

Bibliotec@SWIF

Linee di Ricerca

a cura di
Luciano Floridi

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.

CURATORE

Luciano Floridi [luciano.floridi@swif.it] è professore associato di logica, Università di Bari, e Markle Foundation Fellow in Information Policy, membro del Wolfson College e dei Dipartimenti di Filosofia e di Informatica, Oxford University, dove ha fondato e coordina con Jeff Sanders lo IEG, il gruppo di ricerca interdipartimentale in Philosophy of computing and information. Le sue ricerche riguardano la philosophy of computing and information, la computer ethics, l'epistemologia e la storia dello scetticismo, con oltre cinquanta articoli in riviste internazionali e varie monografie, tra cui: *Scepticism and the Foundation of Epistemology - A Study in the Metalogical Fallacies* (Brill, 1996); *Philosophy and Computing: An Introduction* (Routledge, 1999); *Sextus Empiricus, The Recovery and Transmission of Pyrrhonism* (Oxford University Press, 2002). Ha curato la *Blackwell Guide to the Philosophy of Computing and Information* (Blackwell, 2003). Ha fondato e dirige con Mauro di Giandomenico lo SWIF.

SUPERVISORE TECNICO

Fabrizio Martina [fabrizio.martina@fabriziomartina.it] ha progettato e gestisce il database e i siti di TFO - Tesi Filosofiche Online, di Bibliotec@SWIF, CxC - Calls for Comments, Linee di Ricerca. Titolare di una società di sviluppo e consulenza informatica ed editoriale, esperto in progettazione di database e applicazioni web, collabora con alcune case editrici.

COORDINATORE EDITORIALE

Gian Maria Greco [gianmaria.greco@asi.unile.it]. Laureato in Filosofia e dottorando di ricerca in "Discipline Storico-Filosofiche" presso l'Università di Lecce. Junior Research Associate presso l'Information Ethics Research Group, Computing Laboratory, Università di Oxford. In SWIF è direttore dei progetti Bibliotec@SWIF e TFO - Tesi Filosofiche Online, coordinatore editoriale di Linee di Ricerca, responsabile dell'Ufficio Stampa di FOLDOP.

La revisione editoriale generale di Linee di Ricerca è a cura di Gian Maria Greco.

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.

Linee di Ricerca

Aldo Antonelli

LA LOGICA DEL RAGIONAMENTO PLAUSIBILE

Versione 1.0



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

G. Aldo Antonelli [aldo@uci.edu] è professore associato presso il Dipartimento di logica e filosofia della scienza dell'Università della California, Irvine. Oltre che delle logiche per l'intelligenza artificiale si è occupato di teoria della definizione, logica e teoria dei giochi, teorie degli insiemi senza fondazione, logica modale, logiche libere, e il programma neo-logicista. Ha pubblicato su numerose riviste, fra cui *Journal of Symbolic Logic*, *Journal of Philosophical Logic*, *Notre Dame Journal of Formal Logic*, *Artificial Intelligence*, e *Philosophia Mathematica*.

La revisione editoriale di questo saggio è a cura di Gian Maria Greco.

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

LA LOGICA DEL RAGIONAMENTO PLAUSIBILE

ALDO ANTONELLI

Versione 1.0

INTRODUZIONE

La moderna logica simbolica, così come si è venuta a sviluppare cominciando con la *Begriffsschrift* di Gottlob Frege alla fine del diciannovesimo secolo, e gli ulteriori contributi di Bertrand Russell, Kurt Gödel e Alfred Tarski nel ventesimo, è caratterizzata da una proprietà formale nota come monotonicità. In base a questa proprietà, se una certa conclusione C può essere ottenuta sulla base di un insieme P di premesse, tale conclusione può essere ottenuta anche sulla base di qualsiasi insieme di premesse che includa le premesse contenute in P . In altre parole, una volta che si è ottenuta una certa conclusione P , tale conclusione non può essere invalidata tramite l'aggiunta di ulteriori premesse.

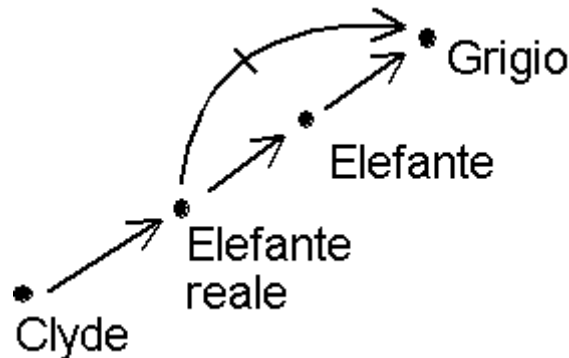
La ragione di tale caratteristica risiede nel fatto che la logica simbolica, e in particolare la sua incarnazione di maggiore successo, la logica del prim'ordine (d'ora in avanti: LPO), utilizza una nozione di conseguenza logica basata sulla nozione di *contro-esempio*. In particolare, si dice che un enunciato E è conseguenza di un insieme S di enunciati se e soltanto se non è possibile re-interpretare il linguaggio —mantenendo fisso il significato delle parole logiche— in modo da

rendere veri gli enunciati in S falsificando al tempo stesso l'enunciato E (una tale re-interpretazione è detta un «contro-esempio»). Ne segue che se E è conseguenza di S , è anche conseguenza di qualsiasi insieme S' che contenga tutti gli enunciati di S , poiché un contro-esempio che verifica S' falsificando E è anche un contro-esempio che verifica S .

La natura monotona della logica simbolica deriva dal fatto che essa fu sviluppata allo scopo di rappresentare in modo puramente formale il ragionamento deduttivo (specialmente matematico). Quando la conclusione di un'argomentazione segue con cogenza deduttiva dalle premesse (quando la conclusione è, in un certo senso, già contenuta nelle premesse), l'insieme delle premesse può essere aumentato a piacere senza con questo invalidare la cogenza dell'argomentazione. Se la conclusione *non può non essere vera* se le premesse sono vere, questo fatto non può venire a meno, non importa quali altre informazioni vengano aggiunte alla base di conoscenza. Ne segue che, nel ragionamento deduttivo, l'insieme delle conclusioni che possono essere derivate da un dato insieme di conoscenze cresce in modo proporzionale all'insieme di conoscenze stesso.

Vi è, tuttavia, un altro tipo di ragionamento, più tipico della vita quotidiana (ma non solo), in cui le conclusioni sono raggiunte solo *provvisoriamente*, riservandosi il diritto di ritrarle alla luce di ulteriori informazioni. In questo tipo di ragionamento, cosiddetto *plausibile* («*defeasible*», nella letteratura di lingua inglese), il soggetto «salta alle conclusioni» in base a informazioni parziali, riservandosi il diritto di rivisitare tali conclusioni quando nuove informazioni siano disponibili.

Un tipico esempio di ragionamento plausibile è dato dalle *tassonomie*. Vi sono molte situazioni in cui vogliamo organizzare le informazioni a nostra disposizione in modo *gerarchico*, in cui le categorie più ampie ricomprendono quelle più ristrette. Un esempio è dato dal seguente diagramma:



Il diagramma rappresenta una tassonomia relativa a un individuo, Clyde. Clyde è un elefante reale, e gli elefanti reali sono (un particolare tipo di) elefanti; e come si sa, gli elefanti tendono a essere grigi. Ma gli elefanti reali sono un'eccezione: non sono grigi, ma — come è noto — bianchi. La connessione con il ragionamento plausibile è la seguente. Supponiamo di avere inizialmente informazioni relative solo al fatto che Clyde è un elefante. Siccome gli elefanti tendono a essere grigi, possiamo plausibilmente inferire che Clyde è grigio. Quando poi veniamo a sapere che Clyde è sì un elefante, ma un elefante reale, la conclusione che Clyde è grigio viene ritratta, per essere rimpiazzata da quella opposta, che Clyde non è grigio.

Abbiamo qui un meccanismo di tipo ereditario in cui le sotto-categorie tassonomiche ereditano le caratteristiche delle sopra categorie. Ma tale meccanismo è raramente stretto, e spesso le informazioni tassonomiche hanno eccezioni: gli uccelli volano, ma i pinguini (un tipo particolare di uccelli) costituiscono un'eccezione. Un modo per catturare questo tipo di inferenze è attraverso una lista

esplicita di tutte le eccezioni. Si tratterebbe di avere una regola della forma «se x è un un uccello, e x non è un pinguino, non è uno struzzo, non è un neonato, non è ..., allora x vola», dove la lista di eccezioni (abbraviata da ..., è potenzialmente molto lunga). Un'alternativa più naturale consiste nell'interpretare la relazione di ciascuna categoria con le proprie sottocategorie in modo non rigido, assumendo che ciascuna sottocategoria erediti i tratti delle categorie sotto cui ricade, *a meno che non ci siano informazioni esplicite al contrario*.

Considerando un altro esempio, quando ci viene detto che Stellaluna è un mammifero, è naturale concludere che Stellaluna non vola, poiché, per la maggior parte, i mammiferi non volano. Ma tale conclusione può venire invalidata dall'informazione che Stellaluna è un pipistrello: i pipistrelli sono un tipo particolare di mammiferi, e le informazioni più specifiche hanno la precedenza su quelle più generiche. È possibile, naturalmente, che anche tale conclusione venga invalidata, ad esempio se venissimo a sapere che Stellaluna è un pipistrello neonato. Si tratta, come si può vedere, di complesse argomentazioni, che non solo sono fuori portata per LPO, ma che richiedono di venire analizzate e regimentate.

Un altro esempio di ragionamento plausibile proviene dalla teoria delle basi dati (*database*). Supponiamo di dover viaggiare da Oshkosh nel Wisconsin a Minsk in Russia. Consultando l'agenzia di viaggi, scopriamo che non vi sono voli diretti: come fa l'impiegato dell'agenzia a sapere che non ci sono voli? In un certo senso, l'impiegato non *sa* che non ci sono voli. La base dati delle linee aeree contiene solo informazioni esplicite riguardanti i voli effettivamente previsti, ma nessuna informazione riguardanti i voli non esistenti. Tuttavia, la base dati incorpora anche quella che si chiama l'*assunzione del mondo chiuso*, che ci dice in pratica che la

base dati è completa, e che quindi se un certo volo non è menzionato esplicitamente, non esiste. È chiaro che la conclusione che non ci sono voli diretti da Oshkosh a Minsk è solo provvisoria, e che viene a cadere quando la base dati è aumentata con certi tipi di informazioni.

In modo simile, i sistemi diagnostici incorporano forme di ragionamento plausibile. Quando un dispositivo complesso subisce un guasto, è naturale supporre che un insieme minimo di componenti ne sia responsabile. Ad esempio, se il venir meno di due componenti scelti fra *A*, *B*, e *C*, può spiegare il guasto, è plausibile supporre che solo due, e non tutti e tre siano venuti a mancare. Ma tale conclusione è solo plausibile, perché può essere ritirata alla luce di ulteriori informazioni (ad esempio, se il guasto permane dopo avere sostituito *A* e *B*).

Questi tipi di ragionamento plausibile sono assai difficili da catturare in modo formale. Nonostante sia sempre stato chiaro, fin dagli inizi, che LPO è del tutto inadeguata a rappresentare formalmente il ragionamento plausibile, I tentativi di sviluppare dei formalismi che fossero al tempo stesso precisi e adeguati non sono cominciati fino a circa il 1980. In quell'anno, infatti, l'autorevole rivista *Artificial Intelligence Journal* pubblicò un numero monografico dedicato alle logiche *non monotone*, le logiche, cioè, dedicate a catturare formalmente il ragionamento plausibile. Lo studio formale del ragionamento plausibile, a partire dalle sue origini nell'informatica e nell'intelligenza artificiale, si è sviluppato fino a diventare un campo di ricerca autonomo, enfatizzando gli aspetti fondazionali e concettuali almeno altrettanto quanto quelli di complessità computazionale.

RELAZIONI DI CONSEGUENZA

Gli esempi della sezione precedente rendono chiaro che la formalizzazione del ragionamento plausibile richiede l'abbandono della monotonicità della relazione di conseguenza logica. Ma tale abbandono è solo un primo passo. Bisogna chiedersi, infatti, quali proprietà della relazione di conseguenza plausibile siano desiderabili, una volta abbandonata la monotonicità. Poiché la relazione di conseguenza connette un insieme Γ di enunciati con singoli enunciati φ , si possono considerare in astratto le proprietà di tale relazione. Ad esempio, la monotonicità può essere espressa dicendo che se φ è una conseguenza di Γ , allora è anche una conseguenza di qualsiasi altro insieme Δ di enunciati che contenga Γ come un sottoinsieme. Le seguenti proprietà sono comunemente considerate proprietà desiderabile di qualsiasi relazione di conseguenza plausibile:

1. **Sopraclassicalità**: se φ è una conseguenza di Γ in LPO, allora è anche una conseguenza plausibile di Γ .
2. **Riflessività**: se φ è uno degli enunciati contenuti in Γ , allora φ è una conseguenza plausibile di Γ .
3. **Taglio**: se φ è una conseguenza plausibile di Γ , e ψ è una conseguenza plausibile di Γ insieme a φ , allora ψ è già una conseguenza plausibile di Γ .

La sopraclassicalità incorpora l'idea che la relazione di conseguenza logica di LPO è in un certo senso la più piccola, o la più conservativa, e che quindi se φ è una conseguenza di Γ in LPO, allora deve anche essere una conseguenza di Γ in qualsiasi altra relazione di conseguenza che sia più liberale di quella di LPO.

Delle altre due condizioni, la più semplice concettualmente è la riflessività. Certamente vogliamo che tutti gli enunciati già contenuti in Γ siano inferibili da Γ , qualunque sia la nostra nozione di conseguenza logica. Non è chiaro in che senso una relazione che non soddisfa la riflessività possa chiamarsi una relazione di conseguenza.

Il taglio, una forma di transitività, è un'altra caratteristica cruciale delle relazioni di conseguenza. Il taglio esprime un principio di conservazione: se si aggiunge a Γ un enunciato che è già una conseguenza di Γ , ciò non conduce ad alcun *incremento* nelle conseguenze di Γ . Un altro modo di considerare il taglio è come una condizione sulla lunghezza delle dimostrazioni, e in particolare come la condizione che la lunghezza di una dimostrazione non deve condizionare il grado con cui le premesse danno sostegno alla conclusione.

Mentre LPO chiaramente soddisfa tutte e tre le proprietà sopramenzionate, il ragionamento di tipo probabilistico, spesso indicato come candidato alla formalizzazione del ragionamento plausibile, non riesce a soddisfare il taglio. La ragione, chiaramente, è che nel ragionamento di tipo probabilistico il sostegno che le premesse danno alla conclusione è inversamente proporzionale alla lunghezza della dimostrazione. Si consideri il seguente esempio. Si abbrevi « α è alto-atesino» con Ax , « α è di madre lingua tedesca» con Bx , e « α è nato in Germania» con Cx . Inoltre, sia Γ l'insieme comprendente gli enunciati «La maggior parte degli A sono B », «La maggior parte dei B sono C », e Ax . Enunciati della forma «La maggior parte degli A sono B » vanno interpretati probabilisticamente, nel senso, che la probabilità condizionale di B dato A è, ad esempio, maggiore del 50%. In modo simile, diciamo che Γ dà sostegno (probabilisticamente) a un enunciato φ se Γ assegna a φ una

probabilità maggiore del 50%. Nel nostro esempio, Γ sostiene Bx (perché Ax insieme a «La maggior parte degli A sono B » assegna a Bx una probabilità maggiore del 50%); e Γ insieme a Bx sostiene Cx (per simili motivi), ma Γ da solo non sostiene Cx , poiché Γ ci dice solo che x è alto-atesino, e la probabilità che un alto-atesino sia nato in Germania è molto bassa o nulla.

Come abbiamo visto, la monotonicità (o la sua mancanza) meritano considerazioni a parte. Vi sono, nella letteratura scientifica, due ulteriori proprietà che sono state proposte come surrogati delle monotonicità, di cui sono casi particolari:

1. **Monotonicità cauta:** Se φ e ψ sono (ciascuna indipendentemente) conseguenze plausibili di Γ , allora ψ è una conseguenza plausibile di Γ insieme a φ .
2. **Monotonicità razionale:** Se la negazione di φ non è conseguenza plausibile di Γ , ma ψ è conseguenza plausibile di Γ , allora ψ è anche conseguenza plausibile di Γ insieme a φ .

La monotonicità cauta ci dice che aggiungere una conseguenza plausibile all'insieme delle premesse non conduce a nessuna *riduzione* nell'insieme delle conseguenze. In quanto tale, la monotonicità cauta è l'inverso preciso del taglio. Insieme, le due proprietà ci dicono che aggiungere una conseguenza plausibile all'insieme delle premesse non conduce né a un aumento né a una diminuzione dell'insieme delle conseguenze. Queste due proprietà — taglio e monotonicità cauta — fissano il carattere *cumulativo* del processo inferenziale: ci dicono che possiamo continuare a trarre conclusioni e usare tali conclusioni come nuove premesse, senza preoccuparci che così facendo alcune conclusioni vengano indebitamente validate o invalidate.

Il caso della mononicità razionale, sebbene superficialmente simile, è molto diverso, come dimostra il controesempio, dovuto a Robert Stalnaker, dei tre compositori: Verdi, Bizet, e Satie. Supponiamo inizialmente di accettare (correttamente, ma senza assoluta certezza) che Verdi è italiano, mentre Bizet e Satie sono francesi. Supponiamo anche di venire a sapere da una fonte attendibile che Verdi e Bizet sono compatrioti. Questa nuova informazione ci porta a ritrattare la conclusione che Verdi è italiano (perché potrebbe essere francese) e che Bizet è francese (perché potrebbe essere italiano), ma non ci dice nulla riguardo a Satie, e quindi continuiamo a credere che Satie è francese. Quindi, l'enunciato «Satie è francese» è una conseguenza plausibile di «Verdi e Bizet sono compatrioti» (assumendo le nostre ipotesi iniziali sulla nazionalità dei tre compositori sullo sfondo).

Notiamo anche che l'informazione che Verdi e Bizet sono compatrioti non ci porta a quella che Verdi e Satie *non* sono compatrioti (perché Verdi, in fondo, potrebbe essere francese). Possiamo quindi considerare cosa succederebbe, dovessimo aggiungere «Verdi e Satie sono compatrioti» al nostro insieme di premesse, che già include «Verdi e Bizet sono compatrioti». Questi due enunciati, insieme, sono equivalenti all'asserzione che tutti e tre i compositori hanno la stessa nazionalità.

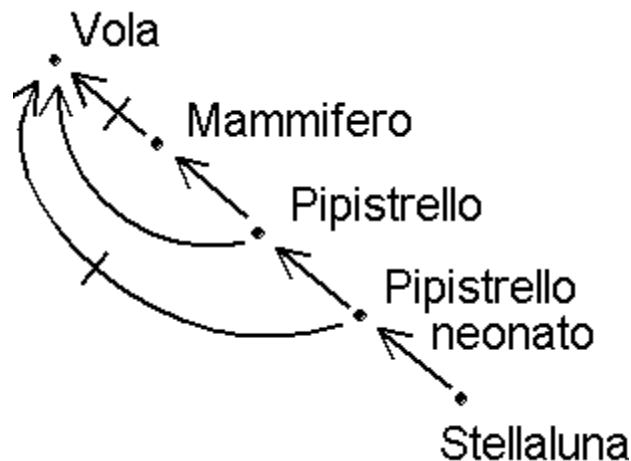
Ne segue che la conclusione che Satie è francese non è preservata dall'aggiunta del second enunciato («Verdi e Bizet sono compatrioti»), contrariamente a quanto richiede la mononicità razionale. Questo ci porta a identificare le quattro proprietà (sopraclassicalità, riflessività, taglio, e mononicità cauta) come caratteristiche desiderabili di una relazione di conseguenza plausibile.

LOGICHE NON MONOTONE

Nel corso degli ultimi 25 anni, sono stati proposti molti formalismi volti alla formalizzazione del ragionamento plausibile, a cominciare dai contributi di John McCarthy, Drew McDermott, Jon Doyle, e Raymond Reiter. È con tali formalismi che la realizzazione dell'inadeguatezza della LPO per la formalizzazione del ragionamento plausibile viene per la prima volta coniugata con un serio tentativo di provvedere un contesto formale in cui cominciare a discutere delle inferenze plausibili in modo preciso.

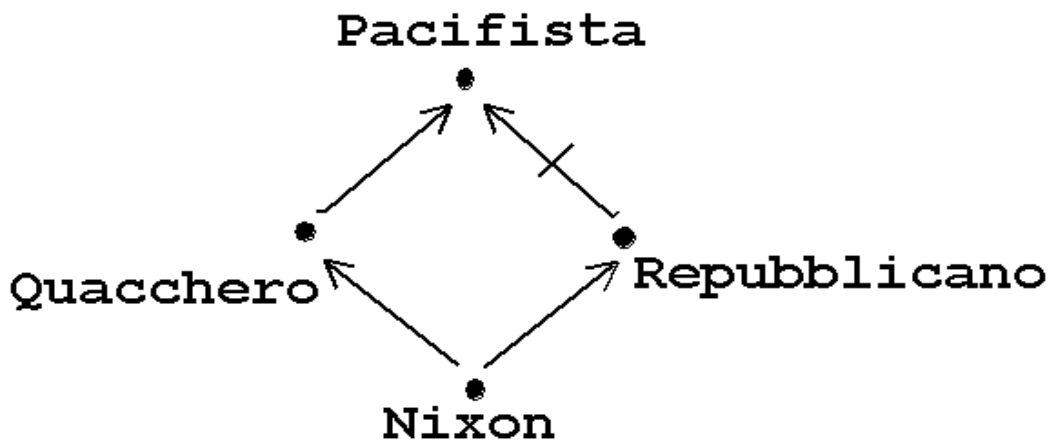
Fra i molti formalismi proposti, ci limitiamo in questa sede a considerare brevemente le reti semantiche non monotone e la logica di *default* di Reiter (per un ampio e completo trattamento si veda, ad esempio, Dov M. Gabbay et al. [1994]).

Come abbiamo visto, ogniqualevolta si abbia un insieme di conoscenze organizzato tassonomicamente, sorge la possibilità di *eccezioni*, e con le eccezioni, anche la possibilità di inferenze contraddittorie (come ad esempio le conclusioni che *Stellaluna vola* — perché è un pipistrello — e che *non vola* — perché è un mammifero). Le relazioni tassonomiche sono rappresentabili in modo naturale tramite reti semantiche, in cui le varie categorie sono usate per etichettare i nodi della rete, che a loro volta sono connessi per mezzo di archi rappresentanti la relazione di inclusione fra categorie. Quando tale relazione di inclusione non è stretta, ma solo plausibile, si hanno reti semantiche non monotone in cui si possono rappresentare le eccezioni. Una rete semantica non monotona si può trovare, ad esempio, nella seguente figura:



Un arco del tipo $A \rightarrow B$ rappresenta l'asserzione che gli A sono, tipicamente, B (mentre l'arco corrispondente in cui la freccia è sbarrata rappresenta l'asserzione che tipicamente gli A non sono B). Il compito principale di una teoria delle reti semantiche non monotone è quello di determinare quali asserzioni siano giustificate sulla base di un tale diagramma. Una speciale considerazione, estranea ad altri formalismi, entra in gioco nel caso delle reti semantiche non monotone, e cioè la *specificità*. È una delle intuizioni principali riguardanti le reti semantiche che nel caso di conflitti fra potenziali conclusioni, le informazioni più specifiche devono avere il sopravvento su quelle più generiche. Nell'esempio qui sopra, ciò comporta che l'asserzione che i pipistrelli volano viene preferita a quella che i pipistrelli non volano perché quest'ultima può essere ottenuta solo mediante il meccanismo dell'ereditarietà dalla sovracategoria dei mammiferi.

Non tutti i conflitti, però, possono essere risolti tramite la specificità. Ci sono conflitti in cui nessuna delle due asserzioni incompatibili è più specifica dell'altra. Si consideri ad esempio il famoso diagramma noto come «Nixon diamond»:



È chiaro che il diagramma sostiene due asserzioni incompatibili, e cioè che Nixon è un pacifista (perché quacchero) e che Nixon non è un pacifista (perché repubblicano), e nessuna delle due è più specifica dell'altra. In questo caso la teoria delle reti semantiche non monotone si trova di fronte a una scelta. Una prima alternativa è quella di adottare un atteggiamento cosiddetto *scettico* e in presenza di conclusioni contrastanti, sospendere il giudizio. Ciò significa, ad esempio, adottare le conclusioni intermedie che Nixon è sia un quacchero che un repubblicano, così come quella che i quaccheri, a differenza dei repubblicani, tendono a essere pacifisti, ma né la conclusione che Nixon è una pacifista, né quella contraria.

L'alternativa a un atteggiamento scettico è invece un atteggiamento cosiddetto *credulone*, che consiste nell'accettare sempre un insieme di conclusioni quanto più grande possibile, soggetto solamente alla condizione della coerenza. In questo caso, ciò comporterebbe l'accettazione di una, e soltanto una, delle due conclusioni incompatibili («Nixon è pacifista», «Nixon non è pacifista»).

La teoria delle reti semantiche usa la nozione di *estensione* allo scopo di rendere queste idee precise. (La nozione di estensione è un concetto ricorrente nella

formalizzazione del ragionamento plausibile lo ritroviamo, ad esempio, nelle logiche cosiddette di *default*.) Come nozione ausiliaria, si definisca un *cammino* attraverso la rete semantica come una successione di archi consecutivi, tutti positivi, tranne possibilmente l'ultimo (un cammino si dice positivo o negativo a seconda del segno dell'ultimo arco). Dato un cammino attraverso una rete, ad esempio $A \rightarrow B \rightarrow C$, si può considerare l'asserzione corrispondente, «Gli A sono tipicamente C » ottenuta dai nodi estremi del cammino. Possiamo ora definire un'*estensione* di una rete semantica come un insieme massimale di cammini tale che nessuna delle asserzioni da essi sostenute risulta incompatibile con nessun'altra. Più specificamente, un'estensione è un insieme massimale di cammini attraverso la rete non contenente alcuna coppia di cammini da uno stesso nodo A a uno stesso nodo B , ma di segno opposto. Una volta ottenuta un'estensione per una rete semantica è possibile considerare l'insieme di asserzioni da essa sostenute: ad esempio, l'asserzione «Gli A sono (oppure: tendono ad essere) B » è sostenuta da una data estensione se e soltanto se l'estensione contiene un cammino dal nodo (etichettato) A al nodo (etichettato) B .

Molte varietà di ragionamento plausibile, specialmente le varianti cosiddette «credulone», ammettono multiple estensioni incompatibili, senza che vi sia alcun modo preferito per sceglierne una privilegiata. Ciò rende difficile risolvere il problema della definizione di una relazione di conseguenza per il ragionamento plausibile. Per tali ragioni, alcuni studiosi, come ad esempio Horty, Thomason e Touretzky [1990] hanno preferito dare un approccio di tipo scettico diretto.

Un approccio di tipo scettico si dice *diretto* se procede alla costruzione di un'estensione privilegiata — corrispondente presumibilmente alle intuizioni

scettiche e quindi non avente il carattere massimale usato sopra — seguita dalla derivazione delle asserzioni giustificate sulla base di tale estensione. Nell'esempio del «Nixon diamond», l'approccio diretto procede a identificare l'unica estensione corrispondente alle intuizioni di tipo scettico, in cui né l'asserzione che Nixon è un pacifista, né quella che non lo è sono giustificate.

Bisogna anche notare che, accanto all'approccio scettico diretto ve ne è anche uno *indiretto*, che procede alla costruzione di *tutte* le estensioni massimali (cioè, «credulone») della rete, e poi considera l'insieme di asserzioni giustificate in ogni estensione. Alcuni studiosi, come ad esempio Makinson e Schlechta [1991] hanno fornito argomentazioni volte a mostrare la superiorità di tale approccio indiretto (tali argomentazioni sono considerate nella prossima sezione).

Accanto alle reti semantiche, le logiche cosiddette di *default* (originariamente proposte da R. Reiter), costituiscono un formalismo particolarmente ricco e flessibile per la rappresentazione del ragionamento plausibile. In tale formalismo, un *default* è semplicemente una regola di inferenza della forma:

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

dove α , β , e γ sono enunciati in dato linguaggio, detti la *premessa*, *giustificazione*, e la *conclusione* della regola (rispettivamente). L'interpretazione intuitiva della regola di default è che se α è conosciuto, e non vi sono ragioni per pensare che β sia falso, allora si può inferire γ . È chiaro che la regola è essenzialmente non monotona: se a un certo punto si dovesse venire a sapere che β è falso, l'inferenza verrebbe bloccata.

Queste intuizioni possono essere rese precise attraverso la nozione di estensione. Dato un insieme Γ di regole di default, un'estensione per Γ è un sottoinsieme Δ di Γ le cui conclusioni implicano le premesse delle regole in Δ e le cui giustificazioni non sono refutabili sulla base delle conclusioni delle regole in Δ . Come nel caso delle reti semantiche, vi sono insiemi Γ per cui non esiste alcuna estensione, o per cui esistono estensioni multiple. Si consideri ad esempio l'insieme Γ comprendente solamente la seguente regola R:

$$\frac{\tau : \beta}{\neg \beta}$$

dove τ è un enunciato analiticamente vero e $\neg \beta$ è la negazione di β . L'insieme Γ non può avere alcuna estensione. Si supponga infatti che Δ sia un'estensione per Γ ; vi sono due casi, a seconda che R sia contenuta in Δ oppure no. Se lo è, allora $\neg \beta$ è una conclusione delle regole in Δ , e quindi la giustificazione di R viene a cadere. Se non lo è, allora si ha che τ è conosciuto (perché analiticamente vero), e non vi sono ragioni per pensare che β sia falso (l'unica ragione possibile è data da $\neg \beta$ che però non è fra le conclusioni delle regole in Δ), e quindi R deve essere in Γ dopo tutto. Siccome in ogni caso si ottiene una contraddizione, bisogna concludere che Γ non ha estensioni.

Si consideri ora il caso di un insieme Γ con estensioni multiple. Γ contiene, ad esempio, le due regole R_1 e R_2 :

$$\frac{\tau : \neg \beta}{\neg \beta}$$

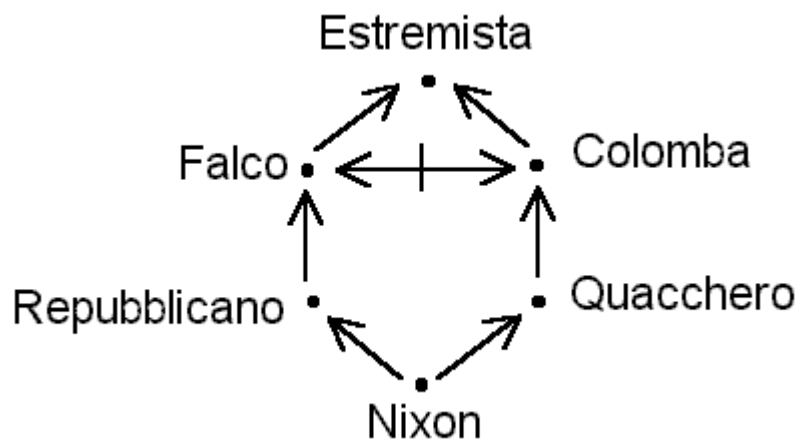
$$\frac{\tau : \neg \beta}{\beta}$$

Un ragionamento simile a quello sopra dimostra che ciascuna estensione di Γ deve contenere esattamente una delle due regole R_1 e R_2 , e quindi che Γ deve avere esattamente due estensioni, una contenente R_1 e una contenente R_2 .

Come nel caso delle reti semantiche, quando si hanno insiemi di regole senza estensioni o aventine più di una, è difficile definire una relazione di conseguenza basata sul concetto di regola di default.

IL PROBLEMA DELLE «FLOATING CONCLUSIONS»

A illustrazione del tipo di problemi che si incontrano nella formalizzazione del ragionamento plausibile si consideri il caso già accennato delle cosiddette «floating conclusions». Per semplicità ci limitiamo a considerare il caso (più semplice) delle reti semantiche. Un enunciato E è una «floating conclusion» di una rete semantica Γ se tale enunciato è derivabile in ogni estensione di Γ , ma non vi è nessuna linea di argomentazione che è comune a tutte le estensioni. Per chiarire il problema si consideri il seguente diagramma:



La rete semantica in questione ha, come è chiaro, due estensioni massimali. Una prima estensione giustifica l'asserzione che Nixon è un falco, mentre l'altra

giustifica l'asserzione che Nixon è una colomba, ed entrambe giustificano l'asserzione che Nixon è un estremista. Tale asserzione, tuttavia, è contenuta nelle due estensioni attraverso due *diversi cammini*, ed è pertanto detta una «floating conclusion».

Bisogna accettare le «floating conclusions»? A prima vista si direbbe di sì: sembrerebbe intuitivamente corretto concludere che Nixon è politicamente estremista, in analogia con il classico modello della prova per casi. Si consideri l'individuo, Nixon; due casi sono possibili: o Nixon è un repubblicano oppure è un quacchero. Nel primo caso Nixon è un falco, e quindi estremista; nel secondo caso è una colomba, e quindi di nuovo estremista. Ne segue che, indipendentemente da quale dei due casi corrisponda a realtà, Nixon è politicamente estremista.

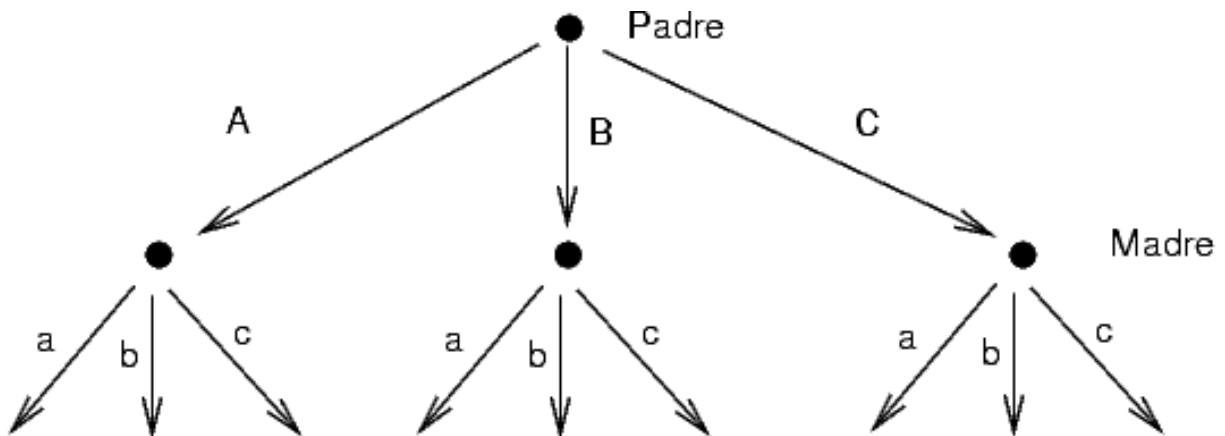
La correttezza intuitiva di tale linea di ragionamento è stata apportata da vari ricercatori (ad esempio Makinson and Schlechta) come un'argomentazione contro gli approcci scettici diretti alla teoria delle reti semantiche. Un approccio diretto, infatti, procedendo all'identificazione di un'estensione non massimale, deve per forza perdere la conclusione che Nixon è politicamente estremista, dato che ciascuno dei due cammini che la giustificano viene eliminato nel ragionamento scettico. Recentemente, tuttavia, John Horty [2002] ha presentato un'argomentazione volta a chiamare in causa questa conclusione apparentemente inattaccabile. Horty ci chiede di considerare il seguente scenario.

Si supponga che I miei genitori abbiano accumulato una fortuna di un milione di dollari, ma che al fine di ridurre la loro esposizione fiscale abbiano scelto un regime di separazione dei beni. Purtroppo, entrambi sono stati colpiti da una malattia rara e incurabile, hanno fatto testamento, e sono caduti in un coma

irreversibile. Supponiamo anche che, per alleviare il mio imminente dolore, io stia considerando un grosso acquisto, ad esempio, uno yacht. Benché lo yacht sia disponibile immediatamente, il prezzo è sufficientemente buono da garantirne la vendita entro un mese. Ora, io posso assicurarmi lo yacht pagando un grosso anticipo, e promettendo di pagare il resto non appena avrò ereditato la mia fortuna. Le mie presenti condizioni finanziarie sono tali che non potrei assolutamente permettermi lo yacht senza la futura eredità, e se essa non si materializzasse dovrei rinunciare al cospicuo anticipo.

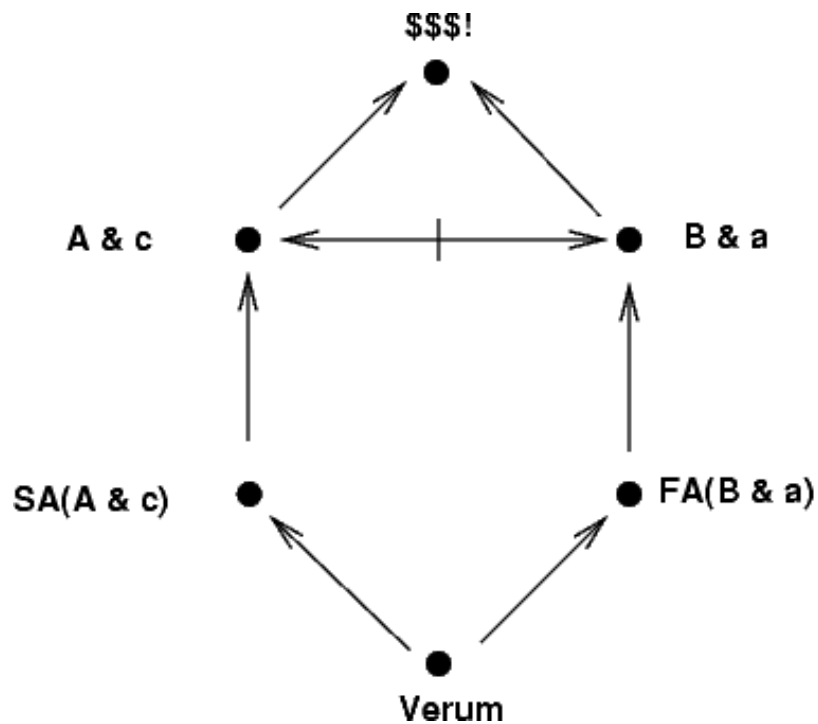
Lasciando da parte i dubbi sulla sincerità del mio dolore, è chiaro che le mie preferenze sono tali che se io fossi convinto di ereditare da almeno uno dei miei genitori allora sarebbe nel mio interesse pagare il deposito, altrimenti sarebbe nel mio interesse non pagare il deposito.

A complicare la situazione, c'è il fatto che io ho un fratello e una sorella, entrambi i quali sono fonti di informazioni (fallibili, come tutti, ma altrimenti) estremamente attendibili. Entrambi sono riusciti a parlare con i miei genitori prima che cadessero nel coma, e adesso mi riportano la seguente situazione. Mio fratello, avendo parlato con mio padre, mi assicura che papà lascerà i soldi a lui, ma che non devo preoccuparmi, perché mamma lascerà la sua fortuna a me. Mia sorella invece, avendo parlato con mia madre, mi assicura che mamma lascerà i soldi a lei, ma che non devo preoccuparmi, perché papà lascerà la sua fortuna a me. Questo è tutto quel che so. I miei genitori sono in coma, e mio fratello e mia sorella sono irraggiungibili. La questione, ovviamente è se io debba concludere che eredito almeno mezzo milione di dollari e quindi se io debba pagare l'anticipo sullo yacht. La situazione è rappresentata in termini di teoria della decisione nel seguente diagramma ad albero:



Gli archi etichettati A, B, e C rappresentano le tre possibili azioni di mio padre (lasciare la fortuna a me, mio fratello, o mia sorella), e in modo simile gli archi etichettati a, b, e c. Mio fratello mi informa che la situazione è data da $B+a$, mentre mia sorella sostiene $A+c$.

È chiaro, sostiene Horty, che in questo caso è del tutto non banale che io riceverò un'eredità. Infatti ci sono buone ragioni per pensare il contrario. Eppure, in questo caso l'asserzione che io avrò almeno mezzo milione di dollari alla fine del mese è una «floating conclusion». Infatti, se la situazione viene rappresentata in termini epistemici piuttosto che di teoria della decisione, abbiamo il seguente diagramma:



che è, come è chiaro, isomorfo a quello usato nell'esempio sull'estremismo politico di Nixon. Dal momento che l'insieme di conclusioni derivabili da una data rete semantica è determinato dalla topologia della rete e non dalle etichette dei nodi, è chiaro che anche nel caso dell'estremismo politico di Nixon la derivabilità della «floating conclusion» deve dipendere da fattori ulteriori rispetto alla rappresentazione per mezzo di una certa rete semantica.

Più in generale, questo esempio è sufficiente a chiamare in questione l'assunzione che le «floating conclusions» siano sempre desiderabili. Un'analisi precisa di questo tipo di conclusioni va oltre i limiti del presente trattamento, ma almeno questo è chiaro, che tale analisi non dovrebbe essere pregiudicata dal particolare formalismo scelto.

CONCLUSIONI

Vi sono due requisiti fondamentali alla base dello sviluppo di un formalismo per il ragionamento plausibile. Il primo è un requisito di *adeguatezza materiale*, il secondo un requisito di *correttezza formale*. Il primo requisito concerne l'adeguatezza del formalismo a rappresentare in maniera corretta un vasto numero di esempi di ragionamento plausibile. In molti casi vi sono chiare intuizioni riguardo al tipo di conclusioni plausibili siano derivabili da un dato insieme di conoscenze, a quali condizioni tali conclusioni siano ritratte e quali conclusioni invece *non* seguano dalle conoscenze disponibili. Noi vogliamo che il formalismo prescelto correttamente rappresenti ciascuno di questi casi, e non pregiudichi quelle circostanze in cui intuizioni precise non sono disponibili.

Il requisito di correttezza formale, invece, concerne le proprietà intrinseche del formalismo. In questo caso, vogliamo che il formalismo dia origine a una relazione di conseguenza plausibile che abbia le necessarie proprietà matematiche, e in particolare che soddisfi i requisiti della riflessività, taglio, e monotonicità cauta.

È possibile che a volte i due requisiti di adeguatezza materiale e correttezza formale entrino in tensione. La spinta a catturare un numero sempre maggiore di esempi può portare a varianti del formalismo che non soddisfano le proprietà formali desiderate, e viceversa il desiderio di mantenere un formalismo matematicamente elegante può portare a volte a perdere i fenomeni che si vogliono rappresentare.

A maggior ragione, la possibilità di questa tensione contribuisce a fare dello studio formale del ragionamento plausibile uno dei campi di ricerca più interessanti, all'intersezione della filosofia con le scienze formali e cognitive.

BIBLIOGRAFIA RAGIONATA

Mentre sono reperibili numerose opere che introducono gentilmente il lettore alla logica classica del prim'ordine (Bencivenga [1984] un esempio per tutti), è difficile trovare opere analoghe per la logica del ragionamento plausibile, specialmente in italiano. A livello introduttivo, ma in lingua inglese, si possono consultare Brewka, Dix, e Konolige [1997], Antonelli [2001] e Antonelli [2003], quest'ultimo disponibile in rete attraverso la *Stanford Encyclopedia of Philosophy* (<http://plato.stanford.edu>). In italiano, ma a livello più avanzato si possono consultare Benzi [1997] e Fischer Servi [2001].

Una pietra miliare nello sviluppo delle logiche per l'intelligenza artificiale fu la pubblicazione, nel 1980, di un numero monografico dell'autorevole rivista *Artificial Intelligence*. Molti dei lavori contenuti in tale volume introducono i nuovi formalismi per il ragionamento plausibile: non solo le già citate logiche di default di Reiter [1980], ma anche formalismi di cui non ci siamo veramente potuti occupare, come ad es. (ma non solo) la "circostrizione" di McCarthy [1980] e l'approccio modale di McDermott e Doyle [1980]. Molti di tali lavori (ed altri, ugualmente fondamentali) sono stati successivamente raccolti nell'antologia di Ginsberg [1987] ormai purtroppo esaurita. Al suo posto va invece segnalata un'altra opera di consultazione, che raccoglie molti eccellenti articoli: il volume curato da Gabbay, Hogger e Robinson [1994], si presenta come il punto di riferimento fondamentale

per chi, a suo agio con trattamenti tecnici, voglia avvicinarsi al campo delle logiche non monotone.

Un tema che invece può essere affrontato anche senza vaste competenze tecniche è quello delle relazioni di conseguenza non monotone. In questo ambito è possibile applicare le proprie intuizioni pre-teoretiche e aggiudicare le varie alternative (naturalmente, l'argomento può anche essere perseguito in maniera altamente tecnica). Il punto di partenza, almeno storicamente, è il lavoro di Gabbay [1985], e contributi fondamentali si devono a Stalnaker [1994] e Makinson [1994] (ma si veda anche Antonelli [1997]).

Infine, un altro argomento che, purché potenzialmente tecnico, può essere affrontato anche a livello informale (grazie all'intuitività dei diagrammi) è quello delle reti semantiche non monotone. Qui il riferimento centrale è Horty [1994], che fornisce un'eccellente sistematizzazione del soggetto delle reti semantiche. Fra i testi originali che propongono l'approccio "scettico" vi sono Thomason, Horty, Touretzky [1987] e Horty, Thomason, Touretzky [1990], mentre Antonelli [1997] affronta il problema delle reti semantiche non monotone *cicliche*.

RIFERIMENTI BIBLIOGRAFICI

- Aldo Antonelli (1997), "Defeasible Inheritance on Cyclic Networks", *Artificial Intelligence* 92, pp. 1-23.
- Aldo Antonelli (1999), "A Directly Cautious Theory of Defeasible Consequence for Default Logic via the Notion of General Extensions", *Artificial Intelligence* 109, n. 1-2, pp. 71-109.
- Aldo Antonelli (2001), "Non-Monotonic Logic", in E. Zalta (a cura di), *The Stanford Encyclopedia of Philosophy* (<http://plato.stanford.edu>).
- Aldo Antonelli (2003), "Logic", in Luciano Floridi (a cura di), *Blackwell Guide to the Philosophy of Computing and Information*, Blackwell, Oxford.
- Ermanno Bencivenga (1984), *Il primo libro di logica*, Bollati Boringhieri, Torino.

- Margherita Benzi (1997), “Il ragionamento incerto. Probabilità e logica in Intelligenza Artificiale”, Angeli, Milano.
- Gerd Brewka, Jürgen Dix, Kurt Konolige (1997), *Nonmonotonic Reasoning: An Overview*, Computational Studies in Logic, Language and Information, Lecture Notes 73, Cambridge, Cambridge University Press.
- David W. Etherington (1987), “Formalizing nonmonotonic reasoning systems”, *Artificial Intelligence*, v. 31 pp. 41-85.
- David W. Etherington, Ray Reiter (1983), “On inheritance hierarchies with exceptions”, in *Proceedings of AAAI-83*, pp. 104-08.
- Gisèle Fischer Servi (2001), “Quando l'eccezione è la regola - Le logiche non monotone”, McGraw-Hill, Milano, 2001.
- Dov M. Gabbay (1985), “Theoretical foundations for nonmonotonic reasoning in expert systems”, in K. Apt (a cura di), *Logics and Models of Concurrent Systems*, pp. 439-459, Springer Verlag, Berlino e New York.
- Dov M. Gabbay, C. J. Hogger, J.A. Robinson (1994), (a cura di) *Handbook of Logic in Artificial Intelligence and Logic Programming*, vol. 3, Oxford University Press, Oxford.
- Matthew Ginsberg (1987), (a cura di) *Readings in Nonmonotonic Reasoning*, Morgan Kaufman, Los Altos, California.
- John F. Horty (1994), “Some direct theories of nonmonotonic inheritance”, in Gabbay, Hogger, Robinson (a cura di) [1994], pp. 111-187.
- John F. Horty (2002), “Skepticism and floating conclusions”, *Artificial Intelligence*, 135, pp. 55-72.
- John F. Horty (2002), “Nonmonotonic logic”, in Lou Goble (a cura di), *Blackwell Guide to Philosophical Logic*, pp. 336-61, Blackwell, Oxford.
- John F. Horty, Richmond H. Thomason, David S. Touretzky (1990), “A skeptical theory of inheritance in nonmonotonic semantic networks”, *Artificial Intelligence*, v. 42, pp. 311-48.
- Richmond H. Thomason, John F. Horty, David S. Touretzky (1987), “A calculus for inheritance in monotonic semantic nets”, in Z.W. Ras and M. Zemankove (a cura di), *Methodologies for Intelligent Systems*, pp. 280-87, Elsevier.
- John McCarthy (1977), “Epistemological problems of artificial intelligence”, in *Proceedings of the Fifth International Joint Conference on Artificial Intelligence*, MIT Press, Cambridge, Mass.
- John McCarthy (1980), “Circumscription – A Form of Non-Monotonic Reasoning”, *Artificial Intelligence* 13, pp. 27-39.
- Drew McDermott, Jon Doyle (1980), “Non-Monotonic Logic I”, *Artificial Intelligence* 13, pp. 41-72.
- David Makinson (1984), “General patterns in nonmonotonic reasoning”, in: Gabbay, Hogger e Robinson (1984).
- David Makinson, Karl. Schlechta (1991), “Floating conclusions and zombie paths: two deep difficulties in the “directly skeptical” approach to defeasible inheritance networks”, *Artificial Intelligence*, v. 48, pp. 199-209.
- Richmond H. Thomason (1989), (a cura di) *Philosophical Logic and Artificial Intelligence*, Kluwer Academic Press, Dordrecht.
- Ray Reiter (1980), A logic for default reasoning, *Artificial Intelligence*, v. 13, pp. 81-132.

- Ray Reiter, Giovanni Criscuolo (1981) “On interacting defaults”, in *Proceedings of the Seventh International Joint Conference on Artificial Intelligence*, pp. 270-276, Vancouver, B.C.
- Robert Stalnaker (1994), “Nonmonotonic consequence relations”, *Fundamenta Informaticæ*, v. 21 pp. 7-21.
- Karl Schlechta (1993) “Directly skeptical inheritance cannot capture the intersection of extensions”, *Journal of Logic and Computation*, v. 3, pp. 455-67.
- David Touretzky (1986), *The Mathematics of Inheritance Systems*, Morgan Kaufmann, Los Altos, California.