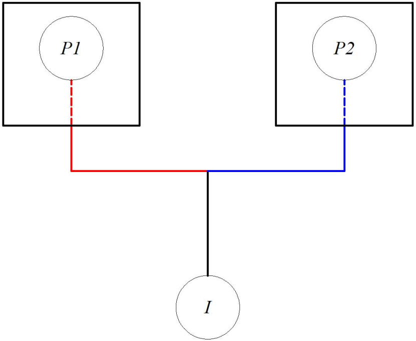
# Gödel, Turing, the Undecidability Results and the Nature of Human Mind

Riccardo Bruni
riccardobruni@hotmail.com

Dipartimento di Filosofia
Università degli Studi di Firenze
(Italy)

*Computability in Europe '06*
Swansea (UK), 29th June - 5th July 2006

## §1 *Turing and the "Mathematical Objection"*




Turing's Imitation Game (1950) – aka Turing's Test

The "objection" tackled by Turing goes as follows:

The "objection" tackled by Turing goes as follows:

1. Given any consistent [Universal Turing] machine M, we know that there are (codes of) questions such that M is uncapable to give an answer to.

The "objection" tackled by Turing goes as follows:

1. Given any consistent [Universal Turing] machine M, we know that there are (codes of) questions such that M is uncapable to give an answer to.
2. On the contrary, any human player, whose activity is not subdued to a given set of instructions, would go freely in search of new tecniques, making blunders and eventually finding an answer to these questions at last.

The "objection" tackled by Turing goes as follows:

1. Given any consistent [Universal Turing] machine M, we know that there are (codes of) questions such that M is uncapable to give an answer to.
2. On the contrary, any human player, whose activity is not subdued to a given set of instructions, would go freely in search of new tecniques, making blunders and eventually finding an answer to these questions at last.
3. That would suffice for the Interrogator to make a distinction between the two players.

The "objection" tackled by Turing goes as follows:

1. Given any consistent [Universal Turing] machine M, we know that there are (codes of) questions such that M is uncapable to give an answer to.
2. On the contrary, any human player, whose activity is not subdued to a given set of instructions, would go freely in search of new tecniques, making blunders and eventually finding an answer to these questions at last.
3. That would suffice for the Interrogator to make a distinction between the two players.
4. Then, no machine can satisfactorily take the role of a human in the Imitation Game [which, according to Turing's own conventions, *means* that no machine can show an intelligent behaviour].

Turing's answer was two-fold:

Turing's answer was two-fold:

- Consistency is not a requirement for Intelligence

Turing's answer was two-fold:

- Consistency is not a requirement for Intelligence

**Lecture to the London Mathematical Society, 20 February 1947**

"... *[I]f a machine is expected to be infallible, it cannot be also intelligent. There are several mathematical theorems which say almost exactly that. But these theorems say nothing about how much intelligence may be displayed if a machine makes no pretence at infallibility*."

Turing's answer was two-fold:

- Consistency is not a requirement for Intelligence

**Lecture to the London Mathematical Society, 20 February 1947**

*"... [I]f a machine is expected to be infallible, it cannot be also intelligent. There are several mathematical theorems which say almost exactly that. But these theorems say nothing about how much intelligence may be displayed if a machine makes no pretence at infallibility."*

- In order to perform tasks requiring intelligence, a new sort of machine is needed.

Turing's answer was two-fold:

- Consistency is not a requirement for Intelligence

**Lecture to the London Mathematical Society, 20 February 1947**

*"... [I]f a machine is expected to be infallible, it cannot be also intelligent. There are several mathematical theorems which say almost exactly that. But these theorems say nothing about how much intelligence may be displayed if a machine makes no pretence at infallibility."*

- In order to perform tasks requiring intelligence, a new sort of machine is needed.

**Lecture to the London Mathematical Society, 20 February 1947**

*"A human mathematician has always undergone an extensive training ... One must therefore not expect a machine to do a very great deal of building up of instructions table on its own ... [T]he machine must be allowed to have contact with human beings in order that it may adapt itself to their standards."*

# Turing's example: the Reward & Punishment machine (1948)

Main features:

# Turing's example: the Reward & Punishment machine (1948)

Main features:

- the machine has a suitable communication device, through which it is in contact with an external educator;

# Turing's example: the Reward & Punishment machine (1948)

Main features:

- the machine has a suitable communication device, through which it is in contact with an external educator;
- instructions are divided into:

# Turing's example: the Reward & Punishment machine (1948)

Main features:

- the machine has a suitable communication device, through which it is in contact with an external educator;
- instructions are divided into:
  - program istructions: fixed in advance and unmodifiable;

# Turing's example: the Reward & Punishment machine (1948)

Main features:

- the machine has a suitable communication device, through which it is in contact with an external educator;
- instructions are divided into:
  - program istructions: fixed in advance and unmodifiable;
  - character instructions: ending in a peculiar, 'listening' state $\mathbf{q_a}$;

# Turing's example: the Reward & Punishment machine (1948)

Main features:

- the machine has a suitable communication device, through which it is in contact with an external educator;
- instructions are divided into:
  - program istructions: fixed in advance and unmodifiable;
  - character instructions: ending in a peculiar, 'listening' state $\mathbf{q_a}$;
- when the machine is in state $\mathbf{q_a}$, the educator is allowed to either 'reward', or to 'punish' it by means of suitable stimula, causing the instruction just performed either to be adopted definitively, or to be substituted by an alternative (to be chosen by the machine itself among a number of them provided in advance).

The crucial conceptual issue underlying Turing's approach is the following:

The crucial conceptual issue underlying Turing's approach is the following:

Can computing agents of the new type have a computational power outperforming that of Turing machines
(namely, can they go beyong the Church-Turing Thesis)?

The crucial conceptual issue underlying Turing's approach is the following:

Can computing agents of the new type have a computational power outperforming that of Turing machines
(namely, can they go beyong the Church-Turing Thesis)?

This seems to be presently at the attention of those pursuing directions of research of which Turing's seems an anticipation (i.e., the study of Cognitive and Intelligent Systems by Wiedermann and van Leeuwen).

# Is mind mechanical? (1)
## Turing's philosophical analysis

# Is mind mechanical? (1)

## Turing's philosophical analysis (sort of)

# Is mind mechanical? (1)

## Turing's philosophical analysis (sort of)

### A. *Turing on mathematical reasoning: 1939*

Mathematical knowledge is based on both

# Is mind mechanical? (1)
## Turing's philosophical analysis (sort of)

### A. *Turing on mathematical reasoning: 1939*

Mathematical knowledge is based on both

- intuition, the ability of making spontaneous judgements

# Is mind mechanical? (1)

## Turing's philosophical analysis (sort of)

### A. *Turing on mathematical reasoning: 1939*

Mathematical knowledge is based on both

- intuition, the ability of making spontaneous judgements
- and ingenuity, the faculty of building rigorous arguments out of them.

# Is mind mechanical? (1)

## Turing's philosophical analysis (sort of)

### A. *Turing on mathematical reasoning: 1939*

Mathematical knowledge is based on both

- intuition, the ability of making spontaneous judgements
- and ingenuity, the faculty of building rigorous arguments out of them.

Metamathematical investigations (feat. Gödel's theorem) show that not both can be reduced to a mere search of the next most effective step in a derivation according to a finite set of given rules (that is: there's no way to eliminate intuition, and leave only ingenuity).

# Is mind mechanical? (1)

## Turing's philosophical analysis (sort of)

**B.** *Turing on intelligence: 1948*

Intelligent behaviour requires both

# Is mind mechanical? (1)
## Turing's philosophical analysis (sort of)

**B.** *Turing on intelligence: 1948*

Intelligent behaviour requires both

- discipline, the ability of obeying orders

# Is mind mechanical? (1)

## Turing's philosophical analysis (sort of)

### B. *Turing on intelligence: 1948*

Intelligent behaviour requires both

- discipline, the ability of obeying orders
- an additional ingredient called initiative

# Is mind mechanical? (1)

## Turing's philosophical analysis (sort of)

### B. *Turing on intelligence: 1948*

Intelligent behaviour requires both

- discipline, the ability of obeying orders
- an additional ingredient called initiative

"*To convert a brain or machine into a universal machine is the extremest form of discipline. ... But discipline is certainly not enough in itself to produce intelligence. ... Our task is to discover the nature of this residue as it occurs in man, and to try and copy it on machines.*"

# Is mind mechanical? (1)
## Turing's philosophical analysis (sort of)

**C.** *Turing's prediction on intelligent machinery: 1950*

*"I believe that in about fifty years' time it will be possible to programme computers ... to make them play the imitation game so well that an average interrogator will not have more than 70 per cent. chance of making the right identification after five minutes of questioning."*

## §2 *Gödel's approach*

# Gödel's disjunction - 1951

# Gödel's disjunction - 1951

1. Given any suitable and consistent formal system of axioms F, Gödel's so-called second incompleteness theorem in its hypothetical form yields

$$CORR_F \Rightarrow \nvdash_F Con_F$$

where $CORR_F$ and $Con_F$ are (some form or another of) F-validity, and the statement representing F-consistency resp.

# Gödel's disjunction - 1951

1. Given any suitable and consistent formal system of axioms F, Gödel's so-called second incompleteness theorem in its hypothetical form yields

$$CORR_F \Rightarrow \nvdash_F Con_F$$

where $CORR_F$ and $Con_F$ are (some form or another of) F-validity, and the statement representing F-consistency resp.

2. In other words, if any of these systems is known to be correct with mathematical certitude, then it is known not to contain all of "*subjective mathematics*" (with the obvious meaning that there exists a statement which is mathematically known to be true, but also unprovable in F).

# Gödel's disjunction - 1951

1. Given any suitable and consistent formal system of axioms F, Gödel's so-called second incompleteness theorem in its hypothetical form yields

$$CORR_F \Rightarrow \nvdash_F Con_F$$

where $CORR_F$ and $Con_F$ are (some form or another of) F-validity, and the statement representing F-consistency resp.

2. In other words, if any of these systems is known to be correct with mathematical certitude, then it is known not to contain all of "*subjective mathematics*" (with the obvious meaning that there exists a statement which is mathematically known to be true, but also unprovable in F).

3. Conversely, if a given system of that sort contains all of subjective mathematics, then it cannot be known to be correct with mathematical certitude.

That is,

That is,

### *Gödel's disjunction*

Either subjective mathematics cannot be comprised in any finite rule, equivalently the human mind infinitely surpasses the power of any Turing machine in the domain of pure mathematics,

That is,

*Gödel's disjunction*

Either subjective mathematics cannot be comprised in any finite rule, equivalently the human mind infinitely surpasses the power of any Turing machine in the domain of pure mathematics,

or

[the mind is equivalent to some finite machine and] there exist absolutely unsolvable mathematical problems of a specified logical type

That is,

### *Gödel's disjunction*

Either subjective mathematics cannot be comprised in any finite rule, equivalently the human mind infinitely surpasses the power of any Turing machine in the domain of pure mathematics,

or

[the mind is equivalent to some finite machine and] there exist absolutely unsolvable mathematical problems of a specified logical type
(where that both the human mind cannot be reduced to a finite rule and that there exist absolutely unsolvable problems of a different logical complexity, is a further alternative as well).

That is,

*Gödel's disjunction*

Either subjective mathematics cannot be comprised in any finite rule, equivalently the human mind infinitely surpasses the power of any Turing machine in the domain of pure mathematics,

or

[the mind is equivalent to some finite machine and] there exist absolutely unsolvable mathematical problems of a specified logical type
(where that both the human mind cannot be reduced to a finite rule and that there exist absolutely unsolvable problems of a different logical complexity, is a further alternative as well).

## $P = NP$ as a solution to the disjunction

Gödel to von Neumann (letter dated March 1956):

## $P = NP$ as a solution to the disjunction

Gödel to von Neumann (letter dated March 1956):

- Take the problem: 'Is a given formula $A$ of first-order predicate logic provable in $n$ steps ($n \in \mathbb{N}$)?'.

# $P = NP$ as a solution to the disjunction

Gödel to von Neumann (letter dated March 1956):

- Take the problem: 'Is a given formula $A$ of first-order predicate logic provable in $n$ steps ($n \in \mathbb{N}$)?'.
- A linear or polynomial (quadratic) solution of it in the input $n$, seems "quite within the realm of possibilities".

# $P = NP$ as a solution to the disjunction

Gödel to von Neumann (letter dated March 1956):

- Take the problem: 'Is a given formula $A$ of first-order predicate logic provable in $n$ steps ($n \in \mathbb{N}$)?'.
- A linear or polynomial (quadratic) solution of it in the input $n$, seems "quite within the realm of possibilities".
- If that were really the case, and $n$ is taken to be so large that the restriction to proofs of lenght $n$ can be regarded as irrelevant for any practical purpose, then a human mathematician dealing with questions with an yes-or-no answer could be really substituted by a machine, despite the recursive unsolvability of the general decision problem.

# Is mind mechanical? (2)
## Gödel's "naturalized" Platonism

**A.** *Against constructivism – 1951*

The existence of absolutely unsolvable mathematical problems "*seems to disprove the view that mathematics is only our own creation, for the creator necessarily knows all properties of his creatures because they can't have any other except those he has given to them. So this alternative seems to imply ... some form or other of Platonism or 'realism' as to mathematical objects*".

# Is mind mechanical? (2)
## Gödel's "naturalized" Platonism

**B.** *Knowing the world of mathematical entities – 1951*

The very same nature of mathematical statetment (which are "*true owing to the (nature of the) concepts occurring in it*"), makes it possible for them to be undecidable.

But then our knowledge of the world of concepts would be "*as limited and incomplete as that of the world of things. This occurs in the paradoxes of set theory, which are frequently alleged as a disproof of Platonism, but, I think, quite unjustly.*"

The End