Cemela Summer School
Mathematics as language
Fact or Metaphor?

John T. Baldwin

June 23, 2007
...I interpret that mathematics is a language in a particular way, namely as a metaphor."

David Pimm, Speaking Mathematically
Alternatively

Scientists, Quine said, put a straitjacket on natural languages

Moses, Radical Equations page 198
Formal languages arose to remedy the lack of precision in natural language.

1. Motivate with classroom examples the reasons for developing a formal language for mathematics.
2. Interweave the definition of a first order language adequate for mathematics.
3. The interplay between natural language, ‘regimented language’, and formal language
Outline

1. Framing the issues
2. structures and languages
3. variables
4. Defining Truth
5. equality
6. Moses
7. Diversions
History of Formal Languages

Foundations: 1850-1900
Boole, Pierce, Frege, Schroeder

Formulation: 1900-1935
Hilbert, Russell-Whitehead, Tarski, Gödel

Applications: post 1950
Linguistics: Chomsky, Montague, Barwise
Computer Science: McCarthy, Codd, Scott
contemporary symbolic logic
USA Today – Gallup Poll

A. Evolution, that is, the idea that human beings developed over millions of years from less advanced forms of life is

<table>
<thead>
<tr>
<th></th>
<th>Definitely true</th>
<th>Probably true</th>
<th>Probably false</th>
<th>Definitely false</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Evolution</td>
<td>18</td>
<td>35</td>
<td>16</td>
<td>28</td>
</tr>
</tbody>
</table>

B. Creationism, that is, the idea that God created human beings pretty much in their present form at one time within the last 10,000 years is

<table>
<thead>
<tr>
<th></th>
<th>Definitely true</th>
<th>Probably true</th>
<th>Probably false</th>
<th>Definitely false</th>
</tr>
</thead>
<tbody>
<tr>
<td>B. Creationism</td>
<td>39</td>
<td>27</td>
<td>16</td>
<td>15</td>
</tr>
</tbody>
</table>
How do you tell a sentence is true?

Compositional theory of truth

The truth of sentence $\phi$ is defined recursively from

1. the syntactic rule constructing $\phi$ from its components
2. the truth value of those components.
We adapt naive realism about mathematical objects.

We are concerned about how students form their concepts.

The ontological status of various concepts is a different subject.
After the last 150 years of research, logicians divide the world into:

- **Semantics**: mathematical objects
- **Syntax**: the language to describe them
Example: Arithmetic

What do the following statements say?
What are they about?

1. $III \oplus IV = VII$
2. $11 + 10 = 111$
3. $\cdots + \cdots = \_$
4. tres mas cuatro es igual a siete.
Natural numbers

\[ N = \langle a, aa, aaa, aaaa, aaaaa, \ldots \rangle \]

Let \( u \) and \( v \) be in \( N \) i.e. strings of \( a \)'s.

operations

\[ u +^N v = uv. \]

Relations

\( u <^N v \) if \( u \) is an initial segment of \( v \).

\( = \) is always interpreted as ‘identity’. 
A language for Arithmetic

operations

Two binary and two 0-ary function symbols (constants):
+ , × , 0 , 1

Relations

Two binary relations symbols:
= , <

Discussion: What are the expressions in this formal language that denote the natural numbers?
A language for Arithmetic

operations

Two binary and two 0-ary function symbols (constants):
+ , × , 0 , 1

Relations

Two binary relations symbols:
= , <

Discussion: What are the expressions in this formal language that denote the natural numbers?

How should we distinguish the symbol and the interpretation?
This distinction is essential for primary school teachers. They confuse the addition algorithm with addition.
Don’t fall into ‘numerals versus numbers’

Do Emphasize algorithms are means not ends.
Is this a correct statement of the Pythagorean Theorem?

\[ a^2 + b^2 = c^2 \]
What about?

\[ c^2 + a^2 = b^2 \]
What’s happening here?

What is the difference between these two equations?

\[ x^2 + 5x + 6 = 0 \]

\[ xy = yx \]
What’s happening here?

What is the difference between these two equations?

\[ x^2 + 5x + 6 = 0 \]

\[ xy = yx \]

\[ (\forall x)(\forall y)xy = yx \]
Free Variables

A variable that is not in the scope of a quantifier is **free**.

\[ x^2 + 5x + 6 = 0 \]

An expression with free variables is a **question**.

What elements can be substituted for the free variables and give a true statement?
Free Variables

A variable that is not in the scope of a quantifier is free.

\[ x^2 + 5x + 6 = 0 \]

An expression with free variables is a question.

What elements can be substituted for the free variables and give a true statement?

\[ x^2 = -1 \]
Bound Variables

If a variable is in the scope of a quantifier then it is bound. If all the variables in a statement are bound, we call it a (closed) sentence. It is either true or false (in a given structure).

$$(\forall x)(\exists y)y > x$$
The Angle Problem

The following statement is taken from a high school trigonometry text.

What does it mean?

sin $A = \sin B$ if and only if $A = B + 360K$ or $A + B = 180 + 360K$. 
The Angle Problem

The following statement is taken from a high school trigonometry text.

What does it mean?

\[
\sin A = \sin B \text{ if and only if } \\
A = B + 360K \text{ or } A + B = 180 + 360K.
\]

\[(\exists K)(A = B + 360K) \text{ or } (\exists K)A + B = 180 + 360K.\]
A vocabulary (or signature) $L$ is a collection of relation and function symbols.

A structure for that vocabulary ($L$-structure) is a set with an interpretation for each of those symbols.
Inductive definition of Language

terms

Every constant symbol is a term.
If \( f \) is an \( n \)-ary function and \( t_1, \ldots t_n \) are terms then
\( f(t_1, \ldots, t_n) \) is a term.
Nothing else is a term.

open sentences or well-formed -formulas

If \( R \) is an \( n \)-ary relation symbol and \( t_1, \ldots t_n \) are terms then
\( (t_1, \ldots, t_n) \) is a wff.
If \( \phi \) and \( \psi \) are wffs so are:

\[ \phi \land \psi, \neg \psi, (\exists x)\psi. \]

Nothing else is a wff.
Inductive definition of truth

Fix an $L$-structure $\mathcal{M} = (M, +, 0, 1, \cdot, =, <)$. Add to $L$ names for each element of $M$. Then each closed term $t$ (no free variables) denotes an element $t^M \in M$. $\mathcal{M} \models t = s$ iff $t^M = s^M$. $\mathcal{M} \models t < s$ iff $t^M < s^M$. For any $\phi, \psi$, $\mathcal{M} \models \phi \land \psi$ iff $\mathcal{M} \models \phi$ and $\mathcal{M} \models \psi$. $\mathcal{M} \models \neg \phi$ iff $\mathcal{M}$ does not model $\phi$. $\mathcal{M} \models (\exists x)\phi(x)$ iff for some $m \in M$, $\mathcal{M} \models \phi(m)$. 
No one actually does mathematics in the formal language. For logicians the fact that a certain concept can be expressed in a formal language is a powerful tool.
Turn to page 158-159 of Khisty-Chval.
On page 158, what decisions has the teacher made about ‘mathematical’ notions that need to be taught?

On page 159, why does Ms. Martinez expand on the students response?
Equality

What goes in the box?

\[ 8 + 4 = \square + 5 \]

What do 6th graders put in the box?
And yet, finally! I was so happy to learn that something from computer science could be reused in my Japanese class. Consider the following snippet of code in VBA:

```vba
if (a = b) then a = b end if
```

That’s right, the same symbol is used for testing equality and for variable assignment. The term is "overloaded operator" and it’s particular use can only be determined based on context. So yesterday’s Japanese class seemed obvious to me (for once) thanks to science.

http://thewayfaringstranger.blogspot.com/search/label/computer%20science
Glencoe: Properties of Equality

1. reflexive property: \( x = x \)
2. symmetric property: \( x = y \rightarrow y = x \)
3. transitive property:
   \[ (x = y \land y = z) \rightarrow x = z \]
4. Addition and subtraction properties:
   \[ (x = y \land y = z) \rightarrow x + z = y + z \text{ and } x - z = y - z \]
5. multiplication and division properties:
   \[ (x = y \land y = z) \rightarrow x \cdot z = y \cdot z \text{ and } x/z = y/z \]
6. substitution property: For all numbers \( a \) and \( b \) if \( a = b \), then \( a \) may be replaced by \( b \) in any equation or expression.
7. distributive property: \( x(y + z) = xy + xz \)
Leibnitz’s Law

For all numbers $a$ and $b$ if $a = b$ then $a$ may be replaced by $b$ in any equation or expression.
Leibnitz’s Law

For all numbers $a$ and $b$ if $a = b$ then $a$ may be replaced by $b$ in any equation or expression.

Formally

For any function symbol $f$ or formula $\phi$:

$$x_1 = y_1 \land \ldots \land x_n = y_n \rightarrow f(x_1, \ldots x_n) = f(x_1, \ldots x_n)$$

$$x_1 = y_1 \land \ldots \land x_n = y_n \rightarrow \phi(x_1, \ldots x_n) \leftrightarrow \phi(x_1, \ldots x_n)$$
Equality Axioms

\[ x = x \]

\[ x = y \rightarrow y = x \]

\[ (x = y \land y = z) \rightarrow x = z \]
Problem

I went to Pompeii and bought the same number of salads and small pizzas. Salads cost two dollars each and pizzas cost six dollars each. I spent $40 all together. Assume that the equation $2S + 6P = 40$ is correct.

What is wrong with the following reasoning? Be as detailed as possible. How would you try to help a student who made this mistake?
Solution

Then,

\[ 2S + 6P = 40. \]

Since \( S = P \), I can write

\[ 2P + 6P = 40. \]

So

\[ 8P = 40. \]

The last equation says 8 pizzas is equal to $40 so each pizza costs $5.
The Moses Analysis: General

1. people talk (natural language)
2. feature talk (regimented language)
3. math talk (formal language)
A Socratic Dialogue
People Talk

Which costs less, a pizza or a salad?
Which costs less, a pizza or a salad?

A salad costs less than a pizza
What feature of pizza and salad are we talking about?
What feature of pizza and salad are we talking about?

Their cost
Write a sentence describing this situation beginning

The cost of a salad
Write a sentence describing this situation beginning

The cost of a salad

is less than the cost of a pizza.
Where is the information about height encoded in

1. A salad costs less than a pizza.
2. The cost of a salad is less than the cost of a pizza.
People Talk vrs Feature Talk

Where is the information about height encoded in

1. A salad costs less than a pizza.
2. The cost of a salad is less than the cost of a pizza.

Let’s abbreviate 2):

C(S) is less than C(P).
How much?

How much less does a salad cost?
How much?

How much less does a salad cost?

$4 less
How much less does a salad cost?

$4 less

people talk: A salad costs $4 less than a pizza.

feature talk: $C(S)$ is $4 less than $C(P)$. 
Syntactical Observation

English adds the amount of change to the linking phrase. Feature talk puts it in a name position.
C(S) compared to C(P) is $4 less.

Abbreviate:
C(S) c/t C(P) is $4 less.
3 variants

\[ C(S) \ c/t \ C(P) \ \text{is } \$4 \text{ less.} \]
\[ C(S) \ c/t \ C(P) \ \text{is } -\$4. \]
\[ C(S) - C(P) = -\$4. \]
Positive numbers wed direction and quantity.

They are displacements.
Positive numbers wed direction and quantity.

They are displacements

Moses continues the analysis with trips on the MTA.

Disposable materials for students are located at http://www.algebra.org/ (programs/curriculum development/algebra)
As in the logical approach I outlined at the beginning:

vocabulary: relations $<$, $=,$ functions $+$, $-$

structure is $(\mathbb{Z}, +, -)$
Two metaphors

The **primary school metaphor** for subtraction is ‘take away’.

The **algebra metaphor** for subtraction is ‘compared to’.
Reflection

One should not justify ‘borrowing’ by:

Since 8 is greater than 7 we cannot subtract (take away) 8 from 7.
Reflection

One should not justify ‘borrowing’ by:

Since 8 is greater than 7 we cannot subtract (take away) 8 from 7.

Note the statement is true in 2nd grade because the term $8 - 7$ cannot be interpreted in $(\mathbb{N}, +, 1, -)$.

Many algebra books have silly statement about ‘closure’ under the operations. The authors didn’t realize that the closure condition was redundant when the intended interpretation is the real or rational numbers.
Moses’s example is about the height of the boys: Coast to Coast (CTC) and Watch Me.

What are three ways this example is different from mine?
Reprise

1. structures and languages
2. the compositional theory of truth
3. truth of a sentence is a structure
We have defined $M \models \phi$.
But what does it mean to say $\phi$ is true?!

Validity

The sentence $\phi$ is valid if it is true in every structure.
A proof system

### Logical Axioms

For any formula $\phi$: $\phi \lor \neg \phi$.

The equality axioms.

$\phi_x(a) \rightarrow (\exists x)\phi$.

### Inference rules

- Expansion: Infer $\phi \lor \psi$ from $\psi$
- Contraction: Infer $\psi$ from $\psi \lor \psi$
- Associative: Infer $(\phi \lor \psi) \lor \chi$ from $\phi \lor (\psi \lor \chi)$
- Cut: Infer $\phi \lor \psi$ from $\phi \lor \chi$ and $\chi \rightarrow \psi$
- Exists introduction: If $x$ is not free in $\phi$, infer $(\exists x)\psi \rightarrow \phi$ from $\psi \rightarrow \phi$. 
A formal proof from a set of axioms $\Phi$ is a sequence of wff’s such that each one

1. is a member of $\Phi$
2. or is a logical axiom
3. or follows from earlier lines by a rule of inference
The completeness theorem

**Gödel I**

There is a proof of $\psi$ if and only $\psi$ is valid.
There is a proof of $\psi$ from $\Phi$ if and only $\psi$ is true in every structure that satisfies each member of $\Phi$.

**Gödel II**

There is no effective way to decide whether a sentence $\phi$ is valid.
The inerrancy of mathematics

There is a procedure to check a proof is correct. There is no procedure to check if a sentence is valid. But the valid sentences are not interesting anyhow. To actually encode mathematics, add nonlogical axioms:
Some important sets of axioms

1. axioms for arithmetic
2. Axioms for the real field \((\mathbb{R}, +, \times, <, = 0, 1)\)
3. axioms for set theory
<table>
<thead>
<tr>
<th><strong>universe</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>0, 1, 2, 3, ...</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>operations</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>+, ×, 0, 1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Relations</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>=, &lt;</td>
</tr>
</tbody>
</table>
Patterns

Consider a circle with $n$ points on it. How many regions will the circle be divided into if each pair of points is connected by a chord?
Consider a circle with $n$ points on it. How many regions will the circle be divided into if each pair of points is connected by a chord?

1. Is this question well-formed? That is does the answer depend on the placement of the points.

2. Variant: What is the maximum number of regions of a circle that you can make by drawing chords between $n$ points on the circumference?

3. Guess the formula? Be very careful.
(a) The anti-realist often claims that the burden of proof is on the epistemological front and challenges the realist by asking: How is it that human beings are able to access information about this mysterious platonic universe of atemporal, acausal mathematical objects?
(b) The realist on the other hand, would often like to place the burden of proof on explaining the universality of mathematics. If mathematics is just a fiction, why is it so useful and universal in nature?
Both the realist and the anti-realist would like to make us believe that answering this question is such an impossible task that the only reasonable conclusion is to adopt their point of view.
Henrik Nordmark (on fom)
What is mathematics?

Quoting "Timothy Y. Chow" tchow at alum.mit.edu Sat, 03 Mar 2007:
Some years ago it occurred to me that a possible definition of mathematics is that anything that is *sufficiently precise* is mathematics.
The term "sufficiently precise" it itself not sufficiently precise to count as mathematical, but perhaps it is sufficiently precise to be a useful idea.
Thus mathematics, unlike most other fields of study, is characterized not so much by its *subject matter* as by a certain *threshold of precision*.
Sazonov
Problem

The Golden Eagle Ferry is 150 feet long. Cars are 20 feet long and Trucks are 30 feet long. Each lane holds the same number of cars and trucks. How many cars and trucks are in each lane?

Solution:

\[20C + 30T = 150.\]

Since \(C = T\), I can write

\[20C + 30C = 150.\]

So

\[50C = 150.\]

The last equation says 50 cars are 150 feet long, so each car is 3 feet long.