014.
- [2] P. Bourdieu, *Outline of a Theory of Practice*, Cambridge University Press, Cambridge, 1977. translated by R. Nice.
- [3] Bourdieu, P *Logic of Practice*. Polity, 1990.
- [4] Brandom, R *Making it Explicit*, Harvard University Press, 1994
- [5] Collins, HM, What is tacit knowledge? In in T Schatzki, K Cetina and E von Savigny (eds). *The Practice Turn in Contemporary Theory*, Routledge, 2001.
- [6] d’Inverno M and M Luck *Understanding Agent Systems*, Springer, 2004
- [7] Dreyfus, H ‘Why Heideggerian AI failed and how fixing it would require making it more Heideggerian’, *Artificial Intelligence* 171 1137-1160. 2007.
- [8] FIPA: Communicative Act Library Specification: SC00037J, Dec. 2002.
- [9] Henrich, J et al, ‘In search of Homo Economicus’, *American Economic Review* 91/2, 2001.
- [10] Hollander, CD,Wu, AS (2011) The current state of normative agent-based systems. *Journal of Artificial Societies and Social Simulation*, 14.
<http://jasss.soc.surrey.ac.uk/14/2/6.html>.
- [11] Kibble, R ‘Speech Act Theory and Intelligent Software Agents’ In Sbisà and Turner (eds), *Pragmatics of Speech Actions*, 2013.
- [12] Kibble, R ‘Regulism, regularism and some limitations of agent-based modelling’. In *Proc of AISB 2013 symposium, Social Path*, University of Exeter April 2013.
- [13] Marx, K ‘Thesen über Feuerbach’, in I Fetscher (ed), *Marx/Engels Studienausgabe I*, Fischer, 1966.
- [14] Poslad, S. Specifying Protocols for Multi-Agent Systems Interaction. *ACM Trans. Autonom. Adapt. Syst.* 2, 4, Article 15 (November 2007)
- [15] Reckwitz, A. ‘Towards a theory of social practices: a development in culturalist theorising’. *European Journal of Social Theory* 2002.
- [16] J. Rouse. Practice theory. Division I Faculty Publications. Paper 43, 2007. http://wescholar.wesleyan.edu/div1_facpubs/43.
- [17] Savarimuthu, BTR, Cranefield, S, Purvis, MA and Purvis, MK, (2010) Obligation norm identification in agent societies. *Journal of Artificial Societies and Social Simulation*, 13.
<http://jasss.soc.surrey.ac.uk/13/4/3.html>
- [18] Schatzki, T, ‘Introduction: Practice Theory’ in T Schatzki, K Cetina and E von Savigny (eds). *The Practice Turn in Contemporary Theory*, Routledge, 2001.
- [19] F. Varela, E. Thompson and E. Rosch, *The Embodied Mind*, 1991.
- [20] Wittgenstein, L *Philosophische Untersuchungen*, Suhrkamp, 1953.