

Here are a couple of pure postulational systems from Douglas Hofstadter's *Gödel, Escher, Bach: An Eternal Golden Braid*.<sup>4</sup>

### Hofstadter's MIU System

Hofstadter created the **MIU** system to illustrate the idea of postulational systems.

The only formation rule of the **MIU** system is: a WFF is any string containing only the letters **M**, **I**, and **U**. There are four transformation rules, and all of them are "one-way" rules (like implicational rules). The rules are:

**Rule I:** Given a WFF whose last letter is **I**, you can add a **U** at the end. For example, from **MMI** you can get **MMIU**.

**Rule II:** Given a WFF of the form **Mx**,<sup>5</sup> you may repeat **x** to get **Mxx**. From **MIU** you can get **MIUIU**, from **MUM** you get **MUMUM** and from **MU** you can get **MUU**, etc.

**Rule III:** If **III** occurs in a WFF, you can replace the **III** with **U**. So, from **UMIIMU** you get **UMUMU** and from **MIIII** you get **MIU** or **MUI**. From **MIII** you can make **MU**.

Note: these are "one-way" rules, so you cannot use Rule III to get **MIII** from **MU**.

**Rule IV:** If **UU** occurs in a WFF, you can drop the **UU**. So from **UUU** you can get **U**. From **MUUIII** you can get **MIII**.

There is one postulate (initial premise or axiom) in this system. It is **MI**. Hofstadter sets his readers a problem that he calls "the **MU**-puzzle." Starting from that one postulate as a given WFF, can you derive **MU**?

Here's an example of a derivation in the **MIU** system. It resembles our propositional logic derivations, except that it doesn't mean anything.<sup>6</sup> The goal is to derive the theorem<sup>7</sup> **MUIIU**.

(1)	<b>MI</b>	Axiom
(2)	<b>MII</b>	1, Rule II
(3)	<b>MIIII</b>	2, Rule II
(4)	<b>MIIIIU</b>	3, Rule I
(5)	<b>MUIU</b>	4, Rule III
(6)	<b>MUIUUIU</b>	5, Rule II
(7)	<b>MUIIU</b>	6, Rule IV

<sup>4</sup> Douglas R. Hofstadter, *Gödel, Escher, Bach: An Eternal Golden Braid (A Metaphorical Fugue on Minds and Machines in the Spirit of Lewis Carroll)* (New York: Basic Books, 1979). This is an amazing, wonderful book. Gödel was a mathematician; Escher was a graphic artist/printmaker/painter; Bach was Johann Sebastian Bach. The book is about mathematics, graphic art, music, philosophy, minds and brains, meanings, computers and artificial intelligence, and lots more.

<sup>5</sup> **x** here stands for any of **M**, **I**, or **U** or any string of them. It works something like our **p**, **q**, and **r** forms. That is, **Mx** is not a WFF – it is the form of a WFF. **Mx** is any WFF whose first letter is **M**.

<sup>6</sup> A former student asked (about this sentence) "Then what's the point?" The point emerges in the section on "Interpretation of Formal Systems," below.

<sup>7</sup> Any WFF that can be derived just from the basic axiom(s) or postulate(s) of a system is a theorem of that system. Remember this definition!!

### Exercise on the MIU System

Play with the **MIU** system. Derive some theorems. Try to derive **MU**. If you cannot derive **MU**, try to prove that it is impossible. The proof of impossibility cannot be done within the **MIU** system; you have to "step outside the system" and do your reasoning in traditional logic or math.

### Hofstadter's pq- System

WFFs in the **pq-** system contain just three kinds of symbols: **p**, **q**, and **-** (the hyphen). There are infinitely many axioms. Hofstadter defines an *axiom-schema*, which is the form of an axiom. Every WFF of that form is an axiom of the system.

**Definition:** **xp-qx-** is an axiom, where **x** consists only of hyphens. Note that **x** is not part of a WFF, but is used to stand for a string of hyphens. The same string of hyphens (i.e., the same number of hyphens in the string) replaces both **xs**.

From this definition, we can see that **-p-q--** is an axiom (substituting a single hyphen for **x**). So are **--p-q---** and **---p-q----** and **----p-q-----**. **Any axiom of the system is also a theorem of the system.**

There is only one **transformation rule** (Hofstadter calls it a "rule of production" or "production rule") in this system.

**Rule:** Suppose **x**, **y**, and **z** stand for particular strings of hyphens. Suppose that **xpyqz** is a theorem. Then **xpy-qz-** is a theorem. If **x** is '--' and **y** is '---' and **z** is '-', then the rule tells us: If **--p---q-** is a theorem (is it?), then **--p----q--** is a theorem.

What about **formation rules**? A derivation must start with an axiom or a theorem. If it begins with a theorem, that theorem will have to have been derived from an axiom or a theorem. Ultimately we arrive at premises that are axioms only. No axiom contains more than one or less than one **p** or **q**. Our transformation rule does not allow us to add a **p** or a **q**. So every WFF must contain exactly one **p** and exactly one **q**. By similar argument, since the transformation rule never permits derivation of a WFF that has fewer hyphens than the WFF it's derived from, and the simplest axiom is **-p-q--**, so every WFF in a derivation will begin with a hyphen, and the **p** and **q** will be separated by at least one hyphen, and the WFF will end with a hyphen. The axiom schema and the transformation rule determine the only allowable formulas in derivations. We don't need explicit formation rules.

### Exercise on the pq- System

Play around with the system. Get an interesting axiom using the schema and use the transformation rule to derive theorems.

### Interpretation of Formal Systems

Now we get to the point of all this. **MI** is an axiom (postulate) of the **MIU** system. Is it true? Is it false? The theorems of the system are derived according to the

transformation rules from the axiom **MI**. In a sense, then, they are conclusions of valid arguments using **MI** as a premise. Are they true or false? **M** and **I** and **U** were never defined. We don't know what these symbols denote or connote. They are undefined terms or **primitives** of the system.

Are the axioms of the **pq**- system true or false? We don't know what **p** or **q** or **-** stand for, so we don't know what (if anything) is stated by "**---p-q----**." The theorems can be validly (i.e., according to the transformation rules of the system) derived from the axioms, but are the axioms true or false?

WFFs in these systems don't make statements until we **interpret** the system.

One system is considered to be an **interpretation** of another if and only if one can map the first system into the other in such a way that each part<sup>8</sup> of the first system corresponds to one part of the other, where "correspond" means that the two parts play similar roles in their respective systems. When such a mapping exists, we say that one system is an **interpretation** of the other.

Hofstadter proposes that we consider the isomorphism or interpretation given by the mapping: **p** ↔ plus; **q** ↔ equals; **-** ↔ one; **--** ↔ two; **---** ↔ three; etc. On this interpretation, **---p-q----** maps to "three plus one equals four." It and all of the axioms and theorems are **true on this interpretation**.

Another interpretation is: **p** ↔ horse; **q** ↔ happy; **-** ↔ apple. Then **-p-q--** translates as "apple horse apple happy apple apple." On this interpretation, axioms and theorems are no more true than non-axioms and non-theorems. Hofstadter says, "A horse might enjoy 'happy happy happy apple horse' (mapped onto **qqq-p**) just as much as any interpreted theorem."

A problem with interpretation is that one might "read too much into" the original system, based on knowledge of the system to which it is mapped. For example, on the "plus, equals, one, two, ..." interpretation, someone might think that **--p--p--q-----** should be a theorem because "two plus two plus two equals six" is true. That would be a mistake. **--p--p--q-----** is not a theorem. It's not even a WFF in the system.

Is "plus, equals, one, two, ..." the interpretation of the **pq**- system? Hofstadter gives the mapping: **p** ↔ equals; **q** ↔ taken from; **-** ↔ one; **--** ↔ two; etc. On this interpretation, **--p--q-----** means "two equals three taken from five." This is another interpretation. Again, all the axioms and every theorem are true on this interpretation. Which is the real meaning of the string? Is there even any sense in asking about the real meaning?

---

<sup>8</sup> "Part" includes the objects that are components of the two systems as well as the relations between and operations on those objects.