Necessity in mathematics II

Wilfrid Hodges
Herons Brook, Sticklepath, Okehampton
August-September 2009
wilfridhodges.co.uk/semantics11.pdf

Bernard Bolzano, 'Wissenschaftslehre' §1, 1837:

> Logic (*Wissenschaftslehre*) is 'the science which instructs us in the representation of the sciences in adequate textbooks' (*zweckmässigen Lehrbüchern*).

He adds: textbooks and not oral instruction, otherwise

> 'it would have to contain rules for discourse with children, congenitally blind or mute persons, etc.
> But investigations of this sort have so far never been conducted in logic. '

This is deep insight, and deeply wrong.
I predict that the next hundred years will show how wrong it is.
The rest of this talk is a footnote to this prediction.

All exposition, both oral and written, rests on the ability of the hearer/reader to understand what is said/written.

To understand what we read, we use many processes that are not in the province of logicians (though in fact logicians do sometimes discuss them).

I listed all the modal words in some pages of Birkhoff and Mac Lane, 'A Survey of Modern Algebra' and Hocking and Young 'Topology'.

In the first 50 pages of Birkhoff and Mac Lane there are 185 such words, i.e. 3.7 per page.

For Hocking and Young there are 226, i.e. 4.5 per page.

At first sight this is paradoxical.
All the mathematics in those pages is easily translatable into ZFC, and ZFC has no modal terms.

Two groups of examples can be explained straightforwardly.

(1) Interaction between writer and reader.

BM (9.23): Prove as many laws on the relation $a \leqslant b$ as you **can**.

HY (20.108): The word “compact” has been defined in so many (related) ways that one **must** be quite careful in reading the literature.

(2) Some modal words indicate what follows from what. So they correspond to the marginal numbers in a formal ZFC proof.

BM (98.335): To prove (L, U) a cut, one **need** only establish that U is the set of all upper bounds of L. [NB a $\diamond$ modality?]

HY (7.46): It follows that $\bigcup_p B_{\alpha(p)} \ldots$ **must** be contained in $\bigcap_{i=1}^n B_{\alpha_i}$.

$$*2{\cdot}14.\quad \vdash . \sim(\sim p) \supset p$$

Dem.

$$\left[\text{Perm}\,\frac{\sim\{\sim(\sim p)\}}{q}\right] \quad \vdash : p \lor \sim\{\sim(\sim p)\} . \supset . \sim\{\sim(\sim p)\} \lor p \qquad (1)$$

$$[(1).*2{\cdot}13.*1{\cdot}11] \quad \vdash . \sim\{\sim(\sim p)\} \lor p \qquad (2)$$

$$[(2).(*1{\cdot}01)] \quad \vdash . \sim(\sim p) \supset p$$

$$*2{\cdot}15.\quad \vdash : \sim p \supset q . \supset . \sim q \supset p$$

Dem.

$$\left[*2{\cdot}05\,\frac{\sim p, \sim(\sim q)}{p, \quad r}\right] \quad \vdash :. q \supset \sim(\sim q) . \supset : \sim p \supset q . \supset . \sim p \supset \sim(\sim q) \qquad (1)$$

$$\left[*2{\cdot}12\,\frac{q}{p}\right] \quad \vdash . q \supset \sim(\sim q) \qquad (2)$$

$$[(1).(2).*1{\cdot}11] \quad \vdash : \sim p \supset q . \supset . \sim p \supset \sim(\sim q) \qquad (3)$$

$$\left[*2{\cdot}03\,\frac{\sim p, \sim q}{p, \quad q}\right] \quad \vdash : \sim p \supset \sim(\sim q) . \supset . \sim q \supset \sim(\sim p) \qquad (4)$$

To minimise (1), (2) and some other minor variants, we restrict to occurrences of modal words in

> axioms, definitions, theorems and exercises.

This reduces to 55 instances in the first 100 pages of Birkhoff and Mac Lane, 13 in Hocking and Young.

Many of these really are paradoxical.
For example the student can't do the exercises without eliminating the modal content.

The paradoxical examples fall into clusters with some sporadic outsiders.

The ‘can be expressed’ cluster (BM 19, HY 1?)

BM Ex p. 20: Prove that any three integers a, b, c have a g. c. d. which **can** be expressed in the form $sa + tb + uc$.

HY Ex p. 22: Show that the plane set consisting of all points (x, y) satisfying $1 \leqslant x^2 + y^2 \leqslant 4$ **can** be given coordinates consisting of a point on the circle $x^2 + y^2 = 1$ and a point on the interval $[0, 1]$.

The 'can be embedded' cluster (BM 3 examples, HY 5)

BM Ex p. 43: **Can** the system J_0 of integers modulo 6 be embedded in a field?

HY Th p. 70: Every completely separable regular space **can** be imbedded in Hilbert coordinate space.

Interlude

The student reads the sentence and then understands what it means.
What is going on?

Classical theory:

The meaning of the sentence is a complex whole whose parts are the meanings of the words or phrases in the sentence.
Composing part meanings into a compound meaning corresponds to composing phrases by a grammatical construction.

Al-Fārābī (10th century Arabic):
[We] compose sentences of expressions signifying parts of the compound affair signified by the sentence. … the imitation of the composition of meanings by the composition of expressions is by [linguistic] convention.

Frege (letter to Jourdain):
Our ability to understand sentences that we have never heard before obviously rests on the fact that we build the sense of a sentence out of parts that correspond to the words.

Under the classical theory, the meaning of ‘can’ is *part* of the meaning of the whole exercise.

Since the exercise could have been written in ZFC, which uses no modal notions, it follows that these textbooks have *incorrectly* expressed their intended meaning.

Speaking to philosophers, I sometimes find that they draw this conclusion automatically.
My guess is that they are (consciously or unconsciously) subscribing to the classical theory.

The conclusion is absurd: these are two excellent textbooks, clear and reliable.
So the classical theory is wrong.

Myself I doubt that we can pinpoint a specific error in the classical theory.
It's just too crude to handle the phenomena of language.
The meanings of the words and the meanings of the grammatical constructions are of course the heart of the raw data that reaches the student's mind.
But this doesn't begin to describe the *processing* of this data — most of which is unconscious.

A better theory will allow the writer to express the same intended meaning in many different ways, with no 1-1 correlation between the meanings of the words.

Pulvermüller et al., ‘Functional links between motor and language systems’, *European Journal of Neuroscience* 21 (2005):

> Words related to actions involving different body parts, such as *pick* and *kick*, activate motor and premotor cortex in a somatotopic fashion … We show for the first time that stimulation of the motor system influences language processing in a category-specific manner, thereby proving an active role of cortical motor systems in word recognition.

In short, the processing of words like ‘embed’ and ‘express’ actively involves the brain circuitry for putting things into things and for speaking.

But those of us who aren't neuroscientists have to make do with more down-to-earth analyses.

We can ask for example:

- What difference would it make to replace the modal expression by a non-modal one?
- Is there a systematic translation that eliminates the modal word and would be understood by any competent English speaker?

We begin with the ‘can be expressed/written as’ cluster.

This usage is overwhelmingly mathematical (try Google). A few examples in linguistics and computer science feel as if they were mathematically influenced. Some others:

- the course of history **can be written** as a tug of war between rule-based and discretion-based methods of organizing and controlling
- Pinyin syllables **can be written** as one string (”zhongguo”) or as separated strings (”zhong guo”).

Nevertheless we never teach our students this usage. It seems to be naturally understood by English speakers.

Example (BM Ex p. 82:)
'Let D be the set of all rational numbers which **can be written as** fractions a/b with a denominator b relatively prime to 6.'

What difference would it make to replace 'can be written as' by 'are' or 'is'?

'Let D be the set of all rational numbers which are fractions a/b with a denominator b relatively prime to 6.'

But:

- Money can be written as the root of all evil.
- Some mushrooms can be expressed as poisonous.

‘Let D be the set of all rational numbers which are fractions a/b with a denominator b relatively prime to 6.’

I think what the paraphrase misses is the implied existential quantifier:

‘rational numbers r such that there are integers a, b with $a/b = r$ and b relatively prime to 6’.
If this is correct, then ‘can be expressed/written as’ is a device for introducing an existential quantifier *after* the subject.

BM Ex p. 20:
'In the Euclidean Algorithm, show by induction on k that each remainder **can be expressed in the form** $r_k = s_k a + t_k b$, where s_k and t_k are integers.'

Paraphrase:
'In the Euclidean Algorithm, show by induction on k that for each remainder there are integers s_k and t_k such that $r_k = s_k a + t_k b$.'

I think this works. But it leaves unexplained why this modal expression does this job.

Next the ‘can be embedded in’ cluster.

(1) HY Th p. 70:
‘Every completely separable regular space **can be imbedded in** Hilbert coordinate space.’

For brevity, ‘Every CSR space can be imbedded in HC space’.

A clear equivalent (and note the existential quantifier after the subject):

(2) ‘For every CSR space *S* there is an imbedding of *S* in HC space.’

Could (1) be paraphrased as follows?

(3) ‘Every completely separable regular space is imbedded in Hilbert coordinate space.’

In practice no. This would mean: the action of embedding each CSR space in HC space is performed.
(Possibly meaning we have in front of us such embeddings.)

This confirms the role of the existential quantifier.

But why ‘can be’ ?
Note first that (1) is not equivalent to:

(4) ‘For every CSR space X, it is possible that X is imbedded in HC space.’

(4) is mathematical nonsense: either X is imbedded in HC space or it isn’t.

Closer analysis shows that ‘can be imbedded in’ is the passive of ‘can imbed in’.
But the verb ‘embed’ is (so far as I know) never defined in mathematical texts.
What is defined is ‘embedding of A in B’.

Pattern:

an embedding of A into B
a mapping of A to B
a piercing of (surface) A by (line) B
a splitting of (group) A into B, C
a splitting of A by B

All are perceived as verbal nouns from action verbs
'embed', 'map', 'pierce', 'split'.
But the nouns are defined, never the verbs.

Compare these quotations (from the mathematical literature):

- ‘the interior of any simple closed curve can be mapped in an angle-preserving way to the open unit disk’
- ‘Each 2-sphere in each 3-manifold can be pierced by a tame arc’
- ‘Every division k-algebra D can be split by a finite Galois extension K/k’

Same pattern:

- ‘points on C can be injected into a proper linear subspace’
- ‘the triangles {11,3,6} and {11,6,1} can be retracted into the path (11,3,6,1)’

So the mathematical usage should be explained in terms of some general phenomena with action verbs and their nominalisations.

But this still doesn't explain why it's natural to use action verbs in order to express existential quantifiers.

Coda

Does all this mean that we need higher linguistics and neurophysiology in order to write 'adequate textbooks of the sciences'?
Obviously not.

- Just as the processes of understanding are inbuilt and largely unconscious,
  so also the processes of making ourselves understood are inbuilt and largely unconscious.
- For some of us who write textbooks,
  a sense of being in direct conversation with the reader is a vital ingredient of the craft.

But knowledge increases possibilities,
and (to repeat) I am aiming to write a footnote to knowledge that I predict is on its way.