

# Kripke: modalità e verità

Achille C. Varzi

[Versione finale pubblicata in *Il genio compreso. La filosofia di Saul Kripke* (a cura di A. Borghini), Roma, Carocci Editore, 2010, pp. 23–78, 186–191]

La prima fase della carriera filosofica di Saul Kripke è legata principalmente, se non esclusivamente, ai suoi contributi in ambito logico. Si tratta di contributi che hanno avuto un impatto enorme soprattutto in due capitoli centrali di questa disciplina, la logica modale e la teoria formale della verità, con conseguenze e ramificazioni che hanno interessato un po' tutta la filosofia analitica contemporanea. In questo capitolo cerchiamo di ricostruirne i tratti principali e di evidenziare la loro portata con particolare riferimento alla filosofia del linguaggio e alla metafisica. La sezione 1 è dedicata alla logica modale. Della teoria della verità, che è un po' più complessa, ci occupiamo nella sezione 2.

## 1. La logica modale

Come già accennato nell'Introduzione, la prima pubblicazione di Kripke venne accettata dal «Journal of Symbolic Logic» quando era ancora adolescente. Si tratta dell'articolo *A Completeness Theorem in Modal Logic* (1959a). A questo fecero seguito cinque altri articoli pubblicati nel giro di pochi anni (1962a, 1963a, 1963b, 1965a, 1965b), tutti dedicati alla logica modale e alle sue applicazioni, e si può ben dire che nel loro insieme questi sei lavori abbiano ridisegnato l'intera mappa della disciplina e le linee principali lungo le quali si è successivamente sviluppata sino ai nostri giorni. Per comprenderne appieno la portata innovativa, è opportuno cominciare con alcune considerazioni generali sullo stato della logica modale negli anni immediatamente precedenti alla loro pubblicazione, e prima ancora sulla nozione stessa di modalità.

### 1.1. Il contesto storico

In termini molto generali, la logica modale è quella parte della teoria logica che si occupa del significato e del comportamento inferenziale di quegli enunciati che non si limitano a presentare un certo stato di cose, come in

(1) I cittadini pagano le tasse.

ma si pronunciano esplicitamente sul suo *modo* di darsi, come in ciascuno degli esempi seguenti:

(2) È *necessario che* i cittadini paghino le tasse.

(3) È *possibile che* i cittadini paghino le tasse.

(4) *Si sa che* i cittadini pagano le tasse.

(5) *Si crede che* i cittadini paghino le tasse.

(6) È *obbligatorio che* i cittadini paghino le tasse.

(7) *Talvolta* i cittadini pagano le tasse.

(8) *Qui da noi* i cittadini pagano le tasse.

Le espressioni in corsivo sono dette *operatori modali* e interagiscono con l'enunciato a cui si applicano, in questo caso (1), dando luogo a enunciati più complessi le cui condizioni di verità differiscono da quelle di partenza. Per esempio, posto che (1) sia vero, cioè che i cittadini paghino davvero le tasse, non ne segue che sia necessariamente così, o che sia un fatto risaputo o obbligatorio, sebbene sia ragionevole concludere che sia almeno una possibilità, o che talvolta sia proprio così. Il compito della logica modale – o di *una* logica modale – è appunto quello di stabilire il sussistere o meno di nessi logici di questo tipo.

Operatori siffatti hanno ampio impiego nel linguaggio ordinario. Soprattutto, essi riflettono distinzioni che appaiono importanti anche sul piano filosofico. Per esempio, la distinzione tra ciò che è necessario e ciò che è meramente possibile occupa una posizione centrale in metafisica, tant'è vero che già Aristotele se ne serviva per caratterizzare la doppia demarcazione tra proprietà essenziali e proprietà accidentali, e tra proprietà attuali e proprietà potenziali. Analogamente, la distinzione tra sapere e credere occupa un ruolo centrale in epistemologia, quella tra ciò che è obbligatorio e ciò che non lo è risulta fondamentale in ambito deontico (o etico in senso lato), e così via. Anche il linguaggio della scienza si serve spesso di nozioni modali. Per esempio, è pratica comune formulare certe leggi sulla base delle proprietà disposizionali degli oggetti a cui si riferiscono, come quando diciamo che lo zucchero è *solubile* nell'acqua o che il ferro *tende* a ossidarsi all'aria: queste affermazioni sono modali, ancorché implicitamente, poiché non si limitano a dire come stanno le cose ma si pronunciano su come potrebbero essere. Inoltre le leggi della fisica non si limitano a registrare delle verità universali ma sono spesso presentate come verità necessarie, cioè inviolabili: è una verità universale, per esempio, che tutti i blocchi di zucchero abbiano una massa inferiore al metro cubo, ma ciò non basta a fare di questa

verità una legge fisica per il semplice fatto che *potremmo* produrre di blocchi di massa maggiore. In breve: la modalità è dappertutto, e tanto basta a definire l'ampio campo d'azione della logica modale.

Ora, nella logica tradizionale questo fatto era largamente riconosciuto, almeno per ciò che concerne alcuni operatori. Per esempio, già Aristotele aveva studiato in maniera sistematica come la modalità di un enunciato interagisca con la sua qualità (affermativa o negativa), giungendo a identificare ciò che è necessario con ciò che «non può non essere». Anche l'interazione tra la modalità di un enunciato e la sua quantità (universale o particolare) era ampiamente studiata, e la sillogistica modale abbozzata da Aristotele aveva raggiunto livelli di sviluppo molto sofisticati tra i logici medievali del dodicesimo e del tredicesimo secolo, al punto da ispirare il motto *De modalibus non gustabit asinus*. I logici medievali erano anche giunti a evidenziare un'importante somiglianza tra gli operatori modali e quelli di quantificazione, suggerendo quell'analogia tra la necessità e l'universalità che successivamente trovò la sua massima espressione nell'identificazione leibniziana di ciò che è necessario con ciò che «vale in tutti i mondi possibili».

Nonostante tutto ciò, la nascita e lo sviluppo della logica contemporanea a partire dalla fine del diciannovesimo secolo coincisero con un deciso mutamento di prospettiva nei confronti della modalità. Per Frege, presentare un enunciato come necessario significava impegnarsi in qualche modo sulla sua giustificazione, e ciò esulerebbe dal raggio d'interesse della logica pura. Per Russell, gli unici modi di essere di un enunciato erano rappresentati dalla verità e dalla falsità, e ogni altra nozione modale andava semplicemente «bandita dalla logica». Anche per Wittgenstein la necessità e la possibilità esulavano dal dominio del «dicibile». In aggiunta, due importanti considerazioni avevano contribuito alla emarginazione delle modalità dal campo d'interesse della nuova logica, o almeno a un certo scetticismo nei confronti di un loro trattamento adeguato.

In primo luogo, l'idea iniziale in base alla quale le modalità rappresenterebbero il «modo di darsi» di uno stato di cose sembrava entrare in diretta collisione con il potere espressivo dei nuovi linguaggi formali, che non pongono alcun limite al numero di operatori che si possono applicare a un dato enunciato. Così, per limitarci al caso della necessità e della possibilità, accanto a enunciati come (2) e (3) una teoria sufficientemente completa dovrebbe fare i conti, non solo con enunciati quali

- (9) *È necessario che sia necessario che* i cittadini paghino le tasse.
- (10) *È possibile che sia possibile che* i cittadini paghino le tasse.

nei quali la modalità è iterata, ma addirittura con enunciati di complessità arbitraria come

- (11) *È necessario che non sia necessario che sia possibile che non sia necessario che sia possibile che . . . sia necessario che* i cittadini paghino le tasse.

dinnanzi ai quali l'intuizione sembra alzare le braccia. Qual è il nesso logico tra enunciati del genere? E in che relazione stanno, sul piano logico, con gli enunciati più semplici in (2) e (3)? Non mancava chi si fosse cimentato nell'impresa di rispondere a queste domande mettendo a punto sistemi di assiomi e regole d'inferenza che nulla avevano da invidiare a quelli comunemente accettati per la logica non modale, a partire da autori come Clarence I. Lewis, Oskar Becker, Cooper H. Langford e Edward J. Lemmon. Ma tali tentativi si erano in breve tempo risolti nello sviluppo di una selva di logiche modali differenti (e fra loro non equivalenti) che anziché chiarire le cose sembravano confermare la natura sfuggente e in apparenza incontrollabile degli operatori modali.

In secondo luogo, l'idea stessa che le modalità corrispondessero a operatori enunciativi veri e propri era stata messa sotto accusa da considerazioni che apparivano ben più gravi delle osservazioni sporadiche e tutto sommato opinabili avanzate da Frege, Russell e Wittgenstein. In particolare, a cominciare dagli anni Quaranta Willard V. O. Quine aveva obiettato che tali operatori soffrirebbero di una inguaribile «opacità referenziale», dichiarando nel contempo che la loro integrazione nella logica dei predicati avrebbe finito col riproporre la «giungla metafisica dell'essenzialismo aristotelico». La prima obiezione si basava sulla constatazione che gli operatori modali determinano contesti linguistici che non sempre ammettono la sostituzione *salva veritate* di termini coreferenziali, contrariamente a quanto vorrebbero le leggi dell'identità. Per usare un esempio dello stesso Quine, sebbene l'espressione «il numero dei pianeti» e il numerale «9» si riferiscano al medesimo numero, sembra proprio che la sostituzione di un termine con l'altro si risolva nel passaggio dalla verità in (12) alla falsità in (13):

- (12) *Necessariamente*, 9 è maggiore di 7.  
(13) *Necessariamente*, il numero dei pianeti è maggiore di 7.

Considerazioni analoghe valgono anche per operatori modali diversi. Per esempio, nonostante Marco Tullio e Cicerone fossero la stessa persona, dal fatto che si sappia o si creda che Marco Tullio abbia fatto certe cose non segue che si sappia o si creda che Cicerone abbia fatto quelle cose. Quindi, nella misura in cui la sostituibilità degli identici – e più in generale la cosiddetta trasparenza

«estensionale» – è considerata un prerequisito irrinunciabile di qualsiasi contesto suscettibile di analisi logica, gli operatori modali definirebbero contesti inaccettabili. Quanto alla seconda accusa avanzata da Quine, essa si basava sul fatto che il trattamento logico delle modalità nell'ambito della teoria della quantificazione ci costringerebbe a prendere posizione in merito a questioni che con la logica non dovrebbero avere nulla a che vedere. Per esempio, come già avevano osservato i logici medievali, c'è una bella differenza tra un enunciato come (14), nel quale l'operatore di necessità precede il quantificatore «ogni», e un enunciato come (15), nel quale si applica invece soltanto alla parte che segue il quantificatore:

- (14) *Necessariamente*, ogni madre ha dei figli.
- (15) Ogni madre ha *necessariamente* dei figli.

In (14) abbiamo a che fare con una modalità *de dicto*, ossia una modalità che riguarda l'intero enunciato «Ogni madre ha dei figli»; in (15) abbiamo invece a che fare con una modalità *de re*, che riguarda cioè delle cose ben precise: tutte quelle cose che ricadono sotto il predicato «madre». E mentre nel primo caso la modalità esprime una verità concettuale che appare del tutto innocua (una madre senza figli è semplicemente un controsenso), nel secondo caso la verità o falsità dell'enunciato complessivo sembrerebbe dipendere da considerazioni metafisiche belle e buone (si può pensare che la maternità sia una proprietà in qualche modo «essenziale» all'identità di coloro che la posseggono, ma si può anche pensare che certe madri – certe donne che di fatto hanno dato vita a uno o più figli – sarebbero potute esistere senza avere figli). Naturalmente non c'è niente di male nel fatto che la verità di certi enunciati possa risultare controversa sul piano metafisico. Tuttavia ciò diventa un problema nel momento in cui si tratta di stabilire le relazioni che tali enunciati intrattengono con altri enunciati sul piano puramente *logico*, per esempio se (14) implichi logicamente (15). Quindi, poiché una volta riconosciuta la legittimità di un operatore non c'è modo di impedire che venga usato per esprimere modalità *de re*, tanto peggio per la logica modale.

## 1.2. Operatori modali e mondi possibili

È appunto in questo contesto che si inseriscono i contributi fortemente innovativi di Kripke. In buona sostanza, tutte le difficoltà che abbiamo ricordato erano sintomatiche di un fatto molto semplice: per quanto il significato intuitivo degli operatori modali potesse inizialmente sembrare chiaro, mancava ancora una sua

caratterizzazione precisa. Mancava cioè un rigoroso apparato *semantico* attraverso il quale analizzare i quesiti che l'intuizione lasciava inevasi e sulla base del quale esaminare l'adeguatezza dei sistemi assiomatici che si erano venuti accumulando sulla scorta dei lavori pionieristici di Lewis e degli altri autori citati. Né ci si poteva accontentare di soluzioni «locali» che spiegassero il significato di *certi* operatori modali senza inquadrarlo in una spiegazione più ampia e sistematica della modalità, tale cioè da consentire un trattamento unificato di quel fenomeno generale che gli esempi in (2)–(8) illustrano soltanto in parte: le modalità aletiche (cioè quelle che esprimono necessità o possibilità in senso metafisico) accanto a quelle epistemiche, deontiche, temporali, spaziali, e così via. Era questo il limite di certe proposte che pure erano state avanzate da autori come John McKinsley, Alfred Tarski, e soprattutto Rudolf Carnap (il cui libro *Significato e necessità*, del 1947, costituì comunque un importante passo in avanti nella direzione di un approccio semantico ad ampio respiro). Ed era questo che, secondo Quine, non era possibile fare senza incappare in problemi insormontabili.

Ebbene, la pubblicazione degli articoli di Kripke citati in apertura venne a colmare precisamente questa lacuna. La teoria proposta da Kripke si configurava infatti come la naturale estensione ai linguaggi modali della semantica modellistica elaborata da Tarski per i linguaggi della logica estensionale tanto cara a Quine, condividendo con quella gli stessi standard di rigore matematico e fornendo al tempo stesso tutto il materiale necessario per separare le questioni strettamente logiche da quelle metafisiche o di altra natura. Per la verità, anche la teoria di Kripke riguardava inizialmente soltanto alcune modalità, ossia quelle rappresentate dagli operatori di necessità e possibilità che tanto avevano fatto discutere. Nella sua sostanza, tuttavia, la proposta presentava caratteristiche di straordinaria flessibilità che la rendevano suscettibile di generalizzazioni e applicazioni a tutto raggio, e nel giro di breve tempo la logica modale nel suo complesso cessò di essere un luogo di periferia disdegnato dai logici di professione per riconquistare una posizione di assoluta centralità nel panorama della logica contemporanea. (Bisogna dire che, contemporaneamente a Kripke, altri autori stavano lavorando a progetti simili. In particolare, alcune pubblicazioni di Jaakko Hintikka, tra cui gli articoli *Modality and Quantification* del 1961 e *The Modes of Modality* del 1963, contengono una caratterizzazione semantica della logica modale molto simile a quella di Kripke, sebbene priva di risultati di adeguatezza rispetto ai sistemi assiomatici esistenti. Anche la tesi di dottorato di Stig Kanger *Provability in Logic*, pubblicata a Stoccolma nel 1957, contiene già un trattamento semantico basato su idee simili, sebbene più complesso e a

sua volta privo di risultati di adeguatezza. Alcune di queste idee sono già prefigurate anche nell'articolo di Richard Montague *Logical Necessity, Physical Necessity, Ethics, and Quantifiers*, pubblicato nel 1960 ma risalente al 1955, e nel libro di Arthur Prior *Time and Modality*, del 1957, per quanto limitato a certe modalità temporali. Infine, è da segnalare che proprio nel 1959, l'anno di pubblicazione del primo articolo di Kripke in cui si fornisce una dimostrazione di completezza per il sistema di Lewis chiamato S5 – di cui parleremo tra poco – Arnauld Bayart pubblicava sulla rivista belga «Logique et Analyse» un articolo nel quale dimostrava un risultato analogo, ma rispetto a una semantica diversa e piuttosto artificiosa. Insomma, la teoria e i risultati erano ormai nell'aria. Ma a Kripke va riconosciuto il merito di aver saputo formulare per la prima volta una proposta unitaria e sistematica, e di averlo fatto nella piena consapevolezza dei requisiti di rigore e intelligibilità imposti dalle critiche di Quine.)

Sul piano tecnico, la difficoltà principale che un adeguato trattamento semantico delle modalità doveva affrontare risiedeva nella natura non vero-funzionale degli operatori modali. Un operatore si dice vero-funzionale se le condizioni di verità o falsità di qualunque enunciato che si ottiene applicando quell'operatore è determinato interamente ed esclusivamente da quelle degli enunciati più semplici a cui si applica. Per esempio, il comune connettivo di negazione, corrispondente all'espressione «non si dà il caso che», o semplicemente «non», è vero-funzionale in quanto il valore di ciò che si ottiene applicando tale operatore a un enunciato qualsiasi è semplicemente l'opposto del valore di quell'enunciato:

(16) «*Non A*» è vero se *A* è falso, ed è falso se *A* è vero.

Analogamente, il connettivo binario per la congiunzione, corrispondente alla parola «e», è vero-funzionale in questo senso preciso, dal momento che la verità o falsità di qualunque congiunzione dipende esclusivamente da quella dei due congiunti di cui è composta:

(17) «*A e B*» è vero se *A* e *B* sono entrambi veri, altrimenti è falso.

Tutti gli operatori logici della logica estensionale standard sono vero-funzionali in questo senso. Ed è per questo motivo che il loro significato viene generalmente rappresentato mediante delle vere e proprie tabelle, o «tavole di verità», che possono essere lette alla stregua di istruzioni precise per calcolare la verità o falsità di un enunciato qualsiasi a partire dal valore degli enunciati «atomici» da cui è composto. È evidente, tuttavia, che da questo punto di vista gli operatori modali si comportano in modo differente. Sul piano sintattico tutti gli esempi

illustrati in (2)–(8) sono assimilabili a ciò che si ottiene applicando a (1) il connettivo di negazione:

(18) *Non si dà il caso che i cittadini paghino le tasse.*

Ma a differenza di quest'ultimo caso, sul piano semantico è chiaro che la falsità o verità degli enunciati in questione non è determinata interamente ed esclusivamente dalla verità o falsità dell'enunciato (1). Come già abbiamo notato, supponendo che (1) sia vero, cioè che i cittadini paghino davvero le tasse, non abbiamo modo di stabilire se è necessario che i cittadini paghino le tasse, o se si crede o si sa che pagano le tasse, o se sia obbligatorio che le paghino, e così via. Analogamente, supponendo che (1) sia falso, non abbiamo modo di stabilire se è almeno possibile che i cittadini paghino le tasse, o se si creda che le paghino (nel qual caso si tratterebbe di una credenza errata), o se sia comunque obbligatorio che le paghino (nel qual caso i cittadini starebbero violando una legge), e così via. In breve, i connettivi modali non si comportano come delle semplici funzioni definite sull'insieme dei valori {vero, falso}: la verità o falsità degli enunciati in cui compaiono dipende almeno in parte da fattori che esulano dalla semplice verità o falsità degli enunciati a cui si applicano. Del resto è ovvio che sia così: tutti gli operatori esemplificati in (2)–(8) sono operatori unari, al pari dell'operatore di negazione esemplificato in (16), e ciascuno ha un significato diverso; tuttavia esistono soltanto quattro funzioni unarie definite sull'insieme {vero, falso}:

(19)  $f_1(A)$  è vero se  $A$  è vero, ed è falso se  $A$  è falso.

(20)  $f_2(A)$  è vero (indipendentemente dalla verità o falsità di  $A$ ).

(21)  $f_3(A)$  è falso (indipendentemente dalla verità o falsità di  $A$ ).

(22)  $f_4(A)$  è vero se  $A$  è falso, ed è falso se  $A$  è vero.

Ne segue che le risorse sarebbero semplicemente insufficienti, oltre che gravemente inadeguate sul piano intuitivo (sicuramente nessuno dei connettivi modali citati corrisponde a una di queste funzioni, l'ultima delle quali non è altro che quella corrispondente al connettivo di negazione).

Ora, l'intuizione di Kripke – e degli altri autori che, come detto, si stavano occupando dello stesso problema – nasce dalla constatazione che almeno in certi casi il fattore aggiuntivo che interviene nella determinazione delle condizioni di verità di un enunciato modale sia piuttosto chiaro: si tratta di prendere in considerazione, non già o non solo il valore reale dell'enunciato di partenza, bensì (anche) il valore di quell'enunciato rispetto a situazioni alternative a quella reale. Quest'idea risulta particolarmente intuitiva nel caso delle modalità co-



siddette aletiche, cioè la necessità e la possibilità, per le quali la nozione di «situazione alternativa» corrisponde in buona sostanza alla nozione leibniziana di «mondo possibile». E infatti Kripke stesso non esitava a presentare la sua caratterizzazione semantica di questi operatori come una versione dell'intuizione di Leibniz richiamata sopra, che potremmo formulare in modo più esplicito nei termini seguenti:

- (23) «Necessariamente  $A$ » è vero (nel mondo attuale) se e solo se  $A$  è vero in ogni mondo possibile.
- (24) «Possibilmente  $A$ » è vero (nel mondo attuale) se e solo se  $A$  è vero in qualche mondo possibile.

Si trattava di riuscire a precisare meglio che cosa sia un «mondo possibile» senza impelagarsi in questioni metafisiche, e soprattutto si trattava di farlo in un modo sufficientemente flessibile da consentire un'eventuale estensione dell'intuizione leibniziana anche in riferimento ai diversi tipi di «situazioni alternative» chiamate in causa dagli altri operatori modali.

Nel primo articolo della serie, quello del 1959 (*A Completeness Theorem in Modal Logic*), la preoccupazione di Kripke era soprattutto di fornire una semantica chiara per le logiche modali aletiche, e per questo motivo soltanto la prima delle due questioni appena citate era affrontata in modo esplicito. La seconda questione, tuttavia, ricevette un trattamento esplicito già a partire dal primo articolo del 1963 (*Semantical Analysis of Modal Logic I. Normal Modal Propositional Calculi*) ed è soprattutto con riferimento alla formulazione più generale fornita in questo secondo lavoro che si è cominciato a parlare di «semantica alla Kripke». È utile, comunque, ripercorrere le tappe nella loro successione storica, perché è proprio nel passaggio dalla prima versione alla seconda che si coglie appieno la portata innovativa della proposta di Kripke. Lo faremo concentrandoci inizialmente sul frammento corrispondente alla cosiddetta logica proposizionale, dove gli unici operatori logici sono i connettivi (modali o vero-funzionali). In una sezione successiva ci occuperemo di illustrare brevemente come si possa estendere la trattazione anche al linguaggio più espressivo della logica dei predicati, inclusiva degli operatori logici di quantificazione.

### **1.3. La semantica di Kripke per le logiche modali proposizionali**

Per procedere in modo sistematico, è opportuno a questo punto introdurre un minimo di notazione formale. Dal punto di vista della logica proposizionale, possiamo pensare che il linguaggio abbia una struttura molto semplice. Innanzi-

tutto, il vocabolario includerà un certo numero di *enunciati atomici*, cioè enunciati che, come «Piove» oppure «Kripke è un filosofo», non includono alcun operatore logico. Dal momento che il loro contenuto non è rilevante, si è soliti rappresentare tali enunciati mediante semplici variabili enunciative, per le quali ci serviremo di lettere minuscole in corsivo: « $p$ », « $q$ », ecc. In secondo luogo, il vocabolario includerà espressioni per gli operatori logici, che in questo caso sono dei semplici connettivi che consentono di formare un numero potenzialmente infinito di enunciati complessi a partire da un numero finito di enunciati atomici. Per gli operatori vero-funzionali possiamo considerare i già citati connettivi per la negazione e per la congiunzione, che rappresentiamo rispettivamente con i simboli « $\sim$ » e « $\wedge$ », unitamente ai connettivi per la disgiunzione « $\vee$ », per il condizionale «se ... allora» e per il bicondizionale «se e solo se», per i quali useremo i simboli « $\rightarrow$ », « $\leftrightarrow$ » e « $\leftrightarrow$ ». Inoltre, assumeremo inizialmente che il linguaggio contenga anche due connettivi unari corrispondenti agli operatori modali «è necessario che» (o «necessariamente»), che rappresentiamo con il simbolo « $\Box$ », ed «è possibile che» (o «possibilmente»), che rappresentiamo con il simbolo « $\Diamond$ ». La sintassi del linguaggio si può quindi riassumere mediante regole ricorsive molto semplici: diremo innanzitutto che ogni variabile enunciativa è una formula ben-formata del linguaggio; e diremo che se  $A$  e  $B$  sono formule ben-formate qualsiasi, allora sono ben-formate anche le formule  $\sim A$ ,  $A \wedge B$ ,  $A \vee B$ ,  $A \rightarrow B$ ,  $A \leftrightarrow B$ ,  $\Box A$  e  $\Diamond A$  (eventualmente contrassegnate da parentesi per evitare ambiguità). Nient'altro è una formula ben-formata.

Ora, abbiamo visto che, sul piano semantico, la caratterizzazione della componente vero-funzionale è semplice. Chiamato *modello* qualunque funzione che assegni il valore V (vero) o il valore F (falso) alle variabili enunciative del linguaggio, possiamo stabilire se una formula ben-formata  $X$  è vera rispetto a un modello  $\alpha$  (in breve:  $\alpha \models X$ ) affidandoci alle seguenti condizioni ricorsive:

- |      |                                      |              |  |
|------|--------------------------------------|--------------|--|
| (25) | $\alpha \models p$                   | se e solo se | $\alpha(p) = V$                                    |
| (26) | $\alpha \models \sim A$              | ” ”          | non $\alpha \models A$                             |
| (27) | $\alpha \models A \wedge B$          | ” ”          | $\alpha \models A$ e $\alpha \models B$            |
| (28) | $\alpha \models A \vee B$            | ” ”          | $\alpha \models A$ o $\alpha \models B$            |
| (29) | $\alpha \models A \rightarrow B$     | ” ”          | se $\alpha \models A$ allora $\alpha \models B$    |
| (30) | $\alpha \models A \leftrightarrow B$ | ” ”          | $\alpha \models A$ se e solo se $\alpha \models B$ |

(Le condizioni (26) e (27) non sono che la controparte formale delle condizioni (16) e (17) richiamate sopra.) Poiché gli operatori modali non sono vero-funzionali, sappiamo che questa semplice procedura non può venire estesa a tutte le formule del linguaggio: la verità o falsità di un enunciato modale rispetto a un

dato modello *trascende* le informazioni fornite dal modello stesso. A questo punto, però, entra in gioco l'intuizione leibniziana. Un modello è, in fondo, l'analogo formale di un mondo possibile, dato che ci dice come stanno le cose rispetto a qualsiasi fatto descrivibile mediante un enunciato atomico. Quindi si potrebbe pensare di sfruttare quell'intuizione fissando le condizioni di verità di enunciati contenenti operatori modali attraverso un'opportuna quantificazione dei modelli rispetto ai quali sono veri gli enunciati a cui si applicano quegli operatori. In *Meaning and Necessity* Carnap proponeva proprio questo, sebbene in una terminologia un po' diversa, suggerendo di completare il quadro con le clausole seguenti:

- (31)  $\alpha \models \Box A$  se e solo se  $\beta \models A$  per ogni modello  $\beta$   
 (32)  $\alpha \models \Diamond A$  " "  $\beta \models A$  per qualche modello  $\beta$

Questa soluzione, tuttavia, è insoddisfacente per un motivo molto semplice: le uniche nozioni di necessità e di possibilità che corrispondono alle condizioni (31) e (32) sono quelle di necessità e possibilità *logica*. In altre parole, dato un modello qualsiasi, gli unici casi in cui un enunciato della forma  $\Box A$  risulterà vero sono quelli in cui  $A$  esprime una verità logica, per esempio una tautologia come

- (33) O piove o non piove,

e gli unici casi in cui un enunciato della forma  $\Diamond A$  risulta falso sono quelli in cui  $A$  esprime una falsità logica, per esempio una contraddizione come

- (34) Piove e non piove.

Ora, per quanto interessanti possano essere, è evidente che questi sono casi molto speciali. D'altra parte, ogni ulteriore qualifica in merito alla gamma dei modelli  $\beta$  su cui quantificare nel *definiens* di (31) e (32) rifletterebbe una presa di posizione ben precisa riguardo a ciò che è necessario o possibile in un senso diverso da quello puramente logico, e come tale esulerebbe dai compiti della logica stessa. Quindi?

È qui che la proposta di Kripke comincia a fare la differenza. L'idea iniziale, sviluppata nell'articolo del 1959, è di superare l'*impasse* definendo la nozione di verità, non rispetto a un modello nel senso suddetto, bensì rispetto a una *coppia* definita da un certo modello e da un insieme di modelli alternativi, ciascuno dei quali corrisponde intuitivamente a un mondo possibile rispetto al modello dato (che a sua volta può essere considerato rappresentativo del mondo attuale). Più precisamente, sia una *struttura-modello* qualunque coppia  $(\alpha, W)$

dove è  $W$  un insieme qualsivoglia di modelli e  $\alpha$  un elemento di  $W$ . Specificare una struttura di questo tipo significa stabilire come vanno le cose nel mondo attuale,  $\alpha$ , e come vanno in certi mondi giudicati possibili rispetto ad  $\alpha$ . In altre parole, significa dire com'è fatto il mondo e come potrebbe essere: la richiesta che  $\alpha$  faccia parte di  $W$  riflette semplicemente l'intuizione (già presente negli scritti di Aristotele) per cui ciò che è attuale è *ipso facto* possibile. Dal punto di vista logico, tuttavia, non siamo tenuti a specificare una struttura-modello particolare più di quanto non siamo tenuti a identificare un certo modello della logica estensionale con il mondo attuale. Dal punto di vista logico, tutto ciò che siamo tenuti a fare è fissare le condizioni di verità di ogni formula ben-formata per qualunque scelta accettabile, e ciò equivale a definire tali condizioni rispetto a strutture-modello arbitrarie, proprio come nella logica vero-funzionale le si definisce rispetto a modelli arbitrari. Chiarito questo aspetto, le clausole definitorie proposte da Kripke dovrebbero a questo punto essere ovvie: si tratta di riscrivere (25)–(32) sostituendo al modello arbitrario  $\alpha$  una struttura-modello arbitraria  $(\alpha, W)$  e relativizzando le condizioni sugli operatori modali ai mondi possibili specificati da quest'ultima. Otteniamo in questo modo la definizione ricorsiva di verità di una formula ben-formata  $X$  rispetto a una struttura-modello  $(\alpha, W)$ . Schematicamente:

(25')	$(\alpha, W) \models p$	se e solo se	$\alpha(p) = V$
(26')	$(\alpha, W) \models \sim A$	” ”	non $(\alpha, W) \models A$
(27')	$(\alpha, W) \models A \wedge B$	” ”	$(\alpha, W) \models A$ e $(\alpha, W) \models B$
(28')	$(\alpha, W) \models A \vee B$	” ”	$(\alpha, W) \models A$ o $(\alpha, W) \models B$
(29')	$(\alpha, W) \models A \rightarrow B$	” ”	se $(\alpha, W) \models A$ allora $(\alpha, W) \models B$
(30')	$(\alpha, W) \models A \leftrightarrow B$	” ”	$(\alpha, W) \models A$ se e solo se $(\alpha, W) \models B$
(31')	$(\alpha, W) \models \Box A$	” ”	$(\beta, W) \models A$ per ogni $\beta$ in $W$
(32')	$(\alpha, W) \models \Diamond A$	” ”	$(\beta, W) \models A$ per qualche $\beta$ in $W$

(Si noti che le prime sei clausole non comportano alcuna differenza rispetto a (25)–(30), dato che il loro *definiens* è del tutto indifferente alla composizione di  $W$ .) Dopo di che possiamo introdurre tutte le comuni nozioni logiche, ivi inclusa quella di *validità*, facendo riferimento alle condizioni di verità così definite:

(35) Una formula ben-formata  $A$  è logicamente valida se e solo se  $A$  è vera rispetto a ogni struttura-modello.

A titolo illustrativo, è utile notare che anche in questo caso tutte le tautologie e, più in generale, tutte le formule valide risultano necessarie rispetto a qualunque struttura-modello. Più precisamente, se  $A$  è logicamente valida, allora

avremo che  $(\beta, W) \models A$  per qualunque scelta di  $W$  e per qualunque (quindi per ogni)  $\beta$  in  $W$ , da cui segue che  $\Box A$  risulterà vera rispetto a qualunque struttura-modello  $(\alpha, W)$ . Tuttavia la conversata non vale: non è detto che  $\Box A$  risulti vera rispetto a una struttura-modello *solo se*  $A$  è logicamente valida, che è come dire che in questa semantica la nozione di necessità non si esaurisce in quella di necessità logica. Per esempio, se  $\alpha$  è il nostro mondo e  $W$  contiene soltanto mondi compatibili con le nostre leggi fisiche, allora la formula corrispondente a un enunciato come

- (36) È necessario che l'attrazione gravitazionale tra due corpi sia direttamente proporzionale al prodotto delle loro masse.

risulterà vero rispetto alla struttura-modello  $(\alpha, W)$  malgrado la legge di gravitazione universale non esprima una tautologia. Per lo stesso motivo, la nozione di possibilità non si esaurisce in quella di possibilità logica: se  $A$  è una contraddizione,  $\Diamond A$  risulta falsa (cioè non vera) in ogni struttura-modello, ma non è detto che  $\Diamond A$  risulti falsa *solo se*  $A$  è una contraddizione. Per esempio, nel modello appena citato la formula corrispondente all'enunciato

- (37) È possibile che l'attrazione gravitazionale tra due corpi sia inversamente proporzionale al prodotto delle loro masse.

risulterà falsa, sebbene la legge di gravitazione universale possa essere stravolta senza con ciò cadere in contraddizione sul piano logico. Tanto basti a illustrare come la semantica in questione consenta di superare i limiti evidenziati in precedenza con riferimento alla semantica carnapiana definita in (25)–(32).

Ora, nel suo articolo Kripke non si limitava a esporre la teoria appena delineata, ma aggiungeva tre osservazioni importanti. La prima è che la teoria stessa è formulata interamente in un metalinguaggio che fa uso soltanto di concetti appartenenti alla teoria degli insiemi e alla logica estensionale dei predicati (come si può constatare leggendo le clausole corrispondenti a ciascun *definiens*). Questa caratteristica, anticipata banalmente dalla teoria di Carnap ma condivisa anche dalle teorie più ricche che esamineremo tra breve, era particolarmente importante nel contesto storico a cui Kripke si rivolgeva, poiché consentiva di rispondere almeno in parte allo scetticismo di coloro i quali, sulla scia di Quine, ritenevano che tali concetti fossero gli unici a cui la logica potesse affidarsi senza indugio. In particolare, la valutazione di enunciati contenenti modalità iterate o complesse, come negli esempi in (9)–(11), non presenta problemi di sorta, dal momento che si tratta semplicemente di applicare le regole in (31') e (32') in modo ricorsivo.

La seconda osservazione è che la semantica in questione risulta adeguata al più forte dei sistemi di logica modale introdotti anni prima da Lewis e dai suoi collaboratori, noto come sistema S5. Più precisamente, Kripke dimostrava che la nozione di validità definita in (35) individua esattamente le formule che risultano dimostrabili in tale sistema, di cui riportiamo gli assiomi specifici e la regola di inferenza:

- A1  $\diamond A \leftrightarrow \neg \Box \neg A$
- A2  $\Box A \rightarrow A$
- A3  $\diamond A \rightarrow \Box \diamond A$
- A4  $\Box(A \rightarrow B) \rightarrow (\Box A \rightarrow \Box B)$
- R Se  $A$  è un teorema, allora si può inferire  $\Box A$

Anche questo fatto è molto importante, non solo perché equivale alla prima dimostrazione di *completezza e coerenza* di un sistema di logica modale (sebbene in contemporanea alla dimostrazione di Bayart), ma anche perché S5 era da molti autori ritenuto il sistema più adatto a rappresentare le nozioni di necessità e di possibilità aletica. Tra le altre cose, S5 include i seguenti quattro teoremi:

- (38)  $\Box \Box A \leftrightarrow \Box A$
- (39)  $\Box \diamond A \leftrightarrow \diamond A$
- (40)  $\diamond \diamond A \leftrightarrow \diamond A$
- (41)  $\diamond \Box A \leftrightarrow \Box A$

da cui segue che le modalità iterate sono in ultima analisi superflue. Quindi non solo la semantica di Kripke consente di valutare enunciati modali di complessità arbitraria; essa ne determina anche la riducibilità alle due modalità basilari.

Infine, e questo è il punto cruciale, nel paragrafo conclusivo Kripke osservava che la semantica in questione può essere opportunamente generalizzata in modo da rendere giustizia anche a intuizioni diverse rispetto a quelle codificate nel sistema S5, ivi incluse le intuizioni di coloro che si opporrebbero alle equivalenze in (38)–(41). È precisamente in questa generalizzazione che, come dicevamo, risiede la portata veramente innovativa della proposta di Kripke, ed è a questo compito che sono dedicati gli articoli successivi, a partire dal già citato *Semantical Analysis of Modal Logic I* del 1963.

#### 1.4. Generalizzazione della semantica proposizionale

L'idea è molto semplice. Nelle strutture-modello considerate sin qui, la nozione di mondo possibile è, per così dire, «assoluta»: dato l'insieme  $W$ , ciascun elemento di  $W$  è giudicato possibile relativamente a ogni altro elemento. Ciò è evi-

dente dalle clausole relative agli operatori modali, nelle quali la valutazione di un enunciato rispetto alla struttura-modello data,  $(\alpha, W)$ , richiede che si esaminino tutte le strutture-modello  $(\beta, W)$  dove  $\beta$  è un elemento di  $W$ . Si potrebbe tuttavia pensare che la nozione di mondo possibile sia «relativa», nel senso che i mondi giudicati possibili dal punto di vista di un certo mondo in  $W$  potrebbero non essere giudicati possibili nel momento in cui ci si ponga dal punto di vista di un altro mondo in  $W$ , o viceversa. Per esempio, noi riusciamo benissimo a immaginare dei mondi nei quali le cose funzionano molto meglio che nel mondo attuale, ma non è detto che in quei mondi si riesca a immaginare un mondo come il nostro. Oppure: a noi sembra di poter immaginare soltanto mondi soggetti alla legge di gravitazione universale, ma non è detto che tutti i mondi che riusciamo a immaginare siano a loro volta legati a questo modo di vedere le cose, ossia che l'esistenza di un mondo che trasgredisce la legge di gravitazione sia un'assurdità anche dal *loro* punto di vista. Ovviamente può essere difficile stabilire a priori tutta la rete di possibilità relative che può caratterizzare un certo insieme di mondi. Ma proprio questo è il punto: la semantica della sezione precedente impone surrettiziamente un punto di vista ben preciso, ossia quello in base al quale tutti i mondi possibili sono sullo stesso piano (col risultato che le modalità iterate risultano sempre ridondanti). Se vogliamo rimanere neutrali in proposito – e nel contesto di una semantica formale questo sembra l'atteggiamento giusto – dobbiamo trovare il modo di relativizzare la nozione di mondo possibile. E il modo più semplice per farlo è quello di richiedere che una struttura-modello specifichi, non già un insieme generico di mondi, ma un insieme di mondi su cui è definita una certa relazione di possibilità relativa, o «accessibilità». Dopo di che possiamo richiedere che le condizioni di verità delle formule ben-formate contenenti operatori modali dipendano esclusivamente dall'ispezione di *certe* alternative: non tutti i mondi possibili, ma solo quei mondi che sono accessibili, rispetto alla relazione in questione, dal mondo che la struttura-modello considera attuale.

È questo l'ingrediente principale della generalizzazione additata da Kripke (e anticipata in vario modo da autori come Kanger e Hintikka). A partire dal 1963, la nozione di struttura-modello viene arricchita da un terzo ingrediente, corrispondente alla relazione di accessibilità fra mondi possibili, e la definizione ricorsiva di verità viene modificata di conseguenza. Più precisamente, per struttura-modello si intende adesso una *tripla*  $(\alpha, W, R)$  – dove  $W$  e  $\alpha$  sono, come prima, rispettivamente un insieme di modelli e un elemento di quell'insieme, e dove  $R$  è una relazione binaria su  $W$  – e le condizioni di verità di una formula ben-formata rispetto a una struttura-modello di questo tipo diventano:

(25'')	$(\alpha, W, R) \models p$	se e solo se	$\alpha(p) = V$
(26'')	$(\alpha, W, R) \models \sim A$	" "	non $(\alpha, W, R) \models A$
(27'')	$(\alpha, W, R) \models A \wedge B$	" "	$(\alpha, W, R) \models A$ e $(\alpha, W, R) \models B$
(28'')	$(\alpha, W, R) \models A \vee B$	" "	$(\alpha, W, R) \models A$ o $(\alpha, W, R) \models B$
(29'')	$(\alpha, W, R) \models A \rightarrow B$	" "	se $(\alpha, W, R) \models A$ allora $(\alpha, W, R) \models B$
(30'')	$(\alpha, W, R) \models A \leftrightarrow B$	" "	$(\alpha, W, R) \models A$ se e solo se $(\alpha, W, R) \models B$
(31'')	$(\alpha, W, R) \models \Box A$	" "	$(\beta, W, R) \models A$ per ogni $\beta$ in $W$ tale che $\alpha R \beta$
(32'')	$(\alpha, W, R) \models \Diamond A$	" "	$(\beta, W, R) \models A$ per qualche $\beta$ in $W$ tale che $\alpha R \beta$

(Anche in questo caso, come già nel passaggio da (25)–(32) a (25')–(32'), la novità diventa significativa soltanto nelle ultime due clausole, cioè quelle relative agli operatori modali. Inoltre è evidente che la semantica definita da (25')–(32') può essere vista come un caso limite della presente, ottenuto richiedendo che  $R$  sia la relazione *universale* su  $W$ , proprio come la semantica definita da (25)–(32) può essere vista come il caso limite ottenuto richiedendo altresì che  $W$  sia l'insieme di *tutti* i modelli.)

L'effetto di questa modifica è tanto immediato quanto straordinario sul piano della flessibilità che ne deriva: nel momento in cui la verità o falsità di una formula modale dipendono dalla relazione  $R$ , studiando le proprietà di questa relazione si potranno individuare le condizioni alle quali certi principi modali risultano validi. Più precisamente, si consideri la seguente generalizzazione della nozione di validità logica definita in (35):

- (42) Una formula ben-formata  $A$  è logicamente valida rispetto a una classe  $C$  di strutture-modello se e solo se  $A$  è vera rispetto a ogni elemento di  $C$ .

Il problema dell'adeguatezza semantica di un sistema di logica modale si traduce, allora, nel problema di identificare un'opportuna classe di strutture-modello rispetto alla quale l'insieme delle formule valide coincide con quello delle formule dimostrabili nel sistema. E questo problema è interamente riconducibile a quello di determinare le condizioni su  $R$  che definiscono la classe in questione. In particolare, Kripke dimostrava che se  $C$  è la classe delle strutture-modello in cui  $R$  è una relazione di equivalenza, ossia soddisfa le condizioni seguenti:

- |      |  |                     |
|------|--|---------------------|
| (43) | per ogni $\beta$ in $W$ : $\beta R \beta$  | <i>Riflessività</i> |
| (44) | per ogni $\beta$ e $\gamma$ in $W$ : $\beta R \gamma$ solo se $\gamma R \beta$                             | <i>Simmetria</i>    |
| (45) | per ogni $\beta, \gamma$ e $\delta$ in $W$ : $\beta R \gamma$ e $\gamma R \delta$ solo se $\beta R \delta$ | <i>Transitività</i> |

allora le formule valide rispetto a  $C$  sono ancora tutti e soli i teoremi del sistema



S5. Ciò significa che una classe siffatta risulta equipollente alla classe delle strutture-modello in cui  $R$  è universale. Kripke però dimostrava anche che se si impongono soltanto le prime due condizioni, cioè se consideriamo la classe  $C$  di tutte le strutture-modello in cui  $R$  è riflessiva e simmetrica, allora le formule  $C$ -valide sono i teoremi di un sistema più debole chiamato  $B$  (o «Brouweriano»), che differisce da  $S5$  per la presenza dell'assioma seguente in luogo di  $A3$ :

$$A3' \quad A \rightarrow \Box \Diamond A$$

Se imponiamo soltanto le condizioni di riflessività e di transitività, allora le formule  $C$ -valide sono i teoremi del cosiddetto sistema  $S4$  di Lewis, che a sua volta differisce da  $S5$  per la presenza dell'assioma seguente in luogo di  $A3$ :

$$A3'' \quad \Box A \rightarrow \Box \Box A$$

E se imponiamo soltanto la condizione di riflessività, allora le formule  $C$ -valide sono i teoremi del cosiddetto sistema  $T$  (dovuto a Robert Feys e, indipendentemente, a Georg von Wright), che differisce da  $S5$  semplicemente per l'assenza di  $A3$ . Possiamo anche osservare che se non imponiamo alcuna condizione su  $R$ , cioè se identifichiamo  $C$  con la classe di *tutte* le strutture-modello  $(\alpha, W, R)$ , allora le formule logicamente valide sono esattamente quelle dimostrabili nel sistema logico definito soltanto dalla regola  $R$  e dagli assiomi  $A1$  e  $A4$ : questo sistema, oggi noto come  $K$  (in onore di Kripke) rappresenta quindi la più debole tra le logiche modali analizzabili nei termini della semantica in esame, e la presenza di  $A1$  conferma l'ineluttabilità del nesso tra necessità e possibilità già catturato nella tesi aristotelica citata in precedenza: necessario è ciò che non può non essere.

Evidentemente questi risultati vennero accolti con estremo interesse, sia perché costituivano altrettante dimostrazioni di completezza e coerenza per alcuni importanti sistemi di logica modale presenti nella letteratura, sia perché consentivano di vedere sotto una nuova luce l'analisi della modalità nel suo complesso: lungi dal dover affrontare ogni caso sulla base della semplice intuizione, l'accettabilità o meno di certi principi modali (tra cui quelli esprimenti la ridondanza delle modalità iterate) si traduceva sistematicamente nell'accettabilità o meno di certe condizioni sulla relazione di accessibilità, e quest'ultima questione non presentava difficoltà particolari. In altre parole, il confronto tra i diversi sistemi di logica modale diventava una questione che poteva essere affrontata a livello semantico affidandosi interamente ai principi della teoria degli insiemi e della logica delle relazioni, e ciò non poteva che riscuotere interesse anche da parte dei logici più scettici. Nel giro di breve tempo, lo studio siste-

matico delle logiche modali divenne un campo di ricerca assai fecondo, al punto che intere famiglie di sistemi logici sono state messe a punto partendo direttamente dallo studio delle condizioni che potrebbero essere soddisfatte da  $R$ . A titolo illustrativo, ecco qualche altro esempio delle corrispondenze – non sempre elementari – stabilite negli anni successivi alla pubblicazione dell’articolo di Kripke:

- |    |  |   |
|----|--|---|
| P1 | $\Box A \rightarrow \Diamond A$          | per ogni $\beta$ in $W$ esiste qualche $\gamma$ in $W$ tale che: $\beta R \gamma$   |
| P2 | $\Box \Box A \rightarrow \Box A$         | per ogni $\beta$ e $\gamma$ in $W$ : $\beta R \gamma$ solo se esiste qualche $\delta$ in $W$ tale che $\beta R \delta$ e $\delta R \gamma$                  |
| P3 | $\Box \Diamond A \rightarrow \Diamond A$ | per ogni $\beta$ in $W$ esiste qualche $\gamma$ in $W$ tale che: $\beta R \gamma$ e, per ogni $\delta$ in $W$ , $\gamma R \delta$ solo se $\delta = \gamma$ |
| P4 | $\Box(\Box A \rightarrow A)$             | per ogni $\beta$ e $\gamma$ in $W$ : $\beta R \gamma$ solo se $\gamma R \gamma$   |

(Queste corrispondenze vanno lette nel modo seguente: sia  $S$  un sistema di logica modale adeguato rispetto a una certa classe  $C$  di strutture-modello. Allora i teoremi dell’estensione di  $S$  ottenuta aggiungendo come assioma uno qualsiasi dei principi indicati a sinistra coincidono esattamente le formule che risultano valide rispetto alla sottoclasse di  $C$  determinata dalla condizione su  $R$  indicata a destra.)

Senza entrare in ulteriori dettagli tecnici, possiamo a questo punto completare il quadro limitandoci a tre considerazioni di carattere generale. Innanzitutto, l’apparato semantico che abbiamo esposto può essere ulteriormente generalizzato astraendo completamente dall’identificazione dei mondi possibili con i modelli della semantica vero-funzionale. Kripke stesso, per evitare ogni discussione di carattere metafisico sulla natura dei *mondes possibles* di memoria leibniziana, evidenziava questo fatto nell’articolo del 1963, proponendo di trattare  $W$  come un insieme qualsiasi e di affidare a una funzione di valutazione *esterna* alla struttura-modello il compito di associare a ciascuna variabile enunciativa uno dei due valori V o F. Questa decisione richiede una leggera modifica alla definizione ricorsiva di verità in (25’)-(32’), che verrebbe a dipendere non soltanto dalla struttura-modello di riferimento ma anche dalla funzione di valutazione prescelta. Sul piano logico, tuttavia, la modifica sarebbe ininfluenza e questo è sicuramente un pregio dell’approccio.

In secondo luogo è importante sottolineare che, a prescindere da tale eventuale modifica, a questo punto si dispone di un apparato formale che in linea di principio si presta all’analisi semantica di qualunque operatore modale, non soltanto di quelli corrispondenti alle modalità aletiche di necessità e possibilità. Per esempio, supponiamo di attribuire a « $\Box$ » il significato dell’operatore epi-

stemico «si sa che» (e a « $\diamond$ » il significato indiretto di «non si sa che non» garantito dall'assioma A1). In tal caso Hintikka suggeriva di concepire i mondi accessibili alla stregua di situazioni epistemicamente accettabili, cioè mondi che coincidono con il mondo attuale per ciò che concerne tutti i fatti di cui si è a conoscenza: asserire un enunciato della forma  $\Box A$  significa asserire ciò che è vero in tutti questi mondi. Dopo di che è naturale pensare che  $R$  debba essere quantomeno riflessiva, e quindi che l'assioma A2 debba risultare valido, in ottemperanza al principio per cui si può sapere qualcosa soltanto se è confermato dalla realtà. Per contro, se attribuiamo a « $\Box$ » il significato dell'operatore «si crede che», possiamo seguire ancora Hintikka nel concepire i mondi accessibili alla stregua di mondi doxasticamente accettabili, cioè mondi che concordano nel verificare le nostre credenze. In tal caso, però, sarebbe irragionevole richiedere che  $R$  sia riflessiva, e quindi che valga l'assioma A2: possiamo benissimo avere credenze che non corrispondono alla realtà dei fatti. Semmai potremmo pensare che le credenze debbano essere coerenti, cioè che non si possa razionalmente credere  $A$  e al contempo credere la negazione di  $A$ . In presenza di A1 ciò equivale ad assumere il principio P1 citato sopra, e quindi  $R$  dovrà soddisfare la condizione di «serialità» che abbiamo indicato accanto a questo principio. Un discorso analogo vale per la lettura deontica di « $\Box$ » come «è obbligatorio che» (e di « $\diamond$ » come «è lecito che»). Seguendo una proposta che risale già a Kanger, possiamo in questo caso concepire i mondi accessibili alla stregua di mondi «ideali» o «deonticamente perfetti» nei quali tutti gli obblighi sono realizzati, e siccome non è detto che il mondo attuale sia tra quelli (anzi, è ragionevole supporre che da noi vi siano diversi obblighi inevasi), la riflessività di  $R$  risulterebbe troppo forte e, con essa, l'assioma A2. Tuttavia si potrebbe pensare che sia comunque obbligatorio adempiere agli obblighi, cioè che valga almeno il principio P4 citato sopra e, di conseguenza, che  $R$  debba soddisfare la condizione di «riflessività secondaria» che abbiamo indicato accanto a questo principio. Inoltre, nella misura in cui pensiamo che gli obblighi, al pari delle credenze razionali, debbano essere coerenti, richiederemo anche in questo caso che  $R$  sia seriale e che valga il principio P1. Insomma, si capisce che a questo punto si apre un intero ventaglio di opzioni, e il pregio della semantica di Kripke è proprio quello di consentire un loro trattamento unitario e formalmente ben fondato. In linea di principio, essa si presta addirittura al trattamento di linguaggi irriducibilmente «multimodali», cioè linguaggi contenenti una molteplicità di operatori modali diversi e mutuamente indipendenti: sarà sufficiente postulare una relazione di accessibilità per ciascun tipo di modalità e studiarne le rispettive proprietà.

L'ultima considerazione riguarda gli eventuali limiti di questo metodo<sup>1</sup>. Abbiamo visto che la semantica di Kripke consente di analizzare in modo uniforme intere famiglie di sistemi assiomatici: basta variare le condizioni sulla relazione di accessibilità. Ma è così in *tutti* i casi? Si può procedere in questo modo per fissare la semantica di *qualsiasi* logica modale di un certo interesse? La risposta iniziale è negativa, ed è stato Kripke stesso ad affrontare la questione nel seguito dell'articolo del 1963, pubblicato nel 1965 col titolo *Semantical Analysis of Modal Logic II. Non-normal Modal Propositional Calculi*. Già ai tempi di Kripke esistevano logiche modali più «deboli» del sistema K che, come abbiamo già notato, rappresenta il sistema minimale tra quelli analizzabili nei termini della semantica in esame. Per esempio, due dei sistemi proposti da Lemmon alla fine degli anni Cinquanta, noti come E2 e E3, avevano la caratteristica di non contenere nessun teorema della forma  $\Box A$ ; ma abbiamo già visto che, quando  $A$  è una formula logicamente valida,  $\Box A$  risulta vera in qualunque struttura-modello e, quindi, valida a sua volta. Sul piano sintattico, ciò si riflette nella regola R di K, che consente di derivare  $\Box A$  da qualsiasi teorema  $A$  (e che non è inclusa nei sistemi E2 e E3). Anche due dei sistemi di Lewis, noti come S2 e S3, avevano la caratteristica di non contenere la regola R, ed è naturale pensare che questa caratteristica debba essere condivisa anche da certe logiche modali non aletiche. Per esempio, dal fatto che  $A$  sia una tautologia non segue che si sappia o si creda che sia una tautologia (potrebbe essere di forma estremamente complessa), e nemmeno che si sappia o si creda che  $A$  sia vera; quindi la validità di  $A$  non sembra implicare quella di  $\Box A$  in nessuno dei due sensi epistemici di « $\Box$ » (salvo presupporre una forma di onniscienza logica che secondo molti autori esula dal nostro concetto di conoscenza o di credenza razionale). Per lo stesso motivo, non tutte le logiche della conoscenza e della credenza includono l'assioma A4, che consentirebbe di derivare sempre  $\Box B$  dalla congiunzione di  $\Box(A \rightarrow B)$  e  $\Box A$ . In breve: non tutte le logiche modali includono il sistema K, ma la regola e gli assiomi di K sono automaticamente soddisfatti nelle strutture-modello di Kripke. In questo senso, l'applicabilità della teoria semantica del 1963 non è illimitata. Ciononostante, sarebbe un errore concludere che l'*approccio* a cui si ispira tale semantica sia intrinsecamente inadatto a trattare queste logiche modali. Si tratta solo di rilassare ulteriormente la nozione di struttura-modello in modo da consentire maggiore flessibilità, e l'articolo del 1965 conteneva due proposte molto semplici in questo senso. Per limitarci a

---

<sup>1</sup> Questa considerazione, di carattere più tecnico, può essere omessa senza pregiudicare la comprensione del resto del capitolo.

quella entrata maggiormente nell'uso comune, Kripke suggeriva di classificare i mondi possibili inclusi nell'insieme  $W$  in due categorie disgiunte: i mondi «normali» e quelli «non normali» (nei quali, per esempio, l'onniscienza logica non vige). Dopo di che è sufficiente decretare che mentre i mondi del primo tipo obbediscono all'intuizione leibniziana per ciò che concerne la valutazione delle formule modali, i mondi del secondo tipo trattano gli enunciati della forma  $\Box A$  come automaticamente falsi (e quelli della forma  $\Diamond A$  come automaticamente veri). Sul piano formale, questo significa arricchire ulteriormente la nozione di struttura-modello aggiungendo un quarto ingrediente. Una struttura-modello sarà cioè una *quadrupla*  $(\alpha, W, R, N)$  dove  $\alpha$ ,  $W$  e  $R$  sono come sopra e  $N$  è un sottoinsieme di  $W$ , da pensarsi appunto come l'insieme dei mondi normali. Tale modifica non ha alcun effetto sulle condizioni di verità degli enunciati non modali, che possono quindi essere importate direttamente da (25'')–(30''). Le condizioni di verità per gli enunciati contenenti uno dei due operatori modali, invece, diventano:

- (31''')  $(\alpha, W, R, N) \models \Box A$  se e solo se  $\alpha$  è in  $N$  e  $(\beta, W, R, N) \models A$  per ogni  $\beta$  in  $W$  tale che  $\alpha R \beta$   
(32''')  $(\alpha, W, N, R) \models \Diamond A$  " "  $\alpha$  non è in  $N$  o  $(\beta, W, R, N) \models A$  per qualche  $\beta$  in  $W$  tale che  $\alpha R \beta$

Su queste basi, formulando opportune condizioni sulla relazione  $R$  e sulla composizione di  $N$ , Kripke dimostrava la completezza di una varietà di logiche più deboli di  $K$ . In particolare,  $E2$  risulta completa rispetto alla classe di tutte le strutture-modello, mentre  $E3$ ,  $S2$  e  $S3$  risultano complete nelle classi di strutture-modello in cui, rispettivamente,  $R$  è transitiva,  $N$  contiene  $\alpha$ , e  $R$  è transitiva e  $N$  contiene  $\alpha$ . La completezza di certe logiche epistemiche, come pure di certe logiche deontiche più deboli di  $K$ , si può dimostrare nello stesso modo.

### 1.5. Estensioni alla logica dei predicati

Con tutto ciò, è difficile sopravvalutare la portata complessiva dei tre articoli di Kripke su cui ci siamo soffermati. All'età di soli 25 anni, questo giovane filosofo che all'asilo aveva imparato da solo a leggere e scrivere in ebraico e che già in quarta elementare aveva divorato le opere di Shakespeare si imponeva di prepotenza all'attenzione della comunità scientifica internazionale, ridisegnando con pochi tratti l'intera mappa della logica modale e collocandola in modo definitivo al centro d'interesse della logica formale contemporanea. Resta solo da aggiungere, e non è cosa da poco, che sebbene qui ci si sia soffermati sulla

logica modale proposizionale, la proposta iniziale di Kripke includeva anche un trattamento estensivo delle logiche modali quantificate, cioè di quelle logiche che derivano dall'aggiunta degli operatori modali al linguaggio di base della logica dei predicati. I dettagli a questo riguardo sono piuttosto complessi e non è il caso di appesantire eccessivamente l'esposizione. D'altra parte, abbiamo visto che i dubbi più seri di Quine riguardavano proprio le logiche modali quantificate, che già negli anni Quaranta e Cinquanta avevano registrato un certo sviluppo per opera di Carnap e soprattutto in una serie di articoli di Ruth C. Barcan. Inoltre, è proprio questa dimensione della teoria che fornisce le basi per alcune delle tesi più interessanti successivamente sviluppate da Kripke in materia di metafisica e di filosofia del linguaggio, alle quali sono dedicati gli altri capitoli del presente volume. Completiamo pertanto questa prima parte del capitolo ricostruendo almeno i tratti essenziali della teoria estesa.

Il linguaggio della logica dei predicati (o logica quantificazionale) si distingue da quello della logica proposizionale sotto due aspetti. Innanzitutto, gli enunciati atomici non sono più trattati alla stregua di entità prive di struttura, ma sono definiti a partire da un vocabolario costituito da espressioni più semplici, che possiamo distinguere in due gruppi: termini singolari – o nomi – come «Saul Kripke» o «Barack Obama», e termini generali – o predicati – di adicITÀ variabile, come «è un filosofo» (monadico) oppure «è più giovane di» (diadico). Se stabiliamo di usare le lettere minuscole dell'alfabeto corsivo per designare espressioni del primo tipo e le lettere maiuscole per designare quelle del secondo tipo, gli enunciati atomici possono allora essere rappresentati da formule che consistono di una lettera maiuscola (predicato) concatenato con, e più precisamente seguito da, il giusto numero di lettere minuscole (nomi). Per esempio, l'enunciato «Saul Kripke è un filosofo» può essere rappresentato dalla formula  $Fs$  e l'enunciato «Barack Obama è più giovane di Saul Kripke» dalla formula  $Gbs$ . In generale, un enunciato atomico sarà quindi una sequenza della forma  $Pt_1\dots t_n$ , dove  $P$  è un predicato  $n$ -adico e  $t_1, \dots, t_n$  sono  $n$  nomi qualsiasi. In secondo luogo, il vocabolario logico è arricchito dalla presenza dei quantificatori, che sono quegli operatori logici che corrispondono alle espressioni «ogni» e «qualche» (e simili). Generalmente questi operatori sono rappresentati dai simboli « $\forall$ » e « $\exists$ », nell'ordine, e vengono usati in combinazione con dei simboli speciali detti variabili individuali, per le quali ci accontenteremo di usare le lettere « $x$ » e « $y$ » (sebbene le variabili possano essere in numero infinito). Sulla base di questo vocabolario, la sintassi del linguaggio si può riassumere mediante tre regole ricorsive molto semplici: innanzitutto, ogni enunciato atomico è una formula ben-formata del linguaggio; in secondo luogo, come nella

logica proposizionale, se  $A$  e  $B$  sono formule ben-formate qualsiasi, allora sono ben-formate anche le formule  $\sim A$ ,  $A \wedge B$ ,  $A \vee B$ ,  $A \rightarrow B$ ,  $A \leftrightarrow B$ ,  $\Box A$  e  $\Diamond A$ ; infine, se  $A$  è una formula ben-formata, allora sono ben-formate anche tutte le formule del tipo  $\forall x A_t^x$  e  $\exists x A_t^x$ , dove  $x$  è una variabile,  $t$  è un nome, e  $A_t^x$  è ciò che si ottiene da  $A$  sostituendo zero o più occorrenze di  $t$  con  $x$ . Per esempio,  $\forall x Fx$  e  $\exists x (Gbx \wedge Fx)$  sono formate in questo modo a partire da  $Fs$  e da  $Gbs \wedge Fs$  (sostituendo in ciascun caso tutte le occorrenze di  $s$ ) e possono essere intese come la traduzione formale degli enunciati quantificati «Ogni cosa è un filosofo» e «Barack Obama è più giovane di qualche filosofo». Si noti che anche in questo caso ci serviamo liberamente delle parentesi per evitare possibili ambiguità.

Fissato così il linguaggio della logica dei predicati, è evidente che la semantica non può essere determinata in modo adeguato direttamente sulla base delle strutture-modello della logica proposizionale, dato che il valore di verità di un enunciato atomico – e derivatamente quello di ogni formula ben-formata – dovrà dipendere dal significato attribuito ai termini di cui è costituito. Se ci limitiamo per un momento al frammento del linguaggio che non contiene operatori modali, ciò significa che un modello non può essere identificato semplicemente con una funzione che assegna uno dei valori V o F agli enunciati atomici in modo arbitrario. Piuttosto, identificheremo un modello con una funzione che, dato un certo dominio di oggetti D, assegnerà a ogni nome un elemento di D e a ogni predicato  $n$ -adico una funzione che, per ogni  $n$ -pla ordinata  $\langle d_1, \dots, d_n \rangle$  di elementi di D, ci dirà se quella  $n$ -pla soddisfa o meno il predicato in questione attribuendole rispettivamente il valore V o il valore F. Per esempio, se D include le persone che esistono nel nostro mondo, possiamo immaginare che il modello assegni al nome  $b$  l'individuo Barack Obama, al nome  $s$  l'individuo Saul Kripke, al predicato unario  $F$  la funzione che attribuisce il valore V a tutti i filosofi esistenti, e al predicato binario  $G$  la funzione che attribuisce il valore V alle coppie  $\langle d_1, d_2 \rangle$  tali che  $d_1$  è più giovane di  $d_2$ . Dopo di che, se indichiamo con  $\alpha_D$  un modello di questo tipo, che chiameremo «modello su D», la caratterizzazione standard delle condizioni di verità per gli enunciati non modali del linguaggio si può formulare nel modo seguente:

- |      |                                    |              |   |
|------|------------------------------------|--------------|---|
| (46) | $\alpha_D \models P t_1 \dots t_n$ | se e solo se | $\alpha_D(P)(\alpha_D(t_1), \dots, \alpha_D(t_n)) = V$                    |
| (47) | $\alpha_D \models \sim A$          | ” ”          | non $\alpha_D \models A$  |
| (48) | $\alpha_D \models A \wedge B$      | ” ”          | $\alpha_D \models A$ e $\alpha_D \models B$                               |
|      | $\vdots$                           |              | $\vdots$  |
| (49) | $\alpha_D \models \forall x A$     | ” ”          | $\beta_D \models A_x^t$ per ogni $t$ -variante $\beta_D$ di $\alpha_D$    |
| (50) | $\alpha_D \models \exists x A$     | ” ”          | $\beta_D \models A_x^t$ per qualche $t$ -variante $\beta_D$ di $\alpha_D$ |

dove  $A'_x$  è ciò che si ottiene da  $A$  sostituendo ogni occorrenza di  $x$  con il termine  $t$ , che per definizione assumiamo essere il primo nome (in ordine alfabetico) che non compare in  $A$ , e dove per  $t$ -variante di  $\alpha_D$  si intende un modello su  $D$  che assegni un valore qualsiasi al termine  $t$  e che per il resto coincide con  $\alpha_D$ .

A titolo illustrativo, supponiamo che  $\alpha_D$  sia il modello parzialmente descritto sopra, dove  $D$  è l'insieme delle persone che esistono nel nostro mondo. In tal caso avremo sicuramente che  $\alpha_D \models Fs$ , dato che  $\alpha_D(s)$ , cioè l'individuo Saul Kripke, è tra quegli individui a cui  $\alpha_D(F)$  assegna il valore  $V$ , cioè i filosofi. Ciò riflette la verità di «Saul Kripke è un filosofo». Per contro, non avremo che  $\alpha_D \models Fb$ , dal momento che  $\alpha_D(b)$ , cioè Barack Obama, non è tra gli individui a cui  $\alpha_D(F)$  assegna il valore  $V$ : l'enunciato «Barack Obama è un filosofo» è falso. Inoltre avremo, per esempio, che  $\alpha_D \models \exists xGxs$  ma non che  $\alpha_D \models \forall xGxs$ : qui il primo nome che non compare in  $Gxs$  è  $a$ , e mentre esiste *qualche*  $a$ -variante di  $\alpha_D$  tale che  $\alpha_D \models Gas$ , come quella che assegna ad  $a$  l'individuo Barack Obama (la coppia  $\langle$ Barack Obama, Saul Kripke $\rangle$  è tra quelle a cui  $\alpha_D(G)$  assegna il valore  $V$ ), ciò non vale per *tutte* le  $a$ -varianti di  $\alpha_D$ , come quella che assegna ad  $a$  lo stesso Saul Kripke (la coppia  $\langle$ Saul Kripke, Saul Kripke $\rangle$  è tra quelle a cui  $\alpha_D(G)$  assegna il valore  $F$ , poiché nessuno è più giovane di se stesso). Ciò corrisponde alla verità di «Qualcuno è più giovane di Saul Kripke» e alla falsità di «Ognuno è più giovane di Saul Kripke», rispettivamente.

Chiarito questo, se adesso passiamo all'intero linguaggio della logica dei predicati con operatori modali, il modo più semplice per implementare la teoria di Kripke consiste nell'affidarsi a strutture-modello costruite a partire da modelli del tipo appena descritto. Abbiamo visto tre modi diversi di costruire queste strutture, di generalità sempre maggiore. Per limitarci alla versione più comune, quella intermedia del 1963, questo significa che una struttura-modello può essere definita come una tripla  $(\alpha, W_D, R)$ , dove  $W_D$  è un insieme di modelli su uno stesso dominio  $D$ ,  $\alpha$  è un elemento di  $W_D$  e  $R$  è una relazione binaria su  $W_D$ , e dove si assume che tutti gli elementi di  $W_D$  concordino sul valore che assegnano a ciascun nome. Quest'ultima assunzione corrisponde all'idea secondo cui i nomi sono dei «designatori rigidi», ovvero designano lo stesso individuo in tutti i mondi possibili contemplati dalla struttura-modello (sebbene possano ovviamente designare individui differenti nei mondi di altre strutture-modello). Si tratta di un'assunzione importante, sulla quale Kripke è tornato ripetutamente nei suoi scritti di filosofia del linguaggio degli anni Settanta (cfr. Capitolo 2), ma che nella fattispecie riflette semplicemente l'intuizione in base alla quale, quando ci rivolgiamo ai mondi che rappresentano situazioni alternative al mondo attuale, continuiamo a chiamare le cose con il loro nome. Nel ca-



so dei predicati, invece, quest'assunzione sarebbe irragionevole, dal momento che è proprio cambiando il loro valore semantico che riusciamo a rappresentare formalmente l'intuizione in base alla quale certi oggetti potrebbero soddisfare predicati diversi da quelli che soddisfano nel mondo attuale, e quindi che il valore di verità di certi enunciati atomici potrebbe variare da un mondo all'altro. Chiarito questo punto, la definizione ricorsiva della nozione di verità di una formula ben-formata rispetto a una struttura-modello qualsiasi estenderà le condizioni in (46)–(50) nel modo ovvio:

- |       |   |              |  |
|-------|---|--------------|--|
| (46') | $(\alpha, W_D, R) \models Pt_1 \dots t_n$ | se e solo se | $\alpha(P)(\alpha(t_1), \dots, \alpha(t_n)) = V$                                     |
| (47') | $(\alpha, W_D, R) \models \sim A$         | ” ”          | non $(\alpha, W_D, R) \models A$   |
| (48') | $(\alpha, W_D, R) \models A \wedge B$     | ” ”          | $(\alpha, W_D, R) \models A$ e $(\alpha, W_D, R) \models B$                          |
|       | $\vdots$                                  |              | $\vdots$   |
| (49') | $(\alpha, W_D, R) \models \forall x A$    | ” ”          | $(\beta, W_D, R) \models A_x^t$ per ogni $t$ -variante $\beta$ di $\alpha$ su $D$    |
| (50') | $(\alpha, W_D, R) \models \exists x A$    | ” ”          | $(\beta, W_D, R) \models A_x^t$ per qualche $t$ -variante $\beta$ di $\alpha$ su $D$ |
| (51') | $(\alpha, W_D, R) \models \Box A$         | ” ”          | $(\beta, W_D, R) \models A$ per ogni $\beta$ in $W_D$ tale che $\alpha R \beta$      |
| (52') | $(\alpha, W_D, R) \models \Diamond A$     | ” ”          | $(\beta, W_D, R) \models A$ per qualche $\beta$ in $W_D$ tale che $\alpha R \beta$   |

Si noti che l'analogia leibniziana tra modalit  e quantificazione diventa a questo punto esplicita: i quantificatori veri e propri spaziano sugli oggetti di  $D$ , le modalit  sui mondi in  $W_D$ .

Questa estensione rappresenta, in effetti, un «mix» della semantica del 1959, le cui strutture-modello erano ancora prive della relazione  $R$ , e della semantica del 1963, che analizzava l'impiego delle strutture-modello relazionali soltanto con riferimento alla logica proposizionale. Si tratta, comunque, di un mix del tutto naturale, tant'  vero che molti libri di testo oggi prendono le mosse proprio da una semantica di questo tipo. Ed   un mix che non   privo di interesse sul piano filosofico, poich  proprio le conseguenze filosofiche di questo modo di procedere indussero Kripke a pubblicare nel 1963 un secondo articolo, *Semantical Considerations on Modal and Intuitionistic Logic*, nel quale le cose sono un po' pi  complesse. (A essere precisi, c'  anche da dire che il linguaggio a cui si riferiva Kripke non conteneva nomi ma soltanto variabili, secondo una pratica piuttosto diffusa in quegli anni, e che le condizioni relative ai quantificatori erano formulate in modo leggermente diverso; si tratta per  di idiosincrasie che nel presente contesto possiamo ignorare.)

La conseguenza principale della semantica appena delineata riguarda la distinzione tra modalità *de dicto* e modalità *de re* che Quine aveva additato come una delle fonti di preoccupazione nell'interazione tra modalità e quantificazione, come notato a proposito degli esempi (14) e (15). In un certo senso, le definizioni in (46')–(52') dimostrano che dal punto di vista formale l'analisi delle modalità *de re* non presenta problemi particolari. Abbiamo una procedura ricorsiva e tant'è: forse verremo inghiottiti nella «giungla metafisica dell'essenzialismo aristotelico» nel momento in cui cercheremo di stabilire quale, tra le infinite strutture-modello messe a disposizione dalla teoria, corrisponda alla giusta mappa della realtà e delle sue alternative possibili; ma come abbiamo già osservato, questo non è il compito della logica. Dal punto di vista logico non siamo tenuti a specificare una struttura-modello particolare più di quanto non siamo tenuti a specificare un certo modello della logica estensionale quale rappresentante del mondo attuale. D'altro canto, è importante sottolineare che a questo riguardo la teoria non è del tutto neutrale. Ciò risulta particolarmente chiaro se consideriamo i due principi seguenti:

$$\begin{array}{ll} \text{BF} & \forall x \Box A \rightarrow \Box \forall x A \\ \text{BF}_c & \Box \forall x A \rightarrow \forall x \Box A \end{array}$$

Questi principi sono noti rispettivamente come «formula di Barcan» e «conversa della formula di Barcan» in quanto presenti – in questa forma o in una forma equivalente – nelle logiche modali quantificate proposte da Ruth C. Barcan a cui abbiamo accennato sopra. Evidentemente si tratta di principi che instaurano un nesso logico molto esplicito tra necessità *de dicto* e necessità *de re*, al punto che la loro congiunzione si risolve nella completa equivalenza tra i due tipi di necessità in presenza del quantificatore universale. E abbiamo già avuto modo di osservare che il nesso tra questi due tipi di modalità sembra dipendere, non solo da considerazioni di ordine logico, ma anche da tesi metafisiche belle e buone. (In effetti, le considerazioni svolte a proposito degli esempi (14) e (15) si applicano pari pari a  $\text{BF}_c$ , posto che  $\forall x A$  rappresenti l'enunciato «Ogni madre ha dei figli».) Ebbene, risulta che la semantica in esame *non* è neutrale rispetto a questi due principi. Al contrario, tanto la formula di Barcan quanto la sua conversa risultano universalmente valide, cioè logicamente valide rispetto a *qualunque* classe di strutture-modello. E se le cose stanno così, allora le preoccupazioni di Quine sembrano motivate: la logica modale quantificata ci impegnerebbe a un punto di vista metafisico ben preciso. Un discorso analogo potrebbe farsi anche riguardo alle varianti esistenziali delle due formule in questione, ossia:

$$\begin{aligned} \text{BF}_\diamond & \quad \diamond \exists x A \rightarrow \exists x \diamond A \\ \text{BF}_{\diamond c} & \quad \exists x \diamond A \rightarrow \diamond \exists x A \end{aligned}$$

Questi due principi risultano equivalenti rispettivamente a BF e BF<sub>c</sub>, da cui segue che, in presenza del quantificatore esistenziale, ogni possibilità *de dicto* si riduce alla corrispondente possibilità *de re*, e viceversa. (Curiosamente, abbiamo invece che, sebbene le due formule seguenti risultino logicamente valide:

$$\begin{aligned} (53) \quad & \exists x \Box A \rightarrow \Box \exists x A \\ (54) \quad & \diamond \forall x A \rightarrow \forall x \diamond A \end{aligned}$$

le loro converse non lo sono, cioè possono risultare false rispetto a determinate strutture-modello:

$$\begin{aligned} (55) \quad & \Box \exists x A \rightarrow \exists x \Box A \\ (56) \quad & \forall x \diamond A \rightarrow \diamond \forall x A \end{aligned}$$

Ciò è giusto ed è confermato dall'intuizione. Se, per esempio, qualcuno è così fortunato da essere necessariamente ricco, cioè ricco in tutti i mondi, allora è necessario che ci sia qualcuno di ricco; e se è possibile che tutti siano ricchi, cioè se c'è un mondo in cui tutti sono ricchi, allora tutti possono essere ricchi. La validità di (53) e (54) non fa che registrare questi esempi in termini generali. Per contro, benché sia necessario che qualcuno vinca la lotteria, non è detto che ci sia qualcuno così fortunato da vincere necessariamente, cioè in tutti i mondi; e anche se tutti possono vincere la lotteria, non è detto che sia possibile che vincano tutti, cioè che ci sia un mondo in cui tutti vincono. Abbiamo quindi dei controesempi alla validità di (55) e (56).)

Come mai ci ritroviamo con una logica impelagata in questioni metafisiche? La risposta, in parole semplici, è che la teoria semantica che abbiamo delineato è *già* impregnata di metafisica. Ma il problema non riguarda il trattamento degli operatori modali: il problema riguarda il trattamento dei quantificatori, ossia le clausole (49') e (50'), che sono definite con esclusivo riferimento al dominio D. Più precisamente, richiedendo che per la valutazione di enunciati quantificati della forma  $\forall x A$  o  $\exists x A$  si considerino *tutte* le *t*-varianti su D di un certo modello *quale che sia* quel modello, si sta assumendo che tutti gli elementi di D rientrino nel campo d'azione dei quantificatori per ciò che concerne qualunque modello. E poiché rientrare nel novero d'azione dei quantificatori equivale a esistere, come diceva Quine, ciò equivale ad associare lo stesso insieme di esistenti, D, a ogni modello, cioè a ogni mondo possibile. Nessun mondo alternativo contiene individui che non esistano già nel mondo attuale, così come

nessun mondo alternativo può essere privo degli individui che esistono in quello attuale. Se le cose stanno così, non c'è da sorprendersi se  $BF$  e  $BF_c$  risultano sempre vere. Se ogni individuo attuale soddisfa necessariamente una certa condizione  $A$  (cioè la soddisfa in ogni mondo accessibile), e se gli individui che abitano gli altri mondi sono né più né meno che quelli attuali, allora non si scappa: necessariamente (cioè in ogni mondo accessibile) ogni individuo soddisferà la condizione  $A$ . D'altro canto, se necessariamente (cioè in ogni mondo possibile) ogni individuo soddisfa la condizione  $A$ , e se gli individui che abitano gli altri mondi sono né più né meno che quelli attuali, allora non si scappa: ogni individuo attuale soddisferà necessariamente la condizione  $A$  (cioè la soddisferà in ogni mondo accessibile). Idem per le varianti esistenziali  $BF_\diamond$  e  $BF_{\diamond c}$ . In breve: la semantica induce una nozione di validità logica che non è del tutto neutrale. Ma questo risultato non è che l'inevitabile conseguenza di una mancanza di neutralità da parte della semantica stessa.

È proprio per questo motivo che nel secondo articolo del 1963, dove l'analisi semantica della logica modale quantificata è studiata con maggior dettaglio, Kripke non si accontenta della formulazione apparentemente ovvia che abbiamo fornito sopra. Per Kripke, come per buona parte dei filosofi, è perfettamente concepibile che certe cose di questo mondo (forse tutte) possano non esistere, così come è perfettamente concepibile che possano esistere delle cose (per esempio: Pegaso) che non figurano tra i cittadini di questo mondo. Più in generale, che tutti i mondi siano popolati dai medesimi individui è una tesi metafisica ben precisa e tutt'altro che pacifica, sulla quale è giusto pensare che la logica non si debba esprimere. Quindi, anche senza volerla scartare a priori, le condizioni di verità degli enunciati quantificati devono essere modificate in modo da non *presupporre* la verità di tale tesi. E ciò richiede che le strutture-modello rispetto alle quali si definiscono condizioni di verità dichiarino in modo esplicito, per ogni mondo possibile, quale sia il novero degli individui in esso esistenti. In termini formali, ciò significa che in aggiunta ai tre elementi iniziali  $\alpha$ ,  $W_D$  e  $R$ , una struttura-modello adeguata alla logica dei predicati dovrà comprendere anche una funzione,  $Q$ , che a ogni elemento  $\beta$  di  $W_D$  assegni un certo dominio di quantificazione  $Q(\beta)$ . Naturalmente questo dominio dovrà essere un sottoinsieme di  $D$ , e se non vogliamo violare uno degli assunti fondamentali della logica classica bisogna anche richiedere che  $Q(\beta)$  non sia vuoto. Per il resto, tuttavia, possiamo lasciare alla funzione  $Q$  libertà assoluta. Chiamiamo dunque *struttura-modello quantificazionale* qualunque quadrupla  $(\alpha, W_D, R, Q)$  che soddisfi questi requisiti minimali, dove  $\alpha$ ,  $W_D$  e  $R$  sono come sopra. Si tratta a questo punto di ridefinire le condizioni di verità (46')–(52') con riferimento a

strutture di questo tipo. Per le clausole relative alle formule atomiche e ai connettivi la modifica è ininfluente. Le clausole relative ai quantificatori, invece, diventano:

$$(49'') \quad (\alpha, W_D, R, Q) \models \forall x A \quad \text{se e solo se} \quad (\beta, W_D, R, Q) \models A'_x \text{ per ogni } t\text{-variante } \beta \text{ di } \alpha \text{ tale che } \beta(t) \in Q(\alpha)$$

$$(50'') \quad (\alpha, W_D, R, Q) \models \exists x A \quad \text{'' ''} \quad (\beta, W_D, R, Q) \models A'_x \text{ per qualche } t\text{-variante } \beta \text{ di } \alpha \text{ tale che } \beta(t) \in Q(\alpha)$$

È questo, in sostanza, il nucleo della semantica proposta da Kripke nell'articolo *Semantical Considerations*. Ed è facile verificare che in questa semantica né la formula di Barcan BF né la sua conversa BF<sub>c</sub> (e quindi nemmeno BF<sub>◇</sub> e BF<sub>◇c</sub>) risultano universalmente valide. Piuttosto, la validità di questi principi rispetto a una certa classe C di strutture-modello quantificazionali dipenderà, come per gli altri principi modali, da eventuali restrizioni specifiche che si potranno imporre sulla relazione R e/o sulla funzione Q. In particolare, si può dimostrare che BF e BF<sub>◇</sub> sono valide rispetto a qualunque classe C di strutture nelle quali Q è «conservativa» rispetto a R, ossia tale che βRγ vale soltanto se il dominio Q(γ) è incluso in Q(β), mentre BF<sub>c</sub> e BF<sub>◇c</sub> sono valide rispetto a qualunque classe C di strutture nelle quali Q è «espansiva» rispetto a R, ossia tale che βRγ vale soltanto se Q(β) è incluso in Q(γ). Ma non è il caso di spingersi oltre in questi dettagli, così come non entreremo nel merito degli interessanti sviluppi discussi nella seconda parte dell'articolo di Kripke, che riguardano l'applicazione della teoria alla logica dell'operatore «è dimostrabile nell'aritmetica di Peano che» e alla logica intuizionistica (sulla quale Kripke è tornato con maggior dettaglio nell'ultimo articolo della serie, *Semantical Analysis of Intuitionistic Logic I*, pubblicato nel 1965). Concludiamo semplicemente notando alcuni aspetti della teoria che possono ancora dar luogo a qualche complicazione.<sup>2</sup>

Per prima cosa, l'ammissione di domini di quantificazione variabili comporta immediatamente la possibilità che l'individuo β(t) assegnato a un nome t da un modello β non appartenga a Q(β). In termini intuitivi, questo significa che un nome può essere interpretato come un individuo che in certi mondi non esiste, proprio come Pegaso, per esempio, non esiste nel nostro mondo. Ciò ha come conseguenza la violazione di uno degli assiomi fondamentali della logica classica, la cosiddetta legge di «Esemplificazione dell'universale»:

---

<sup>2</sup> Le osservazioni che seguono presuppongono una certa familiarità con la filosofia della logica e del linguaggio. Possono essere omesse senza pregiudicare la comprensione del resto del capitolo.

$$(57) \quad \forall x A_i^x \rightarrow A$$

Ne segue che le logiche catturate da questa semantica non possono essere assiomaticizzate semplicemente aggiungendo dei principi modali al calcolo dei predicati classico (mentre le logiche catturate dalla semantica a «domini fissi» sono perfettamente classiche). Si tratta, in effetti, di fare affidamento a una cosiddetta «logica libera», sebbene i dettagli possano differire a seconda delle caratteristiche precise del linguaggio che si considera.

Inoltre, la possibilità che un nome possa venire interpretato come un individuo non esistente in certi mondi solleva la questione, molto discussa in filosofia del linguaggio, delle condizioni di verità degli enunciati che contengono il nome in questione. Nella formulazione della teoria che abbiamo dato qui, la questione è in un certo senso ignorata, dal momento che la clausola relativa agli enunciati atomici consente comunque di determinare in ciascun caso se l'enunciato è vero rispetto alla struttura-modello data (e se non è vero, è falso): basta controllare se la  $n$ -pla di individui che  $\alpha$  assegna ai nomi che compaiono nell'enunciato è tra quelle che ricevono il valore  $V$  dalla funzione che  $\alpha$  assegna al predicato, a prescindere dalla loro effettiva esistenza o non esistenza in  $\alpha$ , cioè dalla loro appartenenza o meno al dominio  $Q(\alpha)$ . In termini intuitivi, questo significa che per valutare un enunciato come

$$(58) \quad \text{Pegaso è bianco.}$$

basta vedere se Pegaso è tra le cose – attuali o meramente possibili – che nel nostro mondo classifichiamo come bianche. Si potrebbe tuttavia ritenere che questo trattamento sia eccessivamente semplificato, per non dire inadeguato. Per esempio, secondo una tradizione che risale almeno a Russell, gli enunciati atomici in cui compaiono termini che non si riferiscono a oggetti esistenti sono da considerarsi falsi; secondo una tradizione che risale invece a Frege, e che negli anni Cinquanta era stata difesa soprattutto da Peter F. Strawson, è preferibile considerare tali enunciati privi di un valore di verità definito. Kripke era consapevole di entrambe le opzioni e non esitava a osservare che in una «esposizione completa della semantica» bisognerebbe prenderle tutte in esame, sottolineando tuttavia che ai fini della logica modale si tratta in ultima analisi di «convenzioni alternative» egualmente sostenibili. Nella letteratura successiva questo compito è stato effettivamente affrontato con una certa sistematicità, anche nel contesto di semantiche dedicate a interpretazioni non aletiche degli operatori modali. Ma è opportuno sottolineare che, mentre la prima opzione non presenta problemi tecnici particolari (per esempio, è sufficiente richiedere che

ogni predicato venga interpretato in ogni modello  $\beta$  come una funzione che assegna il valore V soltanto a  $n$ -ple provenienti da  $Q(\beta)$ , la seconda richiede che si affrontino tutte le difficoltà che di norma si accompagnano all'abbandono del «principio di bivalenza», cioè del principio in base al quale ogni enunciato è o vero o falso. Per esempio, se (58) è indefinito, è ragionevole pensare che lo sia anche la sua negazione. Ma che dire della disgiunzione in (59)? Che dire dell'apparente contraddizione in (60)?

(59) O Pegaso è bianco o Pegaso non è bianco.

(60) Pegaso è bianco e Pegaso non è bianco.

E poi bisognerà chiarire come ci si deve comportare con gli enunciati modali. Continueremo a dire che  $\Box A$  è vera se e solo se  $A$  è vera in ogni mondo accessibile, o sarà sufficiente richiedere che  $A$  non sia falsa in quei mondi?

Infine, è importante notare che l'assunzione in base alla quale i nomi sono dei designatori rigidi ha delle conseguenze importanti per ciò che concerne l'interazione tra modalità e identità. Supponiamo infatti di arricchire la componente logica del linguaggio aggiungendo un predicato binario speciale « $\equiv$ » e fissando una volta per tutte il suo significato in conformità alle leggi classiche dell'identità. In altre parole, supponiamo di estendere la semantica in modo tale che « $a = b$ » risulti vero rispetto a una struttura-modello  $(\alpha, W_D, R, Q)$  se e solo se  $\alpha(a)$  e  $\alpha(b)$  sono il medesimo individuo. Allora è facile rendersi conto che il principio seguente, noto come «Necessità dell'identità», risulterà universalmente valido:

(61)  $a = b \rightarrow \Box(a = b)$

Questo fatto veniva notato *en passant* nell'articolo del 1959 con riferimento alla semantica a domini fissi priva della relazione di accessibilità, ma evidentemente continua a valere anche nella versione finale della semantica che stiamo considerando, rispetto alla quale Kripke osservava che una teoria «più generale» dell'identità potrebbe essere ottenuta «complicando la nozione di struttura-modello quantificazionale». Presumibilmente, questa osservazione era dovuta al fatto che in quegli anni (61) era ritenuto un principio controverso (sebbene fosse un teorema delle logiche modali di Barcan). Per esempio, molti filosofi erano inclini a considerare falso un enunciato come:

(62) Se Espero è identico a Fosforo, allora è necessario che Espero sia identico a Fosforo.

L'idea era che l'identità della stella del mattino e della stella della sera fosse un

fatto contingente, e che quindi l'antecedente di (62) potesse essere vero senza che lo fosse il conseguente. Ebbene, proprio questo costituirà uno dei temi su cui Kripke tornerà nei suoi lavori filosofici più maturi, specialmente nell'articolo *Identity and Necessity* e nel saggio *Naming and Necessity*. E proprio gli argomenti addotti da Kripke a sostegno di (61), basati su un'accurata distinzione tra ciò che è necessario e ciò che è vero a priori, e più in generale tra metafisica ed epistemologia, possono essere considerati tra i contributi più significativi e influenti della produzione filosofica complessiva di Kripke. Ma su questi contributi si torna ampiamente negli altri capitoli del volume, e per il momento possiamo fermarci qui.

## **2. La teoria della verità**

Passiamo adesso all'altro fondamentale contributo di Kripke in ambito logico: la teoria della verità elaborata nell'articolo *Outline of a Theory of Truth*, del 1975, e in due brevi abstract del 1976 (*A Theory of Truth, I e II*). Anche in questo caso si tratta di lavori piuttosto tecnici, ma il tema è evidentemente di grande portata sia per ciò che concerne la logica in senso stretto sia, più in generale, per l'importanza che la nozione di verità riveste sul piano filosofico. Se infatti è compito di qualunque ambito d'indagine approdare a una descrizione veritiera del mondo o pervenire a una ricostruzione veritiera dei fatti, il compito preliminare e per certi aspetti fondamentale di chiarire il concetto stesso di verità – e più precisamente di stabilire le condizioni necessarie e sufficienti in base alle quali *determinare* se una certa affermazione o una certa teoria sia vera – presenta difficoltà filosofiche considerevoli. Non è un'esagerazione affermare che la teoria di Kripke ha, se non risolto pienamente tali difficoltà, rivoluzionato il modo di affrontarle.

### **2.1. Verità e paradosso**

Cominciamo col delineare il contesto storico nel quale Kripke ha elaborato la sua teoria. Il concetto di verità è, ovviamente, al centro della riflessione filosofica sin dall'antichità, ma si può ben dire che nel corso del ventesimo secolo lo sviluppo delle semantiche formali abbia contribuito in modo determinante alla sua precisazione. In particolare, la semantica «standard» della logica dei predicati, che abbiamo brevemente riassunto all'inizio della sezione 1.5 e che si traduce nella definizione ricorsiva delle condizioni di verità schematizzate in (46)–(50), costituisce il punto di arrivo di un lungo e complesso lavoro di analisi che ha determinato un salto di qualità senza precedenti rispetto alla concezione tra-



dizionale della verità come semplice *adaequatio intellectus et rei*, o alla concezione aristotelica secondo cui dire il vero equivale a «dire di ciò che è che è o di ciò che non è che non è». L'articolo di Alfred Tarski *Pojęcie prawdy w językach nauk dedukcyjnych* [Sul concetto di verità nelle scienze deduttive formalizzate], pubblicato in polacco nel 1933 e in tedesco nel 1935, fu in questo senso altrettanto significativo quanto lo sviluppo della stessa logica dei predicati a opera di Frege: tutta la semantica formale dei nostri giorni, inclusa la semantica della logica modale messa a punto da Kripke, può essere considerata un'emanazione della teoria tarskiana.

Purtroppo, accanto a questo salto di qualità, la riflessione contemporanea sul concetto di verità ha anche messo in evidenza una serie di problemi molto seri che affliggono tale concetto. Uno di questi è il problema fondamentale di stabilire a *che cosa* esso si applichi. Agli enunciati di un linguaggio? Alle asserzioni effettuate per mezzo di enunciati appartenenti a un certo linguaggio? Alle proposizioni che vengono espresse effettuando certe asserzioni per mezzo di enunciati appartenenti a un certo linguaggio? Anche supponendo risolto questo problema, per esempio decidendo (come continueremo a fare qui) che il concetto di verità si applica direttamente agli enunciati del linguaggio, ogni tentativo di pervenire a una formulazione sistematica dei *criteri* per la corretta applicazione del concetto si scontra a sua volta con difficoltà molto concrete, a partire da quelle esemplificate dai cosiddetti «paradossi semantici». Il cretese di cui si parla nell'*Epistola a Tito* affermava:

(63) I cretesi non dicono mai la verità.

Nell'ipotesi in cui tutte le altre asserzioni uscite dalla bocca di un cretese fossero effettivamente false, appare impossibile determinare il valore di verità di (63) senza cadere in contraddizione: l'enunciato non può essere vero, perché altrimenti il cretese che lo ha pronunciato avrebbe detto la verità, e tanto basterebbe a decretare la *falsità* di (63); ma nemmeno può essere falso, perché altrimenti quel cretese avrebbe detto una falsità, e tanto basterebbe (data l'ipotesi) a decretare la *verità* di (63). Naturalmente si potrebbe semplicemente concludere che questo ragionamento costituisce una *reductio ad absurdum* dell'ipotesi da cui siamo partiti, ossia che tutte le altre asserzioni uscite dalla bocca di un cretese fossero false. Anche questa conclusione, però, sarebbe inaccettabile: in tal caso avremmo infatti dimostrato, sulla base di un ragionamento puramente logico, che a quei tempi esisteva almeno un cretese che almeno una volta aveva detto la verità. Ma questo è un fatto *empirico*, e la logica non può consentire la dimostrazione di verità empiriche.

Del resto non è difficile produrre varianti di (63) che non ammettono scappatoie di sorta. L'esempio più classico è:

(64) Quest'enunciato è falso.

che un attimo di riflessione rivela essere vero se e solo se falso. È proprio a enunciati di questo tipo che di solito ci si riferisce quando si parla del «paradosso del mentitore». Un altro esempio è costituito dalla seguente coppia di enunciati, che insieme formano un «circolo» paradossale che li rende assimilabili a (64):

(65) Il prossimo enunciato è vero.

(66) L'enunciato precedente è falso.

Si potrebbe pensare che queste considerazioni non facciano che confermare la limitata attendibilità del principio di bivalenza (come già osservato a proposito di quegli enunciati che, come (58), contengono nomi privi di riferimento). Se (63) non può essere né vero né falso, allora vuol dire che *non è* né vero né falso. Idem per (64) e per la coppia (65)–(66). Anche così però il problema è soltanto spostato. Basta infatti sostituire alla parola «falso» la parola «non vero» perché sia (64) sia (65)–(66) risultino paradossali anche in assenza del principio di bivalenza.

Insomma, per quanto fondamentale sia, e per quanto possa a prima vista sembrare del tutto innocuo, il concetto di verità non si lascia addomesticare facilmente e non c'è da sorprendersi se alcuni autori, a partire dallo stesso Tarski, siano giunti a concludere che *salvo porre restrizioni molto precise sul suo uso* si tratta di un concetto incoerente. Anzi, la conclusione di Tarski era più drastica: non vi è modo di porre restrizioni adeguate sul concetto di verità in uso nelle lingue naturali, e soltanto certi linguaggi formalizzati – come quello della logica dei predicati – ammetterebbero una teoria della verità coerente e rigorosa.

Il motivo principale di questa conclusione risiedeva, per Tarski, nel fatto che una lingua naturale come l'italiano *contiene* il suo concetto di verità, ossia contiene delle espressioni («è vero», «è falso») che consentono di applicare il concetto agli enunciati della lingua stessa. È proprio a questa caratteristica che possiamo attribuire la responsabilità degli esempi paradossali in (63)–(66). Per contro, un linguaggio formale come quello della logica dei predicati è esente da questa caratteristica. Un tipico linguaggio siffatto non possiede le risorse per parlare delle proprie espressioni ed è per questo motivo che la teoria semantica che specifica le condizioni di verità dei suoi enunciati, come quella che si rias-

sume nelle condizioni (46)–(50), è formulata in un *altro* linguaggio: l’italiano, o meglio una parte (speriamo coerente) dell’italiano arricchita di un po’ di notazione matematica. Anche le altre definizioni semantiche esaminate nella sezione precedente funzionano così: sia per la logica modale proposizionale sia per quella quantificazionale, abbiamo riassunto le teorie di Kripke formulando in italiano le condizioni di verità per gli enunciati o formule ben-formate di linguaggi formali opportunamente prefissati. In tutti i casi abbiamo cioè fornito una teoria semantica per un determinato *linguaggio oggetto* nell’ambito di un *metalinguaggio* ben più ricco sul piano espressivo (sebbene le formule del linguaggio oggetto possano essere usate per «rappresentare» certi enunciati del metalinguaggio, proprio come gli enunciati della lingua tedesca possono prestarsi alla traduzione di certi enunciati della lingua polacca). Se a questo punto volessimo precisare anche la semantica del metalinguaggio, Tarski ci inviterebbe a farlo collocandoci in un metametalinguaggio ancora più ricco sul piano espressivo, rispetto al quale il metalinguaggio stesso fungerebbe da linguaggio oggetto, e il risultato sarebbe la definizione di una nozione di verità diversa da quella che abbiamo caratterizzato a quel livello. Se volessimo precisare anche la semantica del metametalinguaggio, dovremmo ascendere di un ulteriore livello e definire una nozione ancora diversa. Questa distinzione tra linguaggi di livelli diversi era fondamentale per Tarski, ed è proprio nel rispetto della «gerarchia metalinguistica» e della conseguente stratificazione del concetto di verità che egli individuava le restrizioni necessarie al confinamento di tale concetto entro i limiti della coerenza logica. Nel caso dei linguaggi formali – o di certi linguaggi formali – ciò si rivela possibile e gli esempi che abbiamo citato lo confermano. Che anche le lingue naturali possano venire «irreggimentate» in questo modo era, per Tarski, improbabile e implausibile. E siccome l’incoerenza è una brutta malattia, tanto peggio per le lingue naturali.

Ora, intorno agli anni Settanta questo modo di vedere le cose aveva cominciato a traballare. Da un lato, un crescente numero di filosofi aveva maturato la convinzione che la strumentazione concettuale e formale messa a disposizione dalla filosofia del linguaggio e dal moltiplicarsi di teorie logiche sempre più sofisticate, fra cui una varietà di logiche non classiche, consentisse una «soluzione» dei paradossi e, di conseguenza, un atteggiamento di rinnovato rispetto nei confronti di quelle lingue naturali che al paradosso sembravano condannate. In questo filone si inserisce, per esempio, il lavoro di autori come Richard L. Martin, Charles Parsons e Bas van Fraassen, pur nella diversità talvolta radicale delle rispettive proposte. Dall’altro lato, non mancava chi, nonostante il pessimismo di Tarski, aveva cominciato a nutrire fiducia nei confronti di un progetto

di analisi semantica delle lingue naturali che passasse attraverso una loro irregimentazione gerarchica, proprio come per i linguaggi formali. Anzi, si era andata affermando l'idea che le lingue naturali *fossero* dei linguaggi formali suscettibili di una semantica di tipo tarskiano, un'idea condivisa da autori anche molto diversi fra loro come Donald Davidson e Richard Montague (che proprio nel 1970 pubblicava un importante articolo dal titolo sintomatico *English as a Formal Language*).

È appunto sullo sfondo di questi due opposti programmi di ricerca che prende forma la proposta di Kripke. Per Kripke entrambe le strategie erano sbagliate. Ed erano sbagliate perché muovevano da una diagnosi del problema – quella tarskiana – che si fondava su un'idea ingannevole: l'idea che si trattasse di un problema linguistico, cioè riconducibile alle caratteristiche sintattiche e semantiche delle lingue naturali. Vediamo dunque di chiarire innanzitutto questo punto. Dopo di che passeremo a delineare la parte costruttiva della proposta kripkeana.

## 2.2. La diagnosi di Kripke

Parte dell'inganno, per Kripke, derivava proprio dall'enfasi posta sui paradossi, che a ben vedere sono solo la punta di un iceberg. È vero che il nostro concetto di verità sembra entrare in crisi dinnanzi all'enunciato del mentitore, (64). Ma ci troviamo in imbarazzo anche dinnanzi a enunciati che, come

(67) Quest'enunciato è vero.

non sono propriamente paradossali. Nel caso di (64) abbiamo a che fare con un enunciato che è vero se è falso ed è falso se è vero, sicché qualunque assegnazione di un valore di verità definito sarebbe *incoerente*; nel caso di (67) l'imbarazzo deriva dal fatto che l'enunciato è vero se è vero ed è falso se è falso, sicché assegnargli un valore di verità sarebbe, non incoerente, bensì *arbitrario*. Né possiamo identificare il problema generale nel fatto che in entrambi i casi abbiamo a che fare con enunciati autoreferenziali. La loro disgiunzione è altrettanto autoreferenziale:

(68) Quest'enunciato è vero o falso.

Tuttavia in questo caso le nostre intuizioni semantiche ci dicono chiaramente che si tratta di un enunciato che non può essere falso ma può benissimo essere vero: una valutazione che non sarebbe né incoerente né arbitraria. E quando passiamo a

- (69) Quest'enunciato consta di sei parole.  
(70) Quest'enunciato consta di sette parole.

abbiamo addirittura a che fare con enunciati autoreferenziali il cui valore di verità è perfettamente determinato. Insomma, ci sono delle somiglianze e delle differenze che vanno chiarite, classificate, studiate: non si può pensare di risolvere il problema se prima non si è capito bene *quale* sia il problema. E tanto basta a screditare ogni strategia che si basi sull'idea di una «soluzione» del paradosso del mentitore.

In secondo luogo – e qui il discorso interessa entrambe le strategie risolutive che abbiamo citato – la nozione stessa di autoreferenzialità presenta caratteristiche la cui problematicità non è riconducibile a criteri univoci. Tanto per cominciare, abbiamo appena visto che enunciati come (69) e (70), per quanto autoreferenziali, sono del tutto innocenti, e ciò significa che vietare in modo assoluto ogni trasgressione delle restrizioni imposte dalla gerarchia dei metalinguaggi sarebbe una misura eccessiva. Ma ancora più importante, per Kripke, è rendersi conto del fatto che in certi casi è impossibile stabilire a priori, cioè sulla base delle caratteristiche intrinseche di un enunciato, *se* si tratta di un enunciato autoreferenziale. Per rendersene conto è sufficiente tornare al caso illustrato dalla coppia (65)–(66). Il primo di questi enunciati afferma che il secondo è vero, e il secondo afferma che il primo è falso. Il risultato è che entrambi risultano indirettamente autoreferenziali. Ma questo risultato è del tutto accidentale: se uno dei due enunciati fosse stato diverso – per esempio, se avesse riguardato il numero dei pianeti o il colore della neve – entrambi sarebbero stati del tutto esenti da autoriferimento e l'accertamento della loro verità o falsità non avrebbe presentato alcuna difficoltà. Questo significa che a differenza del mentitore classico esemplificato in (64), come anche degli enunciati in (67)–(70), l'autoreferenzialità di (65) e di (66) è un fatto empirico, non linguistico. Possiamo illustrare questo punto anche considerando enunciati singoli, a partire dal nostro esempio iniziale. (63) afferma che i cretesi non dicono mai la verità: che quest'enunciato sia o meno autoreferenziale dipende dal fatto che sia stato pronunciato da un cretese, proprio come il suo essere o meno paradossale dipende da ciò che hanno davvero detto gli altri cretesi. Un altro esempio potrebbe essere

- (71) Gli enunciati sulla lavagna non sono veri.

Per stabilire se quest'enunciato è autoreferenziale non basta esaminare la sua struttura: bisogna guardare al mondo, e più precisamente alla lavagna. E l'even-

tuale paradossalità di (71) dipenderà da *quali* altri enunciati compaiono sulla lavagna.

Più in generale, è facile rendersi conto che non sempre è possibile stabilire a priori, non solo se un determinato enunciato è autoreferenziale, ma anche quale sia il livello della gerarchia metalinguistica che lo riguarda (e questo costituisce un problema particolare per il secondo tipo di strategia a cui Kripke si rivolgeva). In certi casi, come in

(72) Il numero dei pianeti è maggiore di 7.

il livello è chiaro: ci troviamo alla base, a livello 0, dato che l'enunciato fa riferimento esclusivamente alla realtà extralinguistica. Ma quando abbiamo a che fare con enunciati nei quali si fa uso del concetto di verità (o di altri concetti semantici), il loro livello può dipendere da fattori empirici. Per esempio, se sulla lavagna compaiono soltanto affermazioni di livello 0, come (72), allora (71) sarà di livello immediatamente superiore, cioè di livello 1. Se però compare anche

(73) L'enunciato (72) è vero.

allora (71) sale a livello 2. Con i linguaggi formali ci si può affidare a una stratificazione rigidamente gerarchica della nozione di verità perché possiamo aver cura di assegnare ogni enunciato a un livello ben preciso sulla base della sua struttura sintattica (per esempio il numero di occorrenze iterate del predicato di verità); ma con una lingua come l'italiano non c'è modo di classificare gli enunciati se non sulla base di considerazioni empiriche, e tanto basta a impedire che si possa seguire la stessa strategia.

Si potrebbe continuare, ma questi semplici esempi dovrebbero bastare per cogliere il senso importante della morale complessiva che Kripke ne trae: quali che siano le motivazioni che la ispirano, la speranza che si possano «setacciare» gli usi problematici del concetto di verità dai suoi usi innocui sulla base di considerazioni puramente linguistiche è destinata a fallire. In linea di principio, qualunque enunciato della nostra lingua può rivelarsi problematico, fatta eccezione per quegli enunciati che, come (72), sono interamente *fondati* sulla realtà extralinguistica. In linea di principio, le condizioni di verità di qualunque enunciato dipendono dalle relazioni empiriche, o comunque *contingenti*, che lo legano non solo al mondo, ma anche agli altri enunciati della lingua. Una teoria adeguata e sufficientemente generale deve pertanto rendere conto di questi due dati. E una teoria sufficientemente rigorosa deve poterlo fare sulla base di una caratterizzazione sistematica dei due concetti di cui ci siamo serviti per riassumerli: quello di fondatezza e quello di contingenza delle condizioni di

merli: quello di fondatezza e quello di contingenza delle condizioni di verità di un enunciato.

### 2.3. La teoria

Per delineare i tratti principali della sua teoria, Kripke considera un linguaggio arbitrario  $L$  che non è, naturalmente, un vero e proprio linguaggio naturale come l'italiano: rivolgersi direttamente a quest'ultimo sarebbe un'impresa mastodontica che, oltre a scontrarsi con una infinità di altri problemi che nulla hanno a che vedere con la nozione di verità, non consentirebbe di mettere a fuoco i tratti principali della proposta. Più semplicemente, Kripke ci chiede di immaginare che  $L$  sia un linguaggio che condivide con le lingue naturali la caratteristica di essere sufficientemente ricco da poter esprimere la propria sintassi e la propria semantica, che è la caratteristica responsabile di tutti i problemi che abbiamo visto.

A tal scopo possiamo supporre che  $L$  sia essenzialmente ciò che si ottiene aggiungendo un predicato unario  $T$ , che leggeremo come «è vero», a un comune linguaggio della logica dei predicati come quello che abbiamo definito all'inizio della sezione 1.5, che indicheremo con  $L_0$ . Per ovvi motivi di semplicità, assumeremo che  $L_0$  sia privo di qualsiasi operatore modale. Tuttavia assumeremo che  $L_0$  disponga di tutte le risorse per soddisfare il requisito di espressività sintattica a cui accennavamo, e in particolare che sia possibile associare a ciascuna formula ben-formata di  $L_0$  un nome appartenente al vocabolario dello stesso  $L_0$ . Questo requisito può essere soddisfatto in vari modi. In italiano, il modo forse più comune per assegnare un nome a qualsiasi enunciato consiste nello scrivere quell'enunciato tra virgolette di citazione. Un altro modo, seguito ampiamente in queste pagine, consiste nell'assegnare a ciascun enunciato un certo numerale che lo identifichi in modo univoco, un po' come le cifre stampate sul passaporto identificano in modo univoco ogni cittadino. Così, per esempio, possiamo parlare dell'enunciato «Il numero dei pianeti è maggiore di 7» oppure dell'enunciato (72): in entrambi i casi non stiamo *usando* l'enunciato in questione; lo stiamo *nominando*, e in questo modo possiamo attribuirgli delle proprietà servendoci di certi predicati (per esempio il predicato «è vero») proprio come possiamo attribuire una proprietà a una persona chiamandola per nome e applicando a quel nome un certo predicato (per esempio «è un filosofo»). Tornando a  $L_0$ , Kripke ipotizza che i nomi delle formule vengano fissati nel secondo modo, che è un po' più semplice in quanto non richiede l'aggiunta di funtori che, come le virgolette di citazione, non sono governati dalle leggi della

logica dei predicati. In pratica questo significa ipotizzare che il vocabolario di  $L_0$  includa interamente il vocabolario del comune linguaggio dell'aritmetica e che sia stata definita una procedura per assegnare a ciascuna formula ben-formata di  $L_0$  un numero ben preciso, per esempio una procedura di «aritmetizzazione» della sintassi simile a quella introdotta da Kurt Gödel per la dimostrazione del suo celebre Teorema di Incompletezza del 1931.

Sia dunque  $L_0$  un linguaggio del tipo appena descritto, e sia  $L$  l'estensione di  $L_0$  ottenuta aggiungendo al vocabolario di quest'ultimo il predicato  $T$ . La procedura di aritmetizzazione può ovviamente essere estesa anche a  $L$ . Se  $A$  è una formula ben-formata di  $L$ , scriveremo « $\ulcorner A \urcorner$ » per indicare il numerale del suo «gödeliano», cioè del numero assegnato ad  $A$  dalla procedura in questione. In questo modo potremo intuitivamente leggere  $T\ulcorner A \urcorner$  come una formula che dice, in  $L$ , che la formula  $A$  di  $L$  è vera (e  $T\ulcorner \sim A \urcorner$  come la formula che dice che è vera la negazione di  $A$ , cioè che  $A$  è falsa). Su queste basi, il primo passo verso la formulazione della teoria consisterà evidentemente nella definizione dei modelli di  $L$ . Dopo di che il passo cruciale risiederà nella messa a punto di criteri precisi per individuare i modelli di  $L$  che forniscono un'interpretazione *adeguata* di  $T$ , cioè un'interpretazione che giustifica la lettura intuitiva che desideriamo attribuire a  $T$  («è vero»).

Per quanto riguarda il primo punto, possiamo affidarci tranquillamente alla definizione standard fornita all'inizio della sezione 1.5, con un'unica variante: che l'interpretazione assegnata da un modello  $\alpha_D$  al predicato  $T$  può essere una funzione *parziale*, ossia definita soltanto per alcuni elementi del dominio  $D$ . In altre parole,  $\alpha_D(T)$  assegnerà a certi elementi di  $D$  uno dei due valori V o F, come per qualunque predicato unario, ma è possibile che vi siano degli elementi ai quali  $\alpha_D(T)$  non assegna alcun valore. Nel caso estremo, è anche possibile che la funzione  $\alpha_D(T)$  sia indefinita per *tutti* gli elementi di  $D$ . Per il resto possiamo però supporre che i modelli di  $L$  si comportino esattamente come nella sezione 1.5, che è come dire che le loro restrizioni a  $L_0$  saranno modelli del tutto standard.

Intuitivamente, il motivo della variante appena descritta dovrebbe essere evidente. Nella misura in cui  $T$  deve rappresentare il predicato «è vero», richiedere che  $\alpha_D(T)$  sia una funzione totale equivarrebbe ad assumere che  $T$  obbedisca a una forma del principio di bivalenza. Ma abbiamo già visto che in presenza di certi enunciati, come quello del mentitore, la bivalenza è sufficiente a generare un paradosso. Quindi, se vogliamo sperare che la teoria ci restituisca un predicato di verità che si comporta in modo coerente, dobbiamo lasciare aperta la possibilità che l'applicazione di  $T$  al nome di una formula  $A$  si risolva



in un enunciato né vero né falso, e ciò significa che l'applicazione di  $\alpha_D(T)$  ad  $\alpha_D(\ulcorner A \urcorner)$  deve poter essere indefinita. La cosa importante da notare, a questo punto, è che rilassando la nozione di modello in questo modo non potremo più affidarci alla caratterizzazione standard delle clausole per la valutazione delle formule ben-formate che avevamo riassunto in (46)–(50). Quelle clausole erano formulate sotto l'ipotesi della bivalenza, sicché la non verità di una formula rispetto a un dato modello coincideva sempre con la sua falsità. Se però certi predicati – nella fattispecie: il predicato  $T$  – corrispondono a funzioni parziali, la clausola relativa alle formule atomiche deve essere riformulata più esplicitamente in modo da distinguere tra i casi in cui una formula non è vera rispetto al modello dato e i casi in cui è falsa. E una volta ammessa l'eventualità che una formula atomica non sia né vera né falsa, le clausole relative agli enunciati composti attraverso gli operatori logici dovranno a loro volta essere riformulate in modo da fornire istruzioni precise su come valutare quelle formule che contengono delle parti, cioè delle sotto-formule, che non sono né vere né false. In breve: la possibilità che  $T$  venga interpretato come un predicato non bivalente si traduce immediatamente nella necessità di definire una semantica non bivalente per il linguaggio  $L$ .

Abbiamo già avuto modo di osservare, con riferimento alla semantica della logica modale quantificata, che questo compito presenta delle difficoltà. Tuttavia le proposte non mancano e Kripke stesso ritiene che nel presente contesto la scelta di una soluzione piuttosto che un'altra sia largamente arbitraria. L'unica condizione importante è che la semantica prescelta soddisfi un requisito intuitivo: tanto più è determinata l'interpretazione del predicato  $T$ , tanto più sono determinate le condizioni di verità per il linguaggio  $L$ . Più precisamente, i modelli della semantica devono soddisfare il seguente «requisito di stabilità»:

- (74) Se  $\alpha_D$  differisce da  $\beta_D$  per il solo fatto che  $\alpha_D(T)$  è definita per alcuni elementi di  $D$  per i quali  $\beta_D(T)$  non è definita, allora ciò che è vero in  $\beta_D$  resta vero in  $\alpha_D$  e ciò che è falso in  $\beta_D$  resta falso in  $\alpha_D$ .

Per completezza di esposizione, riportiamo qui di seguito le clausole che definiscono una delle semantiche non bivalenti considerate da Kripke, dovuta al logico Stephen C. Kleene. Queste clausole sono una estensione naturale delle clausole standard fornite in (46)–(50) e differiscono da quelle soltanto per il fatto che, per ciascuna formula ben-formata  $X$ , si forniscono sia le condizioni alle quali  $X$  è vera rispetto a un dato modello  $\alpha_D$  ( $\alpha_D \models X$ ), sia le condizioni alle quali  $X$  è falsa ( $\alpha_D \models\!\!\!\!\!\! \neq X$ ). S'intende che nei casi in cui entrambe le condizioni vengono meno, la formula stessa è da considerarsi né vera né falsa.

(75)	$\alpha_D \models Pt_1 \dots t_n$	se e solo se	$\alpha_D(P)(\alpha_D(t_1), \dots, \alpha_D(t_n)) = V$
	$\alpha_D \models \neg Pt_1 \dots t_n$	” ”	$\alpha_D(P)(\alpha_D(t_1), \dots, \alpha_D(t_n)) = F$
(76)	$\alpha_D \models \sim A$	” ”	$\alpha_D \models A$
	$\alpha_D \models \sim \sim A$	” ”	$\alpha_D \models A$
(77)	$\alpha_D \models A \wedge B$	” ”	$\alpha_D \models A$ e $\alpha_D \models B$
	$\alpha_D \models A \wedge \sim B$	” ”	$\alpha_D \models A$ o $\alpha_D \models \sim B$
	$\vdots$		$\vdots$
(78)	$\alpha_D \models \forall x A$	” ”	$\beta_D \models A'_x$ per ogni $t$ -variante $\beta_D$ di $\alpha_D$
	$\alpha_D \models \sim \forall x A$	” ”	$\beta_D \models A'_x$ per qualche $t$ -variante $\beta_D$ di $\alpha_D$
(79)	$\alpha_D \models \exists x A$	” ”	$\beta_D \models A'_x$ per qualche $t$ -variante $\beta_D$ di $\alpha_D$
	$\alpha_D \models \sim \exists x A$	” ”	$\beta_D \models A'_x$ per ogni $t$ -variante $\beta_D$ di $\alpha_D$

Veniamo adesso al secondo passo nella formulazione della teoria, quello cruciale e profondamente innovativo: l'identificazione dei modelli di  $L$  che forniscono un'interpretazione *adeguata* di  $T$ . Quali sono, fra gli infiniti modelli attraverso i quali possiamo interpretare  $L$ , quelli nei quali  $T$  *rappresenta* propriamente il predicato «è vero»? Evidentemente vorremmo rispondere che i modelli in questione sono ne più né meno che quei modelli nei quali  $T$  riflette le condizioni di verità per gli enunciati di  $L$ , ossia quei modelli  $\alpha_D$  che soddisfano la seguente «condizione di adeguatezza» per ogni formula ben-formata  $A$ :

$$\begin{aligned} \text{CA} \quad & \alpha_D \models A \text{ se e solo se } \alpha_D \models T^r A \\ & \alpha_D \models \sim A \text{ se e solo se } \alpha_D \models \sim T^r A \end{aligned}$$

Il problema è che non c'è nessuna garanzia che esistano modelli siffatti. Se non ce ne fossero, avremmo una conferma della tesi tarskiana secondo cui un linguaggio «semanticamente chiuso» come  $L$  è intrinsecamente incoerente, anche in assenza del principio di bivalenza. Se però riuscissimo a dimostrare l'esistenza di almeno un modello che soddisfa CA, allora avremmo in questo modo dimostrato che è possibile avere una vera e propria teoria della verità anche per linguaggi del genere. Ebbene, il grande contributo di Kripke risiede proprio nella dimostrazione di questo fatto. Più esattamente, risiede nell'aver fornito una dimostrazione di questo fatto alla luce della quale risulta possibile dare una spiegazione convincente del comportamento problematico esibito dalla nozione di verità nei linguaggi naturali di cui  $L$  è un'immagine formale. (Sul piano puramente matematico, l'esistenza di un modello conforme a CA per una semantica abbastanza simile a quella definita in (75)–(79) era appena stato dimostrato anche da Richard L. Martin e Peter W. Woodruff in un saggio intitolato *On Representing «True-in-L» in L*, già disponibile ai tempi della stesura dell'articolo di Kripke sebbene pubblicato soltanto l'anno successivo. Tuttavia l'esistenza di

tale modello non consentiva di dire molto se non, appunto, che il pessimismo tarskiano non era giustificato sul piano matematico.)

Prima di descrivere la struttura matematica della dimostrazione, che per certi aspetti consiste in una vera e propria «costruzione» di modelli per L conformi a CA, può essere utile anticiparne l'idea intuitiva con riferimento a una lingua naturale come l'italiano. Kripke stesso ci invita a considerare una situazione immaginaria che contiene tutti gli elementi essenziali. Supponiamo – dice Kripke – di dover insegnare il significato del predicato «è vero» a una persona che non ne è in possesso. Come procederemmo? Evidentemente diremmo al nostro interlocutore che un parlante competente dell'italiano è autorizzato ad applicare il predicato «è vero» ne più né meno che a quegli enunciati che il parlante stesso è disposto ad asserire. Questa risposta riflette, rispetto alla lingua italiana, esattamente la stessa intuizione che trova espressione nella condizione di adeguatezza CA per il linguaggio L. E possiamo subito intuire quali siano i suoi effetti. Innanzitutto, siccome il nostro interlocutore è disposto ad asserire, poniamo,

(80) La neve è bianca.

la nostra risposta lo metterà subito nelle condizioni di asserire:

(81) «La neve è bianca» è vero.

Possiamo anzi supporre che a questo punto egli sia in grado di applicare (o negare) correttamente il predicato «è vero» a tutti quegli enunciati di cui era in grado di comprendere il significato prima che avesse inizio la nostra conversazione, ossia quegli enunciati che secondo Tarski dovrebbero appartenere al livello 0 della «gerarchia» linguistica: quelli che riguardano la realtà *extralinguistica*. Questo è l'effetto immediato, per così dire, della nostra risposta. Che dire però degli enunciati appartenenti agli altri livelli? Se ha capito fino in fondo le nostre istruzioni, adesso che è disposto ad asserire (81) il nostro interlocutore dovrebbe essere disposto anche ad applicare il predicato «è vero» a quella stessa asserzione, e quindi ad asserire

(82) ««La neve è bianca» è vero» è vero.

Anzi, possiamo supporre che questa considerazione si applichi a tutte le asserzioni contenenti il predicato «è vero» che egli è disposto ad asserire sulla base della considerazione precedente. Il nostro interlocutore dovrebbe cioè sapere come comportarsi non solo davanti a tutti gli enunciati di livello 0, ma anche davanti a tutti gli enunciati appartenenti al livello 1 della gerarchia tarskiana.

Ma se le cose stanno così, allora possiamo ripetere il ragionamento anche con riferimento a (82) e a tutti gli enunciati dello stesso livello: il nostro interlocutore dovrebbe sapere come applicare «è vero» anche nei confronti degli enunciati di livello 2. Si capisce che a questo punto il procedimento può continuare indefinitamente in modo da consentire la valutazione di qualunque enunciato appartenente a qualunque livello  $n + 1$  della gerarchia, e questa è la seconda importante conseguenza della nostra risposta iniziale. In aggiunta, è ragionevole supporre che a questo punto il nostro interlocutore sarà anche disposto ad asserire enunciati che contengono il predicato «è vero» pur senza fare riferimento in modo esplicito a enunciati appartenenti a un livello particolare, come

(83) Qualche enunciato è vero.

che il nostro interlocutore sarà in grado di inferire da (81) o da (82) per generalizzazione esistenziale. Ora, non c'è motivo di pensare che il nostro interlocutore riesca di questo passo a prendere una decisione in merito a *tutti* gli enunciati della lingua italiana che contengono il predicato «è vero», ed è per questo che sarebbe irragionevole imporgli di accettare il principio di bivalenza. Non è escluso che nel giro di poco tempo egli si scontri con enunciati che, come il paradosso del mentitore, non sono *fondati* in questo modo sulla verità degli enunciati di livello 0 da cui è partito, e che quindi non sappia come comportarsi. Tuttavia possiamo pensare che nella maggior parte dei casi il suo uso del predicato «è vero» sia perfettamente conforme a quello di un parlante competente. Anzi – e questo è l'effetto finale della nostra risposta – c'è da pensare che *al limite* di questo processo il suo uso del predicato coincida esattamente con quello di un parlante competente.

La dimostrazione di Kripke è sostanzialmente una ricostruzione formale, rispetto al linguaggio  $L$ , di questo tipo di ragionamento, con una conferma matematica della considerazione finale. Si cominci col considerare un modello del linguaggio  $L$  che interpreta  $T$  come una funzione del tutto vuota, lo si estenda a un modello che interpreta  $T$  limitatamente a tutti (e soli) i gödeliani degli enunciati di  $L_0$ , e si estenda progressivamente questo modello saturando di volta in volta l'interpretazione di  $T$  sulla base dell'interpretazione precedente. Il risultato sarà un modello di  $L$  che interpreta  $T$  come una funzione capace di attribuire il valore  $V$  o il valore  $F$  a tutti gli enunciati che si fondano, per le loro condizioni di verità, sul modello iniziale della serie. E un esame attento delle proprietà matematiche della serie rivelerà che questo modello ha le caratteristiche che cercavamo: l'interpretazione di  $T$  che ne emerge è pienamente conforme alla condizione di adeguatezza CA.

Vediamo dunque con maggiore dettaglio come si possa costruire una serie del tipo appena descritto<sup>3</sup>. Sia  $D$  un dominio prefissato di oggetti che include tutti i numeri naturali, e sia  $\alpha_D$  un certo modello di  $L_0$  su  $D$  che soddisfa tutti i requisiti di un modello standard dell'aritmetica. In particolare,  $\alpha_D$  assegnerà a ogni numerale  $n$  di  $L_0$  il numero corrispondente in  $D$  e a ogni predicato aritmetico di  $L_0$  la funzione corrispondente su  $D$  (eventualmente estesa in modo arbitrario agli elementi di  $D$  che non sono dei numeri). Dato che  $D$  è fissato, conveniamo inoltre di omettere l'indice « $D$ » e di parlare semplicemente del modello  $\alpha$ . A questo punto definiamo, per ogni coppia ordinata  $\langle D_V, D_F \rangle$  di sottoinsiemi di  $D$  che non abbiano elementi in comune, un corrispondente modello di  $L$  su  $D$ , che indicheremo con  $\alpha[D_V, D_F]$ . Si tratta di quel modello che coincide esattamente con  $\alpha$  per ciò che concerne i simboli di  $L_0$ , e che interpreta il predicato aggiuntivo  $T$  in conformità alle seguenti condizioni generali, cioè valide per ogni elemento  $d$  di  $D$ :

$$(84) \quad \begin{aligned} \alpha[D_V, D_F](T)(d) &= V \text{ se e solo se } d \in D_V \\ \alpha[D_V, D_F](T)(d) &= F \text{ se e solo se } d \in D_F \end{aligned}$$

In altre parole,  $\alpha[D_V, D_F]$  è il modello di  $L$  che si ottiene da  $\alpha$  trattando  $D_V$  come l'*estensione* di  $T$ , cioè l'insieme di quegli elementi in  $D$  di cui  $T$  è vero, e  $D_F$  come la *controestensione* di  $T$ , cioè l'insieme di quegli elementi di cui  $T$  è falso. Definiamo adesso una funzione  $\phi$  che a ogni modello  $\alpha[D_V, D_F]$  associa un modello  $\alpha[D_V', D_F']$  come segue:

$$(85) \quad \begin{aligned} D_V' &= \text{l'insieme degli elementi di } D \text{ che sono gödeliani delle formule di } L \\ &\quad \text{vere rispetto ad } \alpha[D_V, D_F]; \\ D_F' &= \text{l'insieme degli elementi di } D \text{ che sono gödeliani delle formule di } L \\ &\quad \text{false rispetto ad } \alpha[D_V, D_F] \text{ o che non sono i gödeliani di alcuna formula di } L. \end{aligned}$$

In altre parole,  $\phi$  associa ad  $\alpha[D_V, D_F]$  quel modello che interpreta il predicato  $T$  per mezzo di una funzione che rispecchia fedelmente, nell'attribuzione (parziale) dei valori  $V$  e  $F$ , le condizioni (parziali) di verità e falsità determinate dallo stesso  $\alpha[D_V, D_F]$ . (La decisione di mettere in  $D_F'$  gli elementi di  $D$  che non sono tra i gödeliani delle formule di  $L$  è del tutto arbitraria, ma riflette l'idea che  $T$  si applica correttamente soltanto a nomi di formule ben-formate; anche in

---

<sup>3</sup> La costruzione si avvale di nozioni di teoria degli insiemi ed è qui inclusa per amore di completezza. Il lettore che non avesse familiarità con tali nozioni può procedere direttamente alla sezione successiva.

italiano non ha molto senso dire che Saul Kripke è vero, o che è vero il numero dei pianeti, e le cose si semplificano molto se anziché lasciare aperta la questione decidiamo una volta per tutte di trattare casi del genere alla stregua di falsità.) A questo punto non è difficile rendersi conto del significato di questa costruzione rispetto alla condizione di adeguatezza CA: affinché un modello  $\alpha[D_V, D_F]$  soddisfi tale condizione, sarà infatti necessario e sufficiente che si abbia  $\alpha[D_{V'}, D_{F'}] = \alpha[D_V, D_F]$ . In altre parole, sarà necessario e sufficiente che il modello  $\alpha[D_V, D_F]$  sia un *punto fisso*, nel senso matematico del termine, della funzione  $\phi$ . E che  $\phi$  debba avere dei punti fissi è proprio ciò che si riesce a dimostrare applicando il ragionamento intuitivo illustrato sopra, cioè considerando il *limite* di una serie di modelli nei quali l'estensione e la controestensione del predicato  $T$  crescono progressivamente.

Sul piano formale – e qui siamo costretti a presupporre una certa dimestichezza con la teoria degli ordinali transfiniti – la serie in questione può essere identificata con la serie di tutti i modelli  $\alpha_\xi$  di  $L$  che emergono dalla seguente definizione induttiva, dove  $\xi$  è un numero ordinale qualsiasi:

- (86)  $\alpha_\xi = \alpha[\emptyset, \emptyset]$  se  $\xi = 0$   
(87)  $\alpha_\xi = \alpha[D_{V'}, D_{F'}]$  se  $\xi = \zeta + 1$  e  $\alpha_\zeta = \alpha[D_V, D_F]$   
(88)  $\alpha_\xi = \alpha[\bigcup_{\zeta < \xi} D_{\zeta V'}, \bigcup_{\zeta < \xi} D_{\zeta F'}]$  se  $\xi$  è un limite e  $\alpha_\zeta = \alpha[D_{\zeta V}, D_{\zeta F}]$  per ogni  $\zeta < \xi$

Intuitivamente, il primo modello della serie,  $\alpha_0$ , corrisponde alle condizioni nelle quali si trovava il nostro interlocutore prima di rivolgersi a noi: il suo uso del predicato «è vero» era nullo, cioè determinava un'estensione e una controestensione di questo predicato pari in ciascun caso all'insieme vuoto  $\emptyset$ . Il modello  $\alpha_1$  corrisponde alle condizioni del nostro interlocutore per effetto immediato della nostra spiegazione: l'estensione del suo uso di «è vero» conteneva a quel punto tutti gli enunciati di livello 0 che era disposto ad asserire, e la controestensione tutti gli enunciati di livello 0 che era disposto a negare. Analogamente,  $\alpha_2$  corrisponde alle condizioni in cui il nostro interlocutore si è venuto a trovare nel momento in cui, per effetto di una nuova applicazione della nostra spiegazione, l'estensione e la controestensione di «è vero» si erano estese anche a tutti gli enunciati di livello 1. In generale,  $\alpha_{n+1}$  corrisponde alla situazione venutasi a creare dopo  $n$  applicazioni della nostra spiegazione. Dopo di che possiamo pensare che  $\alpha_\omega$  corrisponda alle condizioni del nostro interlocutore una volta resosi conto di come questa procedura possa *al limite* essere iterata un numero infinito di volte ( $\omega$  è il primo ordinale infinito): l'estensione di «è vero» includerà ogni enunciato dichiarato vero in almeno uno dei modelli precedenti, e la controestensione ogni enunciato dichiarato falso in almeno uno dei modelli

precedenti. Le definizioni in (86)–(88) non fanno che esplicitare tutto ciò riferendolo al predicato  $T$  di  $L$ , e non solo sino al raggiungimento di  $\alpha_\omega$  ma per tutta la serie degli ordinali transfiniti.

L'esistenza di un  $\alpha_\xi$  tale che  $\phi(\alpha_\xi) = \alpha_\xi$  è a questo punto una semplice conseguenza del fatto che l'insieme delle formule ben-formate di  $L$  ha una certa cardinalità: a furia di ampliare l'estensione e la controestensione di  $T$ , a un certo punto del processo avremo per forza di cose esaurito le formule classificate come vere o come false, e quel punto sarà per definizione un punto fisso di  $\phi$ . Più esattamente, l'esistenza di un punto fisso per la serie definita in (86)–(88) segue dalla cardinalità di  $L$  insieme al fatto che, poiché la semantica in (75)–(79) soddisfa il requisito di stabilità (74), la serie così definita è *conservativa*, cioè obbedisce al seguente principio di *monotonicità*:

- (89) Se  $\alpha[D^*_V, D^*_F]$  è un'estensione di  $\alpha[D_V, D_F]$ , cioè se  $D^*_V$  include ogni elemento di  $D_V$  e  $D^*_F$  include ogni elemento di  $D_F$ , allora, per ogni formula ben-formata  $A$  di  $L$ ,  $\alpha[D_V, D_F] \models A$  solo se  $\alpha[D^*_V, D^*_F] \models A$  e  $\alpha[D_V, D_F] \not\models A$  solo se  $\alpha[D^*_V, D^*_F] \not\models A$ .

#### 2.4. Applicazioni e limiti della teoria

Vediamo dunque di ricapitolare. La teoria di Kripke consta di due parti principali. La prima consiste nella definizione di una semantica formale non bivalente per un linguaggio,  $L$ , che condivide con le lingue naturali la caratteristica di essere sufficientemente ricco da poter esprimere la propria sintassi. Questa parte della teoria non presenta caratteristiche innovative, al punto che le condizioni di verità per le formule di  $L$  sono adottate in modo fondamentalmente arbitrario da una teoria semantica – quella di Kleene – che aveva ampia applicazione negli anni Settanta (per esempio, per il trattamento di fenomeni come la vaghezza o la mancanza di riferimento, inclusi i problemi citati in relazione alla semantica per le logiche modali quantificate basata su domini di quantificazione variabili). In realtà Kripke discute anche opzioni diverse, fra cui la semantica «supervalutazionale» dovuta a Bas van Fraassen, ma il quesito di *quale* sia la semantica migliore per un linguaggio del genere rimane sullo sfondo, purché valga il requisito della stabilità. La seconda parte della teoria è quella che giustifica il titolo dell'articolo di Kripke, perché è lì che si tratta di mostrare che tra i tanti modelli di  $L$  ve ne sono alcuni che autorizzano una lettura del predicato  $T$  come di un vero e proprio predicato di *verità*, e con ciò l'analogia tra  $L$  e una tipica lingua «semanticamente chiusa» come l'italiano. È questa seconda parte che raccoglie la sfida di Tarski. Ed è questa parte che offre una prospettiva nuova e per

certi aspetti rivoluzionaria dalla quale accostarsi allo studio del concetto di verità e dei problemi che lo affliggono. Resta da vedere come ciò si integri con la diagnosi dei problemi che abbiamo sommariamente riassunto nella sezione 2.2, e soprattutto in che misura si possa rivendicare di averli risolti.

Innanzitutto un chiarimento<sup>4</sup>. Abbiamo appena visto che sul piano tecnico il risultato principale consiste nell'identificazione dei modelli adeguati con quei modelli che corrispondono a un punto fisso in una serie del tipo definito in (86)–(88), l'esistenza del quale è garantita dal fatto che si tratta di una serie conservativa. È evidente, tuttavia, che la serie che abbiamo costruito è soltanto una delle tante che soddisfano questo requisito. Per esempio, anziché iniziare la serie con un modello  $\alpha_0$  che assegna a  $T$  un'estensione e una controestensione interamente vuote, cioè identificando  $\alpha_0$  con  $\alpha[\emptyset, \emptyset]$ , avremmo potuto iniziare con un modello che classificava (arbitrariamente) alcuni elementi di  $D$  nell'uno o nell'altro insieme. Per tornare al caso del nostro interlocutore italiano, questa possibilità corrisponde all'idea che il suo uso di «è vero» non fosse *completamente* nullo: per qualche motivo, sapeva *già* che questo predicato si applica correttamente a certi enunciati, per esempio certi enunciati di livello 0. Ora, non è difficile rendersi conto che anche iniziando la serie in questo modo alternativo, e definendo i passi successivi come in (87) e (88), il risultato sarebbe stato comunque una serie conservativa, e quindi avremmo comunque raggiunto un punto fisso. Questo significa che la procedura di Kripke ci consente in effetti di dimostrare l'esistenza, non di uno, ma di una molteplicità di modelli nei quali  $T$  rappresenta la verità. Volendo, una volta raggiunto un punto fisso potremmo ripartire, aggiungendo arbitrariamente qualche elemento all'estensione e alla controestensione di quel punto fisso e usando il modello così ottenuto come elemento iniziale di una nuova serie che terminerà in un nuovo punto fisso. La domanda quindi sorge spontanea: *quale* di questi modelli – quale punto fisso – dobbiamo privilegiare nella scelta di un modello adeguato per  $L$ , posto che  $\alpha$  sia un modello adeguato per  $L_0$ ?

Per Kripke la risposta è relativamente priva di interesse. Si può dimostrare che quello che si ottiene partendo con  $\alpha[\emptyset, \emptyset]$  è il punto fisso *minimo*, cioè tale che ogni altro punto fisso è una sua estensione (nel senso definito in (89)). Si può anche dimostrare che esistono punti fissi *massimali*, cioè tali che le loro estensioni non conducono a ulteriori punti fissi. E tra questi due estremi abbiamo una serie di punti fissi più o meno «ricchi», tra cui quelli che Kripke chiama

---

<sup>4</sup> Questo chiarimento si rivolge al lettore che abbia approfondito i dettagli formali della teoria di Kripke illustrati al termine della sezione precedente.



punti fissi *intrinseci*: punti fissi in cui nessuna formula riceve un valore di verità diverso da quello che riceve in altri punti fissi. Kripke trova che il punto fisso minimo sia la scelta più naturale, come del resto suggerisce anche la situazione immaginaria della persona che inizialmente non conosce affatto il significato di «è vero», e ha una certa simpatia anche per il punto fisso intrinseco più grande, che si dimostra essere unico e che ha la caratteristica interessante di fornire la più ricca interpretazione di  $T$  che non dipende da decisioni *arbitrarie*. Ma non è su queste candidature che Kripke ci invita a riflettere (sebbene proprio qui risieda la differenza cruciale rispetto al risultato di Martin e Woodruff citato sopra: nella terminologia appena introdotta, il modello adeguato di cui loro avevano dimostrato l'esistenza con metodi diversi corrisponde in sostanza a un punto fisso massimale, quindi a uno dei modelli meno «naturali», in quanto saturo di decisioni arbitrarie). Piuttosto, per Kripke il dato importante è che questa molteplicità di soluzioni consente di articolare con precisione alcune di quelle distinzioni concettuali che, come abbiamo visto nella sezione 2.2, erano in buona parte assenti nelle teorie elaborate sino allora. Ci limiteremo a considerare i due casi principali che abbiamo discusso, ma dovrebbero bastare a illustrare il potenziale esplicativo della mappa che si è venuta delineando.

Il primo caso riguarda la varietà di quegli enunciati che in un modo o nell'altro risultano problematici puramente in virtù della loro forma autoreferenziale, come il mentitore classico, (64), che dice di se stesso di essere falso, oppure quello che potremmo chiamare l'assertore, (67), che dice di se stesso di essere vero. Come abbiamo notato, c'è una bella differenza tra i due casi: il primo è vero se è falso ed è falso se è vero, sicché è impossibile assegnargli un valore di verità definito; il secondo è vero se è vero ed è falso se è falso, e assegnargli un valore di verità sarebbe semplicemente arbitrario. Naturalmente il linguaggio  $L$  non contiene una traduzione letterale dei due enunciati in questione, dato che il suo vocabolario non include espressioni indicali analoghe alla parola «questo» che figura sia in (64) che in (67). Tuttavia  $L$  contiene delle formule che posseggono esattamente le stesse caratteristiche semantiche. Per esempio, supponiamo che  $P$  sia un predicato sintattico la cui interpretazione (fissata dal modello standard di  $L_0$ ) assegna il valore  $V$  a un unico elemento del dominio, e precisamente al gödeliano della formula seguente:

$$(90) \quad \forall x(Px \rightarrow \sim Tx)$$

Siccome (90) dice che i  $P$  non sono veri, e siccome l'unico  $P$  è proprio il suo gödeliano, è chiaro che ci troviamo dinnanzi a una formula che dice di se stessa di non essere vera, come il mentitore (anzi, come la versione rafforzata del

mentitore, che non dipende dall'identificazione di «non vero» con «falso»). Analogamente, se l'interpretazione del predicato sintattico  $Q$  assegna  $V$  soltanto al gödeliano di

$$(91) \quad \forall x(Qx \rightarrow Tx)$$

questa formula dice di se stessa di essere vera, proprio come l'assertore. Bene, con riferimento a casi del genere, la teoria di Kripke fornisce una spiegazione molto chiara delle relative somiglianze e differenze. La somiglianza è che nessuna di queste due formule è semanticamente *fondata*, in un senso che adesso possiamo definire con precisione:

$$(92) \quad \text{Una formula ben-formata } A \text{ è fondata se e solo se } A \text{ è o vera o falsa rispetto al punto fisso minimo}$$

(da cui segue, per la monotonicità, che una formula fondata ha lo stesso valore di verità in *tutti* i punti fissi). D'altra parte, è facile rendersi conto della differenza: la formula del mentitore, (90), non è *mai* valutata come vera o falsa in un punto fisso; la formula dell'assertore (91) risulta vera (o falsa) in *qualsunque* punto fisso che includa nell'estensione di  $T$  (o nella sua controestensione) il suo gödeliano. Quindi il mentitore è paradossale, l'assertore no. Specificando un po' meglio il modo in cui  $L$  riesce a esprimere la propria sintassi (su cui abbiamo dovuto sorvolare per non appesantire troppo la presentazione), si potrebbero in questo modo caratterizzare anche altre importanti differenze e somiglianze a cui abbiamo accennato nella sezione 2.2. Per esempio, risulta che la formula corrispondente alla disgiunzione del mentitore classico e dell'assertore, che intuitivamente non è mai falsa ma può essere considerata vera, ha precisamente la caratteristica di risultare vera in un punto fisso intrinseco. E risulta che enunciati autoreferenziali ma del tutto innocui come (69) e (70), che dicono semplicemente di quante parole sono composti, sono addirittura fondati. Va da sé che queste distinzioni non definiscono ancora una tassonomia completa, ma si capisce che siamo sulla strada di una classificazione piuttosto accurata (al cui perfezionamento la letteratura successiva ha dedicato ampie risorse).

Questi esempi consentono di gettare luce anche sul secondo punto enfatizzato da Kripke nella sua diagnosi dei problemi legati al concetto di verità: la natura *contingente*, in molti casi accidentale, di tali problemi. Con riferimento al mentitore, per esempio, è sufficiente ipotizzare che nella formula (90)  $P$  non sia un predicato puramente sintattico per rendersi conto di come l'infondatezza di tale formula non sia determinabile a priori. Se l'interpretazione di  $P$  corrisponde a quella del predicato italiano «è un enunciato sulla lavagna», per esem-

pio, allora (90) dice quello che (71) diceva in italiano: che gli enunciati sulla lavagna non sono veri. E la verità o falsità di quest'enunciato nel punto fisso minimo, e prima ancora la sua autoreferenzialità, dipende dall'esatta composizione dell'estensione e della controestensione di  $P$  in quel modello, e quindi nel modello iniziale  $\alpha_0$ : dipende da quali (altri) enunciati compaiono davvero sulla lavagna secondo il modello. Identico è il caso del cretese dell'*Epistola a Tito*, (63), che corrisponde all'ipotesi in cui l'interpretazione di  $P$  rispecchi quella del predicato «è stato asserito da un cretese». E simile è anche il caso dei «circoli» paradossali illustrati dalla coppia (65)–(66) o dalla loro versione rafforzata, che in  $L$  potremmo di nuovo rappresentare mediante formule come (91) e (90), rispettivamente. Se l'interpretazione di  $P$  assegna il valore  $V$  soltanto al gödeliano di (91), e quella di  $Q$  assegna il valore  $V$  soltanto al gödeliano di (90), allora entrambe le formule risultano infondate e, più precisamente, paradossali: nessun punto fisso assegnerà loro un valore di verità definito. Se però almeno uno tra  $P$  e  $Q$  riceve un'interpretazione diversa, le cose cambiano ed entrambe le formule possono risultare fondate.

Sono solo esempi, ma tanto basti a illustrare l'efficacia e il potere esplicativo della teoria di Kripke. Se torniamo al contesto storico richiamato in apertura, si capisce come la pubblicazione di *Outline of a Theory of Truth* sia stata accolta con estremo interesse non solo dagli addetti ai lavori ma anche da parte di chi nutriva seri dubbi sulla possibilità di pervenire a un'analisi coerente del concetto di verità e del suo impiego nell'ambito di linguaggi non addomesticati alla rigida gerarchia tarskiana. Non è un'esagerazione dire che, da questo punto di vista, il 1976 segna uno spartiacque nella riflessione logico-filosofica su questi temi proprio come il 1959 segna uno spartiacque nello studio delle logiche modali. Certamente si tratta soltanto di un'*outline*, come recita il titolo, cioè di una formulazione sommaria, ed è un peccato che Kripke non abbia mai prodotto la versione più completa che annunciava all'inizio dell'articolo. Ciononostante, nel giro di poco tempo l'*outline* di Kripke è stata sottoposta ad applicazioni e sviluppi anche molto sofisticati, sul piano strettamente logico-matematico come su quello filosofico, e sin dagli inizi si sono moltiplicate varianti e alternative che al di là dei dettagli confermano l'impatto rivoluzionario di questo lavoro. (L'esempio forse più significativo è la cosiddetta teoria «revisionista» proposta da Hans Herzberger e, indipendentemente, da Anil Gupta e Nuel Belnap, autori nel 1993 del poderoso *The Revision Theory of Truth*. Ma il debito nei confronti di Kripke è palese anche nell'opera di autori come Van McGee, Aladdin Yaqūb e Keith Simmons e nelle più recenti teorie avanzate da Hartry Field, Tim Maudlin, Graham Priest e altri, nonché nei trattamenti assiomatici del predicato di

verità avviati da Solomon Feferman sul finire degli anni Settanta e ampiamente studiati nei due decenni successivi.)

A questo punto possiamo concludere con un'osservazione di carattere generale sulla portata della teoria. Kripke stesso non esitava a usare parole di cautela a questo riguardo, dicendo in modo esplicito che non la considerava una soluzione definitiva a tutti i problemi. Tuttavia c'è *un* problema la cui mancata soluzione potrebbe essere considerata indice di un limite intrinseco non solo della teoria così come Kripke l'ha abbozzata, ma dell'intero approccio sul quale essa si regge. Ed è un problema che riguarda la questione di fondo: possiamo davvero considerare L, con la sua bella semantica ricca di punti fissi, alla stregua di una tipica lingua «semanticamente chiusa» come l'italiano? Possiamo davvero dire di avere trovato, studiando L, una dimostrazione di come si possa parlare coerentemente una lingua capace di esprimere al suo interno tutte le verità che la riguardano?

Purtroppo la risposta non è del tutto affermativa. Come Kripke rilevava nelle pagine finali del suo articolo, la coerenza di L richiede per forza di cose un sacrificio sul piano espressivo. Per rendersene conto è sufficiente considerare fino in fondo che cosa succede nel caso di un enunciato che la teoria classifica come paradossale, quale il mentitore classico. Se  $A$  è una formula di L che esprime un enunciato di questo tipo, sappiamo con certezza che  $A$  non può risultare vera rispetto a un modello adeguato di L, cioè rispetto a un punto fisso  $\alpha_\xi$ . Sappiamo anche che  $A$  non può risultare falsa, e ciò equivale a dire che rispetto ad  $\alpha_\xi$  non può essere vera nemmeno la sua negazione (come si può verificare applicando la clausola (76) per la valutazione delle formule negate). Ora, nel linguaggio L, questi due fatti possono venire espressi attraverso le seguenti formule ben-formate:

$$(93) \quad \sim T^r A^r$$

$$(94) \quad \sim T^r \sim A^r$$

Tuttavia, se  $\alpha_\xi$  è un punto fisso, il fatto che  $A$  e  $\sim A$  siano né vere né false rispetto ad  $\alpha_\xi$  significa che nemmeno  $T^r A^r$  e  $T^r \sim A^r$  saranno vere o false rispetto ad  $\alpha_\xi$ , e quindi che non lo saranno nemmeno le rispettive negazioni. Questo segue immediatamente dal fatto che i punti fissi soddisfano la condizione di adeguatezza CA. Quindi (93) e (94) *non* saranno vere rispetto ad  $\alpha_\xi$ , come invece vorremmo dire. E questo significa che, sebbene l'adeguatezza di  $\alpha_\xi$  ci consenta di affermare che L è in grado di esprimere la propria nozione di verità attraverso il predicato  $T$ , e indirettamente la nozione di falsità, lo stesso linguaggio non dispone delle risorse per esprimere la *non verità* e la *non falsità*. L'unico

modo per descrivere correttamente la situazione consiste nell'ascendere al metalinguaggio:

(93') non  $\alpha_{\xi} \models T^r A^r$

(94') non  $\alpha_{\xi} \models T^r \sim A^r$

Ebbene, questo vuol dire che L *non* è perfettamente chiuso sul piano semantico: almeno in certi casi, il ricorso a un metalinguaggio più espressivo di L è necessario per poter esprimere dei fatti semantici che riguardano L.

Che le cose debbano andare in questo modo è del resto evidente se torniamo per un attimo alla versione rafforzata del mentitore, che in italiano fa leva proprio su una delle nozioni in questione:

(95) Quest'enunciato non è vero.

Se (95) è vero, allora i fatti corrispondono a ciò che dice, quindi deve non essere vero. D'altra parte, se (95) non è vero, allora i fatti non corrispondono a ciò che dice, e quindi deve essere vero. Insomma, (95) è vero se e solo se non è vero: una contraddizione. Avevamo già notato che questa versione del paradosso è particolarmente insidiosa perché non dipende dall'assunzione della bivalenza. Ora, sappiamo che (95) si può tradurre in L, per esempio attraverso una formula come (90). È evidente, quindi, che se potessimo tradurre in L anche il ragionamento che abbiamo appena fatto, e in particolare l'affermazione della non verità di (90), ci ritroveremmo anche in L con una contraddizione malgrado la semantica non sia bivalente. L'esistenza di uno scarto tra linguaggio oggetto e metalinguaggio, per quanto contenuto, è quindi *inevitabile* pena incoerenza.

Fino a che punto questo risultato costituisca un limite serio della teoria è a tutt'oggi oggetto di discussione. Sicuramente non è il limite drastico che caratterizza le teorie che impongono un rispetto rigido e assoluto della gerarchia linguistica per ogni uso del predicato di verità, e sul piano pratico possiamo anche dire che si tratta di un limite del tutto irrilevante. Ma che *qualche* uso rimanga illegittimo si presenta sicuramente come un limite considerevole sul piano teorico, e in questo senso si può pensare che il successo della teoria sia solo parziale. Nelle parole dello stesso Kripke: «Lo spettro di Tarski è ancora fra noi».

## Approfondimenti bibliografici

### 1. Logica modale

Per apprezzare la portata filosofica dei contributi di Kripke alla logica modale è sicuramente utile integrare la lettura del testo con quella di opere a carattere più generale. Il libro di J. Melia, *Modality* (McGill-Queens University Press, Montreal, 2003), e quello ancora più introduttivo di A. Borghini, *Che cos'è la possibilità* (Carocci, Roma, 2009), tracciano un quadro piuttosto completo delle tematiche principali. Per un approfondimento dei delicati intrecci tra aspetti logici e aspetti metafisici della modalità si può inoltre consultare l'introduzione all'antologia curata da M. J. Loux, *The Possible and the Actual* (Cornell University Press, Ithaca NY, 1979), mentre l'ampia introduzione al volume collettaneo curato da T. Szabó Gendler e J. Hawthorne, *Conceivability and Possibility* (Clarendon, Oxford, 2002), è consigliata per chi volesse approfondire l'intreccio tra modalità e questioni epistemologiche.

Altrettanto utile, soprattutto per misurare l'impatto innovativo della teoria di Kripke, è l'inquadramento delle tematiche affrontate in una più ampia prospettiva storica. Per una prima panoramica sulla logica modale nell'antichità e nel Medioevo è ancora attuale *The Development of Logic*, di W. C. Kneale e M. Kneale (Clarendon, Oxford, 1962; tr. it.: *Storia della logica*, Einaudi, Torino, 1972), soprattutto le sezioni II.7, III.2, e IV.2–3. Per un approfondimento della teoria aristotelica si possono inoltre consultare le monografie di G. Seel, *Die Aristotelische Modaltheorie* (Walter de Gruyter, Berlino, 1982), J. van Rijen, *Aspects of Aristotle's Logic of Modalities* (Kluwer, Dordrecht, 1989) e R. Patterson, *Aristotle's Modal Logic* (Cambridge University Press, Cambridge, 1995), mentre per le teorie medievali segnaliamo gli studi di S. Knuuttila, *Modalities in Medieval Philosophy* (Routledge, Londra, 1993), H. Lagerlund, *Modal Syllogistics in the Middle Ages* (Brill, Leiden, 2000) e P. Thom, *Medieval Modal Systems: Problems and Concepts* (Ashgate, Aldershot, 2003). Sulla nozione leibniziana di «mondo possibile» un buon punto di partenza è il capitolo 6 dell'*Introduzione alla filosofia di Leibniz* di M. Mugnai (Einaudi, Torino, 2001). Infine, per gli sviluppi più recenti ci si può ancora affidare al citato volume di Kneale e Kneale, specialmente alla sezione IX.4, e successivamente all'articolo di B. J. Copeland, *The Genesis of Possible Worlds Semantics* («Journal of Philosophical Logic», 31, 2002, pp. 99–137), che offre una ricostruzione molto dettagliata del contesto storico nel quale ha finalmente preso forma la teoria semantica di Kripke.

Per quanto concerne gli aspetti più tecnici, abbiamo cercato per quanto possibile di rendere l'esposizione autosufficiente, ma va da sé che una certa conoscenza della logica classica (non modale) costituisce lo sfondo naturale sul quale apprezzare i dettagli della teoria di Kripke, soprattutto qualora si volesse integrare la lettura del capitolo con quella degli articoli originali. A questo riguardo sono disponibili in italiano diverse opere a carattere manualistico, fra cui testi decisamente introduttivi come (in ordine crescente di complessità) *Il primo libro di logica* di E. Bencivenga (Boringhieri, Torino, 1986), *Logica da Zero a Gödel* di F. Berto (Laterza, Roma-Bari, 2007), *Logica* di A. C. Varzi, J. Nolt e D. Rohatyn (McGraw-Hill Italia, Milano, 2007<sup>2</sup>), *Corso di logica. Introduzione elementare al calcolo dei predicati* di D. Palladino (Carocci, Roma, 2002) e *Logica* di M. Mondadori e M. D'Agostino (Bruno Mondadori, Milano, 1997). Più avanzati, ma particolarmente utili per una comprensione delle proprietà matematiche (metalogiche) della logica classica, sono i testi di E. Casari, *Introduzione alla logica matematica* (UTET, Torino, 1997) e di V. M. Abrusci, *Logica matematica* (Laterza, Roma-Bari, 2000) e gli ormai classici *Introduction to Mathematical Logic* di E. Mendelshon (Van Nostrand, New York, 1964; tr. it.: *Introduzione alla logica matematica*, Boringhieri, Torino, 1972) e *Mathematical Logic*, di J. R. Shoenfield (Addison-Wesley, Reading MA, 1967; tr. it.: *Logica matematica*, Boringhieri, Torino, 1980).

Molto ricca è anche la letteratura a carattere manualistico dedicata esplicitamente alle logiche modali, a conferma degli enormi sviluppi della disciplina sotto l'influsso delle teorie semantiche proposte da Kripke. Tra i più classici segnaliamo i due volumi di G. E. Hughes e M. J. Cresswell, *An Introduction to Modal Logic* (Methuen, Londra, 1968; tr. it.: *Introduzione alla logica modale*, Il Saggiatore, Milano, 1973) e *A Companion to Modal Logic* (Methuen, Londra, 1984; tr. it.: *Guida alla logica modale*, Bologna, CLUEB, 1990), benché ormai superati da un terzo volume degli stessi autori intitolato *A New Introduction to Modal Logic* (Routledge, Londra, 1996). Due altri manuali piuttosto agevoli e aggiornati sono *Modal Logic for Philosophers*, di J. W. Garson (Cambridge University Press, Cambridge, 2006), e *Modal Logic: An Introduction to Its Syntax and Semantics*, di N. B. Cocchiarella e M. A. Freund (Oxford University Press, Oxford, 2008). Per un approfondimento della logica modale proposizionale e delle sue applicazioni si possono inoltre consultare i manuali di B. F. Chellas, *Modal Logic: An Introduction* (Cambridge University Press, Cambridge, 1980), di S. Galvan, *Logiche intensionali: Sistemi proposizionali di logica modale, deontica, epistemica* (Angeli, Milano, 1991), e di W. A. Carnielli e C. Pizzi, *Modalità e multimodalità* (Angeli, Milano, 2001), se non quelli più com-

pleti ma assai impegnativi di A. Chagrov e M. Zakharyashev, *Modal Logic* (Oxford University Press, Oxford, 1997) e di P. Blackburn, M. de Rijke e Y. Venema, *Modal Logic* (Cambridge University Press, Cambridge, 2001), mentre per la logica dei predicati segnaliamo soprattutto *First-Order Modal Logic*, di M. Fitting e R. L. Mendelsohn (Kluwer, Dordrecht, 1998). Infine, di recente pubblicazione – ma rivolto principalmente al lettore esperto – è il monumentale *Handbook of Modal Logic* curato da P. Blackburn, J. van Benthem e F. Wolter (Elsevier, Amsterdam, 2007), che fornisce un quadro completo dello stato dell'arte della disciplina per ciò che concerne le sue diverse interpretazioni (aleutica, epistemica, deontica, temporale, ecc.) e le sue numerose applicazioni (ad esempio in ambito linguistico, matematico e informatico).

Venendo al dibattito più propriamente filosofico, buona parte della letteratura è disponibile in forma di articoli apparsi su riviste o volumi collettanei in lingua inglese. Esistono tuttavia diverse raccolte antologiche che, nell'insieme, raccolgono i contributi più significativi e influenti degli ultimi decenni. Tra queste, oltre al già citato volume curato da M. J. Loux, sono essenziali le antologie di B. Linsky, *Reference and Modality* (Oxford University Press, Oxford, 1971; tr. it.: *Riferimento e modalità*, Bompiani, Milano, 1974), di C. Pizzi, *Leggi di natura, modalità, ipotesi* (Feltrinelli, Milano, 1978), e di D. Silvestrini, *Individui e mondi possibili* (Feltrinelli, Milano, 1979). Quest'ultima è particolarmente utile per chi volesse approfondire le questioni filosofiche che, a partire di Quine, hanno accompagnato l'integrazione delle modalità nella logica dei predicati. Esiste inoltre una cospicua letteratura dedicata a temi specifici. In particolare, un testo classico sullo scetticismo quineano è la dissertazione di D. Føllesdal, *Referential Opacity and Modal Logic*, risalente al 1963 ma pubblicata di recente con una nuova e aggiornata introduzione (Routledge, Londra, 2004). Sugli sviluppi del dibattito segnaliamo anche il volume collettaneo *Studies in Essentialism* curato da P. A. French, T. E. Uehling, Jr., e H. K. Wettstein (University of Minnesota Press, Minneapolis, 1986) e la rassegna di M. Della Rocca, *Recent Work on Essentialism* («Philosophical Books», 37, 1996, pp. 1–13 e 81–89). Sulla metafora dei mondi possibili è utile la lettura di J. Divers, *Possible Worlds* (Routledge, Londra, 2002), sebbene le tesi esaminate vadano ben oltre le sporadiche osservazioni di Kripke sino ad abbracciare la posizione dichiaratamente realista, ben più impegnativa e controversa, difesa da D. K. Lewis in *On the Plurality of Possible Worlds* (Blackwell, Oxford, 1986). Per ulteriori approfondimenti su questo tema è consigliabile anche la lettura di C. S. Chihara, *The Worlds of Possibility: Modal Realism and the Semantics of Modal Logic* (Clarendon, Oxford, 1998). Infine, è giusto riconoscere che sul piano fi-



losofico il successo della teoria di Kripke è stato determinato, in parte, anche dai successivi contributi di quegli autori che si sono fatti carico di approfondirne presupposti, implicazioni e aspetti critici sino ad elaborare varianti e teorie alternative, talvolta radicali. Tra questi, oltre al già citato D. K. Lewis, merita menzionare soprattutto A. Plantinga, *Essays on the Metaphysics of Modality* (Oxford University Press, Oxford, 2003), R. C. Stalnaker, *Ways a World Might Be: Metaphysical and Anti-Metaphysical Essays* (Clarendon, Oxford, 2003), e K. Fine, *Modality and Tense. Philosophical Papers* (Clarendon, Oxford, 2005).

## 2. Teoria della verità

Anche a questo riguardo, data la vastità del tema, è sicuramente utile integrare la lettura del capitolo con quella di opere a carattere generale. La monografia di G. Volpe, *Teorie della verità* (Guerini, Milano, 2005), è attualmente la più completa in lingua italiana per ciò che concerne le teorie contemporanee, ma per una panoramica più approfondita si raccomandano i testi di R. L. Kirkham, *Theories of Truth: A Critical Introduction* (MIT Press, Cambridge MA, 1992), e di W. Kühne, *Conceptions of Truth* (Clarendon, Oxford, 2003), possibilmente in congiunzione con l'antologia curata da M. P. Lynch, *The Nature of Truth: Classic and Contemporary Perspectives* (MIT Press, Cambridge MA, 2001). Più specifica, ma particolarmente utile per un esame critico della teoria di Kripke alla luce delle proposte che l'anno preceduta, è la monografia di G. Usberti, *Logica, verità e paradosso* (Feltrinelli, Milano, 1980), mentre per un confronto con le principali varianti e alternative che sono seguite si può partire dal recente studio di L. Bellotti, *Teorie della verità* (ETS, Pisa, 2008). Il libro di S. Soames, *Understanding Truth* (Oxford University Press, Oxford, 1999), è invece consigliabile soprattutto per chi volesse approfondire i presupposti filosofici della teoria, a partire dall'ipotesi che la nozione di verità debba in qualche modo trovare espressione nel linguaggio (per esempio attraverso un *predicato* di verità).

Sulla tematica dei paradossi, che come abbiamo visto costituiscono la sfida con cui ogni teoria deve misurarsi, si può abbinare la lettura del libro di Usberti a quella di F. Rivetti Barbò, *L'antinomia del mentitore da Peirce a Tarski. Studi, testi, bibliografia* (Jaca Book, Milano, 1986), che oltre a un excursus storico include la traduzione di alcuni testi classici tra cui il fondamentale articolo di Tarski più volte citato nel corso dell'esposizione. Le due antologie di R. L. Martin, *The Paradox of the Liar* (Ridgeview, Atascadero CA, 1970) e *Recent Essays on Truth and the Liar Paradox* (Oxford University Press, Oxford,

1984), sono comunque il riferimento essenziale per chi volesse approfondire lo stato del dibattito ai tempi dell'elaborazione della teoria di Kripke (il cui articolo è ristampato nel secondo volume). Per prendere il polso al dibattito più recente si raccomandano invece i contributi inclusi nelle due collettanee curate da J. C. Beall, *Liars and Heaps: New Essays on Paradox* (Clarendon, Oxford, 2003) e *Revenge of the Liar: New Essays on the Paradox* (Clarendon, Oxford, 2008) e nel volume curato da S. Rahman, T. Tulenheimo, ed E. Genot, *Unity, Truth and the Liar: The Modern Relevance of Medieval Solutions to the Liar Paradox* (Springer, Berlin, 2008). Tutti questi testi includono saggi di autori che hanno contribuito alla letteratura anche mediante scritti a più ampio respiro, e fra questi sono molti quelli che, direttamente o indirettamente, si confrontano con l'opera di Kripke. I principali sono V. McGee, *Truth, Vagueness, and Paradox: An Essay on the Logic of Truth* (Hackett, Indianapolis, 1990), A. Gupta e N. Belnap, *The Revision Theory of Truth* (MIT Press, Cambridge MA, 1993), A. M. Yaquūb, *The Liar Speaks the Truth: A Defense of the Revision Theory of Truth* (Oxford University Press, Oxford, 1993), K. Simmons, *Universality and the Liar: An Essay on Truth and the Diagonal Argument* (Cambridge University Press, Cambridge, 1993), T. Maudlin, *Truth and Paradox: Solving the Riddles* (Oxford University Press, Oxford, 2004), G. Priest, *Doubt Truth to Be a Liar* (Oxford University Press, Oxford, 2006), e H. Field, *Saving Truth From Paradox* (Clarendon, Oxford, 2008). Infine, il libro di A. Cantini, *Logical Frameworks for Truth and Abstraction: An Axiomatic Study* (Elsevier, Amsterdam, 1996) costituisce il punto di riferimento obbligato, benché impegnativo, della letteratura dedicata al trattamento assiomatico del predicato di verità alla quale si è accennato *en passant* verso la fine del capitolo.

Sul piano tecnico, abbiamo visto che la formulazione rigorosa della teoria di Kripke presuppone una certa familiarità con le procedure di «aritmetizzazione» della sintassi e con la teoria degli insiemi (inclusa la teoria degli «ordinali transfiniti»). Per quanto riguarda il primo aspetto, il terzo capitolo del libro di F. Berto, *Tutti pazzi per Gödel: la guida completa al Teorema di Incompletezza* (Laterza, Roma-Bari, 2008), contiene un'esposizione chiara ed esauriente, e per ulteriori approfondimenti è sempre attuale il testo ormai classico di D. R. Hofstadter, *Gödel, Escher, Bach: An Eternal Golden Braid* (Basic Books, New York, 1979; tr. it.: *Gödel, Escher, Bach: un'eterna ghirlanda brillante*, Adelphi, Milano, 1984), specialmente al capitolo IX. Per quanto invece riguarda la teoria degli insiemi e degli ordinali transfiniti, ci si può affidare a uno qualsiasi dei numerosi testi a carattere manualistico. Per una prima panoramica possono essere sufficienti il volume di P. Casalegno e M. Mariani, *Teoria degli*

*insiemi: un'introduzione* (Carocci, Roma, 2004) e i due libri di G. Lolli, *Dagli insiemi ai numeri* (Boringhieri, Torino, 1994) e *Guida alla teoria degli insiemi* (Springer Italia, Milano, 2008). Tra i testi più avanzati si possono invece consultare G. Lolli, *Teoria assiomatica degli insiemi* (Boringhieri, Torino, 1973), K. Kunen, *Set Theory* (North Holland, Amsterdam, 1983<sup>2</sup>) e T. Jech, *Set Theory* (Academic Press, New York, 2003<sup>2</sup>). Infine, e su queste basi, l'articolo di M. Fitting, *Notes on the Mathematical Aspects of Kripke's Theory of Truth* («Notre Dame Journal of Formal Logic», 27, 1986, pp. 75–88), è particolarmente importante per una comprensione approfondita degli esatti presupposti matematici su cui si regge la teoria di Kripke (e sui quali Kripke stesso non si pronuncia in modo esplicito).