Informație implicită (Defaults)

Deși FOL este un limbaj de reprezentare foarte expresiv, nu oferă tipuri variate de raționamente. De exemplu, presupunând că într-o bază de cunoștințe KB avem Dog(fido), există doar două posibilități de a ajunge la concluzia Carnivore(fido):

1. acest lucru este menționat explicit în KB;
2. în KB avem implicația universală $\forall x. Dog(x) \supset Carnivore(x)$

Problema este cum să exprimăm ce știm în general și în particular despre indivizi. De exemplu, putem spune că

"viorile au patru coarde"

pentru a le distinge de chitare, care au 6 coarde. Dar nu înseamnă că vrem să spunem

"toate viorile au patru coarde"

deoarece am exclude o vioară ce are o coardă adăugată sau una scoasă.

O variantă ar fi să spunem

"toate viorile care nu sunt P_1 sau P_2 ... sau P_n au patru coarde".

Dorim să facem distincție între proprietățile <u>universale</u> ce au loc pentru toate instanțele și sunt ușor de exprimat în FOL și proprietățile <u>generice</u>, care au loc "în general".

Marea parte a cunoștințelor pare să fie de tip generic, deci este important să considerăm formalismul ce permite manipularea cunoștințelor generale, dar nu universale.

Inferență bazată pe informație implicită

În general, știm despre câini că sunt carnivori. Dacă Fido este un câine, în ce circumstanțe este potrivit să deducem că Fido este carnivor? Putem răspunde astfel:

Știind că un P este în general un Q și dat fiind că P(a) este adevărat, putem conclude că Q(a) este adevărat dacă nu cumva există un motiv bun să nu fie adevărat.

Dacă tot ce se știe despre un individ a este că e o instanță a lui P, atunci nu ne împiedică nimic să deducem că a este o instanță și a lui Q.

De exemplu, dacă știm că un urs polar s-a tăvălit în noroi probabil că nu dorim să tragem vreo concluzie cu privire la culoarea lui. Dar dacă tot ce știm despre un individ este că e urs polar, atunci este rezonabil să deducem că este alb.

Concluzia nu are garanția corectitudinii logice, ci este doar o informație implicită rezonabilă.

Această formă de raționament ce aplică informații generale (nu universale) unui individ particular se numește inferență bazată pe informație implicită.

Exemple de situații când dorim să deducem $Q(a)$ din $P(a)$:

Afirmații generale
- normale: copiilor le place joaca
- prototipice: portocalele sunt portocalii; bufnițele vânează noaptea
- statistice: oamenii din sala de așteptare devin nerăbdători

Lipsa informației contrare
- familiaritate: nicio țară nu are un președinte mai înalt de 2 m
- încredere în grup: copiii învață ușor limbi străine

Convenții
- conversaționale: "cea mai apropiată stație de benzină este la două străzi de aici" - deducerea implicită este că stația e deschisă
- de reprezentare: limita de viteză în oraș

Persistență
- inerție: statut marital
- timp: dimensiunea unui obiect

Lista de exemple nu este exhaustivă, dar sugerează marea varietate de surse de informații implicite.

Intrebare: In absența universalilor, cum putem determina când este potrivit să tragem concluzii implicite?

Raționamentul deductiv obișnuit este monoton, adică faptele noi pot produce doar cunoștințe adiționale. Dacă $KB_1 \models \alpha$, atunci $KB_2 \models \alpha$ pentru orice KB_2 cu $KB_1 \subseteq KB_2$.

Raționamentul implicit este nemonoton, adică faptele noi pot invalida cunoștințe anterioare (putem considera implicit că o pasăre zboară până când aflăm că pasărea respectivă este struț). Din acest motiv, raționamentul implicit se mai numește raționament nemonoton.

I Raționamentul lumii închise (closed-world reasoning)

Este prima (dpv cronologic) și cea mai simplă formalizare a raționamentului implicit. Faptele despre lume sunt reprezentate printr-un vocabular finit de predicate și constante. Însă din totalitatea propozițiilor ce ar putea fi formate, se așteaptă ca doar un număr mic să fie adevărate. O convenție de reprezentare este ca să fie reprezentate explicit propozițiile atomice adevărate, iar orice propoziție nemenționată să fie falsă.

KB
- DirectConnect (cleveland, toronto)
- DirectConnect (toronto, chicago)
- DirectConnect (cleveland, vancouver)

Convenția este ca dacă un zbor între două orașe nu este afișat, atunci el nu există.

In presupunerea lumii închise (close-world assumption CWA) se consideră că dacă nu se știe că o propoziție atomică este adevărată atunci ea se presupune a fi falsă.

O propoziție care se presupune a fi falsă, poate fi ulterior adevăr[ată]

Def $KB^{+} = KB \cup \{\neg p \mid p$ este un atom și $KB \not\models p\}$

Spunem că $KB \models_{c} \alpha$ dnd $KB^{+} \models \alpha$.

În exemplul anterior, KB^+ ar conține propoziții de tipul
$\neg DirectConnect(c_1, c_2)$.

<u>Def</u>. Spunem că o KB are cunoștințe <u>consistente</u> dnd nu există nicio propoziție α pentru care atât α cât și $\neg\alpha$ să fie cunoscute (deduse).
KB are cunoștințe <u>complete</u> dnd pentru orice propoziție α din vocabular, α sau $\neg\alpha$ este cunoscut.

Cunoștințele pot fi incomplete. De exemplu, dacă $KB = \{(p \vee q)\}$, din KB nu se poate deduce nici p nici $\neg p$.

Dar $KB \models_c \alpha$ sau $KB \models_c \neg\alpha$ (se dem prin inducție după lungimea lui α). Deci în CWA, relația de implicație este completă.

Dacă o KB este completă, atunci are proprietatea că dacă ne spune că una din două propoziții este adevărată, trebuie să precizeze care anume.
Ori de câte ori $KB \not\models p$, atunci fie $KB \models \neg p$ direct sau se presupune că $\neg p$ este ceea ce se intenționează a fi și este adăugat la KB.

<u>Evaluarea întrebărilor</u>

Întrebarea $KB \models_c \alpha$ se reduce la o colecție de întrebări despre literalii din α:

1. $KB \models (\alpha \wedge \beta)$ dnd $KB \models \alpha$ și $KB \models \beta$.
2. $KB \models \neg\neg\alpha$ dnd $KB \models \alpha$.
3. $KB \models \neg(\alpha \vee \beta)$ dnd $KB \models \neg\alpha$ și $KB \models \neg\beta$.
4. $KB \models_c (\alpha \vee \beta)$ dnd $KB \models_c \alpha$ sau $KB \models_c \beta$.
5. $KB \models_c \neg(\alpha \wedge \beta)$ dnd $KB \models_c \neg\alpha$ sau $KB \models_c \neg\beta$.
6. Dacă KB^+ este consistentă, atunci $KB \models_c \neg\alpha$ dnd $KB \not\models_c \alpha$.

De exemplu, $KB \models_c ((p \wedge q) \vee \neg(r \wedge \neg s))$ dnd

$$(KB \models_c p \text{ și } KB \models_c q) \text{ sau } KB \models_c \neg r \text{ sau } KB \models_c s.$$

Deci KB este consistentă, nu înseamnă că și KB^+ este consistentă.

$KB = \{(p \vee q)\}$

deoarece $KB \not\models p$ avem că $\neg p \in KB^+$ și similar $\neg q \in KB^+$.

Deci $KB^+ \Rightarrow \{(p \vee q), \neg p, \neg q\}$ și nu este consistentă.

Dar dacă KB este alcătuită doar din propoziții atomice (e.g. DirectConn) sau conține conjuncții de propoziții atomice (e.g. $p \wedge q$) sau disjuncț de literali negativi (e.g. $\neg p \vee \neg q$), atunci KB^+ este consistentă.

<u>Def</u>. Presupunerea generalizată a lumii închise (Generalized CWA)

$KB \models_{GC} \alpha$ dnd $KB^* \models \alpha$, unde KB^* se definește astfel:

$$KB^* = KB \cup \{\neg p \mid \text{pentru toți atomii } q_1, \dots q_n, \text{ dacă } KB \models (p \vee q_1 \vee \dots \vee q_n) \text{ atunci } KB \models (q_1 \vee \dots \vee q_n)\}$$

Deci nu vom presupune că p este fals dacă există o disjuncție dedusă de atomi incluzând p, ce nu poate fi redusă la o disjuncție dedusă ce nu conține p.

De exemplu, dacă $KB = \{(p \vee q)\}$ avem $KB \models (p \vee q)$ dar $KB \not\models q$

Deci $\neg p \notin KB^*$ și similar $\neg q \notin KB^*$.

Dacă luăm un atom r, vom avea $\neg r \in KB^*$ deoarece $KB \models (r \vee p \vee q)$ și $KB \models (p \vee q)$.

Un exemplu de interpretare: știm că există un zbor de la Cleveland către Dallas sau Houston. Deci deducem că există un zbor de la Cleveland către Dallas, Houston sau Austin. Dar deoarece există un zbor către unul din primele două orașe, în GCWA presupunem că nu există niciun zbor către Austin.

Implicațiile în GCWA sunt o submulțime a implicațiilor în CWA.

Dacă $\neg p \in KB^*$ atunci $\neg p \in KB^+$.

Dacă KB nu are cunoștințe disjunctive (adică dacă $KB \models (q_1 \vee \ldots \vee q_n)$ atunci $KB \models q_i$ pentru un i), atunci GCWA și CWA sunt în acord total.

Dacă KB este consistentă, atunci KB^* este consistentă.

Deci GCWA este o versiune mai slabă a CWA ce este în acord cu CWA în absența disjuncțiilor, dar rămâne consistentă în prezența lor.

Închiderea domeniului

Să presupunem că limbajul de reprezentare conține predicatul DirectConnect și constantele $c_1, \ldots c_n$.

Dacă KB conține doar propoziții atomice de forma $DirectConnect(c_i, c_j)$ atunci KB^+ va conține fie $DirectConnect(c_i, c_j)$ fie $\neg DirectConnect(c_i, c_j)$, $\forall\, i, j = 1, n$.

Să presupunem că există smallCity ce nu are aeroport, deci $\forall c_j$ $\neg DirectConnect(c_j, smallCity) \in KB^+$.

Considerăm întrebarea $\neg \exists x\, DirectConnect(x, smallCity)$.

În CWA nu sunt deduse nici întrebarea nici negarea ei. CWA exclude toate orașele $c_1 \ldots c_n$ dar nu exclude orașele nenumite ce nu sunt în legătură cu smallCity. Adică există o interpretare în care domeniul conține un oraș nenumit de niciun c_i, astfel încât orașul și denotarea lui smallCity să se afle în relația DirectConnect.

Deci vrem să presupunem că niciun oraș nu este conectat cu smallCity, nu doar $c_1, \ldots, c_n$.

Def. Presupunerea lumii închise, cu închiderea domeniului:

$$KB \models_{CD} \alpha \text{ dnd } KB^{\diamond} \models \alpha, \text{ unde}$$

$KB^{\diamond} = KB^+ \cup \{\forall x [x = c_1 \vee \ldots \vee x = c_n]\}$, cu $c_1, \ldots, c_n$ toate constantele ce apar în KB.

În KB^c se face presupunerea suplimentară că nu există alte obiecte în afara constantelor denumite $c_1, \ldots c_n$.

Revenind la exemplu, $\neg \exists x$ DirectConnect (x, smallCity) este dedusă în CWA cu închiderea domeniului deoarece $\neg$DirectConnect(c_i, smallCity) este dedusă în CWA pentru orice c_i.

Proprietățile principale ale CWA cu închiderea domeniului sunt:

$KB \models_{CD} \forall x \alpha$ dnd $KB \models_{CD} \alpha^x_c$ pentru orice constantă c din KB;

$KB \models_{CD} \exists x \alpha$ dnd $KB \models_{CD} \alpha^x_c$ pentru un c din KB.

Acum avem $KB \models_{CD} \alpha$ sau $KB \models_{CD} \neg\alpha$ pentru orice α (cu sau fără cuantificatori)

II Circumscrierea

O modalitate de a marca situațiile excepționale în care informația implicită nu trebuie aplicată este prin intermediul unui predicat notat Ab (abnormal).

$$\forall x [\text{Bird}(x) \wedge \neg \text{Ab}(x) \supset \text{Flies}(x)]$$

Dacă în KB am avea în plus faptele:

Bird(chilly), Bird(tweety), (tweety $\neq$ chilly), $\neg$ Flies(chilly)

am dori să tragem concluzia că Tweety zboară, dar Chilly nu.

Dar $KB \not\models$ Flies(tweety) deoarece există interpretări ce satisfac KB și în care Flies(~~tweety~~) este fals. In aceste interpretări, denotarea lui Tweety este inclusă în interpretarea predicatului Ab.

Strategia de minimizare a anormalității: vom considera interpretările ce satisfac KB în care interpretările predicatului Ab sunt (mulțimi) cât se poate de mici. Deci concluziile implicite sunt cele adevărate în modelele (interpretările) KB în care cât mai puțini indivizi sunt anormali.

În exemplul anterior, știm despre Chilly că este o pasăre "anormală", dar nu știm nimic despre Tweety. Interpretarea predicatului Ab include Chilly (că trebuie) dar exclude Tweety (nimic nu dictează că Tweety ar trebui inclus). Această tehnică s.n. circumscrierea predicatului Ab.

În general, se folosește o familie de predicate Ab_i pentru a trata diverse aspecte ale indivizilor. Chilly poate fi în interpretarea lui Ab_1 dar nu și în interpretarea lui Ab_2 ș.a.m.d.

<u>Implicație minimală</u>

Implicația este caracterizată în termenii proprietăților interpretărilor.

Fie P o mulțime fixată de predicate unare Ab. Fie $J_1 = \langle D, I_1 \rangle$ și $J_2 = \langle D, I_2 \rangle$ două interpretări pe același domeniu astfel încât constantele și funcțiile sunt interpretate la fel.

Definim relația $\leq$ astfel:

$$J_1 \leq J_2 \text{ dnd } \forall P_r \in P \text{ avem } I_1[P_r] \subseteq I_2[P_r].$$

$J_1 < J_2$ dnd $J_1 \leq J_2$ și $J_2 \not\leq J_1$.

Spunem că J_1 micșorează extensia (adică interpretarea) predicatelor Ab. Sau, echivalent, J_1 este mai "normală" decât J_2.

<u>Def</u>. Implicația minimală $\models_{\leq}$ se definește astfel:

$KB \models_{\leq} \alpha$ dnd pentru orice interpretare J a.î. $J \models KB$,
fie $J \models \alpha$ sau există $J' < J$ cu $J' \models KB$.

În exemplul anterior, $KB \not\models Flies(tweety)$ dar $KB \models_{\leq} Flies(tweety)$.

Dacă $J \models KB$ dar $J \not\models Flies(tweety)$, atunci $J \models Ab(tweety)$.

Luăm J' identică cu J, cu excepția ștergerii denotării lui tweety din extensia predicatului Ab. Presupunând că $P = \{Ab\}$, avem că $J' < J$ și $J' \models KB$.

Dar $J' \not\models Ab(tweety) \Longrightarrow J' \models Flies(tweety)$
(deoarece J' satisface implicația din KB)

Deci, în modelele minimale ale KB, Tweety este o pasăre normală
$KB \models_{\leq} \neg Ab(tweety)$, de unde putem deduce că Tweety zboară.

Însă în toate modelele KB, Chilly este o pasăre anormală.
Deci în acest raționament, singurul pas implicit a fost să deducem
că Tweety este o pasăre normală (restul este raționament obișnuit).

<u>Obs</u>. Cele mai "normale" modele ale KB nu satisfac exact aceleași
propoziții.

De exemplu, dacă avem:

$$KB \left[\begin{array}{l} Bird(c) \\ Bird(d) \\ (\neg Flies(c) \vee \neg Flies(d)) \\ \forall x [Bird(x) \wedge \neg Ab(x) \supset Flies(x)] \end{array} \right.$$

În orice model al KB, interpretarea lui Ab trebuie să conțină
fie denotarea lui c fie denotarea lui d. Dacă modelul ar conține
și alți indivizi anormali, nu ar mai fi minimal.

Deci în orice $J \models KB$, avem fie $J \models Ab(c)$, fie $J \models Ab(d)$.

Dacă $J \models Ab(c)$, atunci $KB \not\models_{\leq} Flies(c)$.

Similar, dacă $J \models Ab(d)$ atunci $KB \not\models_{\leq} Flies(d)$.

Dar putem deduce implicit că $KB \models_{\leq} Flies(c) \vee Flies(d)$.

<u>Obs</u> CWA și GCWA au un comportament diferit

Deoarece $KB \not\models Ab(c)$ și nici $KB \not\models Ab(d)$, rezultă că

$$KB^{+} \supset KB \cup \{\neg Ab(c), \neg Ab(d)\}.$$

Deci $KB \models_{c} (Flies(c) \wedge Flies(d))$, adică KB^{+} nu este consistentă.

În GCWA însă $\neg Ab(c) \notin KB^{*}$, $\neg Ab(d) \notin KB^{*}$

$$[\, KB \models Ab(c) \vee \neg Bird(c) \vee Flies(c) \text{ dar } KB \not\models \neg Bird(c) \vee Flies(c) \,]$$

Deci în GCWA nu putem deduce nimic legat de Flies(c) sau
Flies(d) (sau disjuncția lor).

In cazul circumscrierii, un model al KB este preferat față de alt model dacă are mai puține obiecte anormale.
Dacă avem următoarele afirmații: Richard Nixon a fost atât quaker (deci implicit pacifist) și republican (implicit nu este pacifist) în KB putem scrie astfel:

$$Republican(nixon) \wedge Quaker(nixon)$$

$$\forall x [Republican(x) \wedge \neg Ab_2(x) \supset \neg Pacifist(x)]$$

$$\forall x [Quaker(x) \wedge \neg Ab_3(x) \supset Pacifist(x)]$$

Dacă circumscriem Ab_2 și Ab_3, avem două modele minimale

$$\mathcal{I}_1 \models Ab_2(nixon) \text{ și } \mathcal{I}_1 \models Pacifist(nixon)$$

$$\mathcal{I}_2 \models Ab_3(nixon) \text{ și } \mathcal{I}_2 \models \neg Pacifist(nixon)$$

Deci $KB \not\models_{\leq} Pacifist(nixon)$ nici $KB \not\models_{\leq} \neg Pacifist(nixon)$

Dacă, de exemplu, acordăm prioritate convingerilor religioase în fața convingerilor politice, putem exprima acest fapt prin circumscriere prioritizată în care preferăm modelul (minimal) în care Ab_3 este minimizat.

III Logica implicită

Oferă un mecanism prin care se specifică explicit propozițiile ce trebuie adăugate la KB, cu îndeplinirea proprietății de consistență.
De exemplu, dacă KB deduce $Bird(t)$, probabil dorim să adăugăm presupunerea implicită $Flies(t)$, dacă se păstrează consistența KB.
O KB are două părți: o mulțime F de propoziții și o mulțime D de reguli implicite, ce specifică ce presupuneri se fac și când.
Rolul logicii ~~descriptive~~ implicite este să specifice:
- mulțimea adecvată de cunoștințe implicite ce incorporează fapte din F.
- cât ~~mai~~ multe presupuneri implicite posibile, plecând de la regulile implicite din D
- implicații logice deduse din cunoștințe și reguli.

Reguli implicite

O regulă se scrie sub forma $<\alpha : \beta/\delta>$, unde α este premisa, β justificarea și δ concluzia. δ este adevărat dacă α este adevărat și presupunerea că β este adevărat este consistentă (adică $\neg\beta$ nu este adevărat).

$$<Bird(tweety) : Flies(tweety)/Flies(tweety)>$$

Regulile în care justificarea este identică cu concluzia se numesc reguli implicite normale și se scriu $Bird(tweety) \Rightarrow Flies(tweety)$.

Regulile pot fi formulate și folosind variabile:

$<Bird(x) : Flies(x)/Flies(x)>$ este valabilă pentru orice termen t fără variabile (în general, nu universal)

Extensii implicite

Fiind dată o teorie implicită $KB = (\mathcal{F}, \mathcal{D})$, ce propoziții ar trebui considerate adevărate?

Mulțimea propozițiilor rezonabile se numește extensie a teoriei și se notează $\mathcal{E}$.

Dacă $\mathcal{E}$ este o extensie a teoriei implicite $(\mathcal{F}, \mathcal{D})$, atunci o propoziție $\pi \in \mathcal{E}$ dnd $\mathcal{F} \cup \{\delta \mid <\alpha : \beta/\delta> \in \mathcal{D}, \alpha \in \mathcal{E}, \neg\beta \notin \mathcal{E}\} \models \pi$.

Deci o mulțime de propoziții este o extensie dacă este mulțimea tuturor implicațiilor din $\mathcal{F} \cup \Delta$, unde Δ este o mulțime de presupuneri aplicabile (o presupunere δ este aplicabilă dacă $<\alpha : \beta/\delta> \in \mathcal{D}$, $\alpha \in \mathcal{E}$, $\neg\beta \notin \mathcal{E}$).

Definiția nu ne spune cum să găsim $\mathcal{E}$, dar $\mathcal{E}$ este caracterizat complet de mulțimea presupunerilor aplicabile Δ.

$$\mathcal{F} = \{Bird(tweety), Bird(chilly), \neg Flies(chilly)\}$$
$$\mathcal{D} = \{Bird(x) \Rightarrow Flies(x)\}.$$

Fie $\mathcal{E} = \mathcal{F} \cup \{Flies(tweety)\}$.

$Flies(tweety)$ este singura presupunere aplicabilă deoarece:

$$\left.\begin{array}{l} Bird(tweety) \in \mathcal{E} \\ \neg Flies(tweety) \notin \mathcal{E} \end{array}\right\} \Rightarrow Flies(tweety) \text{ presupunere aplicabilă} \Rightarrow \mathcal{E} \text{ extensie}$$

Flies(t) nu este aplicabilă pentru niciun alt t (în cazul nostru t ar putea fi doar chilly), deci Flies(tweety) este singura presupunere aplicabilă $\Rightarrow \mathcal{E}$ este singura extensie posibilă.

<u>Extensii multiple</u>

$\mathcal{F} = \{ Republican(nixon), Quaker(nixon) \}$

$\mathcal{D} = \{ Republican(x) \Rightarrow \neg Pacifist(x), Quaker(x) \Rightarrow Pacifist(x) \}$

Fie extensia $\mathcal{E}_1$ caracterizată de presupunerea Pacifist(nixon) și $\mathcal{E}_2$ caracterizată de presupunerea ¬ Pacifist(nixon).

$\mathcal{E}_1$ și $\mathcal{E}_2$ sunt extensii deoarece presupunerile lor sunt aplicabile și nu există alte presupuneri aplicabile pentru $t \neq nixon$.

Mulțimea vidă de presupuneri nu produce o extensie, deoarece atât Pacifist(nixon) și ¬ Pacifist(nixon) ar fi aplicabile.

Pentru alte extensii, presupunerile ar fi de forma Pacifist(t) sau ¬ Pacifist(t), dar niciuna nu este aplicabilă pentru $t \neq nixon$ deoarece nu avem premisele corespunzătoare Quaker(t) sau Republican(t) în $\mathcal{E}$. Deci $\mathcal{E}_1$ și $\mathcal{E}_2$ sunt singurele extensii posibile.

Logica implicită ne spune că putem alege să presupunem că Nixon este pacifist sau că nu este pacifist. Plecând de la KB, ambele presupuneri (luate pe rând) sunt rezonabile.

Avem două opțiuni:

1. Varianta sceptică - considerăm adevărate propozițiile ce se regăsesc în toate extensiile teoriei implicite;
2. Varianta credulă - considerăm adevărate propozițiile ce se regăsesc într-o extensie aleasă la întâmplare.

În unele situații, existența extensiilor multiple poate indica faptul că nu s-a spus suficient pentru a lua o decizie rezonabilă.

In exemplul anterior, putem spune că informația implicită în privința Quaker-ilor ar trebui să se aplice doar indivizilor ce nu sînt activi politic.
Dacă avem regula $\forall x [Republican(x) \supset Political(x)]$, putem înlocui regula implicită $Quaker(x) \Rightarrow Pacifist(x)$ cu regula ne-normală

$$\frac{Quaker(x) : [Pacifist(x) \wedge \neg Political(x)]}{Pacifist(x)}$$

Pentru Quaker-ii simpli, presupunerea este că sînt pacifiști. Dar pentru Quaker-ii republicani ca Nixon, vom presupune că nu sînt pacifiști.

Dacă înlocuim $\forall x [Republican(x) \supset Political(x)]$ cu regula implicită $Republican(x) \Rightarrow Political(x)$, atunci am avea două extensii:
- una caracterizată de presupunerile $\{\neg Pacifist(nixon), Political(nixon)\}$
- cealaltă caracterizată de presupunerea $\{Pacifist(nixon)\}$

Arbitrarea între reguli implicite aflate în conflict este foarte importantă atunci când avem de-a face cu ierarhii.
Dacă KB conține

$$\left[\begin{array}{l} \forall x [Penguin(x) \supset Birds(x)] \\ Bird(x) \Rightarrow Flies(x) \\ Penguin(x) \Rightarrow \neg Flies(x) \\ Penguin(chilly) \end{array}\right.$$

avem două extensii: una în care Chilly zboară și una în care nu zboară.
Implicația că pinguinii nu zboară ar trebui să fie prioritară celei mai generale că păsările zboară.

$$\frac{Bird(x) : [Flies(x) \wedge \neg Penguin(x)]}{Flies(x)}$$