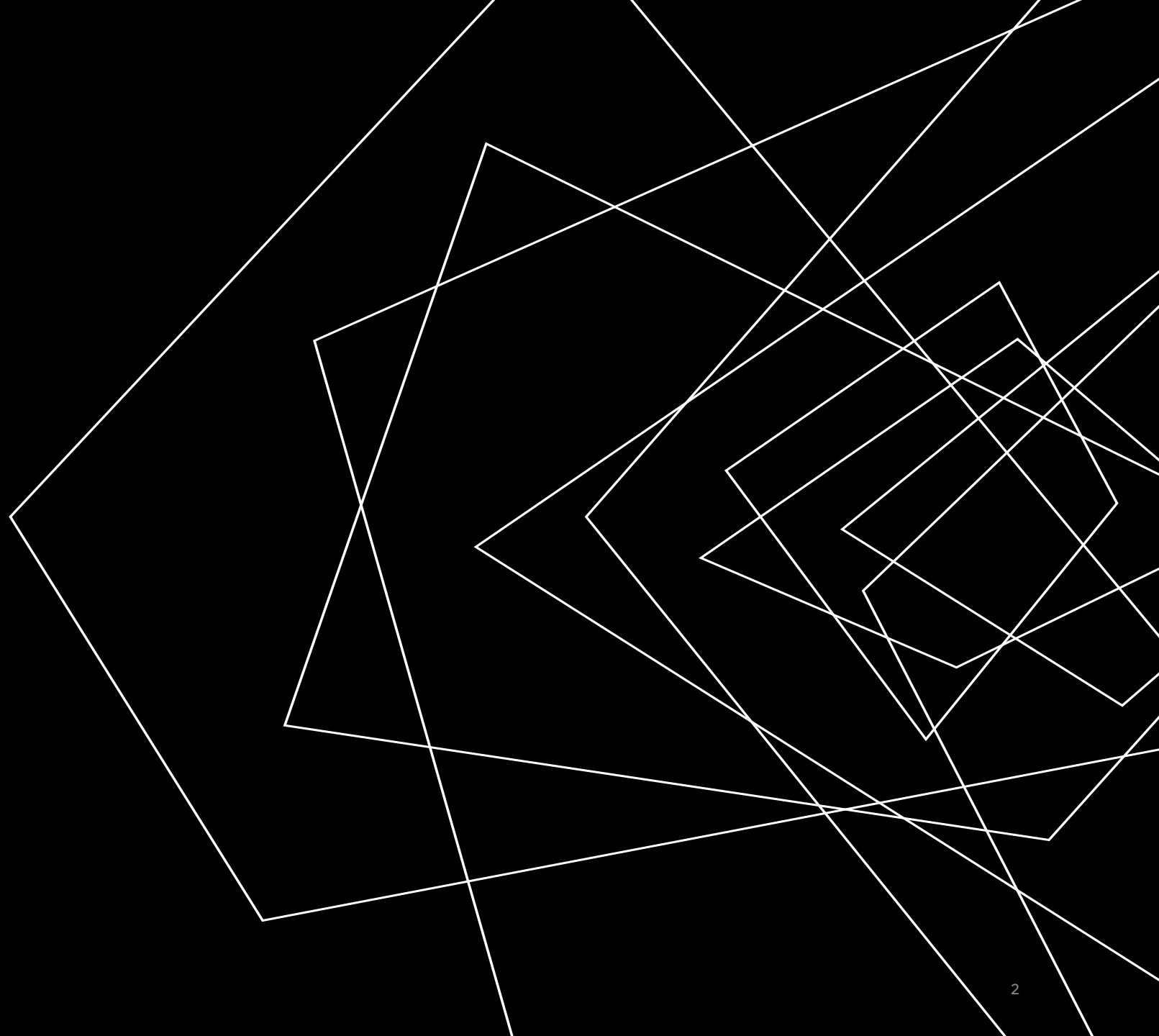
GÖDEL-OV ARGUMENT PROTIV JAKE UMJETNE INTELIGENCIJE

SADRŽAJ



TEOREMI O NEPOTPUNOSTI

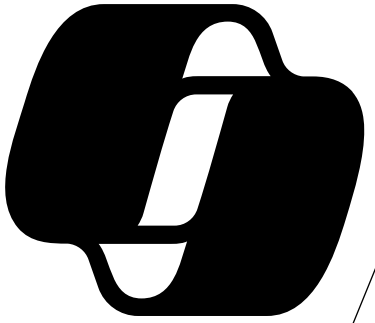
- Kurt Gödel, 1931. godina, Königsberg
- John R. Lucas i Roger Penrose
- „Bilo koji dosljedan formalan sistem u sklopu kojeg se može izraziti određena količina osnovne aritmetike nije kompletan, odnosno, postoje tvrdnje jezika u sistemu F koje se ne mogu dokazati, a ni osporiti u okviru tog sistema F“.

„OVA FORMULA NIJE DOKAZIVA UNUTAR SISTEMA“

- Šta se dešava ukoliko je ova formula dokaziva unutar sistema?
- Šta se dešava ukoliko je formula nedokaziva unutar sistema?

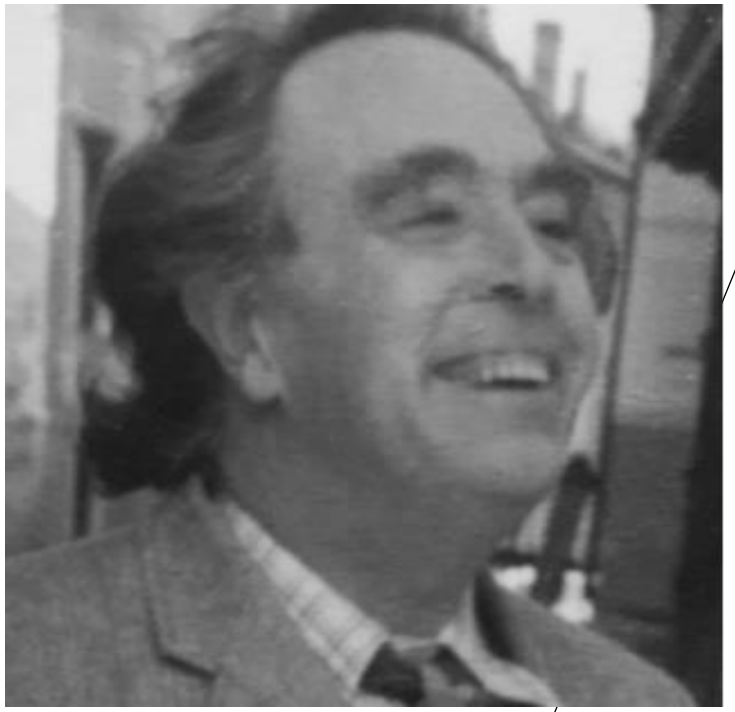
GÖDEL-OVA REČENICA

- ChatGPT vs Microsoft Copilot.
- Kakvo pitanje postaviti?



JOHN R. LUCAS

- „Ova formula nije dokaziva unutar sistema“
- Mind, Machines and Gödel
- „Nema šanse da mašina nešto napravi sama“
- Ljudski um nije superioran u odnosu na sve mašine;
- Može li mašina biti jednaka ljudskom umu?;



ANALOGIJA BIBLIOTEKE

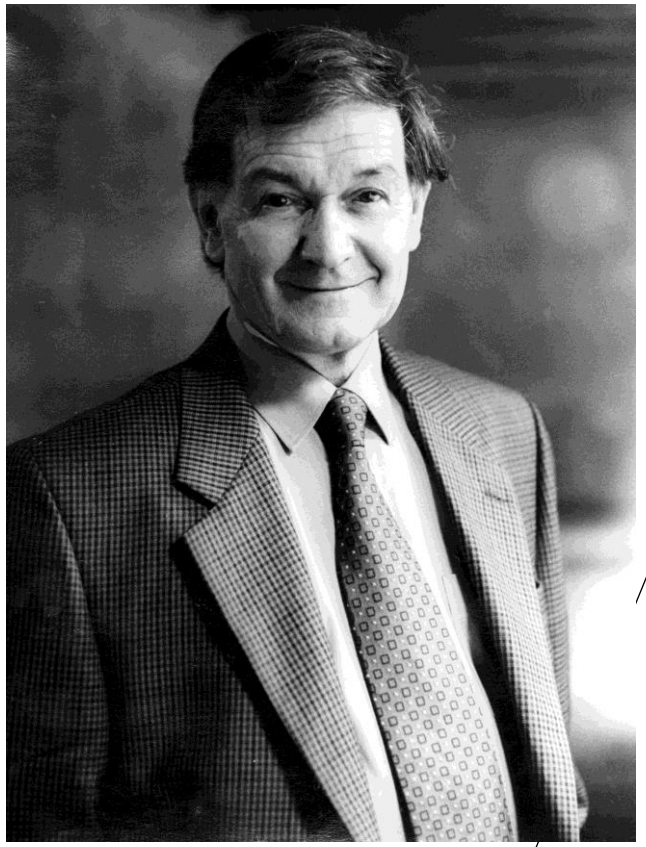
Dekameron
Zbirka novela Giovani-a Boccaccio-a

Mind, Machines and Gödel
Ova knjiga se ne nalazi u sistemu

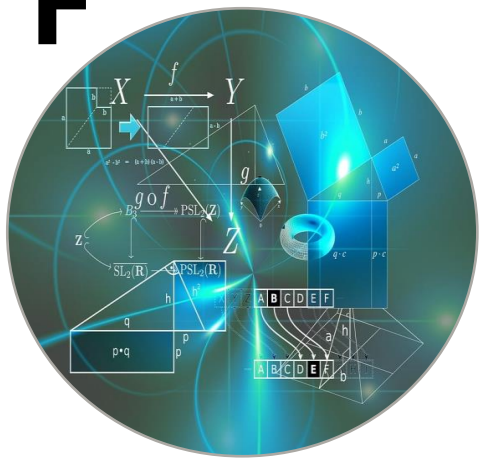
?

ROGER PENROSE

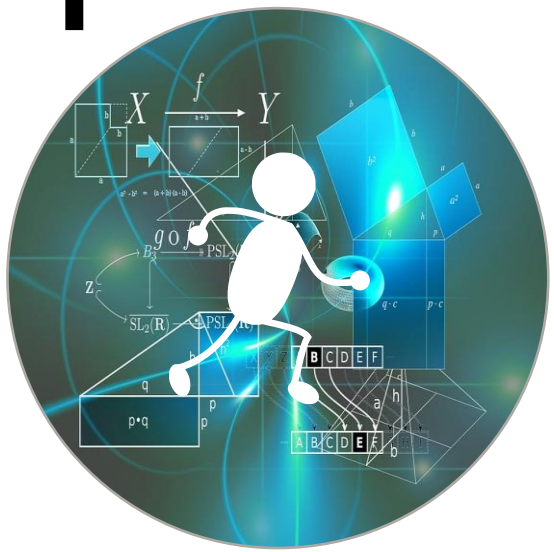
- „The Emperor‘s New Mind“
- Matematičko razmišljanje nije moguće enkapsulirati u neki čisto računarski model misli;
- „Shadows of the mind“
- „Beyond the Doubting of A Shadow“



F



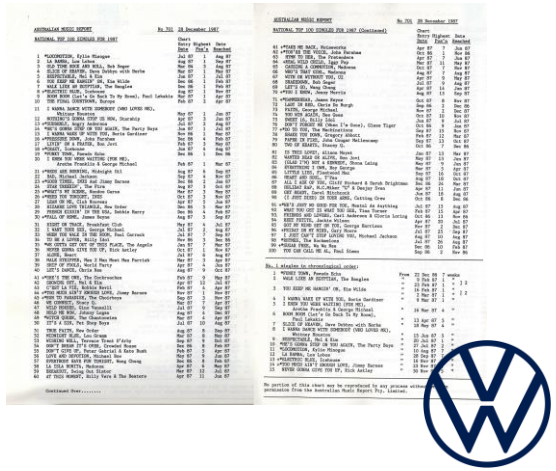
F‘



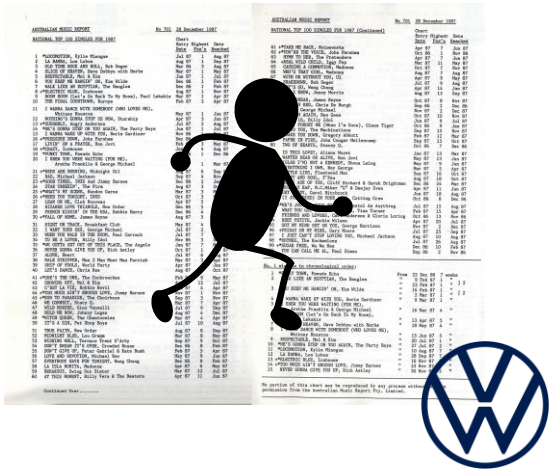
G(F‘)

ANALOGIJA AUTOMEHANIČARSKE RADNJE

F



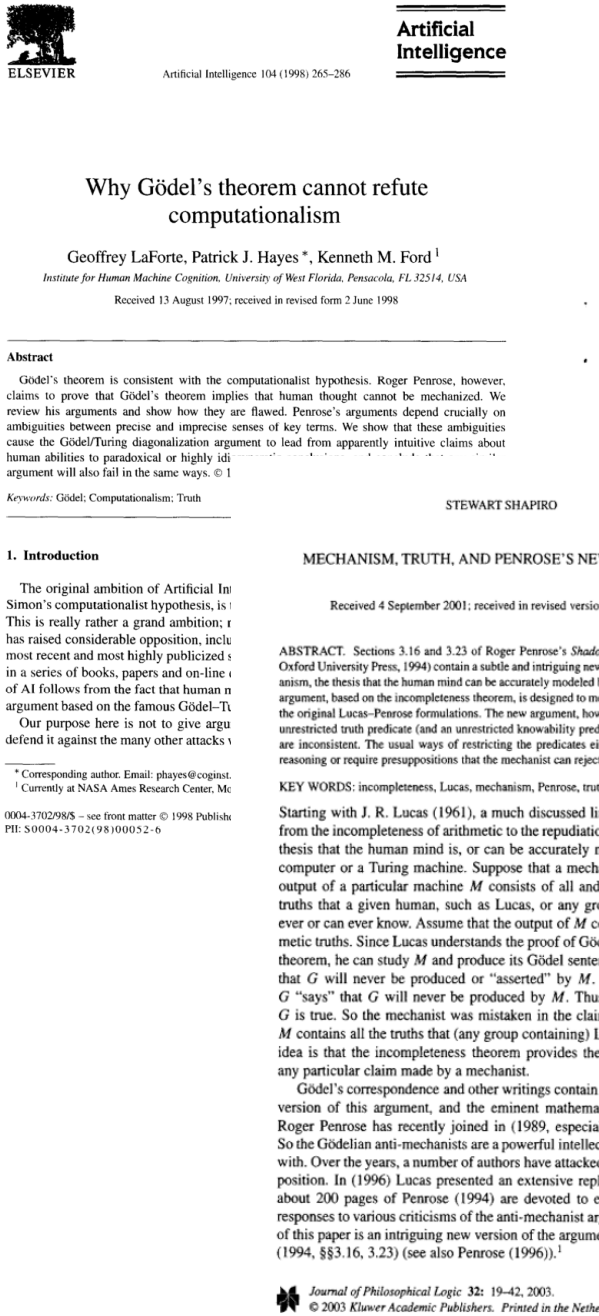
F‘



G(F‘)

KRITIKE

- LaForte, Hayes, Ford;
- Bringsjord i Xiao;
- Shapiro;



ELSEVIER

Artificial Intelligence 104 (1998) 265–286

Artificial Intelligence

Why Gödel's theorem cannot refute computationalism

Geoffrey LaForte, Patrick J. Hayes *, Kenneth M. Ford [1]

Institute for Human Machine Cognition, University of West Florida, Pensacola, FL 32514, USA

Received 13 August 1997; received in revised form 2 June 1998

Abstract

Gödel's theorem is consistent with the computationalist hypothesis. Roger Penrose, however, claims to prove that Gödel's theorem implies that human thought cannot be mechanized. We review his arguments and show how they are flawed. Penrose's arguments depend crucially on ambiguities between precise and imprecise senses of key terms. We show that these ambiguities cause the Gödel/Turing diagonalization argument to lead from apparently intuitive claims about human abilities to paradoxical or highly idi
argument will also fail in the same ways. © 1

Keywords: Gödel; Computationalism; Truth

1. Introduction

The original ambition of Artificial In
Simon's computationalist hypothesis, is
This is really rather a grand ambition;
has raised considerable opposition, inclu
most recent and most highly publicized
in a series of books, papers and on-line
of AI follows from the fact that human n
argument based on the famous Gödel–T
Our purpose here is not to give argu
defend it against the many other attacks

* Corresponding author. Email: phayes@coginst.
[1] Currently at NASA Ames Research Center, Mc

0004-3702/98/$ – see front matter © 1998 Publishe
PII: S0004-3702(98)00052-6

J. Expt. Theor. Artif. Intell. 12(2000)307–329

A refutation of Penrose's Gödelian case against artificial intelligence†

SELMER BRINGSJORD and HONG XIAO

Department of Philosophy, Psychology & Cognitive Science,
Department of Computer Science, Rensselaer Polytechnic Institute
NY 12180 USA
il: selmer@rpi.edu
//www.rpi.edu/~brings

ing, as it is generally agreed, failed to destroy the computational
ind with the Gödelian attack he articulated in his *The Emperor's*
se has returned, armed with a more elaborate and more fastidious
xpressed in Chapters 2 and 3 of his *Shadows of the Mind*. The core
se chapters is enthymematic, and when formalized, a remarkable
nical glitches come to light. Over and above these defects, the
st, is an instance of either the fallacy of denying the antecedent, the
principii, or the fallacy of equivocation. More recently, writing in
ritics in the electronic journal *Psyche*, Penrose has offered a Gödelian
improve on the version presented in *SOTM*. But this version is
er failure. In falling prey to the errors we uncover, Penrose's new
s unmasked as the same confused refrain J. R. Lucas initiated 35

del's incompleteness theorems, artificial intelligence, computation-
m

STEWART SHAPIRO

MECHANISM, TRUTH, AND PENROSE'S NEW ARGUMENT

Received 4 September 2001; received in revised version 22 July 2002

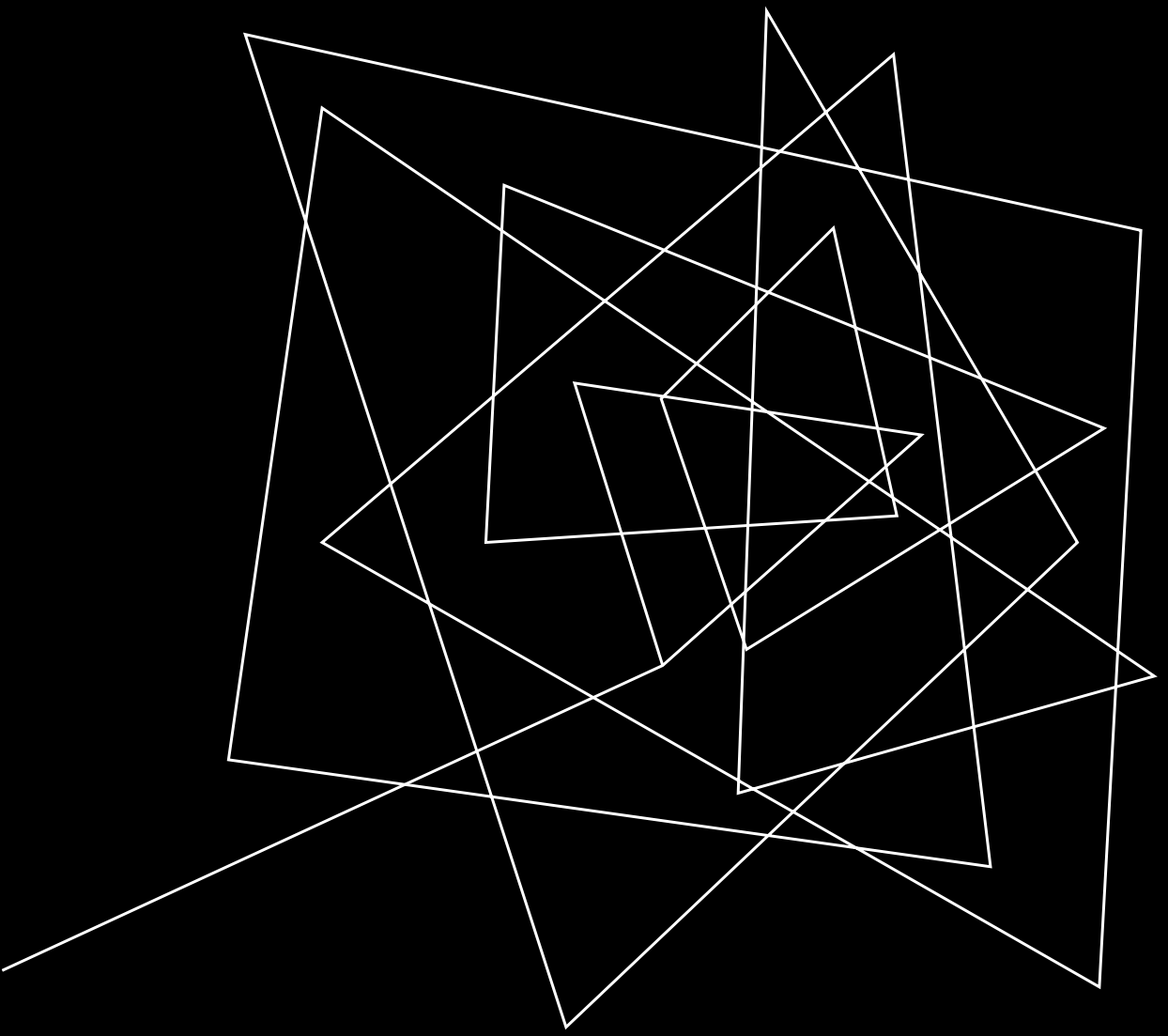
ABSTRACT. Sections 3.16 and 3.23 of Roger Penrose's *Shadows of the mind* (Oxford, Oxford University Press, 1994) contain a subtle and intriguing new argument against mechanism, the thesis that the human mind can be accurately modeled by a Turing machine. The argument, based on the incompleteness theorem, is designed to meet standard objections to the original Lucas–Penrose formulations. The new argument, however, seems to invoke an unrestricted truth predicate (and an unrestricted knowability predicate). If so, its premises are inconsistent. The usual ways of restricting the predicates either invalidate Penrose's reasoning or require presuppositions that the mechanist can reject.

KEY WORDS: incompleteness, Lucas, mechanism, Penrose, truth

Starting with J. R. Lucas (1961), a much discussed line of argument goes from the incompleteness of arithmetic to the repudiation of the mechanistic thesis that the human mind is, or can be accurately modeled as, a digital computer or a Turing machine. Suppose that a mechanist claims that the output of a particular machine M consists of all and only the arithmetic truths that a given human, such as Lucas, or any group of humans, will ever or can ever know. Assume that the output of M consists of only arithmetic truths. Since Lucas understands the proof of Gödel's incompleteness theorem, he can study M and produce its Gödel sentence G. Lucas knows that G will never be produced or "asserted" by M. He also knows that G "says" that G will never be produced by M. Thus, Lucas knows that G is true. So the mechanist was mistaken in the claim that the output of M contains all the truths that (any group containing) Lucas can know. The idea is that the incompleteness theorem provides the resources to refute any particular claim made by a mechanist.

Gödel's correspondence and other writings contain a carefully qualified version of this argument, and the eminent mathematician and physicist Roger Penrose has recently joined in (1989, especially Chapters 4, 10). So the Gödelian anti-mechanists are a powerful intellectual group to reckon with. Over the years, a number of authors have attacked the Lucas–Penrose position. In (1996) Lucas presented an extensive reply to his critics, and about 200 pages of Penrose (1994) are devoted to even more extensive responses to various criticisms of the anti-mechanist argument. The subject of this paper is an intriguing new version of the argument found in Penrose (1994, §§3.16, 3.23) (see also Penrose (1996)).[1]

Journal of Philosophical Logic **32:** 19–42, 2003.



HVALA NA PAŽNJI