8. Logiche modali epistemiche e mondi possibili
8.1 Logica epistemica, onniscienza logica e intelligenza artificiale

La storia del trattamento logico degli atteggiamenti proposizionali, e, in particolare, del credere, può essere fatta risalire al lavoro di Hintikka del 1962 (Hintikka 1962), da cui prese le mosse gran parte della ricerca successiva in questo settore. In quell'opera la logica del credere veniva formulata per mezzo di un linguaggio di tipo modale, cui veniva associata una semantica non ancora definita in termini di mondi possibili, ma (in maniera sostanzialmente equivalente) per mezzo di strutture dette *model set*. La riformulazione in termini di mondi possibili avvenne più tardi, ad opera dello stesso Hintikka (1969), dopo la diffusione dei lavori di Kripke. L'individuazione dell'*onniscienza logica* come di un problema cruciale posto dalla formulazione model teoretiche della semantica per la logica epistemica è vecchio praticamente quanto la logica epistemica stessa, e numerosi tentativi di affrontarlo furono compiuti in ambito filosofico[1].

Con l'affermarsi dell'intelligenza artificiale di impostazione logica, lo studio del problema ricevette nuovi impulsi. In intelligenza artificiale il ragionamento di tipo, in senso lato, epistemico presenta molteplici motivi di interesse. Oltre al problema della rappresentazione del significato di enunciati di atteggiamento proposizionale nell'ambito dell'elaborazione del linguaggio naturale, il ragionamento sulle credenze è coinvolto nell'elaborazione di modelli di agenti artificiali che interagiscano con altri agenti, e che debbano quindi costruirsi rappresentazioni delle proprie e delle altrui credenze, e compiere inferenze su di esse. Nella progettazione di basi di conoscenza complesse è spesso necessario mettere in grado il sistema di ragionare sulle proprie conoscenze e sulle proprie capacità inferenziali. Da un punto di vista più specificamente informatico, le logiche della credenza possono essere applicate allo studio ed alla definizione dei sistemi di calcolo distribuiti ed alla computazione parallela. D'altro canto, poiché lo scopo dell'IA è quello di modellare le "competenze inferenziali" di soggetti razionali finiti e sottoposti a limiti di risorse realistici, le idealizzazioni imposte dal modello classico basato sui mondi possibili, prima fra tutte l'onniscienza logica, risultano del tutto inaccettabili[2]. Da un punto di vista tecnico, le logiche tradizionali per il ragionamento epistemico presentano seri problemi in una prospettiva computazionale. I formalismi che comportano l'assunzione dell'onniscienza logica presentano generalmente caratteristiche computazionali poco allettanti, nel senso che, anche nel caso in cui essi siano decidibili, non risultano in generale computazionalmente trattabili. Quindi, la ricerca di formalismi in cui le capacità inferenziali dei soggetti epistemici risultino più deboli è motivata anche dalla prospettiva di individuare strumenti di rappresentazione e di ragionamento dotati di migliori proprietà computazionali. Dal punto di vista di una teoria globale della razionalità, vi sono considerazioni in base alle quali per un soggetto finito l'onniscienza logica non risulta neppure un obiettivo ideale a cui è auspicabile avvicinarsi quanto più possibile. Gilbert Harman (1986) ad esempio sottolinea che, dal punto di vista logico, ogni credenza implica infinite conseguenze del tutto triviali, che non rivestono alcun interesse per gli scopi di un agente. Per un agente razionale finito sarebbe controproducente "ingombrare" la propria mente con tali credenze inutili. Nella prospettiva di una razionalità limitata, può essere svantaggioso destinare troppe risorse a fare inferenze da ciò che si crede, senza passare all'azione (Cherniak 1986).

I tentativi di affrontare il problema dell'onniscienza logica nell'ambito dell'intelligenza artificiale di impostazione logicista possono essere classificati, in prima approssimazione, in due grandi gruppi[3]. Il primo gruppo comprende quelle che si potrebbero definire soluzioni "sintattiche", il cui denominatore comune consiste nell'inserimento nella semantica della logica di entità di tipo sintattico quali enunciati e regole di derivazione. Il secondo gruppo comprende invece le soluzioni di tipo "semantico" in senso stretto, in cui i modelli sono definiti sulla base di strutture insiemistiche tradizionali. Fra le soluzioni di tipo semantico possono a loro volta essere distinte quelle che utilizzano mondi possibili classici, e quelle che fanno uso di strutture semantiche non classiche, quali mondi incoerenti o incompleti. Esistono poi soluzioni ibride o intermedie, che, ad esempio, adottano mondi possibili classici sovrapponendovi una sorta di "filtraggio" di tipo sintattico. Nei prossimi paragrafi passeremo in rassegna alcuni fra gli esempi più significativi di tali ricerche. Tutte queste linee presentano punti di collegamento con ricerche svolte in ambito filosofico, sulle quali ci soffermeremo di volta in volta[4].

---

[1]Si veda (Bäuerle e Cresswell 1988) per una rassegna.

[2]Questo aspetto è ben sintetizzato da Stalnaker (1991), che afferma "ogni tipo di elaborazione dell'informazione o di computazione è priva di senso se riferita a un agente onnisciente dal punto di vista deduttivo" (p. 429).

[3]Per questa distinzione si vedano ad esempio (Halpern 1986b) e (McArthur 1988).

[4]E' opportuno ricordare che la ricerca sul ragionamento epistemico in IA non è limitata al solo problema dell'onniscienza logica. Vi sono altri aspetti che, per quanto non vengano trattati se non marginalmente nelle pagine seguenti, presentano notevoli punti di interesse logico e filosofico. Un esempio è costituito dal ragionamento di tipo autoepistemico, che rientra nella linea di ricerca sul ragionamento non monotòno, e cui abbiamo già accennato nel par. 6.2. Un altro ambito di ricerca legato alla logica epistemica è quello in cui si studiano le credenze o le conoscenze di gruppi di attori, e le loro interazioni reciproche. Vengono quindi indagati concetti quali la *conoscenza comune* (*common knowledge*) o la conoscenza implicita (*implicit knowledge*) (si veda ad esempio Halpern e Moses 1984b). In generale, sui vari temi e sui problemi posti dal ragionamento epistemico in IA possono essere consultate le seguenti rassegne:

In questo lavoro, per semplicità, ci limiteremo ai soli calcoli proposizionali, in quanto i problemi dell'onniscienza logica si presentano in tutta la loro complessità già a questo livello. Affrontare il caso predicativo comporterebbe prendere in considerazione problemi ulteriori, estremamente complessi e di natura controversa, quali ad esempio il problema dell'interpretazione *de dicto* e dell'interpretazione *de re* dei contesti di atteggiamento proposizionale, e quello della quantificazione all'interno di tali contesti (*quantifying in*)[5]. Si tratta di temi in parte connessi a quello dell'onniscienza logica, ma che richiederebbero una trattazione indipendente, e che amplierebbero troppo i confini della presente ricerca. Inoltre, per alcune delle logiche descritte qui di seguito esiste esclusivamente una trattazione del caso proposizionale, e non è stata elaborata per il momento un'estensione al primo ordine.

Il presente capitolo e i seguenti sono strutturati come segue. Il par. 8.2 comprende un'introduzione sintetica alla semantica tradizionale di Kripke per la logica epistemica basata sui mondi possibili; in tale paragrafo vengono analizzate in dettaglio le fonti del problema dell'onniscienza logica. Il par. 8.3 tratta dei cosiddetti modelli minimali per le logiche modali, o modelli di Scott-Montague, che costituiscono un indebolimento della semantica di Kripke, ancorché elaborato nella tradizione dei mondi possibili classici. L'analisi dei modelli minimali per le logiche epistemiche consente di indagare, in un certo senso, le "possibilità estreme" delle semantiche classiche a mondi possibili nei confronti del problema dell'onniscienza logica. Nel cap. 9 vengono esaminate le cosiddette semantiche enunciative per le logiche della credenza, ossia quelle semantiche basate sull'introduzione di entità di tipo sintattico (ad esempio formule o enunciati) nell'interpretazione. In particolare, verrà presa in considerazione la logica della credenza elaborata da Konolige (1984, 1985a, 1986). Il cap. 10 tratta dell'utilizzo di *situazioni*, ossia di "mondi" non classici, eventualmente incoerenti o incompleti, per la semantica delle logiche della credenza. Il cap. 11, infine, verte sul lavoro di Fagin e Halpern sul ragionamento epistemico limitato. Questi autori argomentano che i vari modelli proposti in letteratura possono essere utilizzati per affrontare aspetti diversi del problema dell'onniscienza logica, e propongono tre tipi di logica che sviluppano differenti possibilità: una logica della consapevolezza che utilizza mondi incompleti, una logica della "consapevolezza generalizzata" che introduce elementi di tipo sintattico in una semantica modellistica tradizionale, e una logica per il ragionamento locale che mostra come alcune classi di modelli minimali possano modellare tipi interessanti di ragionamento limitato.

8.2 Strutture di Kripke per la logica del credere

Nelle logiche epistemiche di tipo modale i contesti epistemici, come quelli di credenza e di conoscenza, sono rappresentati mediante operatori modali, che assumono come argomenti formule del linguaggio della teoria. In questo paragrafo utilizzeremo principalmente due operatori, uno per il credere e uno per il conoscere, che indicheremo rispettivamente per mezzo dei simboli B e K (rispettivamente da *Belief* e *Knowledge*). Data una formula φ del linguaggio, l'espressione $B\varphi$ può essere letta come "è creduto che φ", e l'espressione $K\varphi$ può essere letta come "si sa che φ"[6]. Si è visto che l'idea che sta alla base della semantica dei mondi possibili consiste nel prendere in considerazione, per la valutazione delle espressioni della teoria logica che comprendano operatori di tipo intensionale, oltre che il mondo reale, anche altre possibili situazioni alternative, i *mondi possibili* appunto. Nel caso delle logiche epistemiche, la semantica a mondi possibili si basa sull'idea intuitiva che ad un soggetto epistemico sia associato un insieme di mondi, che corrispondono a tutte le situazioni compatibili con le credenze del soggetto epistemico stesso. Ad esempio, un soggetto epistemico può credere che la capitale d'Italia sia Roma e che la capitale degli Stati Uniti sia New York. In tal caso, in tutti i mondi possibili compatibili con le sue credenze sarà vero che la capitale d'Italia è Roma e che la capitale degli Stati Uniti è New York. Inoltre, lo stesso soggetto potrebbe non avere alcuna opinione circa il fatto che Atene sia o meno la capitale della Grecia. Allora, in alcuni dei mondi compatibili con le sue credenze sarà vero che Atene è la capitale della Grecia, mentre in altri sarà falso. Così, affinché sia vero che è creduta una certa formula α, cioè sia vera la formula $B\alpha$, allora α dovrà essere vera in tutti i mondi che il soggetto epistemico considera possibili. Affinché invece si sappia che α, cioè $K\alpha$ sia vera, α deve risultare vera, oltre che in tutti i mondi ritenuti possibili dal soggetto epistemico, anche nel mondo reale. La conoscenza viene infatti caratterizzata rispetto alla semplice credenza dal fatto di essere credenza vera[7]. Nell'esempio precedente, è plausibile che il soggetto epistemico *sappia* che Roma è la

(Halpern 1986b), (Halpern e Moses 1985), (McArthur 1988), (Parikh 1990b), (Halpern 1992), (Fagin et al. 1993), (Zanaboni 1992). Da alcuni anni a questa parte, la sede di discussione privilegiata per questo tipo di argomenti è il convegno *Theoretical Aspects of Reasoning about Knowledge* (*TARK*), che si tiene ogni due anni negli Stati Uniti. Questi convegni raccolgono contributi ampiamente interdisciplinari (ad esempio di logici, informatici, filosofi, economisti), di cui sono testimonianza i vari volumi di atti (Halpern 1986a; Vardi 1988; Parikh 1990a; Moses 1992).

[5] Su questi problemi in relazione alla credenza e agli atteggiamenti proposizionali si veda ad esempio (Linsky 1971).

[6] Per ora prendiamo in considerazione esclusivamente il caso in cui venga modellato il comportamento di un solo soggetto epistemico; vedremo in seguito come estendere questo tipo di trattamento ad un numero qualunque di soggetti.

[7] Che un certo fatto sia creduto e al tempo stesso sia vero costituisce una condizione necessaria ma non sufficiente per poter affermare che quel fatto è conosciuto (lo si potrebbe credere ad esempio per la ragione sbagliata). Dal punto di vista filosofico, caratterizzare il conoscere rispetto al credere pone problemi complessi: si veda ad esempio (Gettier

capitale d'Italia, mentre non è certamente vero che *sa* che la capitale degli Stati Uniti sia New York (perché ciò è falso nel mondo reale). Vedremo come vengono formalizzate queste intuizioni nella semantica a mondi possibili.

Formalmente, una *Struttura di Kripke*, o *Struttura a Mondi Possibili*, per la logica proposizionale è una terna $M = (W, \varphi, R)$[8]. W è un insieme di mondi possibili. φ è la funzione interpretazione che, per ogni mondo possibile, assegna un valore di verità ad ogni formula primitiva del linguaggio; vale a dire, per ogni mondo possibile $w \in W$ e per ogni formula primitiva p, $\varphi[w,p] \in \{\mathbf{v},\mathbf{f}\}$: $\varphi[w,p] = \mathbf{v}$ sse p è vera nel mondo w, e $\varphi[w,p] = \mathbf{f}$ sse p è falsa nel mondo w. φ assegna *uno ed un solo* valore ad ogni formula primitiva; ogni mondo possibile w è dunque coerente (non può accadere che una formula primitiva sia contemporaneamente vera e falsa in w) e completo (ogni formula primitiva ha un valore di verità in w). R è una relazione binaria di accessibilità fra i mondi di W. La semantica di Kripke per le logiche epistemiche si basa sul fatto che $R(w,w')$ vale se e solo se il mondo w' è compatibile con le credenze del soggetto epistemico nel mondo w. Proprietà diverse degli operatori di credenza o di conoscenza dipendono dalle differenti proprietà di cui può godere la relazione di accessibilità R.

Iniziamo col prendere in considerazione l'operatore modale per la conoscenza K. Definiamo innanzi tutto un linguaggio proposizionale per la conoscenza.

.L'*alfabeto* sia definito nel modo seguente:

a) un insieme Φ infinito numerabile di lettere proposizionali per formule primitive: p,q,r,p_1,p_2, ...;
b) i connettivi verofunzionali $\wedge$ e $\neg$;
c) l'operatore proposizionale K;
d) segni ausiliari: (,).

L'*insieme delle formule* viene definito come il più piccolo insieme tale che:

i. ogni lettera proposizionale è una formula;
ii. se α è una formula, allora $(\neg\alpha)$ è una formula;
iii. se α e β sono formule, allora $(\alpha \wedge \beta)$ è una formula;
iv. se α è una formula, allora $(K\alpha)$ è una formula.

Alcune parentesi si possono omettere imponendo le usuali condizioni di priorità sui connettivi. Gli altri connettivi verofunzionali si possono definire nel modo usuale. In particolare, si avrà:

$$\alpha \to \beta =_{def} \neg(\alpha \wedge \neg\beta)$$
$$\alpha \vee \beta =_{def} \neg(\neg\alpha \wedge \neg\beta)$$
$$\alpha \leftrightarrow \beta =_{def} (\alpha \to \beta) \wedge (\beta \to \alpha).$$

Ovunque nel seguito compaiano formule con connettivi logici diversi da $\wedge$ e $\neg$, le si deve intendere quali abbreviazioni di formule in cui compaiano esclusivamente $\wedge$ e $\neg$.

Data una interpretazione M, un mondo w e una formula α, l'espressione $M,w \models \alpha$ può essere letta "α è vera nel mondo w rispetto all'interpretazione M", oppure "il mondo w soddisfa α rispetto all'interpretazione M". Se si sottintende l'interpretazione, $w \models \alpha$ può essere letta "α è vera nel mondo w", oppure "il mondo w soddisfa α". Definiamo formalmente la relazione $\models$ come segue:

i. $M,w \models \alpha$ (per α appartenente all'insieme Φ delle formule primitive) sse $\varphi[w,\alpha]=\mathbf{v}$;
ii. $M,w \models \neg\alpha$ sse $M,w \not\models \alpha$;
iii. $M,w \models \alpha \wedge \beta$ sse $M,w \models \alpha$ e $M,w \models \beta$;
iv. $M,w \models K\alpha$ sse $M,w' \models \alpha$ in tutti i w' t.c. $R(w,w')$.

L'ultima clausola della definizione afferma che una formula del tipo $K\alpha$ è vera in un mondo possibile w se e soltanto se α è vera in tutti i mondi accessibili da w, cioè, secondo l'intuizione da cui siamo partiti, se α è vera in tutte le situazioni compatibili con le credenze del soggetto epistemico in w.

Diremo che una formula α è *valida* se e solo se $M,w \models \alpha$ per ogni interpretazione M e per tutti i mondi w in W. Diremo che una formula α è *soddisfacibile* se e solo se, per qualche interpretazione M, esiste un w t.c. $M,w \models \alpha$. Se α è

1963). Qui non approfondiremo tuttavia tali aspetti, e ci limiteremo a considerare la verità come una condizione necessaria perché si abbia conoscenza.

[8]Si tratta della stessa definizione data nel paragrafo 1.3, semplificata per il fatto che qui si ha a che fare con il caso proposizionale anziché con quello predicativo. Per questa ragione non viene specificato il dominio dell'interpretazione D.

valida, scriveremo $\models \alpha$. Da questa definizione di validità consegue che α è soddisfacibile sse $\neg\alpha$ non è valida. Possono inoltre essere dimostrate le seguenti proposizioni:

(1) Tutte le tautologie proposizionali sono formule valide.
(2) Date due formule qualunque α e β del linguaggio, la formula $(K\alpha \wedge K(\alpha \rightarrow \beta)) \rightarrow K\beta$ è valida.
(3) Date due formule qualunque α e β del linguaggio, se $\models \alpha$ e $\models \alpha \rightarrow \beta$, allora $\models \beta$.
(4) Per ogni formula α del linguaggio, se $\models \alpha$ allora $\models K\alpha$.

Diremo inoltre che una formula β è *conseguenza logica* di α se, per ogni interpretazione M e per ogni mondo w, ogni qual volta si ha che $M,w \models \alpha$, allora si ha anche che $M,w \models \beta$.

Vediamo ora quali restrizioni debbano essere imposte sulla relazione R affinché l'operatore K rifletta la nozione intuitiva di conoscenza. Abbiamo visto che, affinché il soggetto epistemico *sappia* che α, α deve essere creduta dal soggetto e deve inoltre essere vera. Quindi, affinché $K\alpha$ sia vera nel mondo w, α deve essere a sua volta vera in w. Ciò può essere garantito imponendo che la relazione R di accessibilità sia riflessiva, cioè che, per ogni $w \in W$, debba valere $R(w,w)$. In altri termini, dato qualsiasi mondo possibile, al soggetto epistemico deve sempre essere accessibile anche il mondo in cui esso stesso si trova.

Definiamo ora un sistema di assiomi che caratterizzi questo tipo di logica della conoscenza. Esso comprende i tre schemi di assiomi seguenti:

A1. Tutti gli assiomi del calcolo proposizionale.
A2. $K\alpha \wedge K(\alpha \rightarrow \beta) \rightarrow K\beta$
A3. $K\alpha \rightarrow \alpha$

e le due seguenti regole di inferenza:

R1. da α e $\alpha \rightarrow \beta$ segue β (*Modus ponens*)
R2. da α segue $K\alpha$ (Regola di necessitazione)

La regole R2 viene detta *regola di necessitazione* in quanto essa è stata elaborata originariamente rispetto alle logiche modali aletiche, dove viene utilizzata per introdurre l'operatore modale di necessità. L'assioma A2 può anche essere formulato in modo equivalente come segue:

$$K(\alpha \rightarrow \beta) \rightarrow (K\alpha \rightarrow K\beta),$$

e viene detto usualmente *assioma distributivo*.

Il sistema modale caratterizzato da questo apparato deduttivo va sotto il nome di sistema **T**. Diremo che una formula α è *dimostrabile* in **T**, in simboli $\vdash_{\mathbf{T}} \alpha$ (nel seguito adotteremo la norma di omettere **T** - o, in generale, il nome del sistema formale - quando ciò non crei ambiguità), se e soltanto se α è un assioma di **T**, oppure α è ottenibile mediante le regole R1 o R2 a partire da formule dimostrabili in **T**. Si noti l'analogia fra gli assiomi A1 e A2 e le proposizioni (1) e (2) sopra enunciate. Analogamente, le regole R1 e R2 sono il corrispettivo sintattico delle proposizioni (3) e (4). L'assioma A3 è ciò che caratterizza **T** come una logica della conoscenza: il fatto che una formula α sia conosciuta implica che α sia vera. A3 è il corrispettivo sintattico della riflessività della relazione R di accessibilità. Esso infatti impone sulle interpretazioni di **T** che, per ogni mondo w, ogni qual volta una formula sia vera in tutti i mondi accessibili da w, essa sia vera anche in w stesso. Si noti che la regola R2 asserisce che da α si può derivare $K\alpha$ se α è una formula *dimostrabile*, cioè se è un *teorema* della teoria, ottenibile a partire esclusivamente dagli assiomi logici, e non da eventuali assiomi propri (cioè assiomi specifici che non sono formule valide, ma che sono veri solo in alcune delle interpretazioni). Se si ammettesse che R2 sia applicabile anche a formule derivabili da assiomi non logici, allora R2 comporterebbe una sorta di "onniscienza fattuale" del tutto inaccettabile.

E' possibile dimostrare il seguente teorema:

Teorema: **T** è l'assiomatizzazione corretta e completa delle strutture di Kripke in cui la relazione di accessibilità sia riflessiva.[9].

In altri termini, se indichiamo con $\models_{\mathbf{T}} \alpha$ il fatto che la formula α è valida rispetto alle strutture a mondi possibili in cui R è riflessiva, si ha, per ogni formula α, che $\models_{\mathbf{T}} \alpha$ sse $\vdash_{\mathbf{T}} \alpha$.

[9]Per le dimostrazioni dei risultati enunciati in questo paragrafo si può far riferimento a manuali di logica modale, come ad esempio (Hugues e Cresswell 1968) o (Chellas 1980).

Le proprietà dell'operatore K possono essere ulteriormente specificate a livello sintattico aggiungendo altri assiomi, cui corrispondono, nell'interpretazione, ulteriori restrizioni sulla relazione di accessibilità R. Uno schema di assiomi che spesso viene ritenuto adeguato per la conoscenza è il seguente:

$$\text{A4. } K\alpha \rightarrow KK\alpha.$$

Si tratta del cosiddetto *assioma dell'introspezione positiva*, che impone che, ogni qualvolta il soggetto epistemico conosca α, egli a sua volta sappia di sapere α. Il sistema modale che comprende tutti gli assiomi e le regole di **T** più A4 viene detto **S4**. Semanticamente, A4 equivale ad imporre che R sia transitiva, vale a dire che, per ogni $w,w',w''\in W$, se $R(w,w')$ e $R(w',w'')$, allora $R(w,w'')$. A4 asserisce infatti che condizione necessaria perché $K\alpha$ sia vera in un mondo w è che in w sia vera $KK\alpha$. $KK\alpha$ è vera in w sse α è vera in tutti i mondi $w_1, ..., w_n$ accessibili dai mondi che sono accessibili da w. Quindi non deve essere possibile che α sia falsa in uno dei $w_1, ..., w_n$ e che contemporaneamente $K\alpha$ sia vera in w. Non deve cioè succedere che esista un mondo w_i accessibile da un mondo accessibile da w senza che w_i sia accessibile da w stesso. Nei modelli di **S4** la relazione di accessibilità R deve quindi essere, oltre che riflessiva, transitiva. E' possibili dimostrare il seguente teorema:

Teorema: **S4** è l'assiomatizzazione corretta e completa delle strutture di Kripke in cui la relazione di accessibilità sia riflessiva e transitiva.

Cioè, se indichiamo con $\models_{\mathbf{S4}} \alpha$ il fatto che la formula α è valida rispetto alle strutture a mondi possibili in cui R è riflessiva e transitiva, si ha, per ogni formula α, che $\models_{\mathbf{S4}} \alpha$ sse $\vdash_{\mathbf{S4}} \alpha$.

Un ulteriore schema di assiomi che può essere imposto su K è il seguente:

$$\text{A5. } \neg K\alpha \rightarrow K\neg K\alpha,$$

che impone che il soggetto epistemico sappia di non conoscere ciò che non conosce. A5 è detto usualmente *assioma dell'introspezione negativa*. Il sistema che comprende gli assiomi e le regole di **S4** più A5 viene detto sistema **S5**. Semanticamente, A5 equivale ad imporre che la relazione di accessibilità fra mondi sia euclidea. Una relazione R si dice *euclidea* se vale quanto segue: per ogni x,y,z, se $R(x,y)$ e $R(x,z)$, allora $R(y,z)$. Intuitivamente, il fatto che R sia euclidea serve ad evitare che si verifichino situazioni come quella in fig. 8.1. Nella figura ogni nodo del grafo è un mondo, e gli archi rappresentano la relazione di accessibilità. Nei mondi w_1 e w_2 è vera la formula α, mentre in w_3 è vera $\neg\alpha$. Dato un mondo w, A4 vuole escludere che in w sia vera $\neg K\alpha$, e che al tempo stesso non sia vera $K\neg K\alpha$, il che equivarrebbe ad avere un mondo accessibile da w in cui non valga $\neg K\alpha$. Questo è quanto accadrebbe appunto in fig. 8.1, dove, benché $\neg K\alpha$ sia vera in w, in w_1 è vera $K\alpha$. Se tuttavia si impone che R sia euclidea, w_3 deve essere accessibile a partire anche da w_1, nel quale quindi non sarebbe più vera $K\alpha$.

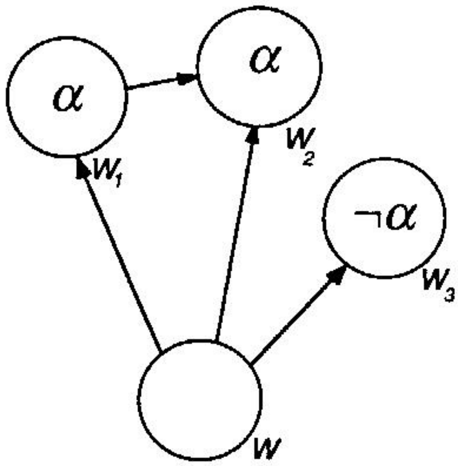

fig. 8.1

Nei modelli di **S5** R deve essere ad un tempo riflessiva, transitiva ed euclidea. Una relazione R si dice *simmetrica* se, per ogni x,y, $R(x,y)$ comporta $R(y,x)$. E' possibile dimostrare che ogni relazione che sia ad un tempo riflessiva ed euclidea è anche simmetrica. D'altro canto, ogni relazione simmetrica e riflessiva è anche euclidea. Una relazione che sia contemporaneamente riflessiva simmetrica e transitiva viene detta una relazione di *equivalenza*. Quindi, nei modelli di **S5** la relazione R di accessibilità è una relazione di equivalenza.

E' possibile dimostrare che:

Teorema: **S5** è l'assiomatizzazione corretta e completa delle strutture di Kripke in cui la relazione di accessibilità sia una relazione di equivalenza.

Cioè, se indichiamo con $\models_{\mathbf{S5}} \alpha$ il fatto che la formula α è valida rispetto alle strutture a mondi possibili in cui R è una relazione di equivalenza, si ha, per ogni formula α, che $\models_{\mathbf{S5}} \alpha$ sse $\vdash_{\mathbf{S5}} \alpha$.

Il comportamento logico dell'operatore modale B per la credenza è fondamentalmente analogo a quello di K, eccetto per il fatto che, perché una formula del tipo $B\alpha$ sia vera in un mondo possibile w, non è richiesto che α sia vera in w (è cioè possibile che il soggetto epistemico creda cose false). Definiamo un linguaggio logico proposizionale per la credenza in maniera del tutto analoga a quello per la conoscenza, eccetto per il fatto che la clausola c) per la definizione dell'alfabeto e la clausola iv. per la definizione delle formule vengono sostituite rispettivamente da due clausole c') e iv'. in cui il simbolo K viene sostituito dal simbolo B. Analogamente, la definizione della relazione $\models$ resta uguale, eccetto per il fatto che la clausola iv. viene sostituita dalla seguente:

iv'. $M,w \models B\alpha$ sse $M,w' \models \alpha$ in tutti i w' t.c. $R(w,w')$.

Le proposizioni (1), (2) e (3) sopra enunciate continuano a valere, mentre al posto della (4) vale la (4'), ottenuta sostituendo nella (4) B a K.

A livello sintattico, di teoria della dimostrazione, gli schemi di assiomi A1 e A2 (in cui tutte le occorrenze del simbolo K siano state sostituite dal simbolo B) sono validi per la logica della credenza. Analogamente, le regole R1 e R2 continuano ad essere corrette rispetto alla semantica data. Poiché, a differenza di ciò che si sa, ciò che si crede non deve necessariamente essere vero, nelle strutture di Kripke per la credenza il mondo reale non deve necessariamente essere accessibile per il soggetto epistemico. Cade quindi il vincolo che la relazione di accessibilità fra mondi possibili sia riflessiva. Analogamente, a livello sintattico, non vale il corrispettivo dell'assioma A3. D'altra parte, se si vuole che il soggetto epistemico non abbia credenze contraddittorie, si deve aggiungere il seguente assioma:

P. $\neg B(\alpha \wedge \neg\alpha)$.

(Si noti che, per come è definita la semantica, un soggetto epistemico che creda una contraddizione crederebbe qualunque cosa; cioè, per ogni formula ben formata α del linguaggio, varrebbe $B\alpha$). Nelle strutture di Kripke, l'assioma P equivale ad imporre che la relazione R di accessibilità sia seriale. Una relazione R è *seriale* sse, per ogni x, esiste un y, tale che $R(x,y)$. Il fatto che la relazione di accessibilità sia seriale impone che da ogni mondo possibile sia accessibile sempre almeno un mondo. Questo, dato che i mondi sono coerenti, evita che il soggetto epistemico possa credere una contraddizione[10].

Chiamiamo **KP** il sistema caratterizzato dagli assiomi A1, A2 e P, e dalle regole R1 e R2. Si può dimostrare che:

Teorema: **KP** è l'assiomatizzazione corretta e completa delle strutture di Kripke in cui la relazione di accessibilità sia seriale.

Lo schema di assiomi A4 dell'introspezione positiva può essere riformulato per l'operatore di credenza B:

A4. $B\alpha \rightarrow BB\alpha$.

Il sistema formale che comprende l'apparato deduttivo di **KP** più A4 viene detto **S4** debole. Come per **S4**, A4 equivale ad imporre che la relazione di accessibilità sia transitiva, per cui si può dimostrare che:

[10]Si noti che, nelle logiche per la conoscenza, l'equivalente di P segue facilmente da A3, e non deve quindi essere aggiunto come assioma:

1) $K(\alpha \wedge \neg\alpha) \rightarrow (\alpha \wedge \neg\alpha)$
2) $\neg(\alpha \wedge \neg\alpha) \rightarrow \neg K(\alpha \wedge \neg\alpha)$
3) $\neg(\alpha \wedge \neg\alpha)$
4) $\neg K(\alpha \wedge \neg\alpha)$.

1) è un esempio di A3; 2) è ottenuto da 1) per contrapposizione; 3) è una tautologia proposizionale, e 4) e ottenuto da 2) e 3) per *Modus Ponens*. Analogamente, per quanto riguarda la semantica, ogni relazione riflessiva è a sua volta seriale: per ogni $w \in W$, il fatto che esista almeno un w' t.c. $R(w,w')$ è soddisfatto in quanto $R(w,w)$. Quindi, nei modelli delle logiche che comprendono lo schema A3, la relazione di accessibilità R è seriale.

Teorema: **S4** debole è l'assiomatizzazione corretta e completa delle strutture di Kripke in cui la relazione di accessibilità sia seriale e transitiva.

Infine, riformulando in termini di credenza lo schema di assiomi per l'introspezione negativa:

$$\text{A5. } \neg B\alpha \rightarrow B\neg B\alpha,$$

e aggiungendolo all'apparato deduttivo di **S4** debole si ottiene il sistema formale **S5** debole. Anche in questo caso, A5 comporta che la relazione di accessibilità sia euclidea, per cui vale il seguente teorema:

Teorema: **S5** debole è l'assiomatizzazione corretta e completa delle strutture di Kripke in cui la relazione di accessibilità sia seriale, transitiva ed euclidea.

In **S5** debole risulta valido lo schema lo schema $BB\alpha \rightarrow B\alpha$, il quale, assieme allo schema di assiomi A4, consente di eliminare tutte le iterazioni dell'operatore modale B. Vale a dire, in **S5** debole tutte le formule del tipo $B \ldots B\ \alpha$ risultano equivalenti a $B\alpha$. Per quanto riguarda le logiche della conoscenza, lo schema corrispondente $KK\alpha \rightarrow K\alpha$ vale già in **K**, in quanto si tratta di un esempio dello schema di assiomi A3.

In ambito filosofico vi è stato un ampio dibattito su quanto i principi di introspezione positiva e negativa fossero adeguati a caratterizzare i concetti di conoscenza e di credenza. Tuttavia, tali problemi non sono rilevanti in questa sede, e rimandiamo chi fosse interessato a (Lenzen 1978).

Quanto visto sino ad ora può essere generalizzato al caso in cui i soggetti epistemici siano in numero maggiore di uno. Il linguaggio di una logica della conoscenza o della credenza per n soggetti epistemici comprenderà n operatori distinti K_i, o, rispettivamente, B_i (con $1 \leq i \leq n$). Una struttura di Kripke per una logica epistemica a n attori è una $n + 2$-pla $M = (W, \varphi, R_1, \ldots, R_n)$, dove $R_1, \ldots, R_n$ sono le relazioni di accessibilità relative a ciascun soggetto. La definizione della relazione di soddisfacibilità $\models$ resta identica alla precedente, con l'unica differenza che la clausola iv. deve essere modificata in modo da tenere conto degli n operatori epistemici. Ad esempio, nel caso della credenza, si avrà:

$$\text{iv''. } M,w \models B_i\alpha \text{ sse } M,w' \models \alpha \text{ in tutti i } w' \text{ t.c. } R_i(w,w'), \text{ con } 1 \leq i \leq n.$$

Cioè, il soggetto i-esimo crede l'enunciato α nel mondo w se e solo se α è vero in tutti i mondi *a lui accessibili* da w, cioè in tutti i mondi che sono nella relazione R_i con w. Analogamente, assiomi e regole di inferenza devono essere generalizzati a n attori. Ad esempio, lo schema A2 per la credenza diventerà:

$$\text{A2'. } B_i\alpha \wedge B_i(\alpha \rightarrow \beta) \rightarrow B_i\beta, \qquad \text{con } i = 1, \ldots, n.$$

Per ogni sistema formale sopra descritto, esistono i corrispettivi sistemi ad n soggetti epistemici. Ad esempio, al sistema **S4** corrispondono i sistemi $\mathbf{S4_n}$, e così via. Per tali sistemi valgono le generalizzazioni dei teoremi di correttezza e completezza sopra enunciati[11].

Il problema con la semantica dei mondi possibili per il trattamento dei contesti epistemici nasce dallo schema di assiomi A2 e dalla regola di necessitazione (R2). Semanticamente, A2 equivale ad affermare che l'insieme delle credenze di un soggetto epistemico è chiuso rispetto alla relazione di conseguenza logica: un soggetto deve credere tutte le conseguenze logiche delle sue credenze. R2 comporta invece che debbano essere credute tutte le formule valide,

[11]Un aspetto particolarmente interessante legato alle logiche del credere che prevedono più soggetti epistemici sono i problemi legati alla *conoscenza comune* (*common knowledge*), che non affronteremo in questa sede, ma che meritano qualche rapido cenno. Data una formula α, si dice che α è conoscenza comune di un gruppo di attori quando vale che:

1) ogni attore sa che α;
2) ogni attore sa che ogni attore sa che α;
3) ogni attore sa che ogni attore sa che ogni attore sa che α;

e così via. L'interesse per i problemi della conoscenza comune è nato con un libro del filosofo David Lewis (1969). Lewis ha mostrato che, affinché in un gruppo valga una convenzione (ad esempio, una convenzione di tipo linguistico), essa deve essere conoscenza comune fra i membri del gruppo. Ciò richiede infiniti livelli di conoscenza, e sorge quindi il problema di come le convenzioni possano essere istituite. Questo tipo di ricerche ha avuto interessanti riscontri in varie discipline, fra cui la teoria dei giochi, l'informatica, l'economia e le scienze sociali. Per maggiori informazioni sui problemi della conoscenza comune si rimanda a (Geanakoplos 1992) e alle varie rassegne sul ragionamento sulla conoscenza citate nel par. 8.1.

ossia, ad esempio, tutte le verità logiche della logica classica. I soggetti epistemici devono quindi essere *logicamente onniscienti*. Il punto è che l'*onniscienza logica* è inevitabile all'interno del paradigma classico della semantica a mondi possibili. Si consideri il sistema formale definito dagli schemi di assiomi e dalle regole seguenti (definito per l'operatore di credenza B e per un unico soggetto epistemico).

Schemi di assiomi:

A1. Tutti gli assiomi del calcolo proposizionale.
A2. $B\alpha \wedge B(\alpha \rightarrow \beta) \rightarrow B\beta$.

Regole di inferenza:

R1. da α e $\alpha \rightarrow \beta$ segue β (*Modus ponens*)
R2. da α segue $B\alpha$ (Regola di necessitazione)

Tale sistema viene chiamato usualmente **K**. Esso comprende, oltre agli assiomi e alle regole (*modus ponens*) per il calcolo proposizionale classico, soltanto lo schema A2 e la regola di inferenza R2, responsabili dell'onniscienza logica. E' possibile dimostrare che:

Teorema: **K** è l'assiomatizzazione corretta e completa delle strutture di Kripke con relazione di accessibilità qualsiasi.

Vale a dire, qualsiasi siano le restrizioni imposte sulla relazione R di accessibilità fra mondi possibili, gli schemi di assiomi A1 e A2 restano validi e le regole R1 e R2 continuano ad essere corrette nelle strutture di Kripke. **K** è per così dire la base da cui partire per ogni sistema di logica epistemica basato sulla semantica dei mondi possibili tradizionale. La causa di tutto questo risiede nel fatto che i mondi sono logicamente coerenti e completi: la funzione φ assegna, per ciascun membro di W, uno ed un solo valore di verità (**v** o **f**) a ciascuna formula atomica. Di conseguenza, tutte le verità logiche sono vere in ogni mondo possibile. Quindi, dato un mondo w, qualunque sia l'insieme dei mondi accessibili da esso, le verità logiche valgono in ciascuno di essi, e, per come è stata definita la semantica di B (e di K), esse devono essere credute da ogni soggetto epistemico. Analogamente, sempre perché i mondi sono coerenti e completi, se in tutti i mondi accessibili da un mondo w è vero α ed è vero al tempo stesso $\alpha \rightarrow \beta$, allora in ciascuno di essi sarà vero anche β. Quindi l'insieme delle credenze di un soggetto epistemico deve essere chiuso rispetto all'implicazione logica.

Connessa con A2 e R2 è la validità in **K** del seguente schema:

$$B\alpha \wedge B\beta \leftrightarrow B(\alpha \wedge \beta).$$

Vale a dire, l'operatore B è distributivo rispetto alla congiunzione. B non è tuttavia distributivo rispetto alla disgiunzione: $B(\alpha \vee \beta)$ non implica $B\alpha \vee B\beta$ in nessuno dei sistemi qui esaminati (anche se l'implicazione inversa è valida già in **K**). Analogamente, $\neg B\alpha$ non implica $B\neg\alpha$ (mentre l'implicazione inversa, $B\neg\alpha \rightarrow \neg B\alpha$, è valida nei sistemi con relazione di accessibilità seriale). Così, mentre $B\alpha \vee \neg B\alpha$ e $B(\alpha \vee \neg\alpha)$ sono entrambe valide in **K**, in nessuno dei sistemi visti è valida $B\alpha \vee B\neg\alpha$ (considerazioni analoghe valgono ovviamente anche per l'operatore K di conoscenza)[12].

Prima di passare a descrivere i vari approcci sviluppati nell'ambito dell'intelligenza artificiale di impostazione logica per affrontare il problema dell'onniscienza logica, dedicheremo il prossimo paragrafo alla trattazione dei *modelli minimali* per le logiche modali. Tali modelli generalizzano la semantica a mondi possibili per le logiche modali, e consentono, per quanto qui ci concerne, di ottenere teorie logiche della credenza in cui il problema dell'onniscienza logica risulti attenuato.

---

[12]Una formulazione alternativa della semantica per le logiche epistemiche è data dalle *knowledge structure* proposte da Fagin, Halpern e Vardi (Vardi 1985, Hamilton e Delgrande 1989, Fagin, Halpern e Vardi 1991). Intuitivamente, una *knowledge structure* costituisce l'analogo di una struttura di Kripke, e consiste in una gerarchia di *livelli*, dove il livello 0 è un modello di ciò che è vero nel mondo reale (nel caso proposizionale corrisponde a un assegnamento di valori di verità alle lettere proposizionali del linguaggio), il livello 1 è un modello di ciò che gli agenti epistemici direttamente credono a proposito del mondo reale, il livello 2 è un modello di ciò che gli agenti credono a proposito delle credenze del livello 1, e così via. Si dimostra che le *knowledge structure* sono equivalenti alle strutture di Kripke per le logiche epistemiche proposizionali. Inoltre, Vardi (1986) estende l'approccio delle *knowledge structure* in maniera da rendere conto delle logiche dei modelli minimali che tratteremo nel prossimo paragrafo, e Hamilton e Delgrande le estendono alle logiche epistemiche proposte da Levesque e Lakemayer che descriveremo nel cap. 10.

8.3 Modelli minimali e ragionamento epistemico

I *modelli minimali* (o *modelli di Scott-Montague*, o *neighborood model*) generalizzano la semantica a mondi possibili per le logiche modali, in maniera tale che le strutture di Kripke ne costituiscano un caso particolare. Essi sono stati studiati originariamente da Montague (1968) e Scott (1970). Montague (1970) li ha proposti come una possibile formalizzazione per una logica della credenza che superasse alcune delle difficoltà che le strutture di Kripke comportano per la logica epistemica. I modelli minimali forniscono le logiche della credenza più deboli dal punto di vista inferenziale che sia possibile elaborare all'interno del paradigma dei mondi possibili classici. Al di là della loro adeguatezza nel descrivere le prestazioni di agenti razionali limitati reali, i modelli minimali sono quindi interessanti per individuare i limiti che l'approccio a mondi possibili pone da questo punto di vista. La presente esposizione è basata fondamentalmente su Chellas (1980). Una formulazione in termini epistemici della logica dei modelli minimali si può trovare in Vardi (1986).

Nella logica dei modelli minimali le credenze di un soggetto epistemico non sono chiuse rispetto alla conseguenza logica, e possono essere creduti enunciati contraddittori. Come vedremo, la parte modale della logica che si ottiene in questo modo è estremamente debole dal punto di vista inferenziale. Di fatto, l'unico vincolo imposto dai modelli minimali sulle credenze è che, se è creduto un certo enunciato α, allora devono essere creduti tutti gli enunciati logicamente equivalenti a α. Logiche meno deboli possono essere ottenute ponendo ulteriori restrizioni sui modelli.

L'idea intuitiva alla base dei modelli minimali per la logica epistemica consiste nel modellare le credenze di un soggetto come insiemi arbitrari di proposizioni, ossia di intensioni di enunciati. Un insieme di proposizioni è associato a ciascun soggetto epistemico in ciascun mondo possibile. Nessun vincolo viene imposto sulle proposizioni che fanno parte di ciascuno di questi insiemi. Abbiamo visto che, nella semantica a mondi possibili, le intensioni sono modellate come funzioni da mondi possibili a oggetti semantici nel dominio (par. 1.3). Dato un elemento del linguaggio (ad esempio una lettera predicativa, un termine individuale o un enunciato), la sua intensione è una funzione che associa ad ogni mondo possibile il suo riferimento in quel mondo. Così, ad esempio, in logica dei predicati, l'intensione di una lettera predicativa ad un argomento P è una funzione ψ_P da mondi possibili a sottoinsiemi del dominio. Per ogni mondo possibile w assunto come argomento, ψ_P produce come valore il sottoinsieme del dominio che costituisce l'estensione di P in w. Poiché l'estensione di un enunciato è costituita da un valore di verità, la *proposizione* corrispondente (vale a dire, l'intensione dell'enunciato) sarà una funzione da mondi possibili a valori di verità; in particolare sarà la funzione che, per ogni mondo w, associa a w il valore **v** (vero) se l'enunciato è vero in w, il valore **f** (falso) se l'enunciato è falso in w. Nel caso di enunciati, l'intensione può quindi essere caratterizzata in modo equivalente come un insieme di mondi possibili: l'insieme dei mondi in cui l'enunciato è vero. Vale a dire, data una struttura a mondi possibili M analoga a quelle descritte nel paragrafo precedente, e dato l'insieme W dei mondi possibili, l'intensione di un enunciato α può essere definita come l'insieme $\{w|\ M,w\models\alpha\}$. Poiché in quanto segue avremo a che fare esclusivamente con proposizioni (intensioni di enunciati), adotteremo questa seconda formulazione, che ci eviterà inutili complicazioni nella formulazione della semantica.

Premesso ciò, possiamo passare alla presentazione formale dei modelli minimali per la logica della credenza. Per semplicità, ci limiteremo al caso in cui si abbia un solo soggetto di credenza. Quanto segue è tuttavia facilmente generalizzabile a un numero qualunque di soggetti epistemici. Assumiamo che il linguaggio sia lo stesso definito nel paragrafo precedente. Seguendo Vardi (1986), chiamiamo *struttura di credenza* un modello minimale per la credenza. Una *struttura di credenza* è una terna $M = (W, \varphi, N)$, dove, come nelle strutture di Kripke, W è l'insieme dei mondi possibili, e φ è una funzione a due argomenti che assegna un valore di verità (**v** o **f**) alle proposizioni primitive del linguaggio in ogni mondo possibile. N è una funzione che assegna al soggetto epistemico un insieme di proposizioni in ciascun mondo possibile. Intuitivamente, le proposizioni assegnate da N in ciascun mondo sono le proposizioni credute dal soggetto in quel mondo. D'ora in avanti, data una formula α, indicheremo con $prop_\alpha$ la sua intensione. Quindi, date le proposizioni $prop_{\alpha_1}, ..., prop_{\alpha_n}$ e un mondo w, $N(w) = \{prop_{\alpha_1}, ..., prop_{\alpha_n}\}$ significa che $\alpha_1, ..., \alpha_n$ sono credute in w. Poiché, come abbiamo visto, ciascuna proposizione è modellata come un insieme di mondi possibili, N sarà una funzione da mondi possibili a insiemi di insiemi di mondi possibili; quindi, N sarà una funzione di tipo $N: W \rightarrow 2^{2^W}$. Definiamo ora la relazione $\models$ come segue:

i. $M,w \models \alpha$ (per α appartenente all'insieme Φ delle formule primitive) sse $\varphi[w,\alpha]=\mathbf{v}$
ii. $M,w \models \neg\alpha$ sse $M,w \not\models \alpha$
iii. $M,w \models \alpha \wedge \beta$ sse $M,w \models \alpha$ e $M,w \models \beta$
iv. $M,w \models B\alpha$ sse $\{w'|\ M,w' \models \alpha\} \in N(w)$.

Le clausole i.-iii. sono identiche a quelle corrispondenti per le strutture di Kripke standard definite nel par. 8.2 La clausola iv. riflette invece l'idea intuitiva che abbiamo enunciato sopra: una formula di tipo $B\alpha$ è vera in un mondo w se e soltanto se la sua intensione fa parte dell'insieme delle proposizioni credute in w.

Le definizioni di soddisfacibilità, di validità e di conseguenza logica sono analoghe a quelle introdotte nel paragrafo precedente per le strutture di Kripke standard.

Secondo Vardi (1986) i modelli minimali costituiscono, in un certo senso, l'analogo, in una prospettiva model teoretica tradizionale, delle semantiche per le logiche della credenza basate su insiemi di enunciati, che descriveremo nel prossimo capitolo. Come vedremo, nelle semantiche basati su enunciati si assume che gli oggetti della credenza non siano proposizioni, ma formule o enunciati, vale a dire oggetti di tipo sintattico. Ciò che è creduto da un soggetto è rappresentato come un insieme di formule che, come caso limite, può essere del tutto arbitrario. Analogamente, dice Vardi, nei modelli minimali le credenze di un soggetto sono rappresentate da un insieme potenzialmente del tutto arbitrario di proposizioni: per ogni w, non viene posta alcuna condizione sull'insieme $N(w)$ delle proposizioni credute. Unico vincolo, che discende direttamente dalla natura stessa della semantica a mondi possibili, è quello per cui, se due formule α e β sono logicamente equivalenti, allora, se in un'interpretazione è vera $B\alpha$, deve essere vera anche $B\beta$. Infatti, se α e β sono logicamente equivalenti, esse sono vere esattamente negli stessi mondi possibili. Vale cioè che:

$$\models \alpha \leftrightarrow \beta \text{ sse } prop_\alpha = prop_\beta .$$

Se α e β sono logicamente equivalenti, allora hanno la stessa intensione, e sono quindi del tutto indiscernibili dal punto di vista della semantica qui formulata.

La logica definita dalla classe di modelli sopra descritta non è chiusa né rispetto all'implicazione, né rispetto all'implicazione valida (cioè, se è creduto α, e se $\models \alpha \rightarrow \beta$, non deve essere creduto necessariamente β). In essa inoltre non devono essere credute tutte le formule valide. Riguardo a quest'ultimo punto, infatti, una formula del tipo $\neg B(\alpha \vee \neg \alpha)$ è soddisfacibile. Perché, in un'interpretazione M, si abbia che $M,w \models \neg B(\alpha \vee \neg\alpha)$ basta infatti che $W \notin N(w)$. L'insieme W di tutti i mondi possibili coincide infatti con l'intensione di tutte le tautologie; si ha quindi $W = prop_{a\vee\neg a}$. Non è valido quindi lo schema di formule seguente:

$$\text{N.} \qquad B(\alpha \vee \neg\alpha)$$

Non è neppure valida la regola R1 di necessitazione:

da α si inferisce $B\alpha$.

D'altra parte, il fatto che tutte le tautologie abbiano la stessa intensione comporta che, in questa logica, data una tautologia φ, se è vero $B\varphi$, allora, per ogni altra tautologia ψ, sia vero $B\psi$. Analogamente, possono essere credute delle contraddizioni: è soddisfacibile, ad esempio, $B(\alpha \wedge \neg\alpha)$. Perché, in un'interpretazione M, si abbia $M,w \models B(\alpha \wedge \neg\alpha)$ si deve avere che $\varnothing \in N(w)$. $\varnothing$ è infatti l'intensione di tutte le contraddizioni (che, appunto, non sono vere in alcun mondo possibile). Anche in questo caso, quindi, data una contraddizione φ, se $B\varphi$ è vera, allora, per ogni altra contraddizione ψ, si avrà che è vero $B\psi$. Inoltre, è possibile che siano credute formule fra loro contraddittorie, senza che venga creduta una contraddizione esplicita. Ad esempio, si può avere che $M,w \models B\alpha$ e $M,w \models B\neg\alpha$ senza che $M,w \models B(\alpha \wedge \neg\alpha)$: è sufficiente che sia $prop_\alpha$, sia $\overline{prop_\alpha}$ (cioè il complemento di $prop_\alpha$ rispetto a W) appartengano a $N(w)$ senza che tuttavia vi appartenga $\varnothing$. Si noti infatti che $prop_{\neg\alpha} = \overline{prop_\alpha}$. Per quanto concerne la chiusura rispetto all'implicazione,

$$B\alpha \wedge B(\alpha \rightarrow \beta) \rightarrow B\beta$$

non è una formula valida: perché essa non sia soddisfatta in un mondo w è sufficiente che $prop_\alpha$ e $prop_{\alpha\rightarrow\beta} \in N(w)$, senza che $prop_\beta$ appartenga a $N(w)$. Considerazioni analoghe valgono per la chiusura rispetto all'implicazione valida. E' facile constatare che anche i due schemi seguenti non sono validi nella logica delle strutture di credenza:

M. $B(\alpha \wedge \beta) \rightarrow (B\alpha \wedge B\beta)$

C. $(B\alpha \wedge B\beta) \rightarrow B(\alpha \wedge \beta)$.

Si noti infatti che, date due formula α e β, l'intensione della loro congiunzione coincide con l'intersezione delle due formule, vale a dire: $prop_{\alpha\wedge\beta} = prop_\alpha \cap prop_\beta$. Ora, dato un mondo $w \in W$, non è affatto detto che, se $prop_{\alpha\wedge\beta}$ appartiene a $N(w)$, vi appartengano anche $prop_\alpha$ e $prop_\beta$; né che, se vi appartengono $prop_\alpha$ e $prop_\beta$, vi appartenga anche $prop_{\alpha\wedge\beta}$. Infine, è possibile che in un mondo w non venga creduto nulla, cioè che, per ogni formula α, si abbia $M,w \not\models B\alpha$. Perché accada ciò, è sufficiente che si abbia $N(w) = \varnothing$ (si noti che, ovviamente, $N(w) = \varnothing$ non è la stessa cosa di $\varnothing \in N(w)$).

La logica delle strutture di credenza quindi, non ponendo vincoli sull'insieme delle proposizioni credute da un soggetto epistemico, corrisponde a un modello di attore epistemico con capacità inferenziali praticamente nulle. Per quanto concerne la sua assiomatizzazione, abbiamo visto che l'unico vincolo sulla parte modale della logica consiste nel fatto che, qualora venga creduta una formula, devono essere credute tutte le formule ad essa equivalenti. Per quanto concerne la parte non modale, infine, è facile constatare sulla base delle regole semantiche date sopra, che essa coincide con il calcolo proposizionale classico.

Possiamo quindi definire un sistema formale, che chiameremo **E** (Chellas 1980), il cui apparato deduttivo sia caratterizzato come segue:

Gli *assiomi* siano tutti gli assiomi del calcolo proposizionale.

Come regole di inferenza, avremo il *modus ponens* e la regola RE così definita:

RE. da $\alpha \leftrightarrow \beta$ si inferisce $B\alpha \leftrightarrow B\beta$

Si può dimostrare il seguente teorema:

Teorema: Il sistema formale **E** è l'assiomatizzazione corretta e completa della logica delle strutture di credenza[13].

Logiche più potenti si possono ottenere ponendo vari vincoli sull'insieme di proposizioni associate al soggetto epistemico in ciascun mondo possibile. Prenderemo ora in considerazione alcune di tali restrizioni.

Un modello minimale M si dice *chiuso rispetto ai soprainsiemi* se e soltanto se, per ogni mondo $w \in W$ e per ogni proposizione (cioè per ogni insieme di mondi) X e Y in M vale la seguente condizione:

(m) se $X \subseteq Y$ e $X \in N(w)$, allora $Y \in N(w)$.

Si noti che la condizione (m) può essere espressa in maniera equivalente come segue, nei termini di intersezione insiemistica:

(m') se $X \cap Y \in N(w)$, allora $X \in N(w)$ e $Y \in N(w)$.

E' facile constatare che nelle strutture di credenza chiuse rispetto ai soprainsiemi è valido lo schema di formule M sopra enunciato; infatti, per ogni mondo possibile w e per ogni modello M, se $prop_{\alpha\wedge\beta} = prop_\alpha \cap prop_\beta \in N(w)$, allora si avrà che $prop_\alpha \in N(w)$ e $prop_\beta \in N(w)$.

Sia **EM** il sistema formale ottenuto aggiungendo M ad **E** come schema di assiomi. Si può allora dimostrare il seguente teorema:

Teorema: Il sistema formale **EM** è l'assiomatizzazione corretta e completa della logica delle strutture di credenza chiuse rispetto ai soprainsiemi.

Un modello minimale M si dice *chiuso rispetto all'intersezione* se e soltanto se, per ogni mondo $w \in W$ e per ogni proposizione (cioè per ogni insieme di mondi) X e Y in M, vale la seguente condizione:

(c) se $X \in N(w)$ e $Y \in N(w)$, allora $X \cap Y \in N(w)$.

Nelle strutture di credenza chiuse rispetto all'intersezione vale lo schema di formule C. Infatti, se $prop_\alpha \in N(w)$ e $prop_\beta \in N(w)$, allora $prop_{\alpha\wedge\beta} \in N(w)$.

Sia allora **EC** il sistema formale ottenuto aggiungendo C ad **E** come schema di assiomi. Si può allora dimostrare il seguente teorema:

Teorema: Il sistema formale **EC** è l'assiomatizzazione corretta e completa della logica delle strutture di credenza chiuse rispetto all'intersezione.

Si dice che un modello minimale *M contiene l'unità* se e soltanto se, per ogni mondo $w \in W$, vale che:

[13]Per la dimostrazione di questo risultato e dei risultati enunciati nel seguito di questo paragrafo si rimanda a (Chellas 1980), cap. 7.

$$\text{(n)} \qquad W \in N(w).$$

Vale a dire, la proposizione costituita da tutti i mondi possibili deve essere creduta in tutti i mondi. Questo significa che in ogni mondo possibile sono credute tutte le formule valide (tutte le formule valide sono infatti logicamente equivalenti). Cioè, nelle strutture di credenza contenenti l'unità vale lo schema di formule N.

Sia allora **EN** il sistema formale ottenuto aggiungendo N ad **E** come schema di assiomi. Si può allora dimostrare che:

Teorema: Il sistema formale **EN** è l'assiomatizzazione corretta e completa della logica delle strutture di credenza che contengono l'unità.

Ovviamente, tali condizioni sui modelli possono essere variamente combinate. In particolare, chiameremo *filtro* un modello minimale che rispetti contemporaneamente le tre condizioni (m), (n) e (c). Sia **EMNC** il sistema formale ottenuto aggiungendo ad **E** gli schemi di assiomi M, N e C. Si dimostra che:

Teorema: Il sistema formale **EMNC** è corretto e completo rispetto all'insieme dei filtri.

E' possibile dimostrare che **EMNC** è un'assiomatizzazione equivalente del sistema modale **K** descritto nel paragrafo precedente. La logica delle strutture di credenza risulta in questo modo essere una generalizzazione della logica delle strutture standard a mondi possibili di Kripke, che ne sono casi particolari. In particolare, **K** risulta corretto e completo rispetto all'insieme dei filtri.

Vediamo, in maniera intuitiva, come gli assiomi e le regole modali di **K** risultino validi rispetto ai filtri. Come abbiamo visto, il sistema modale **K** è caratterizzato dal seguente apparato deduttivo:

Schemi di assiomi:

A1. Gli assiomi del calcolo proposizionale.
A2. $B\alpha \wedge B(\alpha \rightarrow \beta) \rightarrow B\beta$.

Regole di inferenza:

R1. da α e $\alpha \rightarrow \beta$ si inferisce β (*Modus ponens*)
R2. da α si inferisce $B\alpha$ (Necessitazione)

Iniziamo dalla regola di necessitazione. Che essa sia valida rispetto ai filtri segue facilmente dalla condizione (n) sopra enunciata. La necessitazione consente di dedurre $B\alpha$ per ogni formula α valida. Se α è una formula valida, allora essa è vera in ogni mondo possibile, pertanto $prop_\alpha = W$. Ma, in base a (n), in ogni filtro la proposizione W è creduta in ogni mondo possibile, quindi, per ogni $w \in W$, $B\alpha$ è vera in w.

Per quanto riguarda la validità dello schema di assiomi A2, si deve mostrare che, per ogni filtro, ogni qual volta in un mondo possibile w siano vere $B\alpha$ e $B(\alpha \rightarrow \beta)$, allora in w è vero $B\beta$. Si noti innanzi tutto che, date due formule α e β, si ha quanto segue:

$$(*) \qquad prop_{\alpha\rightarrow\beta} = \overline{prop_\alpha} \cup prop_\beta$$

vale a dire, l'intensione di $\alpha \rightarrow \beta$ è l'unione dell'intensione di β con il complemento dell'intensione di α rispetto a W ($\alpha \rightarrow \beta$ è vera in tutti i mondi in cui è vera β oppure è falsa α). Si è detto inoltre che la proposizione di una congiunzione è uguale all'intersezione delle proposizioni dei due disgiunti. Quindi:

$$prop_{\alpha\wedge(\alpha\rightarrow\beta)} = prop_\alpha \cap prop_{\alpha\rightarrow\beta}.$$

Da (*) segue che:

$$prop_\alpha \cap prop_{\alpha\rightarrow\beta} = prop_\alpha \cap prop_\beta.$$

D'altra parte se $B\alpha$ è vera in w, allora $prop_\alpha \in N(w)$. Analogamente, se $B(\alpha \rightarrow \beta)$ è vera in w, allora $prop_{\alpha\rightarrow\beta} \in N(w)$. Quindi, per la condizione (c), $prop_{\alpha\wedge(\alpha\rightarrow\beta)} \in N(w)$. E' poi ovvio che:

$$prop_{\alpha\wedge(\alpha\rightarrow\beta)} = prop_\alpha \cap prop_\beta \subseteq prop_\beta.$$

Possiamo dunque concludere, sulla base di (m) (nella sua formulazione (m')), che

$$prop_\beta \in N(w).$$

Quindi, $B\beta$ è vero in w, e lo schema di assiomi A2 è quindi valido in tutti i modelli di **EMNC**.

Ulteriori condizioni possono essere imposte sui modelli minimali affinché valgano i vari principi modali visti in precedenza. Ad esempio, affinché in una struttura di credenza M valga lo schema di assiomi P:

$$\text{P.} \qquad \neg B(\alpha \wedge \neg\alpha)$$

basta imporre che valga la seguente condizione:

$$\text{(p)} \qquad \text{per ogni } w \in W, \varnothing \notin N(w),$$

in base alla quale in nessun mondo possibile può essere creduta una formula contraddittoria (poiché $\varnothing$ è l'intensione di tutte le contraddizioni). Abbiamo già messo in luce che, in generale, nei modelli minimali è possibile credere formule fra loro in contraddizione senza credere esplicitamente alcuna contraddizione. Quindi, la condizione (p) da sola non è sufficiente di norma a evitare che siano credute formule fra loro in contraddizione. Questo non accade tuttavia nei modelli chiusi rispetto all'intersezione. Ricordiamo che, data una formula α, α e la sua negazione sono credute entrambe in un mondo w se e soltanto se sia $prop_\alpha$ che il suo complemento rispetto a W $\overline{prop_\alpha}$ appartengono a $N(w)$. Poiché $prop_\alpha \cap \overline{prop_\alpha} = \varnothing$, è evidente che (p) e (c) assieme vietano che si abbia sia $prop_\alpha \in N(W)$ che $\overline{prop_\alpha} \in N(W)$. Si noti infine che nelle strutture di credenza chiuse rispetto ai soprainsiemi non è possibile che sia creduta una formula contraddittoria senza che sia creduta qualsiasi formula. Infatti, se per qualche mondo w si ha che $\varnothing \in N(w)$, allora, poiché l'insieme vuoto è sottoinsieme di qualsiasi insieme, in base a (m') si ha che, per ogni $X \subseteq W$, $X \in N(w)$.

Per quanto riguarda l'introspezione positiva, affinché valga lo schema di assiomi:

$$\text{A4.} \qquad B\alpha \rightarrow BB\alpha$$

deve essere imposto sull'interpretazione che, per ogni proposizione $X \subseteq W$ e per ogni $w \in W$, valga quanto segue:

$$\text{(a4)} \qquad \text{se } X \in N(w), \text{ allora } \{w' | X \in N(w')\} \in N(w).$$

Si noti che, se X è l'intensione di una certa formula α, allora $\{w' | X \in N(w')\}$ è l'insieme dei mondi in cui si crede α, cioè l'intensione di $B\alpha$. Quindi la condizione (a4) impone che, per ogni w, se α è creduta in w, allora vi sia creduta anche $B\alpha$.

Affinché valga lo schema di assiomi dell'introspezione negativa:

$$\text{A5.} \qquad \neg B\alpha \rightarrow B\neg B\alpha,$$

si deve imporre la seguente condizione. Per ogni proposizione $X \subseteq W$ e per ogni $w \in W$, deve valere che:

$$\text{(a5)} \qquad \text{se } X \notin N(w), \text{ allora } \{w' | X \notin N(w')\} \in N(w).$$

Analogamente al caso precedente, se X è l'intensione di α, allora $\{w' | X \notin N(w')\}$ è l'intensione di $\neg B\alpha$. Quindi, se $X \notin N(w)$, cioè se α non è creduta in w, allora $\neg B\alpha$ è creduta in w.

Se si intende definire una logica della conoscenza, deve valere l'assioma:

$$\text{A3.} \qquad K\alpha \rightarrow \alpha.$$

A3 è valido nei modelli in cui, per ogni $w \in W$ e per ogni proposizione $X \subseteq W$,

$$\text{(a3)} \qquad \text{se } X \in N(w), \text{ allora } w \in X.$$

Vale a dire, ponendo che X sia l'intensione di una formula α, se nel mondo w si sa che α, allora α deve essere vera anche in w (cioè, il mondo w deve appartenere all'intensione di α).

Ponendo quindi restrizioni opportune sulle strutture di credenza, si possono ottenere logiche per la credenza e la conoscenza che corrispondono a vari tipi di capacità inferenziali di soggetti epistemici limitati. Vedremo nel seguito un esempio del genere (par. 11.3).

Tuttavia, l'assunzione che due enunciati logicamente equivalenti non siano distinguibili dal punto di vista del ragionamento epistemico è una idealizzazione troppo forte, che deve in qualche modo essere mitigata. Come vedremo, in questa direzione è stata sviluppata la maggior parte delle logiche che prenderemo in esame nei prossimi paragrafi.