Indecidibilità e incompletezza

Seconda parte: Incompletezza

Gabriele Lolli

(Con la collaborazione di Stefano Baratella e Domenico Luminati.)

Questa frase è falsa! Il Primo Teorema di Incompletezza

Nel numero 13 di *XlaTangente*, abbiamo presentato alcuni risultati dovuti al matematico Alan Turing ed abbiamo anticipato come tali risultati abbiano legami profondi con il *Primo Teorema di Incompletezza di Gödel* che li ha cronologicamente preceduti. È arrivato quindi il momento di tornare indietro nel tempo....

Il fenomeno dell'*autoreferenza* (o *autoriferimento*) è una fonte inesauribile di paradossi e nello scorso numero di *XlaTangente* avevamo (appena) accennato a come Gödel abbia trasformato questo fenomeno in una tecnica matematica molto feconda.

Per spiegare cosa intendiamo per autoreferenza presentiamo un famoso *paradosso semantico*. Si tratta di una frase che parla di se stessa. La possibilità di autoreferenza è comune a tutti i linguaggi usati nella comunicazione quotidiana, come italiano, tedesco, mandarino, etc... cioè ai cosiddetti *linguaggi naturali*. Ciò che Gödel ha dimostrato nei suoi lavori è che anche molte teorie matematiche hanno un certo grado di autoreferenza.

1 Il paradosso del mentitore

Il *paradosso del mentitore* è provocato da una frase come la seguente tra virgolette “questa frase è falsa”

La frase tra virgolette non può essere né vera né falsa, o meglio, è sia vera che falsa. Infatti, se la frase è vera, allora è vero quello che afferma, e cioè che la frase stessa è falsa. Se invece la frase è falsa, allora è falso quello che afferma, e la frase è dunque vera.

Gödel riconosce che il paradosso del mentitore, ed altri paradossi semantici, sono conseguenza del fatto che, nei linguaggi naturali, si possono costruire frasi che parlano di frasi dello stesso linguaggio. In particolare si possono costruire frasi che parlano di se stesse, come quella che provoca il paradosso del mentitore, cioè frasi *autoreferenziali*.

2 Il Primo Teorema di Incompletezza di Gödel: i preliminari

Per enunciare il Primo Teorema di Incompletezza, dobbiamo introdurre alcune nozioni.

Parliamo prima di tutto di *teorie*. Formulare una teoria (in ambito matematico e scientifico, ma anche filosofico, sociale. . .), significa specificare i *principi* su cui la teoria si fonda (li chiameremo gli *assiomi* della teoria). Una volta fissati gli assiomi c'è un macchinario deduttivo, fatto di *regole di deduzione*, che consente di produrre *dimostrazioni* a partire dagli assiomi. Le conclusioni delle dimostrazioni sono i *teoremi* della teoria.

Strettamente parlando, una teoria è l'insieme dei suoi teoremi ma, dato che le regole di deduzione corrette sono le stesse per tutte le teorie, è conveniente identificare una teoria con i suoi assiomi.

Ci interessano le teorie che non hanno contraddizioni al proprio interno (le cossiddette *teorie consistenti*). Questo perché una fra le regole di deduzione corrette stabilisce che da una contraddizione è possibile dimostrare una qualsiasi asserzione (in latino: *ex falso sequitur quodlibet*). Quindi, una teoria inconsistente dimostra ogni asserzione (ed anche la negazione di ogni asserzione!). Una teoria inconsistente è troppo potente! . . . ma è anche di scarso interesse, proprio perché dimostra “tutto” ed “il contrario di tutto”.

Un'altra proprietà di una teoria, di cui abbiamo già parlato nel Dossier Turing pubblicato nel numero 13 di *XlaTangente*, è la *completezza*. Ricordiamo che una teoria si dice *completa* se ogni asserzione oggetto della teoria è dimostrabile oppure *refutabile* (cioè è dimostrabile la negazione) a partire dagli assiomi. Intuitivamente, le teorie complete sono quelle che non lasciano nessuna asserzione "in sospeso".

Quando si formulano gli assiomi di una teoria, solitamente li si elenca in modo effettivo. Anche per teorie con infiniti assiomi, in molti casi - ma non sempre! - si può produrre una lista degli assiomi in modo effettivo, ad esempio mediante un programma per computer. Senza insistere troppo sugli aspetti tecnici, diciamo che vogliamo distinguere fra le teorie che sono *ricorsivamente assiomatizzabili* (cioè quelle che hanno un insieme di assiomi elencabile in modo effettivo) e quelle che non lo sono.

Possiamo finalmente formulare il

Primo Teorema di Incompletezza di Gödel Ogni teoria matematica che sia

- *abbastanza potente* da esprimere alcune proprietà elementari delle operazioni aritmetiche di somma, prodotto, successore e nella quale è formalizzabile il principio di induzione sui numeri naturali e

- *ricorsivamente assiomatizzabile*,

se è consistente *non è completa*, cioè alcune fra le asserzioni oggetto della teoria non sono *nè dimostrabili nè refutabili* a partire dagli assiomi della teoria.

Soprattutto se le si incontra per la prima volta, le ipotesi del Primo Teorema di Incompletezza sembrano molto restrittive, ma in realtà non lo sono: molte teorie matematiche le soddisfano.

Vogliamo ora mettere in evidenza il legame tra il paradosso del mentitore e la tecnica usata da Gödel per dimostrare il Primo Teorema di Incompletezza.

Gödel si propone di creare una situazione simile a quella del paradosso del mentitore all'interno di una teoria matematica, in particolare egli lo fa per la *teoria dell'aritmetica* **PA** (**PA** per *Peano Arithmetic*, dal nome del matematico torinese Giuseppe Peano). Descriviamo **PA** nella Finestra 2 e ad essa faremo sempre riferimento nel seguito. Per creare una asserzione autoreferenziale all'interno di **PA** (in analogia con quanto avviene nel paradosso del mentitore), dapprima Gödel fissa in modo preciso un linguaggio per l'aritmetica, poi introduce una codifica numerica degli elementi del linguaggio aritmetico (si veda la Finestra 3). Mediante la codifica, i numeri naturali, che sono l'oggetto della teoria, denotano anche elementi del linguaggio come i simboli, i termini, le formule, le sequenze finite di formule (in particolare, quindi, le dimostrazioni) e così via. È proprio in questo modo che Gödel riesce a riprodurre il fenomeno dell'autoreferenza all'interno di **PA**.

D'ora in poi chiameremo *aritmetiche* le formule che si scrivono nel linguaggio di **PA** (descriviamo tale linguaggio nella Finestra 2).

Gödel costruisce poi un *enunciato* γ ("enunciato" è il nome tecnico di una asserzione) che, interpretato non come un enunciato aritmetico, ma come un enunciato autoreferenziale, si legge

"questo enunciato è indimostrabile".

Infine, Gödel prova che l'enunciato γ è *indecidibile*, cioè né dimostrabile né refutabile in una qualunque teoria consistente e ricorsivamente assiomatizzabile che permette di compiere le operazioni sui numeri che codificano le espressioni del linguaggio. In particolare, segue che l'enunciato γ è indecidibile in **PA**. Trovate una traccia della dimostrazione di questo fatto nella Finestra 4.

Anche l'affermazione di consistenza di **PA** è formalizzabile con un enunciato aritmetico, che è indecidibile nella teoria: questo è il contenuto del Secondo Teorema di Incompletezza. In realtà, il secondo Teorema di Incompletezza vale per ogni teoria ricorsivamente assiomatizzabile che sia almeno tanto potente quanto **PA**: ogni teoria con queste proprietà, se consistente, non è abbastanza potente da provare la propria consistenza.

Non si creda comunque che l'indecidibilità dell'enunciato γ sia una conseguenza immediata del ragionamento che si fa con il paradosso del mentitore. Infatti, il concetto di verità ha solo due valori, *vero* e *falso*, mentre quello di dimostrabilità, in una teoria consistente, presenta tre possibilità mutuamente esclusive: se φ è un qualunque enunciato, o φ è dimostrabile, o la sua negazione $\neg\varphi$ è dimostrabile, o nessuno fra φ e $\neg\varphi$ è dimostrabile.

I due teoremi di incompletezza di Gödel sono tra i principali risultati della matematica del XX secolo. Nella prossima Sezione discuteremo brevemente qualche loro conseguenza sui fondamenti e sulla filosofia della matematica.

3 L'importanza del lavoro di Gödel

Per svolgere i passaggi accennati nella Finestra 4, Gödel aveva bisogno che la codifica descritta nella Finestra 2 soddisfacesse alcune proprietà. Le operazioni sintattiche sulle parole hanno un carattere effettivo, quasi materiale, e le operazioni aritmetiche corrispondenti sui numeri di codifica devono avere lo stesso carattere. Inoltre, la realizzazione concreta di tali operazioni deve tradursi nella dimostrazione dei corrispondenti numerici. Quindi le operazioni e le relazioni sintattiche devono essere definite da formule aritmetiche che soddisfino tali condizioni. Ad esempio, con riferimento alla Finestra 3, se m è la codifica di una dimostrazione dell'enunciato codificato da n, allora $Dim(n, m)$ deve essere dimostrabile, e se m non è la codifica di una dimostrazione dell'enunciato codificato da n, allora $Dim(n, m)$ deve essere refutabile.

Da qui è nato l'interesse per le *funzioni effettivamente calcolabili*, di cui abbiamo parlato nell'articolo precedente, e per una loro caratterizzazione precisa che permettesse di provare le proprietà della codifica di cui abbiamo parlato sopra. Il seguito della storia è la nascita della *teoria della calcolabilità*, che nel 1936 ha portato alla definizione delle macchine da parte di Turing e ad altre caratterizzazioni equivalenti da parte di Gödel e di Alonzo Church.

Un altro motivo di interesse nei confronti degli algoritmi in genere dipendeva dall'intento di David Hilbert (1862-1943) di riuscire a dimostrare la non contraddittorietà dell'aritmetica con metodi che fossero di sicuro affidamento - "finitistici" li chiamava Hilbert, per esprimere il concetto vago che aveva in mente.

Il lavoro di Hilbert e di diversi suoi allievi negli anni Venti del secolo scorso si consolidò in un vero e proprio programma di lavoro, nel cui ambito si ottennero diversi risultati importanti, soprattutto sulla decidibilità di problemi e teorie. I risultati negativi di Gödel e Turing segnarono i limiti del programma.

L'esigenza di dimostrare la consistenza dell'aritmetica si era manifestata in seguito allo sviluppo ottocentesco della matematica, che aveva ricondotto la definizione dei sistemi numerici, fra cui i numeri reali, a quella dei numeri naturali, e definito i numeri naturali in termini insiemistici, come fatto da Richard Dedekind. La teoria degli insiemi sviluppatasi alla fine del 1800 forniva un inquadramento soddisfacente della matematica conosciuta, ma era una teoria rivolta esplicitamente allo studio dell'infinito, e basata su assunzioni discutibili, e per alcuni inaccettabili, quali appunto l'esistenza di insiemi infiniti. Nei primi usi entusiasti ed ingenui della teoria degli insiemi si erano incontrati anche preoccupanti paradossi (come quello di Russell, che trovate nella Finestra 1). Ed il bisogno di una dimostrazione di consistenza per la teoria degli insiemi era molto sentito, perché le dimostrazioni di consistenza della matematica superiore erano basate sulla consistenza della teoria degli insiemi.

Un secondo filone di analisi logica riguardava il fatto che, con la scoperta delle geometrie non euclidee e lo studio in generale di nuovi argomenti algebrici e geometrici, l'oggetto della matematica non era più visto come radicato direttamente nella realtà, ma si esprimeva nella organizzazione assiomatica delle diverse teorie: gli assiomi enunciavano solo vincoli reciproci dei concetti rappresentati negli assiomi, e le teorie esploravano le conseguenze di tali assunzioni, in linea di principio arbitrarie (anche se di solito storicamente motivate). La richiesta della non contraddittorietà era il minimo requisito per non vanificare il lavoro di deduzione.

Un'ulteriore proprietà sembrava interessante e desiderabile per le teorie assiomatiche, in particolare per l'aritmetica, quella della completezza (ne abbiamo parlato all'inizio del paragrafo precedente). Per l'aritmetica la completezza avrebbe garantito che il ricorso alla teoria degli insiemi sarebbe stato solo un modo comodo di abbreviare o semplificare i ragionamenti, ma non avrebbe permesso di dimostrare nulla di falso relativamente alle affermazioni concrete riguardanti i numeri, già completamente decise, e quindi in particolare avrebbe garantito la non contraddittorietà.

I teoremi di incompletezza di Gödel, e le dimostrazioni da lui date, hanno risposto negativamente a queste aspettative: l'aritmetica, e tutte le teorie consistenti che permettono una definizione formale della classe delle funzioni effettivamente calcolabili, sono incomplete; e la non contraddittorietà dell'aritmetica non è dimostrabile con strumenti finitistici, visto che che non è dimostrabile neanche con la forza dell'aritmetica stessa. Il lavoro di Gödel ha anche avuto un duplice effetto liberatorio: da una parte ha fatto capire che la matematica superiore non è solo ancella di quella elementare, ma consiste di un genuino arricchimento del pensiero, aperto a indefinite estensioni. (Infatti, ogni teoria che soddisfa le ipotesi del Primo Teorema di Incompletezza dà origine a teorie più forti per la semplice aggiunta dell'assunzione della propria consistenza.) Dall'altra parte, il lavoro di Gödel ha innescato lo studio nuovo dei fenomeni di ca lcolo effettivo, le cui conseguenze sono ora sotto gli occhi di tutti.

Riferimenti bibliografici

[Dawson 2001] J. W. Dawson, *Dilemmi logici*, Bollati Boringhieri, Torino, 2001.

[Gödel 1931] K. Gödel, "Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandte Systeme I", trad. it. in [Gödel 1999, pp. 113-38].

[Gödel 1999] K. Gödel, *Collected Works. Vol. I*, Oxford Univ. Press, Oxford, 1986; trad. it. *Opere*, vol. 1, Bollati Boringhieri, Torino, 1999.

[Hofstadter 1984] D. R. Hofstadter, *Gödel, Escher, Bach: un'eterna ghirlanda brillante*, Adelphi, Milano, 1984.

[Lolli 2004] G. Lolli, *Da Euclide a Gödel*, il Mulino, Bologna, 2004.

[Lolli 2007] G. Lolli, *Sotto il segno di Gödel*, il Mulino, Bologna, 2007.

[Turing 1994] A. M. Turing, *Intelligenza meccanica*, Bollati Boringhieri, Torino, 1994.

Finestra 1

Autoreferenza e paradossi

Ricordiamo che, nell'ambito dei linguaggi naturali, si dice *autoreferenziale* una frase si riferisce a se stessa, direttamente oppure attraverso qualche frase intermedia. Nell'articolo principale abbiamo anche presentato la frase autoreferenziale che genera il paradosso del mentitore. C'è da dire che non sempre l'autoreferenza causa paradossi: considerate, ad esempio, la frase "questa frase è in italiano". (Cosa potete dire della frase "questa frase è in tedesco"?)

L'autoreferenza è presente in vari contesti. Anche un quadro, o un'opera letteraria, possono essere autoreferenziali. Al riguardo, se non lo conoscete già, cercate in internet una riproduzione di "Le trahison des images" ("L'inganno delle immagini") di René Magritte, oppure pensate al libro dal titolo "Steal this book" ("Rubate questo libro") di Abbie Hoffman.

Ora vogliamo però presentare tre *paradossi semantici.* Sebbene non sia evidente come nel caso del paradosso del mentitore, anch'essi sono legati all'autoreferenza. Come per il paradosso del mentitore, si tratta di frasi della lingua italiana che non contengono errori logici evidenti, ma che fanno uso di termini, come "vero", "definibile", "aggettivo", che non sono necessariamente parte del linguaggio matematico. Non appena si dà una definizione matematica dei termini tra virgolette, i paradossi semantici scompaiono. Per questa ragione, non costituiscono una grave minaccia per il matematico. Inoltre sono divertenti!

Paradosso di Richard Jules Richard (1905) considera i numeri reali compresi tra 0 e 1, che si possono identificare con le successioni infinite di 0 e 1, in espansione binaria.

Quelli definibili si possono organizzare in una matrice infinita di questo tipo

$$\begin{array}{ccccccccccc}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\
0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & \dots \\
0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & \dots \\
0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & \dots \\
\vdots & & & & & & & & & &
\end{array}$$

dove ogni riga rappresenta un numero reale definibile.

Se indichiamo con $r_{m,n}$ l'elemento all'incrocio dell'm-esima riga e dell'n-esima colonna; si può definire la funzione

$$1 - r_{m,m},$$

chiamata *antidiagonale*.

La definizione dell'antidiagonale sembra corretta e legittima, in quanto eseguita con operazioni aritmetiche sulla matrice, e quindi rappresenta un numero reale definibile, ma allora deve coincidere con una delle righe della matrice. Se però l'antidiagonale coincidesse, diciamo, con la k-esima riga, per ogni m si avrebbe

$$1 - r_{m,m} = r_{k,m}$$

e si otterrebbe una contraddizione per $m = k$.

Dunque, l'antidiagonale rappresenta un numero reale che è, allo stesso tempo, definibile e non definibile!

Paradosso di Berry Attribuito da Russell a G. C. Berry (1905) è espresso da "il minimo numero non definibile con meno di venticinque sillabe", definito dalla formula appena scritta che ha ventiquattro sillabe.

Paradosso di Grelling Kurt Grelling propose di chiamare un aggettivo autologo se si applica a se stesso, altrimenti lo si dica eterologo. Ad esempio "corto" è corto, mentre "lungo" non è lungo.

Se l'aggettivo "eterologo" si applica a se stesso, allora eterologo è autologo, quindi non eterologo, e quindi non si applica a se stesso; se non si applica a se stesso, allora eterologo è eterologo, e quindi si applica a se stesso.

Ecco invece un esempio di un *paradosso logico*. I paradossi logici non nascono da nozioni non definite, o definite in modo vago. Essi "segnalano" al matematico la necessità di modificare i principi di base (gli assiomi) della teoria con cui sta lavorando, per evitarne l'inconsistenza.

Paradosso di Russell Il *principio di comprensione* della teoria degli insiemi di fine 1800 era l'assioma il cui significato intuitivo è "ogni proprietà matematica definisce un insieme". Per non cadere in un paradosso semantico, bisogna dare una definizione rigorosa di "proprietà matematica". Questo si fa senza problemi, non lo vediamo solo per una ragione di brevità. Vediamo invece un esempio di una proprietà che è, senza ombra di dubbio, matematica. Per ogni insieme, si presentano due possibilità: l'insieme appartiene a se stesso, oppure no. C'è da scommettere che gli insiemi che vi vengono subito in mente non appartengono a loro stessi. Anzi, sembra sia davvero difficile dare un esempio di un insieme che appartiene a se stesso! Comunque, gli insiemi che non appartengono a loro stessi sono esattamente quelli che soddisfano la formula insiemistica $x \notin x$.

Se accettiamo il principio di comprensione ottocentesco, la collezione

$$R = \{x : \ x \notin x\}$$

è un insieme. Inoltre, per quanto abbiamo appena osservato,

$$\forall x (x \in R \leftrightarrow x \notin x),$$

da cui la contraddizione

$$R \in R \leftrightarrow R \notin R.$$

La scoperta del paradosso di Russell ha avuto come effetto l'indebolimento del principio di comprensione ottocentesco nella forma in cui lo utilizziamo oggi: "per ogni proprietà matematica P e per ogni insieme A, la collezione degli elementi di A che soddisfano P è un insieme".

Finestra 2

La teoria **PA** dell'Aritmetica di Peano

Il linguaggio per l'aritmetica ha i simboli $\mathbf{0}$, $=$, $<$, s per la funzione successore, $+$, $\cdot$, le variabili (una lista infinita: $x_1, x_2, \ldots, x_n, \ldots$) e i simboli logici (connettivi come $\neg, \wedge, \vee, \rightarrow \ldots$ e quantificatori come $\forall$ and $\exists$). Ci sono anche gli usuali simboli per le parentesi: (e). Notate che distinguiamo il numero 0 dal simbolo $\mathbf{0}$. Dovremmo anche usare un carattere diverso da $+$ per il secondo simbolo di funzione, così da non creare confusione tra il linguaggio formale che stiamo descrivendo ed il linguaggio matematico in cui $+$ ha già un significato (l'operazione di somma tra numeri naturali, nel nostro contesto). E così dovremmo comportarci con tutti gli altri simboli per cui c'è rischio di confusione: non lo faremo per non appesantire la trattazione.

Con questi simboli si possono scrivere i *termini* (che corrispondono alle espressiono aritmetiche) e le *formule*, ad esempio $x_1 < x_3$; $\neg(x_{11} < 0)$; $\forall x_1(x_1 + \mathbf{0} = x_1)$; $(x_1 < \mathbf{0}) \wedge (x_2 \cdot x_3 = \mathbf{0})$... ed infinite altre.

La teoria **PA** - oggetto del lavoro di Gödel - ha i seguenti assiomi (vicino a ciascuno ne trovate il significato intuitivo):

- $\forall x_1 \, (s(x_1) \neq \mathbf{0})$ (zero non è successore di nessun numero)
- $\forall x_1 \forall x_2 (s(x_1) = s(x_2) \rightarrow x_1 = x_2)$ (la funzione successore è iniettiva)
- $$\forall x_1(x_1 + \mathbf{0} = x_1)$$ $$\forall x_1 \forall x_2 \, (x_1 + s(x_2) = s(x_1 + x_2))$$ (definizione ricorsiva della somma)
- $$\forall x_1 \, (x_1 \cdot \mathbf{0} = \mathbf{0})$$ $$\forall x_1 \forall x_2 \, (x_1 \cdot s(x_2)) = x_1 \cdot x_2 + x_1)$$ (definizione ricorsiva del prodotto)
- per ogni formula φ, l'assioma

  $(\varphi(\mathbf{0}) \wedge \forall x_1(\varphi(x_1) \rightarrow \varphi(s(x_1))) \rightarrow \forall x_1 \varphi(x_1)$ (formalizzazione del principio di induzione)

Dopo aver superato qualche inevitabile difficoltà nella lettura, vi potete rendere conto che tutti gli assiomi sono veri nei numeri naturali.

Finestra 3

Codifica del linguaggio aritmetico

Una possibile assegnazione di numeri ai simboli dell'alfabeto del linguaggio per l'artimetica che abbiamo introdotto nella Finestra 2, si presenta – con varianti inessenziali – nel modo seguente, che usa la scomposizione in fattori primi dei numeri naturali per codificare le informazioni:

x_1	3
x_2	3^2
x_3	3^3
$\vdots$	$\vdots$
$\mathbf{0}$	5
$=$	$7 \cdot 11^2$
$<$	$7^2 \cdot 11^2$
s	$13 \cdot 17$
$+$	$13^2 \cdot 17^2$
$\cdot$	$13^3 \cdot 17^2$
$\neg$	19
$\wedge$	23
$\vdots$	

Ad esempio, il numero 634933 codifica il simbolo $\cdot$ di moltiplicazione, infatti $634933 = 13^3 \cdot 17^2$; l'esponente 3 di 13 ci dice che stiamo codificando il terzo simbolo di funzione, mentre l'esponente 2 di 17 ci dice che la funzione ha due argomenti.

A ogni espressione legale, termini e formule, è associato un numero grazie alla possibilità di codificare con numeri le sequenze finite di numeri. Questi numeri sono detti *gödeliani* o *numeri di Gödel* delle entità linguistiche che rappresentano.

Oltre a saper riconoscere in modo effettivo che un numero è un numero di questo tipo, si devono saper estrarre i numeri delle espressioni componenti: se è èil numero di un termine i numeri degli eventuali sottotermini, se è il numero di una formula i numeri del suo segno logico, delle sottoformule, delle variabili libere, dei termini che vi compaiono.

L'obiettivo si può realizzare assegnando i numeri nel modo seguente:

- se r_1, r_2 sono i gödeliani dei termini t_1, t_2,

$$2^{7\cdot 11^2} \cdot 3^{r_1} \cdot 5^{r_2} \text{ è il gödeliano della formula } t_1 = t_2;$$

$$2^{7^2\cdot 11^2} \cdot 3^{r_1} \cdot 5^{r_2} \text{ è il gödeliano della formula } t_1 < t_2;$$

- se n è il gödeliano della formula φ,

$$2^{19} \cdot 3^n \text{ è il gödeliano di } (\neg\varphi)$$

- se n è il gödeliano di φ e m il gödeliano di ψ,

$$2^{23} \cdot 3^n \cdot 5^m \text{ è il gödeliano di } (\varphi \wedge \psi)$$

- ...

Possiamo anche codificare una dimostrazione, che è una sequenza finita $\varphi_1, \ldots, \varphi_r$ di formule: se $n_1, \ldots, n_r$ sono i gödeliani delle formule che la costituiscono, si può associare alla dimostrazione il numero $2^{n_1} \cdot 3^{n_2} \cdot \ldots \cdot p_r^{n_r}$, dove p_i è l'i-simo numero primo ($p_1 = 2$).

Finestra 4

La dimostrazione di Gödel

Nel linguaggio aritmetico non c'è nulla che corrisponda al "questo" che compare nel paradosso del mentitore (e anche in italiano è discutibile che il significato di "questo" catturi proprio la denotazione che si vuole senza margini di ambiguità).

Per realizzare in modo preciso l'autoriferimento - nella teoria **PA** che abbiamo descritto nella Finesta 2 - Gödel ha dimostrato un risultato generale sui *punti fissi* delle formule. Da questo risultato segue l'esistenza di un numero k che codifica l'asserzione "k non ha dimostrazione". (Qui e nel seguito è da tenere a mente che, mediante la codifica, possiamo associare ad ogni formula il suo numero di Gödel, quindi "k non ha dimostrazione" significa "l'enuciato con numero di Gödel k non ha dimostrazione".)

L'affermazione tra virgolette è, in realtà, un enunciato aritmetico: esiste infatti una formula $Dim(x_1, x_2)$ tale che, per ogni coppia di numeri naturali m, n

- $Dim(m, n)$ è dimostrabile nella teoria **PA** se m codifica una dimostrazione dell'enunciato n;

- $Dim(m, n)$ è refutabile nella teoria **PA** se m non codifica una dimostrazione dell'enunciato n.

Si considera allora la formula $\neg\exists x_2 Dim(x_1, x_2)$ e si conclude che esiste un numero k tale che

$$k \text{ è il numero di Gödel di } \neg\exists y Dim(k, y).$$

In un certo senso, k e "k non ha dimostrazione" sono sostituibili uno all'altro, con tutta l'attenzione richiesta dai continui passaggi da un livello all'altro - quello aritmetico vero e proprio e quello linguistico dell'interpretazione sovrapposta.

Se si dimostrasse "k non ha dimostrazione", cioè se ci fosse una dimostrazione di $\neg\exists y Dim(k, y)$, e se m fosse la codifica di una tale dimostrazione, allora si dimostrerebbe "k ha la dimostrazione m". Si dimostrerebbe dunque $Dim(k, m)$, e si avrebbe una contraddizione.

Se si dimostrasse "k ha una dimostrazione", cioè se ci fosse una dimostrazione di $\neg\exists y Dim(k, y)$, questo andrebbe contro il fatto appena visto che nessun numero m codifica una dimostrazione di k.

Concludiamo che l'enunciato $\neg\exists y Dim(k, y)$ non è nè dimostrabile nè refutabile nella teoria **PA**.