

Argument-Forms Which Turn Invalid Over Infinite Domains: Physicalism as Supertask?

Catherine Legg

Argument-forms exist which are valid over finite but not infinite domains. Despite understanding of this by formal logicians, philosophers can be observed treating as valid arguments which are in fact invalid over infinite domains. In support of this claim I will first present an argument against the classical pragmatist theory of truth by Mark Johnston. Then, more ambitiously, I will suggest the fallacy lurks in certain arguments for physicalism taken for granted by many philosophers today.

1. Introduction

Argument-forms exist which are valid over finite but not infinite domains. One such was discovered by De Morgan, who dubbed it the “syllogism of transposed quantity.” A particularly vivid example articulated by Charles Peirce (1931, §3.288) is the following:

Every Texan kills a Texan,
No Texan is killed by more than one Texan,
Hence, every Texan is killed by a Texan.

Common-sense suggests interpreting this argument over a domain containing (legendary state pride notwithstanding) a finite, albeit large, number of Texans. If so, the argument is valid in the sense that there is no possible world where the premises (as interpreted) are true and the conclusion false. But where the Texans are even countably infinite, there is such a possible world. The argument is an instance of the following argument-form:

$$\begin{aligned} & \forall x \exists y K(x,y) \\ & \forall x \forall y \forall z (K(y, x) \wedge K(z, x)) \rightarrow (y = z) \\ \therefore & \forall x \exists y K(y,x) \end{aligned}$$

That this argument-form is invalid is shown by the following instantiation:

Every number has a successor.

For any three numbers, if the second and the third are successors of the first, then the second is equal to the third.

Therefore, every number is the successor of some number.

When interpreted over a domain consisting of the positive natural numbers, the premises of this argument are true and the conclusion false. However if the domain is restricted to any finite subset of those numbers, this does not hold.

Many logicians are uncomfortable with talk of argument forms “turning” invalid over infinite domains – they would rather say that infinite domains show that such argument forms are invalid *simpliciter*. But I believe it is worth highlighting the fact (however one refers to it) that certain argument-forms can be relied on to deliver true conclusions from true premises if one only interprets them over finite domains, while at the same time if infinite domains are considered they cannot. I wish to argue that an understanding of this fact has still not fully percolated from formal logic into broader philosophical reasoning, insofar as philosophers can be observed treating as valid arguments which are in fact invalid over infinite domains.

In support of this claim, I will first present an argument against the classical pragmatist theory of truth by Mark Johnston, where a fallacy of this form is arguably quite clear-cut. Then, more ambitiously, I will suggest the fallacy lurks in certain arguments for physicalism taken for granted by many philosophers today.

2. The Classical Pragmatist Definition of Truth and Johnston’s “No Ideal Theory” Objection to it

The American pragmatist tradition introduced a new and distinctive internalist position to the landscape of philosophers’ attempts to define truth. This position differs from major (non-deflationary) rivals by not characterising truth in terms of a relationship between *truth-bearers* (such as propositions, sentences or beliefs) and something else. (This “something else” is thought in the case of correspondence theories of truth to be worldly *truth-makers*, and in the case of coherence theories of truth to be *other truth-bearers*). Rather, truth is defined as what is believed by what might be termed *truth-believers* (that is, inquirers, *albeit* often idealised in some way). More precisely, as Peirce originally put it (1931, §5.407), “The opinion which is fated to be ultimately agreed to by all who investigate, is what we mean by the truth.” Much more recently, but in the same lineage, Hilary Putnam (1981) defined truth as an “idealisation of rational acceptability.”

This view attracted the critical scrutiny of Mark Johnston, who attacked it as follows. Johnston (1993a, 88–89)¹ takes for granted that this conception of truth is committed to the existence of an ideal theory, “T.” For such a theory, surely, is what is “fated to be ultimately agreed to by all who investigate,” in Peirce’s terms, and is what would be ideally rationally acceptable, in Putnam’s. Johnston then asks what such an ideal theory might consist in. He suggests that it would be a set of sentences from which every true sentence may be deduced. He then objects that, “[a]s a matter of simple metalogic it could not be true that something is true if and only if it follows from the ideal theory T.”

Johnston’s argument runs as follows. If the language of T is rich enough to express arithmetic, then it will be rich enough for a Goedel argument to be constructed within it. From this it follows that there will be truths expressible in T which are not provable within T, such as the famous “Goedel sentence” which, very roughly, says truly of itself that it is not provable. The best that can be hoped for, he writes, is an ideal *hierarchy* of theories, each closer to the truth than the last, proving to be true statements which were only expressible in the theory directly below it. Thus, instead of the classical pragmatist account of truth, Johnston favours a “Progressive Protagoreanism,” which is protagorean (or internalist) merely regarding which truths one should pay attention to – so that one might avoid wasting epistemic energy on “metaphysics in the pejorative sense” – but not regarding truth itself, which he treats in classical correspondence terms.

3. Why Johnston’s Objection Fails

This is a clever objection, to which it might seem *prima facie* that the classical pragmatist theory of truth must fall. It seems that “the opinion which is fated to be ultimately agreed to by all who investigate” (sometimes referred to as the “final opinion”) cannot contain the Goedel sentence – but neither may it omit the Goedel sentence, surely, if the sentence is true. To conclude that this is a knock-down argument would, however, be premature.

First and most importantly, Peirce never said that every true statement had to be *provable* within the final opinion. All that is required is that the statement be *included* in the final opinion. It may be protested that this is to admit that the final opinion is not “complete” (in the logical as well as the ordinary language sense), and this seems to be a strange and troubling failing for such an idealised notion.

However, just because *any* true statement (such as the Goedel sentence and its many higher-order equivalents) would be believed if inquiry proceeded indefinitely, it does not follow that there has to be *some complete* theory in which they are all believed together. Johnston has failed to recognise that the pragmatist may commit to *any* question receiving an answer at *some* point in time, while not committing to there being any point in time at which *every* question has received an answer: just as the fact that every person has a mother

who loves them does not entail that there is some particular mother who loves every person.

At this point Johnston's construal of the pragmatist theory of truth might seem to involve a simple quantifier scope confusion, and thus to constitute an invalid argument even over finite domains, as follows (letting Qx mean that x is a question and $A(x,t)$ mean that x is answered at t):

$$\begin{aligned} & \forall x \exists t (Qx \rightarrow A(x,t)) \\ \therefore & \exists t \forall x (Qx \rightarrow A(x,t)) \end{aligned}$$

However, a supporter of Johnston might respond that the nature of inquiry is, when all goes well, *cumulative* (or "progressive" as Kuhn put it) so that questions which are answered at any given time tend overall to stay answered.

Indeed, the classical pragmatist's own talk of the 'final opinion' seems to demonstrate a commitment to inquiry being at least sufficiently cumulative for such an opinion to form (however haltingly).² This can be argued to add a new premise to the argument, so that the inference from *any* question receiving an answer at some point in time to *every* question having received an answer at some point in time should nevertheless be valid.

Consider the case where every mother loves not just the child she has just brought into the world, but (in a burst of new maternal affection) all other children brought into the world previously. In such a case it seems that we may infer: if every child has a mother who loves them, then there is some mother who loves every child – the last mother to give birth. Formalised (letting Qx mean that x is a mother, $A(x,y)$ mean that x loves y maternally, and $B(x,y)$ mean that x was born before or simultaneously with y), the argument-form may now be represented thus:

$$\begin{aligned} & \forall y \exists x (Qx \rightarrow A(x,y)) \\ & \forall x \forall y_1 \forall y_2 (B(y_1, y_2) \rightarrow (A(x, y_2) \rightarrow A(x, y_1))) \\ \therefore & \exists x \forall y (Qx \rightarrow A(x,y)) \end{aligned}$$

In other words:

Everyone is loved by some mother.

For any two individuals, if the first is born before or simultaneously with the second, then any mother who loves the second also loves the first.

Therefore there is a mother who loves everyone.

Johnston's argument is isomorphic except for a reversal of the inequality:

Every question receives an answer at some time.

For any two times, if the first is before or simultaneous with the second, then any question which is answered at the first time is answered at the second.

Therefore there is a time at which every question is answered.

These arguments are valid *only over a finite domain* (of mothers and children, or questions and times at which they are answered). The arguments' validity depends upon the existence of a "last" member of the domain. As Peirce put it:

We must look forward to the explanation, not of all things, but of any given thing whatever. There is no contradiction here, any more than there is in ... saying that any future time will sometime be passed, though there never will be a time when all time is past. (1931, §1.405)

Given that the set of questions to be answered, and future times at which they might be answered, is greater than finite, then, Johnston's argument fails. And surely this is the case.³

4. Further Consequences for Physicalism

Physicalism is widely believed in contemporary philosophy. It is often put forward as the natural philosophical heir of the scientific naturalism which provided such spectacular explanatory successes in the physical sciences from the seventeenth century onwards. Such successes often lead physicalists to argue for their position (*albeit* often implicitly) by way of an optimistic meta-induction along lines such as the following. Given that physicists were so successful explaining planetary motion, the behavior of gases, the paths of projectiles, and so on, therefore sooner or later they will achieve similar explanatory successes with respect to all other phenomena.

As an example of this optimistic meta-induction, physicalism is particularly popular in current philosophy of mind, where it is often defined in terms such as the following:

All mental properties, states, and events can be wholly explained in terms of physical properties, states, and events.

What is not often drawn attention to is that this definition of "philosophy of mind physicalism" is ambiguous between two interpretations:

(P1) Any mental property can be given a physical explanation in some physical theory.

- (P2) There is some total physical theory which explains every mental property.

Philosophical physicalists tend naively to assume the stronger (P2). Consider for instance, the following remark in Jackson and Braddon-Mitchell (1996, 13), “The physical story that explains each and every physical happening may very well have irreducibly probabilistic elements.” Jack Smart goes so far as to suggest that that physical story is current “physics of ordinary matter” (1978, 339–340). In many other cases (P2) is not explicitly stated but is implicit in a definition of physicalism as the denial of *emergentism* (see for instance, Pettit 2003 and Stoljar 2001). For it is only against the background of a total physical theory that phenomena can *be* emergent, rather than merely explained by a different physical theory to the theory one is presently considering.

However, induction over the history of explanatory successes in science arguably only establishes (P1) at best. One might attempt to argue that (P2) follows from (P1). An argument such as the following (letting Mx mean that x is a mental property and $E(x,y)$ mean that y provides a physical explanation for x):

$$\begin{aligned} &\forall x \exists y (Mx \rightarrow E(x,y)) \\ \therefore &\exists y \forall x (Mx \rightarrow E(x,y)) \end{aligned}$$

would of course constitute another straight-out quantifier-scope fallacy. Once again, therefore, one might seek to add a further premise to make the argument valid.

A consideration which often occurs to physicalists who make use of Ramsey sentences is the following. Physical explanations can simply be concatenated *via* logical conjunction. If a physical theory ($T1$) explains one mental property ($m1$), and a different physical theory ($T2$) explains another mental property ($m2$), then ($T1 \wedge T2$) will surely be a “theory” which explains both $m1$ and $m2$, and so on for every theory and every mental property until (P2) is satisfied. One might formalize the argument like this:

$$\begin{aligned} &\forall x \exists y (Mx \rightarrow E(x,y)) \\ &\forall x_1 \forall x_2 \forall y_1 \forall y_2 [((Mx_1 \wedge E(x_1, y_1)) \wedge (Mx_2 \wedge E(x_2, y_2))) \rightarrow \exists y_3 (E((x_1 \\ &\quad \wedge x_2), y_3) \wedge (y_3 = (y_1 \wedge y_2)))] \\ \therefore &\exists y \forall x (Mx \rightarrow E(x,y)) \end{aligned}$$

This suggestion flies in the face of much careful work in the philosophy of science regarding scientific theory-holism following, for instance, Quine (1953, 1960), arguing that scientific terms such as ‘force’ and ‘electron’ do not have a theory-independent meaning that would be interpretable across the conjunction of two very different theories jammed together. As an example, try

to imagine a vast conjunction composed of all the claims made about gravity by Newtonian mechanics and Einstein's general relativity. What would such a set of statements mean? What would it "say about gravity"?

If such holism is denied, and theories are considered as nothing but sets of individually-interpretable statements, a further issue arises. What if certain statements in $T1$ and certain statements in $T2$ are logically inconsistent? Obviously the conjunction of the two theories would then give rise to a contradiction, which would arguably not constitute a very good explanation of anything. All explanatory virtue might not be lost if one dipped into some form of relevance logic allowing the isolation of contradictory theory-components from one another, and adherents of paraconsistency would most probably see precisely this kind of case as grist for their mill.

Alternatively, defenders of theory-concatenation might wish to argue that, given the principle of excluded middle, one cannot have a logical inconsistency in a group of statements without having at least one true and one false statement. (In other words, for a contradiction one needs something of logical form p and something of logical form $\sim p$). However, all successful physical explanations are *true*. (For surely it is a necessary condition for successful explanations that they be true?) Thus two statements of differing truth-value will not be found anywhere in $T1$ and $T2$, and thus there will be no contradiction. This reply seems to involve quite a leap of faith however, particularly when one notes that one must claim not only that all successful physical explanations are true but that all *parts* (i.e. component propositions) of all successful explanations are true. How many real-life physical theories could live up to that?⁴

Even leaving that thorny issue aside, however, once again, the argument-form above is not valid. Consider the following instantiation of it:

For every natural number there is a number larger than it.

For every two pairs of natural numbers ($a1$ and $a2$) and ($b1$ and $b2$) for which it is the case that the second number is larger than the first, there exists a number which is larger than ($a1 + b1$), namely ($a2 + b2$).

Therefore, there is a natural number which is larger than all natural numbers.

To sum up then: there is no largest number (due to the infinity of the domain over which the successor relation is defined), there is no last moment in time (due to the infinity of the domain over which the 'later than' relation is defined), the number of different phenomena which might require physical explanation is arguably infinite⁵ and thus the potential emendations of physical theory are infinite, so why should we think there is *a* theory which explains everything?

The physicalist may wish to argue at this point that talk of “the” physical theory which would explain everything is a benign *idealisation*, as harmless in epistemology as the concept of the frictionless plane is in physics. Physical law ensures that we never encounter such a thing as a plane with utterly zero friction. Nevertheless, to assume that such a thing exists greatly simplifies certain physical equations. What harm does it do to assume that all explanations come together in a single theory?

Well, first of all, we have already noted that Johnston was led into a fallacious argument by it. More generally, however, an intriguing literature is emerging concerning so-called “super-tasks,” pointing out that once we idealize a process to infinity, the properties possessed by the limit of the process are not always what we would expect given its finite instances. For instance, Thomson’s lamp, which switches from ‘on’ to ‘off’ and *vice versa* at ever-halving intervals of time, is neither ‘on’ nor ‘off’ at the end of its infinite series of switchings, in apparent violation of physical law.

A recent short exchange in *Analysis* (Friedman and Black 2002) provides an even more nicely counterintuitive example. Friedman describes a supertask involving an infinite number of numerically labeled balls, as follows:

- Step 1: Place balls ‘1’ and ‘2’ on a table. Discard the ball numbered ‘1’.
- Step 2: Place balls ‘3’ and ‘4’ on the table. Next, interchange numbers ‘2’ and ‘3’... Finally, discard the ball that is now numbered ‘2’.
- Step 3: Place balls ‘5’ and ‘6’ on the table. Next interchange numbers ‘3’ and ‘5’. Finally, discard the ball that is now numbered ‘3’...
(p. 344)

If this process is idealised to infinity, the number of balls on the table tends to infinity while the number of labels on the balls tends to zero! In response to this apparent paradox, Black cheerfully notes (p. 346), “you can only think this is a problem if you think the equinumerosity of [the balls] and [the labels] has to be carried over to the limit sets, and it doesn’t.”

A number of standard challenges to physicalism in its many guises and versions exist. For instance, it has been complained that physicalism is in principle unable to explain qualia (Chalmers 1996), and intentionality (Kripke 1982), and that physicalists illicitly convert what should be a methodological into a metaphysical claim. However as far as I know this particular criticism has not been raised.⁶

One might complain at this point that there is an odd structure to the dialectic of this article. It would appear that the very same considerations have been used to defeat both Johnston’s argument that can be no final physical theory, and the physicalist’s argument that there is. Can one really have one’s cake and eat it too in this way?

However, though the two views are indeed opposites:

$$\sim \exists t \forall x (Qx \rightarrow A(x,t)) \text{ (Johnston)}$$

$$\exists t \forall x (Qx \rightarrow A(x,t)) \text{ (the physicalist)}$$

neither one of these two views is the contradiction of Peirce's view:

$$\forall x \exists t (Qx \rightarrow A(x,t))$$

I conclude that despite a certain popular "fast-and-loose" image which pragmatism seems to have picked up, a proper understanding of the classical pragmatist theory of truth requires an exact appreciation of modern logic.

NOTES

1. See also an almost identical discussion in his 1993b, p. 318.
2. Of course the possibility must be faced of inquiry not just *slowing down*, but *backsliding*, such that inquirers stop believing certain true things which they once believed, and start believing certain false things which they did not believe before. Taking this possibility duly into consideration would add more bells and whistles to the formalization which follows, but not alter the objection I am about to make.
3. Further assumptions are of course implicit in the classical pragmatist definition of truth as the end of inquiry, such as that rational beings will continue to exist and to inquire, which may be taken issue with. However this would be a further discussion.
4. Consider Newtonian mechanics once again. According to this theory there is no upper bound on the speed of a moving mass, and an object's acceleration cannot alter its size, but neither of these claims is true according to best scientific theory today.
5. As an argument that the universe's phenomena are infinite not merely in quantity but also in kind consider the infinite number of different kinds of thoughts which arguably might be entertained by the human mind. If however this argument is felt to be too swift, and that the possibility should be left open for some kind of finitist discrete theory of space-time, it may be argued that at least one's philosophical naturalism should not beg such an empirical question.
6. Crane and Mellor (1990) do discuss a worry that the key claim of physicalism is crucially ambiguous – but they identify the ambiguity as holding between the claim that an explanation can be given for any phenomenon by *current physical theory* and by *future idealised physical theory*. Both of these options seem to be conceived as unitary theories. Nevertheless there are probably interesting links to be drawn between the two discussions.

REFERENCES

- Black, Robert. 2002. "Solution of a Small Infinite Puzzle," *Analysis* 62: 345–346.

- Chalmers, David. 1996. *The Conscious Mind*. New York: Oxford University Press.
- Crane, Tim, and Mellor, D. H. 1990. "There is No Question of Physicalism," *Mind* 99: 185–205.
- Friedman, K.S. 2002. "A Small Infinite Puzzle," *Analysis* 62: 344–345.
- Hookway, Christopher. 1985. *Peirce*. London and New York: Routledge.
- Jackson, Frank, and Braddon-Mitchell, David. 1996. *The Philosophy of Mind and Cognition*. Oxford: Blackwell.
- Jackson, Frank. 1997. *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford: Oxford University Press.
- Johnston, Mark. 1993a. "Objectivity Refigured: Pragmatism Without Verificationism," in *Reality, Representation and Projection*, ed. J. Haldane and C. Wright (Oxford: Oxford University Press), pp. 85–130.
- Johnston, Mark. 1993b. "Verificationism as Philosophical Narcissism," *Philosophical Perspectives* 7: 307–330.
- Kripke, Saul. 1982. *Wittgenstein on Rules and Private Language: An Elementary Exposition*. Oxford: Blackwell.
- Laraudogoitia, Jon Perez. 1996. "A Beautiful Supertask," *Mind* 105: 81–83.
- Lombardi, Olimpia, and Labarca, Martín. 2005. "The Ontological Autonomy of the Chemical World," *Foundations of Chemistry* 7: 125–148.
- Misak, Cheryl. 1991. *Truth and the End of Inquiry*. Oxford: Clarendon Press.
- Moore, Adrian. 1991. *The Infinite*. London and New York: Routledge.
- Peirce, Charles Sanders. 1931. *Collected Papers of Charles Sanders Peirce, vol. 1: Principles of Philosophy*, ed. Charles Hartshorne and Paul Weiss. Cambridge, MA: Harvard University Press.
- Peirce, Charles Sanders. 1976. *New Elements of Mathematics*, vols. 3 and 4, ed. Carolyn Eisele. The Hague: Mouton.
- Pettit, Philip. 1993. "A Definition of Physicalism," *Analysis* 53: 213–223.
- Putnam, Hilary. 1981. *Reason, Truth and History*. Cambridge, UK: Cambridge University Press.
- Quine, Willard Van Ormand. 1953. "Two Dogmas of Empiricism," repr. in *From A Logical Point of View* (Cambridge, MA: Harvard University Press), pp. 20–46.

Quine, Willard Van Ormand. 1960. *Word and Object*. Cambridge, MA: Harvard University Press.

Smart, J. J. C. 1978. "The Content of Physicalism," *Philosophical Quarterly* 28: 339–341.

Stoljar, Daniel. 2001. "Physicalism," *Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu/entries/physicalism/>. Accessed 26 March 2008.

Thomson, James F. 1954. "Tasks and Super-Tasks," *Analysis* 15: 1–15.

Catherine Legg
Senior Lecturer
Department of Philosophy and Religious Studies
University of Waikato
Private Bag 3105
Hamilton
New Zealand