The commutative law

Peter J. Cameron
University of St Andrews

John Venn lecture
Hull, 8 March 2018

John Venn

Born 4 August 1834, Hull; died 4 April 1923, Cambridge
He worked on logic and probability, and is best remembered for the "Venn diagram". Later he turned to historical research on Cambridge University and in particular his college, Gonville and Caius. The picture shows the stained glass window commemorating him in Caius hall.

The commutative law

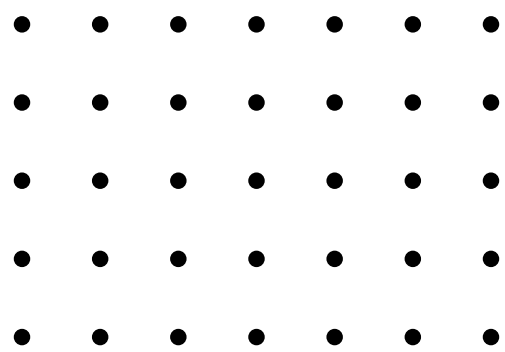
The commutative law states that

$$a \times b = b \times a$$

for all numbers a and b in some number system.
I am going to look at

- why this is true;
- why it matters;
- how and why it breaks down.

Why is $a \times b = b \times a$?



There are a rows with b dots in each row; and there are b columns with a dots in each column. So both $a \times b$ and $b \times a$ count the number of dots.

But . . .

Is that a valid proof?
For most of history, nobody would have had any doubts about that!
But suppose that a and b were each larger than the number of elementary particles in the observable universe. Would you still feel confident in this argument?
What if the geometry of space-time (either its curvature or its granular fine structure) made it impossible to build a very large rectangular array of dots, even in principle?
Should the truth of the commutative law depend on the structure of the universe?

So how do you prove it?

Today mathematics is based on systems of axioms, and proofs should proceed logically from the axioms.
There are two ways to prove the commutative law. One is to use axioms for the natural numbers, such as those formulated by Guiseppe Peano in 1889. The other is to use a more elaborate foundation, the axioms for set theory proposed by Zermelo and modified by Fraenkel in the early 20th century.
I will go through both arguments. Don't give up if you are not a mathematician: think of this as a glimpse of what we spend our time doing.

Proof by induction

All the familiar properties of the natural numbers can be proved from Peano's axioms. The essential tool is mathematical induction. I am leaving a lot out here ...

The product $a \times b$ is defined by $a \times 0 = 0$ and $a \times (b+1) = a \times b + a$. This is a definition by induction; the first statement defines $a \times b$ when $b = 0$, and the second shows how to go from any value b to the next value $b+1$.

We have to prove that $0 \times a = 0$ and $(b+1) \times a = b \times a + a$. Both of these require separate proofs by induction on a. For example, for the first we have $0 \times 0 = 0$ (which starts the induction at $a = 0$) and $0 \times (a+1) = 0 \times a + 0 = 0 + 0 = 0$ (using properties of addition).

Now, if $a \times b = b \times a$, then $(b+1) \times a = b \times a + a = a \times b + a = a \times (b+1)$, and we have won!

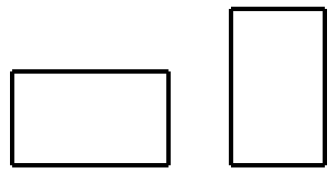
Set theory

The other approach is to use set theory as a basis, as most mathematicians do now. Though the axioms are more complicated, the proof is much simpler, and captures precisely the intuition behind the "dot diagrams".
We define the Cartesian product $A \times B$ of two sets A and B to be the set of all ordered pairs (x, y), for $x \in A$, $y \in B$. (Think of x and y as X- and Y-coordinates in the sense of Descartes.)
Now show that

- the number of elements in $A \times B$ is the product $a \times b$, where a and b are the numbers of elements in A and B respectively.
- There is a natural way of matching up $A \times B$ with $B \times A$ (just turn the ordered pairs around!), so these sets have the same numbers of elements.

Other number systems

We use the natural numbers for counting, and the real numbers for measuring.
The commutative law also holds in the real numbers.



Consider a rectangle with base a and height b. Its area is $a \times b$. Now rotate the rectangle through 90°. The new base is b and height a and the area hasn't changed. So $a \times b = b \times a$.
But there are hidden assumptions here too. If space is non-euclidean, we can't even draw an a by b rectangle. Also, we assume that we can rotate it by 90°; how do we know that?

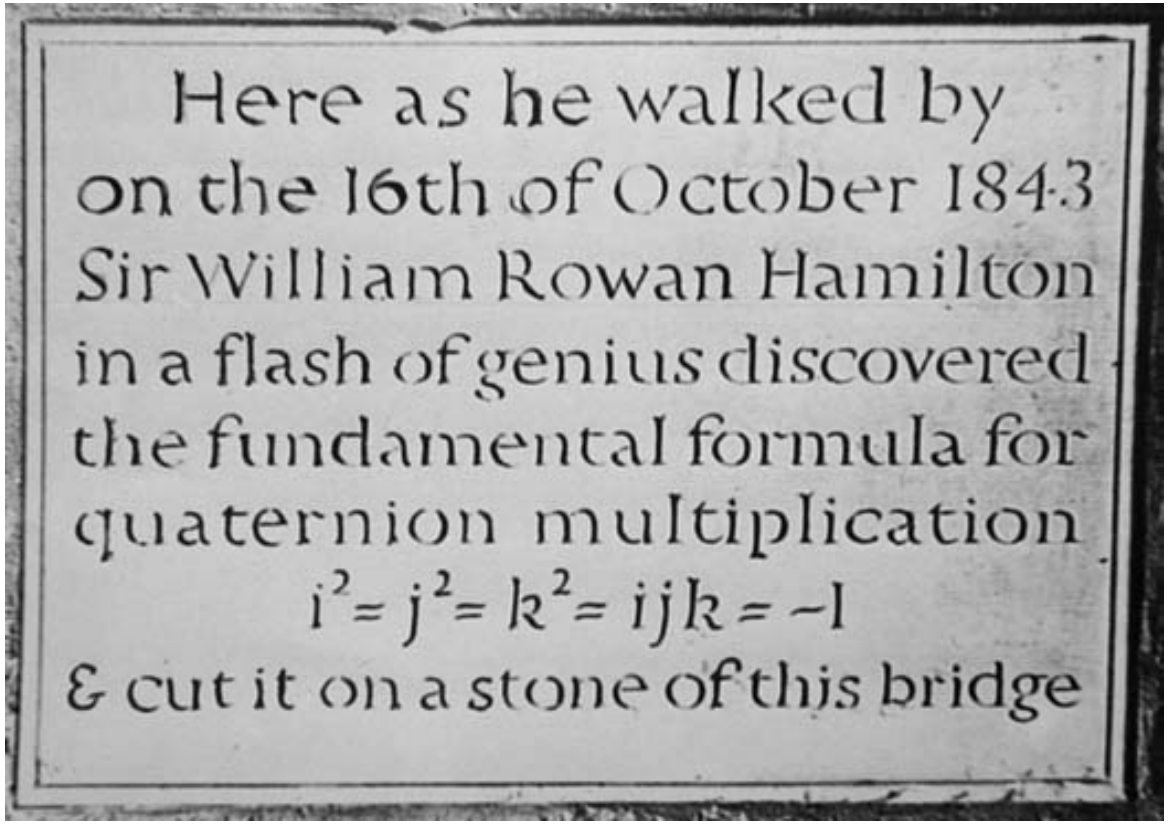
Algebra and physics

After a long struggle, mathematicians admitted the existence of "imaginary numbers" such as the square root of -1. These form a two-dimensional number system which became known as the "complex numbers" including the real numbers, and sharing most of the algebraic properties of the real numbers (including the commutative law). Indeed they have better properties: any polynomial equation over the complex numbers has a solution.
This fact is known as the Fundamental Theorem of Algebra.
The great mathematician Gauss is said to have given ten different proofs of this theorem.
The complex numbers are fundamental to physics, in particular to quantum mechanics, the theory of the very small.

In the 1840s, William Rowan Hamilton in Dublin was trying to extend the complex numbers to a three-dimensional number system, without success.
On 16 October 1843 (a Monday) Hamilton was walking in along the Royal Canal with his wife to preside at a Council meeting of the Royal Irish Academy. Although his wife talked to him now and again Hamilton hardly heard, for the discovery of the quaternions, the first noncommutative algebra to be studied, was taking shape in his mind:-

> And here there dawned on me the notion that we must admit, in some sense, a fourth dimension of space for the purpose of calculating with triples . . . An electric circuit seemed to close, and a spark flashed forth.

He had invented the quaternions, and was so pleased that . . .



It follows from these equations that $ij = k$ and $ji = -k$. So

$$i \times j \neq j \times i.$$

Hamilton had discovered the first non-commutative number system!

Some mathematicians found this hard to accept. According to Melanie Bayley, the Mad Hatter's Tea Party in Lewis Carroll's *Alice's Adventures in Wonderland* was a satire on Hamilton's quaternions.

Operations

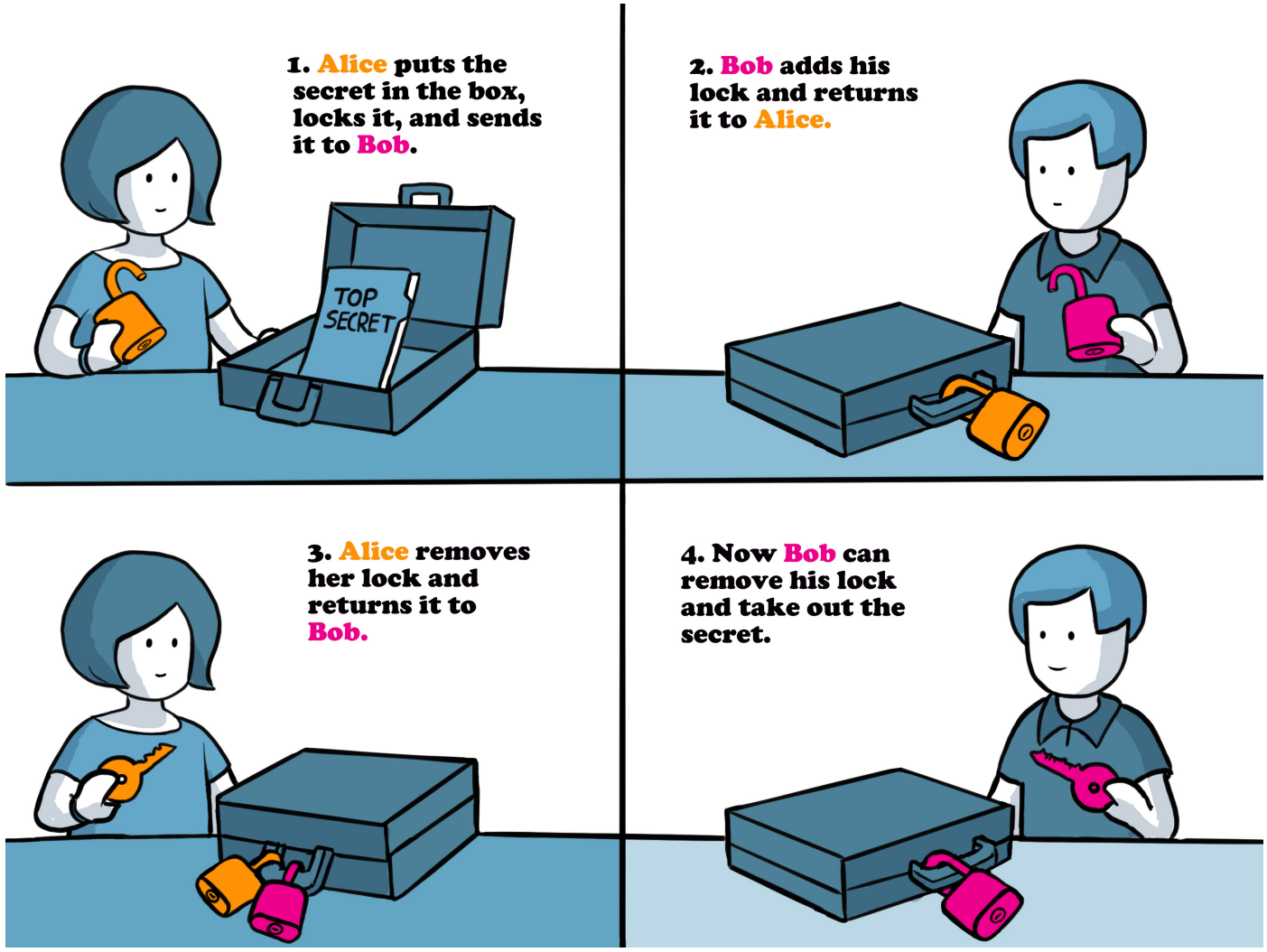
The four-dimensional quaternions were very well suited to describe the four-dimensional world of relativity when that came along half a century later.
But we don't actually use them, because in the meantime an even more flexible algebraic system came along: matrices.
The reason that matrices don't commute is because they represent operations, which are notoriously non-commutative. ("Multiplication" of operations means doing one and then the other.)
Let a be the operation of putting on your socks, and b the operation of putting on your shoes. Obviously a then b is quite different from b then a.

Multiplying and dividing

If you multiply a number x by a and then by b, then to recover x you should divide by b and then divide by a.
However, if multiplication is commutative, you can divide by a and then by b to recover x. If multiplication is not commutative, you can't!
When a and b are operations, then $a \times b$ means "do a, then b"; the example of shoes and socks shows that to undo the effect you have to first undo b and then a, and in general the other way around won't work. (Try taking off your socks before your shoes. It is possible, if your shoes are very loose, but not straightforward!)
We'll now see that this really matters.

Public-key cryptography

One of the great intellectual advances of the 20th century was the invention of public-key cryptography by Diffie and Hellman. I am going to describe to you a simpler precursor of this, Diffie–Hellman key exchange.
Suppose Alice has a secret file that she wants to send to Bob. She knows that her enemy Eve (the eavesdropper) has contacts in the mail service. How can she make sure that the file gets to Bob without Eve reading it first?



Picture by Neill Cameron

Why commutative?

At no time is the box sent through the post without a lock, so Eve is unable to open it and read the file.
It is crucial to this scheme that the operations "Alice locks the box" and "Bob locks the box" commute.
This would fail if, for example, Bob put the box inside a larger box and locked that, since then Alice could not remove her lock.
However, there is another problem as well when we transfer this scheme to electronics; commutativity is not enough.
We use "messages" and "numbers" interchangeably below: any message can be encoded into a big number, and *vice versa*.

Suppose for example that Alice and Bob choose secret keys a and b.

- Alice adds her key a to the message x, producing $x + a$, and sends to Bob.
- Bob adds his key b, and sends $x + a + b$ to Alice.
- Alice removes her key a, giving $(x + a + b) - a = x + b$, and sends to Bob.
- Now Bob can remove his key b and read the message x.

Although x is never transmitted unencoded, all Eve has to do is intercept the three messages $x + a$, $x + a + b$, and $x + b$ that are transmitted, and do a simple sum:

$$(x + a) + (x + b) - (x + a + b) = x.$$

So this is how to do it.

- Alice raises the message x to the power a, producing x^a, and sends to Bob.
- Bob raises this to the power b, and sends x^{ab} to Alice.
- Alice takes the ath root, giving $\sqrt[a]{x^{ab}} = x^b$, and sends to Bob.
- Now Bob can take the bth root and read the message x.

Now there is no straightforward way for Eve to obtain the message. (Do you remember how to extract a square root?)
But when Alice and Bob choose their keys, they do so in such a way that they are able to extract the appropriate roots.
So it is the difficulty of extracting roots without a key that keeps your data safe.

Logic and mathematics

> Reasoning and logic are to each other as health is to medicine, or – better – as conduct is to morality. Reasoning refers to a gamut of natural thought processes in the everyday world. Logic is how we ought to think if objective truth is our goal – and the everyday world is very little concerned with objective truth. Logic is the science of the justification of conclusions we have reached by natural reasoning. My point here is that, for such natural reasoning to occur, consciousness is not necessary. The very reason we need logic at all is because most reasoning is not conscious at all.

Julian Jaynes, *The Origin of Consciousness in the Breakdown of the Bicameral Mind*

Turning logic into algebra

While Hamilton was creating his quaternions, George Boole had the idea of turning logic into algebra.
For two thousand years, it was believed that Aristotle had said the last word on logic. This despite the fact that, starting in the middle ages, logicians (including William of Ockham and Leibniz) had made great advances.
Aristotle's logic was essentially restricted to arguments like this:

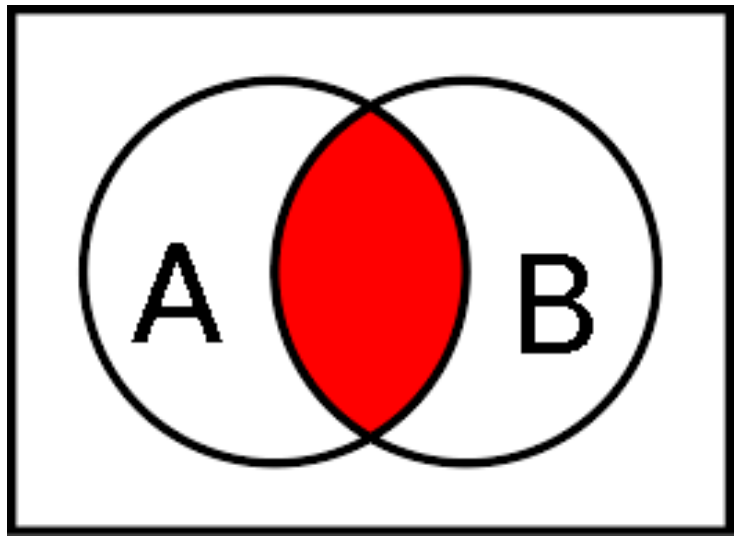
> All men are mortal; Socrates is a man; therefore Socrates is mortal.

A logic book that makes you laugh

In *The Thousand-Petalled Lotus*, Sangharakshita (an Englishman who became a Buddhist monk) relates that, shortly after his ordination as a sramanera (novice monk), he had an interlude at the Benares Hindu University, studying Abhidhamma (Buddhist scripture), Pali (the language in which it was written), and Logic with Bhikkhu Kashyap. He explains the enjoyment he got from Logic:

> Though I had read quite widely in philosophy, for some reason I had neglected [logic] … It was therefore with some trepidation that I set about making good the omission. But I need not have worried. Once I had emerged from the thickets of Formal Logic I found myself in one of the most fascinating stretches of the intellectual terrain … F. C. S. Schiller's *Formal Logic*, a radical empiricist's exposure of the aridities and absurdities of the subject, as traditionally expounded, was undoubtedly one of the most hilarious books I had encountered. While I was reading it there escaped me chuckles–even guffaws–which Kashyapji never heard when I was studying Pali.

In fact Schiller is attacking a straw man, since by his time (early 20th century) logic had awoken from its slumbers. Still, we have to plunge back into the thickets for a while.
Remember George Boole deciding to treat logic as a branch of algebra. If A and B are collections of things, Boole regarded $A + B$ as the collection of things lying in either A or B, and $A \cdot B$ as the collection of things in both A and B.
The easiest way to visualise this is to use John Venn's insight:



The red region is $A \cdot B$. It is clear that $A \cdot B = B \cdot A$.

William of Ockham and Augustus De Morgan

A little more notation. We let A' denote the contradictory opposite of A, so that A' consists of all the things not included in A. Then De Morgan's laws state:

- $(A \cdot B)' = A' + B'$,
- $(A + B)' = A' \cdot B'$.

Venn diagrams make this clear. But already in the 14th century, William of Ockham had stated:

- the contradictory opposite of a copulative proposition is a disjunctive proposition composed of the contradictory opposites of its parts.
- the contradictory opposite of a disjunctive proposition is a copulative proposition composed of the contradictories of the parts of the disjunctive proposition.

Exactly the same thing!

Good and bad

Using Venn diagrams, it is easy to prove things about the system just described. For example,

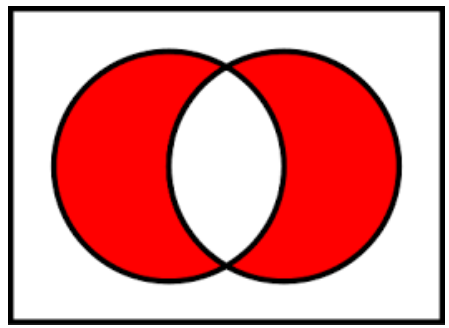
$$(A + B) \cdot C = (A \cdot C) + (B \cdot C),$$

which is the distributive law (the rule for expanding brackets), just what you would expect. But also

$$(A \cdot B) + C = (A + C) \cdot (B + C),$$

which is rather less natural, and indeed is somewhat disturbing!

It turns out to be better to use a different definition of addition: take $A + B$ to be the collection of things lying in either A or B *but not both*:



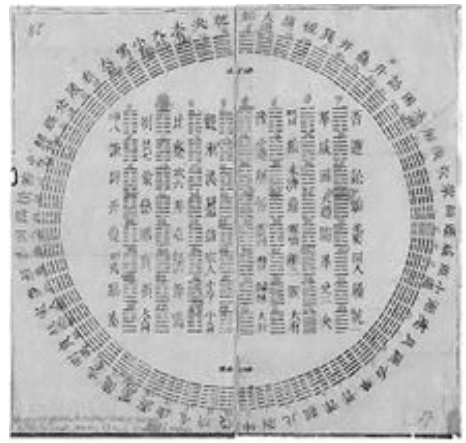
With this choice, the "usual rules of algebra" (commutative, associative and distributive laws, etc.) also apply. But there is still a surprise:

$$A + A = 0.$$

(Nothing can be in one of A and A without being in both!)

From the *I Ching* . . .

Of course, this is one of the laws of binary arithmetic. It is said that Leibniz, who invented the binary system, had been amazed to learn from Jesuit missionaries to China that the Chinese were already using this system.



If there are n things in the universe, then we can represent any collection A by an n-tuple with entries 0 and 1: the ith entry is 1 if the ith object belongs to A, and is 0 otherwise. Then the addition just descibed is just coordinatewise addition modulo 2. There are six bits, and so $2^6 = 64$ different situations can be described by their combinations.

. . . to the planets

This system was used to transmit information in the early days of space exploration. The Mariner probes to Mars used linear functions of five variables to encode information: six bits of data $(a_1, \ldots, a_5, b)$ would be encoded as the linear function $a_1x_1 + \cdots + a_5x_5 + b$. Any two such linear functions differ in at least 16 values, so if seven or fewer errors occur during transmission then the correct message can be deduced.

Modern error correction uses more sophisticated versions based on the same idea.

Is the universe non-commutative?

One of the contenders for a theory of quantum gravity is non-commutative geometry. This is an attempt to complete the 20th century revolution in physics along the following lines:

- General relativity is the theory of curved space-time; according to this, momentum coordinates in different directions don't commute. (If you draw a "parallelogram of vectors" on a curved surface, the ends will not join up.)
- Quantum mechanics says that position and momentum coordinates don't commute with one another; it regards these coordinates as operators, which bring in the non-commutativity.
- Non-commutative geometry would add the assertion that position coordinates in different directions don't commute with each other.

The jury is still out on this.

Quantum logic

Unlike in Venn's case, the logic of quantum mechanics is not commutative. This is because a system is affected in unknown ways when a measurement is made on it.
If we say, "The particle has position x and momentum p", we have carried out two measurements in order; the measurement of momentum means we lose knowledge of position, depending on how accurate our measurement of momentum is.
Thus, "The particle has momentum p and position x" is a completely different assertion.
What would Lewis Carroll have made of that, I wonder?

Quantum computers

A quantum computer with n quantum bits (or qubits) could in principle perform 2^n computations simultaneously (one for each combination of values of the qubits) and only output the result of the one successful calculation.
It is as if we toss the six yarrow sticks for the *I Ching* and all $2^6 = 64$ possibilities are simultaneously realised.
Nobody has built a nontrivial quantum computer yet, but perhaps the universe is a quantum computer inside which we live ...

Tweedledee's view

The last word goes to Tweedledee:

"Contrariwise," continued Tweedledee, "if it was so, it might be; and if it were so, it would be; but as it isn't, it ain't. That's logic."