MATH 195: Gödel, Escher, and Bach (Spring 2001)

Notes and Study Questions for Thursday, April 12

Reading: *READ THE TEXT BELOW*

We're going over the latter half of *Gödel and Mumon* for the third time, because ***it is really important!*** Think back on this semester, the formal systems, the records that couldn't be played, the obscure passages from the *Book of Daniel*... they all pointed to the few pages that we're struggling over now. This, in general terms, is the heart of the book, the heart of the semester: the existence of a sentence G that can refer to itself within TNT. Next week we'll look more closely into the nature of G.

The substance of these few pages is so important that we're trying multiple ways to get it into your heads. This time, Mike will play the role of Joshu, trying to illuminate his fellow sentient beings but in the end speaking in words and symbols. I will play the role of Mumon, interpreting his words, but relying on other words to do so. Bodhidharma came from India into China, some say, to help others gain enlightenment, but in the end:

> *Truly, words have no power.*
> *Even though the mountain becomes the sea,*
> *Words cannot open another's mind.*

So what hope is there? In the end, you must find your own way under your own power:

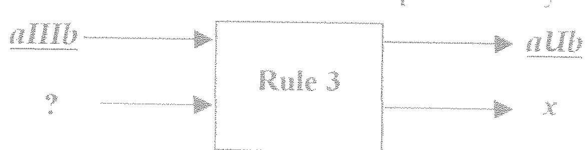
> *To find yourself on it, open yourself wide as the sky.*

We reached a point on Tuesday from which you can understand how what I'm about to write relates to pp.270-271. I'll start with an example from the MIU-system that should be easy to understand. What I want to accomplish is a good understanding of five different versions of the exact same problem. I'll begin by stating the general problem and a specification of this problem in English and symbols of MIU. The specification will be easiest to follow, but you'll need the general version too for what comes next.

20

SQ1. (1) True or False: The string x of the MIU-system is producible as an output from Rule 3 of the MIU-system.

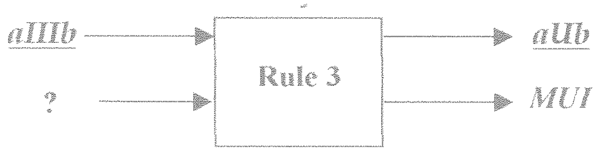
I need to recall, what IS Rule 3? It says: ***aIIIb*** → ***aUb***, that is, a U can replace III. So, what does this question say?



Could there be such a string x? Can you imagine a string that could be produced by Rule 3? Can you imagine a string that could not?

30 **SQ2. (1s) True or False: the string *MUI* is producible as an output from Rule 3 of the MIU-system.**

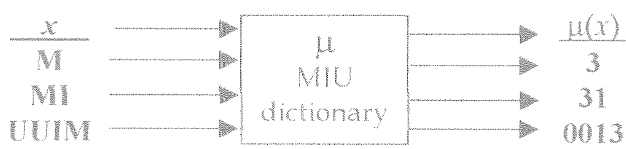
What's the difference between (1) and the specified version (1s)? It's certainly more tangible:



That's more like it. Can Rule 3 produce *MUI*?

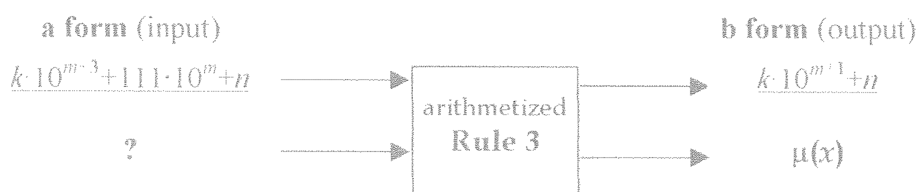
By means of the Gödel-numbering of the MIU-system, which I'll denote by μ (the greek letter "mu"), the same problem can be expressed in English as a problem of number theory. Recall that, at the level of symbols we have $\mu(M) = 3$, $\mu(I) = 1$, and $\mu(U) = 0$.

μ is another machine, one that contains a dictionary inside of it that can translate symbols into numbers:



60 **SQ3. (2) True or False: The number $\mu(x)$ can be written in the "b)-form" for the arithmetic rule that corresponds to Rule 3 of the MIU-system.**

"*The number* $\mu(x)$...". What does that mean? The string, x, goes into box μ and comes out a number. Looking at the results of the box, I see that $\mu(x)$ really is a number. What about "*b)-form*"? And what is "*...the arithmetic rule that corresponds to Rule 3...*"? Well, the rules of **MIU** were arithmetized on p.263, and Rule 3 became:



Can some number able to come out of the μ box be cast in the form of the output of arithmetized Rule 3? Don't know. Try some. Put some **MIU** strings into the μ box and see if you can figure out values of k, m, and n that make that number fit into the **b**-form. Maybe some numbers work, maybe others not?

30 **SQ4. (2s) True or False: The number $\mu(MUI) = 301$ can be written in the "b)-form" for the arithmetic rule that corresponds to Rule 3 of the MIU-system.**

> This says that if you put ***MUI*** through the μ box you get 301. Is that true? Is it also possible for 301 to come out of the arithmetized Rule 3 box?

Next, I'll re-express that problem of number theory using more mathematical-sounding lingo.

> Sounds ominous!

40 **SQ5. (3) True or False: The number $\mu(x)$ can be written in the form $k \cdot 10^m + n$, where k, m, and n are natural numbers and $n < 10^m$.**

> "*The number $\mu(x)$...*"... that's the output of putting a string x through the μ box. "*...can be written in the form $k*10^m + n$...*"... that's the output of the arithmetized Rule 3 box. "*...where k, m, and n are natural numbers and $n < 10^m$.*" Natural numbers, i.e. integers zero and up,... that's easy enough, but why does n have to be less than 10^m? Let's hold off a moment on that one.

30 **SQ6. (3s) True or False: The number 301 can be written in the form $k \cdot 10^m + n$, where k, m, and n are natural numbers and $n < 10^m$.**

> You translate this one yourself. Can 301 be written in that form? Try out values of k, m, and n:

$$\begin{array}{r} k \cdot 10^m \\ +\, n \\ \hline k \cdot 10^m + n \end{array} \qquad \begin{array}{r} \boxed{} \\ + \boxed{} \\ \hline 301 \end{array}$$

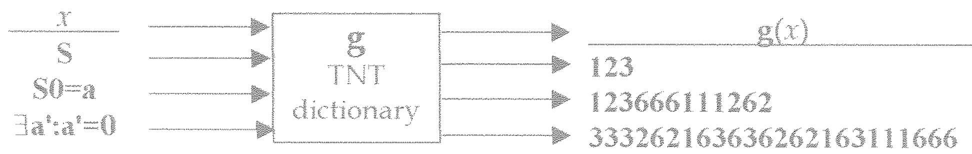
> I think, *k with some number of zeros after it plus n = 301. The upper box (k with zeros) could be 300 and the lower box (n) could be 1. That makes k=3, m=2, n=1. They're all natural numbers and n is less than 10^2.*

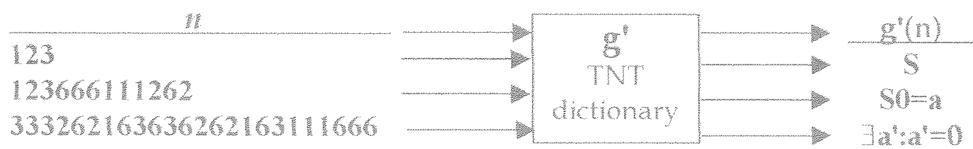
> Now back to $n < 10^m$, suppose that n isn't less than 10^m? What then? Is it possible to fit the form with a very large n? Well, if k=0, then m be anything it wants to be, and the upper box would be... what? What value would n then have? What mayhem results if we allow n to take large values like this? Try plugging in that value of n into the input form and translate it back into MIU language!

With mathematical-sounding lingo in place, an expression of the problem in TNT is not far behind. I'll use g to stand for the Gödel numbering of TNT given on page 268, with g^{-1} as the "inverse of g" taking numbers to symbols of TNT.

> Here's a new wrinkle. It's straightforward to imagine a box called g that has a TNT/Gödel numbering dictionary inside of it and that takes TNT strings x to numbers $g(x)$, just like what the μ box did for strings of MIU. But what about this g^{-1}? The box g, like the box μ goes in only one

direction. If we want to go in the other direction, we need a slightly different box, g':





50 **SQ7. (4) True or False: The string**

$\exists$k:$\exists$m:$\exists$n: < n is less than 10^m $\wedge$ $\mu(x)$ equals ($k \cdot 10^m$ + n) >

is a theorem of TNT.

Can you imagine any numbers k, m, and n such that n is less than 10^m (yes, there are lots like that, e.g. $3 < 10^5$) and $\mu(x)$ equals $(k \cdot 10^m + n)$? That sounds an awful lot like form (3), Study Question 6. If you take an MIU string and run it through box μ to get a number, can you cast that number in the form of the output of arithmetized Rule 3?

70 **SQ8. (4s) True or False: The string**

$\exists$k:$\exists$m:$\exists$n: < n is less than 10^m $\wedge$ 301 equals ($k * 10^m$ + n) >

is a theorem of TNT.

More specifically, can you cast 301 in the form of the output of arithmetized Rule 3? If you can, then the proposed pseudo TNT string is a theorem.

Note that I cheated a little bit in (4), mixing English with symbols of TNT by leaving my variables as numbers, using $\wedge$ for AND, etc. I ask you to imagine the result of actually spelling out the string fully in TNT. For (4s) it would be something like:

∃a:∃a′:∃a″: < ~∃a‴: a″ = () + Sa‴ ∧ S…299 S's…S0 = a·() +a″ >,

where I went ahead and started on "less than" but quit when I reached 10^m, using () to indicate my failure to deal with this expression.

Imagine now that we have the string of TNT described in version 4 of the original problem. Applying g symbol-by-symbol to that string, we arrive at another version of the problem that resides in number theory.

30 **SQ9. (5) True or False: The number**

g(∃a:∃a′:∃a″: < ~∃a‴: a″ = () + Sa‴ ∧ a⁗ = a·() +a″ >)

is a TNT-number.

(5) looks ALMOST identical to (4). What is *k* equivalent to? How about *m*, *n*, and $\mu(x)$? In fact, I could say rewrite sentence (5):

g(sentence 4) is a TNT number

Whoa! What's a TNT number? We seem to have two different definitions for it:

a. A TNT number is a number that is derivable from given numbers through the arithmetized rules of production of TNT.

but from the isomorphism between TNT strings and Gödel numbers:

b. A TNT number is the number you get when you pass a theorem of TNT through box g.

Which is true? Hearken back to Achilles confusion faced with a similar problem (see p.241, 5th paragraph from the bottom).

60 **SQ10. (5s) True or False: The number**

g(∃a:∃a′:∃a″: < ~∃a‴: a″ = () + Sa‴ ^ S…299 S′s…S0 = a·() +a″ >)

is a TNT-number.

Note (5) contains the free variable **a″″** (it isn't quantified) but that applying **g** to the expression does produce a pure number with no variables in sight at all, since **a** with a bunch of primes goes by g to an honest number of form 262 163 163 163 …

The important thing is that from one True/False problem of the MIU-system, we have managed to construct two True/False problems of number theory AND that an answer to any of these problems simultaneously answers all five of them.

One answer gives us all five answers? Is that right? Let's just take (5s). If it is true that numbers **a**, **a'**, and **a"** exist such that the stated conditions hold (what were those conditions?), then what does that tell us about the earlier true/false questions?

Interlude

Now I want to replace the simple MIU-system with TNT itself. I will follow the same line of reasoning developed above for the general problem, using *H* to represent a variable string of TNT.

60 **SQ11. (1) True or False: *H* is a theorem of TNT.** (general problem of TNT)

What is *H*? (a variable string of TNT) Make up some TNT strings. Let's see,… ∃=a<>00S… does it have to be well-formed? Let's say it does. How about **a** = **S0…** is that a theorem of TNT?

40 **SQ12. (2) True or False: The number g(*H*) can be written in the "b)-form" of some arithmetic rule that corresponds to a rule of TNT.**

(general problem of number theory)

> "The number g(*H*)..." means...? It's a number... obtained by sending a string of TNT called *H*... through box g,... which transforms strings into numbers. OK, can THAT number be written in the "b)-form of some arithmetic rule..."... "b)-form" means that it is the output of some arithmetic rule. Which rule? "...that corresponds to a rule of TNT." What ARE the rules of TNT?
>
> TNT incorporates all the rules of the propositional calculus and adds to them several of its own, mostly related to quantifiers. No one's explained how we can translate these typographical rules into arithmetic. Let's take perhaps the simplest rule:
>
> (DOUBLE-TILDE RULE part #1) *The string ~~ can be deleted from any theorem.*
>
> The arithmetized rule would somehow remove the Gödel translation of ~~ (223223) from within a number. Perhaps you can see how to do that, by analogy to Rule #4 of MIU. OK, so let's grant that the rules of TNT can be arithmetized.

Things get a little hazy now, because I'm not exactly sure what those "b)-forms" are ... I will rely on my understanding of the MIU example above and on the power of your imagination.

50 **SQ13. (3) True or False: The number g(*H*) can be written in the form (*whatever the form is*) for some values of the variables representing natural numbers that occur in the description of the form – i.e., for some natural numbers *k*, *m*, and *n*.**

(general problem of number theory)

> I suppose that he's imagining that the output form of the TNT rule we're talking about has variables *k*, *m*, and *n* within it. This question doesn't seem any different from (2), except that it's saying more explicitly what (3) only implied: that the variables within the output form exist.

50 40 50 **SQ14. (4) True or False: The string**

∃k:∃m:∃n: g(*H*) can be written in the desired form

is a theorem of TNT.

(general problem of TNT)

> This is also minimally different from the previous question, except that the statement is written more symbolically. Either way, the statement claims that values for the variables necessary to fit the number into an output form of a rule actually exist.

Note that the string referred to here is again of the form

∃a:∃a′:∃a″: TNT string expressing "g(*H*) equals number of desired form", 40 SQ 15

just as in the MIU example. Once again we're testing whether a string of TNT is a theorem of TNT. With this string in hand, Hofstadter would jump and refer to this version of the problem as a problem of meta-TNT. He is justified in doing so because the string of TNT here in version 4 "talks about" the original string *H* (also a string of TNT). So, as Hofstadter remarks on page 270, we see that "TNT contains strings that talk about other strings of TNT".

We have to go one step further.

SQ16. (5) True or False: The number g(the string of TNT in version 4) is a TNT-number.

(general problem of number theory)

Now the head begins to throb. Let's take this apart. "The number g( )..." ... That means the number that is produced by sending some TNT string through the TNT/number dictionary. What string are we talking about? "the string of TNT in version 4"... version 4 was:

∃k:∃m:∃n: g(*H*) can be written in the desired form

(presumably, the non-TNT symbols are translated into TNT and would appear like the arithmetized rules of MIU, only way more complicated.

And that number, resulting from sending the TNT string of version 4 through g, "... is a TNT-number". What is a TNT number? We've been here before. Most directly, it is a number that can be obtained by pushing one of the 5 given numbers of TNT (the axioms transformed through g) through the arithmetized rules of TNT. By isomorphism, it is a number that results by sending a **theorem** of TNT through g.

So put this all together...

1. *H* is a theorem of TNT [maybe it is, maybe not]

2. If it is, then it g(*H*) can be written as an output of some arithmetized rule of TNT.

3. If that's so, then there must exist appropriate numbers to make the output form of the rule work out for g(*H*)

1'.=4. If that's so, then the TNT statement (call it *H*')

∃k:∃m:∃n: g(*H*) can be written in the desired form

is also a theorem of TNT.

2'. Then g(*H*') is a TNT number

...

You see that the spiral has begun; an original problem of TNT became a problem of number theory which in turn became a problem of TNT which in turn became a different problem of number theory ...

As Hofstadter puts it, Gödel's genius was to find a way to close up the ends of this spiral, producing a two-step strange loop that manages to import the Liar's Paradox into TNT. The string of TNT that accomplishes this feat is the Gödel string G. More about G tomorrow (I hope).

60 **SQ17. Diagram the procedures that take the TNT string *H* to the TNT string *H'*, in the form of a Recursive Transition Network.**