in dispensing with commensurability, perhaps we are merely enabling philosophical analysis to catch up with existing (and correct) human wisdom. All this time, we've all been comparing apples with oranges already.

London School of Economics
London WC2A 2AE, UK
r.d.northcott-alumni@lse.ac.uk

References

Good, I. J. 1961. A causal calculus, parts I and II. *British Journal for the Philosophy of Science* 11: 305–18 and 12: 43–51.

Lewontin, R. 1974. Analysis of variance and analysis of causes. *American Journal of Human Genetics* 26: 400–11.

Matthen, M. and A. Ariew. 2002. Two ways of thinking about fitness and natural selection. *Journal of Philosophy* 99: 55–83.

Miller, R. 1987. *Fact and Method*. Princeton, NJ: Princeton University Press.

Northcott, R. 2004. *Approximate Truth and Causal Strength in Science*. University of London PhD dissertation.

Pearl, J. 2000. *Causality*. New York: Cambridge University Press.

Sober, E. 1988. Apportioning causal responsibility. *Journal of Philosophy* 85: 303–18.

Sober, E., E. O. Wright and A. Levine. 1992. Causal asymmetries. In *Reconstructing Marxism*, 129–75. London: New York: Verso.

Spirtes, P., C. Glymour and R. Scheines. 2000. *Causation, Prediction, and Search* (2nd ed.). Cambridge, MA: MIT Press.

Some more curious inferences

Jeffrey Ketland

1. Boolos's curious inference

In his 1987, George Boolos presented the following inference, I:

(1) $\forall n\, f(n, 1) = s1$
(2) $\forall x\, f(1, sx) = ssf(1, x)$
(3) $\forall n \forall x\, f(sn, sx) = f(n, f(sn, x))$
(4) $D(1)$
(5) $\forall x(D(x) \rightarrow D(sx))$
Therefore,
(6) $D(f(ssss1, ssss1))$

Here, '1' is a constant, '*s*' is a 1-place function symbol, '*f*' is a 2-place function symbol and '*D*' is a 1-place predicate symbol. Interpreted over the positive integers, the function *f* is a rapidly growing Ackermann-like function, defined by a double recursion.[1] The conclusion of I is that $f(5, 5)$ is in the extension of '*D*'. Boolos notes that the integer $f(4, 4)$ is astronomically huge, an exponential stack of 64,536 2's, larger than any integer that might naturally appear in science.

Boolos points out that there is a reasonably short derivation of the conclusion (6) from premises (1)–(5) in a standard deductive system for *second-order logic*. In particular, the short derivation uses the comprehension scheme, 'There is an X such that, for any *n*, X(*n*) iff φ(*n*)'. In more detail, the derivation proceeds by a construction involving the *closure* of whatever '1' denotes under the function denoted by '*s*' (intuitively, these are the positive integers, 1, 2, 3, etc.), and proves that this set is also closed under the function denoted by '*f*'. The notion of closure is not available in first-order logic and must be understood as either set-theoretical or second-order.[2] The second-order proof, with no abbreviations, fills less than one page.[3]

Boolos also points out that although *there exists* ('platonistically', as it were) a first-order derivation Γ of (6) from (1)–(5), the length of Γ is astronomically huge. The size of Γ depends upon the size of the numeral '*ss ... s*1', denoting the number $f(5, 5)$. This numeral is huge. It follows that Γ must be astronomically large: Boolos estimates that the number of symbols in any derivation of (6) from (1)–(5) in a Mates-style deductive system will be at least $f(4, 4)$.

Boolos comments that,

> [T]he fact that we so readily recognize the validity of I would seem to provide as strong a proof as could be asked for that no standard first-order logical system can be taken to be a satisfactory idealization of the psychological mechanisms or processes, whatever they might be, whereby we recognize (first-order!) logical consequences. (Boolos 1987: 380)

The conclusion Boolos draws is thus, in part, *psychological*. The conclusion that we wish to draw in a moment is *epistemological*.

In general, a weak system can have very long proofs for certain formulae, but these formulae become more rapidly provable in a stronger system, which usually has stronger existence assumptions (such as existence axioms for higher-types). We say that the stronger system exhibits

---

[1] For more on the Ackermann function, see Cutland 1980: 46–47.

[2] On the notion of closure, and related second-order concepts, see Shapiro 1991: ch. 5.

[3] See Boolos 1987, Appendix, for details.

'speed-up' over the weaker system. The speed-up phenomenon was first noted by Kurt Gödel (1936).[4] Gödel noted that, where Z_n is nth order arithmetic, then for any recursive function, g, there is an infinite class of formulae φ such that if k is the length of the shortest proof of φ in Z_n and l is the length of its shortest proof in Z_{n+1}, then $k > g(l)$. He added,

> *Thus, passing to the logic of the next higher order has the effect, not only of making provable certain propositions that were not provable before, but also of making it possible to shorten, by an extraordinary amount, infinitely many of the proofs already available.* (Gödel 1936: 397)

2. *Two more curious inferences*

Cardinality statements like 'The number of Fs $= n$' can be re-expressed in a first-order language, via the usual inductive definition of the numerically definite cardinality quantifiers, $\exists_n x F(x)$.[5] Writing variables with unary superscripts (x, x', x'', etc.), $\exists_n x F(x)$ has symbol complexity $O(n^3)$. This can be reduced, however, in two ways.[6] First, we can define 'There are at most n Fs' by $\exists x_1 \ldots \exists x_n \forall y (F(y) \supset y = x_1 \vee \ldots \vee y = x_n)$, and define 'There are at least $(n + 1)$ Fs' by '~(there are at most n Fs)'. For example, 'There are at least 3 Fs' is equivalent to $\forall x' \forall x'' \exists x \ (F(x) \ \& \ x \neq x' \ \& \ x \neq x'')$. Define $\exists_n x F(x)$ by 'There are at most n Fs and at least n Fs'. Then $\exists_n x F(x)$ has complexity $O(n^2)$. Second, code variable subscripts in binary (e.g. x_{100}, x_{101}, etc.). The binary numeral for n has length $O(\log_2 n)$. The symbol complexity of $\exists_n x F(x)$ is then reduced to $O(n \log_2 n)$. To illustrate, on this formalization, the complexity of 'There are at least one million books in Cambridge University library' is roughly 10^7. At roughly 10^3 symbols per metre, a sentence token expressing this would be around 10km long!

Consider the following inference I_2,

---

[4] This holds still when S* is a conservative extension of S. One example concerns Peano arithmetic PA and its conservative second-order extension ACA_0. Every arithmetic statement φ provable in ACA_0 is already provable in PA, but the shortest proof of an arithmetic statement φ in PA may be vastly longer than the shortest proof in ACA_0. For more details on proof-theoretic speed-up results, see Buss 1994.

[5] For completeness, where v is any variable, and w is a variable not appearing in $F(v)$, this is

$\exists_0 v F(v)$	for	$\sim\!\exists v F(v)$
$\exists_{\geq 1} v F(v)$	for	$\exists v F(v)$
$\exists_{\geq n+1} v F(v)$	for	$\exists v (F(v) \ \& \ \exists_{\geq n} w(w \neq v \ \& \ F(w)))$
$\exists_n v F(v)$	for	$\exists_{\geq n} v F(v) \ \& \ \sim\!\exists_{\geq n+1} v F(v)$

[6] Thanks to Samuel Buss for showing me how to do this.

(7)   The number of people in the room is 100.
(8)   The number of houses in the street is 99.
(9)   Each person in the room lives in exactly one house in the street.
Hence,
(10) At least two people in the room share the same house.

The premises and conclusion of I_2 can be expressed 'nominalistically' by the above method. Consider the problem of showing that I_2 is valid. Using rather simple mathematics dealing with numbers, sets and functions, we can prove that the inference I_2 is valid in a few lines. Briefly, the Pigeonhole Principle says that, for sets X and Y, if $\#X > \#Y$ and f is a function from X to Y, then f is non-injective. In set theory, this just follows from the definition of the relation $>$ for cardinalities. The Pigeonhole Principle can also be given an arithmetic formulation, and it is then provable in Peano arithmetic. Now, let P be the set of people, and let $\#P$ be the number of people, i.e. 100. Let H be the set of houses and let $\#H$ be the number of houses, i.e. 99. Let f be the function which maps each person to the unique house they live in. f is a function from P to H. But $100 > 99$, and so, $\#P > \#H$. So, by the Pigeonhole Principle, f is non-injective. So, there are at least two people a, b such that $f(a) = f(b)$.

This proof could be formalized within Frege Arithmetic, so long as we allow comprehension for the predicates 'person in the room', 'house in the street' and 'lives in'. Similarly, we can show that any model of the premises of I_2 makes the conclusion true.

So, the inference I_2 is valid in first-order logic. But the shortest first-order derivation of (10) from (7)–(9), in an austere nominalistic formulation, would be huge, probably more than 10^8 symbols.[7]

Next, consider the following inference I_3,

(11) Every UK citizen has a unique national insurance number.
(12) Different UK citizens have different national insurance numbers.
(13) Every national insurance number is the national insurance number of some UK citizen.
(14) The number of UK citizens is 60 million.
Hence,
(15) The number of national insurance numbers is 60 million.

---

[7] There is a large literature on the complexity of proofs of the pigeonhole principle, particularly when it is reformulated as a propositional formula, PHP_n. In general, the length of proof grows rather rapidly. My desktop computer automated theorem prover (SPASS) took some 21 minutes just to prove the tautology PHP_7. In 1987, Samuel Buss established that there is a polynomial size proof for PHP_n (with an estimated upper bound of n^{20}). However, it is safe to assume that, in general, such proofs are unfeasible for practical purposes.

Premises (11)–(13) say that there is a bijection (one-to-one correspondence) between the UK citizens and the national insurance numbers. The validity of this inference of course rests on the truth of a version of Hume's Principle, which implies that if there is a bijection between sets X and Y, then their cardinal numbers are the same. Some advocates of logicism insist that this principle is 'analytic', or a 'conceptual truth' about cardinal numbers. Perhaps so.

Slightly more formally, the inference I_3 has the form:

The As and Bs are one-to-one correlated. The number of As is n.
Hence, the number of Bs is n.

I_3 can again be formalized in first-order logic, and the conclusion (15) is a logical consequence of premisses (11)–(14). But the shortest derivation would again be practically huge. Formally, the first-order formalization of I_3 looks like this:

(16) $\forall x(A(x) \supset B(f(x)))$
(17) $\forall x_1 \forall x_2(x_1 \neq x_2 \supset f(x_1) \neq f(x_2))$
(18) $\forall y(B(y) \supset \exists x(A(x) \ \& \ y = f(x)))$
(19) $\exists_n x A(x)$
Hence,
(20) $\exists_n y B(y)$.

The point to stress is that although using *mathematics* we can quickly 'see' in the above cases that the conclusion is a logical consequence of the premisses, we could *not* in any feasible manner carry out the first-order derivation for any but the smallest values of the parameters. Thus, the relevant mathematics appears to be indispensable in seeing that these inferences are valid.

We have chosen the two inferences I_2 and I_3 above because of their mathematical simplicity, while Boolos's example turns on the mathematics of rapidly growing recursive functions. Unlike Boolos's example, the inferences I_2 and I_3 do not require *astronomically large* proofs, but they are large enough to be unfeasible in practice. As soon as one gets the gist of the idea behind such inferences, involving finite sets of objects and functions between them, it's easy to come up with examples of first-order representable inferences which are obvious but nonetheless unfeasible. For example, 'There are 20 people here, but I've only made 15 handouts. A few of you will have to share'. Non-mathematicians I have asked reply that such inferences are 'obvious'.

As Boolos notes, there is an issue raised here concerning the *psychological representation* of logical inference. But there is also an important *epistemological* issue, to which we now turn.

3. The unfeasibility problem for nominalism

Nominalism denies the existence of numbers, sets and functions. But a widely discussed problem concerns whether nominalism can account for the *applicability* of mathematics. This is the indispensability argument against nominalism, associated with Gödel, Quine and Putnam.[8] Above we examined examples of the application of mathematics to relationships of logical consequence. It seems to me that the 'speed-up' phenomenon under discussion suggests a modified version of the indispensability argument, based now on unfeasibility considerations. Presumably the nominalist does not wish to deny the validity of the inferences I, I_2, and I_3 under consideration. But there is no feasible direct verification for the above inferences, and the short mathematical derivations involve *practically indispensable* assumptions about numbers, sets and functions.[9] So, how might a nominalist account for our knowledge that such inferences are valid? After all, the anecdotal evidence is that even non-mathematicians find I_2 and I_3 'obvious'.

Boolos commented on the relevant 'psychological mechanisms or processes', but here the point is epistemological: it seems to me that *the nominalist cannot give a nominalistically acceptable reason for thinking that the above valid inferences are indeed valid*.[10]

University of Edinburgh
Edinburgh EH8 9JX, UK
jeffrey.ketland@ed.ac.uk

References

Boolos, G. 1987. A curious inference. *Journal of Philosophical Logic* 16: 1–12. Repr. in G. Boolos 1998, *Logic, Logic and Logic*. Cambridge, MA: Harvard University Press.

Buss, S. 1987. Polynomial size proofs of the propositional Pigeonhole Principle. *Journal of Symbolic Logic* 52: 916–27.

Buss, S. 1994. On Gödel's theorems on lengths of proof I. Number of lines and speed-up for arithmetics. *Journal of Symbolic Logic* 59: 737–56.

[8] The Quine and Putnam versions of this argument are familiar. Gödel's version, in some ways more interesting, may be found in the unpublished essay 'Is mathematics syntax of language?' (Gödel *1953/9–III).

[9] Let me note here (though a fuller discussion is necessary) that it seems to me that Field's point about conservation does not help here. Even if mathematics were conservative (and it isn't, as Gödel showed), the 'nominalistic' proofs are not merely more 'long-winded'; they are practically impossible (again, as Gödel showed).

[10] I would like to thank Samuel Buss, Timothy Chow, Thomas Forster, Torkel Franzén, Brian King, Mary Leng, Adam Rieger, Florian Steinberger and Ed Wallace for comments on the ideas in this article.

Cutland, N. 1980. *Computability: An Introduction to Recursive Function Theory*. Cambridge: Cambridge University Press.

Gödel, K. 1936. Über die Länge von Beweisen. *Ergebnisse eines mathematischen Kolloquiums* 7, 23–24. Page references to the English translation in S. Feferman et al. 1986, *Kurt Gödel: Collected Works*. Volume I (1929–36), 397–99. Oxford: Oxford University Press.

Gödel, K. 1953/9-III. Is mathematics syntax of language? Unpublished essay, in *Kurt Gödel: Collected Works. Volume III. Unpublished Essays and Letters*, ed. S. Feferman et al. 1995, 334–56. Oxford: Oxford University Press.

Shapiro, S. 1991. *Foundations Without Foundationalism: A Case for Second-Order Logic*. Oxford: Oxford University Press.

Why counterpart theory and three-dimensionalism are incompatible

JIM STONE

Suppose that God creates *ex nihilo* a bronze statue of a unicorn; later he annihilates it.[1] The statue and the piece of bronze occupy the same space for their entire career. If God had recast the bronze as a mermaid, the piece of bronze, not the statue, would have survived. As nothing can have and lack the capacity to survive the same change, they are distinct. Yet many philosophers find it incredible that two material things coincide *ever*, not to mention for their entire career. Here we have an apparently irrefutable argument for the apparently impossible conclusion that distinct physical things coincide in space and time.

Counterpart Theory (CT) offers a solution (Lewis 1986: § 4.5; Sider 2001: 113). Suppose that the statue and the bronze are the same enduring three-dimensional object (three-dimensional things persist by existing in their entirety at different times). The statue cannot survive being recast as a mermaid, the bronze can. According to CT, the first claim is true because no statue-counterpart of the statue is mermaid shaped and the second is true because the bronze has a mermaid-shaped bronze counterpart. Counterpart relations are similarity relations. As one thing can have resemblance relations to different sets of things, depending on

---

[1] This is a version of an example from Alan Gibbard (1975): we make a statue by joining two pieces of clay; then we smash the piece, destroying the statue too.