

## 12. Logica, significato, competenza e comprensione

Alla fine del secondo capitolo abbiamo individuato l'insorgere di una sorta di circolo vizioso nei rapporti fra teorie modellistiche del significato e intelligenza artificiale di impostazione logicista. Alcuni filosofi del linguaggio si sono avvicinati all'IA per individuare possibili soluzioni ai problemi lasciati aperti del paradigma model teoretico nello studio del significato. I ricercatori di IA di impostazione logica sostengono l'impiego della semantica logica di tipo model teoretico per dare basi rigorose alla ricerca sui formalismi di rappresentazione e, in generale, sulle tecniche di ragionamento utilizzate in IA. Nei capitoli precedenti abbiamo visto che esistono effettivamente molti punti di contatto fra la ricerca logica in IA e il lavoro in semantica modellistica, sia a livello degli strumenti utilizzati, sia a livello delle problematiche affrontate, sia ancora a livello dei riferimenti culturali. Qui giunti, abbiamo accumulato abbastanza materiale per cercare di tracciare una sintesi, abbozzando al tempo stesso un primo tentativo di rispondere al problema delle relazioni fra questi due ambiti disciplinari. Tale sintesi non può essere ovviamente che provvisoria e, per molti aspetti, imprecisa, a causa di fattori quali lo stato di veloce evoluzione dell'IA negli ultimi anni, evoluzione che è tuttora in pieno corso, e la grande mole di ricerche svolte in questo ambito, che per di più hanno spesso un carattere frammentario. Si tratta dunque di un settore per molti versi ancora in via di assestamento. Tuttavia riteniamo che alcune linee generali possano essere individuate.

### 12.1 Semantica lessicale, simboli e riferimento

E' evidente in primo luogo che l'IA di impostazione logica, e, più in generale l'IA classica di tipo simbolico, non possono offrire una soluzione a *tutti* i problemi lasciati insoluti dalla semantica di tipo modellistico. E' quanto emerge nel caso del problema del significato dei termini primitivi extra logici (o, in termini linguistici, del problema del significato lessicale), come abbiamo già messo in evidenza nel paragrafo 7.2. Questo non significa tuttavia che nulla si sia ottenuto in questo settore. Diego Marconi (1989, 1992b) individua due componenti del significato lessicale, che possono essere caratterizzate come segue.

(1) Da un lato, si ha un complesso insieme di connessioni fra i vari elementi del lessico, che esprimono relazioni di tipo non logico che sussistono fra le parole di una lingua. Esempi di questo genere sono che i cani sono mammiferi, che gli uccelli di solito volano, che un padre è un essere umano di sesso maschile il quale ha dei figli che sono, a loro volta, esseri umani, che se qualcuno corre allora (di solito) si sposta, e così via. Tali informazioni vengono abitualmente utilizzate per eseguire inferenze, e fanno parte, a tutti gli effetti, del significato lessicale: "se uno non sa che, per mangiare, normalmente, si deve aprire la bocca diciamo che non sa cosa *significa* mangiare" (Marconi 1989, p. 79). Questa componente del significato lessicale viene detta da Marconi *competenza inferenziale*.

(2) L'altra componente del significato lessicale consiste nella capacità di mettere in relazione le parole con oggetti, eventi e situazioni nel mondo. Ad esempio, saper dire se un certo animale è un cane, o saperlo distinguere da un gatto, saper riconoscere qualcuno che corre, e così via. Tale capacità è quella che consente, si potrebbe dire, di "calcolare" i valori della funzione interpretazione per i simboli primitivi non logici di un linguaggio, e che consente quindi, in ultima istanza, di individuare le condizioni di verità effettive di un enunciato. Essa è detta da Marconi *competenza referenziale*<sup>1</sup>.

Queste due componenti del significato lessicale sono in parte indipendenti fra loro: può accadere che si conoscano le relazioni che legano una parola ad altre parole della lingua senza per questo saper individuare a che cosa quella parola si riferisca. Oppure, si può essere in grado di individuare il riferimento di un termine senza conoscere tutte le relazioni rilevanti che intercorrono fra esso e le altre parole della lingua.

Di fatto, come abbiamo già visto nel par. 7.2, di questi due problemi soltanto il primo è affrontabile nell'ambito dell'IA logicista. La competenza referenziale, ossia quegli aspetti del significato lessicale che concernono il *mapping* fra elementi del lessico e mondo, non risulta affrontabile nell'IA di impostazione logica più di quanto non sia affrontabile nell'ambito del paradigma model teoretico tradizionale. Tutto quanto è stato detto all'inizio del par. 2.1 sul riferimento dei simboli primitivi in semantica modellistica si applica senza alcuna modifica al paradigma dell'IA logicista. Riprendiamo qui l'argomento per tentare di mostrare che analoghe considerazioni valgono per tutta l'IA classica di impostazione simbolica. Un trattamento esauriente del problema della competenza referenziale eccede tuttavia gli scopi di questo lavoro. Per un'analisi approfondita dell'argomento rimandiamo a Marconi (1992b).

Prima di proseguire, è necessaria una precisazione. A proposito della competenza referenziale, Marconi afferma che:

In generale, sembra che ci sia bisogno di una qualche proiezione, diretta o indiretta, degli item lessicali sugli elementi di un altro sistema, che può essere il "mondo reale" stesso, o più comunemente un sistema che lo rappresenta, come un sistema di percezione, naturale o artificiale; o anche un *altro*

---

<sup>1</sup>Qui e nel caso precedente l'uso da parte di Marconi del termine *competenza* (anziché, ad esempio, *significato*) è dovuto alla volontà di sottolineare che, come vedremo più oltre, questi aspetti della semantica sono indissolubilmente legati al bagaglio cognitivo dei singoli parlanti (sul concetto di competenza cfr. la nota al cap. 2).

*linguaggio*, di cui sia già assicurata la connessione con il dominio di oggetti del primo linguaggio. Non vedo nessuna differenza di principio tra proiettare le parole di una lingua naturale su pattern nel campo visivo di un robot o invece sulle espressioni di, mettiamo, un linguaggio di interrogazione che ha accesso ai dati su cui stiamo ponendo domande in linguaggio naturale: non mi pare ci sia differenza fra una traduzione in un linguaggio davvero interpretato e una proiezione su un sistema non linguistico che sia effettivamente rappresentativo. (Marconi 1989, p. 79-80).

A nostro parere, questo passo rischia di causare fraintendimenti<sup>2</sup>. Riteniamo che, affinché il problema del riferimento dei simboli primitivi di un sistema possa dirsi effettivamente risolto, sia assolutamente necessario che alcuni simboli vengano effettivamente proiettati (direttamente o indirettamente) sul "mondo reale", e che far riferimento a qualche altro "sistema che lo rappresenta" sposti semplicemente il problema senza risolverlo. In particolare, riferirsi a un altro linguaggio "davvero interpretato", "di cui sia già assicurata la connessione con il dominio di oggetti del primo linguaggio" sia particolarmente problematico. O tale linguaggio è a sua volta effettivamente proiettato sul mondo, esiste cioè qualche dispositivo che calcola il riferimento (di alcuni) dei suoi simboli primitivi, e allora, in questo caso, anche i simboli primitivi del primo linguaggio sono proiettati sul mondo (in maniera indiretta, se si vuole), oppure ci si limita a supporre che esso sia interpretato, e in tal caso non si ha nulla di più di una traduzione in un "markerese", del tipo di quelle presupposte dalle teorie semantiche già criticate da Lewis (1972). In altri termini, si rientrerebbe nel concetto di semantica che Woods (1975) ha indicato come tipico dei linguisti, cioè semantica come traduzione in un linguaggio di rappresentazione di qualche tipo. A questo punto, si avrebbe addirittura qualcosa di meno delle funzioni interpretazione della semantica modellistica<sup>3</sup>.

Diversa è la situazione del caso di "un sistema di percezione, naturale o artificiale". Riteniamo che, per risolvere il problema del riferimento dei simboli primitivi di un sistema simbolico, il richiamo ad un sistema di tipo percettivo e motorio sia indispensabile. Se ci si pone da un punto di vista cognitivo, fare riferimento a un sistema di questo genere costituisce l'unica via per affrontare il problema. Ma questo non in quanto i simboli primitivi vengano proiettati *sul* sistema percettivo, quanto piuttosto in quanto un sistema percettivo/motorio di qualche tipo costituisce un *tramite* per proiettare il sistema di simboli sul mondo. A questo proposito è importante chiedersi cosa può significare dire che un sistema di percezione *rappresenta* il mondo. Se si intende che esso contribuisce a costruire un sistema di rappresentazione, e ad ancorarlo al mondo, allora è esattamente quello che intendiamo dicendo che esso è il tramite fra rappresentazioni e mondo reale. Se si intende che sia esso stesso un *sistema di rappresentazioni* che, a loro volta, si collegano al mondo chissà in quale maniera, allora ci sembra che questa sia una visione abbastanza inadeguata del problema della percezione.

In altre parole, riteniamo che il problema del riferimento dei termini lessicali possa ritenersi risolto solo nella misura in cui si possa risolvere il problema dell'*ancoramento dei simboli*<sup>4</sup> (*symbol grounding*) di un sistema computazionale ad oggetti, situazioni ed eventi extralinguistici nel mondo. Riteniamo inoltre che si possa asserire di aver costruito un modello della competenza referenziale se si è in grado di costruire un sistema che superi il *test di Turing totale*<sup>5</sup>, vale a dire, una versione estesa del test di Turing che concerna il comportamento globale, simbolico e non simbolico, di un sistema, e che non sia limitato al comportamento linguistico come avveniva nel test di Turing tradizionale. Ad esempio, si potrà dire di aver costruito un modello adeguato della competenza referenziale relativa al termine "bicchiere" se si sarà in grado di costruire un sistema in grado di reagire in maniera adeguata, ad esempio, a un ordine del tipo "prendi il bicchiere". Per reagire in maniera adeguata si intende, ad esempio, a seconda della situazione, prendere effettivamente il bicchiere, o rispondere "qui non c'è nessun bicchiere", o chiedere "quale bicchiere?" (se i bicchieri sono più di uno), e così via, tutto questo in relazione al mondo reale e non a un mondo simulato<sup>6</sup>.

<sup>2</sup>Non intendiamo quanto segue come una critica a Marconi, le cui posizioni al proposito risultano chiarite in (Marconi 1992b) secondo una linea molto vicina a quella da noi adottata. Riteniamo tuttavia che il passo sopra riportato possa essere considerato sintomatico di alcune confusioni ricorrenti nelle discussioni su questi argomenti.

<sup>3</sup>A questo proposito riteniamo che il problema della competenza referenziale non possa ritenersi risolto neppure ipotizzando la proiezione delle rappresentazioni semantiche su modelli mentali del riferimento nel senso proposto da Johnson-Laird. Come si è visto nel par. 1.2, i modelli mentali costituiscono sistemi di rappresentazione che necessitano a loro volta di essere interpretati sulla realtà. Rispetto al problema della semantica lessicale, riteniamo quindi che il contributo dei modelli mentali si debba ascrivere più all'aspetto della competenza inferenziale che non a quello della competenza referenziale.

<sup>4</sup>Nel senso ad esempio di (Harnad 1989, 1990).

<sup>5</sup>Ancora nel senso di (Harnad 1989, 1990).

<sup>6</sup>Si noti che qui non intendiamo assolutamente entrare nel merito se un sistema del genere sia dotato o meno di *intenzionalità*. Come è ben noto (Agazzi 1967, 1981; Searle 1980) l'intenzionalità è uno degli aspetti più problematici per le tesi dell'IA forte (Searle 1980). Qui intendiamo mantenere distinto il problema della competenza referenziale, che si può considerare risolto da un sistema che superi appunto il test di Turing totale, e che quindi ammette un controllo di tipo intersoggettivo, dal problema dell'intenzionalità. Quest'ultimo potrebbe essere caratterizzato come il problema di

Riteniamo che il problema dell'ancoramento dei simboli al mondo non possa essere affrontato, oltre che con gli strumenti dell'IA di impostazione logica, neppure all'interno del paradigma simbolico dell'IA in senso più generale, almeno se si interpreta l'attributo "simbolico" in senso forte. A questo proposito è necessaria una precisazione. Quando parliamo di paradigma simbolico in senso forte intendiamo fondamentalmente un'accezione di simbolico che coinvolge un'ipotesi semantica sull'interpretazione dei costrutti impiegati nella costruzione di modelli computazionali delle attività cognitive. La distinzione è ben sintetizzata dal seguente passo di Smolensky:

Chiamando l'approccio tradizionale ai modelli cognitivi "paradigma simbolico" intendo porre l'accento sul fatto che in questo approccio le descrizioni cognitive sono costituite da entità che sono simboli sia nel senso semantico di riferirsi ad oggetti esterni, sia nel senso sintattico di essere trattate tramite manipolazione simbolica. (Smolensky 1988, p. 61 della trad. it.).

La accezione semantica di "paradigma simbolico" è ben esemplificata dalla *knowledge representation hypothesis* di Brian Smith, a cui rimandiamo (par. 3.1). Per quanto riguarda l'accezione sintattica, l'ipotesi simbolica si limita a postulare che i costrutti computazionali e le strutture dati mediante cui è realizzato un modello computazionale siano i costrutti tipici della programmazione simbolica (come ad esempio alberi, strutture PROLOG o liste LISP), elaborate da un'architettura computazionale del tipo di quella della macchina di von Neumann. Una buona illustrazione dell'accezione sintattica di "simbolico" è il seguente passo di Levesque:

Userò il termine "simbolo" per indicare una configurazione di stati componenti relativamente passivi, i quali possono essere riconosciuti, distinti fra loro e manipolati dai componenti più attivi di una architettura computazionale. Sono esempi tipici stringhe di bit, di lettere o di cifre. I simboli possono *rappresentare* o meno qualcosa. Perciò, i termini "segno" [*token*] o "carattere" sarebbero forse più appropriati. (Levesque, 1988, p. 385).

E' evidente che questa accezione di simbolo è molto più debole di quella implicata *knowledge representation hypothesis* di Smith.

Qui sosteneremo la tesi secondo cui il problema del riferimento dei simboli primitivi, o il problema del *symbol grounding*, non possono essere affrontati nell'ambito di un paradigma puramente simbolico, se si assume "simbolico" nella sua accezione semantica. Il problema è che, dati i simboli primitivi di un sistema simbolico, se ne vogliono modellare i meccanismi di riferimento restando all'interno del paradigma simbolico inteso in senso semantico, non resta altra via che ipotizzare altre computazioni che a loro volta dovranno fare riferimento a nuovi simboli, dei quali a loro volta si dovrà spiegare il riferimento, entrando così in un regresso all'infinito. In altri termini, la questione può essere posta come segue. Si vuole individuare un dispositivo computazionale che calcoli i valori della funzione interpretazione per i simboli primitivi del sistema. Tuttavia, se per fare ciò, si introducono nuovi simboli, ci si limita a spostare il problema, e si sarà costretti a calcolare il valore di una qualche funzione interpretazione anche per questi ultimi, e così via all'infinito. Unica alternativa è quella di supporre che certi simboli siano connessi ai loro referenti in virtù di qualche qualità intrinseca, cadendo così in una "teoria magica del riferimento" del tipo di quelle già criticate da Wittgenstein (1958) e da Putnam (1981)<sup>7</sup>. Questo argomento è esposto in maggiori dettagli in (Frixione et al. 1989) e in (Frixione e Spinelli 1992), dove si suggerisce che una possibile via di uscita potrebbe consistere nell'ipotizzare un livello di computazione subsimbolico nel senso di Smolensky (1988), ossia di un livello computazionale in cui le unità minime di analisi siano *subsimboli*, cioè entità computazionali cui non sia assegnata una semantica locale specifica, indipendente dal contesto, nel senso che non è possibile immaginare una funzione interpretazione che metta in corrispondenza tali entità su oggetti del mondo, oggetti esterni al sistema, che fanno parte del suo "dominio di competenza". Si noti che tali unità (semanticamente) subsimboliche potrebbero (almeno in linea di principio) essere realizzate mediante entità di tipo simbolico nell'accezione sintattica del termine. Tuttavia, ciò su cui i "simboli" di un meccanismo subsimbolico del genere vanno "interpretati" sono, in ultima analisi, stati interni del sistema stesso, come ad esempio stati di sensori o di attuatori, *pixel* di una matrice *bitmap*, celle di un automa cellulare, valori di attivazione delle unità di una rete neurale, e così via.

Secondo Woods (1980), come in parte abbiamo già visto, uno dei maggiori contributi dell'IA alla teoria semantica consiste nello sviluppo di una semantica di tipo *procedurale*. Una semantica di questo genere estenderebbe la tradizionale semantica modellistica dando corpo alla nozione di condizioni di verità mediante lo sviluppo di procedure

---

come una rappresentazione venga esperita soggettivamente come dotata di significato, e che in quanto tale quindi non può ammettere un controllo intersoggettivo (si veda Frixione e Spinelli 1992).

<sup>7</sup> Z. Pylyshyn, uno dei maggiori teorici della linea simbolica nelle scienze cognitive, concorda sul fatto che i dispositivi che mettono in relazione i simboli con le entità extra-mentali (dispositivi che Pylyshyn chiama *transducer*) devono essere di natura non simbolica. Per Pylyshyn tuttavia ne segue che tali dispositivi dal punto di vista cognitivo devono essere assunti come primitivi: essi hanno natura esclusivamente fisica e non computazionale, e la loro struttura non può essere indagata dalle scienze cognitive (Pylyshyn 1984, cap. 6).

in grado di calcolare il significato dei simboli primitivi del sistema di rappresentazione utilizzato. La semantica modellistica della tradizione tarskiana, nella prospettiva di Woods, definisce le procedure in base alle quali può essere calcolato il valore di verità di una formula complessa a partire dal significato dei suoi componenti. Tuttavia, essa si ferma a livello di costituenti atomici. Per Woods l'approccio procedurale può essere esteso ulteriormente:

Se una persona comprende a pieno un termine che caratterizza una classe di entità, e se le si presenta qualcosa che sia una chiara istanza di un membro di quella classe, allora quella persona può riconoscerlo e dire "sì". Analogamente, può rifiutare qualcosa che chiaramente non sia un membro di quella classe. Se crediamo che nel fare ciò il cervello funzioni come qualche tipo di macchina fisico/elettrica, allora chiaramente *c'è qualche procedura che viene eseguita e che riconosce i membri della classe e rifiuta i non membri.* (Woods 1980, p. 316; corsivo nostro).

Per quanto riguarda la maniera di realizzare tali procedure, Woods afferma:

le primitive procedurali di un linguaggio di programmazione di alto livello come il LISP mettono a disposizione quello che sembra essere un utile insieme di operazioni di base mediante le quali costruire potenziali "funzioni di significato" [...]. Per esempio, il paradigma procedurale permette di costruire "primitive" per misurare la forza di un *pattern matching*, correlazioni statistiche e somme pesate per mezzo delle stesse primitive procedurali di base che si adoperano per caratterizzare il significato dei termini logici E, O, e PER OGNI. (Woods 1980, pp. 316-317).

Tutto questo potrebbe anche dimostrarsi vero. Quello che però si deve sottolineare è che, in base alla distinzione sopra introdotta, nei due casi il LISP svolge due ruoli diversi. Nel caso delle costanti logiche, le strutture (sintatticamente) simboliche del LISP vengono usate per implementare i costrutti simbolici (in senso semantico) di un formalismo logico e le procedure che li elaborano. Nel caso delle operazioni legate al *grounding* dei simboli primitivi, le strutture (sintatticamente) simboliche del LISP vengono usate per implementare processi di tipo subsimbolico, o comunque non di tipo simbolico in senso semantico.

Per quanto riguarda invece la componente del significato lessicale di cui al punto (1), l'IA di tipo simbolico, e l'IA di impostazione logicista in particolare, hanno offerto strumenti che possono risultare preziosi. Come abbiamo visto, i postulati di significato classici sono troppo poco espressivi per catturare molti tipi di relazioni lessicalmente rilevanti che sussistono fra le parole di una lingua. Postulati di significato scritti, ad esempio, in logica *fuzzy* o in qualche logica non monotona<sup>8</sup> possono risultare estremamente utili allo scopo. Sulla rilevanza delle inferenze di tipo non monotono per l'elaborazione di postulati di significato adeguati a rappresentare il lessico delle lingue naturali, la situazione è ben sintetizzata in due recenti articoli di Richmond Thomason (1991a, 1991b). Thomason, ad esempio, afferma che:

Il lavoro logicista in IA ha generalmente riconosciuto l'esigenza di accrescere il quadro logico proposto da Frege in maniera da affrontare i problemi del ragionamento di senso comune. La linea di sviluppo più generalmente condivisa consiste nell'incorporare la non monotonicità nella logica. E questa caratteristica risulta essere precisamente quello che serve per risolvere molti dei problemi che emergono nella semantica lessicale di stile montagoviano. [...] Il fatto che la semantica lessicale sia piena di generalizzazioni che ammettono eccezioni non costituisce certo una sorpresa, e una tecnica generale per esprimere tali generalizzazioni potrebbe estendere in larga misura l'adeguatezza delle teorie logiciste del significato lessicale. (Thomason 1991a, p.4)

Thomason fa l'esempio di inferenze del linguaggio naturale come:

*John sta attraversando la strada.  
Quindi, John avrà attraversato la strada.*

E' ragionevole assumere che un'inferenza di questo tipo sia basata sul significato di "attraversare". Per poterne rendere conto, tuttavia, bisogna risolvere una forma di *frame problem*. Il punto è che quando qualcuno sta attraversando una strada, di solito raggiunge l'altro marciapiede. Però può anche succedere che ciò non si realizzi. Ad esempio, John potrebbe finire sotto una macchina. Inferenze di questo genere, dice Thomason, devono essere considerate dei "*natural language defaults*". Qualche forma di non monotonicità deve essere implicita nel significato del verbo *attraversare*. A proposito della *circumscription* di McCarthy, Thomason (1991b) afferma che essa potrebbe risultare appropriata per affrontare dal punto di vista logico molti problemi rilevanti per la semantica lessicale. Ad esempio il problema di

<sup>8</sup> Oppure in un sistema a *frame* o mediante una rete semantica che ammette eccezioni, che, se si accetta la "tesi di Hayes", è la stessa cosa.

costrutti generici come nelle frasi *gli uccelli volano* oppure *io guido verso il posto di lavoro*, oppure l'informazione che un *cacciavite* è usato normalmente per avvitare viti. O ancora, il problema di fornire una teoria della causalità per interpretare i costrutti linguistici di tipo causale. Gli aspetti connessi alla non monotonicità del ragionamento su azioni (ad esempio, il fatto che le cose continuano ad andare in un certo modo se non accade nulla di anomalo) erano già stati individuati come rilevanti da chi aveva lavorato sulla rappresentazione del significato lessicale dal punto di vista della semantica di Montague, come ad esempio Dowty (1979). Tuttavia non erano ancora state sviluppate le tecniche formali per affrontare il problema in maniera abbastanza sistematica.

Inoltre, Thomason sottolinea la continuità fra le ricerche sulla logica non monotona e la tradizione modellistica in semantica:

Gli approcci alla non monotonicità disponibili potrebbero essere incorporati senza difficoltà nello schema della semantica di tipo montagoviano senza alcun cambiamento alla parte della logica che non ammette eccezioni. (Thomason 1991a, p. 4)

E, più in particolare, a proposito della *circumscription*:

Fra le teorie disponibili del ragionamento con eccezioni che potrebbero essere applicate alla semantica lessicale, la *circumscription* è quella che forse può essere armonizzata in maniera più facile e naturale con il lavoro già esistente nel quadro proposto da Montague, in quanto essa mette a disposizione un modulo logico integrato che si basa su una logica di ordine superiore. [...] Il progetto di sviluppare una spiegazione logica ampiamente efficace delle relazioni semantiche fra gli elementi lessicali di una lingua naturale è, in quanto a scopo, grosso modo comparabile al progetto di sviluppare una teoria di alto livello della conoscenza di senso comune. (1991b, pp. 457)<sup>9</sup>.

## 12.2 Un paradigma modellistico "liberalizzato"

Al di là della compatibilità formale fra i nuovi strumenti logici offerti dall'IA e le tradizionali tecniche della semantica modellistica, vi è tuttavia un punto che merita di essere evidenziato. Abbiamo già messo in luce (par. 2.1) che, rispetto al problema della semantica lessicale, non è in generale possibile una distinzione netta e definitiva fra informazioni semantiche da un lato (nel senso di informazioni di tipo analitico, che dipendono esclusivamente dal significato dei termini, e che sono vere in maniera necessaria) e informazioni fattuali, di tipo contingente, dall'altro<sup>10</sup>. Si consideri la parola "gatto". Certe informazioni, come ad esempio che "gatto" è un nome comune, non hanno un carattere fattuale. Altre, concernenti ad esempio i dettagli dell'apparato riproduttivo dei gatti o le loro abitudini alimentari sono certamente di natura fattuale, e ciò non costituisce un problema in quanto a nessuno verrebbe in mente di affermare che il fatto che i gatti mangino certe cose faccia parte del significato del termine "gatto". Ma si consideri l'informazione che i gatti sono animali. Certamente se qualcuno non sapesse che un gatto è un animale saremmo propensi ad affermare che egli non sa cosa significhi "gatto". Tuttavia, si tratta pur sempre di un'informazione fattuale che, afferma ad esempio Putnam (1970), potrebbe risultare falsa qualora si scoprisse che in realtà i gatti sono robot inviati da Marte. Chi non fosse convinto dall'esperimento mentale di Putnam, pensi ad altri esempi di informazioni dello stesso genere, che risultano inequivocabilmente di natura non analitica e contingente. Ad esempio, che le balene sono mammiferi, o che i coralli o i batteri sono animali. Di norma, le informazioni di tipo prototipico che possono essere espresse mediante postulati di significato non monotoni sono informazioni a carattere fattuale e contingente, ed esse costituiscono la maggior parte delle informazioni relative ai concetti lessicali di senso comune. Questo punto è riconosciuto anche da Thomason, il quale ad esempio afferma:

Conoscere tali significati [i significati lessicali] comporta l'accesso a un ampio spettro di conoscenza rilevante. [...] Molti domini di senso comune offrono trame relativamente blande di ragionamento che ammette eccezioni, trame che sono basate su di un grande numero di concetti connessi fra loro in modo vago. In molti casi è difficile separare quello che è primitivo da quello che è derivato. Mettendo a disposizione abbastanza conoscenza di sfondo è possibile caratterizzare i significati dei termini, ma ciò raramente prende la forma di condizioni necessarie e sufficienti. E' difficile trovare metodi affidabili per articolare la conoscenza di sfondo e maniere generali per applicare tale conoscenza alla caratterizzazione dei significati. (Thomason 1991a, p. 2)

<sup>9</sup>Come esempio di integrazione fra *circumscription* e logiche intensionali Thomason cita (Thomason 1990), in cui regole di tipo non monotono vengono utilizzate per mantenere il coordinamento fra le credenze di sfondo che costituiscono la base comune di una conversazione.

<sup>10</sup>Su questo punto si vedano ancora (Marconi 1989 e 1992b).

Non si può individuare una distinzione netta, fa notare ancora Thomason, fra inferenze del tipo di quella precedente, relativa a John che attraversa la strada, e inferenze come la seguente:

*Alice ha usato il fiammifero per accendere il fuoco.  
Quindi, Alice ha strofinato il fiammifero.*

dove entrano in gioco generalizzazioni *defeasible* di senso comune che riguardano in generale la conoscenza del mondo (come ad esempio il fatto che, di solito, chi usa un fiammifero per accendere qualcosa lo ha prima strofinato), e che sarebbe dubbio se considerare parte della semantica lessicale.

E' interessante confrontare queste affermazioni con i passi sul problema del lessico nella semantica modellistica tratti da Thomason (1974), e riportati nel par. 2.1. Questo confronto dà la misura delle influenze dovute all'IA sul paradigma model teoretico<sup>11</sup>. Secondo le posizioni di Thomason (1974), in sostanza, nella misura in cui certe informazioni rilevanti per la descrizione del significato lessicale erano riconoscibili come informazioni a carattere fattuale, allora esse non dovevano essere considerate rilevanti per la teoria semantica: "i problemi della teoria semantica dovrebbero essere distinti da quelli della lessicografia" in quanto "la lessicografia deve prendere in prestito concetti da tutte le aree del sapere e della pratica". E ancora: "tale evoluzione [della teoria semantica] dovrebbe senza dubbio fermarsi assai prima di ottenere qualcosa che assomigli a un dizionario" in quanto "costruire un dizionario richiede una conoscenza considerevole del mondo".

Questi assunti erano ampiamente condivisi nell'ambito del paradigma modellistico "classico" in semantica: nella misura in cui gran parte del significato lessicale dipende da conoscenze fattuali, allora esso non è competenza della semantica, in quanto altrimenti una teoria semantica dovrebbe comprendere una summa enciclopedica di tutto il sapere umano. E' evidente però che una posizione "minimalista" di questo tipo non è praticabile da chi lavora in IA, che, per costruire sistemi che effettivamente funzionino, deve comunque fornire loro una rappresentazione abbastanza ricca del significato dei termini lessicali. Lo sviluppo di strumenti formali come le logiche non monotone consente di estendere la semantica filosofica in questa direzione, pur mantenendo una forte continuità dal punto di vista delle tecniche formali utilizzate<sup>12</sup>.

Ciò pone però ulteriori problemi, a carattere eminentemente filosofico, rispetto alla visione modellistica classica della semantica. Il significato dei termini del lessico non solo dipende in maniera determinante da informazioni di tipo fattuale, ma è il bagaglio stesso di informazioni fattuali a disposizione dei singoli parlanti a risultare rilevante (Marconi 1989, 1992b). Parlanti diversi associano ai termini del lessico insiemi diversi di conoscenze, e non è di norma possibile individuare un nucleo di conoscenze comuni a tutti i parlanti con il quale identificare il significato del termine. Anche se si ammette che, data una certa parola, esista un'intersezione fra le conoscenze relative al suo significato che ne hanno tutti i parlanti di una lingua, di norma tale intersezione è troppo piccola per costituire il "significato" della parola. Il significato lessicale dipende quindi in maniera determinante dal bagaglio cognitivo dei singoli soggetti: "da un certo punto di vista, non esiste *la* competenza semantica: ci sono *competenze semantiche* al plurale" (Marconi 1989, p. 78). Da questo punto di vista, Marconi individua una asimmetria fra semantica lessicale e semantica strutturale. Quest'ultima infatti si basa su regole di tipo compositivo che si possono considerare con buona approssimazione condivise da tutti i parlanti competenti. Ma ciò non accade con la semantica lessicale. E' altresì evidente che questa concezione del significato lessicale non è compatibile con l'antipsicologismo fregeano tradizionalmente condiviso in ambito model teoretico. Certamente, i significati lessicali così intesi non coincidono del tutto con le "immagini mentali" contro cui polemizzava Frege, che sono contenuti di coscienza del tutto privati e incomunicabili<sup>13</sup>. Tuttavia, è evidente che i significati lessicali così concepiti sono inconciliabili con la concezione fregeana di senso, inteso come entità logica comune a tutti i parlanti, ed anzi dotata di una consistenza ontologica del tutto indipendente da ciò che avviene nelle menti dei singoli soggetti.

Un ulteriore elemento che rende difficilmente compatibili gli strumenti offerti dall'IA logicista con gli assunti fregeani è costituito dall'indebolimento del principio di composizionalità del significato. Per come lo avevamo formulato nel par. 1.1, tale principio impone che il significato di un'espressione complessa sia funzione esclusivamente del

---

<sup>11</sup>Richmond Thomason è un autore che abbiamo incontrato diverse volte nel corso di questo lavoro, e che ancora incontreremo nel seguito. Il suo percorso intellettuale è in un certo senso paradigmatico dell'evoluzione della semantica modellistica posta in contatto con gli stimoli provenienti dall'IA di impostazione logica. Partito dalle posizioni "classiche" della teoria modellistica del significato, si è in seguito avvicinato alla ricerca sulla rappresentazione della conoscenza in IA, ponendosi al tempo stesso in maniera approfondita il problema dei fondamenti degli approcci cognitivi e computazionali al problema del significato, e dei rapporti fra semantica filosofica ed IA.

<sup>12</sup> L'impossibilità di distinguere in maniera netta e definitiva fra informazione semantica e conoscenza fattuale apre la strada al rischio di posizioni semantiche di tipo *olistico*, in base alle quali il significato di ciascuna espressione di un dato linguaggio dipende dal significato di tutte le espressioni del linguaggio stesso. Qui non affronteremo questo aspetto del problema, e rimandiamo a Marconi (1992b) per una discussione.

<sup>13</sup>Sulle differenze fra lo psicologismo contro cui polemizzava Frege e lo psicologismo dell'IA e delle scienze cognitive sono interessanti le osservazioni di Thomason (1979) che prenderemo in considerazione nel seguito (par. 12.4).

significato dei suoi componenti e della sua struttura. In IA sono stati individuati tipi di ragionamento in cui, al contrario, è essenziale che il significato di un'espressione dipenda anche dal contesto costituito dalla teoria in cui esso compare. Ad esempio, a proposito del problema dell'eredità con eccezioni nelle reti semantiche, e della difficoltà di ricondurre tali aspetti ad un trattamento di tipo logico (cfr. par. 7.1), Thomason (1992) individua nell'assunzione classica della composizionalità del significato un limite alla flessibilità della logica nel costruire modelli del ragionamento di senso comune. Egli afferma:

La teoria dei modelli tarskiana [...] costituisce un approccio di grana relativamente grossolana, e questo è dovuto al requisito che le condizioni di verità di espressioni logiche complesse debbano dipendere in qualche modo in maniera uniforme dalle condizioni di verità delle loro parti. Questa caratteristica persiste anche qualora i modelli classici siano estesi in maniera da includere mondi possibili, domini di ordine superiore, valori di verità non classici, e simili. (Thomason 1992, p. 199).

In generale, il principio di composizionalità non vale nel caso delle logiche non monotone. Ad esempio, in base al principio di composizionalità, data una formula del tipo  $\alpha \wedge \beta$ , un'interpretazione  $I = (\varphi, D)$  è un modello di  $\alpha \wedge \beta$  se e soltanto se  $I$  è un modello di  $\alpha$  e al tempo stesso  $I$  è un modello di  $\beta$ . Quindi ogni modello di  $\alpha \wedge \beta$  è anche un modello di  $\alpha$ . La stessa cosa non vale in logica non monotona. Si supponga di operare nell'ambito della *formula circumscription*. Sia data una formula come:

$$(*) \forall x (-ab(x) \rightarrow P(x)),$$

e si supponga di voler circoscrivere il predicato  $ab$ . In questo caso, l'equivalente del fatto che  $I$  sia un modello di  $(*)$  è che  $I$  sia un *modello preferito* di  $(*)$  rispetto al criterio di preferenza sui modelli indotto dalla *formula circumscription* di  $ab$  (nel senso della logica preferenziale di Shoham; si veda il paragrafo 6.3). Perché  $I$  sia un modello preferito di  $(*)$ ,  $I$  deve essere un modello di  $(*)$  in senso tradizionale, e inoltre  $I$  deve essere minimale rispetto ad  $ab$  (cfr. par. 6.1). Il fatto che  $I$  sia minimale rispetto ad  $ab$  dipende in maniera cruciale dall'intera teoria in cui  $(*)$  compare. Se si considera una teoria composta esclusivamente da  $(*)$ , allora  $I$  è minimale se e soltanto se  $\varphi[ab] = \emptyset$  (e, di conseguenza,  $\varphi[P] = D$ ). Se ora consideriamo la congiunzione:

$$(**) \forall x (-ab(x) \rightarrow P(x)) \wedge ab(c),$$

risulta che i modelli preferiti di  $(**)$  sono i modelli di  $(**)$  in cui si ha che  $\varphi[ab] = \{c\}$  (dove  $c$  - in tondo - è l'oggetto del dominio su cui è interpretata la costante  $c$ ), i quali però non sono, a loro volta, modelli preferiti di  $(*)$ . Si potrebbe dire che in una logica non monotona non si può mai parlare di *modello di una formula* singola, ma solo di *modello di una teoria* nel suo complesso. Quindi, si può assegnare un valore semantico a una formula solo nel contesto di una teoria, e, in generale, non è possibile assegnare un valore semantico a una formula a partire dal valore semantico dei suoi componenti atomici.

Quindi, sebbene vi sia una forte continuità fra le tecniche e gli strumenti utilizzati dall'IA logicista da un lato, e la tradizione modellistica in semantica dall'altro, vi sono tuttavia numerosi punti di disaccordo fra la pratica dell'IA e le assunzioni di origine fregeana che sono alla base del paradigma modellistico (par. 1.1). Rispetto agli assunti originari di Frege, l'IA ha comportato una sorta di "liberalizzazione" del paradigma fregeano, che va soprattutto nella direzione dell'abbandono di una posizione "platonista" nei confronti della logica e delle entità semantiche, e pone al centro delle proprie indagini le capacità inferenziali di soggetti razionali finiti. Ad esempio, il problema dello psicologismo, inteso in senso lato, non si pone esclusivamente in relazione alla rappresentazione dei significati lessicali. In molti casi cambia il modo stesso di intendere la natura e lo scopo delle strutture semantico-modellistiche utilizzate. A proposito delle strutture semantiche basate su situazioni incomplete o incoerenti, utilizzate nella logica di Levesque della credenza implicita ed esplicita (par. 10.1) e in altri ambiti di rappresentazione della conoscenza (par. 7.1), Patel-Schneider afferma che le strutture semantiche vanno considerate come stati di conoscenza dell'intero sistema, e non modelli del mondo. Può accadere cioè che le strutture model teoretiche vengano viste non come una rappresentazione astratta di stati "oggettivi" del mondo (reali o possibili), ma come stati del mondo in qualche senso "modellati" da un soggetto cognitivo, o comunque rappresentati dal suo punto di vista. Più in generale, nell'IA logicista, anche se non viene richiesta l'adeguatezza empirica da un punto di vista psicologico delle strutture e dei costrutti utilizzati, tuttavia la logica viene utilizzata non come modello di razionalità astratto e "disincarnato", ma come strumento per rendere conto delle attività inferenziali di soggetti razionali individuali e finiti. Quindi la logica non è vista più (o non è vista più soltanto) come un canone normativo di razionalità assoluta, ma come strumento per costruire modelli di forme di ragionamento limitate, in quanto sottoposte a vincoli di vario genere, quali ad esempio limiti nelle informazioni disponibili, o limiti di tipo computazionale (tempo di calcolo disponibile, dimensioni della memoria). Da un lato, ad esempio, le logiche non monotone hanno a che fare con forme di ragionamento sottoposte a vincoli per quanto riguarda la possibilità di accedere a tutte le informazioni rilevanti, e di conseguenza con la necessità di trarre inferenze sulla base di informazioni incomplete. Per quanto riguarda i vincoli imposti da limiti di tipo computazionale, vanno in questa direzione

l'individuazione di sottoinsiemi trattabili computazionalmente dei calcoli logici (cfr. ad es. il par. 5.3 e il prossimo paragrafo), e, per molti versi, il problema stesso dell'onniscienza logica nel ragionamento epistemico.

Rispetto al problema dello psicologismo, e del rapporto fra IA logicista e settori della scienza cognitiva orientati più psicologicamente, è da notare che in molti casi l'IA logicista ha proceduto in modo "quasi empirico"<sup>14</sup>, nel senso che molti dei problemi che sono stati affrontati e formalizzati in ambito logico erano stati posti originariamente negli ambiti più "cognitivi" e più orientati psicologicamente della disciplina. Un esempio potrebbe essere quello dei *default* e della rappresentazione di concetti mediante tratti prototipici, problema che è stato studiato originariamente da psicologi (cfr. par. 2.2), e che è stato introdotto in IA da un "anti-logico" come Minsky. D'altra parte, nel par. 7.2 abbiamo già messo in luce alcuni parallelismi fra le soluzioni al problema della rappresentazione dei concetti lessicali proposte in psicologia cognitiva e il lavoro svolto sullo stesso problema in IA simbolica e logicista. Sempre a questo proposito, si vedano più oltre la discussione sulla trattabilità computazionale dei formalismi di rappresentazione della conoscenza, e le analogie sottolineate da Levesque (1986) fra *basi di conoscenza vivide* e modelli mentali nel senso di Johnson-Laird.

Si può ipotizzare che le relazioni che sussistono fra IA logicista e scienze cognitive più orientate psicologicamente (o approcci all'IA meno attenti alla dimensione logico-formale) siano per certi versi analoghe alle relazioni che Agazzi (1986) individua fra logica matematica e logica filosofica. Secondo Agazzi, il rapporto che sussiste fra logica matematica e logica filosofica presenta delle analogie con il rapporto fra fisica matematica e fisica sperimentale. Come la fisica sperimentale, la logica filosofica sviluppa spesso tematiche di ricerca in maniera informale, che spetta poi alla logica matematica formalizzare in vista di una loro completa comprensione. In molti casi la logica filosofica individua problemi per affrontare i quali non sono disponibili tecniche formali appropriate. Individuare tali tecniche è compito del logico matematico. E' in questo senso che la logica filosofica può essere ricondotta, se pure con le dovute cautele, a una disciplina sperimentale. Il suo compito è quello di affrontare nuovi campi di indagine, e se, nel fare questo, si trova spesso ad operare senza disporre degli strumenti formali adeguati, ciò è dovuto al fatto che tali strumenti spesso non sono ancora disponibili. Ma questo, ovviamente, nulla toglie al valore delle sue ricerche. Analogamente, superata la contrapposizione frontale fra "logicisti" e "anti-logici", l'IA simbolica di impostazione non strettamente logica, o di ispirazione più psicologica e cognitiva, può essere vista come una sorta di "IA sperimentale". Storicamente, i problemi posti dai settori dell'IA meno orientati verso gli aspetti formali hanno fornito, e presumibilmente continueranno a fornire, nuovi temi di ricerca e nuovi stimoli al lavoro dei logici (e non soltanto a quelli direttamente impegnati nel programma dell'IA logicista), sottoponendo al tempo stesso la ricerca logica a vincoli di adeguatezza (soprattutto di carattere computazionale) che i logici tradizionalmente avevano preso scarsamente in considerazione<sup>15</sup>. Thomason (1988) afferma che un problema generale nello sviluppo della logica è costituito dalla "difficoltà di sviluppare metodologie corrette per motivare le logiche. [...] non è facile articolare principi per controllare delle ipotesi sul ragionamento" (p. 323). Il lavoro logico in IA, con la sua enfasi sulle motivazioni computazionali e sulla costruzione di algoritmi, "promette non solo di suggerire teorie logiche nuove, ma mette a disposizione nuovi metodi per motivare e valutare le alternative teoriche" (p. 324). Gli sviluppi in questa direzione si prefigurano ricchi di ricadute anche per altri settori di applicazione della logica, fra le quali appunto le teorie del significato del linguaggio naturale. Ancora Thomason (1991b) afferma ad esempio: "cercando di mettere in relazione le rappresentazioni a procedure, il lavoro in IA impone vincoli sulle rappresentazioni di cui, io credo, c'è un forte bisogno nella semantica linguistica" (Thomason 1991b, p. 456).

### 12.3 Rappresentazioni finite e trattabilità computazionale

Dati gli scopi di tipo eminentemente computazionale della disciplina, è ovvio che la ricerca logica in IA si sia indirizzata prevalentemente verso l'individuazione di assiomatizzazioni finite dei sistemi di rappresentazione proposti. Per gli scopi dell'IA è infatti essenziale che una *teoria della dimostrazione* venga associata al trattamento model teoretico dei formalismi studiati. Questo tipo di preoccupazione è invece estraneo alle teorie del significato di tipo modellistico classico (per le quali, come vedremo, è problematico fornire un'assiomatizzazione finita). Tuttavia, anche nell'ambito delle teorie filosofiche del significato, esigenze analoghe erano già emerse. L'esigenza che una teoria semantica comportasse una rappresentazione finita dei significati era stata già espressa (pure se in un senso parzialmente diverso da quello richiesto dall'IA) da filosofi del linguaggio di tradizione analitica vicini - anche se esterni - alla tradizione della semantica modellistica in senso stretto, come ad esempio Davidson (1965) e Dummett (1959, 1976). Per

<sup>14</sup>In un'accezione simile a quella usata da Putnam (1975a) rispetto alla filosofia della matematica. Putnam nega il carattere puramente *a priori* della conoscenza matematica, e afferma che "la conoscenza matematica assomiglia alla conoscenza *empirica* - cioè [...] il criterio di verità in matematica, proprio come in fisica, è il successo delle nostre idee nella pratica" (p. 61). Secondo Putnam l'introduzione di nuovi concetti e di nuovi assiomi in matematica è spesso giustificata sulla base del loro successo in relazione a discipline di tipo empirico.

<sup>15</sup>Queste considerazioni possono essere estese, dall'ambito specifico dell'IA, al problema più generale dei rapporti fra logica e *computer science*. Molti degli sviluppi recenti della logica sono stati fortemente influenzati da considerazioni di ordine informatico e computazionale. Basti pensare allo sviluppo della *logica lineare* di Girard (Girard 1987 e 1991; Abrusci 1992).

Davidson, ad esempio, il requisito della finitezza è indispensabile per spiegare come una lingua possa essere appresa da un parlante finito.

L'IA tuttavia, in quanto interessata all'elaborazione di modelli computazionali, pone requisiti più vincolanti della semplice finitezza delle rappresentazioni. Affinché i modelli studiati possano essere effettivamente trasformati in programmi si devono individuare metodi di elaborazione di tipo effettivo, procedure decidibili che elaborano le rappresentazioni semantiche. Inoltre, affinché tali procedure abbiano un reale interesse (non soltanto dal punto di vista applicativo, ma anche in quanto modelli realistici delle prestazioni di soggetti finiti), tali procedure devono sottostare a ulteriori vincoli di trattabilità computazionale. Un modello computazionale non solo deve dare le risposte adeguate in un tempo finito, ma deve anche poterle dare in un lasso ragionevole di tempo.

Nell'ambito dell'IA di impostazione logicista è stato soprattutto Hector Levesque a porre il problema della trattabilità computazionale dei sistemi di rappresentazione della conoscenza<sup>16</sup>. Abbiamo già visto in precedenza (par. 5.3) come Levesque e Brachman abbiano individuato un *tradeoff* fra potere espressivo e trattabilità computazionale nei sistemi di rappresentazione della conoscenza: al crescere del primo la seconda inevitabilmente decresce. Abbiamo visto inoltre che Levesque e la sua scuola hanno studiato la trattabilità della sussunzione nei sistemi terminologici, e abbiamo segnalato come molti dei sistemi ibridi sviluppati a partire da KRYPTON siano stati concepiti in vista di un ragionevole compromesso fra espressività e buone proprietà computazionali. Levesque (1988) motiva il suo punto di vista circa la centralità teorica di questi problemi in IA, ed espone quale dovrebbe essere l'impatto della teoria della complessità computazionale sull'IA e, in generale, sulle scienze cognitive. Il fatto di prendere in considerazione il problema della complessità computazionale consentirebbe di giungere a una visione meno idealizzata della logica, in sintonia con gli scopi dell'IA. Secondo Levesque, la necessità di limitarsi a tipi di elaborazione che possano essere svolti entro limiti di risorse accettabili non è soltanto una questione di efficienza, la cui rilevanza si manifesta esclusivamente sul piano applicativo. Un modello computazionale di un'attività cognitiva che non sia trattabile computazionalmente non può dirci nulla a proposito di come quell'attività possa essere fisicamente realizzata. Le proprietà computazionali di un algoritmo sono in un certo senso "assolute", e non dipendono da come l'algoritmo è stato implementato. In altri termini, se qualcosa non può essere computato in maniera efficiente da una macchina di Turing, allora non può essere computato in maniera efficiente da alcun dispositivo di calcolo. Un modello cognitivo che non tenga conto dei vincoli computazionali ha quindi ben poco valore esplicativo. Di conseguenza, le considerazioni di complessità computazionale devono essere considerate vincolanti per qualsiasi modello, sia esso inteso come psicologicamente motivato o meno. Si potrebbe obiettare che tutti i formalismi interessanti dal punto di vista rappresentazionale risultano computazionalmente intrattabili. Secondo Levesque le cose non stanno necessariamente in questo modo. Innanzi tutto, ad esempio, si deve tenere conto di quali *input* devono essere considerati ammissibili per un dato modello. Non è detto che problemi che risultano in generale intrattabili non possano essere ricondotti a problemi trattabili se ci si restringe ad un insieme più limitato di dati in ingresso. Si tratta di un aspetto che ha anche una forte plausibilità cognitiva. Ad esempio, in ambito linguistico, gli esseri umani riescono facilmente a comprendere enunciati in cui sono presenti subordinate fino ad un certo grado di incapsulamento; tuttavia, al di là di tale grado la capacità di comprensione viene meno. Inoltre, secondo Levesque, esistono formalismi trattabili computazionalmente che risultano interessanti dal punto di vista cognitivo e per la rappresentazione della conoscenza. Un esempio è costituito dal tipo di logica rilevante descritto nel par. 10.1 in relazione al problema dell'onniscienza logica, che, dal punto di vista computazionale, risulta più attraente della logica classica<sup>17</sup>. La tesi di Levesque è che, in generale, le deviazioni dalla logica classica che vanno nella direzione di una migliore trattabilità computazionale sono in stretto rapporto con le modifiche che dovrebbero essere apportate alla logica per renderla psicologicamente più realistica.

Un esempio che potrebbe andare in questa direzione è costituito dalle *basi di conoscenza vivide* (Levesque 1986). Vediamo di che cosa si tratta. Intuitivamente, una base di conoscenza si dice *incompleta* se "ci dice che uno di un

<sup>16</sup>Qui e nel seguito si fa implicitamente riferimento alla teoria della complessità computazionale utilizzata abitualmente in informatica per valutare la trattabilità degli algoritmi. Per semplicità, tuttavia, non affronteremo alcun aspetto tecnico. Ci limitiamo a ricordare che un algoritmo si considera trattabile se il tempo di calcolo cresce al più polinomialmente rispetto alla lunghezza dell'*input*, mentre sono intrattabili gli algoritmi per cui il tempo di calcolo cresce esponenzialmente rispetto alla lunghezza dell'*input*. Per i concetti base si rimanda comunque, ad esempio, a (Garey e Johnson 1979). Una trattazione della teoria della complessità in una prospettiva logica si trova in (Börger 1989).

<sup>17</sup>Mentre infatti le procedure di decisione per la logica proposizionale classica hanno complessità esponenziale, la logica proposizionale utilizzata da Levesque è decidibile in tempo polinomiale. Nell'ambito dell'IA, il problema dell'onniscienza logica è stato spesso indagato in relazione alla trattabilità computazionale delle procedure di decisione per le logiche epistemiche (in questa prospettiva, si veda ad esempio Moses 1988). Da un punto di vista computazionale, l'assunzione di onniscienza logica è tanto più irrealistica, quanto più le procedure di decisione per la logica adottata sono difficilmente trattabili computazionalmente (con il caso estremo in cui la logica non è decidibile, come per la logica epistemica del primo ordine). Sulle proprietà computazionali delle logiche epistemiche proposizionali classiche si vedano (Ladner 1977) e (Halpern e Moses 1992). Vardi (1989) ha studiato le proprietà computazionali delle logiche dei modelli minimali. Un modello del ragionamento epistemico limitato sviluppato nella prospettiva della trattabilità computazionale è quello recentemente proposto da Cadoli e Schaerf (1992), cui faremo cenno più oltre.

certo numero di enunciati è vero, ma non ci dice quali" (Levesque 1986, p. 91). Ad esempio, una base di conoscenza è incompleta se da essa è possibile derivare una formula quantificata esistenzialmente del tipo  $\exists x \alpha(x)$  senza che sia possibile derivare  $\alpha(c)$  per qualche costante individuale  $c$  del linguaggio, oppure se è possibile derivare una disgiunzione del tipo  $\alpha \vee \beta$  senza che sia derivabile  $\alpha$  oppure  $\beta$ . Viceversa, in una base di conoscenza *completa* se, ad esempio, è possibile derivare  $\alpha \vee \beta$ , allora o  $\alpha$  oppure  $\beta$  devono essere derivabili. Le basi di conoscenza complete presentano migliori caratteristiche dal punto di vista computazionale. La conoscenza incompleta è infatti fonte di intrattabilità. Si consideri ancora la disgiunzione. Intuitivamente, se una formula del tipo  $\alpha \vee \beta$  deve essere utilizzata in un processo deduttivo, e non si sa se sia vera  $\alpha$  oppure  $\beta$ , allora entrambi i casi devono essere presi in considerazione separatamente. E' evidente che, nel caso che formule di questo genere siano coinvolte in più passi deduttivi, sorge un problema di esplosione combinatoria. Le basi di conoscenza vivide sono basi di conoscenza complete scritte in una forma particolare, tali da rendere accessibili le informazioni in esse rappresentate in maniera ancora più diretta. Nelle basi di conoscenza vivide esiste innanzi tutto una corrispondenza biunivoca fra costanti del linguaggio e oggetti del dominio. Inoltre, per ogni predicato atomico a  $n$  posti del linguaggio, se la relazione da esso denotata sussiste nel modello fra una  $n$ -pla di individui, allora ciò è rappresentato da una formula atomica corrispondente nella base di conoscenza. Quindi, una base di conoscenza vivida gode della particolare proprietà di essere isomorfa ad un suo modello. E' evidente che, dal punto di vista computazionale, le basi di conoscenza vivide sono particolarmente allettanti: fare inferenze in esse equivale ad "andare a vedere come stanno le cose" in un modello, o, in altri termini, a consultare una tabella. E' altrettanto evidente però che il loro potere espressivo è estremamente limitato: di fatto, le basi di conoscenza vivide non sono altro che un insieme finito di formule atomiche. Secondo Levesque, tuttavia, la conoscenza vivida è estremamente diffusa in molti ambiti di tipo cognitivo. Egli sottolinea l'analogia fra una base di conoscenza vivida e un'immagine. In questo senso le basi di conoscenza vivide sarebbero in parte analoghe ai modelli mentali di Johnson-Laird (Levesque 1988). Inoltre, molta della conoscenza che proviene direttamente dalla percezione sarebbe di tipo vivido. E' possibile utilizzare ragionatori di tipo ibrido (nel senso visto nel par. 5.3), riconducendo la maggior parte possibile della conoscenza in forma vivida, e facendo ricorso a tecniche di ragionamento più complesse solo qualora sia strettamente necessario. Inoltre, anche qualora si disponga di informazione di altro tipo, è possibile utilizzare tecniche di tipo euristico per ricondursi a forme di conoscenza vivida. Una maniera consiste nel rinunciare a parte delle informazioni originarie a favore della vividezza. Ad esempio, se si dispone dell'informazione (incompleta) che Giorgio è cognato di Mario oppure è suo cugino, si può asserire che Giorgio è parente di Mario. Oppure si può introdurre un nuovo predicato a due argomenti (atomico) *cognato\_o\_cugino*. O ancora, si possono utilizzare tecniche di ragionamento non monotono, come l'assunzione di mondo chiuso<sup>18</sup>, o regole di *default*. Ad esempio, sapendo che Tweety sta nuotando oppure sta volando, e che Tweety è un uccello, poiché gli uccelli di solito volano si può asserire che Tweety sta volando.

In generale, l'obiettivo della trattabilità computazionale per i sistemi di rappresentazione della conoscenza può essere perseguito mediante due diverse strategie<sup>19</sup>. Da un lato, è possibile indebolire il potere espressivo del linguaggio di rappresentazione, in modo da disporre di procedure inferenziali complete e trattabili. Dall'altro, si può mantenere la piena espressività del linguaggio e indebolire l'apparato deduttivo, ottenendo quindi la trattabilità computazionale a scapito della completezza. Abbiamo visto entrambi le soluzioni all'opera nel caso dei formalismi terminologici. In taluni casi si è tentato di individuare formalismi poco espressivi per cui il calcolo della sussunzione completa fosse trattabile. In altri si è mantenuta una espressività forte, rinunciando alla completezza degli algoritmi di sussunzione. Anche la logica della credenza esplicita di Levesque può essere vista in questa seconda prospettiva. La logica del *tautological entailment* da lui impiegata infatti può essere vista come un indebolimento dell'apparato deduttivo della logica proposizionale classica, in cui non vale la regola del *modus ponens* (e che quindi risulta incompleta rispetto alla semantica bivalente classica)<sup>20</sup>. Un ulteriore sviluppo di questa strategia è costituito dalle ricerche sul *ragionamento approssimato*. L'idea è di fare in modo che, in un processo inferenziale, si possano ottenere risposte via via "migliori" quanto maggiore è lo sforzo computazionale impiegato. Si vuole cioè che, fissati certi vincoli di risorse computazionali, il sistema sia in grado di fornire comunque risposte semanticamente fondate, le quali approssimino le prestazioni di un ragionatore classicamente corretto e completo<sup>21</sup>.

<sup>18</sup>L'assunzione di mondo chiuso (*closed world assumption*) è una tecnica di ragionamento non monotono che consiste nell'assumere che gli oggetti esplicitamente rappresentati in una base di conoscenza siano tutti e soli gli oggetti del dominio (Reiter 1978; Minker 1982; cfr. anche il cap. 7 di Lukaszewicz 1990). Essa può essere ricondotta a una forma particolare di *circumscription*. L'assunzione di mondo chiuso è molto utilizzata nell'ambito dei *data base*, ed è incorporata nel linguaggio di programmazione logica PROLOG. Essa è legata al concetto di *negazione come fallimento* (*negation as failure*): si assume che valga  $\neg\alpha$  qualora fallisca il tentativo di dimostrare  $\alpha$ .

<sup>19</sup>A questo proposito, si veda ad esempio (Cadoli 1993).

<sup>20</sup>Per tale modo di interpretare questo tipo di logiche si veda (Frisch 1987).

<sup>21</sup>Cadoli e Schaerf (1991) ad esempio hanno sviluppato un modello in base a cui sia possibile trarre conseguenze approssimate da una base di conoscenza in logica proposizionale. Ciò viene ottenuto definendo due successioni di relazioni di conseguenza logica (le prime sempre corrette, le seconde sempre complete), che convergono verso la

## 12.4 I paradossi della competenza semantica

Nel caso di una teoria del significato le richieste di finitezza e di trattabilità computazionale entrano tuttavia in conflitto con alcune difficoltà concernenti la possibilità stessa di una rappresentazione finita dei significati. L'argomento che riteniamo più importante contro la possibilità di ottenere una rappresentazione finita del significato dal punto di vista cognitivo è esposto in un lavoro di Thomason (1979) contro il progetto di una semantica cognitiva e, in generale, contro lo psicologismo in semantica. Si tratta di un tema che meriterebbe di essere esaminato in dettaglio, anche se, per la sua complessità, in questa sede non avremo lo spazio per esaurirne tutti gli aspetti. Un motivo collaterale di interesse del lavoro di Thomason consiste nel fatto che esso consente anche di chiarire alcuni aspetti del rapporto fra le teorie semantiche di impostazione cognitiva e i temi dell'antipsicologismo come erano stati formulati da Frege.

Thomason esamina l'antipsicologismo di Frege, prendendo in considerazione alcune delle motivazioni fregeane contro una fondazione di tipo psicologico della matematica. Nei *Grundlagen* Frege (1884) mette in evidenza alcune conseguenze inaccettabili del fatto di assumere che gli oggetti matematici siano definiti come entità di tipo mentale. Ad esempio, afferma Frege, se una dimostrazione fosse un processo mentale, nella dimostrazione del teorema di Pitagora si dovrebbe tenere conto del contenuto di fosforo del cervello di chi la esegue. Inoltre, ciascuno ha idee differenti sui numeri. Quindi, se i numeri fossero rappresentazioni mentali, dovrebbero esserci numeri diversi per ciascun essere umano: ad esempio, si dovrebbe ammettere che esiste un numero due differente per ciascuno di noi. Da un altro punto di vista, però, ci sarebbero troppo pochi numeri, in quanto ogni essere umano, nel corso della sua vita, può pensarne solo un numero finito. Abbiamo già accennato allo stretto legame in Frege fra l'antipsicologismo in semantica e l'antipsicologismo rispetto al problema dei fondamenti della matematica. Queste considerazioni sono molto simili a quelle che Frege adopera ad esempio per motivare la natura extrapsicologica ed oggettiva dei sensi (cfr. par. 1.1). Tuttavia, afferma Thomason, questo tipo di argomentazioni si può considerare superato dalle posizioni della moderna psicologia cognitivista, e dalle teorie psicologiche di impostazione computazionale. Tali teorie si basano infatti sulla distinzione fra competenza ed esecuzione<sup>22</sup>. La competenza viene caratterizzata nei termini di stati computazionali astratti, definiti funzionalmente. Si assume ad esempio che sia possibile caratterizzare la competenza in maniera indipendente dalle proprietà fisico-anatomiche del sistema nervoso, quali il contenuto di fosforo del cervello (questo è proprio quanto asseriscono le impostazioni di tipo funzionalista sulla natura degli stati mentali - cfr. par. 3.1). Ciò risponde alla prima obiezione. Per quanto riguarda il fatto che esseri umani diversi hanno idee diverse sui numeri, tali differenze possono essere limitate all'ambito dell'esecuzione, e non riguardano una teoria della competenza. Se caratterizziamo la competenza nei termini di una macchina astratta, ad esempio una macchina di Turing, allora, data la rappresentazione di un numero su due realizzazioni fisiche diverse della stessa macchina, si può caratterizzare in modo esatto la loro equivalenza in termini funzionali. Rispetto al problema della finitezza, infine, una teoria della competenza formulata nei termini di dispositivi di calcolo astratti e idealizzati (come appunto le macchine di Turing) consente di ipotizzare, ad esempio, una memoria potenzialmente espandibile all'infinito, in maniera che possa essere generato qualsiasi numero<sup>23</sup>.

Anche adottando uno "psicologismo della competenza", sorgono tuttavia, secondo Thomason, ulteriori problemi per una fondazione psicologica della semantica. Tali problemi costituiscono un ostacolo per la posizione degli obiettivi comuni fra semantica modellistica e semantica cognitiva, e impediscono di considerare le teorie modellistiche del significato (come ad esempio la teoria di Montague) alla stregua di modelli (di almeno una parte) della competenza linguistica (come è stato parzialmente sostenuto ad esempio da Partee 1979b). Thomason (1979) definisce il suo obiettivo come quello di "sostituire gli argomenti di Frege contro lo psicologismo dell'esecuzione in logica e in matematica con argomenti contro lo psicologismo della competenza in semantica" (Thomason 1979, p. 5). Un primo aspetto dell'argomentazione di Thomason può essere riassunto informalmente come segue. Le teorie semantiche di tipo cognitivo assumono che il significato delle espressioni linguistiche debba essere ricondotto a rappresentazioni di tipo mentale, e tali rappresentazioni mentali vanno intese come rappresentazioni simboliche di qualche tipo, che devono poter essere elaborate computazionalmente. La relazione che sussiste fra un'espressione linguistica e il suo valore

---

relazione di conseguenza logica proposizionale classica. Man mano che si procede lungo le due successioni, aumentano rispettivamente completezza e correttezza, e si ottengono quindi sempre migliori approssimazioni della relazione di conseguenza logica classica. A partire da questo modello, Cadoli e Schaerf (1992) hanno elaborato una logica epistemica per ragionatori non onniscienti in cui alle due successioni di relazioni di conseguenza viste sopra corrispondono due famiglie di operatori epistemiche modali. Tali operatori costituiscono approssimazioni via via più precise dell'operatore di conoscenza classico, e corrispondono a vincoli computazionali più o meno restrittivi sulle capacità inferenziali di un agente epistemico. La logica risultante, la cui semantica è data mediante una generalizzazione delle strutture a mondi possibili di Kripke, costituisce uno sviluppo delle logiche basate su situazioni non classiche (cap. 10) e sulla logica dell'*awareness* di Fagin e Halpern (par. 11.1).

<sup>22</sup>Per tale distinzione si veda la nota al cap. 2.

<sup>23</sup>Questo vale, ovviamente, per i numeri naturali, ma non funziona per i numeri reali. Ciò pone ulteriori problemi che si collegano al problema di Benacerraf in filosofia della matematica (cfr. oltre).

semantico (sia a livello intensionale che a livello estensionale) fa ovviamente parte del significato di quell'espressione. Tuttavia questo genere di relazioni semantiche non sono rappresentabili computazionalmente in quanto, in generale, non sono ricorsive, e il supporre di poterle rappresentare conduce a una contraddizione. Prendiamo ad esempio la nozione di verità. Essa è innegabilmente rilevante per caratterizzare il significato linguistico. Poiché, per i sostenitori della semantica cognitiva, il significato deve essere rappresentato mentalmente in termini computazionali, allora la nozione di verità dovrebbe essere rappresentabile computazionalmente. Tuttavia, in generale, la nozione di verità non è ricorsiva, e non è quindi rappresentabile in termini computazionali. E' ben noto dal teorema di Tarski (Tarski 1936) che un linguaggio logico ragionevolmente espressivo in cui si tenti di introdurre un predicato di verità per il linguaggio stesso diventa inconsistente.

Posta in questi termini, l'argomentazione di Thomason lascia spazio a molte perplessità, e sembra che non colga nel segno, fraintendendo gli scopi stessi del paradigma cognitivo in semantica. L'ipotesi che sta alla base della semantica cognitiva, in base alla quale il parlante deve "conoscere" i significati delle espressioni linguistiche, non comporta che il parlante debba incorporare una teoria semantica, ossia non comporta che il parlante debba conoscere la *teoria* del significato, nel senso di possedere una *rappresentazione* (simbolica) di come le espressioni della lingua si collegano ai propri significati extralinguistici (ed extramentali). Se, computazionalmente, possiamo considerare la competenza del parlante (o almeno una parte di essa) modellata nei termini di una teoria di qualche tipo (in senso *proof*-teoretico), allora i concetti di verità e di riferimento avranno a che fare con la metateoria semantica di quella. Questo ovviamente non significa accettare la posizione secondo cui i concetti di verità e di riferimento, e i concetti semantici più in generale, non sono rilevanti per una teoria cognitiva del significato. Vuol dire che la teoria della verità non deve essere *rappresentata nel* modello computazionale, ma piuttosto che al modello computazionale deve essere associata una teoria della verità (intesa in senso lato, come una teoria che metta in relazione le rappresentazioni cognitive con entità esterne al modello stesso), e che le computazioni devono operare in accordo con essa. E' chiarificatore in questo senso il commento di Levesque e Brachman alla *knowledge representation hypothesis* di Brian Smith che abbiamo riportato nel par. 3.2. Da questo punto di vista, l'argomento di Thomason sembra per certi versi paragonabile alle critiche di parte cognitivista al paradigma modellistico che abbiamo presentato nel par. 2.2, in base alle quali l'implausibilità cognitiva della semantica modellistica veniva intesa come implausibilità delle strutture modellistiche in quanto costrutti cognitivi. In entrambi i casi sembra che ci si trovi di fronte a una confusione fra due dei sensi di "semantica" identificati da Woods (1975) (par. 4.4), cioè fra il concetto di semantica "dei linguisti", intesa come traduzione in un sistema di rappresentazione esplicito di qualche tipo, e il concetto di semantica "dei filosofi", intesa come costruzione di una corrispondenza fra rappresentazioni ed entità extrasimboliche.

Anche qualora si accetti che un modello computazionale del linguaggio, nel suo complesso, debba rendere conto anche della competenza referenziale del parlante (nel senso del *symbol grounding*), e che quindi esso sia in grado, in un certo senso, di "calcolare il riferimento" delle proprie rappresentazioni, anche in questo caso ciò non comporta che debba esistere nella mente del parlante una *rappresentazione* simbolica della relazione fra le rappresentazioni cognitive e il loro riferimento, che il soggetto cognitivo cioè debba disporre nella "sua" teoria di qualcosa di analogo a un predicato di verità, o a un equivalente sintattico della funzione interpretazione (anzi, questo, lungi dal costituire una soluzione, costituirebbe soltanto un modo di rinviare il problema). Ed il fatto che un modello cognitivo sia in grado di associare un riferimento nel mondo esterno alle proprie rappresentazioni non dovrebbe costituire alcun problema rispetto ai risultati limitativi quali il teorema di Tarski, non fosse altro perché i dispositivi computazionali demandati a fissare il riferimento delle rappresentazioni non potrebbero essere altro, per loro stessa natura, che dispositivi di tipo parziale (ad esempio, nel senso che tali dispositivi sarebbero collegati soltanto ad alcuni dei simboli primitivi del sistema di rappresentazione interno - cioè a quelli che corrispondono a concetti di tipo "osservativo").

Vi è tuttavia un secondo aspetto dell'argomentazione di Thomason, che risulta molto più problematico, e che qui ci interessa in maniera diretta, in quanto si ricollega al problema della possibilità di ottenere un sistema di rappresentazioni interne sufficientemente espressivo da essere adeguato alla rappresentazione dei significati linguistici (o alle esigenze del ragionamento di senso comune), e che sia allo stesso tempo finitamente assiomaticizzabile. Nel par. 9.1 abbiamo visto che un risultato analogo al teorema di Tarski può essere esteso al trattamento al primo ordine mediante predicati delle modalità e di atteggiamenti proposizionali come la conoscenza. Montague (1963) ha dimostrato che una formalizzazione della conoscenza mediante predicati in un linguaggio del primo ordine, che comprenda l'equivalente dell'apparato deduttivo del sistema modale **T** (assioma della conoscenza, assioma distributivo e regola di necessitazione), e che comprenda inoltre gli assiomi dell'aritmetica di Robinson (e che quindi sia in grado di autoriferimento) conduce all'inconsistenza attraverso un meccanismo simile a quello del paradosso del mentitore nel teorema di Tarski. Sempre nel paragrafo 9.1, abbiamo visto che Thomason (1980) ha dimostrato che un risultato in parte analogo vale anche per le formalizzazioni al primo ordine della credenza. Vi sono quindi *paradossi epistemici* simili per struttura ai paradossi semantici. Inoltre, abbiamo constatato che le usuali assiomaticizzazioni modali al primo ordine delle logiche epistemiche (quelle cioè basate sull'uso operatori modali proposizionali anziché di predicati per esprimere i concetti epistemici come credere e sapere) sono esenti da problemi di inconsistenza in quanto il loro potere espressivo è più limitato di quello delle formalizzazioni al primo ordine basate sull'uso di predicati epistemici. Infatti, nelle assiomaticizzazioni al primo ordine di tipo modale non è possibile quantificare sugli oggetti degli atteggiamenti proposizionali. Il punto è che per poter quantificare al primo ordine sugli oggetti della credenza (e, in generale, di un

atteggiamento proposizionale) si deve poter disporre di termini che denotano gli oggetti della credenza stessa, di "nomi" degli oggetti dell'atteggiamento proposizionale. Inoltre, perché possa essere garantita la possibilità di effettuare le necessarie manipolazioni formali sulla struttura di tali termini, deve essere disponibile nel linguaggio un apparato di operatori sintattici opportuni. Ciò dà origine alla possibilità di costruire formule autoreferenziali, e quindi dà luogo ai paradossi. Si noti che non è necessario assumere che gli oggetti degli atteggiamenti proposizionali siano oggetti sintattici, non è necessario assumere cioè che i termini cui si applicano ad esempio predicati come *know* o *bel* denotino formule. Il paradosso insorgerebbe anche qualora si assumesse che tali termini denotano oggetti di tipo intensionale, ad esempio proposizioni. E' sufficiente che tali termini abbiano una struttura compositiva analoga a quella delle formule. Quindi, l'insorgere del paradosso non è legato direttamente agli approcci di tipo sintattico o citazionale agli atteggiamenti proposizionali, nel senso utilizzato nel capitolo 9 (cfr. par. 9.1)<sup>24</sup>.

Il problema è che, per fornire una semantica delle espressioni del linguaggio naturale, e, in generale, per le esigenze della formalizzazione del ragionamento di senso comune, è necessario disporre di un potere espressivo tale da consentire la quantificazione sugli oggetti degli atteggiamenti proposizionali<sup>25</sup>. Una maniera per ottenere ciò senza incorrere nei paradossi consiste nell'utilizzare una logica modale (che esprima cioè i costrutti epistemiche non come predicati, ma come operatori proposizionali), il cui linguaggio comprenda un insieme di variabili (di un tipo sintattico distinto da quello delle variabili individuali), i cui valori varino sull'insieme degli oggetti degli atteggiamenti proposizionali (ad esempio sull'insieme delle proposizioni). In questo modo non è necessario introdurre nel linguaggio termini che denotino gli oggetti degli atteggiamenti proposizionali, e meccanismi che consentano l'autoriferimento. Questa è la soluzione adottata nelle logiche intensionali, come ad esempio nella teoria di Montague. Tuttavia, il problema consiste nel fatto che così facendo si esce dalla logica del primo ordine, e si passa ad una logica di ordine superiore, con le note difficoltà che ne derivano. In particolare, si esclude in questo modo la possibilità di un'assiomatizzazione completa finita della logica<sup>26</sup>.

Dunque l'esigenza di una rappresentabilità finita dei significati linguistici incontra difficoltà che, dal punto di vista intuitivo, non erano facilmente prevedibili. Thomason (1979, 1986) interpreta tali difficoltà come un sintomo della difficoltà di conciliare gli obiettivi della semantica filosofica e del programma modellistico da un lato, con gli scopi della semantica cognitiva e dell'intelligenza artificiale dall'altro. In particolare, secondo Thomason, è problematico conciliare la tesi cognitivista dell'afferrabilità (*graspability*) dei significati da parte dei parlanti, e la concezione di significato di tipo modellistico<sup>27</sup>. Sembra impossibile interpretare le teorie semantiche alla Montague come modelli della competenza di un parlante idealizzato (ma, ovviamente, finito). Non solo, ma il concetto stesso di competenza idealizzata sembra problematico, in quanto sembra condurre a contraddizione. I risultati di Montague e di Thomason sui paradossi epistemiche si basano sull'assunto che valga l'onniscienza logica: essi sono ottenuti a partire dall'ipotesi di disporre degli equivalenti, formulati al primo ordine, dell'assioma distributivo e della regola di necessitazione, che (come abbiamo visto nel par. 8.2) sono i principali responsabili dell'onniscienza logica. Quindi, abbandonando l'onniscienza logica, verrebbe meno anche la possibilità di derivare il paradosso. Ora, è certamente vero che i soggetti epistemiche reali non sono logicamente onniscienti. Tuttavia, sembra difficile poter accettare che l'impossibilità di avere l'onniscienza logica derivi dalla teoria semantica stessa. Sembra difficile cioè accettare che sia una questione di logica, e non una semplice questione fattuale, che non si possa assumere che esistano soggetti epistemiche onniscienti. Lo stesso

---

<sup>24</sup>Sulle relazioni fra modelli sintattici degli atteggiamenti proposizionali e insorgere dei paradossi si veda Thomason (1977). Sarebbe possibile introdurre nel linguaggio termini che denotano proposizioni (e permettere quindi la quantificazione su di esse) evitando al tempo stesso l'insorgere di paradossi a patto di assumere che tali termini non abbiano struttura, che siano tutti, in un certo senso, atomici, e a patto di non poter esprimere nel linguaggio a quali formule le proposizioni denotate da tali termini sono associate. Ma, per ovvie ragioni, tale soluzione non può risultare soddisfacente. Su questi punti si veda ad esempio (Asher e Kamp 1986).

<sup>25</sup>Ad esempio per rappresentare il significato di un enunciato del tipo "C'è qualcosa di falso in quello che crede Giorgio". Si noti che in un enunciato come questo si ripropone il problema di rappresentare il predicato di verità all'interno del linguaggio. Tuttavia, il motivo per cui qui il problema si pone non è lo stesso che abbiamo criticato in precedenza. Qui si tratta del fatto che "vero" e "falso" sono parole della lingua di cui si intende fornire una semantica, e quindi devono essere rappresentate come un atteggiamento proposizionale fra gli altri.

<sup>26</sup>D'altro canto, Perlis (1988) dimostra che l'inconsistenza può insorgere anche in una logica modale basata su operatori proposizionali qualora si introducano opportuni operatori modali che consentano di ottenere certe forme di autoriferimento. Sulle condizioni a cui possono insorgere paradossi epistemiche anche nelle logiche modali si vedano anche (Koons 1988, 1991) e (Grim 1993).

<sup>27</sup>Thomason individua un'analogia fra questo problema e il problema di Benacerraf in filosofia della matematica (Benacerraf 1973). Il problema di Benacerraf può essere visto come il problema conciliare ontologia ed epistemologia della matematica. Come è possibile una "competenza matematica" da parte di un soggetto matematico finito (anche se idealizzato), se gli enti matematici, in molti casi, eccedono i limiti della ricorsività? (Ad esempio, i numeri reali generabili da una macchina di Turing non sono che un sottoinsieme proprio di tutti i numeri reali).

dicasi per le capacità aritmetiche di cui dispone il parlante ideale. Non sembra ammissibile che dalla teoria semantica segua che un parlante ideale deve sapere meno aritmetica di quella necessaria per derivare il paradosso.

Thomason (1979) sembra considerare i paradossi epistemici come una sorta di *reductio ad absurdum* delle ipotesi alla base del programma della semantica cognitiva (e delle scienze cognitive in generale)<sup>28</sup>. Quello che è certo è che questi risultati sembrano mettere in serie difficoltà il programma logicista in IA (almeno se lo si interpreta in maniera forte, intendendolo cioè come un programma di formalizzazione logica del ragionamento di senso comune in una prospettiva computazionale, e non riducendolo all'elaborazione di una serie di tecniche a carattere informatico). E, in ogni caso, sembrerebbe messa in discussione la plausibilità della posizione degli obiettivi comuni fra semantica modellistica e semantica cognitiva. Tuttavia, in anni recenti, sono stati sviluppati diversi tentativi per superare i problemi posti dall'assiomatizzazione delle modalità e delle logiche epistemiche. Molti di essi si basano ad esempio sulla teoria della verità proposta da Kripke (1975). L'idea centrale di Kripke è quella di evitare l'insorgere dei paradossi semantici rinunciando al principio classico di bivalenza, in base al quale ad ogni enunciato deve essere assegnato uno ed un solo valore di verità, il vero o il falso. In base alla teoria di Kripke, agli enunciati "problematici" come ad esempio le varie forme del paradosso del mentitore, o come il cosiddetto enunciato del *truth teller* (cioè un enunciato del tipo "questo enunciato è vero"), non viene assegnato alcun valore di verità. Questo evita l'insorgere delle inconsistenze. A partire dai lavori di Kripke, e da altri lavori ad esso collegati, come ad esempio la teoria di Gupta (1982), sono state sviluppate diverse proposte per ottenere formalizzazioni al primo ordine degli atteggiamenti proposizionali che superassero i problemi sopra descritti. Un esempio è costituito dai lavori di Perlis (1985, 1988). Le strategie utilizzate sono di diversi tipi: indebolire la logica utilizzata adottando qualche forma di logica non classica, limitare le forme di autoriferimento consentite, e così via<sup>29</sup>. Un'analisi approfondita di questi temi non è possibile in questa sede. In generale, comunque, anche se nessuna soluzione definitiva è stata individuata, si tratta tuttavia di un ambito di ricerca ricco e vitale, da cui potrà emergere qualche chiarificazione su questi temi cruciali dal punto di vista delle basi teoriche della linea logicista in IA e nelle scienze cognitive. Rispetto a questo problema, lo stesso Thomason (1986) sembra mostrarsi meno drastico rispetto a quanto scriveva nel 1979. Egli afferma ad esempio:

Penso che la reazione corretta sia che la metafora del parlante ideale debba essere riconsiderata quando si assume che la semantica sia una parte della competenza del parlante ideale. Mi sembra che le relazioni fra teorie del linguaggio e parlanti siano molto più problematiche di quanto la metafora suggerisca. In ogni caso, credo che se le conseguenze di aderire all'obbiettivo dell'afferrabilità [del significato] verranno elaborate con attenzione, condurranno a forme forti di costruttivismo, o ad approcci paraconsistenti, o ad altri approcci che comportano una revisione dei fondamenti logici. Ogni valutazione delle forme di cognitivismo forte dovrebbe tenere conto di queste conseguenze. [...] Possiamo aspettarci [...] che verranno sviluppate varie soluzioni tecniche, e che queste ci aiuteranno a chiarire le migliori alternative. (Thomason 1986, pp. 234-5).

## 12.5 Significato, competenza, comprensione

A prescindere dai problemi evidenziati nel paragrafo precedente, esistono conflitti e tensioni fra le varie linee della ricerca logica in IA. In particolare, alcune ricerche sembrano ignorare le esigenze di finitezza delle rappresentazioni e di trattabilità computazionale che abbiamo individuato come una delle caratteristiche centrali dell'uso della logica in questa disciplina. Il caso più appariscente è probabilmente quello del ragionamento non monotono. Nel paragrafo 6.1 abbiamo visto come la formalizzazione delle varie forme di *circumscription* richieda l'introduzione di assiomi del secondo ordine, con la conseguenza di escludere la possibilità di una assiomatizzazione completa con strumenti finiti. La logica preferenziale di Shoham (par. 6.3) è stata sviluppata esclusivamente su base semantica, tralasciando gli aspetti di teoria della dimostrazione. Connesso alle difficoltà di assiomatizzazione del ragionamento non monotono, vi è un problema messo originariamente in evidenza da David Israel (1980) in relazione alla *default logic*. E' noto dai risultati di Church che l'insieme dei teoremi di una teoria logica del primo ordine non è in generale decidibile. Tuttavia tale insieme è ricorsivamente enumerabile: si può definire un algoritmo che generi tutti i teoremi della teoria semplicemente definendo un ordine opportuno con cui, a partire dagli assiomi, si applicano ricorsivamente le regole di inferenza ai teoremi che via via vengono ottenuti. Se si estende una teoria del primo ordine con regole di *default*, questa caratteristica va persa: in generale, l'insieme delle conseguenze non monotone di una teoria di *default* non è neppure ricorsivamente enumerabile. Il problema risiede nel fatto che ciò che fa sì che l'insieme dei teoremi di una teoria del primo ordine sia semidecidibile dipende dal fatto che le regole di inferenza sono applicabili in modo effettivo. Al contrario, le regole di *default* della *default logic* non godono di questa proprietà. Si consideri una teoria di *default T* che comprenda la seguente regola:

<sup>28</sup> Anche se, in sede conclusiva, afferma che gli psicologi cognitivi non dovrebbero preoccuparsi più di tanto di questi problemi a carattere filosofico e fondazionale, e proseguire nel loro programma di ricerca.

<sup>29</sup> Oltre ai lavori di Perlis, su questi problemi si vedano anche (Asher e Kamp 1986) e (Asher 1988, 1990). Una rassegna si trova in (Thomason 1986).

$$\frac{\alpha : \beta}{\gamma}$$

Supponendo di disporre di  $\alpha$ , per derivare  $\gamma$  bisogna stabilire se è consistente credere  $\beta$ , se cioè  $\beta$ , unito alla teoria  $T$ , dia luogo o meno a una contraddizione. Ma, in generale, la consistenza di una teoria del primo ordine non è decidibile mediante un algoritmo, e quindi, di conseguenza, la regola in questione non è applicabile in modo effettivo. Analoghe considerazioni valgono per le logiche non monotone di tipo modale, come la logica autoepistemica. Dal punto di vista delle proprietà computazionali sembra quindi che le logiche non monotone dell'IA siano meno attraenti delle usuali teorie logiche del primo ordine. E sembra che vi sia una tensione fra le tecniche utilizzate per modellare soggetti inferenziali limitati dal punto di vista dell'informazione disponibile (e che quindi devono trarre conseguenze rivedibili disponendo di informazioni incomplete), e soggetti inferenziali limitati dal punto di vista delle risorse computazionali<sup>30</sup>. Più in generale, si ha talvolta l'impressione di uno scollamento fra gli obiettivi computazionali della disciplina e alcune delle formalizzazioni elaborate.

Il punto è che i modelli proposti e le ricerche sviluppate nell'ambito dell'IA logicista si collocano spesso a livelli di analisi differenti. Dal nostro punto di vista, nella prospettiva cioè dei possibili apporti della ricerca logica in IA per una teoria del significato, e delle relazioni che sussistono fra questi due ambiti, la questione può essere posta nei termini seguenti. Il problema di una teoria del significato e delle sue relazioni con il parlante può essere posto a livelli diversi. Schematicamente, possono essere individuati i tre livelli seguenti<sup>31</sup>.

a) Al livello più astratto, una teoria del significato può essere concepita come una teoria delle relazioni fra espressioni del linguaggio e mondo (o meglio, fra espressioni del linguaggio e modelli formali della realtà), prescindendo completamente da come i significati possano essere compresi ed elaborati da coloro che utilizzano quel linguaggio. Dal punto di vista logico, questo livello di analisi corrisponde a un trattamento puramente semantico-modellistico di un sistema di rappresentazioni, che trascura gli aspetti legati alla teoria della dimostrazione. È il punto di vista classico della semantica model teoretica, di cui l'esempio più autorevole è costituito dalla teoria di Montague. A questo livello, l'oggetto principale di interesse è costituito dalle condizioni di verità degli enunciati e dai nessi di conseguenza logica che li concernono.

b) Il secondo livello è costituito da una teoria della *competenza* semantica, nel senso di una teoria della rappresentazione dei significati che possa essere accessibile a un parlante finito (anche se idealizzato). In termini logici, una teoria della competenza intesa in questo senso si colloca a livello di teoria della dimostrazione: comporta cioè l'individuazione di assiomatizzazioni finite che siano in grado di rendere conto dei nessi sintattici che corrispondono ai nessi semantici di conseguenza logica rilevanti per il punto a). Il fatto di porsi a livello di teoria della competenza non comporta quindi che siano irrilevanti gli aspetti semantico-modellistici dei sistemi di rappresentazione utilizzati. Di norma, comunque, una teoria della competenza comporta ulteriori vincoli sulle strutture modellistiche utilizzate. Ne costituisce un esempio il problema dei paradossi epistemici trattato nel paragrafo precedente. Una teoria della comprensione non concerne tuttavia il problema della capacità di un parlante di decidere della verità di un enunciato date certe premesse, o di individuare i nessi di conseguenza logica fra enunciati. L'indecidibilità del calcolo dei predicati del primo ordine renderebbe impossibile una nozione di competenza così intesa. Una teoria della competenza ha a che fare con ciò che un parlante (idealizzato) conosce, mantenendo ferma però la distinzione fra ciò che il parlante sa e le proprietà che sono determinate da quello che il parlante sa. Ad esempio, chi conosca una assiomatizzazione del calcolo dei predicati del primo ordine "conosce" in un certo senso la logica del primo ordine (in quanto ne possiede una rappresentazione finita). Ma vi sono proprietà che sono determinate da ciò che egli conosce (ad esempio l'estensione dell'insieme delle formula valide, oppure, date due formule qualunque, il fatto che una implichi logicamente l'altra o meno) che egli di fatto non può conoscere perché indecidibili (Partee 1982). In questo senso la nozione di competenza è indipendente dalla nozione di processo di giudizio o di comprensione<sup>32</sup>.

<sup>30</sup>Queste osservazioni valgono in generale. Vi sono tuttavia forme di ragionamento non monotone che non danno problemi computazionali. Abbiamo visto ad esempio che il PROLOG incorpora una forma di *closed world assumption*. Sono inoltre in corso ricerche per ricondurre i formalismi non monotoni a sistemi computazionalmente più attraenti. Abbiamo già detto (par. 6.1) delle riformulazioni di varie forme di *circumscription* in logica del primo ordine. Sui sistemi di deduzione automatica basati su logiche non monotone si veda ad esempio il capitolo 7 di (Brewka 1991). Vi è chi, come Selman e Kautz (Selman e Kautz 1988; Kautz e Selman 1991), ha studiato dal punto di vista della trattabilità computazionale forme limitate di formalismi non monotoni. Per una rassegna dei risultati di complessità computazionale relativi ai sistemi di logica non monotona si veda (Cadoli e Schaerf 1993).

<sup>31</sup>Questa classificazione, soprattutto nella distinzione fra i punti b) e c), si basa su (Partee 1982).

<sup>32</sup>Anche la nozione di semantica procedurale, secondo Barbara Partee, va oltre i limiti di una teoria della competenza semantica, nella misura in cui verte sui processi utilizzati nella produzione di giudizi validi da parte dei parlanti.

c) Il terzo livello, infine, concerne una teoria della *comprensione (understanding)*, intesa come una teoria che abbia per oggetto i giudizi di validità e il riconoscimento dei nessi di implicazione logica effettuati dal parlante, e che abbia come scopo quello di modellare procedure effettive che possano rendere conto di tali giudizi.

Generalizzando questa classificazione dall'ambito della semantica del linguaggio naturale ai temi della rappresentazione della conoscenza e del ragionamento, le ricerche svolte nell'ambito dell'IA logicista si possono collocare, di volta in volta, a livelli differenti di tale gerarchia, e da ciò dipendono probabilmente i conflitti evidenziati all'inizio del paragrafo. Ad esempio, le teorie circoscrivite nella loro formulazione originaria, o la *preference logic* di Shoham si collocano in una prospettiva molto vicina a quella del punto a). Per quanto riguarda il punto b), si collocano a questo livello quelle ricerche in cui si pone l'enfasi sulla possibilità di assiomatizzare in maniera finita i formalismi proposti. Un esempio è costituito dalle ricerche volte ad individuare forme di *circumscription* assiomatizzabili al primo ordine. Si pongono su questa linea anche le ricerche menzionate nel paragrafo precedente, volte ad individuare formalizzazioni al primo ordine degli atteggiamenti proposizionali che superino i problemi individuati da Montague e Thomason, come ad esempio i lavori di Perlis (1985, 1988). Sono assimilabili al punto c), ad esempio, le ricerche di Levesque e della sua scuola, che pongono l'enfasi sugli aspetti legati alle proprietà computazionali dei formalismi di rappresentazione, e che mirano ad individuare formalismi, o sottoinsiemi di formalismi, che siano trattabili dal punto di vista computazionale. In questo senso, non sono probabilmente casuali le analogie individuate da Levesque fra basi di conoscenza vivide ed il concetto di modello mentale, nato nell'ambito della semantica procedurale. In questa prospettiva si potrebbe vedere forse anche gran parte del lavoro svolto nell'ambito della dimostrazione automatica di teoremi<sup>33</sup>. Una linea di ricerca che concerne direttamente quella che abbiamo chiamato la teoria della comprensione, e che si pone esplicitamente in alternativa con le proposte di Levesque, è quella individuata da Jon Doyle, secondo il quale lo strumento idoneo a studiare e modellare le inferenze effettivamente eseguite da un soggetto va individuato nelle teorie della razionalità economica e della scelta razionale<sup>34</sup>. Doyle e Patil (1991) criticano le posizioni di Levesque, tese a ridurre il potere espressivo dei linguaggi di rappresentazione in cambio della trattabilità computazionale. Questo atteggiamento comporta in molti casi un impoverimento espressivo inaccettabile, tale da distruggere di fatto la generalità del sistema di rappresentazione. Doyle e Patil esemplificano le loro affermazioni sui linguaggi terminologici derivati da KL-ONE, mostrando come un gran numero di costrutti che risultano fondamentali nella descrizione di termini siano stati espunti da questi formalismi. La soluzione corretta, secondo Doyle, non risiede nel ridurre il potere espressivo dei formalismi di rappresentazione, ma di tenere conto degli scopi delle attività inferenziali di un soggetto razionale, e questo può essere ottenuto appunto mediante gli strumenti delle teorie della scelta razionale. In altri termini, per gli obiettivi di quella che abbiamo chiamato una teoria della comprensione gli strumenti logicisti non sarebbero di per sé sufficienti: "In logica, ogni insieme consistente di credenze è buono quanto un altro, e l'unica guida che il logicismo sembra offrire a un ragionatore è la regola *se è corretto, fallo!*" (Doyle 1990a, p. 1095). Gli strumenti logici andrebbero integrati con gli strumenti della razionalità economica. Lo scopo delle teorie della razionalità economica è quello di formalizzare i processi di decisione e di scelta fra diverse alternative. In questo caso, si tratta di formalizzare le scelte fra percorsi inferenziali diversi. In questo senso, le teorie della scelta razionale dovrebbero fornire un sostituto formale alle tecniche euristiche di tipo informale usate in IA. In questa prospettiva, "le nozioni di razionalità logica ed economica soddisfano esigenze complementari. La razionalità logica serve a descrivere le possibilità del ragionamento e dell'azione, mentre la razionalità economica serve a prescrivere scelte fra di esse" (Doyle 1990a, p. 1096).

L'individuazione di questi livelli di analisi costituisce una sorta di "ricostruzione razionale" della ricerca attualmente svolta nell'IA di impostazione logica, nel senso che tali livelli appaiono spesso mescolati e sovrapposti nella pratica effettiva, e non sempre risulta facile (o lecito) distinguerli. Vi sono spesso tensioni fra i livelli, e non è facile comprendere come le esigenze che vengono poste a livelli differenti possano convivere. Ne è un esempio il problema degli atteggiamenti proposizionali e dai paradossi epistemici.

Tornando ai problemi posti dalla ricerca in teoria del significato, è evidente che ciascuno di questi livelli di indagine è pienamente legittimo, e risponde ad esigenze ed obiettivi in parte diversi. E' altrettanto evidente tuttavia che il vantaggio che deriva da un approccio come quello suggerito dall'IA di impostazione logica risiede nel fatto di affrontare tutti questi livelli all'interno di una prospettiva teorica comune, di un paradigma unitario caratterizzato dalla centralità dell'utilizzo degli strumenti della logica. In questo senso, non si avrebbe necessariamente a che fare con una posizione degli scopi comuni: gli scopi potrebbero restare, in linea di principio, distinti (anche se, come vedremo, con dei punti di contatto). Non si tratta però neppure di una posizione separatista (nel senso che i vari ambiti siano destinati a restare inconfondibili, senza alcuna possibilità di comunicazione reciproca). Si tratta piuttosto di una prospettiva che postula un quadro teorico generale comune, una comunanza di strumenti e di tecniche, e che quindi rende possibile l'utilizzo reciproco dei risultati.

E' evidente l'utilità di una comunicazione "dall'alto verso il basso" fra i livelli, in maniera tale cioè che ad ogni livello si possa tenere conto di ciò che è stato fatto ai livelli superiori, e usufruire dei risultati ottenuti. Ad esempio, una teoria della comprensione può essere sviluppata a partire dai modelli sviluppati a livello di teoria della competenza. Vi è inoltre almeno un caso in cui sembra indispensabile effettuare il cammino inverso, in cui cioè i livelli più alti sembrano

<sup>33</sup>Si veda ad esempio (Bundy 1983).

<sup>34</sup>Sulle posizioni di Doyle a questo proposito si vedano ad esempio (Doyle 1983, 1988a, 1988b, 1990a, 1990b).

dover tenere conto delle analisi sviluppate ai livelli inferiori. Si tratti degli enunciati di *atteggiamento proposizionale* e del problema dell'onniscienza logica. Anche una spiegazione del significato che si collochi al livello di astrazione maggiore, che sia interessata cioè esclusivamente alle condizioni di verità degli enunciati, prescindendo da ogni riferimento alla competenza o ai processi di comprensione del parlante, deve coinvolgere anche i livelli più bassi di analisi qualora si debba rendere conto, ad esempio, di enunciati di tipo epistemico. Come ha sottolineato Barbara Partee (1982) (cfr. par. 2.1), una teoria semantica che accetti l'assunzione dell'onniscienza logica è inadeguata anche come teoria del linguaggio in quanto "parlato da Dio", perché obbliga ad assumere che *tutti* i soggetti di credenza del dominio siano a loro volta ragionatori perfetti. Una teoria semantica che voglia rendere conto dei valori di verità e dei nessi di implicazioni fra enunciati di atteggiamento proposizionale deve quindi incorporare, o far riferimento a, una teoria della comprensione relativa ai soggetti epistemici. Ad esempio, le condizioni di verità di un enunciato come "Tizio crede che  $\alpha$ " devono tenere conto necessariamente della capacità di Tizio di rappresentarsi  $\alpha$  e di giudicarlo vero. Inoltre, se ad esempio  $\alpha$  implica  $\beta$ , per sapere se da "Tizio crede che  $\alpha$ " segue che "Tizio crede che  $\beta$ " si deve sapere se Tizio è in grado di riconoscere la validità di  $\alpha \rightarrow \beta$ . Il fatto che, nell'IA di tradizione logica, i problemi che si collocano al livello di teoria della comprensione vengano affrontati con strumenti "omogenei" a quelli dei livelli superiori (nel senso che si tratta di strumenti di tipo logico, o di estensioni di strumenti di tipo logico), può far ritenere che questo paradigma sia particolarmente adatto ad affrontare il problema. Tutti i modelli logici della credenza che si basano su una distinzione fra conoscenza implicita e conoscenza esplicita (capp. 10 e 11) possono essere visti come tentativi di introdurre aspetti diversi di una teoria della comprensione nella semantica dei formalismi epistemici proposti. In generale, la distinzione fra conoscenza implicita e conoscenza esplicita potrebbe essere caratterizzata definendo la conoscenza esplicita come conoscenza implicita *più* una teoria della comprensione. I tentativi compiuti sino ad ora sono per molti aspetti ancora rudimentali. Ci sembra tuttavia che una prospettiva in senso lato computazionale, come appunto quella dell'IA logicista, costituisca una via proficua per affrontare il problema.