A LOGICAL PARADOX
AND THE FINITENESS OF NATURAL LANGUAGE

Arthur Skidmore
University of Kansas

Two powerful and contradictory intuitions govern a priori reflections on whether natural languages contain finitely or infinitely many sentences   The first of these is connected with the criterion that a sentence of a natural language be recognizable as such by a native speaker   Simple considerations connected with the finiteness of human capacities and ultimately of human life insistently suggest that the number of sentences of a natural language, though undoubtedly very large, must be finite

This point is made rather pungently in Olmstead 1967 (304), "The assumption that English sentences are not always less than a million words in length smells of the computer center rather than of the field situation " Even if the native speaker of our natural language were possessed of a monumental memory and were heroically loquacious, it hardly makes sense to speak of him as uttering or recognizing sentences of such stupendous length This unless he were idealized to such an extent as to equip him for conversation only with God

On the other hand, the existence of recursive grammatical constructions in natural languages leads one to the opposite conclusion   For if there were only finitely many sentences of a natural language, then there must be one of maximal length Applying such a construction to this sentence would then produce a still longer sentence   Hence the number of sentences in a natural language is infinite

This intuition is rather clearly expressed in

Grinder and Elgin 1973 (24)

    .. while any particular sentence of English
is finite in length, the set of all well-
formed sentences in English is infinite.
There can be no "longest possible sentence "
To any given sequence constituting a
grammatical English sentence, you can add
another relative clause, and no natural
limit can be set at which such structures
become incomprehensible

    The aim of what follows is to show that
considerations of the sort just mentioned are
specious   My method of attack will consist in
showing the close similarities between the
recursiveness argument and an important logical
paradox

    There are many different ways in which the
paradox I am concerned to expose may be formulated
It can in particular be formulated for virtually
any concrete empirical predicate which admits of
a class of clear instances and also of a class of
clear counterinstances   I have selected the
predicate 'is poor' primarily for reasons of
technical convenience.

    Consider the following argument

(Al) (Pl) An American family with a yearly income
      of only 10 dollars is poor
      (Px) For any x, if an American family with
      a yearly income of only x dollars is poor,
      then an American family with a yearly income
      of only x+5 dollars is poor   Therefore,
      (Cl) An American family with a yearly income
      of only 15 dollars is poor

    I am presuming it to be uncontroversial that
(Pl) is in fact true   To be sure, one could imagine
bizarre circumstances in which (Pl) might fail to
hold completely invariably   But I am supposing that
such unusual circumstances do not in fact obtain

There is even less difficulty in accepting (Px) as
true, since it amounts merely to the claim that enrich-
ing a poor family by five dollars will not remove them
from their poverty   Five dollars is too small a sum to
make enough of a difference to remove a family from the
class of poor families and firmly establish them in the
class of the non-poor

Next, according to the principles of modern
quatifier logic (and the additional tacit premiss
'10+5=15'), argument (A1) is <u>valid</u>   if its premisses
are true then its conclusion cannot be false   Since its
premisses are in fact obviously true, we are committed
by modern logic to the truth of (C1)   And we see in
this case that (C1) is in fact also true

Suppose we now make conclusion (C1) the first
premiss of a new argument (A2)

(A2)   (P2) An American family with a yearly income
          of only 15 dollars is poor
          (Px) For any x, if an American family with a
          yearly income of only x dollars is poor, then
          an American family with a yearly income of
          only x+5 dollars is poor   Therefore, (C2)
          An American family with a yearly income of only
          20 dollars is poor

Exactly the same reflections appropriate to argu-
ment (A1) continue to apply to argument (A2)   Addi-
tionally, however, we note that the truth of (P2) is
guaranteed by virtue of its having been the conclusion
of argument (A1), and so we have the stronger result
which is expressed by argument (A2)

(A2')   (P1) An American family with a yearly income
          of only 10 dollars is poor
          (Px) For any x, if an American family with a
          yearly income of only x dollars is poor, then
          an American family with a yearly income of
          only x+5 dollars is poor   Therefore, (C2)
          An American family with a yearly income of only
          20 dollars is poor

But now the paradox arises. For there is nothing to prevent us from continuing to construct a long list of arguments, using a computer or a bevy of graduate students if need be, which includes such arguments as (A998*) and (A3998*).

(A998*)   (P1)   An American family with a yearly income of only 10 dollars is poor.
          (Px)   For any x, if an American family with a yearly income of only x dollars is poor, then an American family with a yearly income of only x+5 dollars is poor   Therefore, (C998) An American family with a yearly income of only 5000 dollars is poor

(A3998*) (P1).   (Px)   Therefore,
          (C3998) An American family with a yearly income of only 20000 dollars is poor

But here we are obviously in trouble   For (C3998) is surely false. And yet the validity of each argument in the series guarantees the truth of each conclusion, and hence the truth of (C3998)

Notice incidentally that none of the arguments is excessively long   That is, we are not faced here with the problem of requiring that the arguments in the series contain premisses of indefinite length   If we were to continue the series of arguments indefinitely, though, this problem might well arise, since the numerals in the premisses and conclusions would become excessively long

This paradox may not be lightly dismissed   For surely the principles of logic must not become suspect merely because applied too often   And if argument (A1) is valid, then argument (A3998*) is also valid, since each of the previous arguments in the series is valid. If we reject the reasoning which has led us to (C3998), then we must be able to explain where the reasoning breaks down, and why   The force of the paradox is that it seems impossible to point the finger of blame at any one of the arguments in the series.

Solving the paradox by rejecting either premiss
(P1) or (Px) is of course something which is open to us
But these premisses are obviously true, and I doubt
whether it would occur to anyone to doubt them were it
not for the fact that they lead us to the false conclu-
sion (C3998)

I wish now to set out what seems to me to be the
most coherent solution to the paradox   It would be
wrong to dismiss out of hand other proposed solutions,
the great thing about a good paradox is that it opens
up various lines of inquiry for tracing precisely what
are the hidden premisses of a piece of reasoning   A
really good paradox, such as I think ours is, fascinates
us because the hidden premisses are so well hidden and
so few

Although modern quantifier logic requires very
little of the arguments to which it is to be applied,
it does make several stipulations or requirements   The
two most important of these may be expressed as the
requirements of bivalence and non-contradiction

The requirement of non-contradiction is that the
sentences to which logical theory is applied may not be
both true and false   This requirement does not play a
role in the paradox

The requirement of bivalence is that the sentences
to which logical theory is applied are to be either
true or false   It is not that modern logic assumes
that every declarative sentence is either true or false
It is just that modern quantifier logic has little to
say about arguments constructed from sentences which
fail to satisfy the bivalence requirement

The bivalence requirement is normally taken for
granted, due in part to the desire of logicians to deal
with the form of sentences as abstracted from the con-
crete circumstances of their use   Normally, that is,
all that a logician will require is that the arguments
he examines be composed of sentences having a trans-
parent logical structure

The arguments in the paradox have a sufficiently transparent logical structure to make the validity of their common form quite evident   And under normal circumstances the logical form is all that matters But the paradox reveals that more than the form of the arguments is at issue

Modern logic does guarantee that none of the separate arguments (A1)-(A3998) permits us to pass from true premisses to a false conclusion   But it does not rule out the possibility that we may pass from true premisses to an indeterminate conclusion, that is, a conclusion which is neither true nor false, or from premisses one of which is indeterminate to a false conclusion   And this is, I would suggest, precisely what has happened during the series of arguments   Somewhere in the region of (A998) the conclusions have become indeterminate

As a result, the argument (A3998*) may well be expected to have true premisses and a false conclusion, since by transitivity of entailment the indeterminate conclusions have all dropped out   Only an underlying presumption that all of the sentences under consideration are either true or false has led to the expectation that all of the conclusions must be true

The most common source of failure of a sentence to satisfy the bivalence requirement is the application of an empirical predicate to a borderline case   It has become a philosophical truism to observe that empirical predicates do not have sharp boundaries   Normally there are some cases to which it clearly applies, some cases to which it clearly does not apply, and a region of borderline cases to which it neither clearly applies nor clearly does not apply   And the most coherent interpretation of sentences which result from applying an empirical predicate, such as 'is poor', to a borderline case is I would suggest to treat them as neither true nor false, and thus as failing to satisfy the bivalence requirement

It should be noted that one seldom is able to say precisely where the borderline cases begin   It seems

clear enough to me, for example, that (C998) is neither
true nor false   Yet I am not prepared to say just
where the cases become borderline   An attempt to be
more precise will often result in a rather arbitrary
stipulation, such as a guideline laid down by the
Department of Health, Education, and Welfare

     If the empirical predicate in question is made
sharp, then paradox is avoided in another way, namely,
by observing that for such predicates the second premiss
(Px) is false   Indeed, I think it is clear that it is
precisely the existence of indeterminate cases which
guarantees the plausible truth of (Px)   A strict
definition of poverty, such as 'having a yearly income
of less than 5000 dollars', does make it possible some-
times for five dollars to remove a family from poverty,
no doubt to their dismay

     Being unable to say precisely where the indeter-
minate cases begin and end must not blind us to their
existence   For an inability to take account of border-
line cases may lead us into a very common fallacy,
which might be called the Where do you draw the line
fallacy   Noticing correctly that there is no sharp
distinction between cases of poverty and cases of non-
poverty, but rather a continuous blending of the one
into the other through an indeterminate region, we may
if we are not careful fall into the trap of concluding
that there is no difference between the two conditions

     Similarly, our inability in general to say
precisely which straw made the camel's burden too heavy
to bear should not lead us to conclude that it could
bear any burden   A poignant image for us to respect
in this regard may be derived from Truffaut's great
film, The 400 Blows   There is no single point in the
film where the young boy has been injured too badly to
recover, but he has clearly suffered irreversible damage
at the end

     Although the point to be made against the recursive-
ness argument is probably clear enough by now, I shall
nevertheless spell it out in order to make sure

The recursiveness argument can be rephrased in such
a way that its being an instance of the paradox under
discussion is made evident   For consider the following
argument

(Arl)   (Prl)   There is an English sentence containing
         10 words.
         (Prx)   For any x, if there is an English sentence
         containing x words, then there is an English
         sentence containing x+5 words.   Therefore,
         (Crl)   There is an English sentence containing
         15 words

Premiss (Prl) is obviously true.   Premiss (Prx) is
also clearly true, if one takes into account, for example,
the prefixing of the words 'It is not true that' to any
English sentence   The conclusion (Crl) is logically
entailed by the premisses, and it is also obviously true.

We can now reconstruct the paradox, leading us
successively to such conclusions as

(Cr998)   There is an English sentence containing 5000
           words.
(Cr3998)   There is an English sentence containing 20000
           words.
(Crl99998)   There is an English sentence containing
             1000000 words

But now we are not compelled by logical grounds
alone to the truth of (Crl99998)   For now we under-
stand what we may demand, namely, assurance that the
sentences composing the arguments all satisfy the bi-
valence requirement.   That is, we are not being
irrational in assenting to the truth of (Prl) and (Prx)
and yet questioning the truth of (Crl99998) or even
(Cr3998).   For as we have seen, this method of chaining
arguments is paradoxical

Of course to say that the reasoning just illus-
trated is paradoxical does not mean that the conclusions
need be false   Our assurance that (Crl99998), for
example, is false must rest on independent grounds.

I do think, however, that the considerations mentioned as the first intuition are compellingly persuasive regarding the finiteness of natural language, providing only that 'is a sentence of a natural language' is being used in an <u>empirical</u> sense And there is some evidence that linguists whose theories appear to commit them to the existence of infinitely many sentences of natural language are sensitive to this point   I would note for example the distinction drawn in Bach 1974 (25) between the infinitely many <u>grammatical</u> sentences of English as opposed to the finitely many <u>acceptable</u> sentences of English

Although I would quarrel with Bach's grounds for making his distinction, I find no fault with his con-clusion, which I take to be that there need be no exact fit between an abstract <u>theory</u> and the empirical domain to which it is intended to be applied   In particular, the abstract theory may well concern an infinite domain

At the risk of simplifying the formal linguist's task beyond all recognition, let me attempt to explain a bit the point just raised   Suppose, following Quine 1961 (54), we take the task of the formal linguist to be to construct a purely formal device which generates precisely the sentences of English   Reasons of theoretical economy may well lead him to produce an infinitely large set S   But this will not matter very much, as long as S corre-sponds reasonably well to the examined sentences of English   For problems of applying the theory to sequences of unmanageable length are by hypothesis not going to arise

What I have tried to show is that the puzzle con-cerning the finiteness or otherwise of natural language is not grounded in special considerations involving only linguists   It is just one manifestation of a logical paradox which may pop up almost anywhere

Where the presumption of bivalence obtains, as in arithmetic, we may take a proof of the claim

(1)   There is no largest prime.

to be equivalent to the claim

(2)   There are infinitely many primes

This is so because the series of arguments constituting
the paradox may not pass from true premisses to a false
conclusion, and since the only alternative to being
false is being true, the arguments all have true
conclusions

    But in a domain where bivalence may not be presumed,
as in any empirical domain, presumably including
linguistics, the claim

(3)   There is no longest English sentence

is not strictly equivalent to the claim

(4)   There are infinitely many English sentences.

For here the series of arguments constituting the
paradox may pass from true premisses to conclusions
which are neither true nor false, and from these to
conclusions which are false.

    We may if we wish coherently avoid identifying
sentences (3) and (4) on the analogy of the natural
equivalence between sentences (1) and (2)   We may even
assert, as I believe to be the case, that sentence (3)
is true and sentence (4) is false

BIBLIOGRAPHY

Bach, Emmon   1974   Syntactic theory   New York   Holt,
    Rinehart & Winston

Grinder, John T   and Suzette Haden Elgin. 1973   Guide
    to transformational grammar   New York   Holt,
    Rinehart & Winston

Olmstead, D L   1967   On some axioms about sentence
    length. Language 43 303-305

Quine, W V O   1961   From a logical point of view
        Cambridge   Harvard University Press