Filosofia Unisinos
Unisinos Journal of Philosophy
18(3):195-200, sep/dec 2017
Unisinos – doi: 10.4013/fsu.2017.183.11

PHILOSOPHY SOUTH

No rationality through brute-force

Danilo Fraga Dantas[1]

ABSTRACT

All reasoners described in the most widespread models of a rational reasoner exhibit logical omniscience, which is impossible for finite reasoners (real reasoners). The most common strategy for dealing with the problem of logical omniscience is to interpret the models using a notion of beliefs different from explicit beliefs. For example, the models could be interpreted as describing the beliefs that the reasoner *would hold* if the reasoner were able reason indefinitely (stable beliefs). Then the models would describe maximum rationality, which a finite reasoner can only approach in the limit of a reasoning sequence. This strategy has important consequences for epistemology. If a finite reasoner can only approach maximum rationality in the limit of a reasoning sequence, then the efficiency of reasoning is epistemically (and not only pragmatically) relevant. In this paper, I present an argument to this conclusion and discuss its consequences, as, for example, the vindication of the principle 'no rationality through brute-force'.

Keywords: finite reasoning, logical omniscience, efficient reasoning, asymptotic analysis, computational complexity.

Introduction

Rationality is often studied as if it were independent from the limitations of the cognitive structures that implement it. This is an example of this widespread attitude:

It can be epistemically rational for a person *S* to believe even that which, given his circumstances or given his limitations as a believer, he cannot believe. It also can be epistemically rational for *S* to believe that which, given his circumstances, or given his limitations as a believer, he cannot help but believe (Foley, 1987, p. 13).

An example of such limitations is the (limited) amount of (cognitive) resources available for reasoning (specially, memory and time)[2]. I agree that rationality is independent from the possession of *specific* amounts of cognitive resources. For example, I did not become less rational because my memory has worsened in the last years nor would I become more rational if I took a thinking-faster pill[3]. However, (human) epistemology is especially concerned with human rationality and it seems to be an essential feature of human rationality that humans have *finite* amounts of cognitive resources. For this reason, the study of rationality should acknowledge the fact that humans are finite reasoners[4].

That humans are finite reasoners is not often acknowledged in the literature because epistemologists are often concerned, not with reasoning, but only with the final product of reasoning:

[1] University of California, Davis. Philosophy Department. 1240 Social Science and Humanities, 1 Shields Avenue, Davis, CA 95616, USA. E-mail: dfdantas@ucdavis.edu

[2] Time is a *cognitive* resource in the sense of being related to how fast a reasoner can execute inferences and to how long a reasoner can reason (e.g. life span).

[3] Intelligence, on the other hand, seems to be tied to the possession of specific amounts of cognitive resources.

[4] Informally, a finite reasoner is a reasoner with cognitive limitations such as having only finite space in memory and being able to execute only finitely many inferences in a finite time interval (non-instantaneous inferences).

(sets of) beliefs. The most widespread formal model of a rational reasoner is Hintikka's model, based on modal epistemic logic (Hintikka, 1962). All reasoners described in Hintikka's model are logically omniscient in the sense of believing all the logical consequences of their beliefs (see Jago, 2006, for other notions of logical omniscience)[5]. Logically omniscient reasoners believe all logical tautologies (supposedly, tautologies are logical consequences of any set of beliefs). But logical omniscience is impossible for finite reasoners (e.g. humans), among other things, because there are infinitely many tautologies. Supposing that the adoption of each (explicit) belief demands some space in memory, finite reasoners cannot believe infinitely many tautologies because they have only finite space in memory. Supposing that the adoption of each (explicit) belief demands the execution of inferences, finite reasoners cannot believe infinitely many tautologies because they are able to execute only finitely many inferences in a finite time interval. This sort of inadequacy is known as the *problem of logical omniscience* (see Stalnaker, 1991; Duc, 1995; Jago, 2013; Artemov and Kuznets, 2014).

The most common strategy for dealing with the problem of logical omniscience is to interpret the models using a notion of beliefs different from explicit beliefs (e.g. implicit beliefs in Hintikka, 1962, p. 38). Consider a model of a reasoner that provides a clear definition for a notion of beliefs that may be used in dealing with the problem of logical omniscience (see Dantas, 2016, Ch. 1). In the model, a reasoner is composed of a language ($\mathcal{L}$), a knowledge base (KB), and a pattern of inference (π), where KB is a set of sentences in $\mathcal{L}$ that models the explicit beliefs of the reasoner and π: $2^{\mathcal{L}} \times \mathbb{Z}^+ \to 2^{\mathcal{L}}$ is a function for updating KB that models the pattern of inference of the reasoner. A fact about the pattern of inference of a reasoner is that the reasoner can execute different inferences from the same premises. In the model, this fact is expressed using a function π that has a numeric parameter (integer) in addition to the parameter for KB. In this context, π(KB, 1) models one inference from KB, π(KB, 2) models another inference from KB, etc. Then function π determines a reasoning sequence KB_0, KB_1,..., KB_i,..., where KB_0 is the initial set of explicit beliefs and $KB_{i+1} = \pi(KB_i, i+1)$. Supposing that the numeric parameter models an ordering of intention, a reasoning sequence models how the reasoner would reason if it could reason indefinitely. The set of stable beliefs, the beliefs that the reasoner would hold in the limit of a reasoning sequence, is $KB_\omega = \bigcup_i \bigcap_{j \geq i} KB_j$. The problem of logical omniscience could be avoided if Hintikka's model, for example, were interpreted in terms of stable beliefs. The model would describe maximum rationality, which a finite reasoner can only approach in the limit of a reasoning sequence[6].

This strategy has important consequences for epistemology. If a finite reasoner can only approach maximum rationality in the limit of a reasoning sequence, then the efficiency of patterns of inference is epistemically (and not only pragmatically) relevant. In the first section ("The argument"), I present an argument to this conclusion. In the second section ("Discussion"), I discuss the consequences of this conclusion. The main consequence is the vindication of the principle 'no rationality through brute-force'[7]. Rationality would be related to efficiency: the good use of (scarce) cognitive resources.

The argument

The efficiency of a pattern of inference may be measured in different ways, such as: (m1) the relative number of (explicit) beliefs at each stage of a reasoning sequence; (m2) the relative number of inferential steps executed until each stage of a reasoning sequence; (m3) the relative number, at each stage of a reasoning sequence, of (explicit) beliefs that will be retracted at later stages of the sequence (Kelly, 1988).

Measure m1 is related to memory, m2 is related to time, and m3 is related to both. The following argument is stated in terms of m2, but similar considerations might be done in terms of m1 or m3 (see Kelly, 1988, for an argument in terms of m3). An inference is a sequence of inferential steps, where an inferential step is the execution of an inference rule for a group of sentences. For example, concluding that q from $p \to q$ and p using modus ponens is an inferential step. Supposing that each inferential step demands time and that a finite reasoner has an upper bound for time (e.g. life span), executing relatively more direct inferences allows a reasoner to reach farther in a reasoning sequence because otherwise it would reach its upper bound at some earlier stage of the sequence[8].

Absolute efficiency (e.g. the absolute number of inferential steps) is usually said to be relative to implementations.

[5] In Hintikka's model, a reasoner is described as a set of possible worlds and an accessibility relation. The reasoner believes those sentences that are true in all accessible possible worlds. Possible worlds are maximally consistent (see Menzel, 2016). Then if a group of sentences are all true in all accessible possible worlds so are all their logical consequences. Then the reasoner believes all the logical consequences of its beliefs and is logically omniscient.

[6] I use 'maximum rationality' in the sense of the highest possible level of rationality for finite reasoners.

[7] In cryptography, a brute-force attack consists of an attacker trying many passwords with the hope of eventually guessing correctly. Accordingly, if there exists a procedure for checking guesses, a reasoner that is able to execute inferences instantaneously (unlimited time) may solve any problem (instantaneously) simply by generating and checking random guesses successively. This brute-force behavior does not model rationality for finite reasoners.

[8] For example, concluding that q from $p \to q$ and p using modus ponens (one inferential step) is a more direct inference than concluding q from a reduction to absurdity of the supposition that $\neg q$ (four inferential steps). The number of inferential steps is relative to the underlying logic, but the notion of more direct inferences is clear enough.

For example, the number of inferential steps is relative to the underlying logic. I will measure efficiency by using asymptotic analysis and, consequently, by comparing classes of computational complexity (those classes are usually said to be invariant over implementations – see the appendix about the computational complexity of algorithms).

This is the argument (in the remaining of this section, I will defend each of its premises):

- p1: If a finite reasoner with a polynomial pattern of inference had increasingly more cognitive resources, it would tend to reach infinitely farther in a reasoning sequence (in comparison to if it had an exponential pattern of inference).
- p2: If p1, then, under certain conditions, having a polynomial pattern of inference enables a finite reasoner to reach closer to maximum rationality (in comparison to having an exponential pattern of inference).
- p3: If p1 and p2, then the computational complexity of patterns of inference is relevant to epistemology.
- ∴ : The computational complexity of patterns of inference is relevant to epistemology.

P1

The claim in p1 is that if a finite reasoner with a polynomial pattern of inference had increasingly more cognitive resources, it would tend to reach infinitely farther in a reasoning sequence (in comparison to if it had an exponential pattern of inference)[9]. Let $r(i)$ be the 'resource function' of a reasoner, a function that measures the amount of some cognitive resource (e.g. time) necessary for the reasoner to reach the ith stage of a reasoning sequence ($r(i)\geq 0$ because the amount of cognitive resources necessary for reasoning is always nonnegative). In the following, I use $poln(x)$ as a predicate for denoting polynomial functions ('$f(x)$ is $poln(x)$' means that $f(x)$ is in the extension of $poln(x)$). The same holds for $exp(x)$ (exponential) and $log(x)$ (logarithmic, see the appendix about the computational complexities of algorithms for those classes)[10]. Consider theorem t1 (t1 and t2 are classical results in analysis, see Conrad, 2016, for related proofs):

$$\lim_{i\to\infty} \frac{exp(i)}{poln(i)} = \infty \qquad \text{(t1)}$$

Proof of t1: I will show that $\lim_{x\to\infty} \frac{a^x}{x^b} = \infty$, where $a>1$ and $b\geq 0$ are constants. If the proof holds for $\lceil b \rceil$, it holds for b. Then I will assume that b is an integer and use induction on b. Theorem t1 is true for b=0 because, in this case, $a^x \to \infty$ ($a>1$) and x^b is a constant. Suppose that t1 fails for some b and choose the minimal b for which t1 fails. Since $b\geq 1$, $\lim_{x\to\infty} x^b = \infty$. Since t1 holds for b-1, it follows that $\lim_{x\to\infty} \frac{a^x}{x^{b-1}} = \infty$. Let $f(x) = a^x$ and $g(x) = x^b$. L'Hopital's entails that $\lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{x\to\infty} \frac{f'(x)}{g'(x)} = \lim_{x\to\infty} \frac{ln(a)a^x}{bx^{b-1}} = \frac{ln(a)}{b} \lim_{x\to\infty} \frac{a^x}{x^{b-1}} = \frac{ln(a)}{b}\infty = \infty$. ■

Theorem t1 may be interpreted as stating that advancing in the reasoning sequence tends to demand infinitely more cognitive resources if the reasoner has an exponential pattern of inference (in comparison to an exponential pattern of inference). Now, consider a finite reasoner with a resource function $r(i)$ and a fixed upper bound $u\geq 0$ for some cognitive resource (e.g. time). Then the reasoner can reach the ith stage of a reasoning sequence iff $r(i)\leq u$. Consider t2, where $max(i|r(i)\leq u)$ denotes the maximum i such that $r(i)\leq u$ (the farthest stage in a reasoning sequence that the reasoner can reach):

$$\lim_{u\to\infty} \frac{max(i|poln(i)\leq u)}{max(i|exp(i)\leq u)} = \infty \qquad \text{(t2)}$$

Proof of t2: It is easy to see that $max(i|poln(i)\leq u)$ is at most $poln(u)$ and $max(i|exp(i)\leq u)$ is $log(u)$. Then I will show that $\lim_{u\to\infty} \frac{poln(u)}{log(u)} = \infty$. In other words, that $\lim_{x\to\infty} \frac{x^a}{log_c(x)^b} = \infty$, where $a\geq 0$, $b\geq 0$, $c>0$ are constants. Since the ratio $\frac{log_c(x)^b}{ln(x)^b}$ is a constant $log_c(e)^b > 0$, I will restrict myself to the natural logarithm ln. By the same reasoning used in the proof of t1, I will assume that b is an integer and use induction on b. By the same reasoning used in the proof of t1, t2 is true for b=0. Suppose that t2 fails for some b and choose the minimal b for which it fails. Since $b\geq 1$, $\lim_{x\to\infty} ln(x)^b = \infty$. Since t2 holds for b-1, $\lim_{x\to\infty} \frac{x^a}{ln(x)^{b-1}} = \infty$. Let $f(x)=x^a$ and $g(x)=ln(x)^b$. L'Hopital's entails that $\lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{x\to\infty} \frac{f'(x)}{g'(x)} = \lim_{x\to\infty} \frac{ax^{a-1}}{bx^{b-1}1/x} = \frac{a}{b} \lim_{x\to\infty} \frac{x^a}{ln(x)^{b-1}} = \frac{a}{b}\infty = \infty$.

Theorem t2 is more difficult to interpret. For any finite reasoner, we may conceive a series of otherwise identical (hence, similar) finite reasoners with increasingly larger (but still finite) upper bounds for some cognitive resource (e.g. time). Then t2 may be interpreted as stating that, as we consider two series of similar finite reasoners with increasingly larger upper bounds us, those with polynomial patterns of inference tend to reach infinitely farther in the reasoning sequence (in comparison to those with exponential patterns of inference). But, if t2 may be interpreted in terms of series of (merely possible) similar finite reasoners, it may also be interpreted in terms of series of counterfactual versions of a finite reasoner. Then theorem t2 may be interpreted as stating that if a finite reasoner with a polynomial pattern of inference had increasingly more cognitive resources, it would tend to reach infinitely farther in a reasoning sequence (in comparison to if it had an exponential pattern of inference), which is p1.

[9] Solving 2-satisfiability using truth-tables is an example of exponential pattern of inference. Using Krom's algorithm for the same end is an example of polynomial pattern of inference (see Krom, 1967). The equations in this section may be interpreted in terms of the task 'solve 2-satisfiability for each suitable sentence in the language'.

[10] Then 'polynomial pattern of inference' denotes a pattern of inference with resource function of the form *poln*(*i*) and 'exponential pattern of inference' denotes a pattern of inference with resource function of the form *exp*(*i*).

P2

The claim in p2 is that if p1, then, under certain conditions, having a polynomial pattern of inference enables a finite reasoner to reach closer to maximum rationality (in comparison to having an exponential pattern of inference). Maximum rationality can (only) be approached in the limit of the reasoning sequence. Then if a polynomial reasoner tending to reach further in the sequence (in comparison to an exponential reasoner) entails that it tends to reach closer to the limit of the sequence, then it may be said that p1 entails that, under certain conditions, a polynomial reasoner tends to reach closer to maximum rationality (in comparison to an exponential reasoner)[11]. Reaching closer to the limit of an infinite sequence does *not* make sense for any finite difference of positions in the sequence, but p1 states that the difference of positions between a polynomial and an exponential reasoner tends to infinity. In this case, I think that it may be said that p1 entails that, under certain conditions, a polynomial reasoner reaches closer to maximum rationality (in comparison to an exponential reasoner), where the conditions in question are 'at the limit of a reasoning sequence if it had increasingly more cognitive resources.'

Premise p1 also suggests that having a polynomial pattern of inference is *the* feature that enables a polynomial reasoner to reach farther in the sequence (closer to maximum rationality) because having a polynomial pattern of inference is the only difference between the polynomial reasoner and its exponential counterpart. In this case, it may be said that p1 entails that, under certain conditions, having a polynomial pattern of inference is what enables a polynomial reasoner to reach closer to maximum rationality (in comparison to an exponential reasoner). Then it may be said that if p1, then, under certain conditions, having a polynomial pattern of inference enables a finite reasoner to reach closer to maximum rationality (in comparison to having an exponential pattern of inference), which is p2.

P3

The claim in p3 is that if p1 and p2, then the computational complexity of patterns of inference is relevant to epistemology. If p1 and p2, then (by modus ponens) it follows that, under certain conditions, having a polynomial pattern of inference enables a finite reasoner to reach closer to maximum rationality (in comparison to having an exponential pattern of inference). I regard as a general principle of (meta-)epistemology that if, under certain conditions, some feature enables a reasoner to reach closer to maximum rationality and those conditions are relevant to epistemology, then whether a reasoner possesses that feature is relevant to epistemology.

Performing more reasoning (and having the necessary cognitive resources for doing so) usually enables a reasoner to be in a better epistemic position. Then the conditions 'at the limit of a reasoning sequence if it had increasingly more cognitive resources' are relevant to epistemology. Then it follows from the general principle that whether a pattern of inference is polynomial or exponential (i.e. its computational complexity) is relevant to epistemology. If p1 and p2, then the computational complexity of patterns of inference is relevant to epistemology, what is p3.

Discussion

If p1, p2, and p3 are all true, then (by two modus ponens) it follows that (∴) the computational complexity of patterns of inference is relevant to epistemology. The question now is how to interpret this conclusion. What would be the role of computational complexity in epistemology? I think that the preceding discussion suggests an epistemic norm of the form 'a rational reasoner should have a polynomial pattern of inference (if possible)'. The clause 'if possible' is in place because (most probably, if $PN \neq P$) it is not possible for a finite reasoner to deal with some problems (e.g. NP-complete problems) using a polynomial pattern of inference (see the appendix about the computational complexity of problems).

In the literature on computer science, exponential patterns of inference are often correlated with brute-force search whereas polynomial patterns of inference are correlated with deep understanding:

> *The motivation for accepting this requirement is that exponential algorithms typically arise when we solve problems by exhaustively searching through a space of solutions, what is often called a brute-force search. Sometimes brute-force search may be avoided through a deeper understanding of a problem, which may reveal polynomial algorithms of greater utility (Sipser, 2012, p. 285).*

But, if this correlation is correct, to require rational reasoners to have polynomial patterns of inference (if possible) is to require rational reasoners to approach maximum rationality through deep understanding and not through brute-force (if possible). This seems to be a vindication of the principle 'no rationality through brute-force.'

Appendix

For more information about theory of computational complexity, see Sipser (2012, p. 273).

[11] 'Polynomial reasoner' denotes a reasoner with a polynomial pattern of inference and 'exponential reasoner' denotes a reasoner with an exponential pattern of inference.

Complexity of algorithms

The analysis of the complexity of algorithms is a part of computational complexity theory. The complexity of an algorithm is often understood as the rate in which the running time (time complexity) or the memory requirements (space complexity) of the algorithm grows in relation to the size of the input. The absolute complexity of an algorithm usually requires assumptions about implementation. In order to abstract from these assumptions, an asymptotic analysis is used.

In an asymptotic analysis, the complexity of the algorithm is determined for arbitrarily large inputs. The asymptotic analysis of an algorithm is often done under some extra abstractions. The first step is to abstract over the input of the algorithm, using some parameter to characterize the size of the input. The second step is to use some parameters to characterize the running time or memory requirements of the algorithm. In addition, the complexity of an algorithm is usually measured in relation to its worst-case scenario (the most demanding case for the algorithm).

Nevertheless, the exact complexity of an algorithm is often a complex expression. In most cases, the big-O notation is used for simplification (e.g. $f(n) = O(g(n))$, where n is the size of input). In big-O notation, only the highest order term of the expression of the complexity of the algorithm is considered and both the coefficient of that term and any lower order terms are disregarded. The idea is that, when the input grows arbitrarily, the highest order term dominates the other terms. For example, the function $f(n) = 7n^3 + 5n^2 + 10n + 34$ has four terms and the highest order term is $7n^3$. Disregarding the coefficient 7, we say that $f(n)=O(n^3)$.

Definition 1 (Big-O notation ($O(g(n))$)). Let f and g be functions $f, g\colon \mathbb{N}\rightarrow\mathbb{R}^+$. Then $f(n)=O(g(n))$ iff there exist positive integers c and n_0 such that for every integer $n\geq n_0$, $f(n)\leq c\cdot g(n)$ (Sipser, 2012, p. 227).

In this framework, algorithms are often classified under two categories: polynomial and exponential complexity. A polynomial algorithm is an algorithm whose complexity may be expressed using a function of the kind $O(n^c)$, where c is a constant. A species of polynomial algorithm are the logarithmic algorithms, which may be expressed using a function of the kind $O(log_c(n))$. An exponential algorithm is an algorithm whose complexity must be expressed using a function of the kind $O((2^n)^c)$, where c is a constant greater than 0. While the actual complexity of an algorithm depends on low-level encoding details, where an algorithm falls on the polynomial/exponential dichotomy is independent of almost all such choices (for reasonable models of computation).

Complexity of problems

In computational complexity theory, a complexity class is a class of problems of related resource-based computational complexity. The most important complexity classes are P, NP, and NP-complete. The class P (polynomial time) is the class of decision problems[12] that are solvable by a deterministic Turing machine in polynomial time. Meanwhile, NP (nondeterministic polynomial-time) is the class of decision problems for which a solution may be checked by a deterministic Turing machine in polynomial time, even if the solution cannot be found in polynomial time[13].

Since all problems solvable in polynomial time may be checked in polynomial time, P ⊆ NP. Whether P = NP is an open question[14]. Of special interest for the solution of this question is the class NP-complete. NP-complete is the class of decision problems such that all decision problems in NP are reducible to these problems in polynomial time[15]. In this context, if a problem in NP-complete is solvable in polynomial time, then all problems in NP are solvable in polynomial time and P=NP. An example of NP-complete problem is the satisfiability problem.

Acknowledgements

The author thanks Rodrigo de Lima for the assistance with the mathematical proofs.

References

ARTEMOV, S.; KUZNETS, R. 2014. Logical Omniscience as Infeasibility. *Annals of Pure and Applied Logic*, **165**(1):6-25. https://doi.org/10.1016/j.apal.2013.07.003

CONRAD, K. 2016. Orders of Growth. Available at: http://www.math.uconn.edu/~kconrad/blurbs/analysis/growth.pdf. Accessed on: January 11, 2018.

DANTAS, D. 2016. *Almost Ideal: Computational Epistemology and the Limits of Rationality for Finite Reasoners*. Davis, California. PhD dissertation. University of California, 160 p.

DUC, H. 1995. Logical Omniscience vs. Logical Ignorance on a Dilemma of Epistemic Logic. *In: Proceedings of the 7th Portuguese Conference on Artificial Intelligence*. London, UK, p. 237-248. https://doi.org/10.1007/3-540-60428-6_20

FOLEY, R. 1987. *The Theory of Epistemic Rationality*. Cambridge, Harvard University Press, 335 p. https://doi.org/10.4159/harvard.9780674334236

HINTIKKA, J. 1962. *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Ithaca, Cornell University Press, 148 p.

[12] A decision problem is a problem with a yes-or-no answer.

[13] To solve a decision problem is to return the correct answers for all instances of the problem; to check the solution of a decision problem is to solve for all 'yes' instances of the problem.

[14] Nevertheless, most theorists work on the assumption that P ≠ NP.

[15] To reduce a problem to another is to create an algorithm which maps the solution for the first problem to the solution for the second.

JAGO, M. 2006. Hintikka and Cresswell on Logical Omniscience. *Logic and Logical Philosophy*, **15**(3):325-354.

JAGO, M. 2013. The Problem of Rational Knowledge. *Erkenntnis*, **79**(S6):1-18.

KELLY, K. 1988. Artificial Intelligence and Effective Epistemology. *In:* J. FETZER (ed.), *Aspects of Artificial Intelligence*. Amsterdam, Springer, p. 309-322. https://doi.org/10.1007/978-94-009-2699-8_11

KROM, M. 1967. The Decision Problem for a Class of First-order Formulas in which All Disjunctions are Binary. *Mathematical Logic Quarterly*, **13**(1-2):15-20. https://doi.org/10.1002/malq.19670130104

MENZEL, C. 2016. Possible Worlds. *In:* E. ZALTA (ed.), *The Stanford Encyclopedia of Philosophy*. Stanford, Metaphysics Research Lab.

SIPSER, M. 2012. *Introduction to the Theory of Computation.* Boston, Cengage Learning, 458 p.

STALNAKER, R. 1991. The Problem of Logical Omniscience, I. *Synthese*, **89**(3):425-440. https://doi.org/10.1007/BF00413506555

Submitted on November 20, 2017

Accepted on February 02, 2018