

6. Frame, reti semantiche e ragionamento non monotono

E' facile constatare che un tipo di traduzione logica quale quella utilizzata per KL-ONE non è adeguata nei confronti dei sistemi a *frame* e delle reti semantiche che ammettono eccezioni. Si consideri ad esempio la coppia di *frame tigre* e *tigre da circo* del paragrafo 4.3. Lo *slot* *è un* del *frame tigre da circo* può essere plausibilmente tradotto per mezzo di un'implicazione quantificata di questo genere:

$$(1) \forall x (tigre_da_circo(x) \rightarrow tigre(x)),$$

cioè, tutte le tigri da circo sono tigri. Adottando una tecnica analoga a quella utilizzata per i ruoli di KL-ONE, ai due *slot habitat* dei *frame* corrisponderebbero le due seguenti formule di calcolo dei predicati:

$$(2) \forall xy (tigre(x) \wedge habitat(x,y) \rightarrow giungla(y))$$
$$(3) \forall xy (tigre_da_circo(x) \wedge habitat(x,y) \rightarrow circo(y)).$$

Tuttavia, da (1)-(3) è ora possibile dedurre che:

$$\forall xy (tigre_da_circo(x) \wedge habitat(x,y)) \rightarrow (giungla(y) \wedge circo(y)),$$

che certamente non corrisponde all'intuizione che si intendeva catturare. Se inoltre si assume che valga il ragionevole assioma:

$$(4) \forall x \neg(circo(x) \wedge giungla(x)),$$

(1)-(4) rendono contraddittorio assumere che una tigre da circo abbia un habitat.

Nel caso di una rete semantica in cui compaia un arco di tipo *isa* che ammetta delle eccezioni, tale arco non può essere tradotto mediante una implicazione quantificata universalmente come nel caso degli archi di superconcetto in KL-ONE. Prendiamo ad esempio una rete in cui compaia un arco *isa* fra il concetto *uccello* e il concetto *vola* (come nella fig. 6.1). Supponiamo di tradurlo mediante la formula seguente:

$$(5) \forall x (uccello(x) \rightarrow vola(x)).$$

In tal caso, l'esistenza di un uccello che non sia in grado di volare (ad esempio perché è un pinguino) genererebbe immediatamente una contraddizione. I tradizionali sistemi logici non consentono di affrontare in modo soddisfacente il problema. Si potrebbe pensare di aggirare la difficoltà indebolendo la premessa della (5), trasformandola ad esempio in:

$$(5') \forall x (uccello(x) \wedge \neg pinguino(x) \rightarrow vola(x)).$$

Tuttavia tale soluzione pone immediatamente difficoltà ulteriori. In primo luogo, i pinguini non sono certo l'unico caso di uccelli che non sono in grado di volare. Ci sono, ad esempio, anche gli struzzi. C'è il dodo. Ci sono gli uccelli appena nati. Ci sono gli uccelli legati, gli uccelli morti, gli uccelli con le ali spezzate, quelli paralitici, quelli ipnotizzati, e così via. In generale, non è possibile enumerare tutte le possibili eccezioni che ammette una regola che esprime una connessione di carattere "prototipico" fra due concetti¹. Inoltre, anche assumendo che questo sia possibile, il problema non sarebbe comunque risolto. Supponiamo infatti di avere ottenuto la regola seguente:

$$(5'') \forall x (uccello(x) \wedge \neg\alpha(x) \rightarrow vola(x)),$$

dove $\alpha(x)$ sia la disgiunzione di tutte le possibili "eccezioni", di tutti i possibili casi cioè in cui qualcosa è un uccello e tuttavia non vola. Ebbene, in tal caso, l'informazione che, poniamo, Tweety è un uccello non è sufficiente per inferire che Tweety vola. Per poter concludere che Tweety vola si deve disporre di tutte le informazioni relative al fatto che non si tratti di un uccello anomalo. Si deve cioè sapere che Tweety non è un pinguino, che non è uno struzzo, che non ha le ali rotte, e così via. Ciò non è certamente adeguato dal punto di vista di una forma di ragionamento basato su prototipi, che è appunto il tipo di ragionamento che *frame* e reti semantiche con eccezioni intendono catturare. Ciò che si desidera è che, in mancanza di informazioni più specifiche, si possa assumere che valgano le proprietà del prototipo. Ad esempio, sapendo che Tweety è un uccello, in mancanza di altre informazioni più specifiche, si vuole assumere che Tweety è un uccello tipico e quindi poter inferire che vola. Qualora diventassero disponibili altre informazioni più specifiche in contrasto con tale inferenza (ad esempio che Tweety è un pinguino) si vuole poter eliminare la conclusione tratta

¹Tale problema è una forma del cosiddetto *qualification problem* - si veda più oltre in questo paragrafo.

precedentemente. Questo tipo di esigenza è in contrasto con una caratteristica condivisa dalla quasi totalità dei sistemi logici tradizionali, vale a dire con la cosiddetta *proprietà di monotonia*. In un sistema logico monotono l'insieme delle conclusioni cresce monotonicamente al crescere del numero delle premesse. Vale a dire, aggiungendo nuove premesse l'insieme delle conclusioni può eventualmente crescere, ma non può diminuire; non può cioè succedere che aggiungendo nuove premesse debbano essere eliminate alcune delle conclusioni tratte precedentemente. In maniera più formale, la proprietà di monotonia può essere caratterizzata dicendo che un sistema logico L è *monotono* se vale quanto segue. Data una formula α e due insiemi Δ e Γ di formule di L , se α è derivabile in L da Δ , e se $\Delta \subset \Gamma$, allora α è derivabile in L da Γ . Dall'esempio sopra riportato risulta invece evidente che il ragionamento basato su prototipi è tipicamente non monotono.

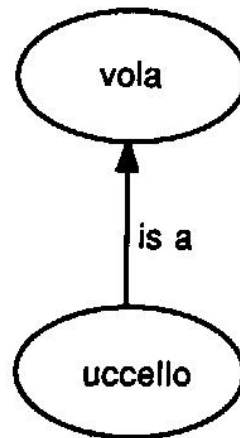


fig. 6.1

Lo sviluppo di sistemi logici che consentano di formalizzare il ragionamento non monotono è stato, fin dalla seconda metà degli anni settanta, uno dei settori di ricerca più attivi e più originali nell'ambito dell'IA di impostazione logica. Ciò è dovuto al fatto che la non monotonicità è una delle caratteristiche più diffuse del ragionamento di senso comune. Oltre al problema della rappresentazione prototipica di concetti, vi è ad esempio il caso del ragionamento su eventi e azioni. In questo ambito sorge il cosiddetto *frame problem*², messo in evidenza originariamente da McCarthy e da Hayes (McCarthy e Hayes 1969). In uno dei suoi aspetti, il *frame problem* riguarda la possibilità di fare previsioni sul futuro senza prendere in considerazione ogni evento passato. Si consideri una regola del tipo "se qualcosa è vero in un certo istante, allora qualcosa d'altro sarà vero ad un istante successivo". Anche qui abbiamo a che fare con il *qualification problem*: in ogni situazione reale, l'antecedente di una regola del genere dovrebbe essere di lunghezza praticamente infinita: ad esempio, per essere certi che una palla da biliardo rotoli in una certa direzione non basta sapere che è stata colpita in un certo modo e con una certa forza, ma anche che non ci siano fori o irregolarità sul piano, che sulla sua traiettoria non si vengano a trovare altre palle o altri oggetti, che rotolando non cambi la sua struttura fisica, e così via. Quindi, se non si vuole presupporre irrealisticamente una conoscenza perfetta del mondo, si deve utilizzare una forma di ragionamento non monotono: in mancanza di altre informazioni, si inferisce che, tipicamente, la palla andrà in una certa direzione; nel momento in cui nuove informazioni fossero disponibili (presenza di buchi sul piano, o altro) si devono poter eliminare le conclusioni tratte in precedenza. Si noti che, sebbene il *frame problem* sembri distante dal problema della rappresentazione del significato lessicale, di cui qui ci stiamo occupando, tuttavia è in relazione con esso. Ad esempio, nel momento in cui si cerca di definire il significato di molti verbi, ci si trova di fronte ad aspetti del *frame problem*. Si vedano gli esempi citati nel par. 4.2 a proposito della teoria di Schank: se si definisce il significato di spingere, si vuole poter rappresentare che quando si spinge qualcosa, di solito quella cosa si sposta, ma vi sono casi in questo può non accadere. Esistono quindi delle relazioni fra *frame problem*, valori di *default* e significato lessicale.

Tali problemi hanno portato all'elaborazione e allo studio di numerosi sistemi di logica non monotona³. Il settore è estremamente ricco e complesso, ed è del tutto impossibile in questa sede darne anche soltanto una panoramica soddisfacente. Ci limiteremo quindi a prendere in esame, nel prossimo paragrafo, il formalismo della *circumscription*.

²Si noti che l'espressione *frame problem* storicamente non ha nulla a che fare con i *frame* quali sono stati esposti nel paragrafo precedente: sebbene vi possano essere relazioni fra questi due concetti, l'espressione *frame* è stata introdotta nei due casi in maniera del tutto indipendente.

³La diffusione dell'interesse per le logiche non monotone nella comunità dei ricercatori di intelligenza artificiale può essere fatto risalire a un numero speciale della rivista *Artificial Intelligence* del 1980 (*Artificial Intelligence* 1980). Sulle logiche non monotone in generale si possono consultare (Lukaszewicz 1990); (Ginsberg 1987); (Brewka 1991); (Reinfrank et al. 1988).

Tale analisi ovviamente non pretende di essere esaustiva, ma ha esclusivamente lo scopo di mettere a disposizione alcuni elementi necessari per lo sviluppo della nostra argomentazione. Nel paragrafo successivo passeremo rapidamente in rassegna altri due fra gli approcci più diffusi alla logica non monotona: la *default logic* e le logiche non monotone modali. Daremo infine qualche cenno sulla *preferential semantics* di Shoham, una teoria che intende fornire una visione unificata dal punto di vista model teoretico dei vari approcci al ragionamento non monotono.

6.1 La formalizzazione del ragionamento non monotono: la *circumscription*

La *circumscription* è un approccio al ragionamento non monotono proposto originariamente da John McCarthy (1977, 1980, 1984)⁴. Vi sono predicati che valgono solo, per così dire, in casi eccezionali, mentre di solito non si applicano in situazioni tipiche. Nel ragionamento di senso comune normalmente si assume implicitamente che tali predicati non valgano a meno che non si sappia esplicitamente che non è così. Un esempio potrebbe essere il predicato *albino*. Le persone albine sono un sottoinsieme molto ristretto degli esseri umani. Solitamente nessuno, ragionando su di una persona che non conosce, prende in considerazione la possibilità che sia albino, a meno che non abbia qualche informazione esplicita in proposito. L'idea di fondo che sta alla base del formalismo della *circumscription* è che, in una teoria logica, si "circoscrivano" i predicati di questo genere, dove circoscrivere un predicato P in una teoria T vuol dire assumere che P abbia l'estensione più piccola possibile compatibilmente con le informazioni presenti in T stessa. Ad esempio, si supponga di avere la seguente teoria (scritta nel linguaggio del calcolo dei predicati del primo ordine con identità):

$$T = \{albino(Gigi), Piero \neq Gigi, Pippo \neq Gigi, Pippo \neq Piero\}.$$

Circoscrivere il predicato *albino* in T significa assumere che Gigi sia l'unica persona albina nel dominio della teoria. Circoscrivendo *albino* in T si deve cioè poter derivare:

$$\forall x(albino(x) \rightarrow x=Gigi).$$

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

$$\begin{aligned} &\neg albino(Piero) \\ &\neg albino(Pippo), \end{aligned}$$

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

$$T' = T \cup \{albino(Piero)\}.$$

Circoscrivendo *albino* in T' si otterrà:

$$\forall x(albino(x) \rightarrow x=Gigi \vee x=Piero),$$

per cui non sarà più possibile ottenere $\neg albino(Piero)$.

Per rappresentare regole non monotone mediante la *circumscription* si ricorre all'introduzione di un predicato *ab*, dove *ab* sta per *abnormal* (McCarthy 1984; Grosz 1984). Ad esempio, che gli uccelli di solito volano viene rappresentato mediante la formula seguente:

$$\forall x (uccello(x) \wedge \neg ab(x) \rightarrow vola(x)),$$

vale a dire, se qualcosa è un uccello e non è atipico, allora vola. Dopo di che, per ottenere le conseguenze volute, si circoscrive il predicato *ab*. Si consideri ad esempio la teoria seguente:

$$T = \{\forall x (uccello(x) \wedge \neg ab(x) \rightarrow vola(x)), uccello(Tweety)\}.$$

Date le informazioni in T , di nessun individuo si riesce a derivare che è *abnormal*. Circoscrivendo *ab* in T si otterrà quindi che:

⁴Sull'uso della *circumscription* per formalizzare *frame* e reti semantiche con eccezioni si vedano ad esempio (Brewka 1987); (Haugh 1988); (Krishnaprasad et al. 1989).

$$\forall x \neg ab(x).$$

Di conseguenza si avrà:

$$\forall x (uccello(x) \rightarrow vola(x))$$

e

$$vola(Tweety).$$

Supponiamo ora di passare da T alla seguente teoria T' :

$$T' = T \cup \{uccello(Fred), \neg vola(Fred), Fred \neq Tweety\}.$$

Da T' è possibile derivare

$$ab(Fred).$$

Quindi, circoscrivendo ab in T' avremo:

$$\forall x (ab(x) \rightarrow x=Fred),$$

e, di conseguenza,

$$\forall x (uccello(x) \wedge x \neq Fred \rightarrow vola(x)).$$

Si possono utilizzare predicati di *abnormality* a più di un argomento. Ad esempio, utilizzando un predicato a due argomenti $ab(x,y)$, si può esprimere una regola non monotona come la seguente

$$\forall xy (compleanno(x) \wedge amico(x,y) \wedge \neg ab(x,y) \rightarrow fa_regalo(y,x))$$

che può essere letta nel modo seguente: "se è il compleanno di x , se x e y sono amici, e se la loro non è una relazione di amicizia atipica, allora y fa un regalo a x ". Inoltre, in una teoria si possono circoscrivere più predicati contemporaneamente.

Dal punto di vista formale, effettuare la *circumscription* di alcuni predicati in una teoria T equivale ad aggiungere a T un opportuno assioma del secondo ordine. Esistono numerosi tipi diversi di *circumscription*, con diverso potere espressivo, e idonei a formalizzare tipi diversi di ragionamento non monotono. Iniziamo descrivendo brevemente la *predicate circumscription* (McCarthy 1980), che costituisce la forma di *circumscription* più semplice e meno potente. Introduciamo la seguente convenzione notazionale. Siano U e V due espressioni che denotano due predicati ad n argomenti. Indicheremo con $U \leq V$ il fatto che l'estensione di U è un sottoinsieme dell'estensione di V o coincide con essa. In altri termini, la relazione \leq può essere definita come segue:

$$U \leq V =_{def} \forall x_1, \dots, x_n (U(x_1, \dots, x_n) \rightarrow V(x_1, \dots, x_n)).$$

Con $U = V$ indicheremo che $U \leq V$ e $V \leq U$; cioè, $U = V$ equivale a $\forall x (U(x_1, \dots, x_n) \leftrightarrow V(x_1, \dots, x_n))$.

Data una teoria T del primo ordine, scriveremo $T(P_1, \dots, P_n)$ per mettere in evidenza che P_1, \dots, P_n sono alcuni dei simboli predicativi del linguaggio di T (che possono eventualmente anche non comparire in T).

Sia $T(P)$ una teoria del primo ordine. Supponiamo di voler circoscrivere P in $T(P)$. Ciò che si vuole ottenere è che l'estensione di P sia minimale rispetto a T , cioè che sia la più piccola possibile senza violare T . Ciò può essere ottenuto aggiungendo a T il seguente assioma del secondo ordine:

$$(\dagger) \forall \Phi (T(\Phi) \wedge \Phi \leq P \rightarrow P \leq \Phi),$$

dove $T(\Phi)$ è la formula ottenuta congiungendo tutti gli assiomi di T e sostituendo la variabile predicativa Φ ad ogni occorrenza della lettera predicativa P . Intuitivamente, (\dagger) dice che, per ogni predicato Φ , se:

- (a) Φ può essere sostituito a P in $T(P)$ senza violare la teoria,
- (b) Φ ha estensione minore di P ,

allora si assume che Φ e P coincidano. Fra i vari predicati Φ che soddisfano (a) e (b) ci sarà anche quello con l'estensione minimale compatibile con $T(P)$.

Quanto detto può essere facilmente generalizzato a un numero qualsiasi di predicati. Sia $T(P_1, \dots, P_n)$ una teoria del primo ordine in cui si vogliono circoscrivere P_1, \dots, P_n . Ciò si ottiene aggiungendo a $T(P_1, \dots, P_n)$ l'assioma:

$$(\dagger\dagger) \forall \Phi_1, \dots, \Phi_n (T(\Phi_1, \dots, \Phi_n) \wedge \Phi_1 \leq P_1 \wedge \dots \wedge \Phi_n \leq P_n \rightarrow P_1 \leq \Phi_1 \wedge \dots \wedge P_n \leq \Phi_n),$$

dove, analogamente a sopra, $T(\Phi_1, \dots, \Phi_n)$ è la formula ottenuta congiungendo tutti gli assiomi di $T(P_1, \dots, P_n)$ e sostituendovi rispettivamente tutte le occorrenze di P_1, \dots, P_n con le variabili predicative Φ_1, \dots, Φ_n .

Chiameremo *predicate circumscription* di P_1, \dots, P_n in T (in simboli: $CIRC_{PR}(T; P_1, \dots, P_n)$) l'insieme di formule $T \cup \{(\dagger\dagger)\}$. Diremo che una formula α è una *conseguenza* della *predicate circumscription* di P_1, \dots, P_n in T se e soltanto se $CIRC_{PR}(T; P_1, \dots, P_n) \models \alpha$, dove \models è la relazione di conseguenza logica della logica dei predicati del secondo ordine classica.

Vediamo un semplice esempio. Consideriamo la seguente teoria:

$$T = \{albino(Gigi)\}.$$

Vogliamo circoscrivere il predicato *albino* in T . L'assioma di *circumscription* per $T(albino)$ è il seguente:

$$(\ddagger) \forall \Phi (\Phi(Gigi) \wedge \forall x(\Phi(x) \rightarrow albino(x)) \rightarrow \forall x(albino(x) \rightarrow \Phi(x)),$$

(ottenuto sostituendo *albino* a P in (\dagger) , ed esplicitando \leq in base alla definizione). Il nostro scopo è minimizzare l'estensione di *albino* in T ; vogliamo cioè ottenere:

$$(\ddagger\ddagger) \forall x(albino(x) \rightarrow x=Gigi).$$

In altri termini vogliamo assumere che il predicato *albino* risulti estensionalmente equivalente al predicato "essere Gigi" (in simboli, usando la λ -notazione⁵, $\lambda x.x=Gigi$). Sostituendo $\lambda x.x=Gigi$ al posto di Φ in (\ddagger) , otteniamo:

$$(\ddagger\ddagger\ddagger) (Gigi=Gigi \wedge \forall x(x=Gigi \rightarrow albino(x)) \rightarrow \forall x(albino(x) \rightarrow x=Gigi).$$

$(\ddagger\ddagger\ddagger)$ è chiaramente una conseguenza di (\ddagger) . D'altro canto,

$$(\ddagger\ddagger\ddagger\ddagger) Gigi=Gigi \wedge \forall x(x=Gigi \rightarrow albino(x))$$

è una conseguenza classica (monotona) di T . Poiché $(\ddagger\ddagger)$ è una conseguenza classica del primo ordine di $(\ddagger\ddagger\ddagger)$ e di $(\ddagger\ddagger\ddagger\ddagger)$, ne segue che:

$$CIRC_{PR}(T; albino) = T \cup \{(\ddagger)\} \models (\ddagger\ddagger),$$

che è quanto si desiderava.

McCarthy (1980) ha fornito una semantica in termini di teoria dei modelli per la *predicate circumscription*. Tale semantica si basa sul concetto di *modello minimale* di una teoria rispetto a un certo insieme di predicati. Introduciamo le seguenti definizioni.

⁵La λ -notazione è stata introdotta da Alonzo Church per eliminare un'ambiguità dell'usuale notazione matematica inerente le funzioni. Ad esempio, l'espressione $sen(x)$ viene usata abitualmente in maniera ambigua, in quanto denota talvolta la funzione seno (come nella frase " $sen(x)$ è periodica"), talvolta un generico valore della funzione seno (come in " $sen(x)$ è compreso fra 1 e -1"). Con la λ -notazione tale ambiguità viene risolta indicando con $\lambda x.sen(x)$ la funzione, e riservando $sen(x)$ per il valore della funzione per l'argomento x . In logica, un'ambiguità simile sorge fra formule aperte e predicati. Ad esempio,

$$(*) cane(x) \vee gatto(x)$$

può indicare sia una formula aperta con la variabile x libera, sia il predicato "essere un cane o un gatto". Utilizzando la λ -notazione, si usa $(*)$ per indicare la formula, e si scrive $\lambda x.(cane(x) \vee gatto(x))$ per indicare il predicato.

Definizione: sia $T(P_1, \dots, P_n)$ una teoria del primo ordine, e siano $M=(D, \varphi)$ e $M'=(D', \varphi')$ due modelli di $T(P_1, \dots, P_n)$. Sia inoltre $\bar{P} = (P_1, \dots, P_n)$. Diremo che M' è un \bar{P} -sottomodello di M della teoria T (in simboli: $M' \leq^{\bar{P}} M$), se e soltanto se:

- (i) $D' = D$;
- (ii) $\varphi'[P_i] \subseteq \varphi[P_i]$, per ogni i tale che $1 \leq i \leq n$;
- (iii) $\varphi'[K] = \varphi[K]$ per ogni costante (predicativa, funzionale o individuale) del linguaggio di T che non sia un membro di \bar{P} .

Definizione: diciamo che M è un modello \bar{P} -minimale di T se e soltanto se ogni modello di T che è un \bar{P} -sottomodello di M è identico a M .

Si può dimostrare allora il seguente:

Teorema: per ogni teoria T e per ogni enupla \bar{P} di costanti predicative del linguaggio di T , i modelli di $CIRC_{PR}(T; \bar{P})$ sono tutti e soli i modelli \bar{P} -minimali di T .

La semantica dei modelli minimali cattura appunto l'intuizione che i modelli della *circumscription* di una teoria $T(P_1, \dots, P_n)$ siano quei modelli di T in cui i predicati P_1, \dots, P_n abbiano la più piccola estensione possibile.

Sin qui abbiamo dato per la *circumscription* una definizione di conseguenza semantica, senza specificare un concetto corrispondente di derivazione sintattica. Il punto è che, come è ben noto, per la logica dei predicati del secondo ordine non è in generale disponibile una teoria della dimostrazione completa, e questo rende problematica una caratterizzazione a livello sintattico della *circumscription*. Nella formulazione originaria della *predicate circumscription*, McCarthy evitava il problema sostituendo all'assioma del secondo ordine ($\dagger\dagger$) uno schema di assiomi del primo ordine ($\dagger\dagger'$). ($\dagger\dagger'$) si ottiene eliminando la quantificazione universale sulle variabili predicative da ($\dagger\dagger$). Se indichiamo con $CS(\dagger\dagger')$ l'insieme delle formule (del primo ordine) che si ottengono sostituendo in maniera appropriata in ($\dagger\dagger'$) espressioni predicative esprimibili al primo ordine alle variabili Φ_1, \dots, Φ_n , allora si possono definire i teoremi della *circumscription* dei predicati P_1, \dots, P_n nella teoria T come quelle formula α tali che:

$$T \cup CS(\dagger\dagger') \vdash \alpha,$$

dove \vdash è la relazione di derivabilità del calcolo dei predicati del primo ordine classico. Il problema è che questa formulazione è meno potente di quella che utilizza un assioma del secondo ordine. Infatti, in ($\dagger\dagger$) si quantifica su *tutti* i predicati, mentre con ($\dagger\dagger'$) si catturano solo quei predicati che sono esprimibili in un linguaggio del primo ordine, e che sono, come è noto, un sottoinsieme proprio dei primi. Di conseguenza, la formulazione al primo ordine della *predicate circumscription* è corretta ma non completa rispetto alla semantica dei modelli minimali. E' tuttavia possibile definire una semantica rispetto alla quale la *predicate circumscription* al primo ordine sia corretta e completa. Tale semantica è basata sul concetto di *modelli minimali deboli* (*weakly minimal model*) (Besnard et al. 1989). In generale, questo tipo di problemi si ripropone per tutte le forme più potenti di *circumscription* che vedremo in seguito, e il problema di individuare forme di *circumscription* esprimibili al primo ordine è uno fra i più studiati in letteratura.

La *predicate circumscription* è la forma più semplice e meno potente di *circumscription*. In particolare, non è abbastanza potente per trattare regole non monotone formulate per mezzo di predicati di *abnormality* (cfr. ad es. Etherington et al. 1985). Il problema è che nella *predicate circumscription* soltanto ai predicati che vengono minimizzati è consentito variare la propria estensione durante il processo di minimalizzazione. Viceversa, minimizzando il predicato ab in una regola come:

$$\forall x(\text{uccello}(x) \wedge \neg ab(x) \rightarrow \text{vola}(x)),$$

si vuole che anche il predicato vola cambi la propria estensione (l'idea intuitiva è che quanti meno sono gli uccelli *abnormal*, tanti più saranno gli uccelli in grado di volare). Ma questo non è possibile in *predicate circumscription*. Si consideri ad esempio la teoria:

$$T = \{ \forall x(\text{uccello}(x) \wedge \neg ab(x) \rightarrow \text{vola}(x)), \text{uccello}(\text{Tweety}) \}.$$

La *predicate circumscription* di $T(ab)$ si ottiene aggiungendo a T il seguente assioma:

$$(\S) \forall \Phi(\text{uccello}(\text{Tweety}) \wedge \forall x(\text{uccello}(x) \wedge \neg \Phi(x) \rightarrow \text{vola}(x)) \wedge \Phi \leq ab \rightarrow ab \leq \Phi).$$

Si vorrebbe che ab avesse estensione nulla, ossia si vorrebbe sostituire a Φ il predicato $\lambda x.false$. Sostituendo però $\lambda x.false$ al posto di ab in (§) si ottiene:

$$(\S) \text{uccello}(Tweety) \wedge \forall x(\text{uccello}(x) \rightarrow \text{vola}(x)) \wedge \forall x(\text{false} \rightarrow \text{ab}(x)) \rightarrow \forall x(\text{ab}(x) \rightarrow \text{false});$$

ma (§) non consente di ottenere $\forall x(\text{ab}(x) \rightarrow \text{false})$. In particolare, non si può usare lo stesso tipo di ragionamento usato nell'esempio precedente: l'antecedente di (§) non segue da T (in quanto da T non segue $\forall x(\text{uccello}(x) \rightarrow \text{vola}(x))$).

Che la *predicate circumscription* non consenta di ottenere i risultati voluti nel caso della teoria T è chiaro ragionando dal punto di vista semantico. Si consideri una interpretazione $M=(D,\varphi)$, in cui $D=\{\text{Tweety}\}$, e in cui si abbia: $\varphi[\text{Tweety}]=\text{Tweety}$, $\varphi[\text{uccello}]=\varphi[\text{ab}]=\{\text{Tweety}\}$ e $\varphi[\text{vola}]=\emptyset$. M è certamente un modello di T . Il punto è che, in base alla definizione data sopra, M è anche un modello ab -minimale di T . Non si può infatti ridurre l'estensione di ab lasciando intatta l'estensione di tutti gli altri predicati (in particolare di vola). Quindi, poiché esiste un modello ab -minimale di T in cui $\text{vola}(\text{Tweety})$ è falsa, $\text{vola}(\text{Tweety})$ non segue da $\text{CIRC}_{PR}(T;ab)$. Il problema è appunto che la *predicate circumscription* non consente di modificare l'estensione di alcun altro predicato eccetto ab .

In casi di questo genere, per ottenere i risultati voluti è necessaria una forma di *circumscription* più potente, che McCarthy ha chiamato *formula circumscription* (McCarthy 1984). Ciò che caratterizza la *formula circumscription* rispetto alla *predicate circumscription* è appunto il fatto che nella *formula circumscription* si consente ad altri predicati, oltre a quelli che vengono minimizzati, di variare la propria estensione nel corso del processo di minimalizzazione. Nella esposizione seguente useremo una formulazione che si discosta per certi versi da quella originale di McCarthy, ed è più vicina alla *second order circumscription* di Lifschitz (1985, 1986), la quale tuttavia risulta essere sostanzialmente una variante notazionale della *formula circumscription* originale.

Sia $T(P_1, \dots, P_n, Q_1, \dots, Q_m)$ una teoria del primo ordine, dove (P_1, \dots, P_n) e (Q_1, \dots, Q_m) sono disgiunti. Si vogliono circoscrivere P_1, \dots, P_n lasciando Q_1, \dots, Q_m liberi di variare durante il processo di minimalizzazione. Lo schema di assiomi del secondo ordine che si deve aggiungere a $T(P_1, \dots, P_n, Q_1, \dots, Q_m)$ per ottenerne la *formula circumscription* è il seguente:

$$(\P) \forall \Phi_1, \dots, \Phi_n, \Psi_1, \dots, \Psi_m (T(\Phi_1, \dots, \Phi_n, \Psi_1, \dots, \Psi_m) \wedge \Phi_1 \leq P_1 \wedge \dots \wedge \Phi_n \leq P_n \\ \rightarrow P_1 \leq \Phi_1 \wedge \dots \wedge P_n \leq \Phi_n),$$

dove, analogamente a sopra, $T(\Phi_1, \dots, \Phi_n, \Psi_1, \dots, \Psi_m)$ è la formula ottenuta congiungendo tutti gli assiomi di $T(P_1, \dots, P_n, Q_1, \dots, Q_m)$ e sostituendovi rispettivamente tutte le occorrenze di $P_1, \dots, P_n, Q_1, \dots, Q_m$ con le variabili predicative $\Phi_1, \dots, \Phi_n, \Psi_1, \dots, \Psi_m$.

Chiameremo *formula circumscription* di P_1, \dots, P_n in T con Q_1, \dots, Q_m variabili (in simboli: $\text{CIRC}_{FC}(T; P_1, \dots, P_n; Q_1, \dots, Q_m)$) l'insieme di formule $T \cup \{(\P)\}$. Diremo che una formula α è una *conseguenza* della *formula circumscription* di P_1, \dots, P_n in T con Q_1, \dots, Q_m variabili se e soltanto se $\text{CIRC}_{FC}(T; P_1, \dots, P_n; Q_1, \dots, Q_m) \models \alpha$, dove \models è la relazione di conseguenza della logica dei predicati del secondo ordine classica.

Riprendiamo l'esempio di Tweety. Sia:

$$T = \{\forall x (\text{uccello}(x) \wedge \neg \text{ab}(x) \rightarrow \text{vola}(x)), \text{uccello}(\text{Tweety})\}.$$

Facendo la *formula circumscription* di ab in T con il predicato vola libero di variare, si risolve il problema visto in precedenza. Si vuole che ab abbia estensione nulla. Questo può essere ottenuto sostituendo Φ con $\lambda x.false$ e Ψ con $\lambda x.\text{uccello}(x)$ nell'assioma di *formula circumscription* per $T(ab; \text{vola})$. Questa seconda sostituzione risolve i problemi che sorgevano con la *predicate circumscription*. In particolare, si avrà che:

$$\begin{aligned} \text{CIRC}_{FC}(T; ab; \text{vola}) &\models \forall x (\text{ab}(x) \rightarrow \text{false}); \\ \text{CIRC}_{FC}(T; ab; \text{vola}) &\models \neg \text{ab}(\text{Tweety}); \\ \text{CIRC}_{FC}(T; ab; \text{vola}) &\models \text{vola}(\text{Tweety}). \end{aligned}$$

La semantica basata sui modelli minimali può essere estesa alla *formula circumscription*. E' sufficiente modificare le precedenti definizioni nel modo seguente.

Definizione: sia $T(P_1, \dots, P_n, Q_1, \dots, Q_m)$ una teoria del primo ordine, e siano $M=(D,\varphi)$ e $M'=(D',\varphi')$ due modelli di $T(P_1, \dots, P_n, Q_1, \dots, Q_m)$. Siano inoltre $\bar{P} = (P_1, \dots, P_n)$ e $\bar{Q} = (Q_1, \dots, Q_m)$. Diremo che M' è un $(\bar{P}; \bar{Q})$ -sottomodello di M della teoria T (in simboli: $M' \leq^{(\bar{P}; \bar{Q})} M$), se e soltanto se:

- (i) $D' = D$;
- (ii) $\phi[P_i] \subseteq \phi[P_i]$, per ogni i tale che $1 \leq i \leq n$;
- (iii) $\phi[K] = \phi[K]$ per ogni costante (predicativa, funzionale o individuale) del linguaggio di T che non sia un membro di \bar{P} o di \bar{Q} .

Definizione: diciamo che M è un modello $(\bar{P}; \bar{Q})$ -minimale di T se e soltanto se ogni modello di T che è un $(\bar{P}; \bar{Q})$ -sottomodello di M è identico a M .

Si può dimostrare allora il seguente:

Teorema: per ogni teoria T e per ogni \bar{P}, \bar{Q} disgiunti, i modelli di $CIRC_{FC}(T; \bar{P}; \bar{Q})$ sono tutti e soli i modelli $(\bar{P}; \bar{Q})$ -minimali di T .

Anche la *formula circumscription* non risolve tuttavia tutti i problemi. Consideriamo una rete come quella in fig. 6.2, da leggersi nel modo seguente: "gli uccelli, tipicamente, volano; gli struzzi sono uccelli; gli struzzi, tipicamente, non volano". Si vuole che il legame di tipo *is not* fra i concetti *struzzo* e *vola* venga letto come un arco che ammette eccezioni. Si vuole cioè che si mantenga aperta la possibilità che esistano struzzi atipici in grado di volare. Tradurre tale rete in calcolo dei predicati utilizzando la tecnica dei predicati di anormalità darà quindi luogo al seguente insieme di formule:

- (i) $\forall x (uccello(x) \wedge \neg ab_2(x) \rightarrow vola(x))$
- (ii) $\forall x (struzzo(x) \wedge \neg ab_1(x) \rightarrow \neg vola(x))$
- (iii) $\forall (x) (struzzo(x) \rightarrow uccello(x))$.

Va notato innanzi tutto che nelle formule (i) e (ii), che traducono i due archi delle rete che ammettono eccezioni, sono stati utilizzati due predicati di *abnormality* distinti: ab_1 e ab_2 . Questo è necessario per evitare conclusioni indesiderate: se in (i) e (ii) si fosse usato lo stesso predicato ab , aggiungendo la formula

$$(iv) struzzo(Fred)$$

si avrebbe avuto

$$(i) - (iv) \vdash \neg ab(Fred) \rightarrow vola(Fred) \wedge \neg vola(Fred),$$

che non corrisponde certamente all'intuizione che si intendeva formalizzare.

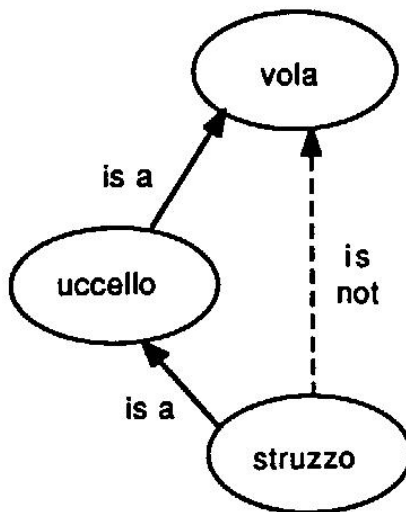


fig. 6.2

Sia dunque $T = \{(i), (ii), (iii), (iv)\}$. Minimizzare in T l'estensione di ab_1 è in conflitto con la minimizzazione di ab_2 . Assumere che ab_1 abbia l'estensione vuota comporta che valga $ab_2(Fred)$, e viceversa. Quindi, praticando la *formula circumscription* di $T(ab_2, ab_1; vola)$, il massimo che si può ottenere è che:

$$\neg ab_2(Fred) \vee \neg ab_1(Fred);$$

cioè, per quanto riguarda la capacità di Fred di volare:

$$vola(Fred) \vee \neg vola(Fred),$$

che è certamente troppo poco. Intuitivamente, una rete come quella della fig. 6.2. avrebbe dovuto consentire di derivare che Fred, in quanto struzzo, in mancanza di informazioni più precise, non vola. Infatti, abbiamo visto che in questo tipo di reti si assume che le informazioni a carattere più specifico godano una sorta di "priorità" su quelle a carattere più generale. Se Fred è uno struzzo, e se gli struzzi di solito non volano, si vuole concludere, in assenza di altre informazioni, che Fred non vola, in quanto l'informazione che Fred è uno struzzo dice qualcosa di più rispetto all'informazione che Fred è genericamente un uccello.

Per risolvere questo tipo di problemi mediante la *circumscription*, Vladimir Lifschitz ha introdotto una tecnica, la *prioritized circumscription* (Lifschitz 1985, 1986) che consente appunto di istituire un ordine di priorità nel circoscrivere i predicati di una teoria.

Consideriamo il caso più semplice, in cui in una teoria $T(P_1, P_2; S)$ si abbiano due soli predicati da circoscrivere, P_1 e P_2 , e si intenda circoscrivere P_1 con priorità su P_2 . Come nella *formula circumscription*, si vuol lasciare il predicato S libero di variare durante il processo di minimalizzazione.

E' necessario innanzi tutto introdurre una relazione d'ordine fra predicati più complessa della relazione \leq utilizzata fino a ora. Sia $\leq\leq$ una relazione fra coppie di predicati definita come segue:

$$(U_1, U_2) \leq\leq (V_1, V_2) =_{def} U_1 \leq V_1 \wedge (U_1 = V_1 \rightarrow U_2 \leq V_2).$$

Si noti che $(U_1, U_2) \leq\leq (V_1, V_2)$ non comporta che si abbia $U_2 \leq V_2$. Questo vale soltanto nel caso che $U_1 = V_1$.

Per ottenere la *prioritized circumscription* di $T(P_1, P_2; S)$ con priorità di P_1 su P_2 e S variabile (che indicheremo con $CIRC_{PC}(T; P_1 > P_2; S)$) si aggiunge a T il seguente assioma del secondo ordine:

$$(PC) \forall \Phi_1, \Phi_2, \Psi (T(\Phi_1, \Phi_2, \Psi) \wedge (\Phi_1, \Phi_2) \leq\leq (P_1, P_2) \rightarrow (P_1, P_2) \leq\leq (\Phi_1, \Phi_2)).$$

Sia $CIRC_{PC}(T; P_1 > P_2; S)$ l'unione degli assiomi di T con (PC).

Vediamo ora come la *prioritized circumscription* possa risolvere il problema dell'esempio visto sopra. Di nuovo, sia $T = \{(i)-(iv)\}$. Per ottenere $CIRC_{PC}(T; ab_1 > ab_2; vola)$ si aggiunge a T lo schema di assiomi (PC) in cui P_1 e P_2 sono rispettivamente ab_1 e ab_2 . Per ottenere i risultati voluti si devono praticare le seguenti sostituzioni delle variabili predicative. Si deve sostituire $\lambda x.false$ a Φ_1 , $\lambda x.struzzo(x)$ a Φ_2 e $\lambda x.uccello(x) \wedge \neg struzzo(x)$ a Ψ . E' facile constatare che tali sostituzioni risultano consistenti in T . Inoltre, in base alla definizione di $\leq\leq$, si ha che:

$$(\lambda x.false, \lambda x.struzzo(x)) \leq\leq (ab_1, ab_2).$$

Si ha quindi che:

$$CIRC_{PC}(T; ab_1 > ab_2; vola) \models (ab_1, ab_2) \leq\leq (\lambda x.false, \lambda x.struzzo(x));$$

cioè, in base alla definizione di $\leq\leq$:

$$CIRC_{PC}(T; ab_1 > ab_2; vola) \models ab_1 \leq \lambda x.false \wedge (ab_1 = \lambda x.false) \rightarrow (ab_2 \leq \lambda x.struzzo(x)).$$

Si ha quindi che:

$$CIRC_{PC}(T; ab_1 > ab_2; vola) \models \forall x \neg (ab_1(x));$$

e quindi:

$$CIRC_{PC}(T; ab_1 > ab_2; vola) \models \neg vola(Fred).$$

D'altra parte:

$$CIRC_{PC}(T; ab_1 > ab_2; vola) \models \forall(x) (ab_2(x) \leftrightarrow struzzo(x));$$

da (i) non segue quindi che $vola(Fred)$.

L'assioma (PC) può essere generalizzato a un assioma di *prioritized circumscription* (PC') che consenta di circoscrivere un numero qualsiasi di predicati. Qui, per brevità, non definiremo (PC'), limitandoci a dare la caratterizzazione semantica della *prioritized circumscription* nel caso più generale. Sia T una teoria del primo ordine, e siano $\bar{P}^1, \dots, \bar{P}^n, \bar{Q}$ insiemi ordinati disgiunti di costanti predicative del linguaggio di T . Indicheremo con $CIRC_{PC}(T; \bar{P}^1 > \dots > \bar{P}^n; \bar{Q})$ la *prioritized circumscription* di $\bar{P}^1, \dots, \bar{P}^n$ con i predicati \bar{Q} liberi di variare, e con priorità di \bar{P}^i su \bar{P}^{i+1} per ogni i tale che $1 \leq i < n-1$. Analogamente agli altri tipi di *circumscription*, $CIRC_{PC}(T; \bar{P}^1 > \dots > \bar{P}^n; \bar{Q})$ sarà ottenuta aggiungendo l'assioma (PC') agli assiomi della teoria T . Le conseguenze della *prioritized circumscription* di $(T; \bar{P}^1 > \dots > \bar{P}^n; \bar{Q})$ sono tutte e sole le formule che seguono da $CIRC_{PC}(T; \bar{P}^1 > \dots > \bar{P}^n; \bar{Q})$ in base alla relazione di conseguenza logica classica per la logica del secondo ordine.

La semantica della *prioritized circumscription* può essere sviluppata generalizzando ulteriormente la semantica dei modelli minimali nella maniera seguente:

Definizione: siano $\bar{P}^1, \dots, \bar{P}^n, \bar{Q}$ insiemi ordinati disgiunti di costanti predicative del linguaggio di una teoria T . Siano $M = (D, \varphi)$ e $M' = (D', \varphi')$ due modelli di T . Diremo che M' è un $(\bar{P}^1 > \dots > \bar{P}^n; \bar{Q})$ -sottomodello di M (in simboli $M' \leq_{(\bar{P}^1 > \dots > \bar{P}^n; \bar{Q})} M$) se e soltanto se vale quanto segue:

- (i) $D' = D$
- (ii) $\varphi'[k] = \varphi[k]$ per ogni costante k non compresa in $\bar{P}^1, \dots, \bar{P}^n, \bar{Q}$;
- (iii) $\varphi'[P] \subseteq \varphi[P]$ per ogni $P \in \bar{P}^1$, e, per ogni $1 < i \leq n$, se

$$\varphi'[P] = \varphi[P] \text{ per ogni } P \in (\bar{P}^1, \dots, \bar{P}^{i-1}),$$

allora

$$\varphi'[P] \subseteq \varphi[P] \text{ per ogni } P \in \bar{P}^i.$$

Definizione: Un modello M di T è $(\bar{P}^1 > \dots > \bar{P}^n; \bar{Q})$ -minimale se e soltanto se ogni modello di T che sia un $(\bar{P}^1 > \dots > \bar{P}^n; \bar{Q})$ -sottomodello di M è identico a M .

Anche in questo caso si può dimostrare il seguente:

Teorema: per ogni teoria T e per ogni $\bar{P}^1, \dots, \bar{P}^n, \bar{Q}$ disgiunti, i modelli di $CIRC_{PC}(T; \bar{P}^1 > \dots > \bar{P}^n; \bar{Q})$ sono tutti e soli i modelli $(\bar{P}^1 > \dots > \bar{P}^n; \bar{Q})$ -minimali di T .

6.2 Altre logiche per il ragionamento non monotono

In questo paragrafo daremo alcuni rapidi cenni a due approcci al ragionamento non monotono diversi dalla *circumscription*, vale a dire la *default logic* e le logiche non monotone modali. Assieme, *circumscription*, *default logic* e logiche modali non monotone costituiscono un quadro abbastanza completo delle principali linee di ricerca sul ragionamento non monotono.

La *default logic* (Reiter 1980; Etherington 1988) è un sistema di logica non monotona che si basa sull'estensione dell'apparato deduttivo di una teoria classica del primo ordine per mezzo di regole di inferenza non monotone, dette *default rule*. Nella caso più generale, una *default rule* ha la forma seguente:

$$(R1) \frac{\alpha: \beta_1, \dots, \beta_n}{\gamma}$$

dove $\alpha, \beta_1, \dots, \beta_n$ e γ sono formule del primo ordine classiche (aperte). Il significato intuitivo di una regola del genere è: se può essere derivato α , e se non è contraddittorio assumere che β_1, \dots, β_n (cioè, se non possono essere derivate $\neg\beta_1, \dots, \neg\beta_n$), allora si derivi γ . In una regola di *default* come (R1) α è detta la *prerequisito*, β_1, \dots, β_n sono dette

giustificazioni e γ il conseguente. Il fatto che gli uccelli tipici possono volare può essere espresso in *default logic* per mezzo di una regola del genere:

$$(R2) \frac{uccello(x) : vola(x)}{vola(x)}$$

che esprime appunto che si vuole assumere che un uccello voli a meno che non si disponga dell'evidenza esplicita del contrario.

Una teoria di *default* (*default theory*) è una coppia $\Delta = (D, W)$, dove D è un insieme di regole di *default* del tipo della (R1), e W è un insieme di formule del primo ordine. Intuitivamente, una formula è derivabile da una teoria di *default* Δ se è derivabile a partire dalle formule di W per mezzo delle regole di inferenza della logica classica e per mezzo delle *default rule* in D . Il concetto di insieme di formule derivabili da una teoria di *default* può essere formalizzato per mezzo del concetto di *estensione* di una teoria di *default*. Una estensione E di una teoria $\Delta = (D, W)$ ⁶ è un insieme minimale di formule tale che:

- 1) $W \subseteq E$;
- 2) E è chiuso rispetto alla deduzione classica;
- 3) E rispetta le regole in D , cioè, per ogni regola (R1) in D , se $\alpha \in E$, e se $\neg\beta_1, \dots, \neg\beta_n \notin E$, allora $\gamma \in E$.

Vi sono casi in cui una teoria di *default* ammette più di un'estensione. Consideriamo tuttavia un semplice esempio in cui l'estensione è unica. L'insieme D abbia come unico elemento la regola (R2) sopra riportata, e si abbia:

$$W = \{uccello(Tweety), \forall x(pinguino(x) \rightarrow \neg vola(x))\}.$$

Poiché dalle formule di W non è possibile derivare $\neg vola(Tweety)$, allora, in base alla regola (R2), all'estensione di questa teoria appartiene la formula $vola(Tweety)$. Aggiungendo tuttavia a W la premessa $pinguino(Tweety)$, allora è possibile derivare $\neg vola(Tweety)$, e l'inferenza di $vola(Tweety)$ mediante (R2) viene bloccata.

Storicamente, la *default logic* è stata la prima logica non monotona utilizzata per ridurre alla logica le reti semantiche che ammettono eccezioni all'ereditarietà. Etherington e Reiter (1983) e Reiter e Criscuolo (1980, 1983) propongono una traduzione in *default logic* delle reti semantiche tipo NETL. In base a tale traduzione, gli archi di tipo *isa* e di tipo *is not* che ammettono eccezioni vengono tradotti mediante regole di *default* (mentre la traduzione degli archi che non ammettono eccezioni consiste di formule della logica del primo ordine) (si veda anche Etherington 1988, cap. 4). Si consideri ad esempio il "rombo di Nixon" (fig. 4.9). L'arco fra *quacchero* e *pacifista* e quello fra *repubblicano* e *pacifista* vengono tradotti rispettivamente mediante le due regole seguenti:

$$(Nix1) \frac{quacchero(x) : pacifista(x)}{pacifista(x)}$$

e

$$(Nix2) \frac{repubblicano(x) : \neg pacifista(x)}{\neg pacifista(x)}$$

mentre gli archi *isa* dal nodo *Nixon* ai concetti *quacchero* e *repubblicano* corrispondono rispettivamente alle formule:

$$(Nix3) quacchero(Nixon)$$

e

$$(Nix4) repubblicano(Nixon).$$

Si noti che la *default theory* che ne deriva, in cui $D = \{(Nix1), (Nix2)\}$ e $W = \{(Nix3), (Nix4)\}$, ammette due estensioni, una che comprende $pacifista(Nixon)$, e un'altra che comprende $\neg pacifista(Nixon)$. Quale sia l'estensione ottenuta dipende dall'ordine in cui vengono applicati i *default* a partire dalle formule di W . Ciò è dovuto alla sostanziale ambiguità della rete della fig. 4.9. Se, d'altra parte, si volesse che, ad esempio, le tendenze repubblicane prevalessero su quelle quacchere, basterebbe sostituire a (Nix1) la seguente regola di *default*, più debole:

$$(Nix1') \frac{quacchero(x) : pacifista(x) \wedge \neg repubblicano(x)}{pacifista(x)},$$

⁶La presente definizione di estensione si applica a teorie di *default chiuse*, in cui cioè i prerequisiti, giustificazioni e conseguenti di tutte le regole di *default* non contengono alcuna variabile libera. Per applicare la definizione a teorie che comprendono regole in cui compaiono variabili libere (come ad esempio la regola R2), tali regole devono essere considerate schemi, che devono essere rimpiazzate da tutte le loro possibili istanziazioni.

che fa sì che, in mancanza di informazioni più specifiche, venga preferita l'estensione che comprende $\neg pacifista(Nixon)$.

Se si vuole tradurre in *default logic* la rete di fig. 6.2, l'arco fra *struzzo* e *uccello*, che non ammette eccezioni, viene tradotto con la formula:

$$(Bird1) \forall x (struzzo(x) \rightarrow uccello(x)).$$

Gli archi che ammettono eccezioni vengono tradotti con regole di *default*. L'arco fra *struzzo* e *vola* corrisponde alla regola:

$$(Bird2) \frac{struzzo(x) : \neg vola(x)}{\neg vola(x)},$$

mentre l'arco fra *uccello* e *vola* viene tradotto dalla regola:

$$(Bird3) \frac{uccello(x) : vola(x) \wedge \neg struzzo(x)}{vola(x)}.$$

Cioè, se qualcosa è un uccello, ed è consistente credere che voli e che non sia un struzzo, allora si assume che voli. Il fatto che nella giustificazione vada specificato che sia consistente credere che non si tratta di uno struzzo, corrisponde, in un certo senso, alla necessità di circoscrivere ab_1 con priorità su ab_2 nel trattamento di questo stesso esempio mediante la *prioritized circumscription*. Se, invece di (Bird3), si utilizzasse infatti la regola:

$$(Bird3') \frac{uccello(x) : vola(x)}{vola(x)}.$$

Non si otterrebbero le conseguenze volute. Infatti, una teoria di *default* che avesse $D = \{(Bird2), (Bird3')\}$ e $W = \{(Bird1), struzzo(Tweety)\}$, ammetterebbe due estensioni, una con $vola(Tweety)$, e l'altra con $\neg vola(Tweety)$.

Un ulteriore approccio al problema dell'inferenza non monotona è costituito dalle logiche non monotone modali. In tali logiche il linguaggio viene esteso mediante opportuni operatori modali che esprimono che una certa informazione è creduta, o che è consistente con le credenze del soggetto epistemico rappresentato. Questa linea di ricerca è stata inaugurata dalla *non monotonic logic* di McDermott e Doyle (McDermott e Doyle 1980; McDermott 1982). La *logica autoepistemica* è una logica modale elaborata da Moore (1985) a partire dal lavoro di McDermott e Doyle, ed è stata la logica più ricca di sviluppi in questa tradizione⁷. Il linguaggio della logica autoepistemica comprende un operatore modale L il cui significato inteso è "è creduto che ...". Lo scopo è quello di costruire un modello logico di un agente epistemico che ragioni introspettivamente sulle proprie credenze. Una teoria in logica autoepistemica rappresenta l'insieme delle credenze di un tale agente. Se una formula α segue dalla teoria, allora α è una delle credenze dell'agente epistemico. Per quanto riguarda le formule modali, il significato intuitivo di una formula del tipo $L\alpha$ è che l'agente crede di sé stesso di credere α . Per cui, se una certa formula α è derivabile dalla teoria, allora l'agente introspettivamente inferisce di credere α (inferisce cioè che valga $L\alpha$). L'aspetto non monotono della logica autoepistemica deriva dal fatto di supporre inoltre che le formule derivabili dalla teoria costituiscano *tutto* ciò che l'agente crede. Quindi, se una formula α non è derivabile, si assume che l'agente introspettivamente inferisca di non credere α (derivando $\neg L\alpha$). Questa, ovviamente, è una forma di inferenza non monotona: se supponiamo di aggiungere α alla base di conoscenza, non deve più essere possibile ottenere $\neg L\alpha$. Questa idea intuitiva viene formalizzata per mezzo del concetto di *espansione stabile* (*stable expansion*) di un insieme di formule T . Dato T , l'espansione stabile di T è l'insieme delle conseguenze logiche classiche del seguente insieme di formule:

$$T \cup \{L\alpha \mid \alpha \in T\} \cup \{\neg L\alpha \mid \alpha \notin T\}.$$

Alla logica autoepistemica è possibile associare una semantica a mondi possibili basata su strutture di Kripke.

In logica autoepistemica l'equivalente di una regola di *default* come (R2) è dato dalla formula:

$$(6) \alpha \wedge \neg L \neg \beta_1 \wedge \dots \wedge \neg L \neg \beta_n \rightarrow \gamma,$$

⁷Su logica autoepistemica e sistemi non monotoni modali si vedano anche (Halpern e Moses 1984a) e (Levesque 1990).

che può essere letta come segue: "se α , e se non sono credute le negazioni di β_1, \dots, β_n , allora γ ". Il fatto che gli uccelli, salvo informazioni contrarie, volino, può essere formalizzato come segue:

$$(7) \text{uccello}(x) \wedge \neg L \neg \text{vola}(x) \rightarrow \text{vola}(x).$$

E' importante notare che in logica autoepistemica le regole di *default* vengono espresse come formule del linguaggio, e non come regole di inferenza metateoriche come accade in *default logic*. Questo consente ad esempio di esprimere *default* che vertono su altri *default* mediante iterazione dell'operatore L , il che non è possibile in *default logic*. Ad esempio l'equivalente di formule come:

$$\begin{array}{c} \neg LL\neg\alpha \\ \neg L\neg(\neg L\neg\alpha \rightarrow \beta) \end{array}$$

non è esprimibile direttamente nella logica di *default*. Sull'uso della logica autoepistemica per tradurre reti semantiche con eccezioni si veda Gelfond e Przymusinka (1989). Altri lavori sulla formalizzazione delle reti con eccezioni in logica non monotona sono ad esempio (Doherty 1989) e (Gregoire 1989).

6.3 Semantica preferenziale delle logiche non monotone

Esistono varie relazioni fra i diversi tipi di logica non monotona, e vari risultati di equivalenza sono stati dimostrati, sui quali qui non ci soffermeremo⁸. E' tuttavia interessante ai nostri scopi prendere in considerazione la proposta di Shoham (1987, 1988), che ha elaborato un quadro generale di tipo semantico allo scopo di offrire una visione unitaria del ragionamento non monotono, in maniera da ricondurvi e da generalizzare le logiche non monotone esistenti⁹. La proposta di Shoham si basa sul concetto di *modello preferito* (*preferred model*). Dal punto di vista intuitivo, l'idea di fondo è che, ragionando in maniera non monotona, si prende in considerazione soltanto un sottoinsieme dei modelli della teoria, quelli che godono di alcune caratteristiche particolari, che sono cioè "preferiti" sotto qualche punto di vista. Ciò è evidente nella *circumscription*, dove sono preferiti quei modelli in cui i predicati da circoscrivere hanno estensione minore. Di fatto, la teoria di Shoham nasce proprio come una generalizzazione della semantica dei modelli minimali di McCarthy per la *circumscription*. Criteri di preferenza diversi danno origine a logiche non monotone diverse. La teoria che ne risulta viene detta teoria della *semantica preferenziale* (*preferential semantics*). In generale, il motivo per cui restringendosi a un sottoinsieme di modelli preferiti di una teoria si ottengano conseguenze non monotone è semplice. Siano Δ e Γ insiemi di formule. In logica monotona, una formula α segue da Δ se e soltanto se α è vera in tutti i modelli di Δ . Poiché tutti i modelli di $\Delta \cup \Gamma$ sono anche modelli di Δ , allora, se α segue da Δ , α segue anche da $\Delta \cup \Gamma$. In una logica non monotona, α segue (non monotonicamente) da Δ se e soltanto se α è vera in tutti i modelli *preferiti* di Δ . Ma nulla ci assicura che i modelli preferiti di $\Delta \cup \Gamma$ siano un sottoinsieme dei modelli preferiti di Δ . Quindi non è detto che α segua da $\Delta \cup \Gamma$.

Shoham formula la sua teoria in termini molto generali. Sia \mathbf{L} una logica standard. Per logica standard Shoham intende una logica monotona classica, ad esempio la logica proposizionale o la logica dei predicati del primo ordine, oppure una logica modale, proposizionale o predicativa. Si ottiene una logica non monotona a partire da \mathbf{L} definendo un ordinamento di preferenza sulle *interpretazioni* di \mathbf{L} . Per *interpretazione* Shoham intende "qualunque cosa che stia alla sinistra di \models ", dove \models è l'usuale simbolo di conseguenza semantica. Quindi, si deve intendere come interpretazione sia una interpretazione in senso stretto (ad esempio, una interpretazione tarskiana per la logica del primo ordine), sia, ad esempio, nel caso di una logica modale, una coppia (I, w) , dove I è una struttura di Kripke e w è un mondo possibile. Questo consente, ad esempio, di ricondurre al quadro della semantica preferenziale anche le logiche autoepistemiche, che come abbiamo visto, si basano su un linguaggio modale e hanno una semantica a mondi possibili. Sia dunque \preceq un ordinamento parziale stretto definito sulle interpretazioni di \mathbf{L} , e M_1 e M_2 siano interpretazioni di \mathbf{L} . $M_1 \preceq M_2$ significa che M_2 è *preferita* rispetto a M_1 . \mathbf{L} assieme a \preceq definiscono una nuova *logica preferenziale* L_{\preceq} . Siano α e β formule del linguaggio di \mathbf{L} . I concetti di *soddisfazione* e di *conseguenza logica* per L_{\preceq} sono definiti come segue.

Definizione: un'interpretazione M soddisfa preferenzialmente α (in simboli, $M \models_{\preceq} \alpha$) se $M \models \alpha$, e se non esiste nessuna altra interpretazione M' tale che $M \preceq M'$ e $M' \models \alpha$. In tal caso si dice che M è un *modello preferito* di α .

Definizione: β è *conseguenza preferenziale* di α (in simboli, $\alpha \models_{\preceq} \beta$) se, per ogni interpretazione M , se $M \models_{\preceq} \alpha$, allora $M \models \beta$ (o, in altri termini, se i modelli di β , preferiti o meno, sono un soprainsieme dei modelli preferiti di α).

⁸Si vedano ad esempio (Imielinsky 1987), (Etherington 1987), (Konolige 1988, 1989).

⁹Si veda anche (Lin e Shoham 1992).

E' agevole constatare che i vari tipi di *circumscription* risultano casi particolari di logica preferenziale ottenuti a partire dalla logica dei predicati del primo ordine. Le relazioni di ordinamento fra i modelli introdotte dalla semantica dei modelli minimali possono essere viste come relazioni di preferenza in una semantica preferenziale. Shoham ad esempio fornisce la seguente definizione di preferenza per la *predicate circumscription*. Fare la *predicate circumscription* di un predicato P ad un argomento significa adottare un ordinamento per cui si ha $M_1 \preceq M_2$ se:

- 1) M_1 e M_2 coincidono per l'interpretazione delle costanti diverse da P ;
- 2) per ogni x , se $M_2 \models P(x)$, allora $M_1 \models P(x)$;
- 3) esiste un y tale che $M_1 \models P(y)$ ma $M_2 \not\models P(y)$.

E' facile constatare che tale definizione è sostanzialmente equivalente a quella di McCarthy data in precedenza. Oltre che alla *circumscription*, la semantica preferenziale è stata applicata ad altri sistemi non monotoni, fra cui la logica autoepistemica e la *default logic* (si veda anche Lin e Shoham 1990).