FALLIBLE RATIONALISM AND MACHINE TRANSLATION

Geoffrey Sampson

Department of Linguistics & Modern English Language
University of Lancaster
LANCASTER LA1-4YT, G.B.

ABSTRACT

Approaches to MT have been heavily influenced by changing trends in the philosophy of language and mind. Because of the artificial hiatus which followed the publication of the ALPAC Report, MT research in the 1970s and early 1980s has had to catch up with major developments that have occurred in linguistic and philosophical thinking; currently, MT seems to be uncritically loyal to a paradigm of thought about language which is rapidly losing most of its adherents in departments of linguistics and philosophy. I argue, both in theoretical terms and by reference to empirical research on a particular translation problem, that the Popperian "fallible rationalist" view of mental processes which is winning acceptance as a more sophisticated alternative to Chomskyan "deterministic rationalism" should lead MT researchers to redefine their goals and to adopt certain currently-neglected techniques in trying to achieve those goals.

1.   Since the Second World War, three rival views of the nature of the human mind have competed for the allegiance of philosophically-minded people. Each of these views has implications for our understanding of language.

The 1950s and early 1960s were dominated by a behaviourist approach tracing its ancestry to John Locke and represented recently e.g. by Leonard Bloomfield and B.F. Skinner. On this view, "mind" is merely a name for a set of associations that have been established during a person's life between external stimuli and behavioural responses. The meaning of a sentence is to be understood not as the effect it has on an unobservable internal model of reality but as the behaviour it evokes in the hearer.

During the 1960s this view lost ground to the rationalist ideas of Noam Chomsky, working in an intellectual tradition founded by Plato and reinaugurated in modern times by Rene Descartes. On this view, stimuli and responses are linked only indirectly, via an immensely complex cognitive mechanism having its own fixed principles of operation which are independent of experience. A given behaviour is a response to an internal mental event which is determined as the resultant of the initial state of the mental apparatus together with the entire history of inputs to it. The meaning of a sentence must be explained in terms of the unseen responses it evokes in the cognitive apparatus,

which might take the form of successive modifications of an internal model of reality that could be described as "inferencing".

Chomskyan rationalism is undoubtedly more satisfactory as an account of human cognition than Skinnerian behaviourism. By the late 1970s, however, the mechanical determinism that is part of Chomsky's view of mind appeared increasingly unrealistic to many writers. There is little empirical support, for instance, for the Chomskyan assumptions that the child's acquisition of his first language, or the adult's comprehension of a given utterance, are processes that reach well-defined terminations after a given period of mental processing -- language seems typically to work in a more "open-ended" fashion than that. Within linguistics, as documented e.g. by Moore & Carling (1982), the Chomskyan paradigm is by now widely rejected.

The view which is winning widespread acceptance as preserving the merits of rationalism while avoiding its inadequacies is Karl Popper's fallibilist version of the doctrine. On this account, the mind responds to experiential inputs not by a deterministic algorithm that reaches a halt state, but by creatively formulating fallible conjectures which experience is used to test. Typically the conjectures formulated are radically novel, in the sense that they could not be predicted even on the basis of ideally complete knowledge of the person's prior state. This version of rationalism is incompatible with the materialist doctrine that the mind is nothing but an arrangement of matter and wholly governed by the laws of physics; but, historically, materialism has not commonly been regarded as an axiom requiring no argument to support it (although it may be that the ethos of Artificial Intelligence makes practitioners of this discipline more than averagely favourable towards materialism).

As a matter of logic, fallible conjectures in any domain can be eliminated by adverse experience but can never be decisively confirmed. Our reaction to linguistic experience, consequently, is for a Popperian both non-deterministic and open-ended. There is no reason to expect a person at any age to cease to improve his knowledge of his mother-tongue, or to expect different members of a speech-community to formulate identical internalized grammars; and understanding an individual utterance is a process which a person can

execute to any desired degree of thoroughness --
we stop trying to improve our understanding of a
particular sample of language not because we reach
a natural stopping-place but because we judge that
the returns from further effort are likely to be
less than the resources invested.

For a Chomskyan linguist, divergences between
individuals in their linguistic behaviour are to be
explained either in terms of mixture of "dialects"
or in terms of failure of practical "performance"
fully to match the abstract "competence" possessed
by the mature speaker. For the Popperian such
divergences require no explanation; we do not
possess algorithms which would lead to correct
results if they were executed thoroughly. Indeed,
since languages have no reality independent of
their speakers, the idea that there exists a
"correct" solution to the problem of acquiring a
language or of understanding an individual sent-
ence ceases to apply except as an untheoretical
approximation. The superiority of the Popperian
to the Chomskyan paradigm as a framework for
interpreting the facts of linguistic behaviour is
argued e.g. in my Making Sense (1980), Popperian
Linguistics (in press).

2. There is a major difference in style between
the MT of the 1950s and 1960s, and the projects of
the last decade. This reflects the difference
between behaviourist and deterministic-rationalist
paradigms. Speaking very broadly, early MT
research envisaged the problem of translation as
that of establishing equivalences between observ-
able, surface features of languages: vocabulary
items, taxemes of order, and the like. Recent MT
research has taken it as axiomatic that successful
MT must incorporate a large AI component. Human
translation, it is now realized, involves the
understanding of source texts rather than mere
transliteration from one set of linguistic con-
ventions to another: we make heavy use of infer-
encing in order to resolve textual ambiguities.
MT systems must therefore simulate these inferenc-
ing processes in order to produce human-like out-
put. Furthermore, the Chomskyan paradigm incorp-
orates axioms about the kinds of operation char-
acteristic of human linguistic processing, and MT
research inherits these. In particular, Chomsky
and his followers have been hostile to the idea
that any interesting linguistic rules or processes
might be probabilistic or statistical in nature
(e.g. Chomsky 1957: 15-17, and cf. the controversy
about Labovian "variable rules"). The assumption
that human language-processing is invariably an
all-or-none phenomenon might well be questioned
even by someone who subscribed to the other tenets
of the Chomskyan paradigm (e.g. Suppes 1970), but
it is consistent with the heavily deterministic
flavour of that paradigm. Correspondingly, recent
MT projects known to me seem to make no use of
probabilities, and anecdotal evidence suggests
that MT (and other AI) researchers perceive pro-
posals for the exploitation of probabilistic tech-
niques as defeatist ("We ought to be modelling
what the mind actually does rather than using
purely artificial methods to achieve a rough
approximation to its output").

3. What are the implications for MT, and for AI
in general, of a shift from a deterministic to a
fallibilist version of rationalism? (On the
general issue see e.g. the exchange between
Aravind Joshi and me in Smith 1982.) They can be
summed up as follows.

First, there is no such thing as an ideal
speaker's competence which, if simulated mechanic-
ally, would constitute perfect MT. In the case of
"literary" texts it is generally recognised that
different human translators may produce markedly
different translations none of which can be con-
sidered more "correct" than the others; from the
Popperian viewpoint literary texts do not differ
qualitatively from other genres. (Referring to
the translation requirements of the Secretariat of
the Council of the European Communities, P.J.
Arthern (1979: 81) has said that "the only quality
we can accept is 100% fidelity to the meaning of
the original". From the fallibilist point of view
that is like saying "the only kind of motors we
are willing to use are perpetual-motion machines".)

Second, there is no possibility of designing
an artificial system which simulates the actions
of an unpredictably creative mind, since any
machine is a material object governed by physical
law. Thus it will not, for instance, be possible
to design an artificial system which regularly
uses inferencing to resolve the meaning of given
texts in the same way as a human reader of the
texts. There is no principled barrier, of course,
to an artificial system which applies logical
transformations to derive conclusions from given
premisses. But an artificial system must be
restricted to some fixed, perhaps very large, data-
base of premisses ("world knowledge"). It is
central to the Popperian view of mind that human
inferencing is not limited to a fixed set of pre-
misses but involves the frequent invention of new
hypotheses which are not related in any logical
way to the previous contents of mind. An MT
system cannot aspire to perfect human performance.
(But then, neither can a human.)

Third: a situation in which the behaviour of
any individual is only approximately similar to
that of other individuals and is not in detail
predictable even in principle is just the kind of
situation in which probabilistic techniques are
valuable, irrespective of whether or not the pro-
cesses occurring within individual humans are
themselves intrinsically probabilistic. To draw
an analogy: life-insurance companies do not con-
demn the actuarial profession as a bunch of cop-
outs because they do not attempt to predict the
precise date of death of individual policyholders.
MT research ought to exploit any techniques that
offer the possibility of better approximations to
acceptable translation, whether or not it seems
likely that human translation exploits such tech-
niques; and it is likely that useful methods will
often be probabilistic.

Fourth: MT researchers will ultimately need
to appreciate that there is no natural end to the
process of improving the quality of translation
(though it may be premature to raise this issue

at a stage when the best mechanical translation is still quite bad). Human translation always involves a (usually tacit) cost-benefit analysis: it is never a question of "How much work is needed to translate this text 'properly'?" but of "Will a given increment of effort be profitable in terms of achieved improvement in translation?" Likewise, the question confronting MT is not "Is MT possible?" but "What are the disbenefits of translating this or that category of texts at this or that level of inexactness, and how do the costs of reducing the incidence of a given type of error compare with the gains to the consumers?"

4. The value of probabilistic techniques is sufficiently exemplified by the spectacular success of the Lancaster-Oslo-Bergen Tagging System (see e.g. Leech et al. 1983). The LOB Tagging System, operational since 1981, assigns grammatical tags drawn from a highly-differentiated (134-member) tag-set to the words of "real-life" English text. The system "knows" virtually nothing of the syntax of English in terms of the kind of grammar-rules believed by linguists to make up the speaker's competence; it uses only facts about local transition-probabilities between form-classes, together with the relatively meagre clues provided by English morphology. By late 1982 the output of the system fell short of complete success (defined as tagging identical to that done independently by a human linguist) by only 3.4%. Various methods are being used to reduce this failure-rate further, but the nature of the techniques used ensures that the ideal of 100% success will be approached only asymptotically. However, the point is that no other extant automatic tagging-system known to me approaches the current success-level of the LOB system. I predict that any system which eschews probabilistic methods will perform at a significantly lower level.

5. In the remainder of this paper I illustrate the argument that human language-comprehension involves inferencing from unpredictable hypotheses, using research of my own on the problem of "referring" pronouns.

My research was done in reaction to an article by Jerry Hobbs (1976). Hobbs provides an unusually clear example of the Chomskyan paradigm of AI research, since he makes his methodological axioms relatively explicit. He begins by defining a complex and subtle algorithm for referring pronouns which depends exclusively on the grammatical structure of the sentences in which they occur. This algorithm is highly successful: tested on a sample of texts, it is 88.3% accurate (a figure which rises slightly, to 91.7%, when the algorithm is expanded to use the simple kind of semantic information represented by Katz/Fodor "selection restrictions"). Nevertheless, Hobbs argues that this approach to the problem of pronoun resolution must be abandoned in favour of a "semantic algorithm", meaning one which depends on inferencing from a data-base of world knowledge rather than on syntactic structure. He gives several reasons; the important reasons are that the syntactic

approach can never attain 100% success, and that it does not correspond to the method by which humans resolve pronouns.

However, unlike Hobbs's syntactic algorithm, his semantic algorithm is purely programmatic. The implication that it will be able to achieve 100% success -- or even that it will be able to match the success-level of the existing syntactic algorithm -- rests purely on faith, though this faith is quite understandable given the axioms of deterministic rationalism.

I investigated these issues by examining a set of examples of the pronoun it drawn from the LOB Corpus (a standard million-word computer-readable corpus of modern written British English -- see Johansson 1978). The pronoun it is specially interesting in connexion with MT because of the problems of translation into gender-languages; my examples were extracted from the texts in Category H of the LOB Corpus, which includes governmental and similar documents and thus matches the genres which current large-scale MT projects such as EUROTRA aim to translate. I began with 338 instances of it; after eliminating non-referential cases I was left with 156 instances which I examined intensively.

I asked the following questions:

(1) In what proportion of cases do I as an educated native speaker feel confident about the intended reference?

(2) Where I do feel confident and Hobbs's syntactic algorithm gives a result which I believe to be wrong, what kind of reasoning enabled me to reach my solution?

(3) Where Hobbs's algorithm gives what I believe to be the correct result, is it plausible that a semantic algorithm would give the same result?

(4) Could the performance of Hobbs's syntactic algorithm be improved, as an alternative to replacing it by a semantic algorithm?

It emerged that:

(1) In about 10% of all cases, human resolution was impossible; on careful consideration of the alternatives I concluded that I did not know the intended reference (even though, on a first relatively cursory reading, most of these cases had not struck me as ambiguous). An example is:

The lower platen, which supports the leather, is raised hydraulically to bring it into contact with the rollers on the upper platen ... (H6.148)

Does it refer to the lower platen or to the leather (la platina, il cuoio!)? I really don't know. In at least one instance (not this one) I reached different confident conclusions about the same case on different occasions (and this suggests that there are likely to be other cases which I have confidently resolved in ways other than the writer intended). The implication is

that a system which performs at a level of success much above 90% on the task of resolving referential it would be outperforming a human, which is contradictory: language means what humans take it to mean.

(2) In a number of cases where I judged the syntactic algorithm to give the wrong result, the premisses on which my own decisions were based were propositions that were not pieces of factual general knowledge and which I was not aware of ever having consciously entertained before producing them in the course of trying to interpret the text in question. It would therefore be quixotic to suggest that these propositions would occur in the data-base available to a future MT system. Consider, for instance:

Under the "permissive" powers, however, in the worst cases when the Ministry was right and the M.P. was right the local authority could still dig its heels in and say that whatever the Ministry said it was not going to give a grant. (H16. 24)

I feel sure that it refers to the local authority rather than the Ministry, chiefly because it seems to me much more plausible that a lower-level branch of government would refuse to heed requests for action from a higher-level branch than that it would accuse the higher-level branch of deceit. But this generalization about the sociology of government was new to me when I thought it up for the purpose of interpreting the example quoted (and I am not certain that it is in fact universally true).

(3) In a number of cases it was very difficult to believe that introduction of semantic considerations into the syntactic algorithm would not worsen its performance. Here, an example is:

... and the Isle of Man. We do by these Presents for Us, our Heirs and Successors institute and create a new Medal and We do hereby direct that it shall be governed by the following rules and ordinances ... (H24.16)

Hobbs's syntactic algorithm refers it to Medal, I believe rightly. Yet before reading the text I was under the impression that medals, like other small concrete inanimate objects, could not be governed; while territories like the Isle of Man can be, and indeed are. Syntax is more important than semantics in this case.

(4) There are several syntactic phenomena (e.g. parallelism of structure between successive clauses) which turned out to be relevant to pronoun resolution but which are ignored by Hobbs's algorithm. I have not undertaken the task of modifying the syntactic algorithm in order to exploit these phenomena, but it seems likely that the already-good performance of the algorithm could be further improved.

It is also worth pointing out that accepting the legitimacy of probabilistic methods allows one to exploit many crude (and therefore cheaply-exploited) semantic considerations, such as Katz/Fodor selection restrictions, which have to be left out of a deterministic system because in practice they are sometimes violated. As we have seen, Hobbs suggested that only a small percentage improvement in the performance of his pure syntactic algorithm could be achieved by adding semantic selection restrictions. Rules such as "the verb 'fear' must have an [+animate] subject" almost never prove to be exceptionless in real-life usage: even genres of text that appear soberly literal contain many cases of figurative or extended usage. This is one reason why advocates of a "semantic" approach to artificial language-processing believe in using relatively elaborate methods involving complex inferential chains -- though they give us little reason to expect that these techniques too will not in practice be bedevilled by difficulties similar to those that occur with straightforward selection restrictions. However, while it may be that the subject of 'fear' is not always an animate noun, it may also be that this is true with much more than chance frequency. If so, an artificial language-processing system can and should use this as one factor to be balanced against others in resolving ambiguities in sentences containing 'fear'.

6. To sum up: the deterministic-rationalist philosophical paradigm has encouraged MT researchers to attempt an impossible task. The fallible-rationalist paradigm requires them to lower their sights, but may at the same time allow them to attain greater actual success.

REFERENCES

Arthern, P.J. (1979) "Machine translation and computerized terminology systems". In Barbara Snell, ed., Translating and the Computer. North-Holland.

Chomsky, A.N. (1957) Syntactic Structures. Mouton.

Hobbs, J.R. (1976) "Pronoun resolution". Research Report 76-1. Department of Computer Sciences, City College, City University of New York.

Johansson, S. (1978) "Manual of information to accompany the Lancaster-Oslo/Bergen Corpus of British English, for use with digital computers". Department of English, University of Oslo.

Leech, G.N., R. Garside, & E. Atwell (1983) "The automatic grammatical tagging of the LOB Corpus". ICAME News no. 7, pp. 13-33. Norwegian Computing Centre for the Humanities.

Moore, T. & Christine Carling (1982) Understanding Language. Macmillan.

Sampson, G.R. (1980) Making Sense. Oxford University Press.

Sampson, G.R. (in press) Popperian Linguistics. Hutchinson.

Smith, N.V., ed. (1982) Mutual Knowledge. Academic Press.

Suppes, P. (1970) "Probabilistic grammars for natural languages". Synthese vol. 22, pp. 95-116.