

Aristotle's *Phronesis* and Normative Algorithmic Computability

In the *Nicomachean Ethics* (VI 1139a), Aristotle introduces practical wisdom as the virtue of the rational faculty that deals with the necessary, but variant and context-sensitive truths of the practical domain (what he calls the λογιστικόν). That is, practical wisdom is the virtue that deals with context-sensitive knowledge of what is to be done. Practical truths, Aristotle warns us, are not invariant, absolute and unchangeable and involve both universal truths (καθόλου) and particular truths (καθ' ἑκαστα) (1141b).

For Aristotle's account of ethics, there are no exceptionless, universal practical truths on the basis of which we can compute in deductive fashion what is to be done across contexts. Hence, practical truths are in principle irreducible to algorithmic computability. With 'algorithmic irreducibility', I mean deductive, rule-based, procedural knowledge. Call this '*the computational irreducibility of practical normativity*'. Instead, Aristotle makes clear that phronesis induces virtuous intuition that enables the *phronimos* (i.e. the virtuous) agent to grasp 'with the mind's eye' (1144a) basic practical principles and what is to be done in a given practical context. Indeed, he explicitly draws an analogy between virtuous practical intuition and how mathematical intuition grasps basic, self-evident mathematical truths (VI 1141b-1142b, 1144a, VII 1150b-1151a). Call this '*the virtue intuitionist thesis*'.

Aristotle's aretaic account of practical knowledge has serious implications about the prospects of AI developing practical intelligence. For if practical normativity is computationally irreducible and we require virtuous intuition embedded in a particular context in order to grasp what is to be done, then the prospects of AI developing normative intelligence should be considered dim. This is both because AI relies on algorithmic computability and because AI systems lack the cognitive prerequisites for rational intuition (e.g. conscious mental states). This implies that current AI cannot develop human-like, autonomous, artificial practical intelligence across contexts.

Of note, is that virtuous human intelligence can, all other things equal, address such novel practical challenges. A virtuous agent who has developed phronesis, that is, a virtuous sensitivity that informed by the peculiarity of a novel practical situation can deliberate about, or even straightaway intuit, what is to be done in the given circumstances. Think for example of the first virtuous agent that intuitively appreciated that s/he ought to help a friend who went bankrupt due to the collapse of stock market. The practical challenge was novel and called for an act that went over and beyond anything previously known.

It might be objected with some justification that this understanding of AI relies on outdated, top-down, deductive symbolic AI (so-called GOF AI) and that sophisticated forms of bottom-

up, statistical, machine learning can address the problem of normative knowledge. That is, we can machine train AI on the basis of huge datasets on how to best act in novel practical situations that pose original unprecedented moral challenges. Given the sheer enormity of the number of moral challenges that we can machine train an AI system and that no human can face and train himself with such a great number, an AI system can be better prepared to face the myriad of moral challenges than any human being. Indeed, there is some empirical evidence that with regard to at least nine basic moral challenges, LLMs such as Chat-GPT can fare better than the average human agent (cf. Khandelwal (2024)).

However, even the machine learning approach cannot develop autonomous, fully reliable normative knowledge across contexts. First, machine learning relies on the quantity and quality of the given datasets and this is a serious constraint. Second, virtuous human agents often need not rely on any previous reflection or deliberation. They can rely on virtuous intuition, but AI is a mechanical system that lacks conscious mental states, such as intuitions, virtues and feelings.

Third, the computational complexity of the normative is in principle infinite and no matter how large the finite datasets are and how many moral situations and challenges they include, they will never be enough to completely anticipate the infinity of possible practical challenges. It is always possible that a practical challenge will be so original and unprecedented that it will outrun the extant algorithms mined from the datasets and, therefore, surprise the algorithms. In contrast, like human linguistic ability, human normative ability is, in principle, infinite. If developed correctly from early childhood via habituation and become virtuous, it could be highly reliable across contexts. It could be practically wise.

Again, some may object that AI systems need not fare worse than human agents because no human agent, no matter how virtuous, can be fully reliable in knowing right from wrong across contexts (cf. Turing (1950) for a similar point). True enough, no one is a fully virtuous agent. But the difference between virtuous human intelligence and artificial machine learning is that the first is unconstrained and can always get to the truth via dialogical reasoning, argument and intuition. Instead, machine learning will always be constrained by their finite datasets and their algorithms. I conclude that artificial practical intelligence should be always supervised by virtuous human intelligence.