

Linee di Ricerca

Marcello Frixione

LOGICA E INTELLIGENZA ARTIFICIALE

Versione 1.0 - 2004



SWIF - Sito Web Italiano per la Filosofia
Rivista elettronica di filosofia - Registrazione n. ISSN 1126-4780

Linee di Ricerca – SWIF

Coordinamento Editoriale: Gian Maria Greco

Supervisione Tecnica: Fabrizio Martina

Supervisione: Luciano Floridi

Redazione: Eva Franchino, Federica Scali.

AUTORE

Marcello Frixione [frix@dist.unige.it] è professore associato di Filosofia del linguaggio presso il Dipartimento di Scienze della Comunicazione dell'Università di Salerno, dove insegna anche logica. Le sue ricerche si collocano nel settore delle scienze cognitive. Tra le altre cose, ha studiato vari aspetti teorici dei modelli computazionali della mente, si è occupato dei rapporti tra teorie logico-formali del significato e formalismi di rappresentazione della conoscenza in intelligenza artificiale, ha collaborato allo sviluppo di modelli computazionali di visione artificiale e di rappresentazione delle azioni. Ha pubblicato numerosi articoli su riviste nazionali e internazionali, e il volume *Logica, significato e intelligenza artificiale* (Franco Angeli, Milano, 1994). Assieme a Dario Palladino ha scritto un'introduzione alla teoria della computabilità per i tipi di Carocci.

La revisione editoriale di questo saggio è a cura di Federica Scali.

LdR è un e-book, inteso come numero speciale della rivista SWIF. È edito da Luciano Floridi con il coordinamento editoriale di Gian Maria Greco e la supervisione tecnica di Fabrizio Martina.

LdR - Linee di Ricerca è il servizio di Bibliotec@SWIF finalizzato all'aggiornamento filosofico. LdR è un e-book in progress, in cui ciascun testo è un capitolo autonomo. In esso l'autore o l'autrice, presupponendo solo un minimo di conoscenze di base, fornisce una visione panoramica e critica dei temi principali, dei problemi più importanti, delle teorie più significative e degli autori più influenti, nell'ambito di una specifica area di ricerca della filosofia contemporanea attualmente in discussione e di notevole importanza. Il fine è quello di fornire al pubblico italiano un'idea generale su quali sono gli argomenti di ricerca di maggior interesse nei vari settori della filosofia contemporanea oggi, con uno stile non-storico, accessibile ad un pubblico di filosofi non esperti nello specifico settore ma interessati ad essere aggiornati.

Tutti i testi di Linee di Ricerca sono di proprietà dei rispettivi autori. È consentita la copia per uso esclusivamente personale. Sono consentite, inoltre, le citazioni a titolo di cronaca, studio, critica o recensione, purché accompagnate dall'idoneo riferimento bibliografico. Per ogni ulteriore uso del materiale presente nel sito, è fatto divieto l'utilizzo senza il permesso del/degli autore/i.

Per quanto non incluso nel testo qui sopra, si rimanda alle più estese norme sui diritti d'autore presenti sul sito Bibliotec@SIWF, www.swif.it/biblioteca/info_copy.php.

Per citare un testo di Linee di Ricerca si consiglia di utilizzare la seguente notazione:

AUTORE, *Titolo*, in L. Floridi (a cura di), *Linee di Ricerca*, SWIF, 2003, ISSN 1126-4780, p. X, www.swif.it/biblioteca/lr.

SWIF – LINEE DI RICERCA

LOGICA E INTELLIGENZA ARTIFICIALE¹

MARCELLO FRIXIONE

Versione 1.0

1. LOGICA E RAPPRESENTAZIONE DELLA CONOSCENZA

I rapporti tra logica e intelligenza artificiale (d'ora in poi IA) sono complessi ed intricati. Ovviamente non intendo affrontarne qui tutti gli aspetti. Mi limiterò a delineare alcune importanti (a mio avviso) linee di influenza dell'IA sulla logica. A tal fine prenderò le mosse dai dibattiti sull'opportunità o meno di impiegare la logica in IA che hanno caratterizzato la ricerca sulla rappresentazione della conoscenza soprattutto nel corso degli anni '70.

Un assunto generalmente condiviso in IA (almeno nella sua forma “classica”) è che una simulazione computazionale dei comportamenti intelligenti presupponga la manipolazione di rappresentazioni simboliche sulla base di regole di tipo formale. Il settore della disciplina che si prefigge di individuare e studiare sistemi di rappresentazione con queste caratteristiche va sotto il nome di *rappresentazione delle conoscenze (knowledge representation)*.

In generale un sistema di rappresentazione della conoscenza come viene concepito in IA comprende un insieme di simboli primitivi, o atomici, a partire dai quali, mediante opportune regole di composizione, si costruisce un insieme di

¹ Ringrazio Roberto Cordeschi, Luciano Floridi e Dario Palladino per aver letto, discusso e commentato versioni precedenti di questo lavoro.

simboli complessi. Vi sono poi procedure e regole computazionali che consentono di manipolare i simboli in modo da ricavarne le conseguenze volute.

In quanto *rappresentazioni*, a tali strutture deve essere associato un significato. Ciò però non può avvenire in maniera arbitraria: deve essere possibile determinare il significato di ogni simbolo complesso a partire dalla sua struttura sintattica e dal significato dei simboli che lo compongono. In altre parole, i simboli devono avere una *semantica compositiva*.

Le procedure e le regole computazionali che manipolano i simboli devono operare in accordo con il significato di questi ultimi. Tali regole e procedure devono però essere puramente formali, le loro operazioni devono cioè dipendere esclusivamente dalla forma sintattica delle rappresentazioni, senza alcun tipo di accesso al loro significato.

Gli obiettivi della rappresentazione della conoscenza in IA classica sono ben sintetizzati dalla cosiddetta *knowledge representation hypothesis* formulata da Brian Smith [1982]. Secondo Smith, ogni “sistema intelligente” deve incorporare un insieme di strutture di tipo, in senso lato, linguistico, tali che:

1. queste strutture, viste da un osservatore esterno al sistema, possono essere interpretate come la rappresentazione della conoscenza di cui il sistema dispone;
2. indipendentemente da tale “attribuzione semantica” data dall'esterno, tali strutture devono essere manipolabili formalmente, in modo da giocare un ruolo causale nel determinare il comportamento del sistema.

I sistemi elaborati dai logici godono esattamente di queste proprietà. È quindi naturale che si sia pensato di utilizzarli come sistemi per la rappresentazione della conoscenza. Tuttavia l'uso della logica in IA pone vari problemi. In passato su questo tema si sono scatenate lunghe e accese dispute, che hanno caratterizzato il dibattito teorico di questa disciplina soprattutto nel corso degli anni '70.

John McCarthy è stato uno dei più autorevoli sostenitori dell'uso della logica in IA. Già fin dagli anni cinquanta aveva sostenuto che la logica avrebbe dovuto svolgere nei confronti dell'informatica, e in particolare dell'IA, la funzione che l'analisi infinitesimale svolge nei confronti della fisica². Queste posizioni non raccolsero tuttavia il consenso dell'intera comunità dei ricercatori. Si affermarono due punti di vista, e la ricerca si sviluppò lungo linee divergenti. McCarthy e Marvin Minsky possono essere considerati i sostenitori più rappresentativi delle due opposte tendenze³.

McCarthy ritiene che lo scopo dell'IA sia progettare programmi per calcolatore che funzionino in accordo coi dettami della logica matematica, a prescindere dal fatto che questo sia o meno il modo in cui ragionano gli esseri

² Si tenga presente che l'uso della logica in IA costituisce solo un aspetto del problema più ampio dell'uso della logica in informatica. Vi sono applicazioni della logica all'informatica teorica che presentano ampie sovrapposizioni con temi di IA, come la programmazione logica o l'uso della logica per le basi di dati. In altri casi i legami con l'IA sono meno forti. Un esempio di questo secondo tipo è costituito dall'impiego di logiche costruttive (come la logica intuizionista) per la generazione automatica di programmi: a differenza di quanto avviene in logica classica, la dimostrazione costruttiva dell'esistenza di un certo oggetto richiede che venga esibita una costruzione effettiva dell'oggetto stesso, per cui la dimostrazione può essere usata come punto di partenza per la sintesi di un programma. Questi aspetti tuttavia non hanno avuto una grande influenza sulle ricerche di IA.

³ John McCarthy e Marvin Minsky sono unanimemente considerati due tra i padri fondatori dell'IA. McCarthy fornì vari contributi anche ad altri settori dell'informatica (ad esempio alla teoria dei sistemi operativi). Tra le altre cose, sviluppò il LISP, uno dei linguaggi di programmazione più antichi, e a tutt'oggi uno dei principali linguaggi impiegati in IA. Riporto le URL delle *home page* di McCarthy (<http://www-formal.stanford.edu/jmc/index.html>) e di Minsky (<http://web.media.mit.edu/~minsky/>). Da entrambe si possono scaricare vari articoli, tra cui numerosi testi classici del settore.

umani. Minsky pensa che l'approccio migliore sia fare sì che i calcolatori imitino il modo in cui funziona la mente umana, che, ritiene Minsky, è certamente diverso dal modo in cui funziona la logica matematica.

Questa contrapposizione ha avuto luogo in vari settori dell'IA, ad esempio nella dimostrazione automatica di teoremi. Tuttavia è in rappresentazione della conoscenza che questo tema ha ricevuto spazio maggiore. Nei prossimi paragrafi riassumerò per sommi capi i termini del problema.

2. CONTRO LA LOGICA

Il versante “anti-logico” in rappresentazione della conoscenza è identificabile con i sostenitori di sistemi alternativi alla logica, quali i *frame*, le reti semantiche o i sistemi a regole di produzione. Gli avversari della logica mettono in evidenza le differenze fra logica matematica e ragionamento del senso comune (o *commonsense reasoning*): la logica è stata sviluppata per fare ciò che il senso comune non è in grado di fare in quanto vago e contraddittorio, ma se vogliamo che un calcolatore “pensi” come un essere umano, dobbiamo rivolgerci a quelle caratteristiche tipiche del ragionamento del senso comune che non sono state prese in considerazione, o che sono state deliberatamente rifiutate dalla logica⁴.

Lo scopo della logica consiste non nel modellare il modo in cui le persone pensano, ma nel costruire modelli astratti dei procedimenti inferenziali al fine di giustificarne la validità. I logici si prefiggono di elaborare modelli di inferenza che non soffrano delle lacune e dei limiti del ragionamento del senso comune. Nei

⁴ In filosofia dibattiti analoghi hanno avuto come oggetto la possibilità di impiegare la logica per la semantica delle lingue naturali. Si ricordino ad esempio le critiche di Strawson [1950] all'analisi logica delle descrizioni definite di Russell.

confronti dei sistemi inferenziali da loro studiati, sono interessati soprattutto a proprietà metateoriche come la correttezza e la completezza. Secondo Minsky però l'applicazione di regole di inferenza corrette e complete non è di importanza fondamentale nel ragionamento del senso comune. Il ragionamento del senso comune può condurci da premesse vere a conclusioni false, oppure consentirci di ottenere conclusioni vere impiegando regole tutt'altro che corrette. Vi è una profonda differenza tra le *inferenze* intese come processi psicologici, e il concetto logico-formale di *deduzione*. Le inferenze spesso sono tutt'altro che deduzioni *valide* dal punto di vista logico. Un frammento di informazione ottenuto mediante inferenza da premesse vere non è mai da intendersi come vero in modo certo, ma solo plausibile ad un certo grado, e sempre in linea di principio rivedibile.

Una delle esposizioni più articolate di quelle che, secondo i suoi avversari, sono le debolezze della logica in rappresentazione della conoscenza è l'appendice all'articolo di Minsky sui *frame* (Minsky [1975]). Egli trae le seguenti conclusioni:

1. Il ragionamento “logico” non è abbastanza flessibile da servire come base per il pensiero: io preferisco considerarlo come un gruppo di metodi euristici, che diventano efficaci solo quando vengono applicati a piani schematici rigidamente semplificati. Altrimenti, la coerenza che la logica richiede non è di solito disponibile - *e probabilmente neppure desiderabile* - perché i sistemi coerenti probabilmente sono troppo deboli.
2. Io dubito della possibilità di rappresentare effettivamente la conoscenza ordinaria nella forma di molte brevi proposizioni, vere indipendentemente l'una dall'altra.

3. La strategia di separare completamente la conoscenza specifica dalle regole generali di inferenza è troppo radicale. Abbiamo bisogno di una maniera più diretta per collegare frammenti di conoscenza alle intuizioni su *come* usarla.
4. È stato creduto a lungo che fosse essenziale rendere tutta la conoscenza accessibile alla deduzione nella forma di asserzioni dichiarative, ma questo sarebbe molto meno urgente nella misura in cui imparassimo a manipolare descrizioni strutturali e procedurali. (Minsky [1975], p. 262 di Brachman e Levesque [1985])

È fuori di dubbio che queste osservazioni sono pertinenti e centrate, e che coinvolgono anche quelle impostazioni che non mirano all'adeguatezza psicologica dei modelli computazionali. Tuttavia, come vedremo, esse non comportano necessariamente l'abbandono della logica. Esaminiamo più in dettaglio alcuni punti specifici delle critiche anti-logiche.

2.1. Concetti e prototipi

In logica i concetti devono essere definiti in termini di condizioni necessarie e sufficienti. Tuttavia nel caso dei concetti del ragionamento del senso comune le condizioni necessarie e sufficienti sono l'eccezione piuttosto che la regola. La maggior parte dei concetti ordinari sono caratterizzati in termini prototipici: i concetti non sono definiti mediante insiemi di tratti che devono valere per tutte le istanze, ma sono caratterizzati nei termini di somiglianza con istanze tipiche del concetto, i *prototipi* appunto. Gli insiemi *basati su esempi* sembrano essere una caratteristica tipica del pensiero umano, anche dal punto di vista psicologico. Le proprietà del prototipo valgono nei casi tipici, ma non devono necessariamente

valere per *tutte* le istanze del concetto. Ma questo crea un problema per il trattamento delle eccezioni. Dice Minsky:

Si consideri un fatto come “gli uccelli sono in grado di volare”. Se pensate che il ragionamento del senso comune [*commonsense reasoning*] sia simile al ragionamento logico, allora dovete credere che vi siano principi generali che affermano “se Joe è un uccello e gli uccelli sono in grado di volare, allora Joe è in grado di volare”. Ma supponete che Joe sia uno struzzo o un pinguino. Bene, possiamo assiomatizzare ciò e dire che se Joe è un uccello, e Joe non è uno struzzo o un pinguino, allora Joe è in grado di volare. Ma se Joe è morto? O se ha i piedi nel calcestruzzo? Il problema con la logica è che una volta che voi avete dedotto qualcosa, non potete più sbarazzarvene.

Vi sono cioè dei valori di *default*, vale a dire valori che vengono attribuiti alle istanze di un concetto in assenza di altre informazioni esplicite. Essi rappresentano le proprietà del prototipo e delle istanze tipiche di un concetto; è tuttavia possibile accettare istanze del concetto che hanno proprietà differenti. Ad esempio, gli elefanti tipici sono grigi ed hanno quattro zampe: queste sono proprietà del prototipo di elefante, assunte come vere per *default* per tutti gli elefanti “normali”. Vi possono essere tuttavia istanze del concetto “elefante” che violano tali attributi (ad es. elefanti bianchi o elefanti con tre zampe). Questo aspetto rimanda al problema del trattamento delle eccezioni e alla non-monotonia del ragionamento del senso comune.

Tradizionalmente la logica è *monotona*: se un enunciato segue logicamente da un insieme di premesse, allora, a maggior ragione, esso deve seguire da un insieme di premesse più vasto che includa le precedenti (si veda a questo proposito

Antonelli [2003]). I concetti caratterizzati in termini di prototipi conducono invece a inferenze non monotone: dalla premessa che Joe è un uccello si trae provvisoriamente la conclusione che Joe è in grado di volare; ma se si aggiunge l'ulteriore premessa che Joe è un pinguino allora si deve rivedere la conclusione tratta in precedenza.

2.2. Inferenze a partire da informazioni incomplete

La non monotonia del ragionamento del senso comune è un problema molto generale, che va al di là della caratterizzazione dei concetti in termini di prototipi. Gli esseri umani sono agenti cognitivi limitati, che non possono accedere a tutte le informazioni rilevanti per i problemi che devono risolvere. Per questa ragione devono continuamente trarre conclusioni sulla base di informazioni incomplete, ed essere disposti a rivedere tali conclusioni qualora esse non siano più in accordo con le nuove conoscenze disponibili.

Ciò accade ad esempio nel ragionamento su azioni ed eventi: sappiamo che, se si agisce in un certo modo, allora, *salvo imprevisti*, si verificherà un certo evento *e*. Ma potrebbe sempre accadere un imprevisto che fa sì che l'evento *e* non si verifichi. Si consideri un tavolo da biliardo su cui si trovano una boccia rossa e una boccia bianca. Basta un po' di fisica e un po' di geometria per formulare generalizzazioni del tipo: *se si colpisce la boccia rossa con una certa forza e una certa inclinazione, allora la boccia rossa colpirà quella bianca e la manderà in buca*. Ma generalizzazioni di questo tipo ammettono sempre un numero imprecisabile di eccezioni (un po' come *Gli uccelli sono in grado di volare*). È facilissimo immaginare situazioni, dalle più plausibili alle più fantasiose, in cui la

boccia rossa è stata colpita con quella forza e quella inclinazione, ma la boccia bianca non finisce in buca. Ad esempio, c'è una piccola imperfezione sul piano del biliardo lungo la traiettoria di una delle due bocce. Oppure qualcuno scontra il tavolo mentre le bocce rotolano. Oppure un gatto intercetta una delle due bocce per giocarci. Oppure si verifica una scossa di terremoto. Oppure improvvisamente crolla il pavimento. Eccetera eccetera.

Questo è un esempio di un problema ben noto in IA, che va sotto il nome di *problema della cornice*, o *frame problem* (McCarthy e Hayes [1969]; questo nome peraltro non rimanda alla nozione di *frame* proposta da Minsky). Se difficoltà di questo genere sorgono per un dominio ben noto e tutto sommato ben delimitato come il gioco del biliardo si può facilmente immaginare il loro peso in contesti meno strutturati. Il *frame problem* si pone ogni qual volta si ragiona sulle precondizioni o sulle conseguenze di azioni ed eventi in ambienti non completamente prevedibili, e costituisce pertanto un problema ineludibile quando si tratta di progettare sistemi artificiali in grado di pianificare le proprie azioni, come nel caso di sistemi di IA applicati alla robotica (si veda ad esempio Pylyshyn [1987]; sul *frame problem* vedi anche Paternoster [2003], §2.2).

2.3. Limiti di risorse nell'elaborazione dell'informazione

Gli esseri umani non sono ragionatori perfetti: in generale non sono in grado di ricavare dalle informazioni di cui dispongono tutte le conseguenze che se ne potrebbero derivare, in quanto non dispongono di risorse di elaborazione sufficienti (tempo di calcolo, spazio di memoria, attenzione, eccetera). Da questo punto di vista la logica costituisce un'idealizzazione inadeguata delle capacità degli agenti

cognitivi. I sistemi logici sono chiusi rispetto alla conseguenza logica. In logica tutte le conseguenze di un certo insieme di premesse sono poste sullo stesso piano, e non è possibile distinguere tra conseguenze che richiedono gradi diversi di sforzo computazionale per essere derivate.

Dal punto di vista cognitivo ciò è decisamente poco realistico. Il problema risulta particolarmente evidente nel caso delle logiche epistemiche, ossia quei sistemi logici in cui si analizzano ragionamenti su ciò che vari agenti credono o conoscono a proposito del mondo e su ciò che è creduto o conosciuto da altri agenti. Nelle logiche epistemiche tradizionali gli agenti cognitivi sono tenuti a credere o a conoscere tutte le verità logiche e tutte le conseguenze logiche delle loro credenze o conoscenze. Si tratta di un aspetto del cosiddetto problema dell'*onniscienza logica* (si vedano ad es. Frixione [2002] e Fagin *et al.* [1995], cap. 9).

Il fatto che i modelli logici tradizionali non tengano conto dei limiti di risorse cui devono sottostare gli agenti cognitivi reali è comunque un problema di portata molto generale, che va al di là del problema specifico delle logiche epistemiche. Nella misura in cui non può tenere conto dei vincoli sulle risorse di calcolo, la logica non può fornire un modello computazionale realistico per il ragionamento.

2.4. Inferenze “giuste” contro inferenze “sbagliate”

Gli esseri umani eseguono inferenze che sono spesso sistematicamente “sbagliate” dal punto di vista della logica⁵. Inoltre, spesso accade che le credenze di un soggetto siano contraddittorie senza tuttavia che questo abbia gli effetti disastrosi che una

⁵In questo settore molte ricerche sono state condotte in psicologia del ragionamento; si veda ad esempio Girotto [1994].

contraddizione comporta in logica (un sistema logico classico in cui sia presente una contraddizione consente di derivare qualsiasi cosa).

In molti casi gli “errori” di ragionamento possono non costituire un limite per un soggetto razionale finito. Per un soggetto limitato (sia dal punto di vista dell’accesso alle informazioni rilevanti, sia dal punto di vista delle capacità inferenziali) trarre certe inferenze in modo logicamente “corretto” potrebbe essere impossibile o troppo dispendioso in termini di risorse computazionali. Ai fini di un comportamento globale efficiente ed adeguato può quindi essere più vantaggioso e più “razionale” saltare alle conclusioni anche se in modo azzardato. Anche il problema del ragionamento non monotono nasce, in un certo senso, da questo ordine di problemi.

2.5. Ragionamento e associazioni mentali

Non è plausibile che le informazioni siano rappresentate nella memoria umana come insiemi non strutturati di formule logiche. Presumibilmente sono organizzate in maniera più complessa, sulla base del loro contenuto, in modo da rendere agevole l'accesso alle informazioni rilevanti per lo svolgimento dei vari compiti cognitivi (ragionamento, comprensione del linguaggio, soluzione di problemi, eccetera). Secondo Minsky, ad esempio, le informazioni nella nostra memoria costituiscono blocchi più estesi e più organizzati di singole formule, e un sistema di rappresentazione della conoscenza adeguato dovrebbe consentire di definire strutture analoghe.

Nel modellare numerose attività cognitive sono rilevanti i legami associativi di “prossimità semantica” fra i concetti, legami che non sono riconducibili a nessi di

tipo logico. Ad esempio, il concetto *abete* è più “vicino” al concetto *Natale* di quanto non lo sia il concetto *calamaro*, ma ciò non rispecchia nessuna relazione logica specifica. Per simulare il modo di ragionare degli esseri umani i sistemi di rappresentazione della conoscenza dovrebbero simulare questa struttura associativa, e la logica non può offrire nulla in questo senso.

3. A FAVORE DELLA LOGICA

I logicisti non negavano la rilevanza dei problemi citati nel paragrafo precedente. Lo testimonia, ad esempio, l'enfasi che essi hanno posto da sempre sul ragionamento non monotono. Essi si limitavano a negare l'adeguatezza delle soluzioni proposte dai loro avversari, ritenendo troppo frettolosa da parte loro la liquidazione della logica. I problemi posti dagli oppositori della logica erano reali. Tuttavia i loro argomenti, almeno nei termini in cui erano stati originariamente formulati, non offrivano alternative rigorose. Sebbene lo studio dei sistemi logico-formali e delle loro proprietà metateoriche abbia avuto storicamente origine nell'ambito delle indagini sui fondamenti della matematica, la formalizzazione logico-matematica è stata progressivamente estesa a comprendere modalità, eventi, azioni, contesti spazio-temporali, rivelando una notevole flessibilità nel rendere conto di classi sempre più ampie di inferenze. Le esigenze dell'IA potrebbero indurre un'ulteriore evoluzione di questo tipo. Vi sono inoltre solidi argomenti che i sostenitori della logica possono opporre alle critiche dei suoi detrattori. Vediamone alcuni.

3.1. Necessità di una semantica

Molti formalismi proposti in IA come alternative alla logica erano stati sviluppati senza una semantica, eliminando in questo modo ogni possibilità di seri confronti. Senza la possibilità di assicurare la correttezza dei costrutti utilizzati per mezzo di un'adeguata semantica, programmi lunghi e complessi rischiano di diventare una massa inestricabile di segni senza significato. Il fatto che un sistema di rappresentazione della conoscenza sia implementato da un programma scritto in LISP o in qualche altro linguaggio di programmazione non garantisce nulla sul suo significato e sulla sua correttezza.

Se un formalismo deve costituire un sistema di *rappresentazione della conoscenza*, allora gli deve essere associata una semantica, cioè, deve essere definito qualche tipo di corrispondenza fra le espressioni che lo compongono e l'universo di oggetti su cui verte, in modo tale che abbia senso chiedersi se il mondo è realmente come quelle espressioni lo descrivono, e se le manipolazioni formali che operano sui simboli rispettano i vincoli imposti dal loro significato. Gli strumenti disponibili per questo tipo di analisi semantica sono quelli elaborati in logica: nel momento in cui ci si pone il problema di studiare con strumenti formali la semantica di un sistema di rappresentazione, si sta facendo (qualche tipo di) logica.

3.2. Livello dell'espressione e livello dell'implementazione

Vari autori hanno insistito sul fatto che lo studio dei sistemi di rappresentazione della conoscenza coinvolge due aspetti diversi, che sono stati denominati rispettivamente *livello dell'espressione* e *livello dell'implementazione*. Altri hanno parlato in modo analogo di *adeguatezza espressiva* contrapposta a *efficacia*

notazionale. Il primo aspetto concerne la progettazione di un linguaggio formale per esprimere conoscenze, il suo potere espressivo. Il secondo aspetto ha a che fare con le strutture dati e gli algoritmi che implementano su un calcolatore un linguaggio di rappresentazione, con la forma effettiva delle rappresentazioni impiegate, e l'impatto di queste strutture sulle operazioni del sistema. Riguardano questo secondo aspetto la sinteticità delle rappresentazioni e la facilità di modificarle.

Molte discussioni sull'uso della logica derivano da una confusione tra questi aspetti. Il ruolo della logica si colloca al livello espressivo, e non a quello dell'implementazione. La logica non è uno stile di programmazione, e non comporta alcuna scelta a livello delle strutture dati e del tipo degli algoritmi. Sulla stessa nozione di inferenza e sullo stesso linguaggio di rappresentazione possono essere basati sistemi diversi dal punto di vista degli algoritmi e dell'implementazione. Un sistema di rappresentazione può essere l'implementazione di una logica anche se non si presenta come un classico dimostratore di teoremi che manipola insiemi di formule.

Analoga è la distinzione di John McCarthy (McCarthy [1977]; McCarthy e Hayes [1969]) tra problemi di tipo *epistemologico* e problemi di tipo *euristico* dell'IA. I problemi epistemologici riguardano il tipo di conoscenza da rappresentare, la semantica e il potere espressivo dei linguaggi di rappresentazione, la correttezza delle regole di inferenza adottate. I problemi euristici riguardano invece la scelta delle strutture dati che implementano il linguaggio di rappresentazione, degli algoritmi che implementano le regole di inferenza, le strategie per esplorare gli spazi di ricerca, e così via. Secondo i sostenitori dell'approccio logico, molte delle obiezioni poste dai detrattori della logica non sono rilevanti a livello epistemologico,

poiché riguardano esclusivamente gli aspetti euristici. Ad esempio, il fatto di raggruppare la conoscenza in unità complesse e strutturate come i *frame*, i vantaggi che offrono *frame* e reti semantiche nella consultazione di una base di conoscenza, o il problema di rappresentare la vicinanza semantica fra concetti sono riconducibili al problema euristico di come rendere facilmente accessibile la conoscenza rappresentata agli algoritmi che eseguono le inferenze, e non sono incompatibili con l'impiego della logica al livello epistemologico.

3.3. Logiche per il ragionamento del senso comune

I logicisti concordano che il ragionamento umano è una cosa diversa rispetto alla deduzione logica. Ad esempio, sono consapevoli che la storia della logica è dominata dalla ricerca di tipi di inferenza che preservino la verità, mentre spesso il ragionamento del senso comune non rispetta questo vincolo. Tuttavia si possono usare gli strumenti della logica per costruire modelli, più o meno astratti, del ragionamento del senso comune. Molti teorici dell'IA pensano che non ci sia nulla da opporre all'uso di linguaggi logici standard per affrontare questi problemi. La logica dei predicati del primo ordine è molto flessibile, e pone pochissime pregiudiziali “ontologiche” sulla struttura della realtà da rappresentare. L'assunzione principale incorporata nei linguaggi del primo ordine è che il mondo possa essere descritto nei termini di individui e delle loro proprietà e relazioni. Ma la logica dei predicati non impone alcun limite su cosa debbano essere gli individui. Non solo oggetti fisici, ma anche intervalli o istanti di tempo, eventi, tipi, organizzazioni, mondi ed enunciati possono essere trattati come individui dal punto di vista logico.

C'è poi la possibilità di impiegare caratteristiche “non standard”, come ad esempio operatori intensionali o di ordine superiore. Oppure si può scegliere di usare sistemi non classici, come le logiche paraconsistenti o quelle condizionali. Anche se si fanno scelte di questo tipo, ci si muove comunque all'interno di un orizzonte di tipo logico.

Vi sono problemi che non erano mai stati affrontati dai logici, ma che non escludono a priori un trattamento logico. Un esempio è costituito dal ragionamento non monotono. Tutte le logiche “tradizionali” sono monotone, sia la logica classica, sia le logiche non classiche (ad esempio quella intuizionista). Tuttavia è stato possibile sviluppare vari sistemi di logica non monotona (si veda Antonelli [2003] e, più oltre, il paragrafo 4.1.).

La maggior parte della ricerca logica in IA è finalizzata ad elaborare modelli formali delle attività inferenziali di agenti razionali finiti, e limitati da diversi punti di vista (conoscenza, risorse di calcolo, attenzione). Il ragionamento non monotono è un caso particolare in questo senso. Un'altra direzione è quella legata alla possibilità di individuare sistemi logici “deboli”, adeguati a modellare le prestazioni di agenti con capacità computazionali limitate. Su questi punti torneremo nel seguito.

Dopo i vivaci (e talora violenti) dibattiti degli anni '70, a partire dalla prima metà del decennio successivo la contrapposizione tra tendenze logiciste e anti-logiche si è andata attenuando. Gran parte dei sistemi di rappresentazione simbolici proposti come alternative alla logica (*frame*, reti semantiche) sono stati ricondotti a notazioni alternative e a varianti di sistemi logici. La contrapposizione tra sostenitori e avversari della logica è stata caso mai sostituita dai dibattiti tra modelli simbolici e

altri tipi di modelli (connessionismo, IA situata, eccetera – si vedano Cordeschi [2003a, 2003b] e Paternoster [2003], §3).

Certamente, nell'ambito dell'IA simbolica, il filone logicista è stato ricco di risultati. Esso ha inoltre stimolato vari sviluppi inediti nel contesto della logica. Nelle pagine che seguono tenterò di identificare alcune linee di evoluzione della logica che sono state stimulate dai dibattiti cui ho fatto cenno.

4. CHE COSA PUÒ OFFRIRE L'IA ALLA LOGICA?

Il connubio con l'IA non è stato privo di conseguenze per la logica – e ha avuto ripercussioni anche sulla ricerca logica in filosofia. Le ricerche logiche svolte in IA presentano elementi di continuità con la ricerca logico-filosofica, ma suggeriscono al tempo stesso aspetti innovativi. Un denominatore comune tra le logiche filosofiche e l'uso della logica che viene fatto in IA consiste nell'intento di estendere l'applicazione della logica a contesti più vasti rispetto all'ambito originario delle inferenze matematiche e della formalizzazione delle teorie scientifiche. Gran parte degli sviluppi della ricerca logica in filosofia si sono indirizzati verso lo studio di forme di argomentazione estranee all'ambito della matematica e delle scienze in generale. Si pensi alle logiche modali aletiche e deontiche, alle logiche temporali, alle logiche rilevanti e paraconsistenti, e, in generale, a tutte le ricerche sulla formalizzazione del linguaggio naturale. In questa prospettiva, la ricerca logica in IA può essere letta come un'ulteriore estensione della ricerca logico-filosofica. Il progetto di formalizzazione del ragionamento del senso comune dell'IA si pone dunque in forte continuità con la ricerca logica in filosofia, ma al tempo stesso ha

posto in evidenza aspetti e temi che erano sfuggiti (o erano stati sottovalutati) dai logici di matrice filosofica.

La continuità tra i due settori peraltro è riscontrabile anche sul piano delle tecniche formali impiegate e dei riferimenti culturali: sfogliando la bibliografia di molti lavori di logica in IA ci si rende conto che i riferimenti alla letteratura logico-filosofica sono quasi altrettanto frequenti dei riferimenti a lavori interni al settore. Numerose ricerche poi si collocano per così dire a metà strada, e risultano difficilmente collocabili in modo univoco in uno dei due ambiti.

Dati gli scopi computazionali dell'IA, un suo apporto alla logica filosofica può essere individuato nell'introduzione di criteri metodologici e di vincoli di adeguatezza più stringenti per la ricerca logica in filosofia. Ad esempio, secondo Richmond Thomason [1988], un problema metodologico generale per la logica consiste nella "difficoltà di sviluppare metodologie corrette per motivare le logiche". Infatti "non è facile articolare principi per controllare delle ipotesi sul ragionamento" (p. 323). L'enfasi posta dall'IA sugli aspetti computazionali "promette non solo di suggerire teorie logiche nuove, ma mette a disposizione nuovi metodi per motivare e valutare le alternative teoriche" (p. 324). L'IA mette quindi a disposizione un banco di prova sperimentale che in molti casi può risultare prezioso per il logico di impostazione filosofica.

La ricerca logica in IA affronta molteplici ambiti diversi, variamente correlati e con numerose intersezioni: ragionamento sul tempo, sugli eventi e sulle azioni, ragionamento del senso comune sul mondo fisico (la cosiddetta *fisica ingenua*), ragionamento spaziale, ragionamento sulle credenze e sulla conoscenza, ragionamento distribuito in gruppi di agenti, formalizzazione logica degli aspetti

pragmatici e della comunicazione. Ne risulta un panorama ricco e articolato e, a prima vista, un po' eterogeneo.

Tuttavia, volendo tentare di identificare un tratto specifico che accomuni le ricerche logiche dell'IA e che le caratterizzi rispetto a ciò che viene fatto in altri settori della ricerca logica, lo si potrebbe individuare probabilmente nell'esigenza di modellare la capacità inferenziali di agenti razionali finiti e limitati. In ultima analisi, lo scopo dell'IA consiste nello studio e nella progettazione di sistemi computazionali in grado di operare disponendo di risorse realistiche. La logica quindi non viene impiegata al fine di elaborare canoni di razionalità assoluta e "disincarnata". Si tratta piuttosto di definire sistemi che, per quanto astratti, siano comunque modelli della *competenza inferenziale* di agenti individuali che possono disporre di risorse limitate e di un limitato accesso alle informazioni.

Si possono individuare due tipi di limiti cui sottostanno gli agenti cognitivi reali:

- (1) limiti epistemologici relativi all'impossibilità di accedere a tutte le informazioni che in linea di principio sarebbero rilevanti per un dato compito;
- (2) limiti nelle risorse computazionali disponibili (tempo di calcolo e spazio di memoria).

I limiti di cui al punto (1) determinano la necessità di prendere in considerazione forme di ragionamento in cui, come nel ragionamento non monotono, in assenza di tutta l'informazione rilevante, si salta alle conclusioni, salvo poi mantenere aperta la possibilità di dover ritrarre tali conclusioni qualora si

rendessero disponibili nuove informazioni incompatibili con le conseguenze derivate. Si tratta del problema già visto nel paragrafo 2.1. In linea di principio, per stabilire se un dato uccello è in grado o meno di volare, o per prevedere cosa succede se una palla da biliardo viene colpita in un certo modo, dovremmo disporre di una mole enorme di informazioni che non ci sono accessibili, e che, anche qualora lo fossero, sarebbe del tutto impossibile da gestire in maniera efficiente. Perciò gli agenti cognitivi traggono conclusioni sulla base di informazioni incomplete e di alcune assunzioni implicite di tipicità; tali conclusioni vengono poi riviste qualora si rivelino determinanti informazioni che non erano disponibili o che erano state trascurate.

Il punto (2) ha a che fare con l'esigenza di individuare forme di ragionamento che siano effettuabili entro limiti di risorse computazionali ragionevoli. Fondamentali a questo riguardo sono i risultati della teoria della complessità computazionale⁶. Sull'importanza teorica di questi aspetti per l'IA, e, in generale, per le scienze cognitive si vedano ad esempio Levesque [1988] e Frixione [2001]. Secondo Levesque, la necessità di concentrarsi su tipi di inferenze che possano essere svolti entro limiti di risorse accettabili non è una mera questione di

⁶ La teoria della complessità computazionale studia l'ammontare di risorse di calcolo (tempo e spazio di memoria) necessarie per risolvere i problemi computazionali. Si assume che un problema sia trattabile computazionalmente se richiede risorse che crescono al più polinomialmente rispetto alle dimensioni dell'input, mentre si ritengono non trattabili quei problemi che richiedono risorse che crescono in maniera esponenziale. Due classi di problemi particolarmente importanti da questo punto di vista sono le classi P e NP. P è la classe dei problemi che possono essere risolti in tempo polinomiale da una macchina di Turing usuale (e che quindi sono trattabili computazionalmente). NP è la classe dei problemi che possono essere risolti in tempo polinomiale da una macchina di Turing *non deterministica*. Benché nessuno sia riuscito a dimostrarlo, si ritiene che vi siano dei problemi in NP che non fanno parte di P. Si tratta della cosiddetta congettura $P \neq NP$, che costituisce probabilmente il problema aperto più importante dell'informatica teorica. Se la congettura $P \neq NP$ è vera, vi sarebbero dei problemi in NP (quelli che non fanno parte di P) che non sono trattabili computazionalmente. Frixione [2001] comprende una introduzione elementare a questi temi; Harel [2000] è un volume divulgativo che espone questo argomento in maniera chiara e piacevole.

efficienza, rilevante soltanto sul piano applicativo. Un modello computazionale di un'attività cognitiva che non sia computazionalmente trattabile non può dire nulla di come quell'attività possa essere fisicamente realizzata. Le proprietà computazionali di un dato problema sono in un certo senso assolute, e non dipendono dai possibili algoritmi per risolverlo, né da come tali algoritmi sono stati implementati, o dal tipo di *hardware* utilizzato. Un modello cognitivo che non tenga conto dei vincoli computazionali ha quindi ben poco valore esplicativo.

Per i sistemi di rappresentazione della conoscenza e per il ragionamento vale un *tradeoff* tra potere espressivo e trattabilità computazionale: al crescere dell'espressività la trattabilità inevitabilmente decresce. È ben noto che, assumendo che sia vera la congettura $P \neq NP$, già le procedure di decisione per la logica proposizionale richiedono tempi di calcolo che, nel caso peggiore, crescono esponenzialmente al crescere delle dimensioni dell'*input*: infatti decidere se una formula della logica proposizionale segue logicamente da un insieme dato di formule equivale al problema di decidere se una formula proposizionale p è una tautologia, e ciò richiede, nel peggiore dei casi, un numero di passi di calcolo esponenziale rispetto al numero di lettere enunciative diverse che compaiono in p . Se si passa poi alla logica dei predicati del primo ordine, essa non è neppure decidibile.

I punti (1) e (2) possono essere sintetizzati come segue: nella prospettiva di elaborare modelli della competenza inferenziale di agenti cognitivi limitati, i calcoli logici (classicamente) corretti e completi consentono di inferire ad un tempo (da punti di vista differenti) troppo e troppo poco. I sistemi di logica non monotona consentono di inferire *di più* rispetto ai sistemi logici tradizionali: le inferenze non

monotone sono infatti inferenze non corrette dal punto di vista tradizionale. Se traggo la conclusione che (a) *Titti vola* a partire dalle premesse che (b) *gli uccelli di solito volano* e che (c) *Titti è un uccello*, non si tratta di un'inferenza logicamente corretta. Infatti (a) non è vera in *tutti* i modelli di (b) e di (c); (a) è falsa ad esempio in quei modelli di (b) e di (c) in cui Titti è un pinguino. Viceversa, la ricerca di sistemi logici computazionalmente trattabili porta ad individuare calcoli logici più deboli, che consentono di inferire *di meno* dei calcoli completi rispetto alla semantica della logica classica.

Dunque:

- (1') La necessità di modellare forme di inferenza a partire da conoscenza incompleta comporta la necessità di studiare *forme di inferenza (classicamente) non corrette*;
- (2') La necessità di tenere conto di limiti di risorse realistici e dei vincoli di complessità computazionale comporta la necessità di prendere in considerazione *forme di inferenza (classicamente) incomplete*.

Si tratta di due tendenze complementari, che non sono tra loro incompatibili. Basti pensare per esempio ai *data base*: per un verso essi costituiscono dei sottoinsiemi della logica dei predicati del primo ordine estremamente poco espressivi e poco potenti dal punto di vista inferenziale, ma facilmente trattabili computazionalmente; per un altro verso spesso prevedono forme di inferenza non monotona quali la negazione come fallimento⁷. Un altro esempio è offerto dal PROLOG, che si basa

⁷ La negazione come fallimento è un tipo di inferenza non monotona in base alla quale se da un insieme di formule Δ non è possibile derivare una certa formula p , allora si assume che da Δ segua $\neg p$. Quando si consulta una base di dati si usa la negazione come fallimento se si ritiene che la base di dati sia ragionevolmente completa. Se ad esempio consultando l'orario ferroviario non trovo alcun treno per Milano che parta da Genova tra le dieci e le undici, assumo che *non ci sia* alcun treno da

su un sottoinsieme del calcolo dei predicati con buone proprietà computazionali (le clausole di Horn – si veda oltre il paragrafo 4.2.) e che incorpora anch'esso la negazione come fallimento. Non sempre però l'integrazione dei due tipi di esigenze si presenta priva di difficoltà; talvolta i sistemi non monotoni hanno proprietà poco entusiasmanti dal punto di vista della trattabilità computazionale. Varie ricerche sono state effettuate in questa direzione (si veda ad esempio Cadoli [1995]).

4.1. Sistemi non corretti rispetto alla logica classica

Le esigenze espresse ai punti (1') e (2') non comportano la necessità di rinunciare in assoluto a requisiti metateorici quali la completezza o la correttezza dei calcoli. Il fatto che i calcoli individuati non risultino corretti o completi rispetto alla semantica della logica classica non significa che per essi non possano valere altri criteri metateorici di adeguatezza. Essi possono risultare corretti o completi rispetto a strutture logiche non classiche di vario tipo (formalizzabili peraltro con strumenti omogenei a quelli tradizionalmente usati dai logici, ad esempio con gli strumenti insiemistici della semantica modellistica).

Forme di inferenza classicamente non corrette come la logica non monotona possono risultare corrette rispetto a definizioni di correttezza opportune. È il caso della *circumscription*, una formalizzazione del ragionamento non monotono proposta da John McCarthy [1977, 1980, 1986]. Essa si basa sulla seguente idea intuitiva. Vi sono predicati che valgono solo, per così dire, in casi eccezionali, mentre non si applicano in situazioni tipiche. Nel ragionamento del senso comune

Genova per Milano tra le dieci e le undici. La non monotonia dell'inferenza consiste nel fatto che, se si aggiungono nuove premesse a \tilde{A} , può darsi che diventi possibile derivare p ; in tal caso la conclusione $\neg p$ dovrà essere ritirata.

normalmente si assume implicitamente che tali predicati non valgano a meno che non si sappia esplicitamente che non è così. Un esempio è il predicato *daltonico*. Le persone daltoniche sono un sottoinsieme abbastanza ristretto dell'insieme degli esseri umani. Solitamente nessuno, ragionando su di una persona che non conosce, prende in considerazione la possibilità che sia daltonico, a meno che non abbia qualche informazione esplicita in proposito. L'idea di fondo che sta alla base della *circumscription* è che i predicati di questo genere vengano “circoscritti”, dove circoscrivere un predicato P in una teoria T vuol dire assumere che P abbia l'estensione più piccola possibile compatibilmente con le informazioni disponibili in T stessa. Ad esempio, si consideri la seguente teoria (scritta nel linguaggio del calcolo dei predicati del primo ordine con identità):

$$T = \{\text{daltonico}(\text{Achille}), \text{Roberto} \neq \text{Achille}, \text{Carlo} \neq \text{Achille}, \text{Carlo} \neq \text{Roberto}\}.$$

Circoscrivere il predicato *daltonico* in T significa assumere che Achille sia l'unica persona daltonica nel dominio della teoria, perché Achille è l'unico di cui si afferma esplicitamente che è daltonico. Circoscrivere *daltonico* in T equivale cioè ad assumere:

$$\forall x(\text{daltonico}(x) \rightarrow x = \text{Achille})$$

In questo modo si potranno derivare a loro volta le formule:

$$\neg \text{daltonico}(\text{Carlo})$$

$$\neg \text{daltonico}(\text{Roberto})$$

che non seguono dalla sola teoria T . È facile constatare che la *circumscription* dà luogo a forme di inferenza non monotona. Si supponga di venire a sapere che anche Carlo è daltonico. In tal caso si passerà dalla teoria T alla teoria T' così definita:

$$T' = T \cup \{daltonico(Carlo)\}$$

Circoscrivendo *daltonico* in T' si avrà:

$$\forall x(daltonico(x) \rightarrow x = Achille \vee x = Carlo)$$

per cui non sarà più possibile ottenere $\neg daltonico(Carlo)$.

McCarthy ha proposto una semantica modellistica per la *circumscription* detta *semantica dei modelli minimali* rispetto alla quale la *circumscription* risulta corretta. Intuitivamente, i modelli minimali di una teoria sono quei modelli che riducono al minimo il numero dei casi “anomali” o “eccezionali”. Nell’esempio precedente i modelli minimali della teoria T sono quelli in cui l’interpretazione del predicato *daltonico* è il più piccolo insieme possibile compatibilmente con le informazioni in T , cioè l’insieme formato dal solo Achille. Nella *circumscription* un’inferenza non monotona è corretta se la conseguenza è vera in tutti i modelli minimali della teoria. Yohav Shoham ha generalizzato l’idea di McCarthy a forme di ragionamento non monotono diverse dalla *circumscription*. Egli ha definito un quadro generale di tipo modellistico per la semantica delle logiche non monotone detto *semantica preferenziale*⁸. In base alla semantica preferenziale, ad ogni tipo di logica non monotona è associato un criterio di “preferenza”, che impone un ordinamento fra i modelli delle teorie. Data una teoria, le inferenze non monotone corrette sono quelle in cui la conseguenza è vera in tutti i modelli “preferiti” della teoria (in questa prospettiva, la semantica dei modelli minimali per la *circumscription* costituisce un esempio particolare di semantica preferenziale, ottenuta applicando un particolare criterio di preferenza tra i modelli).

⁸ Per una introduzione elementare alla semantica dei modelli minimali e alla semantica preferenziale si veda Frixione [1994], 6.1 e 6.3, con i relativi riferimenti.

4.2. Sistemi non completi rispetto alla logica classica

Per quel che concerne il punto (2'), l'obiettivo della trattabilità computazionale per i sistemi di rappresentazione della conoscenza e di ragionamento può essere perseguito mediante due diverse strategie (Cadoli [1995]). La prima consiste nell'indebolire il potere espressivo del linguaggio, in modo da disporre di procedure inferenziali semanticamente complete (rispetto a tale linguaggio ridotto) e computazionalmente trattabili. La seconda consiste nel mantenere la piena espressività del linguaggio e indebolire l'apparato deduttivo, ottenendo quindi la trattabilità computazionale a scapito della completezza (rispetto alla semantica della logica classica; vedremo che tali sistemi possono risultare completi rispetto a semantiche di tipo non classico). La strategia di indebolire l'apparato deduttivo comporta direttamente la perdita della completezza. L'indebolimento del potere espressivo del linguaggio consente di salvaguardare la completezza, ma solo rispetto a *quel* linguaggio. Il calcolo risulta semanticamente incompleto qualora venga immerso in un linguaggio con il pieno potere espressivo della logica classica (in un certo senso, quindi, l'incompletezza semantica viene spostata dall'apparato deduttivo al linguaggio).

Un esempio della prima strategia è offerto dalle *clausole di Horn*. Le clausole di Horn costituiscono un sottoinsieme del linguaggio della logica classica. Una clausola di Horn è una formula chiusa del tipo:

$$\forall x_1 \dots \forall x_n (p_1 \wedge \dots \wedge p_k \rightarrow p_{k+1})$$

dove n e k sono maggiori o uguali a 0, e dove ogni p_i è una formula atomica. Nel caso in cui $k = 0$ una clausola di Horn coincide con una formula atomica chiusa.

Nelle clausole di Horn proposizionali non ci sono variabili e quantificatori, e $p_1 \dots p_{k+1}$ sono singole lettere proposizionali. Sulle clausole di Horn predicative si basa il linguaggio di programmazione PROLOG.

Ebbene, nel caso delle *clausole di Horn proposizionali* il problema di decidere se una data formula segue o meno da un insieme di formule \tilde{A} ha complessità lineare rispetto alle dimensioni di \tilde{A} . Nel caso delle clausole di Horn predicative (senza simboli di funzione) il problema di stabilire se una data formula segue o meno da un insieme di formule è decidibile (contrariamente a quanto accade se si impiega il linguaggio della logica dei predicati del primo ordine completo), e, a determinate condizioni, risulta computazionalmente trattabile.

Un esempio della seconda strategia, basata sull'indebolimento dell'apparato deduttivo della logica classica, è costituito dai calcoli proposizionali sviluppati da Hector Levesque a partire da sistemi di logica rilevante, in particolare da una logica sviluppata da Alan Ross Anderson e Nuel D. Belnap, che la battezzarono logica del *tautological entailment*. Queste logiche risultano complete rispetto a strutture semantiche non classiche di tipo opportuno. A partire dalla logica del *tautological entailment* Levesque ha elaborato una logica modale epistemica che non risente del problema dell'onniscienza logica, in cui, in particolare, le credenze non risultano chiuse rispetto alla conseguenza logica (si vedano Levesque [1988], il cap. 12 di Levesque e Lakemeyer [2001], e Frixione [2002] per una presentazione sintetica).

L'idea alla base del *tautological entailment* consiste nel sostituire le interpretazioni della logica classica (che devono essere coerenti e complete, devono cioè associare a ciascun enunciato uno ed un solo valore di verità) con delle strutture semantiche non classiche, dette *situazioni*, che possono essere sia incoerenti (può

cioè accadere che in una data situazione certi enunciati ricevano sia il valore di verità vero, sia il valore di verità falso), sia incomplete (può accadere che in una situazione certi enunciati non ricevano alcun valore di verità)⁹.

Indichiamo con S l'insieme delle situazioni. Nella semantica per la logica della credenza implicita ed esplicita di Levesque, in ciascuna situazione s di S , a ogni lettera enunciativa p è associato un sottoinsieme di $\{\mathbf{v}, \mathbf{f}\}$. Cioè, a differenza di quanto accade nella semantica della logica proposizionale classica (in cui si danno due soli casi: data un'interpretazione, a p viene associato il valore \mathbf{v} oppure il valore \mathbf{f}), ora i casi sono quattro: oltre ai due precedenti può accadere che a p in s sia associato l'insieme vuoto \emptyset (ossia p non è né vera né falsa nella situazione s) e che a p in s sia associato l'insieme $\{\mathbf{v}, \mathbf{f}\}$ (ossia p è sia vera che falsa nella situazione s).

Intuitivamente, l'uso di situazioni incomplete dovrebbe modellare la parte di realtà rilevante per le credenze di un soggetto epistemico, lasciando indeterminato tutto il resto. “Si consideri – dice Levesque – la situazione in cui io siedo al mio terminale al lavoro. Potremmo dire che questa situazione giustifica il fatto che io sono al lavoro, che qualcuno è al mio terminale, che c'è un libro oppure un terminale sulla mia scrivania, e così via. D'altro canto, essa non motiva l'opinione che mia moglie è a casa, che non è fuori a fare acquisti, e neppure l'opinione che essa è oppure non è a casa. Benché quest'ultimo fatto sia certamente vero, che io sieda al mio terminale non ha nulla a che fare con tutto ciò”. L'uso di situazioni incoerenti viene motivato sulla base del fatto che un soggetto epistemico può avere concezioni

⁹ Il termine *situazioni* richiama la *situation semantics*, un'impostazione allo studio formale della semantica delle lingue naturali proposta da Barwise e Perry (si veda ad esempio Perry [1998]). Le situazioni di Barwise e Perry hanno in comune con quelle di Levesque il fatto di poter corrispondere a stati del mondo incompleti. Per il resto le due impostazioni sono molto diverse tra loro.

o informazioni sbagliate che lo portano a credere possibili stati di cose che di fatto non lo sono.

Data una situazione s , non accade (come accade invece nelle interpretazioni classiche) che una formula α è vera in s se e soltanto se α non è falsa in s . Quindi per caratterizzare il valore di verità di una formula in una situazione bisogna trattare separatamente il caso in cui la situazione rende vera la formula e quello in cui la rende falsa. Si avrà dunque che:

- (i) $\neg \alpha$ è vera in una situazione s se e solo se α è falsa in s
- (ii) $\neg \alpha$ è falsa in s se e solo se α è vera in s
- (iii) $\alpha \wedge \beta$ è vera in s se e solo se α e β sono entrambe vere in s
- (iv) $\alpha \wedge \beta$ è falsa in s se e solo se almeno una tra α e β è falsa in s
- (v) $\alpha \rightarrow \beta$ è vera in s se e solo α è falsa oppure β è vera in s
- (vi) $\alpha \rightarrow \beta$ è falsa in s se e solo α è vera e β è falsa in s

e così via. Nella logica classica la seconda clausola di ognuna di queste coppie diventa ridondante. Non è così invece, per le ragioni esposte sopra, nel caso generale della logica qui descritta.

In questa logica può succedere che una situazione s renda vera le formule α e $\alpha \rightarrow \beta$ senza rendere vera al contempo anche la formula β . Supponiamo infatti che una situazione s renda vere α e $\alpha \rightarrow \beta$. Supponiamo inoltre che in s α sia al contempo vera e falsa, e che sia falsa β . La situazione s rende vere sia α , sia

$\alpha \rightarrow \beta$ (quest'ultima perché α è falsa). In altri termini dunque questa logica non è chiusa rispetto al *modus ponens*.

In modo analogo si può constatare che non tutte le tautologie risultano vere in ogni situazione. Ad esempio vi sono situazioni in cui non è vera la formula $\alpha \vee \neg \alpha$: basta che una situazione non assegni ad α alcun valore di verità. Analogamente, se in una situazione s α è vero ma β non ha alcun valore di verità, allora in s non è vero $\alpha \wedge (\beta \vee \neg \beta)$. Può quindi accadere che due enunciati che sono logicamente equivalenti rispetto alla logica classica non abbiano lo stesso valore di verità in una situazione. Per ragioni analoghe da una contraddizione non segue qualsiasi formula (in una situazione s può essere vero $\alpha \wedge \neg \alpha$ senza che per questo in s sia vera qualsiasi formula).

Questo tipo di logica ha interessanti proprietà dal punto di vista computazionale. Ho già ricordato che, data la congettura $P \neq NP$, in logica proposizionale classica il problema di decidere se una certa formula segue logicamente da un insieme dato di formule non è trattabile computazionalmente. Nella logica del *tautological entailment* il problema di decidere la relazione di conseguenza logica ha caratteristiche computazionali migliori.

Più precisamente, nella logica proposizionale classica stabilire se una formula α segue da un insieme di formule \tilde{A} è un problema co-NP-completo; nella logica della credenza esplicita di Levesque stabilire se una credenza α segue da un insieme di credenze \tilde{A} richiede un tempo di calcolo che cresce linearmente rispetto alla lunghezza di α e di \tilde{A} (misurata nei termini del numero di lettere proposizionali distinte che vi compaiono) a condizione però che α e \tilde{A} siano in forma normale

congiuntiva. (Se \mathcal{A} e $\tilde{\mathcal{A}}$ sono formule proposizionali qualunque, allora il problema è co-NP-completo anche nella logica della credenza esplicita).

Uno sviluppo interessante di questo tipo di logiche è costituito da modelli di ragionamento basati, per così dire, su approssimazioni successive. Nel processo inferenziale si ottengono risposte via via “migliori” quanto maggiore è lo sforzo computazionale, in modo che, all’aumentare delle risorse di calcolo impiegate, si approssimano le prestazioni di un ragionatore classicamente completo. Un sistema di questo tipo per la logica proposizionale è stato proposto da Marco Cadoli e Marco Schaerf (Schaerf e Cadoli [1995], si veda anche Cadoli [1995], cap. 3). Questo modello, che impiega una semantica del tipo di quella utilizzata da Levesque, si basa su due successioni di relazioni di conseguenza logica, le prime sempre corrette, le seconde sempre complete, che convergono verso la relazione di conseguenza logica proposizionale classica. Man mano che si procede lungo le due successioni, aumentano rispettivamente completezza e correttezza, e si ottengono quindi approssimazioni sempre migliori della relazione di conseguenza logica classica. Quanto maggiore è lo “sforzo computazionale” impiegato, tanto migliore è l’approssimazione ottenuta.

5. SVILUPPI SUCCESSIVI

Il consenso sull’utilità della logica nella costruzione di modelli computazionali delle prestazioni cognitive è tutt’altro che unanime. Basti pensare alle posizioni dei sostenitori della cosiddetta *nuova IA* (connessionismo, IA situata, algoritmi genetici, e così via – si veda Cordeschi [2003a, 2003b] e, più in generale sulle scienze cognitive “post-classiche”, Paternoster [2003], §3) che spesso negano ogni

ruolo alla nozione stessa di rappresentazione, e che comunque sono estremamente critici nei confronti dell'IA simbolica in generale. Ma anche un autorevole sostenitore dell'impostazione logicista in IA come Drew McDermott a suo tempo ha espresso dubbi circa l'impiego della logica in IA (McDermott [1987]). Resta il dato di fatto che l'IA ha stimolato molteplici sviluppi inediti della ricerca in logica, che hanno a loro volta avuto ricadute in altri settori, come ad esempio in filosofia.

Le linee di sviluppo delineate nel paragrafo precedente, indirizzate all'elaborazione di modelli logici di agenti razionali finiti e limitati, hanno trovato applicazione e hanno fornito spunti per ulteriori sviluppi in numerosi settori dell'IA. Farò qualche rapido cenno ad alcuni di tali sviluppi, fermo restando che le considerazioni dei paragrafi precedenti sono abbastanza generali da applicarsi in qualche misura a ciascuno di essi.

Ho accennato a come in robotica il *frame problem* costituisca una sfida per i pianificatori e per i sistemi robotici basati sulla rappresentazione simbolica di azioni e di eventi. In origine questo tipo di impostazione presentava indubbie carenze, ad esempio dal punto di vista dell'integrazione della pianificazione e del ragionamento con i compiti di livello più basso come percezione e controllo motorio. La reazione a questo stato di cose ha portato allo sviluppo della cosiddetta *nuova robotica* o *robotica situata*. I sostenitori di questa impostazione (il più noto è Rodney Brooks) ritengono che un sistema robotico debba soprattutto essere in grado di reagire in maniera veloce ed efficiente alle caratteristiche del proprio ambiente, e questo non richiederebbe ricche rappresentazioni o forme complesse di inferenza (si veda ancora Cordeschi [2003a, 2003b] e Paternoster [2003], §3.2). Tale strategia incontra tuttavia severi limiti nella progettazione di sistemi demandati a svolgere compiti

complessi. I recenti sviluppi della cosiddetta *robotica cognitiva* (si veda ad esempio Carlucci Aiello *et al.* [2000])¹⁰ mirano a rispondere alle sfide poste dalla nuova robotica attraverso un rinnovato sviluppo di sistemi di rappresentazione basati sulla logica e finalizzati al ragionamento su azioni e su eventi (si veda ad esempio Reiter [2001], che sviluppa il formalismo del *situation calculus* originariamente proposto da John McCarthy).

Vi sono altri aspetti che presentano punti di contatto con i temi che ho menzionato nel paragrafo precedente, e ai quali l'IA ha dato maggiore enfasi rispetto all'impostazione tradizionale della logica. Ad esempio, nel momento in cui la logica non viene utilizzata come canone di razionalità astratta ma con l'intento di costruire modelli della competenza inferenziale di singoli agenti individuali diventa rilevante analizzare quegli aspetti del ragionamento che riguardano l'interazione tra agenti diversi. In IA sono stati estesamente studiati modelli logici del ragionamento sulle credenze e sulla conoscenza in gruppi di agenti (Fagin *et al.* [1995]). Formalismi di tipo logico vengono utilizzati per la progettazione di *sistemi multiagente* (Wooldridge [2000] e [2002]). Spesso sono state utilizzate a questo scopo logiche basate sulla formalizzazioni degli atteggiamenti proposizionali (come credere, sapere, desiderare, eccetera). In questa tradizione rientrano i cosiddetti sistemi BDI, basati sugli atteggiamenti di credenza, desiderio e intenzione (BDI sta appunto per *Belief, Desire, Intention*).

In questo settore esistono interessanti punti di contatto tra gli sviluppi della logica in IA e alcune evoluzioni recenti delle sulle teorie della razionalità pratica,

¹⁰ Si veda anche la pagina web del gruppo di ricerca di robotica cognitiva all'Università di Toronto: <http://www.cs.toronto.edu/cogrobo/>.

come la teoria della scelta razionale e la teoria dei giochi. Quest'ultima entra in campo soprattutto nel momento in cui ci si pone il problema della coordinazione tra più agenti (si veda ad esempio Bicchieri [1993]), e sorge quindi l'esigenza di formalizzare nozioni come quella di conoscenza comune (*common knowledge* - Lewis [1969]). Anche in questo ambito si è posta l'esigenza di impiegare modelli di agenti razionali limitati, e talvolta sono stati impiegati strumenti quali logiche non monotone o logiche per il ragionamento sulle credenze e sulla conoscenza (Fagin *et al.* [1995])¹¹.

Gli aspetti qui illustrati non esauriscono certo la ricchezza della ricerca logica in IA. Ad esempio, il settore delle logiche *fuzzy* ha dato un nuovo impulso allo studio delle logiche polivalenti. Molta ricerca è stata fatta su logica e ragionamento probabilistico.

Questi sviluppi, oltre a fornire materiali inediti per l'indagine logico-filosofica, possono fornire strumenti per affrontare problemi filosofici preesistenti. Si possono a ricordare, a titolo di esempio, due problemi di tipo logico sorti nell'ambito della filosofia del linguaggio, per i quali le linee di ricerca sopra delineate possono fornire utili strumenti. I modelli di inferenza a partire da informazioni incomplete e il ragionamento non monotono possono risultare rilevanti per un trattamento più adeguato degli aspetti inferenziali della semantica lessicale. Il significato della maggior parte degli elementi del lessico delle lingue naturali è caratterizzabile nei termini di tratti tipici piuttosto che di condizioni necessarie e sufficienti. Da questo punto di vista, le logiche non monotone potrebbero consentire

¹¹ Una sede deputata alla discussione di questi problemi è il convegno TARK (*Theoretical Aspects of Rationality and Knowledge* - <http://www.tark.org/>) che si tiene ogni due anni; si tratta di incontri interdisciplinari che coinvolgono, tra gli altri, logici, filosofi, economisti e informatici.

di elaborare sistemi di postulati di significato più flessibili e adeguati rispetto agli strumenti logici tradizionali¹². I modelli del ragionamento sottoposto a vincoli di risorse possono essere adeguati per elaborare modelli più realistici dei contesti di atteggiamento proposizionale. Ho già ricordato (nel par. 2.3) che le usuali logiche epistemiche, incorporando l'assunto di onniscienza logica, costituiscono idealizzazioni difficilmente accettabili delle prestazioni di agenti epistemici reali. Forme di inferenza approssimata sono state impiegate per elaborare sistemi epistemici più realistici da questo punto di vista (su entrambi i punti si veda Frixione [1994]).

BIBLIOGRAFIA RAGIONATA

Un testo introduttivo sull'IA che rivolge particolare attenzione alla logica è Carlucci Aiello e Cialdea Mayer [1995].

Per gli aspetti tecnici relativi alla logica in IA si possono consultare i vari manuali della disciplina. Burattini e Cordeschi [2001] è un manuale introduttivo pensato per studenti di formazione umanistica, in particolare per le discipline della comunicazione; su logica e rappresentazione della conoscenza si veda il terzo capitolo. Fum [1994] è un manuale di IA destinato agli psicologi; la logica e la rappresentazione della conoscenza vi sono affrontate nei capp. 3, 5 e 6. Tra i manuali più avanzati disponibili in italiano ricordiamo Nilsson [1998] (soprattutto le parti III e IV) e Russell e Norvig [1994] (soprattutto i capp. dal 6 al 10). Nel sito di Stuart Russell (uno degli autori dell'ultimo dei manuali citati) c'è una ricca pagina

¹² Su logiche non monotone e semantica lessicale si vedano ad esempio i lavori di Richmond Thomason reperibili a questo indirizzo: <http://www.eecs.umich.edu/~rthomaso/documents/nls/index.html>.

di link a siti di IA; in particolare su logica e rappresentazione della conoscenza si veda: <http://www.cs.berkeley.edu/%7Erussell/ai.html#logic>.

Per un inquadramento storico dell'IA in generale si veda Cordeschi [1996]. Brachman e Levesque [1985] è una raccolta di articoli classici sulla rappresentazione della conoscenza in IA, che ben documentano, tra le altre cose, il dibattito tra sostenitori e avversari della logica.

Per una prima introduzione alla rappresentazione della conoscenza disponibile *on line*, in cui si parla anche di logica, si veda http://www.dif.unige.it/epi/hp/frixione/appunti_KR.pdf. Il sito dell'Associazione Americana per l'Intelligenza Artificiale (AAAI) comprende un'introduzione ai vari settori dell'IA utile per i numerosi link (<http://www.aaai.org/Pathfinder/html/welcome.html>); vi sono trattati vari temi interessanti per il nostro argomento (si vedano in particolare le voci *Representation* e *Reasoning*). Su vari aspetti dei rapporti tra logica e IA si veda Frixione [1994]. Sul ragionamento non monotono rimando alle indicazioni bibliografiche di Antonelli [2003].

Una recente rassegna su logica e IA dovuta a Richmond Thomason si trova all'URL <http://www.eecs.umich.edu/~rthomaso/documents/lai/index.html>.

Sui sistemi multiagente, Wooldridge [2002] fornisce una trattazione introduttiva. L'autore ha predisposto una pagina web (<http://www.csc.liv.ac.uk/~mjw/pubs/imas/>) in cui si possono trovare varie informazioni, tra cui collegamenti ad altri siti e materiali didattici, comprese le trasparenze per un corso. Sugli agenti si veda anche Colombetti *et al.* [2003], con la

relativa pagina di collegamenti: <http://lgxserver.uniba.it/lei/ai/networks/03-1/colombettilinks.html>

Infine, per un trattamento approfondito dei vari aspetti tecnici dell'impiego della logica in IA si vedano i volumi di Gabbay *et al.* [1993-1998]. I primi due volumi sono dedicati alle nozioni di base, il terzo volume è dedicato al ragionamento non monotono e al ragionamento incerto, il quarto al ragionamento epistemico e temporale, il quinto alla programmazione logica.

RIFERIMENTI BIBLIOGRAFICI

Antonelli A. (2003), "La logica del ragionamento plausibile", *Sito Web Italiano per la Filosofia (SWIF), Linee di Ricerca* (<http://www.swif.it/biblioteca/lr/intro.php>).

Bicchieri C. (1993), *Rationality and Coordination*, Cambridge University Press, 2a ed. 1997. Tr. it., (1998), *Azione collettiva e razionalità sociale*, Feltrinelli, Milano.

Brachman R., Levesque H. (a cura di) (1985), *Readings in Knowledge Representation*, Morgan Kaufmann, Los Altos, CA.

Burattini E., Cordeschi R. (a cura di) (2001), *Intelligenza artificiale*, Carocci, Roma.

Cadoli M. (1995), *Tractable Reasoning in Artificial Intelligence*, Springer Verlag, Berlino.

Carlucci Aiello L., Cialdea Mayer M. (1995), *Invito all'intelligenza artificiale*, Franco Angeli, Milano.

Carlucci Aiello L., Nardi D., Pirri F. (2000), "Case studies in cognitive robotics", in V. Cantoni *et al.* (a cura di), *Human and Machine Perception 3: Thinking, Deciding and Acting*, Kluwer, Dordrecht (NL).

Colombetti M., Fornara N., Verdicchio M. (2003), “Linguaggio e realtà sociale nei sistemi di agenti artificiali”, *Networks*, 1 (<http://lgxserver.uniba.it/lei/ai/networks/>).

Cordeschi R. (1996), “Intelligenza artificiale”, in L. Geymonat (1996), *Storia del Pensiero Filosofico e Scientifico. Il Novecento*, Garzanti, Milano, vol. 8, tomo III.

(2003a), “Vecchi problemi filosofici per la nuova intelligenza artificiale”, *Networks*, 1 (<http://lgxserver.uniba.it/lei/ai/networks/>).

(2003b), "Filosofia dell'Intelligenza Artificiale", *Sito Web Italiano per la Filosofia (SWIF), Linee di Ricerca* (<http://www.swif.it/biblioteca/lr/intro.php>).

Fagin R., Halpern J.Y., Moses Y., Vardi M.Y. (1995), *Reasoning about Knowledge*, MIT Press, Cambridge, Mass.

Frixione M. (1994), *Logica, significato, intelligenza artificiale*, Franco Angeli, Milano.

(2001), “Tractable competence”, *Minds and Machines*, 11(3), 379-397. (per una versione preliminare: http://www.dif.unige.it/epi/hp/frixione/minds_machines.pdf)

(2002), “Contesti di credenza e onniscienza logica”, in *La svolta contestuale*, a cura di C. Penco, McGraw-Hill Italia, Milano.

Fum D. (1994), *Intelligenza artificiale*, Il Mulino, Bologna.

Gabbay D.M., Hogger C.J., Robinson J. A. (a cura di) (1993-1998), *Handbook of Logic in Artificial Intelligence and Logic Programming*, 5 voll., Oxford University Press, Oxford.

Giroto V. (1994), *Il ragionamento*, Il Mulino, Bologna.

Harel D. (2000), *Computers Ltd. What they really can't do*, Oxford University Press, Oxford. Tr. it., (2002), *Computer a responsabilità limitata. Dove le macchine non riescono ad arrivare*, Einaudi, Torino.

Levesque H.J. (1988), "Logic and the complexity of reasoning", in *Journal of Philosophical Logic*, 17(4), pp. 355-389.

Levesque H.J., Lakemeyer G. (2001), *The Logic of Knowledge Bases*, MIT Press, Cambridge, Mass.

Lewis D. (1969), *Convention. A Philosophical Study*, Harvard University Press, Cambridge, Mass. Tr. it., (1974), *La convenzione. Studio filosofico*, Bompiani, Milano.

McCarthy J. (1977), "Epistemological problems of artificial intelligence", Proc. 5th IJCAI, Cambridge, MA, pp. 1038-1044. Anche in Brachman e Levesque 1985. (<http://www-formal.stanford.edu/jmc/epistemological.html>)

(1980), "Circumscription - A form of non-monotonic reasoning", in *Artificial Intelligence*, 13, 27-39. (<http://www-formal.stanford.edu/jmc/circumscription.html>)

(1986), "Applications of circumscription to formalizing common sense knowledge", in *Artificial Intelligence*, 28, 89-116. (<http://www-formal.stanford.edu/jmc/applications.html>)

McCarthy J., Hayes P. (1969), "Some philosophical problems from the standpoint of artificial intelligence", in B. Meltzer e D. Michie (a cura di), *Machine Intelligence 4*, Edinburgh University Press, Edinburgh. (<http://www-formal.stanford.edu/jmc/mcchay69.pdf>)

McDermott D. (1987), "A critique of pure reason", in *Computational Intelligence*, 3, pp. 151-160.

Minsky M. (1975), "A framework for representing knowledge", in P. Winston, McGraw-Hill (a cura di), *The Psychology of Computer Vision*. Anche in Brachman e Levesque 1985. (<http://web.media.mit.edu/~minsky/papers/Frames/frames.html>)

Nilsson N.J. (1998), *Artificial Intelligence: A New Synthesis*, Morgan Kaufmann, San Francisco, CA. Tr. it. di S. Gaglio, (2002), *Intelligenza artificiale*, Apogeo, Milano.

Paternoster A. (2003), "I fondamenti epistemologici della 'nuova' scienza cognitiva. Il funzionalismo tra bancarotta e rifondazione", *Sito Web Italiano per la Filosofia (SWIF), Linee di Ricerca* (<http://www.swif.it/biblioteca/lr/intro.php>).

Perry J. (1998), "Semantics, model-theoretic", in *Routledge Encyclopedia of Philosophy*, Routledge, Londra (cfr. in particolare: <http://www-esli.stanford.edu/~john/PHILPAPERS/sitsem.pdf>).

Pylyshyn Z.W. (1987) (a cura di), *The Robot's Dilemma: The Frame Problem in Artificial Intelligence*, Albex, Norwood, NJ.

Reiter R. (2001), *Knowledge in Action. Logical Foundations for Specifying and Implementing Dynamical Systems*, MIT Press, Cambridge, Mass.

Russell S.J., Norvig P. (1994), *Artificial Intelligence. A Modern Approach*, Simon & Schuster, Englewood Cliffs, NJ. Tr. it. di Luigia Carlucci Aiello, (1998), *Intelligenza artificiale. Un approccio moderno*, UTET, Torino.

Schaerf M., Cadoli M. (1995), "Tractable reasoning via approximation", in *Artificial Intelligence*, 74, pp. 249-310. (file postscript compresso: <http://www.dis.uniroma1.it/pub/ai/papers/scha-cado-93.ps.gz>).

Smith B. (1982), Prologo di *Reflection and Semantics in a Procedural Language*, in Brachman e Levesque 1985.

Strawson P.F. (1950), "On referring", *Mind*, 59, 320-344. Tr. it. in A. Bonomi (a cura di), (1973), *La struttura logica del linguaggio*, Bompiani, Milano.

Thomason R. (1988), "Philosophical logic and artificial intelligence", in *Journal of Philosophical Logic*, 17(4), pp. 321-327.

Wooldridge M. (2000), *Reasoning about Rational Agents*, MIT Press, Cambridge, Mass.

(2002), *An Introduction to MultiAgent Systems*, John Wiley & Sons, Chichester, UK.