

Review of: Idealizations of Uncertainty, and Lessons from Artificial Intelligence

Reviewer: David Hales, University of Szeged, Hungary, July 2015

I would recommend publication subject to minor corrections if the editors think it is in scope of their issue.

I was asked to do the following:

“I would appreciate if you could prepare a brief report focussing on two questions: (i) Is the contribution of the paper potentially significant? (ii) Is the analysis correct? Referees are only expected to explain the paper's main strengths and weaknesses, NOT to rewrite the paper or change its main thrust. Please note that all referee reports for Economics are published on the journal's website (please indicate whether anonymous or attributed). If you want to submit comments which are meant to be shared with me or the author alone, you should provide them in a cover letter.”

Find report below:

(i) Is the contribution of the paper potentially significant?

The paper takes the form of a critique of the existing approaches of AI as models of the way humans actually behave in the context of uncertainty. It includes some basic background on those approaches which would be of value to readers who do not have much knowledge in the area – specifically those from a purely economics background. The critical stance is instructive and valid in my opinion – and I am someone with a background in AI. However, I am sure that some others within the AI field might disagree with some aspects of the critique.

Specifically the main contribution could be to disabuse those who through lack of knowledge, and exposure to hype, believe that AI techniques are more advanced than they actually are with respect to modelling actual human level reasoning.

I think the identification of “wishful mnemonics” and their effect on misunderstands is very relevant and interesting (not just in AI but in modelling generally, including economic and agent-based modelling). The analysis given of how language itself develops based-on and reinforcing such confusions is particularly thoughtful and insightful.

(ii) Is the analysis correct?

Overall the factual aspects are correct and the critique is persuasive.

There are some curious gaps in emphasis however. For example in section 3.2, AI models of learning, we jump straight into (sub symbolic) connectionism as the main paradigm without mention of the symbolic more traditional (and huge area) of machine learning that predates much of the connectionist work. Specifically classifier systems that induce actual symbolic human readable rules. This is minor

issue in the context of the critique but could mislead the reader into thinking that basically AI learning is connectionism or pure statistics – rather than attempts to induce meaningful symbolic representations that make sense also to humans.

In a few instances the author makes very general statements claiming the status of obvious fact when these are contentious and debatable. For example it is stated that:

“It is true that at some (neural) level human thinking is mechanical, and explicable by equations of some sort”

Why is it true? Is there any proof of this? This is an assumption rather than a truth. It might be useful to qualify these kinds of statements in some way because they detract from the plausibility of the main argument.

The main conclusion to the argument is that due to the problems of existing AI models one must look at how people actually behave. However little, if any, attention is given to practical ways forward in this regard or indeed past successes or failures following this approach which surely exist?

In fact the final parts of the paper are weaker than the earlier parts because it is rather vague as to what is to be done. For example the author states:

“Considering the realities of the in vivo social communication of ideas may lead to understanding of the juxtaposition of representations from those shared ideas, which may provide a basis for understanding innovation of ideas themselves”

It is not clear what this means or how it would inform practical work.

Caveats aside the paper is persuasive in its general thrust and valid in its claims and could be very useful for those coming new to many of the “wishful mnemonics” that AI researchers love to use and the hype often associated with them.

Notes for author:

Typos:

Top of page 3 “rationale” -> “rationality”

Top of page 7 “funding for AI researcher” -> “researchers”

Bottom page 21 “size of date” -> “data”

Section 5 “Lane and Mayfield” -> “Maxfield”

Penultimate paragraph page 24 “aid in constructive” -> “construction”

Page 29: Koestler, A. (1763). The Act of Creation. -> “(1963)”