A PLEA FOR AUTOMATED LANGUAGE-TO-LOGICAL-FORM CONVERTERS

by

Joseph S. Fulda

> There is no algorithm that, given a syntactic string in a language, cranks out its unique logical form or semantic structure. (Stephen C. Levinson)

In this article, we will (a) argue that by Levinson's own lights, the unmistakable truth of his statement that 'There is no algorithm that, given a syntactic string in a language, cranks out its unique logical form or semantic structure' (Levinson 2000:8) does not invalidate our plea for automated language-to-logical-form converters; (b) provide a sketch of what we think it would take to make such converters a reality; and (c) suggest three (intended-to-be compelling) reasons, one linguistic, one philosophical, and one practical, which might justify the one-time, but nevertheless enormous expense involved in developing such converters.

1. Introduction

In this article, we will (a) argue that by Levinson's own lights, the unmistakable truth of his statement that 'There is no algorithm that, given a syntactic string in a language, cranks out its unique logical form or semantic structure' (Levinson 2000:8) does not invalidate our plea for automated language-to-logical-form converters (hereinafter automated converters); (b) provide a sketch of what we think it would take to make such converters a reality; and (c) suggest three (intended-to-be compelling) reasons, one linguistic, one philosophical, and one practical, which might justify the one-time, but nevertheless enormous expense involved in developing such converters.

2. Why Levinson's Unmistakably True Statement Is of No Concern Here

Before addressing this section's question, we first address its pre-supposition, that Levinson's statement is, in fact, 'unmistakably' true. After all, *very* few meta-logical truths which are not theorems of mathematics are 'unmistakably' true. What makes Levinson's truth unmistakable are the key words 'its unique logical form', for as everyone who has ever taught logic at any level tells his classes, there simply *is* no unique logical form corresponding to any proposition. 'Flies fly', for example, can be symbolized as simply F or as F_1f_2 (quantified appropriately – and variously) or as F_2f_1 (quantified appropriately – and variously) or relationally as $F_if_{i'}x$ (quantified appropriately – and variously over either or both values of i and i', with x free, and referring to the modality of flight – wings, a vehicle, etc.). What determines the choice of logical form – including the level of detail – is its intended use in (typically) the argument the logical form is supposed to disclose and render provable: Context, in a word.

This view is unmistakably Levinsonian. In the selfsame introduction from which the text quoted above is drawn, Levinson (2000:7-8) writes, *inter alia*,

> Aspects of semantic content ... can be specified by a recursive truth definition, but this is unlikely to have a direct cognitive counterpart. ... [Yet,] truth-conditional semantics viewed in the realist way – as a direct veridical mapping of semantic structures onto states of affairs (bypassing the head as it were) – is useful as a yardstick of human performance. ... So we can have our cake and eat it, too; we can use the insights of truth-conditional semantics without buying Realism, and without caring that it obviously fails to meet any criteria for adequacy as a cognitive model.

In other words, truth-functional logic 'works' in the relevant sense, and that is enough.[1]

I would add to Levinson's insight (that truth-functional logic works in the relevant sense), that, as I have discussed at great length elsewhere, it works far better than is commonly supposed – even on the troublesome cases surrounding the conditional. I would further add that it is a *very* simple formalism, although it is also one that is *extremely* subtle: Propositional logic can be easily and profitably well-taught to eighth-graders; predicate logic can be easily and profitably well-taught to eleventh-graders – although neither audience can genuinely appreciate many of the subtleties. Although a recent monograph (Bennett 2003) probably reflects the consensus of the philosophical community that conditional probabilities work better as a model of the linguistic entity than does truth-functional semantics, such a treatment adds far more in complexity of analysis than the additional light it sheds on the underlying linguistic and argumentative structures it models. It is my guess that this theory of conditionals, as applied to this domain – the modelling of linguistic and argumentative structures, can be easily and profitably well-taught only to at least upper-division college students. Moreover, notwithstanding all the difficulties the conditional poses in the truth-functional model, it remains a far easier 'sell' than this theoretical alternative – probably because the latter is so much further from being anything at all like any plausible model of cognitive processes.

Automated converters need be no more precise than truth-functional logic itself: They merely have to 'work' in the relevant sense. Heuristics will do what algorithms cannot do: The measure of the success of any human artifact – whether mathematical[2] or a software package – is whether it serves its intended purpose.

3. What an Automated Converter Would Require

One application of such a converter (see 4.1, below) is the automatic production of abstracts from logical form or, alternatively and equivalently, a highlighting system which marks parts of a text for subsequent human skimming, as described in detail and at length in Fulda (2006a, 2006b). Although we do suggest two other applications below, this is the only application I have thought through sufficiently to address most fully this section's concern. That said, I would be surprised if much-more-than-minor modifications to what follows were required for the other proposed applications (4.2 and 4.3, below).

Automatic translation into logical form sufficient for this purpose requires[3]:

(1) The identification of non-propositional content:

(a) Interrogatives – these typically either begin with one of a set of words (usually called 'wh-words') or a reemphasized declarative, ending with a '?' (but see (2d) below);

Exs.: What time is it now?

Speaker: He ate it.
Interlocutor: He [actually] ate it?

(Of course, it is the interlocutor's utterance that lacks propositional content, not the original speaker's).

(b) Interjections;

Ex.: Ugh!

(c) Parenthetical phrases within sentences;[4]

Ex.: Mathematics (the only truly easy subject) proved too difficult for him.

(d) Beginning-of-sentence conjunction-marker per (3a) and (3f) below;[5]

Exs.: That is his account. And I believe it. (after (3a))

That is his account – and I believe it. (after (3f))

(e) Rhetorical requests[6];

Ex.: Let us proceed.

(f) Endorsements (or contradictions) of cited materials;

Exs.: As Jacob Mey reminds us, context is the quintessential pragmatic concept.

Contrary to Aristotle's dismissive remarks in his *Politics* III:9, 1280a:34-1280b:33, a state, properly understood, is, indeed, 'established for the prevention of mutual crime and for the sake of exchange'.[7]

(2) The identification of distinct propositions:

(a) Sentences, separated by periods, usually;

Ex.: She is good. She is true. She is genuinely beautiful.

(b) Independent clauses separated by semicolons, usually;

Ex.: He appears to be a complicated man; the appearance is merely a sign of his utter simplicity.

(c) Occasionally, the clause preceding or following a colon, particularly if both clauses are independent;

Ex.: The Bible is inerrant: It does not have a human author.

(d) Following a rhetorical question, the answer (*i.e.*, the proposition asserted by the rhetorical question) must be taken as the affirmation or negation of a declarative rephrased from the interrogative – an exercise we all learned in grammar school;

Ex.: Do all of us not say 'We know' when we mean merely 'We believe'.
(The proposition is: We all say 'We know' when we mean merely 'We believe').

(3) The identification of connective-markers:

(a) And, But, Also, Additionally, Moreover, In addition to, Plus, With, Together with, etc. for conjunction;

Ex.: That is his weakness and his strength.

(b) Or, Alternatively, Either ... Or, Neither ... Nor, Whether … Or, Except, Unless, In exception to, On the other hand, Or else, etc. for disjunction, inclusive and exclusive;

Exs.: That can be seen as either his weakness or as his strength. (inclusive)

Depending on how and when used, it is his weakness or his strength. (exclusive)

(c) Thus, Consequently, Therefore, Then, If, Only if, If... Then, As a result, Accordingly, So, Hence, Else, Otherwise, etc. for the conditional and (sometimes the) biconditional;

Ex.: Machines do not err, properly so-called, although they do fail; so men are not machines.

(d) This means, I.e., That is to say, In other words, Equivalently, Alternatively (sometimes), Just the same as, Whenever (sometimes), Wherever (sometimes), ... and conversely, Is translated as, Is rendered as, Is, As, Is as, Or (sometimes), etc. for equivalence or the identity predicate;

Ex.: A capability is a developed ability.

(e) No, Not, – n't, Hardly, Scarcely, None, Never, Nor is it, Neither is it, On the contrary, Contrariwise, It fails as, etc. for negation;

Ex.: Sincerity palpably fails as a test for truth.

(f) Punctuation indicating conjunction – dashes, commas, and semicolons, colons (occasionally);

Ex.: Joseph S. Fulda, the present author, has never written for *RASK* before.

(Indicating that said author is (a) now writing for *RASK* AND (b) has not done so before.)

(g) Punctuation possibly indicating a conditional – the colon;

Ex.: The Bible is inerrant: It cannot have a human author. (*Cf.* (2c) above.)

(h) Punctuation possibly indicating disjunction – commas;

Exs.: The proper objects of literary analysis and study today are widely in dispute: Shakespeare's work, the writings of Queen Elizabeth I, and the ordinary telephone directory may all qualify. Each has form and content; each is copyrightable and, legally, an 'original' work of authorship; each is unquestionably a proper object of linguistic analysis and study. The debate therefore centers entirely on the meaning of the work 'literary'.[8]

(i) Punctuation possibly indicating equivalence or the identity predicate – dashes and sometimes colons, or even sometimes commas;

Ex.: Nero, the Emperor of Rome, was likely mad.

(4) The identification of quantification-markers as discussed at length in the notes to Fulda (2006a, 2006b).

(5) The resolution of pronominal constructs.

All of these tasks are well within the scope of such commercial programs as GRAMMATIK™ (let alone demonstration programs, which generally have greater abilities if on smaller domains), although GRAMMATIK's performance on task (5) leaves much to be desired. Still, as shown repeatedly in Fulda (2006a, 2006b) – see nn. (35), (46), (58), (60),[9] correct resolution of a reference may make no difference in the calculation, particularly if the reference is embedded. The method, as applied to the texts in the two experiments in Fulda (2006a, 2006b), proved remarkably robust, regardless of translational choices or inability to fully resolve ambiguities.

Yet, to claim that the development of an automated converter is likely easily doable with shallow parsing techniques – such as, say, those in GRAMMATK™ – on the basis of the apparent robustness of the technique is *most emphatically not to say it is easy* in the practical sense. Yes, the technology is there, but building a sophisticated grammar-checker or any other system at that level (as I believe an automated converter would be) would take probably tens of man-years (requiring either the industrial resources of a Corel or a Microsoft or the academic efforts of a dozen graduate students). It is easy in the sense that likely no new ideas are needed to bring it about, but it is *not at all easy* in practical terms. (Remember the mathematician who says, in a house aflame: 'Assume a hose. Problem solved'.) To justify an effort of such magnitude, and notwithstanding its being a one-time effort, compelling reasons must be advanced.

4.1. A Linguistic Application.

In 1970, Montague published his 'Universal Grammar', a very technical piece, which nonetheless opened with the bold declaration:

> There is in my opinion no important theoretical difference between natural languages and the artificial languages of logicians; indeed I consider it possible to comprehend the syntax and semantics of both kinds of languages within a single and mathematically precise theory. (Montague 1970:373)

Ever since Montague's 'mathematically precise theory' being his eponymous universal grammar, linguists have been looking for mechanical means – means underlying language rather than just languages, and means which, when understood properly, would allow, among other things, the transformation of a proposition in one natural language into an equivalent proposition in another natural language –

any other natural language. If there is a universal grammar, one method of mechanizing translation between natural languages would be simply to apply said grammar. If there is no such grammar, as is the majority opinion (including my own admittedly untutored opinion), machine translation could follow the path a good human translator takes, the two-step process of full *natural language understanding* of the proposition in the source language followed by the *natural language generation* of the same proposition in the target language. Of the two steps, it is the first step that is clearly the far more difficult one. One way of conceptualizing this two-step process is that of taking the material in natural language to full logical form, with predicates, objects, and quantifiers, and then taking the result back to natural language using the target language's linguistic mapping between words and grammatical structures, and logical form. This would be one end justifying the construction of automated converters.

Other ends besides translation between natural languages may involve this same two-step process, translation to logical form followed by natural language generation. In Fulda (2006a, 2006b), I considered two (intimately related) such other ends: the production of abstracts, and preprocessing for human skimming.

I gave a single proof-of-concept in Fulda (2006a) and another in Fulda (2006b) of these two (closely related) other ends. Proofs-of-concept are quite valuable as evidence suggestive of the sort of generalities scientists usually seek, although they can never serve as anything like an actual proof. This may explain why proofs-of-concept no longer seem much-valued by most of the AI community. One consequence of what might – notwithstanding this apparent consensus – in fact be a serious misjudgment is that perhaps the AI community has not kept what Nilsson (1995) referred to as its 'eye on the prize' of general intelligence, and has instead focused more on niche systems that perform outstandingly in limited domains. The 'toy systems' of early AI, considered collectively, were in my view quite impressive as demonstrations that the prize was attainable. Nilsson (1995) suggests that more general 'habile' systems will emerge that combine the virtues of both types of systems. This, of course, would be the technological equivalent of a 'proof', rather than a 'proof-of-concept'.

Nevertheless, *any* work claiming to have scientific credibility *must* be reproducible. Is my work in Fulda (2006a, 2006b) reproducible in the *most* relevant sense or is, instead, an automated converter absolutely required? I believe the answer to this two-pronged question is 'Yes' and 'No', respectively; yet notwithstanding that belief, it is also my view that automated converters nevertheless remain essential.

First, the explanation of my answer to the two-pronged question: A computer system which implemented the *a priori* methodology for translation to logical form would do so just as idiosyncratically as a logician would – with the particular choices made (procedurally or declaratively) by the designer of the system. Thus, it would bring to bear a consistent set of analyzing principles (defined by the code) to the text under translation, but almost surely no two such systems designed by different teams would translate text into logical form in the same way. Such a system would reproduce its own results only, something a logician – at least some logicians – can do as well. What is crucial for scientific validity – the sort of reproducibility that bears on the scientific value of Fulda (2006a, 2006b) – is whether my work *or* that of my system can and would be reproduced by someone else *or* his system. The reproduction need not be identical, of course, but it should be close enough that the main results hold, *regardless of which scholar's or system's translation is used.*

Thus, and second, the claim that the *a priori* method advanced in Fulda (2006a, 2006b) produces abstracts usefully can *only* be verified by having people *or* systems independently checking the specifics on the same texts. That is why I provided a very detailed translation of even those propositions which on my account were obviously not meaningful, and it is also why I annotated my work with 70 exegetical notes. In other words, the kind of reproducibility required for scientific validity cannot be guaranteed by a programmed implementation – that would only ensure the system's internal consistency (what is usually called *coherence*) not its external validity (what is usually called *correspondence*). Rather, what is needed is independent peer review of the experimental content itself, so that other scholars working independently of me are able to produce, in the main, the same results on the same text. In the case of the two texts analyzed in Fulda (2006a, 2006b), the peer-review process involved three reviewers, who, from what I can divine, performed both deliberated and random spot-checking, the most that can possibly be expected of human beings.

So while in the *most* relevant sense, the experiments I performed are likely, on the available evidence, reproducible, as already conceded both here and in Fulda (2006a, 2006b), they do not 'prove' the more general claim advanced – tentatively – that the method works on the types of texts it is intended to work on, as described in Fulda (2006a, 2006b), even subject to the limitations on that tentative claim made therein.

The astute reader will have already noticed that the present paper's title and, with notable exceptions, its text, speaks of 'automated converters', in the plural; this is not an accident. A single converter would be proof, even if used on a large variety of texts, only that the method as implemented by that particular converter works; at least two independently programmed converters would be needed for the machine equivalent of 'peer review'. Two converters, producing differing translations but finding more-or-less the same utility in the method, would provide both types of reproducibility – internal and external.

Automated converters would also obviate the need for the enormous expenditure of time and effort needed to translate such long texts as those analyzed in Fulda (2006a, 2006b). Moreover, listing the text and its translation in full, with dozens of exegetical notes carefully recording every translational and computational decision would no longer be needed. All that would be necessary for a complete specification are the original texts, the abstracts produced (or, alternatively and equivalently, text extracts for skimming), and the exact software release used! Fulda (2006a, 2006b), in contrast, required almost fifty (50) pages of journal space by an author deeply committed to economy of language, because a simple reference to a software release was precluded by the absence of any such system up to the task.

A final note. Even with both types of reproducibility, one cannot yet claim that the *a priori* method would make a contribution to AI as that term (in my view, unfortunately, but largely) is currently understood: That would not only depend on whether it generalizes, but on whether, if it does, it also outperforms other techniques directed towards the same goal, an empirical question we have yet to address – even with regard to the texts analyzed in Fulda (2006a, 2006b).

4.2. A Philosophical Application

Searle's (1980) Chinese Room *Gedankenexperiment*, as he himself labels it, depends, as all *Gedankenexperiments*[10] do, on the experiment being 'in principle possible', a point often overlooked. The argument against 'strong AI' that Searle makes depends, as does Fulda (2006a, 2006b), on a proof-of-concept: A man manipulating formal symbols formally, thereby producing Chinese output (answers to questions about a story) from Chinese input (the story), plainly does not understand (the) Chinese (story). *Ergo*, neither does a machine doing just that. But what if formal manipulations of formal symbols *cannot* do this task? The presupposition is that they can, because some such thing was, indeed, done by Roger Schank's team at Yale – in English – as cited by and discussed in Searle (1980). But if we adopt the (majority) view that a universal grammar does not exist, this presupposition might be unjustified. In that case, Searle's argument would no longer qualify as a *Gedankenexperiment*, although it might carry equally grave consequences – not for 'strong AI' (although it would still have very significant force even there; it would simply require some refinement and modification), but for the Church-Turing thesis (roughly stating the equivalence of all of a certain class of mechanisms to the Turing machine) and, therefore, for the philosophical foundations of computability theory.[11]

An automated converter could help decide the empirical question of whether such a formal symbol manipulation process as Searle proposes is, in principle, possible, although it could only decide it in the affirmative, akin to what is known in theoretical computer science as a semi-decidable problem. I tend to believe rather strongly that Searle is right both in his conclusion as to 'strong AI' and in the presupposition of his *Gedankenexperiment*, for which reason I don't find the prospect of obtaining a definitive answer pointing in only one direction disturbing.

4.3. An AI Application

One of the most productive areas of AI research has been automated theorem proving, or, more generally, automated reasoning. Indeed, a long-open conjecture's proof by machine[12] made the front page of the New York *Times* (1996). Another long-open conjecture, the four-color theorem[13], has thus far been provable *only* by machine, although that proof is not what is normally intended by 'automated theorem proving': I mention it only because automated theorem proving may progress to the point where it, like the different computational approach involved in the four-color theorem, can prove things no human being can prove. (The complete veracity of the proof is in some dispute, but the experts are agreed that it is, in the main and in its theoretical underpinnings, correct.)

However, there is a lacuna in the state-of-the-art. Theorems as well conjectures are expressed in a natural language, the very different (perhaps 'unnatural') natural language that mathematicians use involving a mixture of ordinary language, symbols, meta-symbols, and so on. Presently, this language is translated by hand into whatever particular form the automated reasoning program requires as input, thus making the process only semi-automatic, a man-machine process. Automated converters such as those proposed here hold the promise of making the process fully automatic. This can be of substantial importance, provided as in 4.1, above, there are more than one such system.

Schubert's famous steamroller problem, given in full in Walther (1985), and proved by his automated system using many-sorted resolution, as described in the pages of AI's leading journal, *Artificial Intelligence*, is not provable (*i.e.*, the conclusion does not follow from the premises) on what I – translating it manually, just as Walther did – regard as the best translation of the English-language statement of the problem. The success of the mechanical proof depends on a particular (and what I would regard as an inferior) translation of the English into logical form. Walther himself stated in correspondence with me years ago that he agrees with the preceding sentence, save the parenthetical insert – and, of course, the technical excellence of his methodology and program for many-sorted resolution in no way hinges on the translation fed to it by hand. Since Schubert's steamroller problem was an artificial challenge problem, it really doesn't matter: Walther's point was fully carried regardless of whether the system was or was not fed the best input. But as used today – on genuine mathematical conjectures or theorems – slight differences in the statement may change the truth value of something *inherently* important, and *that* is a *serious* matter. Having multiple, independently developed automated converters reduces this risk, and is yet another reason for expending the *one-time* effort needed to make them a reality.

701 West 177th Street, #21
New York, NY 10033
U.S.A.
fulda@acm.org

Acknowledgments

I would like to thank Professor Mey for inviting me to write this article for the *RASK* readership, my sister, Aviva, for her substantial help with the manuscript's preparation, and my one-time editor, Ron Arnold, for his learned assistance. This article is dedicated to Patrick J. Vincent.

Notes

1. Its insufficiency as a cognitive model is demonstrated by a large body of psychological literature on conditionals and reasoning in human beings. Conditionals are such a fundamental concept that they have generated large (and largely disjoint) literatures in philosophy, psychology, and linguistics, and smaller literatures in almost every other field: They have interested almost every thinking man since the beginning of recorded thought.
2. I am not here taking a position on Platonism. The human artifact is not the mathematics *per se* of truth-functional logics (*i.e.*, Boolean algebra(s)) which was discovered but not created by human beings, but instantiations of it (generally termed symbolic logic(s)) to model linguistic entities, and, as we

have already discussed, there is no unique (a necessary but not sufficient condition for Platonic existence) and therefore no perfect model.

3. Note that I do not claim 'requires and only requires'; such a claim could be made, if at all, only after the completion of the development of an actual converter.
4. For an explication of why such an apparently meaningful but parenthetical comment should not be considered propositional content *in the context of an automated converter* proposed for the purposes of Section 4.1, see note (5) in Fulda (2006a).
5. It is the 'and' that lacks propositional content; for a detailed explanation, see note (6) in Fulda (2006a).
6. But non-rhetorical imperatives are better treated as conditional declaratives after the manner described in Fulda (1995).
7. The propositional content begins with 'Context is' and '[A] state, properly understood,'.
8. The disjunction is 'Shakespeare's work may qualify as a proper object of literary study' OR 'Queen Elizabeth I's writings may …' OR 'The ordinary telephone directory may ….'

   This example contains numerous other propositions as well; as Levinson (2000:6, 135-153) convincingly argues, prolixity of this sort is a metalinguistic indication that the matter being discussed is far from simple.
9. The notes are examples of extremely remote references, which are certainly beyond the ken of any method (using merely the technical means proposed in Section 3) to resolve. Yet, this proves to make no difference whatsoever *for the intended purpose*, which does not, by any means – indeed, that is the whole point of Fulda (2006a, 2006b) – require genuine understanding.
10. The McGraw-Hill Dictionary of Scientific and Technical Terms (6th ed., 2003:887) gives this definition: 'A Gedanken ('thought') experiment is a hypothetical experiment which is possible in principle and is analyzed (but not performed) to test some hypothesis'.
11. In fairness to Searle, in his reply to Schank on the merits of 'weak AI', he discloses, as an aside (1980:453), that he is uncertain about whether the Chinese question-answering task he proposes can be done. But it doesn't seem, especially since the context of the aside is a discussion of 'weak AI', that he has at all grasped the larger implications of this for the foundation he builds against 'strong AI'.
12. This was widely considered the *first* truly creative proof discovered by machine. The proof of what was known as the Robbins Conjecture is due to William McCune and his system, developed at the Argonne National Laboratory. It is now, of course, no longer a conjecture.

13. The four-color theorem states that a planar map with *n* distinct regions can always be colored so that no two boundary-sharing regions on the map have the same color, regardless of *n*, provided one is allowed to use four colors. The reader wishing to get some intuition behind this problem is referred to Doerr and Levasseur (1985:224-225) on the map of Euler Island, for which three colors suffice, and to the (much-easier than the four-color case) proof of the five-color theorem in Doerr and Levasseur (1985:225-226).

References

Bennett, Jonathan. 2003. A Philosophical Guide to Conditionals. New York: Oxford University Press.

Doerr, Alan & Kenneth Levasseur. 1985. Applied Discrete Structures for Computer Science. Chicago: Science Research Associates.

Fulda, Joseph S. 1995. Reasoning with Imperatives Using Classical Logic. Sorites 3.7-11.

Fulda, Joseph S. 2006a. Abstracts from Logical Form: An Experimental Study of the Nexus between Language and Logic I. Journal of Pragmatics 38(5).778-807. http://dx.doi.org/10.1016/j.pragma.2005.10.001

Fulda, Joseph S. 2006b. Abstracts from Logical Form: An Experimental Study of the Nexus between Language and Logic II. Journal of Pragmatics 38(6).925-943. http://dx.doi.org/10.1016/j.pragma.2006.01.008

Levinson, Stephen C. 2000. Presumptive Meanings: The Theory of Generalized Conversational Implicature. Cambridge, Mass.: Bradford Books, MIT Press.

Montague, Richard. 1970. Universal Grammar. Theoria: A Swedish Journal of Philosophy 36(3).373-398.

New York Times. Discovery of proof announced on p. A1, with details on pp. C1, C8, December 10, 1996.

Nilsson, Nils. 1995. Eye on the Prize. AI Magazine 16(2).9-17. See also my review in Computing Reviews 37(11).595.

Searle, John R. 1980. Minds, Brains, and Programs. The Behavioral and Brain Sciences 3.417-457. The article itself runs from pp. 417-424, followed by 'open peer commentary' by leading scholars – one of BBS's outstanding features – and the author's reply, ending at p. 457. See also my 1989 review (of the article only) in Computing Reviews 30(3).163.

Walther, Christoph. 1985. A Mechanical Solution of Schubert's Steamroller by Many-Sorted Resolution. Artificial Intelligence 26.217-224.