Sull'Indimostrabilità Interna della Verità e sull'Incompletezza dei Sistemi Formali.
Studio, Critica alle Correnti Anti-Gödeliane e Teorizzazione

Enrico Maria Bufacchi

Indice

Prefazione **6**

1 Impianto Metateorico 7

1 Completezza, Compattezza, Incompletezza **8**

1.1 Sintassi e semantica: FOL, modelli e soddisfacibilità . . . 9
- 1.1.1 Sintassi . . . 9
- 1.1.2 Strutture e interpretazioni . . . 9
- 1.1.3 Soddisfacibilità tarskiana . . . 10
- 1.1.4 Conseguenza semantica e sintattica . . . 10
- 1.1.5 Definibilità ed equivalenza elementare . . . 11

1.2 Diagonalizzazione e sentenza di Gödel . . . 11

1.3 Condizioni di derivabilità e Löb . . . 12

1.4 Formalizzazione dell'essenzialismo semantico . . . 14

1.5 Teorema di completezza di Gödel (1930) e compattezza . . . 14
- 1.5.1 Costruzione di Henkin . . . 14
- 1.5.2 Compattezza . . . 14

1.6 Arifmetizzazione e induzione: verso la *incompleteness* . . . 15
- 1.6.1 Codici di Gödel e rappresentabilità . . . 15
- 1.6.2 Il predicato di provabilità . . . 15
- 1.6.3 Induzione e forza aritmetica . . . 15

1.7 *Incompleteness* I–II di Gödel, Rosser, Löb . . . 15
- 1.7.1 Il miglioramento di Rosser . . . 15
- 1.7.2 Secondo teorema di Gödel . . . 16

1.7.3 Corollari e fenomenologia Π_1^0 17
1.8 Teorema di indefinibilità della verità di Tarski 17
1.9 Church–Turing: non-decidibilità e limiti algoritmici 18
1.9.1 Tesi e schema metodologico 18
1.9.2 Il problema della fermata 18
1.9.3 Rice e completezza r. e. 19
1.9.4 Conseguenze per la provabilità 19

2 Figure e Strutture: da Hilbert a Tarski 20
2.1 Hilbert e il programma finitista 20
2.1.1 Teoremi di conservatività e battuta d'arresto 20
2.2 Gödel: completezza e incompletezza 21
2.2.1 *Derivability conditions* e aritmetizzazione 21
2.3 Tarski: verità e metalinguaggio 22
2.3.1 Indefinibilità e gerarchie semantiche 22

2 Correnti Gödeliane 23

3 Provabilità, interpretabilità, riflessione 24
3.1 GL di Löb e completezza di Solovay 24
3.1.1 Significato aritmetico e condizioni HBL 24
3.1.2 Fissazione modale e diagonalizzazione 24
3.1.3 Completezza aritmetica di Solovay 25
3.2 Assiomi di riflessione di Feferman e gerarchie T_α 26
3.2.1 Riflessione locale, uniforme e iterata 26
3.3 Logiche di interpretabilità di Visser 27
3.3.1 Principi tipici e correttezza aritmetica 27

4 Indefinibilità, informazione, modelli 28
4.1 Tarski e predicati di verità parziali 28
4.2 Kripke: punti fissi e verità parziale 29
4.3 Chaitin: complessità e limiti informativi 29

5 Letture filosofiche di Gödel 31
5.1 Realismo semantico e oggettività del contenuto 31
5.1.1 Deflazionismo, minimalismo e limiti 33
5.2 Incompletezza come argomento anti-meccanicista 33

5.2.1 Tesi meccanicista: formulazioni 33
5.2.2 Disgiunzione di Gödel e forma dell'argomento 34
5.2.3 Lucas–Penrose e formalizzazione del soggetto razionale . 34
5.2.4 Caveat epistemici 35
5.2.5 Ruolo delle verità Π_1^0 35
5.3 Note su Putnam, Penrose, Boolos 35
5.3.1 Putnam e il realismo modellistico 35
5.3.2 Penrose e l'anti-meccanicismo 36
5.3.3 Boolos, Berry e complessità 36

3 Critica Strutturale 38

6 La Tesi Verità-Provabilità: Analisi e Confutazione 39
6.1 *Proof-theoretic semantics* 39
6.2 Curry–Howard e la confusione ontologica 40
6.3 LEM-DNE: limiti dell'incompletezza classica 41
6.4 Disgiunzione-Esistenza: limiti della costruttività 41

7 Superare Gödel? 43
7.1 Intuizionismo, rilevanza, paraconsistenza: invarianti d'incompletezza 43
7.2 Secondo ordine e categoricità: il costo del metalinguaggio 44
7.3 Topoi e verità interna: dualismo interno-esterno 45
7.4 *Type theory* e Girard: coerenza ed espressività 45

4 Critica Meta-Logica 46

8 Generalizzazioni dell'Incompletezza 47
8.1 Schema gödeliano generale e secondo teorema astratto 47
8.2 Löb, GL e logiche della provabilità 48

8.3 Interpretabilità e riflessione: Feferman e Montagna 49

9 Indefinibilità e Gerarchie **51**

9.1 Tarski: nessun predicato di verità interno 51

9.2 Kripke e i punti fissi parziali 52

9.3 Verità esterna e modelli non-standard 52

5 Ricostruzione Trascendentale **54**

10 Dalla Provabilità alla Verità **55**

10.1 Separare semantica e metateoria della dimostrazione 55

10.2 Verità come condizione del giudizio epistemico 56

10.3 Riconoscimento intersoggettivo del valore di verità 57

11 Fenomenologia del Giudizio e Fondamento del Vero **59**

11.1 Lemma di Bufacchi: enunciato e conseguenze 59

11.1.1 Estensioni composizionali conservative e non-determinazione 62

11.1.2 Formulazione categoriale dell'errore di fondamento 63

11.2 Fondamento non-fondato e performatività 63

11.3 Soggetto razionale ideale e accordo minimo 64

6 Applicazioni Critiche **66**

12 Applicazioni critiche della teoria della verità **67**

12.1 Goodstein, Kirby–Paris, Hydra: vero ma non-dimostrabile in PA 67

12.1.1 Codifica ordinale e base variabile 67

12.1.2 Trasferimento agli ordinali e principio di terminazione . . 68

12.1.3 Gioco dell'Hydra e crescita non-primitiva 69

12.1.4 Limite del metodo aritmetico interno 71

12.2 CH, forzatura e ZFC: verità modellistiche 71

12.2.1 Cardinalità e continuità 71

12.2.2 Struttura di ZFC e assiomi 71

12.2.3 Idea della forzatura 71

12.2.4 Indipendenza di CH 72

12.2.5 Pluralismo modellistico 73
12.3 Dimostrazioni automatiche e limiti d'incompletezza 73
12.3.1 Verifica meccanica 73
12.3.2 Automazione e ricerca 73
12.3.3 Limiti teorici 73
12.3.4 Aporie epistemiche 74
12.3.5 Vie di ricerca 74
12.3.6 Convergenza delle aporie del vero 74

13 Approfondimenti e sviluppi avanzati **75**
13.1 Gerarchie di riflessione e stabilizzazione parziale 75
13.2 Sistemi di verità parziale e non-invarianza 76
13.3 Limiti *proof-theoretic* e ordinali 76
13.4 Logiche modali di provabilità e interpretabilità 77
13.5 Logica interna dei topos 78
13.6 Limiti computazionali e Goodstein 78

14 Conclusione **79**
14.1 Irriducibilità del vero e dualità giudizio-verità 79
14.2 Prospettive di ricerca avanzate 80

Prefazione

Identificare la verità con la dimostrabilità è negare l'essenzialismo semantico che i teoremi di incompletezza e di indefinibilità mettono a nudo.

— Enrico Maria Bufacchi

Questo lavoro argomenta una *confutazione sistematica* di ogni tentativo di ridurre la definizione di verità alla definizione di provabilità interna ad un sistema formale. Si mostra come tale riduzione, nelle sue varianti proof-theoretic, costruttiviste, pragmatistiche o revisioniste, sia internamente incapace di definire il limite semantico che i risultati di Gödel, Tarski, Löb, Feferman e le analisi sulla complessità algoritmica aprono. Si mettono in evidenza gli invarianti metateorici che sopravvivono alla variazione della logica e si ricostruisce una definizione di verità come *condizione trascendentale di possibilità del giudizio epistemico* che va oltre ogni schema puramente derivazionale.

La tesi centrale è la seguente. Ogni sistema formale T che soddisfa condizioni minime di rappresentazione aritmetica, come interpretazione di Q, chiusura per *modus ponens*, enumerabilità dei teoremi, genera un *essenzialismo semantico* E_T. Questo E_T non è internamente comprensibile né da estensioni aggiuntive conservative né da riformulazioni della definizione di dimostrazione in senso computazionale o costruttivo. L'essenzialismo si manifesta in tre strati: *indeterminatezza interna*, *indefinibilità*, *irriducibilità riflessiva*.

Parte 1

Struttura Metateorica e Risultati Fondativi (***Pars Expositiva***)

1

Completezza, Compattezza, Incompletezza

Si definiscono preliminarmente gli elementi minimi: linguaggi, strutture, soddisfacibilità, schemi di derivazione. Segue la trasformazione dalla *completezza semantica* alla *incompletezza aritmetica* attraverso l'arifmetizzazione del meta-discorso.

Definizione 1.1 (Linguaggio del primo ordine)**.** *Un linguaggio del primo ordine* $\mathcal{L}$ *è una tupla* $(\mathcal{F}, \mathcal{P}, \mathcal{C}, =)$*. L'insieme* $\mathcal{F}$ *contiene simboli funzionali con arietà finita. L'insieme* $\mathcal{P}$ *contiene simboli predicativi. L'insieme* $\mathcal{C}$ *contiene simboli di costante e può essere vuoto. Gli enunciati si generano da queste componenti tramite le regole sintattiche standard.*

Definizione 1.2 (Teoria)**.** *Una teoria* T *in* $\mathcal{L}$ *è un insieme ricorsivamente enumerabile di formule chiuse, ovvero assiomi.* T *è chiusa per conseguenza logica rispetto a un calcolo deduttivo fissato. Si considera il calcolo hilbertiano, il calcolo naturale e il calcolo dei sequenti. Si richiede a* T *di interpretare l'aritmetica di Robinson* Q *e di rappresentare funzioni primitive ricorsive di base.*

Assioma 1.1 (Minimalità Aritmetica)**.** *Ogni teoria* T *soddisfa:*

i. $Q \subseteq T$.

ii. T *è consistente (non dimostra* $\perp$*).*

iii. T *è* omega-consistente *se richiesto per rafforzare forme di incompleteness rosseriano.*

Proposizione 1.1 (Completezza e Incompletezza)**.** *La completezza di Gödel (1930) garantisce che se un enunciato è logicamente valido allora è dimostrabile. L'Incompletezza (1931) mostra, invece, che per teorie aritmeticamente sufficienti esistono enunciati veri nel modello standard dei naturali ma non-dimostrabili in* T*. Ciò definisce la nascita dell'essenzialismo semantico* E_T*.*

1.1 Sintassi e semantica: FOL, modelli e soddisfacibilità

1.1.1 Sintassi

Sia $\mathcal{L} = (\mathcal{F}, \mathcal{P}, \mathcal{C}, =)$ un linguaggio del primo ordine. Gli *insiemi di variabili* sono denotati da $\mathsf{Var} = \{x_1, x_2, \dots\}$. I *termini* sono definiti come segue:

i. Ogni variabile $x \in \mathsf{Var}$ è un termine.

ii. Ogni costante $c \in \mathcal{C}$ è un termine.

iii. Se $f \in \mathcal{F}$ è un simbolo funzionale di arietà n e $t_1, \dots, t_n$ sono termini, allora $f(t_1, \dots, t_n)$ è un termine.

Le *formule* sono definite come segue:

i. Se t_1, t_2 sono termini, allora $t_1 = t_2$ è una formula.

ii. Se $P \in \mathcal{P}$ è un simbolo predicativo di arietà n e $t_1, \dots, t_n$ sono termini, allora $P(t_1, \dots, t_n)$ è una formula.

iii. Se φ e ψ sono formule, allora $\neg\varphi$, $\varphi \wedge \psi$, $\varphi \vee \psi$, $\varphi \to \psi$ sono formule.

iv. Se φ è una formula e $x \in \mathsf{Var}$, allora $\forall x\, \varphi$ e $\exists x\, \varphi$ sono formule.

1.1.2 Strutture e interpretazioni

Una $\mathcal{L}$*-struttura* è una tupla $\mathcal{M} = (M, \{f^{\mathcal{M}}\}_{f \in \mathcal{F}}, \{P^{\mathcal{M}}\}_{P \in \mathcal{P}}, \{c^{\mathcal{M}}\}_{c \in \mathcal{C}})$ dove:

i. $M \neq \emptyset$ è il dominio.

ii. Per ogni simbolo funzionale f di arietà n, $f^{\mathcal{M}} : M^n \to M$ è l'interpretazione di f.

iii. Per ogni simbolo predicativo P di arietà n, $P^{\mathcal{M}} \subseteq M^n$ è l'interpretazione di P.

iv. Per ogni costante c, $c^{\mathcal{M}} \in M$ è l'interpretazione di c.

Una *interpretazione* $s : \mathsf{Var} \to M$ assegna a ogni variabile un elemento del dominio. L'interpretazione estende ai termini come segue:

$$\|t\|_s^{\mathcal{M}} = \begin{cases} s(x) & \text{se } t \text{ è una variabile } x \\ c^{\mathcal{M}} & \text{se } t \text{ è una costante } c \\ f^{\mathcal{M}}(\|t_1\|_s^{\mathcal{M}}, \dots, \|t_n\|_s^{\mathcal{M}}) & \text{se } t \text{ è } f(t_1, \dots, t_n) \end{cases}$$

1.1.3 Soddisfacibilità tarskiana

La *la relazione di soddisfacibilità* $\mathcal{M}, s \models \varphi$ è definita per induzione strutturale:

i. $\mathcal{M}, s \models t_1 = t_2$ se e solo se $\|t_1\|_s^{\mathcal{M}} = \|t_2\|_s^{\mathcal{M}}$.

ii. $\mathcal{M}, s \models P(t_1, \dots, t_n)$ se e solo se $(\|t_1\|_s^{\mathcal{M}}, \dots, \|t_n\|_s^{\mathcal{M}}) \in P^{\mathcal{M}}$.

iii. condizioni standard per $\neg, \wedge, \vee, \to$.

iv. $\mathcal{M}, s \models \forall x\, \varphi$ se e solo se per ogni $a \in M$ vale $\mathcal{M}, s[x \mapsto a] \models \varphi$. Analogamente per $\exists$.

Si scrive $\mathcal{M} \models \varphi$ per le frasi chiuse, ovvero nessuna variabile libera, e $\models \varphi$ per la validità logica, ovvero vera in ogni struttura. Per un insieme di frasi T, $\mathcal{M} \models T$ indica che $\mathcal{M}$ soddisfa tutte le frasi in T.

1.1.4 Conseguenza semantica e sintattica

La *conseguenza semantica* è $T \models \varphi$ se e solo se per ogni $\mathcal{M}$, $\mathcal{M} \models T$ implica $\mathcal{M} \models \varphi$. La *conseguenza sintattica* $T \vdash \varphi$ è relativa a un calcolo fissato. Valgono i seguenti risultati fondamentali:

Proposizione 1.2 (Correttezza). *Se $T \vdash \varphi$, allora $T \models \varphi$.*

Teorema 1.1 (Completezza). *Se $T \models \varphi$, allora $T \vdash \varphi$.*

Corollario 1.1 (Compattezza). *Se ogni sottoinsieme finito $T_0 \subseteq T$ soddisfa $T_0 \models \varphi$, allora $T \models \varphi$.*

Corollario 1.2 (Teorema di Löwenheim-Skolem)**.** *Se T ha un modello infinito, allora ha modelli di ogni cardinalità infinita.*

1.1.5 Definibilità ed equivalenza elementare

Sia $\mathcal{M}$ una $\mathcal{L}$-struttura. Un sottoinsieme $A \subseteq M^n$ è *definibile* se esiste una formula $\varphi(\bar{x})$ tale che $A = \{\bar{a} \in M^n : \mathcal{M} \models \varphi(\bar{a})\}$. Due strutture $\mathcal{M}, \mathcal{N}$ sono *elementarmente equivalenti*, $\mathcal{M} \equiv \mathcal{N}$, se soddisfano le stesse frasi di $\mathcal{L}$. Queste definizioni articolano il dibattito fra verità interna al modello e verità esterna standard che vi sarà nelle forme forti di incompletezza.

1.2 Diagonalizzazione e sentenza di Gödel

Lemma 1.1 (Lemma di Diagonalizzazione)**.** *Per ogni formula $\varphi(x)$ del linguaggio aritmetico a una variabile libera esiste una frase θ tale che*

$$T \vdash \ \theta \leftrightarrow \varphi(\ulcorner \theta \urcorner).$$

Dimostrazione. Si lavora in una teoria T r. e. che rappresenta tutte le funzioni primitive ricorsive e si usa la rappresentabilità della funzione di sostituzione. Esiste una funzione primitiva ricorsiva $\mathrm{Sub}(y, x)$ tale che $\mathrm{Sub}(\ulcorner \psi(z) \urcorner, \overline{n})$ è il numero di Gödel della formula ottenuta sostituendo il numerale $\overline{n}$ alla variabile libera z in ψ. Per rappresentabilità, esiste una formula binaria $\mathrm{Sub}^*(y, x)$ che rappresenta Sub in T.

Si definisce ora la formula a due variabili

$$\beta(y, x) \equiv \exists u \Big(\mathrm{Sub}^*(y, x, u) \wedge u = \ulcorner \chi \urcorner \to \chi \Big),$$

intesa come schema metalinguistico che unisce il risultato della sostituzione alla formula corrispondente. Per il Lemma della Rappresentazione Diagonale standard, esiste $\delta(y)$ tale che

$$T \vdash \ \delta(y) \leftrightarrow \varphi\Big(\mathrm{Sub}(y, \ulcorner y \urcorner)\Big).$$

Si pone ora $\theta := \delta(\ulcorner\delta\urcorner)$. Allora, sostituendo y con $\ulcorner\delta\urcorner$ nella bicondizionale, si ottiene

$$T \vdash \theta \leftrightarrow \varphi\Big(\mathrm{Sub}(\ulcorner\delta\urcorner, \ulcorner\ulcorner\delta\urcorner\urcorner)\Big).$$

Ma per definizione di Sub si ha $\mathrm{Sub}(\ulcorner\delta\urcorner, \ulcorner\ulcorner\delta\urcorner\urcorner) = \ulcorner\theta\urcorner$. Pertanto

$$T \vdash \theta \leftrightarrow \varphi(\ulcorner\theta\urcorner).$$

□

Teorema 1.2 (Gödel I, schema minimale)**.** *Sia T consistente, r. e., con $Q \subseteq T$. Esiste una formula G_T tale che, nel modello standard $\mathbb{N}$,*

i. $\mathrm{True}_{\mathbb{N}}(G_T)$*,*

ii. $T \nvdash G_T$*.*

Dimostrazione. Si fissa un predicato di provabilità $\mathrm{Prov}_T(x)$ che rappresenta in T la nozione di dimostrabilità. Usando il Lemma 1.1, si costruisce la formula G_T tale che

$$T \vdash G_T \leftrightarrow \neg\,\mathrm{Prov}_T(\ulcorner G_T\urcorner).$$

La verità di G_T nel modello standard segue dalla sua interpretazione: se fosse falso, allora sarebbe dimostrabile, contraddicendo la consistenza di T. La non-dimostrabilità di G_T in T è diretta dalla sua definizione.

□

1.3 Condizioni di derivabilità e Löb

Sia $\mathrm{Prov}_T(x)$ un predicato di provabilità per T. Le condizioni di Hilbert–Bernays–Löb (HBL) richiedono, internamente a T:

i. $T \vdash \mathrm{Prov}_T(\ulcorner A\urcorner) \to \mathrm{Prov}_T(\ulcorner \mathrm{Prov}_T(\ulcorner A\urcorner)\urcorner)$.

ii. $T \vdash \mathrm{Prov}_T(\ulcorner A \to B\urcorner) \to (\mathrm{Prov}_T(\ulcorner A\urcorner) \to \mathrm{Prov}_T(\ulcorner B\urcorner))$.

iii. Se $T \vdash A$, allora $T \vdash \mathrm{Prov}_T(\ulcorner A\urcorner)$.

Teorema 1.3 (Löb)**.** *Se $T \vdash \mathrm{Prov}_T(\ulcorner A\urcorner) \to A$, allora $T \vdash A$.*

Dimostrazione. Per il Lemma 1.1 esiste B tale che

$$T \vdash \; B \leftrightarrow (\mathrm{Prov}_T(\ulcorner B \urcorner) \to A).$$

Si mostrano i seguenti passi interni a T.

i. Dalla metà $B \to (\mathrm{Prov}_T(\ulcorner B \urcorner) \to A)$ e da HBL si ottiene

$$T \vdash \; \mathrm{Prov}_T(\ulcorner B \urcorner) \to \mathrm{Prov}_T\Big(\ulcorner \mathrm{Prov}_T(\ulcorner B \urcorner) \to A \urcorner\Big).$$

ii. Per HBL si ha

$$T \vdash \; \mathrm{Prov}_T(\ulcorner B \urcorner) \to \mathrm{Prov}_T\Big(\ulcorner \mathrm{Prov}_T(\ulcorner B \urcorner) \urcorner\Big).$$

iii. Applicando la condizione di distribuzione (HBL) a $\mathrm{Prov}_T(\ulcorner B \urcorner) \to A$ si ottiene

$$T \vdash \; \mathrm{Prov}_T\Big(\ulcorner \mathrm{Prov}_T(\ulcorner B \urcorner) \to A \urcorner\Big) \to \Big(\mathrm{Prov}_T\Big(\ulcorner \mathrm{Prov}_T(\ulcorner B \urcorner) \urcorner\Big) \to \mathrm{Prov}_T(\ulcorner A \urcorner)\Big).$$

iv. Combinando i punti (i) e (ii) con (iii) per *modus ponens* si ricava

$$T \vdash \; \mathrm{Prov}_T(\ulcorner B \urcorner) \to \mathrm{Prov}_T(\ulcorner A \urcorner).$$

v. Per ipotesi di Löb $T \vdash \mathrm{Prov}_T(\ulcorner A \urcorner) \to A$, quindi da (iv) segue

$$T \vdash \; \mathrm{Prov}_T(\ulcorner B \urcorner) \to A.$$

vi. Dalla definizione di B si ha $T \vdash \; (\mathrm{Prov}_T(\ulcorner B \urcorner) \to A) \to B$, dunque

$$T \vdash \; \mathrm{Prov}_T(\ulcorner B \urcorner) \to B.$$

vii. Per HBL e (vi) si ottiene

$$T \vdash \; \mathrm{Prov}_T\Big(\ulcorner \mathrm{Prov}_T(\ulcorner B \urcorner) \urcorner\Big) \to \mathrm{Prov}_T(\ulcorner B \urcorner).$$

Con (ii) si conclude $T \vdash \; \mathrm{Prov}_T(\ulcorner B \urcorner)$. Infine, per (v) si deduce $T \vdash A$.

□

Corollario 1.3. *Se T è consistente, allora $T \nvdash \mathrm{Con}(T)$.*

1.4 Formalizzazione dell'essenzialismo semantico

Definizione 1.3 (Essenzialismo Semantico)**.** *Per una teoria T come sopra, si definisce l'*essenzialismo semantico *E_T come la classe degli enunciati ψ tali che* $\mathrm{True}_{\mathbb{N}}(\psi)$*, con $T \nvdash \psi$ e, per ogni estensione r. e. $T \subseteq T'$, se T' è conservativa per enunciati Π^0_1, allora ancora $T' \nvdash \psi$.*

Teorema 1.4 (Irriducibilità di E_T)**.** *Se T è consistente e interpreta Q, allora $E_T \neq \emptyset$ e, per ogni estensione r. e. conservativa su Π^0_1, E_T non viene svuotato. Inoltre, non esiste predicato $\mathrm{Tr}(x) \in \mathcal{L}$ tale che, per tutte le frasi χ,*

$$T \vdash \ \chi \leftrightarrow \mathrm{Tr}(\ulcorner \chi \urcorner).$$

Dimostrazione. La prima parte segue dal teorema di Gödel I e dalle proprietà delle estensioni conservative. La seconda parte è il teorema di Tarski sull'indefinibilità della verità, che si dimostra tramite diagonalizzazione e l'argomento del paradosso del mentitore.

□

1.5 Teorema di completezza di Gödel (1930) e compattezza

1.5.1 Costruzione di Henkin

Si estende il linguaggio $\mathcal{L}$ a $\mathcal{L}'$ aggiungendo per ogni formula esistenziale $\exists x\, \varphi(x)$ una nuova costante c_φ e l'assioma di Henkin $\varphi(c_\varphi)$. Si costruisce una teoria T' in $\mathcal{L}'$ che estende T e che è massimale consistente. Si definisce quindi una struttura $\mathcal{M}$ il cui dominio è l'insieme dei termini chiusi di $\mathcal{L}'$ modulo l'equivalenza indotta da T'. Si verifica che $\mathcal{M} \models T'$ e quindi $\mathcal{M} \models T$.

1.5.2 Compattezza

Se ogni sottoinsieme finito di T ha un modello, allora T ha un modello. La dimostrazione usa Henkin, in quanto l'aggiunta di costanti e assiomi mantiene la coerenza su ogni parte finita e quindi sull'insieme intero.

1.6 Arifmetizzazione e induzione: verso la *incompleteness*

1.6.1 Codici di Gödel e rappresentabilità

Si fissa una codifica di Gödel che associa a ogni simbolo, termine, formula e dimostrazione un numero naturale unico. Una funzione $f : \mathbb{N}^k \to \mathbb{N}$ è *rappresentabile* in T se esiste una formula $\varphi_f(x_1, \dots, x_k, y)$ tale che per ogni $n_1, \dots, n_k, m \in \mathbb{N}$:

$$f(n_1, \dots, n_k) = m \implies T \vdash \varphi_f(\overline{n_1}, \dots, \overline{n_k}, \overline{m}),$$

$$f(n_1, \dots, n_k) \neq m \implies T \vdash \neg\varphi_f(\overline{n_1}, \dots, \overline{n_k}, \overline{m}).$$

1.6.2 Il predicato di provabilità

Si definisce $\mathrm{Prov}_T(x) := \exists y\, \mathrm{Proof}_T(y, x)$ con Proof_T primitiva ricorsiva. Le condizioni HBL si dimostrano verificando che Prov_T rispetta le condizioni di derivabilità interne a T.

1.6.3 Induzione e forza aritmetica

La forza di T dipende dallo schema di induzione. Per PA lo schema completo consente di lavorare con definizioni Δ_0^0 con parametri e di estendere la rappresentabilità a funzioni ricorsive totali. Teorie più deboli, come Q o $I\Sigma_1$, limitano la rappresentabilità e la forza dei risultati di incompletezza.

1.7 *Incompleteness* I–II di Gödel, Rosser, Löb

1.7.1 Il miglioramento di Rosser

Teorema 1.5 (Rosser). *Sia T una teoria r. e. che interpreta Q e sia consistente. Allora esiste una frase R_T tale che $T \nvdash R_T$ e $T \nvdash \neg R_T$.*

Dimostrazione. Si definisce il predicato primitivo ricorsivo $\mathrm{Proof}_T(y, x)$: "y codifica una dimostrazione in T della formula con codice x" e si considera la formula binaria

$$\mathrm{Rosser}(x) \; : \; \forall y \Big(\mathrm{Proof}_T(y, x) \to \exists z \leq y \, \mathrm{Proof}_T(z, \ulcorner \neg x \urcorner)\Big).$$

Per la diagonalizzazione esiste R tale che $T \vdash R \leftrightarrow \mathrm{Rosser}(\ulcorner R \urcorner)$. Si verifica ora che $T \nvdash R$ e $T \nvdash \neg R$ sotto sola consistenza di T.

Si suppone per assurdo $T \vdash R$. Allora esiste n con $\mathrm{Proof}_T(\bar{n}, \ulcorner R \urcorner)$. Dalla definizione di Rosser e dall'equivalenza interna si ottiene

$$T \vdash \exists z \leq \bar{n} \, \mathrm{Proof}_T\Big(z, \ulcorner \neg R \urcorner\Big),$$

cioè $T \vdash \mathrm{Prov}_T(\ulcorner \neg R \urcorner)$. Per la correttezza Σ^0_1 standard, esiste dunque effettivamente $m \leq n$ che codifica una dimostrazione di $\neg R$ in T. Ma allora T dimostra sia R sia $\neg R$, contro la consistenza.

Si assume ora $T \vdash \neg R$. Allora esiste n con $\mathrm{Proof}_T(\bar{n}, \ulcorner \neg R \urcorner)$. Si distinguono due casi. Se $\exists z \leq \bar{n} \, \mathrm{Proof}_T\Big(z, \ulcorner R \urcorner\Big)$, allora T è inconsistente. Altrimenti $\neg \exists z \leq \bar{n} \, \mathrm{Proof}_T\Big(z, \ulcorner R \urcorner\Big)$. Ma allora, per definizione di Rosser, risulta $T \vdash R$, ancora contro la consistenza. In ogni caso, da $T \vdash \neg R$ segue l'inconsistenza, quindi $T \nvdash \neg R$.

Si conclude che $R_T := R$ soddisfa $T \nvdash R_T$ e $T \nvdash \neg R_T$ senza assumere ω-consistenza.

□

1.7.2 Secondo teorema di Gödel

Teorema 1.6 (Gödel II)**.** *Se T è consistente, sufficientemente forte e r. e., allora $T \nvdash \mathrm{Con}(T)$.*

Dimostrazione. Si definisce la consistenza in T con il predicato Prov_T che soddisfa le condizioni HBL e si formalizza $\mathrm{Con}(T)$ come $\neg\mathrm{Prov}_T(\ulcorner \bot \urcorner)$. Per logica proposizionale in T vale

$$\mathrm{Con}(T) \to (\mathrm{Prov}_T(\ulcorner \bot \urcorner) \to \bot).$$

Se per assurdo $T \vdash \mathrm{Con}(T)$, allora $T \vdash \mathrm{Prov}_T(\ulcorner \bot \urcorner) \to \bot$. Applicando il teorema di Löb con $A \equiv \bot$, dalle condizioni HBL segue $T \vdash \bot$, contro la consistenza. Quindi $T \nvdash \mathrm{Con}(T)$.

□

1.7.3 Corollari e fenomenologia Π^0_1

Proposizione 1.3. *Se T è consistente e rappresenta la computazione primitiva ricorsiva, esistono enunciati veri Π^0_1 indipendenti da T.*

Esempi classici si ottengono tramite gerarchie di funzioni di Goodstein e risultati come Kirby–Paris. Tali enunciati sono veri nel modello standard, ma non-dimostrabili in T.

1.8 Teorema di indefinibilità della verità di Tarski

Teorema 1.7 (Tarski)**.** *Non esiste formula* $\mathrm{Tr}(x)$ *del linguaggio aritmetico tale che, per ogni frase χ, valga lo schema tarskiano*

$$T \vdash \ \chi \leftrightarrow \mathrm{Tr}(\ulcorner \chi \urcorner).$$

Dimostrazione. Si procede per assurdo. Si assume che esista $\mathrm{Tr}(x)$ tale che per ogni frase χ si abbia $T \vdash \chi \leftrightarrow \mathrm{Tr}(\ulcorner \chi \urcorner)$. Si applica il Lemma 1.1 alla formula $\neg\mathrm{Tr}(x)$ e si ottiene una frase θ con

$$T \vdash \ \theta \leftrightarrow \neg\mathrm{Tr}(\ulcorner \theta \urcorner).$$

Per l'ipotesi tarskiana applicata a $\chi \equiv \theta$ si ha anche

$$T \vdash \ \theta \leftrightarrow \mathrm{Tr}(\ulcorner \theta \urcorner).$$

Combinando le due bicondizionali per sostituzione di equivalenti si ricava

$$T \vdash \ \mathrm{Tr}(\ulcorner \theta \urcorner) \leftrightarrow \neg\mathrm{Tr}(\ulcorner \theta \urcorner).$$

Da cui $T \vdash \ \perp$. Si ottiene dunque una contraddizione con la consistenza di T. Ne segue che una tale $\mathrm{Tr}(x)$ non esiste.

□

La conseguenza concettuale è che la verità supera ogni definizione interna uniforme. Questo motiva l'introduzione di E_T.

1.9 Church–Turing: non-decidibilità e limiti algoritmici

1.9.1 Tesi e schema metodologico

Si assume la Tesi di Church–Turing come principio metodologico di identificazione tra calcolabilità effettiva e computabilità meccanica. Si intende per funzione calcolabile ogni funzione calcolabile da una macchina di Turing, oppure equivalente in uno dei modelli classici. La Tesi non è un teorema e agisce come tramite fra calcolo e formalizzazione matematica.

1.9.2 Il problema della fermata

Teorema 1.8 (Indecidibilità di HALT)**.** *L'insieme* $\mathrm{K} = \{\langle M, x\rangle : M \textit{ ferma su } x\}$ *non è decidibile.*

Dimostrazione. Si procede per assurdo. Si suppone che esista un decider totale H tale che, per ogni coppia $\langle M, x\rangle$, $H(\langle M, x\rangle) = 1$ se M ferma su x e $H(\langle M, x\rangle) = 0$ altrimenti. Sia U una macchina universale fissata e sia $\langle y, y\rangle$ una codifica standard della coppia (y, y).

Si definisce una macchina D che, su input y, computa $H(\langle y, y\rangle)$ e:

i. se $H(\langle y, y\rangle) = 1$, allora $D(y)$ entra in ciclo infinito.

ii. se $H(\langle y, y\rangle) = 0$, allora $D(y)$ ferma immediatamente.

La costruzione di D è effettiva, dunque esiste un codice d per D. Sia $D(d)$. Se $H(\langle d, d\rangle) = 1$, allora per definizione (i) $D(d)$ non ferma, in contraddizione con d che ferma su d. Se $H(\langle d, d\rangle) = 0$, allora per (ii) $D(d)$ ferma, in contraddizione con d che non ferma su d. In entrambi i casi si ottiene una contraddizione. Di conseguenza H non esiste e K non è decidibile.

□

1.9.3 Rice e completezza r. e.

Teorema 1.9 (Rice). *Ogni proprietà non-banale del linguaggio computato da un programma è indecidibile.*

Segue che la maggior parte delle proprietà semantiche dei programmi sono indecidibili e che non esiste un algoritmo generale per verificare l'equivalenza di due programmi.

1.9.4 Conseguenze per la provabilità

Si osserva che il *set* dei teoremi di una teoria r. e. è r. e. Se la teoria T interpreta Q, la teoria della verità aritmetica è indecidibile. Ne segue che il problema dell'appartenenza di una formula al diagramma dei teoremi di T è in generale indecidibile. Ciò vincola ogni tentativo di identificare verità e provabilità, e dimostra l'esistenza dell'essenzialismo semantico E_T.

2

Figure e Strutture: da Hilbert a Tarski

2.1 Hilbert e il programma finitista

2.1.1 Teoremi di conservatività e battuta d'arresto

Si vuole fondare l'uso di concetti *ideali* mostrando che ogni dimostrazione che li usa è *conservativa* sui contenuti *reali* finitamente giustificati. Si propone una metateoria con metodi finiti nella quale formalizzare dimostrazioni di coerenza e risultati di eliminazione.

Strumenti tecnici sono il calcolo ε di Hilbert e l'eliminazione dell'operatore ε, nonché i teoremi di *eliminazione* e di *conservatività definizionale*: l'aggiunta di simboli abbreviativi o di definizioni esplicite non crea nuovi teoremi in un dato frammento aritmetico.

Proposizione 2.1 (Schema di conservatività)**.** *Se T' si ottiene da T per estensioni puramente definizionali e regole giustificate da eliminazione del taglio, allora T' è conservativa su classi aritmetiche basse rispetto a T.*

L'idea hilbertiana è mostrare la coerenza di teorie come PA per via finita. Tuttavia, il secondo teorema di Gödel lo limita dall'interno.

Teorema 2.1 (Gödel II, limite al finitismo)**.** *Se T è r. e., sufficientemente forte e consistente, allora $T \nvdash \mathrm{Con}(T)$.*

La coerenza di T non è ottenibile *dentro* T con mezzi di T stesso: ciò pone un limite concettuale al programma finitista.

Gentzen e l'induzione transfinita

Gentzen supera il limite hilbertiano dimostrando la coerenza di PA usando l'induzione transfinita fino a ε_0. Tale metodo non è finito in senso strettamente hilbertiano, ma è accettabile da una prospettiva più ampia. La dimostrazione di Gentzen utilizza il calcolo dei sequenti e l'eliminazione del taglio per ridurre dimostrazioni di $\perp$ a dimostrazioni senza taglio, analizzando la complessità delle formule coinvolte.

Battuta d'arresto e lascito La battuta d'arresto è concettuale: una teoria non può dimostrare internamente la propria coerenza se soddisfa le ipotesi minime. Il lascito è invece metodologico: si ottengono riduzioni *proof-theoretic*, misure ordinali e teoremi di conservatività che definiscono il limite tra contenuto reale e ideale.

2.2 Gödel: completezza e incompletezza

2.2.1 *Derivability conditions* e aritmetizzazione

Si deve distinguere la completezza logica dalla incompletezza aritmetica e si deve aritmetizzare il metalinguaggio codificando formule e derivazioni. Si introduce $\mathrm{Prov}_T(x)$ come predicato Σ^0_1 e si dimostrano le condizioni di Hilbert–Bernays–Löb. Applicando il lemma di fissazione per costruire G_T, si ricava l'incompletezza.

Lemma 2.1 (Fissazione)**.** *Per ogni $\varphi(x)$ esiste θ tale che $T \vdash \theta \leftrightarrow \varphi(\ulcorner\theta\urcorner)$.*

Le condizioni HBL sostengono il teorema di Löb e il formato della riflessione schematica. Segue il secondo teorema di Gödel formalizzando $\mathrm{Con}(T)$ come negazione della provabilità di $\perp$.

2.3 Tarski: verità e metalinguaggio

2.3.1 Indefinibilità e gerarchie semantiche

Si formalizza l'idea tarskiana di verità come bicondizionale di Tarski–T schema nel metalinguaggio e si mostra l'impossibilità di definire internamente un predicato totale di verità per il linguaggio stesso dell'aritmetica, introducendo la gerarchia di linguaggi $\mathcal{L}_0 \subset \mathcal{L}_1 \subset \mathcal{L}_2 \subset \dots$. Ogni livello $\mathcal{L}_{n+1}$ contiene un predicato di verità per $\mathcal{L}_n$.

Teorema 2.2 (Indefinibilità)**.** *Non esiste* $\mathrm{Tr}(x)$ *in* $\mathcal{L}$ *tale che per ogni frase* χ *si abbia* $\chi \leftrightarrow \mathrm{Tr}(\ulcorner \chi \urcorner)$.

Le soluzioni parziali portano a predicati di verità parziali e semantiche a punti fissi. Esse non negano la necessità del metalinguaggio per avere la verità totale, ma corroborano l'argomento per l'essenzialismo semantico E_T.

Parte 2

Correnti Gödeliane: Estensioni, Letture, Impatti

3

Provabilità, interpretabilità, riflessione

3.1 GL di Löb e completezza di Solovay

Si considera il linguaggio modale con connettivi classici e l'operatore $\Box$. La logica della provabilità GL è data dagli assiomi di K più lo schema di Löb

$$\Box(\Box A \to A) \to \Box A$$

e dalla regola di necessitazione: se $\vdash A$, allora $\vdash \Box A$.

Un'*interpretazione aritmetica* associa a $\Box$ il predicato di provabilità Prov_T di una teoria aritmeticamente adeguata T e trasforma ogni formula modale in una formula aritmetica A^*.

3.1.1 Significato aritmetico e condizioni HBL

Le condizioni di Hilbert–Bernays–Löb per Prov_T mostrano la correttezza aritmetica di GL. Se $\vdash_{\mathsf{GL}} A$, allora per ogni interpretazione aritmetica $*$ vale $T \vdash A^*$. La dimostrazione è per induzione sulla derivazione usando necessitazione e chiusura di Prov_T sotto implicazione interna.

3.1.2 Fissazione modale e diagonalizzazione

La costruzione di enunciati autoreferenziali ha un analogo modale. Per ogni contesto $C(p)$ con p proposizionale esiste una formula G tale che $\mathsf{GL} \vdash G \leftrightarrow$

$C(\Box G)$. Questo lemma di fissazione modale usa la diagonalizzazione aritmetica e consente di dimostrare il teorema di Löb a livello modale.

3.1.3 Completezza aritmetica di Solovay

La completezza usa modelli aritmeticamente costruiti da funzioni di Solovay che mostrano la corrispondenza fra verità modale e provabilità aritmetica. L'idea fondamentale è definire una trasformazione controfattuale che forza il fallimento aritmetico quando la derivabilità modale fallisce. Segue che se $\mathsf{GL} \nvdash A$, allora esiste T e un'interpretazione $*$ con $T \nvdash A^*$.

Teorema 3.1 (Solovay)**.** *Per ogni formula modale A, si ha $\vdash_{\mathsf{GL}} A$ se e solo se per ogni teoria T r. e. che interpreta Q e per ogni interpretazione aritmetica $*$, vale $T \vdash A^*$.*

Dimostrazione. La prima implicazione è banale. Sia $\mathsf{GL} \nvdash A$. Per completezza modale di GL esiste un modello finito noetheriano $\mathcal{M} = (W, R, V)$ e $w \in W$ tali che $\mathcal{M}, w \nVdash A$. Si costruisce una realizzazione aritmetica che associa a ogni $u \in W$ una formula Sol_u con:

$$T \vdash \ \mathrm{Sol}_u \to \Big(p^* \leftrightarrow (\mathcal{M}, u \Vdash p)\Big) \quad \text{per ogni proposizionale } p,$$

e per l'operatore modale:

$$T \vdash \ \mathrm{Sol}_u \to \Big(\mathrm{Prov}_T(\ulcorner \varphi^* \urcorner) \leftrightarrow \bigwedge_{uRv} \Big(\mathrm{Sol}_v \to \varphi^*\Big)\Big).$$

La costruzione si fa definendo funzioni di Solovay come metodo di ricerca di confutazioni. Un'induzione su φ mostra che $T \vdash \mathrm{Sol}_u \to (\varphi^* \leftrightarrow (\mathcal{M}, u \Vdash \varphi))$. In particolare, per $u = w$ si ha $T \nvdash A^*$, altrimenti si ottiene $\mathcal{M}, w \Vdash A$, che è un assurdo. Quindi, se per ogni $T, *$ vale $T \vdash A^*$, segue $\vdash_{\mathsf{GL}} A$.

$\Box$

Proposizione 3.1. *Sia T una teoria r. e. che interpreta Q e sia $\mathrm{Prov}_T(x)$ un predicato di provabilità per T che soddisfa le condizioni HBL. Allora per ogni formula modale A vale:*

$$\vdash_{\mathsf{GL}} A \quad \textit{se e solo se} \quad T \vdash A^*,$$

dove $*$ è l'interpretazione aritmetica che manda $\Box$ in Prov_T.

La dimostrazione segue quella di Solovay, adattando la costruzione delle funzioni di Solovay al predicato Prov_T.

3.2 Assiomi di riflessione di Feferman e gerarchie T_α

Per una classe di formule Γ, lo *schema di riflessione* per T è

$$\mathrm{RFN}_\Gamma(T): \quad \{\, \mathrm{Prov}_T(\ulcorner\varphi\urcorner) \to \varphi : \varphi \in \Gamma \,\}.$$

La riflessione *locale* riguarda singole formule, quella *uniforme* riguarda famiglie di formule con parametri.

Proposizione 3.2. *Per molte teorie T che interpretano Q, l'aggiunta di $\mathrm{RFN}_{\Pi^0_1}(T)$ è consistente se e solo se T è consistente e guadagna forza proof-theoretic strettamente oltre T.*

Dimostrazione. **Consistenza:** L'istanza $\mathrm{Prov}_T(\ulcorner\bot\urcorner) \to \bot$ appartiene a $\mathrm{RFN}_{\Pi^0_1}(T)$ poiché $\bot$ è Π^0_1. Se $T + \mathrm{RFN}_{\Pi^0_1}(T)$ fosse inconsistente, allora T stesso lo sarebbe per monotonia. Quindi la consistenza di T implica la consistenza di $T + \mathrm{RFN}_{\Pi^0_1}(T)$.

Aumento di forza: In $T + \mathrm{RFN}_{\Pi^0_1}(T)$ si ha in particolare $\mathrm{Prov}_T(\ulcorner\bot\urcorner) \to \bot$, e dunque per logica classica $\neg\mathrm{Prov}_T(\ulcorner\bot\urcorner)$, cioè $\mathrm{Con}(T)$. Per il secondo teorema di Gödel, $T \nvdash \mathrm{Con}(T)$. Quindi $T + \mathrm{RFN}_{\Pi^0_1}(T)$ dimostra una nuova verità Π^0_1 non dimostrabile in T ed è strettamente più forte.

□

3.2.1 Riflessione locale, uniforme e iterata

La riflessione locale $\mathrm{RFN}^{\mathrm{loc}}_\Gamma(T)$ definisce singole necessità di formule in Γ dimostrate in T. La riflessione uniforme $\mathrm{RFN}^{\mathrm{unif}}_\Gamma(T)$ quantifica, invece, sui parametri della classe e corrobora sensibilmente la teoria. L'iterazione transfinita definisce $T_0 = T$, $T_{\alpha+1} = T_\alpha + \mathrm{RFN}_\Gamma(T_\alpha)$ e $T_\lambda = \bigcup_{\beta<\lambda} T_\beta$ per λ limite.

Feferman introduce progressioni transfinite $\{T_\alpha\}$ ottenute iterando schemi di riflessione lungo notazioni ordinali ricorsive. Per $\alpha < \varepsilon_0$ si ottengono teorie strettamente crescenti di forza, che realizzano finemente il passaggio dato dalla riflessione su classi sintattiche sempre più ampie.

3.3 Logiche di interpretabilità di Visser

La *interpretabilità* tra teorie è una relazione che rafforza la conservatività: U è interpretabile in T se esiste un'interpretazione che manda teoremi di U in teoremi di T. Le *logiche di interpretabilità* estendono GL con un connettivo modale binario $\rhd$ che codifica l'interpretabilità.

Proposizione 3.3 (Orey–Hájek)**.** *Per teorie T, U r. e. che interpretano Q, vale: T interpreta U se e solo se ogni Π^0_1-conseguenza di U è conseguenza di T e alcune condizioni tecniche di coerenza sono soddisfatte.*

Dimostrazione. Se U è interpretabile in T tramite τ, allora per ogni Π^0_1-frase χ, da $U \vdash \chi$ si ottiene $T \vdash \tau(\chi)$. La preservazione Π^0_1 implica $T \vdash \chi$. Supponendo $\mathrm{Th}_{\Pi^0_1}(U) \subseteq \mathrm{Th}_{\Pi^0_1}(T)$ e le condizioni standard di coerenza, si definisce per simboli di U un'interpretazione Δ^0 in T che rende veri in T gli assiomi interpretati di U. L'induzione su derivazioni e la rappresentabilità in T completano la costruzione dell'interpretazione.

$\square$

La logica IL e le sue estensioni, come ILM, assiomatizzano principi generali di interpretabilità che risultano completi rispetto a interpretazioni aritmetiche fra teorie r. e. che interpretano Q.

3.3.1 Principi tipici e correttezza aritmetica

Tra i principi caratteristici vi sono schemi che uniscono provabilità e interpretabilità come $\Box A \to (B \rhd A)$ e varianti di principio di Montagna. La correttezza aritmetica di tali principi si dimostra costruendo interpretazioni che preservano le classi Π^0_1 e usando la riflessione controllata.

4

Indefinibilità, informazione, modelli

4.1 Tarski e predicati di verità parziali

Si definiscono *predicati di verità parziali* quelli che soddisfano bicondizionali tarskiani solo per una classe ristretta di formule e rispettano condizioni compositionali. Esempi includono verità per formule Δ_0^0 con parametri, schemi composizionali come $\mathsf{CT}^-(T)$ senza induzione completa sul predicato di verità, o classi TB con bicondizionali per tutte le frasi ma con metateoria esterna.

Proposizione 4.1. *Se T interpreta Q, esistono estensioni conservative T' che ammettono un predicato $\mathrm{Tr}_\Gamma(x)$ composizionale corretto per una classe Γ chiusa per verità booleana e quantificazionale, senza ottenere un predicato di verità totale.*

Dimostrazione. Si pone $T' = T + \mathsf{CT}^-_\Gamma(\mathrm{Tr})$ con condizioni composizionali per Γ. Per induzione strutturale $T' \vdash \varphi \leftrightarrow \mathrm{Tr}(\ulcorner\varphi\urcorner)$ per ogni $\varphi \in \Gamma$. L'interpretazione di eliminazione che sostituisce $\mathrm{Tr}(\ulcorner\varphi\urcorner)$ con φ su formule prive di Tr mostra la conservatività: se $T' \vdash \chi$ senza Tr allora $T \vdash \chi$.

$\square$

Tali predicati definiscono la struttura composizionale della verità senza violare l'indefinibilità globale.

4.2 Kripke: punti fissi e verità parziale

Si considerano semantiche a tre valori con l'algebra di Kleene forte. Data una base di valutazioni per atomi del linguaggio, l'operatore che assegna valori di verità secondo condizioni del predicato T è monotono. Per il lemma di Knaster–Tarski esiste un punto fisso minimo che assegna verità alle frasi *grounded* e lascia *undefined* i paradossi.

Teorema 4.1 (Kripke)**.** *Esiste il minimo punto fisso dell'operatore di valutazione associato a T in semantica di Kleene forte. In tale punto fisso tutte le necessità composizionali sono soddisfatte e il predicato T è parzialmente definito senza contraddizioni.*

Dimostrazione. Si definisce l'operatore monotono F sull'algebra di Kleene forte: per atomi $F(I)(p) = I(p)$, per connettivi e quantificatori secondo le condizioni composizionali. La monotonia segue per induzione sulla struttura. Per Tarski–Knaster, F ammette un punto fisso minimo $\mathrm{lfp}(F) = \bigvee_\alpha F^\alpha(\bot)$. In $\mathrm{lfp}(F)$ il predicato T soddisfa le condizioni composizionali senza contraddizione e lascia indefinite le frasi non fondate.

□

I predicati di verità parziale ottenuti per punti fissi mostrano modi coerenti di estendere il linguaggio con un predicato di verità limitato, compatibile con l'indefinibilità totale.

4.3 Chaitin: complessità e limiti informativi

Definizione 4.1 (Complessità di Kolmogorov)**.** *La complessità di Kolmogorov $K(x)$ di una stringa binaria x è la lunghezza del più breve programma per una macchina universale di Turing che produce x come output.*

Una stringa è *incomprimibile* se $K(x)$ è grande rispetto alla sua lunghezza. La complessità di Kolmogorov non è calcolabile: non esiste un algoritmo che, dato x, determini $K(x)$ esattamente.

Teorema 4.2 (Chaitin)**.** *Per ogni T come sopra esiste c_T con la proprietà che se $T \vdash K(\bar{x}) > c_T$ per qualche x, allora T è incoerente. In particolare T non dimostra nessuna istanza vera sufficientemente forte di incomprimibilità.*

Dimostrazione. Sia U una macchina universale prefisso-libera. Per ineguaglianza di Kraft, $|\{x : K(x) \leq n\}| \leq 2^n - 1$. Se T provasse $K(\bar{x}) > n$ per più di $2^n - 1$ stringhe, almeno una sarebbe in realtà di complessità $\leq n$, falsificando T. Esiste dunque c_T tale che nessuna affermazione vera $K(\bar{x}) > c_T$ è provabile in T senza incoerenza. Formalizzando l'argomento in T si ottiene la tesi.

□

Corollario 4.1. *Per ogni teoria r. e. T che interpreta Q, esistono stringhe incomprimibili che T non può riconoscere come tali.*

Il teorema di Chaitin fornisce un'interpretazione informazionale dell'incompletezza: esistono limiti interni alla quantità di informazione che una teoria formale può prendere riguardo alla complessità delle stringhe.

5

Letture filosofiche di Gödel

5.1 Realismo semantico e oggettività del contenuto

La sostituzione della logica classica con varianti intuizioniste, di rilevanza o paraconsistenti non apre una via dall'incompletezza, ma rende più evidente l'irriducibile nucleo strutturale dell'argomento diagonale: ciò che produce l'essenzialismo semantico non è il principio del terzo escluso, ma la meccanicità enumerativa dell'apparato formale insieme alla capacità di aritmetizzare i propri processi, in modo tale che in ogni contesto dove una teoria r. e. interpreta Q e rappresenta funzioni primitive ricorsive, la sintassi può riferirsi a sé stessa definendo una formula che, affermando la propria non dimostrabilità, pone la frattura fra codice e verità come tratto invariabile.

Proposizione 5.1 (Incompletezza Intuizionista)**.** *L'aritmetica intuizionista* HA *è incompleta: esiste* G_{HA} *vero nello standard e non-dimostrabile in* HA*, e se* HA *è consistente allora* $\mathrm{HA} \nvdash \mathrm{Con(HA)}$*. Dunque* $\mathrm{True}_{\mathbb{N}}(G)$ *e* $\mathrm{HA} \nvdash G$*. Per* Con(HA)*, si adotta* $\neg\mathrm{Prov}_{\mathrm{HA}}(\ulcorner\bot\urcorner)$*. Se* $\mathrm{HA} \vdash \mathrm{Con(HA)}$*, allora* $\mathrm{HA} \vdash \mathrm{Prov}_{\mathrm{HA}}(\ulcorner\bot\urcorner) \to \bot$ *e, per Löb interno,* $\mathrm{HA} \vdash \bot$*, assurdo. Quindi* $\mathrm{HA} \nvdash \mathrm{Con(HA)}$.

Dimostrazione. In HA si rappresentano PR, vale il lemma di diagonalizzazione e si formalizzano le condizioni HBL per $\mathrm{Prov}_{\mathrm{HA}}$. Sia G tale che $\mathrm{HA} \vdash G \leftrightarrow \neg\mathrm{Prov}_{\mathrm{HA}}(\ulcorner G\urcorner)$. Se $\mathrm{HA} \vdash G$, allora per HBL (iii) $\mathrm{HA} \vdash \mathrm{Prov}_{\mathrm{HA}}(\ulcorner G\urcorner)$, contro G. Se $\mathrm{HA} \vdash \neg G$, nello standard $\mathrm{Prov}_{\mathrm{HA}}(\ulcorner G\urcorner)$, ancora contro G. Dunque $\mathrm{True}_{\mathbb{N}}(G)$

e HA $\nvdash G$. Per Con(HA), si adotta $\neg\mathrm{Prov}_{\mathrm{HA}}(\ulcorner\bot\urcorner)$. Se HA $\vdash$ Con(HA), allora HA $\vdash \mathrm{Prov}_{\mathrm{HA}}(\ulcorner\bot\urcorner) \to \bot$ e, per Löb interno, HA $\vdash \bot$, assurdo. Quindi HA $\nvdash$ Con(HA).

□

Proposizione 5.2 (Incompletezza Generale)**.** *Se T è r. e., di rilevanza o paraconsistente, interpreta Q e rappresenta PR, allora esiste G_T tale che* $\mathrm{True}_{\mathbb{N}}(G_T)$*,* $T \nvdash G_T$ *e, se T è* Π^0_1*-corretta,* $T \nvdash \neg G_T$*.*

Dimostrazione. Si definisce Prov_T nel linguaggio di T e si costruisce G per diagonalizzazione con $T \vdash G \leftrightarrow \neg\mathrm{Prov}_T(\ulcorner G\urcorner)$. Se $T \vdash G$ allora, per HBL (iii), $T \vdash \mathrm{Prov}_T(\ulcorner G\urcorner)$, contro G. Se T è Π^0_1-corretta e $T \vdash \neg G$, nello standard $\mathrm{Prov}_T(\ulcorner G\urcorner)$, ancora contro G. Pertanto $T \nvdash G$ e, assumendo Π^0_1-correttezza, $T \nvdash \neg G$.

□

La variazione del calcolo influisce sul piano inferenziale ma non sul nucleo di autoreferenzialità che valida la trascendenza semantica rispetto alla lista meccanica dei teoremi. La verità è indipendente dalle nostre procedure dimostrative e si basa su modelli e interpretazioni. I teoremi di Tarski e di Gödel mostrano che la definizione interna di verità e la sua riduzione a provabilità falliscono miseramente.

Il passaggio al secondo ordine con semantica piena offre un'alternativa ontologica dell'unicità della struttura dei naturali: la categoricità blocca la molteplicità di modelli e sembrerebbe decomporre l'oscillazione interpretativa, ma la conseguenza di tale apparente compiutezza è la rinuncia all'effettività, in quanto la verità di secondo ordine non è comprensibile da un sistema di deduzione ricorsivamente enumerabile completo e corretto, e così l'incompletezza non collassa ma passa dal livello della molteplicità strutturale a quello dell'inaccessibilità operativa del metalinguaggio necessario per definire la verità.

Questo è un realismo del contenuto: esistono enunciati con valore di verità determinato nello standard che non dipende dalla nostra capacità di provarli in T.

5.1.1 Deflazionismo, minimalismo e limiti

Si considera una posizione deflazionista secondo cui la verità è un concetto logico che non implica un impegno ontologico forte. Tuttavia, i teoremi di Tarski e Gödel mostrano che la verità non può essere ridotta a provabilità in teorie formali r. e.

Se T interpreta Q ed è consistente, allora non esiste una definizione interna di verità che corrisponda con la provabilità in T.

Nel contesto categoriale dei topoi con oggetto dei naturali, la logica interna ha forma intuizionista e permette la ricostruzione aritmetica, ma il dualismo interno-esterno mostra che la teoria interna r. e. che interpreta Q resta incompleta dal punto di vista meta esterno: l'ambiente categoriale non annulla il residuo semantico, ma ne chiarisce solo la posizione fenomenologica, distinguendo verità interna e verità esterna. Questo implica che, anche nella visione deflazionista, la verità matematica ha una dimensione che supera le capacità formali di dimostrazione.

5.2 Incompletezza come argomento anti-meccanicista

Si valuta la tesi per cui l'incompletezza limita spiegazioni puramente meccaniche del ragionare matematico.

5.2.1 Tesi meccanicista: formulazioni

Si definisce la tesi meccanicista in due forme: meccanicista e onnisciente. La variazione semantica-logica non altera l'invariante diagonale che produce l'incompletezza.

Definizione 5.1 (Competenza meccanicista)**.** *Un soggetto razionale è meccanicista se la sua competenza coincide con una teoria r. e. T che interpreta Q.*

Le teorie dei tipi sono esempi di sistemi formali meccanicisti con una sintassi stratificata dove il contrasto fra espressività e coerenza è radicale: l'impredicatività totale decade nel paradosso di Girard e porta a gerarchie di universi o dispositivi di predicatività, e tuttavia non appena un sistema tipato r. e. consistente rappresenta PR e interpreta Q, la struttura diagonale si ripresenta e nega la derivabilità della propria consistenza, mostrando che la regolazione interna non basta a garantire la coerenza. L'argomento anti-meccanicista intende mostrare che la mente umana non può essere meccanica in tal senso.

Definizione 5.2 (Competenza Π^0_1)**.** *Un soggetto razionale è Π^0_1-corretto se accetta solo affermazioni Π^0_1 vere nello standard.*

La correttezza Π^0_1 è una formalizzazione debole della capacità di riconoscere verità matematiche. Nei sistemi pratici, la coerenza dipende da assunzioni meta gerarchiche che funzionano come limiti trascendenti alla portata dimostrativa interna.

5.2.2 Disgiunzione di Gödel e forma dell'argomento

La disgiunzione di Gödel afferma che o la mente non è meccanica oppure esistono enunciati indecidibili veri che la mente non può riconoscere come tali.

Proposizione 5.3 (Schema di disgiunzione)**.** *Per ogni teoria r. e. consistente T che interpreta Q esiste una frase G_T tale che $\mathrm{True}_{\mathbb{N}}(G_T)$ e $T \nvdash G_T$. Se un soggetto razionale ha competenza non meno di T e riconosce la verità di G_T, allora la sua competenza supera T. In caso contrario, esiste un vero limite interno alla sua capacità dimostrativa.*

Dimostrazione. Per diagonalizzazione si costruisce G_T con $T \vdash G_T \leftrightarrow \neg\mathrm{Prov}_T(\ulcorner G_T \urcorner)$ e si verifica che $\mathrm{True}_{\mathbb{N}}(G_T)$ e $T \nvdash G_T$ come nel primo teorema. Se un soggetto razionale con competenza almeno quanto T riconosce la verità di $\mathrm{True}_{\mathbb{N}}(G_T)$, allora ha competenza su enunciati che superano T. Se non lo riconosce, allora G_T costituisce un vero limite interno alla sua capacità dimostrativa.

□

5.2.3 Lucas–Penrose e formalizzazione del soggetto razionale

Se la mente è un sistema formale consistente T, allora esiste G_T vero e non-dimostrabile in T. Concludere che la mente vede la verità di G_T richiede una metateoria che supera T, e l'argomento si indebolisce se il soggetto razionale reale non coincide con un sistema consistente, oppure se l'accesso al vero è limitato.

Proposizione 5.4 (Ipotesi nascoste)**.** *L'argomento richiede:*

i. identificazione del soggetto razionale con una singola T r. e. consistente.

ii. correttezza almeno Π^0_1 del soggetto razionale.

iii. consapevolezza della consistenza di T o accesso a una metateoria che la garantisce.

La non-soddisfazione di uno dei tre punti invalida la conclusione forte.

La sensibilità dell'argomento a queste ipotesi spiega la sua natura controversa.

5.2.4 Caveat epistemici

La conoscenza matematica reale è fallibile e parziale, e la conclusione anti-meccanicista forte risulta non-forzata dai teoremi, che definiscono limiti interni per sistemi formali ricorsivi. Rimane tuttavia solida la tesi di essenzialismo semantico rispetto a procedure effettive fissate.

5.2.5 Ruolo delle verità Π^0_1

Molti argomenti si riducono a verità Π^0_1. Se il soggetto razionale è Π^0_1-corretto ma non-onnisciente, la disgiunzione si applica in forma debole e mostra limiti informativi senza risultati ontologici su mente e calcolo.

5.3 Note su Putnam, Penrose, Boolos

Si richiamano posizioni note: l'anti-meccanicismo di Penrose, il realismo guidato dai modelli di Putnam, e le analisi di Boolos su verità e complessità.

5.3.1 Putnam e il realismo modellistico

Il realismo modellistico di Putnam sostiene che la verità matematica è relativa a modelli ideali. L'indefinibilità e la non-categoricità del primo ordine rafforzano l'idea che la verità non si identifichi con schemi derivazionali interni.

Proposizione 5.5 (Sottodeterminazione modellistica)**.** *Per teorie prime d'ordine non-categoriche, modelli non-isomorfi soddisfano gli stessi teoremi.*

Dimostrazione. Se T non è categorica, ammette due modelli non-isomorfi $\mathcal{M} \ncong \mathcal{N}$. Per definizione, ogni teorema di T è vero in ogni modello di T e dunque sia in $\mathcal{M}$ sia in $\mathcal{N}$. Quindi modelli non-isomorfi soddisfano gli stessi teoremi di T.

□

Il fenomeno di Löwenheim–Skolem mostra rigorosamente la distanza tra conseguenza formale e fissazione del riferimento.

5.3.2 Penrose e l'anti-meccanicismo

L'argomento di Penrose integra incompletezza e fisica quantistica per negare la meccanicità della mente. Il risultato richiede ipotesi forti sulla formalizzabilità del soggetto razionale ideale e resta fondamentalmente controverso.

Proposizione 5.6 (Vincolo metateorico)**.** *Qualunque passaggio dalla verità di G_T alla sua conoscenza da parte del soggetto razionale presuppone una metateoria che garantisce la consistenza di T. Il vincolo è una necessità del secondo teorema di Gödel e limita le conclusioni anti-meccaniciste.*

Dimostrazione. Sia G_T una frase di Gödel per T. Dal secondo teorema di Gödel, $T \nvdash \mathrm{Con}(T)$. Per inferire $\mathrm{Con}(T)$ e quindi G_T occorre assumere o uno schema di riflessione (almeno Π^0_1) per T oppure lavorare in una metateoria T' che dimostri $\mathrm{Con}(T)$. In assenza di tale assunzione esterna, il soggetto razionale identificato con T non può dimostrare internamente la conoscenza di $\mathrm{True}_{\mathbb{N}}(G_T)$.

□

5.3.3 Boolos, Berry e complessità

Le analisi di Boolos collegano fenomeni di incompletezza a vincoli di complessità e a paradossi definitori come Berry.

Teorema 5.1 (Schema stile Berry)**.** *Per ogni misura ragionevole di lunghezza sintattica, esistono enunciati veri che affermano l'incomprimibilità definitoria di certi numeri e che non sono dimostrabili in teorie aritmeticamente giuste.*

Dimostrazione. Si fissa una misura di lunghezza sintattica $\ell(\cdot)$ primitiva ricorsiva e si definisce il predicato aritmetico $\mathrm{Def}_{<k}(n)$: "esiste una formula φ con $\ell(\varphi) < k$ che definisce univocamente n". Tale predicato è aritmetizzabile.

Sia $\beta(k)$ la formula: "il più piccolo numero non-definibile da alcuna formula di lunghezza $< k$ ha valore n". Si considera la frase B_k che afferma che $\beta(k)$ denota un numero non-definibile da formule più corte di k. Se una teoria T r. e. sufficientemente forte provasse B_k per k abbastanza grande, allora B_k sarebbe essa stessa una definizione di quel numero con lunghezza $< k$, ovvero un paradosso. Dunque per k grande B_k è vera ma non-dimostrabile in T.

$\square$

Parte 3

Critica Strutturale alle Vie Anti-Gödeliane (*Pars Destruens I*)

6

La Tesi Verità-Provabilità: Analisi e Confutazione

6.1 *Proof-theoretic semantics*

Si considera la *proof-theoretic semantics* che definisce il significato dei connettivi logici tramite regole inferenziali e si esamina se tale semantica interna a una teoria aritmeticamente adeguata T possa fornire un predicato di verità totale.

Proposizione 6.1. *Per ogni teoria aritmeticamente adeguata T r. e., la proof-theoretic semantics interna a T non determina un predicato* $\mathrm{Tr}(x)$ *che soddisfi lo schema tarskiano* $\chi \leftrightarrow \mathrm{Tr}(\ulcorner\chi\urcorner)$ *per tutte le frasi χ di $\mathcal{L}$.*

Dimostrazione. Se la semantica proof-theoretic determinasse internamente a T un predicato $\mathrm{Tr}(x)$ con $T \vdash \chi \leftrightarrow \mathrm{Tr}(\ulcorner\chi\urcorner)$ per ogni frase χ, allora Tr sarebbe un predicato di verità totale interno. Il teorema di Tarski dimostrato esclude l'esistenza di tale Tr in teorie r. e. che interpretano Q. Dunque la semantica *proof-theoretic* non può identificare verità e provabilità in senso tarskiano. □

La *proof-theoretic semantics* è rilevante come *criteriologia* di regole ammissibili e di giustificazione dei connettivi, ma tuttavia non identifica verità e provabilità: rimane un essenzialismo semantico E_T individuato dagli enunciati veri e non-dimostrabili, invarianti sotto trasformazioni *proof-theoretic conservative.*

Proposizione 6.2 (Conservatività)**.** *Sia T' un'estensione ottenuta da T tramite aggiunta di regole armoniche e ammissibili che preservano l'eliminazione del* *taglio. Se l'estensione è r. e. e non introduce nuovi assiomi aritmetici, allora T' è conservativa su frasi Π^0_1 rispetto a T.*

Ne segue che le procedure *proof-theoretic* non svuotano E_T. Esse chiariscono *perché* una derivazione vale, non *che cosa* sia vero nel senso tarskiano.

6.2 Curry–Howard e la confusione ontologica

L'esistenza di un termine chiuso di un tipo non *è* la verità dell'enunciato in senso tarskiano: è un fatto sulla derivabilità nel calcolo e, se il calcolo è fortemente normalizzante, sulla terminazione dei programmi ben tipati. La verità modellistica, invece, è una relazione tra formule e strutture.

Proposizione 6.3. *La corrispondenza di Curry–Howard non fornisce un predicato interno di verità totale per l'aritmetica, né identifica verità con esistenza di termini in un sistema di tipi ragionevole.*

Dimostrazione. Se il sistema tipato ammette ricorsione generale, la normalizzazione fallisce e l'abitabilità di un tipo non coincide con la totalità computazionale. Se si impone normalizzazione forte, la classe di funzioni rappresentabili si restringe e non coincide con tutte le funzioni totali aritmetiche. In entrambi i casi, anche postulando un predicato interno $\mathrm{Tr}(x)$ che si allinei con abitazione-terminazione, il teorema di Tarski vieta che Tr soddisfi lo schema tarskiano per tutte le frasi aritmetiche. Quindi non esiste, internamente, un predicato di verità totale e la verità modellistica non si identifica con l'esistenza di termini. □

Un esempio sono le affermazioni Π^0_1 vere ma indipendenti da PA: esistono programmi che effettivamente terminano su ogni input, ossia definiscono funzioni totali, la cui totalità non è dimostrabile in PA. Ciò mostra che la sola esistenza di un termine o di un programma tipato non coincide con la verità aritmetica nello standard: la terminazione è un fatto semantico che può superare la capacità dimostrativa del sistema formale.

6.3 LEM-DNE: limiti dell'incompletezza classica

Si considera l'aritmetica di Heyting HA, che rifiuta i principi classici del terzo escluso (LEM) e della doppia negazione (DNE).

Proposizione 6.4. *Per ogni estensione r. e. consistente dell'aritmetica di Heyting che interpreta Q, esiste una frase G_T tale che* $\mathrm{True}_{\mathbb{N}}(G_T)$ *e* $T \nvdash G_T$.

Dimostrazione. In HA esistono rappresentabilità PR, lemma di diagonalizzazione e condizioni HBL adeguate. Sia G tale che $\mathrm{HA} \vdash G \leftrightarrow \neg\mathrm{Prov}_{\mathrm{HA}}(\ulcorner G \urcorner)$. Se $\mathrm{HA} \vdash G$, allora $\mathrm{HA} \vdash \mathrm{Prov}_{\mathrm{HA}}(\ulcorner G \urcorner)$, contro G. Se $\mathrm{HA} \vdash \neg G$, allora nello standard $\mathrm{Prov}_{\mathrm{HA}}(\ulcorner G \urcorner)$, ancora contro G. Quindi $\mathrm{True}_{\mathbb{N}}(G)$ e $\mathrm{HA} \nvdash G$. L'analogo del secondo teorema segue dall'impossibilità di provare Con(HA) in HA se questa è consistente.

□

L'incompletezza vi è anche in assenza di LEM e DNE: la verità aritmetica nello standard supera la capacità dimostrativa di teorie r. e. consistenti che interpretano Q.

6.4 Disgiunzione-Esistenza: limiti della costruttività

Si richiamano le proprietà di disgiunzione ed esistenza. Se T è costruttiva e dimostra $\varphi \vee \psi$, allora si può estrarre una dimostrazione di φ o di ψ. Se T dimostra $\exists x\, \varphi(x)$, allora si può estrarre un termine t tale che T dimostra $\varphi(t)$.

Proposizione 6.5 (Disgiunzione-Esistenza e incompletezza)**.** *Per ogni estensione r. e. consistente dell'aritmetica di Heyting che interpreta Q e soddisfa le proprietà di disgiunzione ed esistenza, esiste una frase G_T tale che* $\mathrm{True}_{\mathbb{N}}(G_T)$ *e* $T \nvdash G_T$.

Dimostrazione. In T vi sono rappresentabilità PR e lemmi di diagonalizzazione. Si costruisce G_T con $T \vdash G_T \leftrightarrow \neg\mathrm{Prov}_T(\ulcorner G_T \urcorner)$. Se $T \vdash G_T$, allora per HBL, $T \vdash \mathrm{Prov}_T(\ulcorner G_T \urcorner)$, contro G_T. Se $T \vdash \neg G_T$, allora nello standard $\mathrm{Prov}_T(\ulcorner G_T \urcorner)$, ancora contro G_T. Le proprietà di disgiunzione ed esistenza non forniscono una dimostrazione di G_T da $\neg\neg G_T$. Dunque $\mathrm{True}_{\mathbb{N}}(G_T)$ e $T \nvdash G_T$.

$\square$

Enunciati veri Π^0_1 possono rimanere non-dimostrabili anche in teorie che rispettano disgiunzione ed esistenza. Ad esempio, la dimostrazione di $\neg\neg G_T$ non implica la dimostrazione di G_T.

7

Superare Gödel? Varianti non-classiche e critica strutturale

7.1 Intuizionismo, rilevanza, paraconsistenza: invarianti d'incompletezza

Si analizzano tre varianti logiche: intuizionismo, logiche di rilevanza e logiche paraconsistenti. Il punto fondamentale è che gli argomenti diagonalizzanti e le rappresentazioni ricorsive non dipendono in modo essenziale dal principio del terzo escluso, ma dalla ricorsività dell'assiomatizzazione e dalla possibilità di aritmetizzare la sintassi.

Proposizione 7.1 (*Incompleteness in HA*)**.** *L'aritmetica intuizionista HA è incompleta: esiste* G_{HA} *tale che* $\mathrm{True}_N G_{\mathrm{HA}}$ *e* $\mathrm{HA} \nvdash G_{\mathrm{HA}}$*. Inoltre, se* HA *è consistente, allora* $\mathrm{HA} \nvdash \mathrm{Con}(\mathrm{HA})$*.*

Proposizione 7.2 (Rilevanza e paraconsistenza)**.** *Sia T una teoria r. e. non-banale, formulata in una logica di rilevanza o paraconsistente, che rappresenta le funzioni primitive ricorsive e interpreta Q nel senso appropriato. Allora esiste G_T tale che $T \nvdash G\_T$ e $T \nvdash \neg G\_T$.*

Dimostrazione. Si costruisce G e si rappresenta Prov_T nel linguaggio di T. Se $T \vdash G$, allora $T \vdash \mathrm{Prov}_T(\ulcorner G \urcorner)$ contro G. Se $T \vdash \neg G$, la Π^0_1-correttezza relativa nello standard fornisce $\mathrm{Prov}_T(\ulcorner G \urcorner)$, di nuovo contro G. Pertanto $T \nvdash G$ e $T \nvdash \neg G$.

$\square$

In tutti i casi, l'invariante è: teoria r. e., aritmetizzazione efficace, sufficiente forza per definire la computazione. La scelta della logica varia la metateoria e il calcolo delle dimostrazioni, non il nucleo dell'argomento diagonale.

7.2 Secondo ordine e categoricità: il costo del metalinguaggio

La PA del secondo ordine con semantica piena è categorica, ma il prezzo è l'abbandono dell'effettività: non esiste un calcolo ricorsivo completo e corretto per la validità del secondo ordine. L'incompletezza non è meccanica: si sposta nella semantica non-effettiva.

Proposizione 7.3 (Nessuna completezza effettiva). *Non esiste un sistema di dimostrazione r. e., corretto e completo per la validità piena del secondo ordine.*

Dimostrazione. Se esistesse un sistema di dimostrazione r. e., corretto e completo per la validità piena del secondo ordine, l'insieme delle validità del secondo ordine sarebbe ricorsivamente enumerabile. Ma la validità del secondo ordine con semantica piena non è r. e.: infatti, per interpretazione ovvia, ogni validità del primo ordine è validità del secondo ordine e l'insieme delle non-validità del secondo ordine non è r. e. Inoltre, i risultati classici di Mostowski e Henkin, mostrano che una completa assiomatizzazione r. e. per la semantica piena non esiste. Segue che un tale sistema non può esistere.

$\square$

La categoricità piena fissa la struttura, ma non fornisce un criterio meccanico di derivazione: l'incompletezza ricompare come indecidibilità del metalinguaggio.

7.3 Topoi e verità interna: dualismo interno-esterno

Nelle topoi elementari con oggetto dei naturali (NNO), la logica interna è intuizionista. Si può sviluppare aritmetica interna simile a HA. Tuttavia, rispetto alla metateoria esterna classica, la teoria interna r. e. che interpreta Q resta incompleta.

Proposizione 7.4 (Interno-esterno)**.** *Se una teoria aritmetica interna T^{int} in una topos con NNO è r. e. e interpreta Q, allora esiste $G_{T^{\text{int}}}$ tale che* $\text{True}_N G_{T^{\text{int}}}$*, esternamente, e* $T^{\text{int}} \nvdash G_{T^{\text{int}}}$*.*

Dimostrazione. La teoria interna T^{int} è r. e. e interpreta Q grazie all'NNO e alla struttura del topos. Dunque, per lo schema gödeliano generale, esiste $G_{T^{\text{int}}}$ tale che esternamente $\text{True}_{\mathbb{N}}(G_{T^{\text{int}}})$ e internamente $T^{\text{int}} \nvdash G_{T^{\text{int}}}$. La costruzione usa la codifica sintattica nel metalinguaggio esterno e la rappresentazione PR garantita dall'NNO.

□

La distinzione tra verità interna, relativa alla logica del *topos*, e verità esterna, meta-giudizio sulla totalità, mostra che l'incompletezza si manifesta come differenza tra il sistema formale interno e la sua osservazione esterna.

7.4 *Type theory* e Girard: coerenza ed espressività

Le teorie dei tipi sono sistemi formali stratificati che combinano logica e teoria delle categorie. Sistemi con universi impredicativi totali portano al paradosso di Girard, ma sistemi stratificati mantengono coerenza.

Proposizione 7.5 (Gödel II per teorie dei tipi)**.** *Ogni teoria dei tipi r. e., consistente e in grado di rappresentare PR e interpretare Q non dimostra la propria consistenza.*

Nei sistemi pratici, la coerenza è data da stratificazioni e ipotesi meta, ma non è derivabile internamente in modo totale senza superare i limiti gödeliani.

Parte 4

Critica Meta-Logica e Generalizzazioni (Pars Destruens II)

8

Generalizzazioni dell'Incompletezza

8.1 Schema gödeliano generale e secondo teorema astratto

L'incompletezza non è un errore specifico di PA o della logica classica, ma un'invariante metastrutturale: ogni teoria T ricorsivamente enumerabile che interpreta Q e rappresenta le funzioni primitive ricorsive possiede i mezzi per aritmetizzare la propria sintassi e dunque per costruire un enunciato che, affermando la propria non-dimostrabilità in T, definisce la differenza fra codice e contenuto come tratto semantico irriducibile.

Proposizione 8.1 (Schema Gödeliano Generale)**.** *Per ogni teoria r. e. consistente T che interpreta Q esiste una frase G_T tale che* $\mathrm{True}_{\mathbb{N}}(G_T)$ *e* $T \nvdash G_T$*. Se T è Π^0_1-corretta allora $T \nvdash \neg G_T$.*

Dimostrazione. Sia Prov_T Σ^0_1. Per diagonalizzazione, $T \vdash G_T \leftrightarrow \neg\mathrm{Prov}_T(\ulcorner G_T \urcorner)$. Se $T \vdash G_T$ allora $T \vdash \mathrm{Prov}_T(\ulcorner G_T \urcorner)$, che è un assurdo. Se T è Π^0_1-corretta e $T \vdash \neg G_T$, allora nello standard $\mathrm{Prov}_T(\ulcorner G_T \urcorner)$, che è anch'esso un assurdo. Quindi $\mathrm{True}_{\mathbb{N}}(G_T)$, $T \nvdash G_T$ e, se T è Π^0_1-corretta, $T \nvdash \neg G_T$.

□

Proposizione 8.2 (Secondo Teorema, forma astratta)**.** *Se T è r. e., consistente e interpreta Q, allora $T \nvdash \mathrm{Con}(T)$. Se $T \vdash \mathrm{Con}(T)$, allora T è inconsistente.*

Dimostrazione. Formalizzando $\mathrm{Con}(T)$ come $\neg\mathrm{Prov}_T(\ulcorner\bot\urcorner)$, se $T \vdash \mathrm{Con}(T)$ allora per Löb interno si ottiene $T \vdash \mathrm{Prov}_T(\ulcorner\mathrm{Con}(T)\urcorner)$ e, combinando con G_T, si arriva a una contraddizione con la consistenza. Dunque $T \nvdash \mathrm{Con}(T)$. Se invece $T \vdash \mathrm{Con}(T)$, T non è consistente.

□

L'incompletezza vi è in ogni teoria r. e. che interpreta Q e rappresenta PR: la struttura diagonale è un'invariante metastrutturale.

8.2 Löb, GL e logiche della provabilità

La logica modale della provabilità (GL) formalizza il comportamento del predicato Prov_T sotto condizioni di derivabilità: l'assioma di Löb $\Box(\Box p \to p) \to \Box p$ non è un prodotto della logica modale, ma la formalizzazione della verticalità riflessiva per cui, se una teoria riconosce che la dimostrabilità di p implica p, allora ha già la possibilità di dimostrare p stessa.

Proposizione 8.3 (Correttezza provabilistica)**.** *Se $\vdash_{\mathsf{GL}} \varphi$, allora per ogni T r. e. che soddisfa le condizioni di Hilbert–Bernays–Löb vale $T \vdash \varphi^*$.*

Dimostrazione. Si fissa un'interpretazione aritmetica $*$ che manda $\Box$ in Prov_T e interpreta proposizionali e connettivi in modo omomorfo. Si dimostra per induzione sulla lunghezza delle derivazioni in GL. Per gli assiomi proposizionali, φ è banalmente vero nello standard, dunque $T \vdash \varphi^*$ in quanto T dimostra tutte le forme veritative proposizionali. Per la regola di *modus ponens*, dall'ipotesi induttiva su $\psi \to \chi$ e ψ si ottiene per deduzione in T che $T \vdash \chi$ e quindi $T \vdash \chi^*$. Per la componente modale, si usano le condizioni HBL interne a T:

i. $T \vdash \mathrm{Prov}_T(\ulcorner\alpha \to \beta\urcorner) \to (\mathrm{Prov}_T(\ulcorner\alpha\urcorner) \to \mathrm{Prov}_T(\ulcorner\beta\urcorner))$.

ii. $T \vdash \mathrm{Prov}_T(\ulcorner\alpha\urcorner) \to \mathrm{Prov}_T(\ulcorner\mathrm{Prov}_T(\ulcorner\alpha\urcorner)\urcorner)$.

iii. se $T \vdash \alpha$ allora $T \vdash \mathrm{Prov}_T(\ulcorner\alpha\urcorner)$.

La regola di necessitazione di GL è data da (iii). L'assioma di K, $\Box(p \to q) \to (\Box p \to \Box q)$, è derivabile da (i). L'assioma di Löb, $\Box(\Box p \to p) \to \Box p$, è derivabile da (ii) e (iii). Dunque, per induzione, ogni teorema di GL diventa un teorema di T sotto l'interpretazione aritmetica $*$.

□

Proposizione 8.4 (Completezza di Solovay)**.** *Le formule modali valide sotto interpretazione provabilistica sono esattamente i teoremi di GL.*

Dimostrazione. Segue dal Teorema di Solovay: correttezza e completezza aritmetica equivalgono esattamente ai teoremi di GL.

$\square$

La gerarchia $\square^n p$ riflette i livelli di riflessione: le catene di teorie $T_n = T + \mathrm{RFN}^n(T)$ corrispondono a livelli di potenza dimostrativa crescente, ma nessun livello raggiunge la completezza.

8.3 Interpretabilità e riflessione: Feferman e Montagna

L'interpretabilità tra teorie fornisce un criterio di confronto della forza dimostrativa: $T \rhd S$ indica che T contiene un'immagine strutturale di S e permette di trasportare risultati di consistenza e riflessione. Le progressioni di riflessione di Feferman definiscono catene T_α che aumentano la forza senza mai raggiungere la completezza.

Proposizione 8.5 (Progressioni riflessive)**.** *Dato $T_0 = T$, si definiscono $T_{\alpha+1} = T_\alpha + \mathrm{RFN}(T_\alpha)$ e $T_\lambda = \bigcup_{\beta<\lambda} T_\beta$ per λ limite. Ogni livello è più forte del precedente ma per nessun ordinale ricorsivo α si ottiene completezza.*

Dimostrazione. Si dimostra per induzione:

i. Passo base: $T_0 = T$ è r. e. e consistente per ipotesi.

ii. Passo induttivo: se T_α è r. e. e consistente, allora $\mathrm{RFN}(T_\alpha)$ è r. e. e $T_{\alpha+1}$ è r. e. Inoltre, per il secondo teorema di Gödel, $T_{\alpha+1}$ è più forte di T_α ma non completo.

iii. Passo limite: se λ è limite e ogni T_β per $\beta < \lambda$ è r. e. e consistente, allora T_λ è l'unione di r. e. consistenti teorie, quindi è r. e. e consistente. Inoltre, T_λ non è completo poiché ogni T_β non lo è.

Dunque, per ogni ordinale ricorsivo α, T_α è r. e., consistente e non completo.

$\square$

Proposizione 8.6 (Interpretabilità e consistenza)**.** *Se $T \rhd S$ e T è consistente, allora T dimostra un'interpretazione fedele di* $\mathrm{Con}(S)$*. Se l'interpretazione preserva verità* Π^0_1*, T simula la forza riflessiva di S.*

Dimostrazione. Se $T \rhd S$ tramite τ, da $S \vdash \perp$ seguirebbe $T \vdash \tau(\perp)$. Dunque, se T è consistente, $T \vdash \tau\Big(\neg\mathrm{Prov}_S(\ulcorner \perp \urcorner)\Big)$. Se l'interpretazione preserva verità Π^0_1 vale, allora $T \vdash \mathrm{Con}(S)$ in senso forte e T riproduce la forza riflessiva di S.

□

La riflessione e l'interpretabilità creano approssimazioni direzionate alla verità standard, ma il limite semantico supera ogni livello di potenza dimostrativa.

9

Indefinibilità e Gerarchie

9.1 Tarski: nessun predicato di verità interno

L'identificazione della verità con la provabilità fallisce nel modo più totale: non esiste predicato aritmetico interno che soddisfi lo schema tarskiano per tutte le frasi del linguaggio, e dunque la verità si mostra come condizione meta che supera ogni formalizzazione interna.

Teorema 9.1 (Tarski, Indefinibilità)**.** *Per ogni teoria aritmeticamente giusta T r. e., non esiste un predicato* $\mathrm{Tr}(x)$ *definibile in T tale che per ogni frase χ valga* $\chi \leftrightarrow \mathrm{Tr}(\ulcorner\chi\urcorner)$ *nello standard.*

Dimostrazione. Per assurdo, si suppone definibile $\mathrm{Tr}(x)$ con $T \vdash \chi \leftrightarrow \mathrm{Tr}(\ulcorner\chi\urcorner)$ per ogni frase χ. Per il lemma di diagonalizzazione, esiste ψ tale che $T \vdash \psi \leftrightarrow \neg\mathrm{Tr}(\ulcorner\psi\urcorner)$. Applicando lo schema tarskiano a ψ si ottiene anche $T \vdash \psi \leftrightarrow \mathrm{Tr}(\ulcorner\psi\urcorner)$. Combinando le due equivalenze si ricava $T \vdash \mathrm{Tr}(\ulcorner\psi\urcorner) \leftrightarrow \neg\mathrm{Tr}(\ulcorner\psi\urcorner)$ e quindi $T \vdash \bot$, contro la consistenza. Dunque un tale $\mathrm{Tr}(x)$ non esiste in T.

$\square$

La verità non è comprimibile in un predicato interno: essa è una condizione che trascende ogni formalizzazione interna.

9.2 Kripke e i punti fissi parziali

Kripke mostra che predicati di verità parziale possono essere costruiti come punti fissi di operatori monotoni sulla struttura dei sottoinsiemi di frasi: si ottiene una semantica stratificata che evita la contraddizione considerando zone di indeterminatezza.

Proposizione 9.1 (Punto fisso)**.** *Dato un operatore monotono* Φ *su una struttura completa* $\mathcal{L}$*, esistono minimo e massimo punto fisso. Per l'operatore semantico della verità si ottiene un predicato parziale* Tr_∞.

Dimostrazione. Per il teorema di Knaster–Tarski, ogni operatore monotono su una struttura completa ha punti fissi e il minimo punto fisso è dato da $\mathrm{lfp}(\Phi) = \bigvee\{X \in \mathcal{L} : \Phi(X) \leq X\}$. L'operatore semantico Φ che, dato un'assegnazione parziale di verità, ne calcola l'espansione secondo le condizioni composizionali è monotono. L'iterazione transfinita $\Phi^0(\bot) = \bot$, $\Phi^{\alpha+1}(\bot) = \Phi(\Phi^\alpha(\bot))$, $\Phi^\lambda(\bot) = \bigvee_{\alpha<\lambda} \Phi^\alpha(\bot)$ per λ limite converge al minimo punto fisso Tr_∞, che definisce un predicato di verità parziale.

□

La verità parziale indica che l'alternativa al paradosso è l'accettazione dell'incompletezza semantica controllata.

9.3 Verità esterna e modelli non-standard

Per teorie del primo ordine non-categoriche, esistono modelli non-isomorfi che soddisfano gli stessi teoremi. La verità dipende dal modello e la provabilità interna non determina univocamente il contenuto. La verità esterna allo standard funziona come criterio di realtà semantica non riducibile al puro calcolo.

Proposizione 9.2 (Sottodeterminazione modellistica)**.** *Se* T *non è categorica nel primo ordine, allora esistono* $\mathcal{M} \not\cong \mathcal{N}$ *con* $\mathcal{M} \equiv \mathcal{N}$*. I teoremi di* T *non individuano una sola struttura.*

Dimostrazione. Sia T' un completamento coerente di T. Poiché T non è categorica, esistono modelli non-isomorfi di T'. Due qualunque modelli $\mathcal{M}, \mathcal{N} \models$

T' sono elementarmente equivalenti per completezza di T', ma non-isomorfi per ipotesi. Di conseguenza $\mathcal{M} \equiv \mathcal{N}$ e $\mathcal{M} \ncong \mathcal{N}$.

$\square$

La verità modellistica dipende dalla struttura considerata: la provabilità interna non determina univocamente il contenuto semantico.

Parte 5

Ricostruzione Trascendentale della Verità (*Pars Construens*)

10

Dalla Provabilità alla Verità

10.1 Separare semantica e metateoria della dimostrazione

Si fissa il linguaggio aritmetico $\mathcal{L}$ e l'insieme delle frasi $\mathrm{Sent}(\mathcal{L})$. Sia T una teoria r. e. che interpreta Q e rappresenta le funzioni primitive ricorsive. Si utilizza la notazione $\ulcorner \cdot \urcorner$ per il codice di Gödel e Prov_T per il predicato di dimostrabilità interno standard.

Definizione 10.1 (Schema composizionale interno). *Un* predicato interno di verità composizionale *per T è una coppia* $(\mathrm{Tr}(x), \Gamma)$ *dove* $\mathrm{Tr}(x)$ *è una formula di $\mathcal{L}$ e $\Gamma \subseteq \mathrm{Sent}(\mathcal{L})$ è chiusa per connettivi e quantificatori, tale che in T valgano le condizioni composizionali su Γ e per ogni $\varphi \in \Gamma$ si abbia*

$$T \vdash \varphi \leftrightarrow \mathrm{Tr}(\ulcorner \varphi \urcorner).$$

Si richiede inoltre correttezza Σ_1^0: per ogni Σ_1^0-frase χ, $T \vdash \chi \to \mathrm{Tr}(\ulcorner \chi \urcorner)$.

Definizione 10.2 (Invarianza interpretazionale). *Un predicato* $\mathrm{Tr}(x)$ *definibile in T è* interpretazionalmente invariante *se per ogni due modelli $\mathcal{M}, \mathcal{N} \models T$ e per ogni interpretazione aritmetica τ di T in T si ha, per ogni $\varphi \in \mathrm{Sent}(\mathcal{L})$,*

$$(\mathcal{M}, \mathrm{Tr}^{\mathcal{M}}) \models \mathrm{Tr}(\ulcorner \varphi \urcorner) \quad \textit{se e solo se} \quad (\mathcal{N}, \mathrm{Tr}^{\mathcal{N}}) \models \mathrm{Tr}(\ulcorner \tau(\varphi) \urcorner).$$

In particolare, l'estensione di Tr *non dipende dal modello né dalla scelta di una interpretazione interna di T.*

L'obiettivo trascendentale è mostrare che la verità semantica soddisfa condizioni che nessun predicato interno può soddisfare, mantenendo così la distinzione tra verità e provabilità.

10.2 Verità come condizione del giudizio epistemico

Si formalizza il collegamento normativo fra verità e giudizio come relazione fra teorie con predicato di verità composizionale e insiemi di frasi giudicate vere.

Definizione 10.3 (Politica di giudizio)**.** *Sia T r. e., consistente e tale da interpretare Q. Una* politica di giudizio *per T è una teoria S nel linguaggio $\mathcal{L} \cup \{\mathrm{Tr}\}$ tale che:*

i. *S estende T e contiene uno schema composizionale per* Tr *su una classe $\Gamma \supseteq \mathrm{Sent}(\mathcal{L})$ chiusa per connettivi e quantificatori con bicondizionali $\varphi \leftrightarrow \mathrm{Tr}(\ulcorner \varphi \urcorner)$ per ogni $\varphi \in \Gamma$.*

ii. *S è conservativa su $\mathcal{L}$, cioè se $S \vdash \chi$ con $\chi \in \mathrm{Sent}(\mathcal{L})$, allora $T \vdash \chi$.*

iii. *Si definisce il* giudizio indotto *da S come l'insieme*

$$J_S \; := \; \{\, \chi \in \mathrm{Sent}(\mathcal{L}) : S \vdash \mathrm{Tr}(\ulcorner \chi \urcorner) \,\}.$$

Proposizione 10.1 (Allineamento derivativo)**.** *Per ogni politica S e per ogni $\varphi \in \Gamma$, vale $S \vdash \mathrm{Tr}(\ulcorner \varphi \urcorner)$ se e solo se $S \vdash \varphi$. In particolare, $J_S \cap \Gamma = \{\varphi \in \Gamma : S \vdash \varphi\}$. Se S è conservativa su $\mathcal{L}$, allora $J_S \cap \mathrm{Sent}(\mathcal{L}) = \{\chi \in \mathrm{Sent}(\mathcal{L}) : T \vdash \chi\}$.*

Dimostrazione. La prima equivalenza è per bicondizionale composizionale e logica classica in S. L'uguaglianza segue banalmente. La seconda affermazione è conseguenza della conservatività: per χ senza Tr, $S \vdash \chi$ implica $T \vdash \chi$ e quindi caratterizza J_S sulle frasi aritmetiche.

□

Corollario 10.1 (Non-determinazione su indipendenti)**.** *Se σ è indipendente da T, allora esistono politiche S_1, S_2 soddisfacenti le condizioni sopra tali che $\sigma \in J_{S_1}$ e $\sigma \notin J_{S_2}$. In particolare, il giudizio interno non determina il valore di verità di σ.*

Dimostrazione. Per il lemma di conservatività e consistenza di $T+\sigma$ e $T+\neg\sigma$, le teorie $S+\sigma$ e $S+\neg\sigma$ sono consistenti per ogni S conservativa. Per bicondizionali composizionali, in un modello di $S+\sigma$ si ha $\mathrm{Tr}(\ulcorner\sigma\urcorner)$ e in un modello di $S+\neg\sigma$ si ha $\neg\mathrm{Tr}(\ulcorner\sigma\urcorner)$. Quindi esistono politiche che includono σ e politiche che lo escludono.

□

Teorema 10.1 (Limite normativo interno)**.** *Non esiste politica S r. e., composizionale e conservativa su $\mathcal{L}$ per cui J_S coincida con l'insieme delle frasi vere nello standard $\{\chi \in \mathrm{Sent}(\mathcal{L}) : \mathrm{True}_{\mathbb{N}}(\chi)\}$. In particolare, per ogni tale S esiste χ con $\mathrm{True}_{\mathbb{N}}(\chi)$ ma $\chi \notin J_S$.*

Dimostrazione. Per la proposizione, su frasi aritmetiche J_S coincide con $\{\chi : T \vdash \chi\}$. Per incompletezza esistono frasi vere χ con $T \nvdash \chi$. Tali frasi non appartengono a J_S.

□

10.3 Riconoscimento intersoggettivo del valore di verità

Si formalizza l'accordo inter-soggettivo come condizione di coerenza riflessiva tra politiche di giudizio diverse.

Definizione 10.4 (Accordo su base riflessiva)**.** *Siano S_A, S_B politiche per T come sopra. Si dice che S_A e S_B sono* coerenti con riflessione Π^0_1 *se entrambe estendono $T + \mathrm{RFN}_{\Pi^0_1}(T)$. Si definisce il* dominio di accordo derivativo *come*

$$D_{\Pi^0_1}(T) \ := \ \mathrm{Th}_{\Pi^0_1}\Big(T + \mathrm{RFN}_{\Pi^0_1}(T)\Big).$$

Proposizione 10.2 (Accordo minimo su $D_{\Pi^0_1}(T)$)**.** *Se S_A e S_B estendono $T + \mathrm{RFN}_{\Pi^0_1}(T)$ e sono conservative su $\mathcal{L}$, allora per ogni $\chi \in D_{\Pi^0_1}(T)$ si ha $\chi \in J_{S_A} \cap J_{S_B}$. Inoltre, per ogni $\chi \in D_{\Pi^0_1}(T)$ non si ha $\neg\chi \in J_{S_A} \cup J_{S_B}$.*

Dimostrazione. Se $\chi \in D_{\Pi^0_1}(T)$, allora $T + \mathrm{RFN}_{\Pi^0_1}(T) \vdash \chi$. Per estensione, $S_i \vdash \chi$ per $i \in \{A, B\}$ e per bicondizionali composizionali $S_i \vdash \mathrm{Tr}(\ulcorner\chi\urcorner)$, quindi

$\chi \in J_{S_i}$. Non può valere $\neg\chi \in J_{S_i}$ in quanto ciò darebbe $S_i \vdash \neg\chi$ e, per conservatività su $\mathcal{L}$, anche $T \vdash \neg\chi$, che contraddice $T + \mathrm{RFN}_{\Pi^0_1}(T) \vdash \chi$.

$\square$

Proposizione 10.3 (Limite dell'accordo)**.** *Se σ è indipendente da T, esistono politiche conservative S_A, S_B tali che $\sigma \in J_{S_A}$ e $\neg\sigma \in J_{S_B}$. Quindi l'accordo non si estende oltre* $\mathrm{Th}(T)$ *se non si fanno ipotesi esterne sulla verità.*

Dimostrazione. Segue banalmente dalla non-determinazione su indipendenti e dalla costruzione di modelli di $S + \sigma$ e $S + \neg\sigma$ combinata con bicondizionali composizionali.

$\square$

L'accordo intersoggettivo forte richiede ipotesi esterne sulla verità, come la correttezza Π^0_1 dello standard. Tali ipotesi sono invarianti trascendentali ma non sono definibili internamente come predicato totale, in accordo con i risultati di indecidibilità e con il teorema di indefinibilità.

11

Fenomenologia del Giudizio e Fondamento del Vero

11.1 Lemma di Bufacchi: enunciato e conseguenze

Lemma 11.1 (Ipotesi minimali di composizionalità). *Sia* $(\mathrm{Tr}(x), \Gamma)$ *uno schema composizionale interno per* T*. Non è necessario assumere* $\Gamma = \mathrm{Sent}(\mathcal{L})$*. È sufficiente che esista* $\psi \in \Gamma$ *e una frase aritmetica* σ *indipendente da* T *tali che* $T \vdash \psi \leftrightarrow \sigma$*. Allora, per i modelli* $\mathcal{M} \models T + \sigma$ *e* $\mathcal{N} \models T + \neg\sigma$*, si ha*

$$(\mathcal{M}, \mathrm{Tr}^{\mathcal{M}}) \models \mathrm{Tr}(\ulcorner\psi\urcorner) \quad e \quad (\mathcal{N}, \mathrm{Tr}^{\mathcal{N}}) \models \neg\mathrm{Tr}(\ulcorner\psi\urcorner).$$

In particolare, l'invarianza interpretazionale fallisce già con $\tau = \mathrm{id}$*.*

Dimostrazione. Per bicondizionali composizionali su Γ, in ogni modello di T vale $\mathrm{Tr}(\ulcorner\psi\urcorner) \leftrightarrow \psi$. Per $T \vdash \psi \leftrightarrow \sigma$ e correttezza semantica, in $\mathcal{M}$ si ha ψ e in $\mathcal{N}$ si ha $\neg\psi$. Quindi le asserzioni su Tr seguono e l'invarianza fallisce.

□

Proposizione 11.1 (Non-invarianza uniforme). *Sia* T *r. e., consistente e tale da interpretare* Q*. Per ogni predicato definibile* $\mathrm{Tr}(x)$ *composizionale con bicondizionali su* $\Gamma \supseteq \mathrm{Sent}(\mathcal{L})$ *e per ogni interpretazione aritmetica interna* τ *di* T *in* T*, esiste una frase aritmetica* ρ *indipendente da* T *e modelli* $\mathcal{M}, \mathcal{N} \models T$

tali che

$$(\mathcal{M}, \mathrm{Tr}^{\mathcal{M}}) \models \mathrm{Tr}(\ulcorner \rho \urcorner) \quad e \quad (\mathcal{N}, \mathrm{Tr}^{\mathcal{N}}) \models \neg\mathrm{Tr}(\ulcorner \rho \urcorner).$$

Di conseguenza, nessun Tr *così definito è interpretazionalmente invariante in senso uniforme (rispetto a tutti i modelli e a tutte le interpretazioni interne).*

Dimostrazione. Si fissa ρ indipendente da T e si prendono $\mathcal{M} \models T + \rho$ e $\mathcal{N} \models T + \neg\rho$. Per i bicondizionali su tutte le frasi, in $\mathcal{M}$ vale $\mathrm{Tr}(\ulcorner \rho \urcorner) \leftrightarrow \rho$ e in $\mathcal{N}$ vale $\mathrm{Tr}(\ulcorner \rho \urcorner) \leftrightarrow \rho$. Dunque $\mathrm{Tr}(\ulcorner \rho \urcorner)$ ha valori opposti in $\mathcal{M}$ e $\mathcal{N}$. L'invarianza fallisce già per $\tau = \mathrm{id}$ e pertanto non può valere in forma uniforme per interpretazioni arbitrarie.

□

Lemma 11.2 (Lemma di Bufacchi — Indefinibilità forte e non-invarianza uniforme della verità)**.** [1]

Sia T r. e., consistente e tale da interpretare Q. Non esiste un predicato $\mathrm{Tr}(x)$ *di $\mathcal{L}$ definibile in T che sia contemporaneamente:*

i. Composizionale su una classe Γ chiusa per connettivi e quantificatori con $\mathrm{Sent}(\mathcal{L}) \subseteq \Gamma$.

ii. Tale che $T \vdash \varphi \leftrightarrow \mathrm{Tr}(\ulcorner \varphi \urcorner)$ per ogni $\varphi \in \Gamma$.

iii. Interpretazionalmente invariante.

Inoltre, valgono le seguenti generalizzazioni:

a. (Ipotesi minimali) Non è necessario assumere $\Gamma = \mathrm{Sent}(\mathcal{L})$: è sufficiente che Γ contenga ψ con $T \vdash \psi \leftrightarrow \sigma$ per qualche σ indipendente da T (11.1).

b. (Forma uniforme) Per ogni interpretazione aritmetica interna τ di T in T, esiste una frase aritmetica ρ indipendente da T e modelli $\mathcal{M}, \mathcal{N} \models T$ tali che $\mathrm{Tr}(\ulcorner \rho \urcorner)$ *assume valori opposti in $\mathcal{M}$ e $\mathcal{N}$ (11.1).*

Dimostrazione. Si assume per assurdo l'esistenza di $\mathrm{Tr}(x)$ con le proprietà (i)–(iii). Per incompletezza di Gödel applicata a T, esiste una frase aritmetica σ tale che $T \nvdash \sigma$ e $T \nvdash \neg\sigma$. Per completezza semantica, esistono modelli

[1] Esso è una riformulazione dell'indefinibilità di Tarski con le indipendenze modellistiche. La dimostrazione utilizza frasi indipendenti e bicondizionali composizionali. Le generalizzazioni (a) e (b) sono invece pure teorizzazioni.

$\mathcal{M} \models T + \sigma$ e $\mathcal{N} \models T + \neg\sigma$. Per (ii) e (i), in $\mathcal{M}$ si ha

$$(\mathcal{M}, \mathrm{Tr}^{\mathcal{M}}) \models \mathrm{Tr}(\ulcorner \sigma \urcorner) \leftrightarrow \sigma,$$

mentre in $\mathcal{N}$ si ha

$$(\mathcal{N}, \mathrm{Tr}^{\mathcal{N}}) \models \mathrm{Tr}(\ulcorner \sigma \urcorner) \leftrightarrow \sigma.$$

Poiché $\mathcal{M} \models \sigma$ e $\mathcal{N} \models \neg\sigma$, ne segue

$$(\mathcal{M}, \mathrm{Tr}^{\mathcal{M}}) \models \mathrm{Tr}(\ulcorner \sigma \urcorner) \quad \text{e} \quad (\mathcal{N}, \mathrm{Tr}^{\mathcal{N}}) \models \neg\mathrm{Tr}(\ulcorner \sigma \urcorner).$$

Ma allora (iii) è violata anche per l'interpretazione identica $\tau = \mathrm{id}$, giacché i due modelli di T assegnano valori diversi a $\mathrm{Tr}(\ulcorner \sigma \urcorner)$ pur soddisfacendo le stesse condizioni assiomatiche. L'assunto iniziale è di conseguenza impossibile.

□

Corollario 11.1 (Non-unicità modellistica dei predicati di verità interni)**.** *Sia T come già definito e $(\mathrm{Tr}(x), \Gamma)$ uno schema composizionale interno con Γ chiusa sotto connettivi. Allora esistono modelli $\mathcal{M}, \mathcal{N} \models T$ e una formula $\psi \in \Gamma$ tali che*

$$(\mathcal{M}, \mathrm{Tr}^{\mathcal{M}}) \models \mathrm{Tr}(\ulcorner \psi \urcorner) \quad \textit{mentre} \quad (\mathcal{N}, \mathrm{Tr}^{\mathcal{N}}) \models \neg\mathrm{Tr}(\ulcorner \psi \urcorner).$$

In particolare, le condizioni composizionali interne non determinano univocamente la verità per T, in quanto il valore dipende dall'interpretazione sottostante.

Dimostrazione. Sia σ indipendente per T e siano $\mathcal{M} \models T + \sigma$ e $\mathcal{N} \models T + \neg\sigma$ i modelli corrispondenti. Poiché Γ è chiusa per connettivi, contiene una formula logicamente equivalente a σ, come ad esempio $\sigma \wedge (0 = 0)$. La tesi segue banalmente dalle bicondizionali composizionali in $\mathcal{M}$ e $\mathcal{N}$.

□

Proposizione 11.2 (Invarianza Π^0_1 esterna)**.** *Sia T r. e. e Π^0_1-corretta nello standard. Per ogni Π^0_1-frase χ vera nello standard si ha $T \nvdash \neg\chi$ e, per ogni modello standard $\mathbb{N}$, $\mathrm{True}_{\mathbb{N}}(\chi)$ indipendentemente dall'interpretazione interna. Quindi la* verità esterna Π^0_1 *è un'invariante trascendentale rispetto a interpretazioni e completamenti di T.*

Dimostrazione. Per Π^0_1-correttezza esterna, se χ è vera nello standard, allora nessuna teoria T corretta su Π^0_1 può dimostrare $\neg\chi$. L'affermazione su $\mathbb{N}$ è banale: la verità esterna nello standard non dipende da alcuna scelta di predicati interni.

□

Si ottiene così una ricostruzione in due livelli: (a) interno, dove si caratterizzano condizioni composizionali e riflessive come strumenti di *giustificazione* inferenziale senza identificazione con la verità totale. (b) esterno, dove la verità agisce come criterio trascendentale, invariante almeno sulle classi Π^0_1, che fonda il giudizio senza essere definibile internamente.

11.1.1 Estensioni composizionali conservative e non-determinazione

Lemma 11.3 (Conservatività che implica la non-determinazione)**.** *Sia S una estensione consistente di T nel linguaggio $\mathcal{L} \cup \{\mathrm{Tr}\}$ tale che per ogni frase $\chi \in \mathrm{Sent}(\mathcal{L})$ vale il bicondizionale composizionale*

$$S \vdash \chi \leftrightarrow \mathrm{Tr}(\ulcorner\chi\urcorner).$$

Se S è conservativa su $\mathcal{L}$, allora per ogni σ indipendente da T si hanno le seguenti proprietà:

i. $S + \sigma$ è consistente.

ii. $S + \neg\sigma$ è consistente.

Dimostrazione. Per assurdo, si suppone che $S + \sigma$ sia inconsistente. Allora $S \vdash \neg\sigma$ e per bicondizionale composizionale $S \vdash \neg\mathrm{Tr}(\ulcorner\sigma\urcorner)$. Per conservatività su $\mathcal{L}$, si ha $T \vdash \neg\sigma$, che è un assurdo per l'indipendenza di σ. Analogamente, se si suppone che $S + \neg\sigma$ sia inconsistente, si ottiene $T \vdash \sigma$, che è anch'esso un assurdo per la stessa condizione.

□

Corollario 11.2 (CT^-, KF minimale)**.** *Se S è una teoria di verità composizionale* conservativa *su T per frasi di $\mathcal{L}$, come ad esempio una versione debole tipo CT^- o un punto fisso minimale alla Kripke su base aritmetica, allora per ogni*

σ indipendente da T esistono modelli $\mathcal{M}, \mathcal{N} \models S$ tali che $(\mathcal{M}, \mathrm{Tr}^{\mathcal{M}}) \models \mathrm{Tr}(\ulcorner\sigma\urcorner)$ e $(\mathcal{N}, \mathrm{Tr}^{\mathcal{N}}) \models \neg\mathrm{Tr}(\ulcorner\sigma\urcorner)$. Quindi lo schema composizionale non determina univocamente il valore di verità di σ.

Dimostrazione. Per il Lemma 11.3, $S + \sigma$ e $S + \neg\sigma$ sono consistenti, dunque ammettono modelli $\mathcal{M}$ e $\mathcal{N}$. Le tesi su Tr seguono dalla bicondizionale composizionale per σ nei rispettivi modelli.

□

11.1.2 Formulazione categoriale dell'errore di fondamento

Proposizione 11.3 (Non-invarianza lungo i punti di un topos)**.** *Sia $\mathcal{E}$ un topos con oggetto dei naturali e sia T^{int} la teoria interna aritmetica r. e. che interpreta Q. Siano $p, q : \mathcal{E} \to \mathbf{Set}$ due punti tali che i pullback $p^*(T^{\mathrm{int}})$ e $q^*(T^{\mathrm{int}})$ diano modelli classici $\mathcal{M}, \mathcal{N}$ con $\mathcal{M} \models T + \sigma$ e $\mathcal{N} \models T + \neg\sigma$ per qualche σ indipendente da T. Se* Tr *è un predicato interno composizionale per T^{int}, allora*

$$p^*(\mathrm{Tr})(\ulcorner\sigma\urcorner) \quad e \quad q^*(\mathrm{Tr})(\ulcorner\sigma\urcorner)$$

assegnano valori differenti. Quindi Tr *non è invariante rispetto ai punti e non può essere un criterio trascendentale di verità.*

Dimostrazione. Per ipotesi sui punti, $p^*(T^{\mathrm{int}})$ e $q^*(T^{\mathrm{int}})$ corrispondono a modelli classici di T che determinano la verità di σ in modo opposto. La composizionalità interna pone i bicondizionali lungo i *pullback*. Pertanto i valori di $p^*(\mathrm{Tr})(\ulcorner\sigma\urcorner)$ e $q^*(\mathrm{Tr})(\ulcorner\sigma\urcorner)$ devono riflettere le verità dei rispettivi modelli, che sono opposte. Dunque, l'invarianza fallisce.

□

11.2 Fondamento non-fondato e performatività

Si caratterizza il *fondamento interno* come metodo di giustificazione del valore di verità che sia composizionale e interpretazionalmente invariante.

Definizione 11.1 (Fondamento interno)**.** *Sia T r. e., consistente e tale da interpretare Q. Un* fondamento interno per *T è una coppia* $(\mathrm{Tr}(x), \Gamma)$*, con*

$\mathrm{Tr}(x)$ *definibile in* T *e* $\Gamma \subseteq \mathrm{Sent}(\mathcal{L})$ *chiusa per connettivi e quantificatori, tale che:*

i. (Composizionalità) in T *valgono le bicondizionali tarskiane per ogni* $\varphi \in \Gamma$.

ii. (Invarianza) Per ogni due modelli $\mathcal{M}, \mathcal{N} \models T$ *e per ogni interpretazione aritmetica interna* τ *di* T *in* T *si ha*

$$(\mathcal{M}, \mathrm{Tr}^{\mathcal{M}}) \models \mathrm{Tr}(\ulcorner \varphi \urcorner) \iff (\mathcal{N}, \mathrm{Tr}^{\mathcal{N}}) \models \mathrm{Tr}(\ulcorner \tau(\varphi) \urcorner).$$

Proposizione 11.4 (Impossibilità dell'auto-fondazione)**.** *Per nessuna teoria* T *esiste un fondamento interno. In particolare, non esiste* $\mathrm{Tr}(x)$ *definibile in* T *che soddisfi contemporaneamente composizionalità e invarianza interpretazionale.*

Dimostrazione. Segue direttamente dal Lemma di Bufacchi 11.2 applicato con $\Gamma \supseteq \mathrm{Sent}(\mathcal{L})$ o, per l'estensione minima, con $\psi \in \Gamma$ equivalente in T a una frase indipendente.

□

Corollario 11.3 (Metateoria della verità)**.** *Ogni giustificazione del valore di verità che superi l'indeterminatezza interna richiede ipotesi metateoriche esterne, come schemi di riflessione* $\mathrm{RFN}_{\Pi^0_1}(T)$*, oppure il riferimento alla verità esterna nello standard. Nessun predicato di verità definibile in* T *può essere fondamento interno.*

11.3 Soggetto razionale ideale e accordo minimo

Si definisce il soggetto razionale come una politica di giudizio applicata a T con predicato composizionale di verità.

Definizione 11.2 (Soggetto razionale ideale)**.** *Un* soggetto razionale ideale su T *è descritto da una teoria* S *nel linguaggio* $\mathcal{L} \cup \{\mathrm{Tr}\}$ *tale che:*

i. S *estende* T *e contiene uno schema composizionale per* Tr *su una classe* $\Gamma \supseteq \mathrm{Sent}(\mathcal{L})$ *chiusa per connettivi e quantificatori con bicondizionali* $\varphi \leftrightarrow \mathrm{Tr}(\ulcorner \varphi \urcorner)$ *per ogni* $\varphi \in \Gamma$.

ii. S *è conservativa su* $\mathcal{L}$.

iii. Il giudizio indotto da S è

$$J_S \;:=\; \{\, \chi \in \mathrm{Sent}(\mathcal{L}) : S \vdash \mathrm{Tr}(\ulcorner \chi \urcorner)\,\}.$$

Proposizione 11.5 (Allineamento su $\mathcal{L}$)**.** *Per ogni $\chi \in \mathrm{Sent}(\mathcal{L})$, se S è conservativa su $\mathcal{L}$, allora $\chi \in J_S$ se e solo se $T \vdash \chi$.*

Dimostrazione. Per bicondizionali composizionali $S \vdash \mathrm{Tr}(\ulcorner \chi \urcorner) \leftrightarrow \chi$. Se $\chi \in J_S$ allora $S \vdash \chi$ e per conservatività $T \vdash \chi$. Il viceversa è banale in quanto $T \subseteq S$.

□

Proposizione 11.6 (Limite cognitivo interno)**.** *Se σ è indipendente da T, allora esistono estensioni conservative S_1, S_2 con $\sigma \in J_{S_1}$ e $\sigma \notin J_{S_2}$. Di conseguenza il giudizio interno non determina univocamente il valore di verità delle frasi indipendenti.*

Dimostrazione. Per il Lemma di conservatività e consistenza di $T+\sigma$ e $T+\neg\sigma$, $S+\sigma$ e $S+\neg\sigma$ sono consistenti per ogni S conservativa. La tesi segue dai modelli delle due estensioni e dalle bicondizionali composizionali.

□

Corollario 11.4 (Accordo minimo razionale)**.** *Se S_A e S_B estendono $T + \mathrm{RFN}_{\Pi^0_1}(T)$ e sono conservative su $\mathcal{L}$, allora per ogni $\chi \in D_{\Pi^0_1}(T)$ si ha $\chi \in J_{S_A} \cap J_{S_B}$. Inoltre, per ogni $\chi \in D_{\Pi^0_1}(T)$ non si ha $\neg\chi \in J_{S_A} \cup J_{S_B}$.*

Parte 6

Applicazioni Critiche della Teoria della Verità (*Pars Applicata*)

12

Applicazioni critiche della teoria della verità

12.1 Goodstein, Kirby–Paris, Hydra: vero ma non-dimostrabile in PA

12.1.1 Codifica ordinale e base variabile

Le *sequenze di Goodstein* si costruiscono rappresentando un numero naturale n in base $b \geq 2$ con una notazione che utilizza le potenze: si espande n in forma di somma $\sum_i c_i b^{e_i}$ con $0 < c_i < b$ e si ripete ricorsivamente l'espansione degli esponenti e_i nella stessa base, ottenendo la *forma di Goodstein* $G_b(n)$, ovvero la rappresentazione di n in base b con potenze esponenziali annidate. Il passo di decremento della sequenza consiste nel sostituire globalmente b con $b+1$ nella forma $G_b(n)$ ottenendo un termine $G'_{b+1}(n)$ e poi decrementando di 1:

$$n_{b+1} := \mathrm{val}(G'_{b+1}(n_b)) - 1.$$

La sequenza $\langle n_b \rangle_{b \geq 2}$ è la sequenza di Goodstein che parte da $n_2 = n$.

Definizione 12.1 (Forma di Cantor normalizzata). *Un ordinale $\alpha < \varepsilon_0$ si dice in* forma di Cantor normalizzata *se è definito come somma finita di potenze di ω con coefficienti naturali:*

$$\alpha \;=\; \omega^{\beta_1} c_1 + \omega^{\beta_2} c_2 + \ldots + \omega^{\beta_k} c_k,$$

con $c_i \in \mathbb{N} \setminus \{0\}$ *e* $\beta_1 \geq \beta_2 \geq \ldots \geq \beta_k$.

12.1.2 Trasferimento agli ordinali e principio di terminazione

Si effettua un *lifting* sostituendo la base b con ω in ogni occorrenza di $G_b(n)$ e ottenendo un termine ordinale $G_\omega(n)$ in forma di Cantor normalizzata. Il passo di sostituzione $b \mapsto b+1$ corrisponde sul lato ordinale alla identità, mentre il decremento -1 diventa un decremento ordinale. Si ottiene così una funzione di misura $\mu(n_b)$ in un ordinale $< \varepsilon_0$ che decresce finitamente: nessuna catena discendente infinita esiste sotto induzione transfinita fino a ε_0.

Teorema 12.1 (Goodstein)**.** *Ogni sequenza di Goodstein termina, ovvero raggiunge* 0 *in un numero finito di passi.*

Dimostrazione. Si fissa una sequenza di Goodstein $(n_b)_{b \geq 2}$ con $n_2 = n$ e per ogni base b sia $G_b(n_b)$ la forma di Goodstein di n_b. Si definisce una trasformazione τ che sostituisce ogni occorrenza della base b in $G_b(n_b)$ con l'ordinale ω, ottenendo un termine ordinale $\tau(G_b(n_b))$ scritto in forma di Cantor normalizzata. Sia $\mu(n_b) := \tau(G_b(n_b))$ l'ordinario associato a n_b. Si mostra che la sequenza $(\mu(n_b))_{b \geq 2}$ è una catena discendente strettamente $< \varepsilon_0$:

i. $\mu(n_b) < \varepsilon_0$ per ogni b.[1]

ii. $\mu(n_{b+1}) < \mu(n_b)$.

Per (ii), la sostituzione globale $b \mapsto b+1$ in $G_b(n_b)$ seguita da decremento di 1 sul valore numerico, corrisponde sul lato ordinale alla sola operazione di decrescita dell'ultimo termine non nullo nella forma di Cantor. La sostituzione della base non cambia infatti la struttura ordinale, In quanto tutte le basi numeriche si traducevano in ω. Il decremento numerico, dopo la sostituzione, elimina o riduce l'ultimo blocco ordinale, ottenendo un ordinale strettamente minore per la relazione $<$. Si formalizza per induzione sulla complessità della forma: se $G_b(n_b) = b^{e_1}c_1 + \cdots + b^{e_k}c_k$ con ulteriori espansioni di esponenti, τ produce $\omega^{\tau(e_1)}c_1 + \cdots + \omega^{\tau(e_k)}c_k$. Il passo $b \mapsto b+1$ non modifica $\tau(e_i)$. Il -1 si realizza come riduzione dell'ultimo segmento $\omega^{\tau(e_k)}c_k$ o rimozione se $c_k = 1$ e nessun termine successivo. In ogni caso si ottiene un ordinale strettamente minore,

[1] Di fatto, la forma di Goodstein $G_b(n_b)$ è costruita con potenze di b e coefficienti naturali. Sostituendo b con ω si ottiene una somma di potenze di ω con esponenti e coefficienti naturali, che è esattamente la forma di Cantor normalizzata per ordinali $< \varepsilon_0$.

come dimostrato dal lemma classico sulla corrispondenza decremento in somma di Cantor. Quindi la sequenza $(\mu(n_b))_{b\geq 2}$ è una catena discendente strettamente $< \varepsilon_0$. Poiché $< \varepsilon_0$ è ben-fondato, per induzione transfinita sull'ordinario ε_0 la catena $(\mu(n_b))$ è finita. Dunque esiste b_0 tale che $\mu(n_{b_0}) = 0$ e quindi $n_{b_0} = 0$. La sequenza termina.

□

Proposizione 12.1 (Indipendenza da PA)**.** *Il teorema di Goodstein non è dimostrabile in PA se PA è consistente.*

Dimostrazione. Sia $G(n)$ la lunghezza della sequenza di Goodstein che inizia da n. Per Kirby–Paris si mostra che G controlla ogni funzione primitiva ricorsiva: per ogni funzione primitiva ricorsiva f esiste n_0 tale che $f(n) < G(n)$ per ogni $n \geq n_0$. La dimostrazione utilizza la corrispondenza $n \mapsto \mu(n_b)$ dove la decrescita della misura ordinale supera tutte le induzioni oltre tutti gli ordinali primitivamente ricorsivi e raggiunge ε_0. Sia per assurdo che PA dimostri il teorema di Goodstein. Allora in PA si avrebbe che per ogni n la sequenza termina e si potrebbe dimostrare, tramite induzione su formule che codificano la strategia di decrescita ordinale, l'induzione transfinita fino a ε_0. Ma Gentzen ha mostrato che la coerenza di PA, e dunque induzione transfinita fino a ε_0, non è provabile in PA stessa se PA è consistente: l'induzione transfinita fino a ε_0 implicando Con(PA) porterebbe a violare il secondo teorema di Gödel. Dunque dalla dimostrazione interna della terminazione di tutte le sequenze di Goodstein si formalizza in PA un principio che esclude catene discendenti di ordinali $< \varepsilon_0$, equivalente all'induzione transfinita su $< \varepsilon_0$ per formule primitive ricorsive. Tale principio implica Con(PA) ed è quindi non dimostrabile se PA è consistente, ovvero è una contraddizione. Pertanto PA non dimostra il teorema di Goodstein se consistente.

□

12.1.3 Gioco dell'Hydra e crescita non-primitiva

Il *gioco dell'Hydra* di Kirby–Paris definisce una riscrittura su alberi ben fondati: abbattendo un nodo dell'albero si applica una regola di crescita che duplica o espande sottorami a profondità controllata. Nonostante la crescita apparente, ogni partita termina in un numero finito di mosse.

Teorema 12.2 (Kirby–Paris). *Ogni partita del gioco dell'Hydra termina in un numero finito di mosse.*

Dimostrazione. Sia ogni configurazione dell'Hydra un albero radicato finito. Si definisce una misura ordinale $\mu(H)$ assegnando a ogni nodo a profondità d un contributo ordinale della forma ω^{α_d} con esponenti decrescenti ogni volta che si risale verso la radice, costruendo una somma di Cantor finita $\mu(H) < \varepsilon_0$. Le regole del gioco: quando si taglia un ramo, secondo le regole di Kirby–Paris, si sostituisce un sottoalbero con copie o estensioni che aumentano localmente la dimensione. La costruzione delle copie, però, è accompagnata dalla scelta di un'esponente più basso per i nuovi nodi, in modo che la misura globale, riscritta in forma di Cantor normalizzata, decresca strettamente. Formalmente si definisce μ per induzione sulla struttura: se la radice ha sottoalberi $T_1, \ldots, T_k$, si pone $\mu(H) = \omega^{\beta_1} c_1 + \cdots + \omega^{\beta_k} c_k$ con $\beta_1 > \cdots > \beta_k$ e ciascun β_i derivato dalla profondità del corrispondente sottoalbero. Operazioni di taglio sostituiscono un sottoalbero con una famiglia di sottoalberi di profondità minore, riducendo l'ultimo termine della somma ordinale o spostandolo a un esponente inferiore. Si verifica che ogni mossa produce $\mu(H_{t+1}) < \mu(H_t)$. Essendo ε_0 ben-fondato, nessuna catena discendente infinita esiste e il gioco termina.

□

Proposizione 12.2 (Indipendenza dell'Hydra). *La terminazione generale del gioco dell'Hydra non è dimostrabile in PA, supponendo la sua consistenza.*

Dimostrazione. Analogamente al caso Goodstein, sia $H(x)$ la funzione che assegna alla configurazione iniziale codificata dal numero naturale x la lunghezza minima di una partita terminante. Si mostra, utilizzando la codifica ordinali $< \varepsilon_0$ della misura, che H controlla ogni funzione primitiva ricorsiva: per ogni funzione primitiva ricorsiva f esiste x_0 tale che $f(x) < H(x)$ per $x \geq x_0$. Se PA dimostrasse la terminazione per ogni configurazione codificata, si potrebbe formalizzare internamente un principio che esclude la possibilità di catene discendenti ordinali $< \varepsilon_0$, implicando induzione transfinita fino a ε_0 e quindi Con(PA) per la riduzione di Gentzen. Per il secondo teorema di Gödel, assumendo consistenza di PA, tale principio non è ivi dimostrabile. Dunque la terminazione universale dell'Hydra non è dimostrabile in PA.

□

12.1.4 Limite del metodo aritmetico interno

Goodstein e Hydra mostrano che schemi di induzione interna in PA non sono sufficienti a comprendere tutte le verità numeriche. L'uso di ordini transfiniti e metodi esterni sono necessari per dimostrare proprietà vere ma non dimostrabili internamente, mostrando i limiti epistemici dell'aritmetica formale.

12.2 CH, forzatura e ZFC: verità modellistiche

12.2.1 Cardinalità e continuità

Sia $\aleph_0$ la cardinalità di $\mathbb{N}$ e $2^{\aleph_0}$ quella di $\mathcal{P}(\mathbb{N})$. L'*Ipotesi del Continuo* (CH) afferma che non esiste cardinalità intermedia: $2^{\aleph_0} = \aleph_1$.

12.2.2 Struttura di ZFC e assiomi

ZFC determina una teoria r. e. della teoria degli insiemi con assiomi di estensionalità, coppie, unione, infinito, potenza, schemi di comprensione e sostituzione, regolarità e assioma di scelta.

Definizione 12.2 (Modello transitivo)**.** *Un* modello transitivo *di ZFC è una struttura set-teorica $M = (|M|, \in)$ tale che ogni elemento di $|M|$ è un insieme di $|M|$ e $M \models \mathrm{ZFC}$.*

12.2.3 Idea della forzatura

Data una struttura M e un *poset* $\mathbb{P} \in M$, si considerano ultrafiltri generici $G \subseteq \mathbb{P}$ incontrando tutti i densi in M. Si costruisce $M[G]$ interpretando nomi $\dot{x}$ come valutazioni $\dot{x}^G$, mentre verità di formule si caratterizzano tramite la relazione di forzatura $p \Vdash \varphi$.

Lemma 12.1 (Estensione generica)**.** *Se M è modello transitivo di ZFC e $\mathbb{P} \in M$ è un poset, allora per ogni filtro G generico su $\mathbb{P}$ si ha $M \subseteq M[G]$ e $M[G] \models \mathrm{ZFC}$.*

Dimostrazione. Sia M un modello transitivo di ZFC e $\mathbb{P} \in M$ un *poset.* Si costruisce l'estensione tramite la definizione di nomi: per ogni $p \in \mathbb{P}$ e insieme di coppie $(\dot{x}, q)$ con $q \leq p$ definito in M, il nome $\dot{y}$ è un insieme di coppie di nomi e condizioni. Data una $G \subseteq \mathbb{P}$ filtro generico, si definisce l'interpretazione $\dot{x}^G = \{\dot{y}^G : (\dot{y}, p) \in \dot{x} \text{ con } p \in G\}$. Si procede per induzione sul rank del nome, definito come sup dei rank dei nomi contenuti più uno. La transitività assicura che ogni insieme appartenga a M. Si mostra che $M \subseteq M[G]$: per ogni $a \in M$ il nome canonico $\check{a} = \{(\check{b}, 1_{\mathbb{P}}) : b \in a\}$ soddisfa $\check{a}^G = a$. Si verificano gli assiomi: l'estensionalità vale in quanto l'interpretazione preserva l'uguaglianza di elementi. *Pairing* e *union* si ottengono costruendo nomi che combinano coppie e unioni dei nomi già interpretati. L'infinito si preserva in quanto $\omega \in M$ e $\check{\omega}^G = \omega$. La potenza, dato $x \in M[G]$, si possono raccogliere tutti i sottinsiemi definibili da nomi di $\mathbb{P}$. La separazione e sostituzione si verificano in quanto le formule con parametri in $M[G]$ si interpretano mediante la relazione di forzatura $p \Vdash \varphi$ definita in M e l'induzione sul rank. La regolarità si mantiene per la definizione di rank dei nomi: ogni elemento ha rank strettamente minore di quello del contenente. La scelta: se ZFC contiene AC e M è transitivo con una funzione di scelta F, allora la sua interpretazione preserva una funzione di scelta globale, o si estende costruendo un nome per il ben-ordinamento. Quindi $M[G]$ soddisfa ZFC e contiene M.

□

12.2.4 Indipendenza di CH

Gödel mostrò $V = L \models$ CH e Cohen costruì modelli $M[G]$ con ¬CH aggiungendo reali, ovvero sottinsiemi di $\mathbb{N}$, tramite forzatura. Combinando si ottiene che ZFC non dimostra né CH né ¬CH se consistente.

Teorema 12.3 (Indipendenza di CH)**.** *Se ZFC è consistente, allora ZFC+*CH *e ZFC+*¬CH *sono consistenti.*

Dimostrazione. Si suppone ZFC consistente. (i) Gödel: la costruzione dell'universo di Gödel $L = \bigcup_{\alpha \in \mathrm{Ord}} L_\alpha$, definendo $L_{\alpha+1}$ come l'insieme dei sottoinsiemi definibili di L_α con parametri e $L_\lambda = \bigcup_{\alpha<\lambda} L_\alpha$ per ordinale limite, produce un modello di ZFC. In L si dimostra GCH e dunque CH, quindi se ZFC è consistente lo è ZFC+CH. (ii) Cohen: si prende un modello transitivo $M \models$ ZFC e un *forcing* $\mathbb{P}$ che aggiunge almeno $\aleph_2$ nuovi reali evitando collasso di cardinale. Costruendo un filtro generico G su $\mathbb{P}$ si ottiene $M[G] \models$ ZFC e si verifica $2^{\aleph_0} \geq \aleph_2$ in $M[G]$, dunque ¬CH. La preservazione di cardinale è assicurata dal fatto che $\mathbb{P}$ soddisfa la c.c.c. evitando l'aggiunta di nuove sequenze di ordinali di lunghezza ω_1. Così ZFC+¬CH è consistente se ZFC lo è. Si conclude che ZFC non dimostra né CH né ¬CH.

□

12.2.5 Pluralismo modellistico

La verità di CH è *modellistica*: dipende da scelte di universo V o da estensioni generiche $V[G]$. Ciò mostra la non-unicità del valore di verità interno con schemi composizionali limitati, analogo alla non-invarianza uniforme del Lemma di Bufacchi 11.2 applicata a strutture set-teoriche.

12.3 Dimostrazioni automatiche e limiti d'incompletezza

12.3.1 Verifica meccanica

Sistemi come Coq, Lean, HOL Light formalizzano teoremi riducendo la loro validità alla correttezza del *kernel*. Il predicato *provabile-da-sistema* non coincide necessariamente con il predicato metafisico di verità matematica.

12.3.2 Automazione e ricerca

Procedure di *superposition*, *SMT solving* e *model finding* estendono la capacità di esplorazione. Tuttavia generano dimostrazioni dipendenti da euristiche non completamente giustificate internamente.

12.3.3 Limiti teorici

L'incompletezza di Gödel limita la volontà di sistemi automatici di provare tutte le verità matematiche meccanicamente: non esiste algoritmo totale che elenchi tutte le verità aritmetiche. L'indecidibilità di problemi come halting, definisce limiti strutturali ai *tool* di prova.

Proposizione 12.3 (Non-completezza algoritmica)**.** *Nessun sistema automatico di dimostrazione r. e. può generare tutte e sole le frasi vere di $\mathcal{L}$ se include l'aritmetica elementare consistente.*

Dimostrazione. Sia T consistente, r. e. e che interpreta Q. Si suppone per assurdo che esista un algoritmo r. e. che elenchi esattamente l'insieme $\mathrm{Th}_{\mathbb{N}}$ delle frasi vere nell'aritmetica standard. Allora esisterebbe un predicato r. e. $V(x)$ tale che per ogni frase chiusa φ, $V(\ulcorner\varphi\urcorner)$ se e solo se φ è vera in $\mathbb{N}$. Poiché

V è r. e., è rappresentabile in T da una formula $v(x)$. Sia la formula ψ ottenuta per diagonalizzazione che afferma "non v vale sul mio numero di Gödel": $\psi \leftrightarrow \neg v(\ulcorner \psi \urcorner)$. Se ψ è vera, allora $v(\ulcorner \psi \urcorner)$ è falsa, quindi la rappresentazione non è corretta. Se ψ è falsa, allora $v(\ulcorner \psi \urcorner)$ è vera e dunque ψ è vera. Pertanto è una contraddizione. Questo è analogo all'argomento del mentitore e mostra che un predicato r. e. non può coincidere con la verità aritmetica totale. Quindi nessun sistema automatico r. e. elenca tutte e sole le verità: la completezza algoritmica è impossibile.

□

12.3.4 Aporie epistemiche

La dipendenza da strumenti aumenta l'affidabilità locale mantenendo un vincolo metateorico criticamente non fondato: la correttezza del *kernel*. Ciò mostra un livello di *fiducia* esterno al sistema, generando un paradosso di tipo epistemico simile a quello del fondamento interno.

12.3.5 Vie di ricerca

Linee di ricerca includono: riflessione *proof assistant* interna, come la *type-theoretic reflection*, estrazione di modelli certi da dimostrazioni automatiche, e studio di limiti computazionali in relazione a gerarchie di verità.

12.3.6 Convergenza delle aporie del vero

Le aporie del vero in aritmetica, teoria degli insiemi e dimostrazioni automatiche hanno una struttura comune: la non-determinazione del valore di verità interno a causa di limiti formali e la necessità di ipotesi metateoriche esterne per fondare valori di verità, confermando i risultati del Lemma di Bufacchi 11.2.

13

Approfondimenti e sviluppi avanzati

13.1 Gerarchie di riflessione e stabilizzazione parziale

Si definiscono $T^{(0)} := T$ e $T^{(n+1)} := T^{(n)} + \mathrm{RFN}_{\Pi^0_1}(T^{(n)})$ e sia $T^{(\omega)} := \bigcup_{n<\omega} T^{(n)}$. Ogni livello aggiunge verità Π^0_1 non contenute in quello precedente.

Proposizione 13.1 (Crescita non-stazionaria)**.** *Per ogni n, se $T^{(n)}$ è consistente allora $T^{(n)} \nvdash \mathrm{RFN}_{\Pi^0_1}(T^{(n)})$ e $T^{(n)} \subsetneq T^{(n+1)}$.*

Dimostrazione. Per ogni n, $T^{(n)}$ è r. e. e contiene T. Se $T^{(n)} \vdash \mathrm{RFN}_{\Pi^0_1}(T^{(n)})$ allora usando lo schema di riflessione per frasi Π^0_1 che includono la codifica della non dimostrabilità di $\perp$ si ottiene $T^{(n)} \vdash \mathrm{Con}(T^{(n)})$. Ma per il secondo teorema di Gödel, se $T^{(n)}$ è consistente e sufficientemente forte, non dimostra la propria consistenza. Dunque $T^{(n)} \nvdash \mathrm{RFN}_{\Pi^0_1}(T^{(n)})$ e per definizione si ottiene un'estensione stretta $T^{(n+1)}$.

□

Lemma 13.1 (Non-invarianza forte delle gerarchie di riflessione)**.** *Per ogni frase Π^0_1 vera nello standard esiste n con $T^{(n)} \vdash \chi$ se T è Π^0_1-corretto. Non esiste n che comprenda tutte le verità Π^0_1.*

Dimostrazione. Poiché T è Π^0_1-corretto, ogni frase Π^0_1 vera è tale che la sua verità implica la sua non confutabilità in T. Per una data verità Π^0_1 χ si considera $T + \chi$. Per compattezza sintattica si può mostrare che $\mathrm{RFN}_{\Pi^0_1}(T)$ più un numero finito di istanze consente di derivare χ. Si procede per iterazioni. Esiste n con $T^{(n)} \vdash \chi$. Se esistesse n che comprende tutte le verità Π^0_1, allora $T^{(n)}$ sarebbe Π^0_1-completo e, codificando tale completezza, si avrebbe un predicato di verità Π^0_1 interno, violando Gödel-Tarski. Quindi nessun singolo livello comprende tutte le verità Π^0_1.

□

13.2 Sistemi di verità parziale e non-invarianza

Si considerano CT^-, KF minimale, PT stratificato e gerarchie FS. Essi sono conservativi su $\mathcal{L}$ aritmetica, ma differiscono sulla risoluzione dei paradossi e sulla gestione dell'autoreferenza.

Proposizione 13.2 (Non-invarianza dei sistemi di verità parziale)**.** *Sia S uno schema conservativo di verità parziale e σ indipendente per T. Allora esistono modelli di S con valori opposti per* $\mathrm{Tr}(\ulcorner\sigma\urcorner)$*.*

Dimostrazione. Sia σ indipendente per T. Poiché gli schemi di verità parziale — CT^-, KF, PT, FS — sono costruiti per essere conservativi su $\mathcal{L}$ aritmetico, l'aggiunta di σ o $\neg\sigma$ produce estensioni consistenti, assumendo consistenza di T, i cui modelli differiscono sul valore di σ. Le bicondizionali per la parte coperta dallo schema non forzano σ, quindi esistono modelli $\mathcal{M}_1, \mathcal{M}_2$ di S tali che $\mathcal{M}_1 \models \mathrm{Tr}(\ulcorner\sigma\urcorner)$ e $\mathcal{M}_2 \models \neg\mathrm{Tr}(\ulcorner\sigma\urcorner)$. Ne segue banalmente la non-invarianza.

□

13.3 Limiti *proof-theoretic* e ordinali

L'ordinale di PA è ε_0. Estensioni con riflessione iterata portano verso ordinali maggiori e la misura ordinale non è definibile internamente da un predicato aritmetico totale.

Proposizione 13.3 (Non-definibilità della misura)**.** *Non esiste predicato* $\mathrm{OrdBound}(x)$ *definibile in T che assegni a ogni derivazione una codifica completa della sua misura ordinale.*

Dimostrazione. Si supponga l'esistenza di $\mathrm{OrdBound}(x)$ che associa a ogni derivazione una codifica completa della misura ordinale di complessità usata nelle riduzioni di taglio. Allora si potrebbe decidere per ogni derivazione se supera un dato ordinale critico e costruire un predicato $\mathrm{True}(\ulcorner\varphi\urcorner)$ verificando se esiste una derivazione di φ con misura ordinaria limitata (le verità Π^0_1 corrispondono a derivazioni entro ordinali specifici). Ciò fornirebbe un predicato interno che soddisfa bicondizionali per tutte le frasi, contraddicendo Tarski: la verità aritmetica non è definibile. Quindi tale predicato non esiste.

□

13.4 Logiche modali di provabilità e interpretabilità

La logica GL assiomatizza $\Box$ come provabilità, e sistemi di interpretabilità, come quelli Japaridze e Visser, studiano relazioni tra teorie. L'invarianza forte non può essere soddisfatta per $\Box_T$.

Proposizione 13.4 (Non-invarianza di $\Box_T$)**.** *Per ogni estensione consistente T di PA, l'operatore $\Box_T$ non rappresenta alcun predicato composizionale interno uniformemente invariante.*

Dimostrazione. Sia T consistente estensione di PA e σ indipendente. L'operatore modale $\Box_T$ codifica provabilità in T. In due modelli $\mathcal{M}_1, \mathcal{M}_2$ di T con assegnazioni diverse per l'interpretazione di eventuali costanti (o in estensioni conservativi $T + \sigma$, $T + \neg\sigma$) si ha che $\Box_T\sigma$ non cambia (dipende da esistenza di una dimostrazione finita), mentre il valore di verità di σ cambia. Per un predicato composizionale uniformemente invariante si richiederebbe coerenza del valore di $\Box_T$ con la verità di σ su tutti i modelli che preservano struttura sintattica; ciò fallisce in presenza di indipendenza. Applicando il Lemma di Bufacchi 11.2, la classe di formule dove $\Box_T$ coincide con un predicato di verità uniforme non può includere indipendenti. Pertanto $\Box_T$ non è internamente composizionale e uniformemente invariante.

□

13.5 Logica interna dei topos

Nel topos $\mathcal{E}$ l'oggetto verità Ω classifica sotto-oggetti. Punti $p : \mathcal{E} \to \mathbf{Set}$ inducono valutazioni classiche divergenti per frasi indipendenti e la non-invarianza lungo i punti ripete l'argomento aritmetico.

Lemma 13.2 (Dipendenza dai punti del topos)**.** *Esistono p, q punti con $p^*(\mathrm{Tr})(\ulcorner\sigma\urcorner) \neq q^*(\mathrm{Tr})(\ulcorner\sigma\urcorner)$ per σ indipendente.*

Dimostrazione. Sia $\mathcal{E}$ un topos con almeno due punti geometrici distinti $p, q : \mathcal{E} \to \mathbf{Set}$. Si considera un oggetto che interpreta la struttura aritmetica interna e una frase σ indipendente in quella struttura. Le forzature interne — logica intuizionistica — consentono modelli differenti in cui σ assume valori classici diversi dopo il *pullback* lungo i punti, in quanto i punti preservano *colimits* e valutano la logica intuizionistica in quella classica dei *set*. Quindi $p^*(\mathrm{Tr})(\ulcorner\sigma\urcorner) \neq q^*(\mathrm{Tr})(\ulcorner\sigma\urcorner)$. Questo realizza la dipendenza e non-invarianza.

□

13.6 Limiti computazionali e Goodstein

Si considera la funzione di lunghezza di Goodstein $G : \mathbb{N} \to \mathbb{N}$ che associa a n il numero di passi necessari per terminare la sequenza di Goodstein che parte da n.

Proposizione 13.5 (*Lower bound* di Goodstein)**.** *Qualunque procedura totale che decide $J(\sigma)$ per tutte le σ con lunghezza di Goodstein limitata richiede un tempo superiore a ogni funzione primitiva ricorsiva.*

Dimostrazione. Per ogni funzione primitiva ricorsiva f esiste n_0 con $f(n) < G(n)$ per $n \geq n_0$. Se esistesse una procedura primitiva ricorsiva che decidesse $J(\sigma)$ limitandosi alle frasi la cui valutazione richiede esplorare sequenze di Goodstein fino a lunghezza $f(n)$ per qualche f primitiva ricorsiva, tale procedura fallirebbe per input con lunghezza reale $G(n) > f(n)$. La crescita di G oltre ogni primitiva ricorsiva implica che il tempo necessario supera ogni *bound* primitivo ricorsivo: si deduce il *lower bound.*

□

14

Conclusione

14.1 Irriducibilità del vero e dualità giudizio-verità

Si distingue tra livello *deontologico* del giudizio — come le regole di inferenza, schemi composizionali e riflessioni ammissibili — e livello *ontologico* della verità — come valore che permane sotto variazioni modellistiche e interpretative. La tesi è l'**irriducibilità**: nessuna struttura interna di regole prescrittive realizza una comprensione esaustiva e invariabile della verità totale.

Definizione 14.1 (Dualità giudizio-verità)**.** *Il* giudizio interno J *derivato da una teoria estesa* $S \supseteq T$ *con predicato composizionale è l'insieme delle frasi per cui* $S \vdash \mathrm{Tr}(\ulcorner \varphi \urcorner)$*. La* verità ontologica esterna *su classi stabili, come* Π^0_1*, è indipendente da scelte di* S *e modelli di* T*.*

Proposizione 14.1 (Irriducibilità del vero)**.** *Se* T *è r. e., consistente e interpreta* Q*, allora non esiste teoria interna* S *tale che* J *coincida con la verità ontologica esterna e sia interpretazionalmente invariante, quantunque mantenendo composizionalità e conservatività.*

Dimostrazione. Sia T consistente, r. e. e che interpreta Q. Si suppone che esista una teoria estesa $S \supseteq T$ con predicato composizionale di verità tale che il giudizio interno J coincida con la verità ontologica esterna su una classe stabile di frasi, ad esempio Π^0_1, e sia interpretazionalmente invariante. Per il Lemma di Bufacchi 11.2, esistono frasi σ indipendenti per T tali che in modelli diversi di T il valore di verità di σ cambia, mentre il giudizio J rimane fisso in quanto derivato da S. Ciò implica che J non coincide con la verità ontologica esterna per tutte le frasi, contraddicendo l'ipotesi. Di conseguenza non esiste tale teoria S che realizzi questa coincidenza non violando l'invarianza interpretazionale e mantenendo composizionalità e conservatività su T.

$\square$

La componente deontologica gestisce l'*agire inferenziale* senza risolvere l'ontologia della verità. Il risultato evita sia un riduzionismo semanticista, sia una decomposizione relativista: la verità appare come *invariante trascendentale parziale* e il giudizio come *regolazione inferenziale contingente*.

14.2 Prospettive di ricerca avanzate

Segue un elenco di direzioni di ricerca avanzate che estendono i risultati presentati:

- Studio di gerarchie di riflessione trasfinite e loro limiti di stabilizzazione.
- Analisi di sistemi di verità parziale più complessi e loro impatto sull'invarianza.
- Esplorazione di logiche modali avanzate per la provabilità e l'interpretabilità.
- Applicazioni della teoria della verità a teorie non-classiche e logiche sub-classiche.
- Approfondimento delle connessioni tra teoria della verità e teoria dei topos.
- Studio dei limiti computazionali in relazione a funzioni non-primitive ricorsive.
- Investigazione di approcci categoriali alla verità e al giudizio.

Bibliografia

i. Artemov, Sergei N. «Logic of Proofs». In: *Annals of Pure and Applied Logic* 67 (1994). Logica dei certificati di prova (LP), pp. 29–59.

ii. Beklemishev, Lev D. «Reflection Principles and Provability Algebras in Formal Arithmetic». In: *Russian Mathematical Surveys* 58.4 (2003). Principi di riflessione e gerarchie modali, pp. 655–722.

iii. Boolos, George. *The Logic of Provability.* Sintesi su GL e riflessione. Cambridge University Press, 1993.

iv. Buchholz, Wilfried e Stanley S. Wainer. «Provably Computable Functions and Ordinals». In: *Handbook of Proof Theory (capitolo)* (1987). Analisi ordinale e funzioni gerarchiche.

v. Bufacchi, Enrico Maria. «Indefinibilità Forte e Non-Invarianza Uniforme della Verità». In: *Manoscritto* (2025). Lemma di Bufacchi; formulazione originale nello studio presente.

vi. Cantor, Georg. *Beiträge zur Begründung der transfiniten Mengenlehre.* German. Origini teoria insiemi e cardinali. J. C. Hinrichs, 1895.

vii. Cohen, Paul J. «The Independence of the Continuum Hypothesis». In: *Proceedings of the National Academy of Sciences USA* 50 (1963). Indipendenza di CH, pp. 1143–1148.

viii. – «The Independence of the Continuum Hypothesis II». In: *Proceedings of the National Academy of Sciences USA* 51 (1964). Seconda parte forzatura CH, pp. 105–110.

ix. Feferman, Solomon. «Transfinite Recursive Progressions of Axiomatic Theories». In: *Journal of Symbolic Logic* 27.3 (1962). Progressioni di riflessione, pp. 259–316.

x. Fraenkel, Adolf. «Zu den Grundlagen der Cantorschen Mengenlehre». In: *Mathematische Annalen* 86 (1922). Aggiunte agli assiomi insiemistici, pp. 230–237.

xi. G"odel, Kurt. «Die Vollst"andigkeit der Axiome des logischen Funktionenkalk"uls». German. In: *Monatshefte f"ur Mathematik und Physik* 37 (1930). Teorema di completezza, pp. 349–360.

xii. – «Uber formal unentscheidbare S"atze der Principia Mathematica und verwandter Systeme I». German. In: *Monatshefte f"ur Mathematik und Physik* 38 (1931). Teoremi di incompletezza, pp. 173–198.

xiii. Gentzen, Gerhard. «Die Widerspruchsfreiheit der reinen Zahlentheorie». German. In: *Mathematische Annalen* 112 (1936). Consistenza via induzione transfinita fino a ε_0, pp. 493–565.

xiv. Goodstein, Reuben L. «On the Restricted Ordinal Theorem». In: *Journal of Symbolic Logic* 9.2 (1944). Sequenze di Goodstein, pp. 33–41.

xv. Grothendieck, Alexandre, Michael Artin e Jean-Louis Verdier. *Théorie des topos et cohomologie des schémas (SGA4)*. Origini della teoria dei topos. Springer, 1972.

xvi. Hardy, G. H. «A Theorem Concerning the Infinite Cardinal Numbers». In: *Quarterly Journal of Mathematics* 35 (1904). Prime analisi di crescita e ordini infinitari, pp. 87–94.

xvii. Hilbert, David e Paul Bernays. *Grundlagen der Mathematik I*. German. Finitismo e fondamenti. Springer, 1934.

xviii. – *Grundlagen der Mathematik II*. German. Condizioni di derivabilità. Springer, 1939.

xix. Japaridze, Giorgi. «The Polymodal Provability Logic». In: *Annals of Pure and Applied Logic* 35 (1987). Logica di provabilità polimodale (GLP), pp. 139–156.

xx. Jech, Thomas. *Set Theory (3rd ed.)* Manuale avanzato: forcing, cardinali, indipendenze. Springer, 2002.

xxi. Johnstone, Peter T. *Sketches of an Elephant: A Topos Theory Compendium.* Opera di riferimento su teoria dei topos. Oxford University Press, 2002.

xxii. Kirby, Laurie e Jeff Paris. «Accessible Independence Results for Peano Arithmetic». In: *Bulletin of the London Mathematical Society* 14 (1982). Hydra e indipendenze elementari, pp. 285–293.

xxiii. Kripke, Saul A. «Outline of a Theory of Truth». In: *Journal of Philosophy* 72.19 (1975). Punti fissi e verità parziale, pp. 690–716.

xxiv. Kunen, Kenneth. *Set Theory*. Trattazione moderna di Set Theory e forcing. College Publications, 2011.

xxv. L"ob, Martin Hugo. «Solution of a Problem of Leon Henkin». In: *Journal of Symbolic Logic* 20.2 (1955). Teorema di L"ob, base di GL, pp. 115–118.

xxvi. Lawvere, F. William. «Adjointness in Foundations». In: *Dialectica* 23 (1969). Fondamenti categoriali e logica interna, pp. 281–296.

xxvii. Montagna, Franco. «On the Arithmetic Hierarchy in Provability Logics». In: *Journal of Symbolic Logic* 56.3 (1991). Aspetti gerarchici interpretabilità e riflessione, pp. 997–1011.

xxviii. Paris, Jeff e Leo Harrington. «A Mathematical Incompleteness in Peano Arithmetic». In: *Handwritten Manuscript / Later JSL summary* (1977). Principio di Paris–Harrington.

xxix. Rosser, J. Barkley. «Extensions of Some Theorems of G"odel and Church». In: *Journal of Symbolic Logic* 1.3 (1936). Miglioramento dell'incompletezza (Rosser), pp. 87–91.

xxx. Sch"utte, Kurt. *Proof Theory*. Sviluppi dell'analisi proof-theoretic. Springer, 1977.

xxxi. Schoenfield, Joseph R. *Mathematical Logic*. Manuale classico su logica e modelli. Addison-Wesley, 1967.

xxxii. Skolem, Thoralf. «"Uber einige Satz der Elementaren Arithmetik». In: *Skrifter Videnskaps-Ak. Oslo* (1922). Paradosso di Skolem e modelli non standard.

xxxiii. Smullyan, Raymond M. *G"odel's Incompleteness Theorems*. Presentazione divulgativa e tecnica dei teoremi. Oxford University Press, 1992.

xxxiv. Solovay, Robert M. «Provability Interpretations of Modal Logic». In: *Israel Journal of Mathematics* 25 (1976). Completezza aritmetica di GL, pp. 287–304.

xxxv. Takeuti, Gaisi. *Proof Theory*. Analisi proof-theoretic e ordinali. North-Holland, 1975.

xxxvi. Tarski, Alfred. «Pojécie prawdy w jézykach nauk dedukcyjnych». Polish. In: *Prace Towarzystwa Naukowego Warszawskiego* 34 (1933). Versione originale del saggio sull'indefinibilità della verità, pp. 1–114.

xxxvii. – «The Semantic Conception of Truth and the Foundations of Semantics». In: *Philosophy and Phenomenological Research* 4.3 (1944). Esposizione divulgativa della concezione semantica di verità, pp. 341–376.

xxxviii. Tarski, Alfred e Bronisław Knaster. «Contributions to the Theory of Lattice-Ordered Groups». In: *Fundamenta Mathematicae* 41 (1955). Teorema dei punti fissi (applicato in Kripke), pp. 194–218.

xxxix. Tierney, Myles. «Sheaf Theory and the Continuum Hypothesis». In: *Lecture Notes (SGA 4 1/2)* (1972). Collegamenti tra forcing e topos.

xl. Visser, Albert. «An Overview of Interpretability Logic». In: *Advances in Modal Logic (preprint) / Annals of Pure and Applied Logic* volume vari (1989). Logiche di interpretabilità, pp. 1–40.

xli. Zermelo, Ernst. «Untersuchungen "uber die Grundlagen der Mengenlehre I». In: *Mathematische Annalen* 65 (1908). Fondamenti assiomatici insiemi, pp. 261–281.