Intelligenza Artificiale II

Oltre la logica classica

Marco Piastra

Oltre la logica classica?

- Per logica **classica** si intende:
  - La logica predicativa del primo ordine L_{PO}
  - La logica proposizionale L_P (che è contenuta, in senso proprio, in L_{PO})

- Una logica **non classica** adotta regole diverse
- Perchè?
  - Per rappresentare altre forme di ragionamento
    - Non solo deduttivo ma anche *abduttivo* ed *induttivo* (*vedi oltre*)
    - Forme speciali, legate ad obiettivi specifici, come logiche modali o temporali
  - Per esigenze applicative
    - Frammenti (sottoinsiemi) di L_{PO}, più efficacemente automatizzabili (p.es. Jess)

Regole diverse

a) Logica classica in sistemi logici diversi

Esempi:

assert e *retract* in Jess e in Prolog

Closed-World Assumption (CWA)

Negation As Failure (NAF) in Prolog

b) Altre estensioni (non monotone) della logica classica

Esempi:

Ragionamento plausibile

c) Ragionamento non deduttivo

Esempi:

Ragionamento *abduttivo*

Ragionamento *induttivo*

Regole diverse

d) Rappresentazione di nozioni speciali

Esempi:

Logiche modali

Logiche temporali

e) Estensione dei principi base della logica classica

Esempi:

Logiche multivalenti

Fuzzy Logics

Logiche probabilistiche

- Numerose correlazioni

  I diversi sistemi logici, malgrado le apparenze, sono fortemente correlati

- Un fattore comune

  Tutti i sistemi logici che vedremo possono essere *rappresentati* in L_{PO}

Logiche e sistemi logici

In ambito teorico (p.es. in logica matematica) una **logica** è definita da:

a) Linguaggio formale
b) Semantica del linguaggio formale
c) Relazioni $\models$ (*conseguenza*) e $\vdash$ (*derivazione*)

- In intelligenza artificiale

E` utile vedere un **sistema logico** come un *agente ragionatore*

- Basato su una **logica** di riferimento (p.es. L_{PO})
- Adotta una determinata strategia di calcolo (p.es. *SLD depth-first*)
- Può avere risorse limitate (di tempo o memoria)

Si ha quindi il concetto di **derivabilità** in un **sistema logico**

Notazione: $\Gamma \vdash_{<SysLog>} \varphi$ dove $<SysLog>$ indica un **sistema logico** particolare

Esempi:

Strategia SLD *fair* (definita solo per le clausole di Horn)

$$\Gamma \vdash_{LPO} \varphi \neq \Gamma \vdash_{SLD\ fair} \varphi \neq \Gamma \vdash_{SLD} \varphi$$

↑ Derivabilità generale in L_{PO}

Strategia SLD qualsiasi (e.g. *depth-first*)

In linea di principio, la strategia di calcolo di $<SysLog>$ può essere qualsiasi cosa:
p.es. $\Gamma \vdash_{NN} \varphi$ una rete neurale che stabilisce se φ è (NN) derivabile da Γ

Ragionamento plausibile

- In generale

Un ragionamento dove la **relazione** tra premesse e conseguenza
è <u>razionalmente plausibile</u> ma non necessariamente <u>corretta</u> (in senso logico-classico)

Notazione:
$\Gamma \mathrel{|\!\!\sim}_{<SysLog>} \varphi$ indica che φ è una derivazione **plausibile** (con premesse Γ) in un sistema $<SysLog>$

Principi generali della relazione $\mathrel{|\!\!\sim}_{<SysLog>}$:
$\Gamma \mathrel{|\!\!\sim}_{<SysLog>} \varphi \Rightarrow \Gamma \not\mathrel{|\!\!\sim}_{<SysLog>} \neg\varphi$ (plausibile)
$\Gamma \vdash_{<SysLog>} \varphi \Rightarrow \Gamma \mathrel{|\!\!\sim}_{<SysLog>} \varphi$ (razionale)
$\Gamma \mathrel{|\!\!\sim}_{<SysLog>} \varphi \not\Rightarrow \Gamma \vdash_{<SysLog>} \varphi \; (\Rightarrow \Gamma \models \varphi)$ (non necessariamente corretta)

Molto frequente in pratica:

L'orario ferroviario non riporta un treno per Milano alle 06:55,
quindi si assume che tale treno <u>non</u> esista

In generale, un database contiene solo informazione positiva (p.es. i treni esistenti)
L'informazione negativa è ricavata 'per difetto'

... anche *Defeasible Reasoning*

- Inferenza non monotona

  In generale, non vale la proprietà di monotonia

  $\Gamma \vdash_{<SysLog>} \varphi \;\not\Rightarrow\; \Gamma \cup \Delta \vdash_{<SysLog>} \varphi$

  L'arrivo di nuova informazione può infatti falsificare una precedente inferenza

  p.es. l'annuncio di un treno straordinario …

- Inferenza sistemica

  Nel caso del *modus ponens*, si ha uno schema di inferenza particolare di valore generale

  $\varphi \to \psi, \varphi \vdash \psi$

  E` sempre applicabile, non dipende dal contesto

  Al contrario, in generale, le inferenze plausibili dipendono da un'intera teoria Γ

  Tipicamente, la teoria rappresenta un sistema completo di conoscenze (p.es. un database)

  Qualsiasi modifica di Γ, in generale, può cambiare la relazione $\Gamma \vdash_{<SysLog>} \varphi$

Closed-World Assumption (CWA)

$\{\Gamma \not\models \alpha\} \vdash_{CWA} \neg\alpha$ (α atomo base – *ground atom*)

Esempio:

$\Pi \equiv \{\{Filosofo(socrate)\}, \{Filosofo(platone)\}, \{Mortale(felix)\}\}$

I fatti del programma Π possono essere riscritti in L_{PO} come:

$\forall x\ ((x = socrate) \rightarrow Filosofo(x))$
$\forall x\ ((x = platone) \rightarrow Filosofo(x))$
$\forall x\ ((x = felix) \rightarrow Mortale(x))$

La *CWA* equivale al *completamento* (plausibile) di Π:

$\forall x\ ((x = felix) \leftrightarrow Mortale(x))$
$\forall x\ ((x = socrate \vee x = platone) \leftrightarrow Filosofo(x))$

Notare la doppia implicazione

Inferenza plausibile

In questo caso:

$\Pi \vdash_{CWA} \neg Mortale(socrate)$
$\Pi \vdash_{CWA} \neg Mortale(platone)$
$\Pi \vdash_{CWA} \neg Filosofo(felix)$

CWA e regole

Esempio:

$$\Pi \equiv \{\{Umano(x), \neg Filosofo(x)\}, \{Mortale(y), \neg Umano(y)\},$$
$$\{Filosofo(socrate)\}, \{Filosofo(platone)\}, \{Mortale(felix)\}\}$$

La riscrittura dei fatti del programma Π è identica al caso precedente:

$$\forall x\ ((x = socrate) \rightarrow Filosofo(x))$$
$$\forall x\ ((x = platone) \rightarrow Filosofo(x))$$
$$\forall x\ ((x = felix) \rightarrow Mortale(x))$$

Il *completamento* di Π deve tener conto delle regole:

$$\forall x\ ((x = felix \lor x = socrate \lor x = platone) \leftrightarrow Mortale(x))$$
$$\forall x\ ((x = socrate \lor x = platone) \leftrightarrow Filosofo(x))$$
$$\forall x\ ((x = socrate \lor x = platone) \leftrightarrow Umano(x))$$

Inferenza plausibile

In questo caso:

$$\Pi \mid\!\sim_{CWA} \neg Umano(felix)$$
$$\Pi \mid\!\sim_{CWA} \neg Filosofo(felix)$$

Negazione come fallimento (*Negation as Failure* - *NF*)

Come noto, non tutte le fbf sono traducibili in clausole di Horn

In particolare:

$(\neg A \land B) \to C$

$\neg(\neg A \land B) \lor C$

$A \lor \neg B \lor C$ (clausola con due letterali positivi)

- Si legga il connettivo ¬ nella premessa della regola come 'non è noto'

  Esempio:

  $(\neg A \land B) \to C$

  Se A non è noto e B è vero, allora C

  Salvo *"diversa comunicazione"* e *"oggi è venerdì"*, allora *"c'è lezione di IA1"*

Risoluzione SLDNF

- Traduzione speciale di regole con premesse negative

  $(\neg A \wedge B) \rightarrow C$ si traduce come

  $\backslash + \neg A \vee \neg B \vee C$ (il simbolo \+ indica la *NF*)

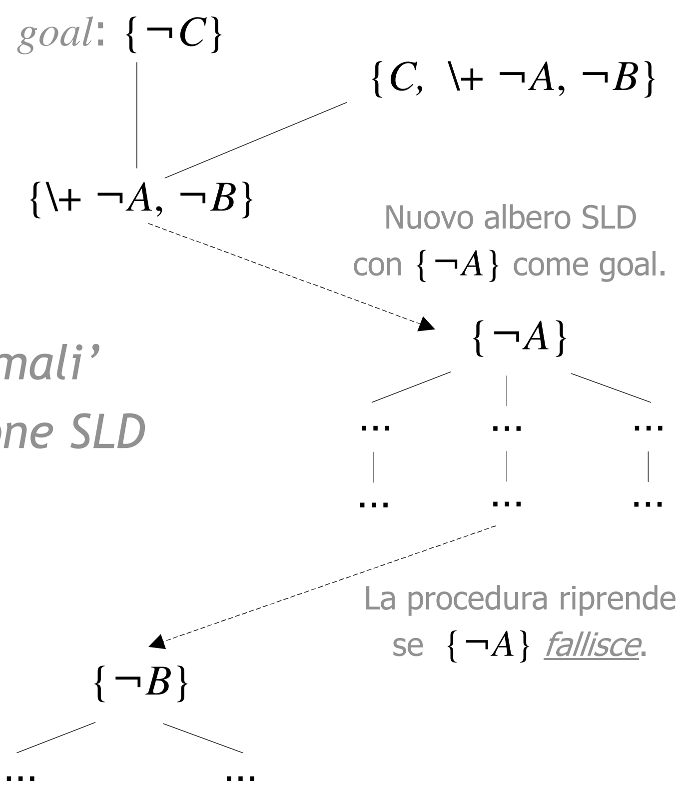
- Metodo

  Le premesse negative sono interpretate come goal la cui derivazione SLD deve *fallire*

  Il metodo SLDNF procede come SLD per i goal 'normali'

  Incontrando un $\backslash + \neg A$ *, si apre una nuova derivazione SLD*

  Il goal $\backslash + \neg A$ *ha successo se* $\neg A$ *fallisce*



Risoluzione SLDNF e completamento

- Completamento
  1) Dato un insieme Γ di clausole di Horn, si aggregano le regole in modo che ciascun atomo compaia al massimo in una implicazione

     Esempio: $\Gamma \equiv \{\{C\}, \{B, \neg F\}, \{B, \neg E\}, \{B, \neg D\}\}$

     Più comodo scrivere: $\{C, F \to B, E \to B, D \to B\}$

     $\Gamma' \equiv \{C, D \vee E \vee F \to B\}$
  2) Per ogni atomo φ non definito si aggiunge $false \to \varphi$ ($false$ indica la *contraddizione*)

     Nell'esempio: $\Gamma'' \equiv \{C, D \vee E \vee F \to B, false \to D, false \to E, false \to F\}$
  3) Si sostituisce l'implicazione $\to$ con l'equivalenza $\leftrightarrow$

     $Comp(\Gamma) \equiv \{C, D \vee E \vee F \leftrightarrow B, false \leftrightarrow D, false \leftrightarrow E, false \leftrightarrow F\}$ (completamento di Γ)

- Correttezza di SLDNF (Clark, 1974)

  Il goal SLDNF φ ha successo per $\Gamma \quad \Rightarrow \quad Comp(\Gamma) \models \varphi$

  Esempio:

  $\Gamma \equiv \{\{C\}, \{B, \neg F\}, \{B, \neg E\}, \{B, \neg D\}\}$

  Il goal $\neg B$ ha successo SLDNF perchè $\neg B$ *fallisce* in SLD

  In effetti: $\{C, D \vee E \vee F \leftrightarrow B, false \leftrightarrow D, false \leftrightarrow E, false \leftrightarrow F\} \models \neg B$

Lo strano caso della SLDNF

- Non è completa

$$Comp(\Gamma) \models \neg\varphi \ \not\Rightarrow \ \Gamma \vdash_{SLDNF} \neg\varphi$$

Infatti la SLDNF può andare in loop

Esempio:

$\Gamma \equiv \{Q \rightarrow P, \neg Q \rightarrow P, Q \rightarrow Q\}$

$Comp(\Gamma) \equiv \{Q \vee \neg Q \leftrightarrow P, Q \leftrightarrow Q\}$ $\qquad Comp(\Gamma) \models P$

Ma la SLDNF va in loop: si dovrebbe mostrare che la SLD *fallisce* per $\Gamma \cup \{\neg P\}$

Non dipende dalle *selection function*, anche sia *fair*

- Non è più logica classica

Non vale più la proprietà di monotonia sintattica

$$\Gamma \vdash_{SLDNF} \neg\varphi \ \not\Rightarrow \ \Gamma \cup \Delta \vdash_{SLDNF} \neg\varphi$$

Esempio:

$\{C, D \vee E \vee F \rightarrow B\} \vdash_{SLDNF} \neg B$

$\{C, D \vee E \vee F \rightarrow B\} \cup \{D\} \nvdash_{SLDNF} \neg B$

Aggiungendo informazione, il risultato della NF può cambiare

Relazione tra CWA, NAF e SLDNF

- *Closed-World Assumption* (CWA)

  $\{\Gamma \not\models \alpha\} \vdash_{CWA} \neg\alpha$ (α atomo base – *ground atom*)
  Notare che $\Gamma \not\models \alpha$ non è decidibile in L_{PO}, quindi nemmeno la relazione $\vdash_{CWA}$

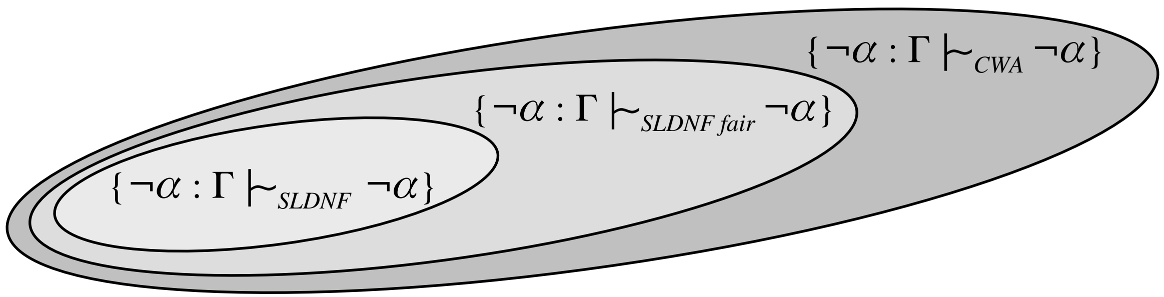
- *Negation as Failure* (NAF)

  $\{\alpha \in FF(\Gamma)\} \vdash_{SLDNF\ fair} \neg\alpha$
  Se per α esiste **un albero SLD finito** di $\Gamma \cup \{\neg\alpha\}$ che fallisce, si assume $\neg\varphi$
  (solo una procedura SLD *fair*, cioè completa, lo trova certamente)

- SLDNF

  $\{\alpha \in FF_{SLD}(\Gamma)\} \vdash_{SLDNF} \neg\alpha$
  Se la prova di $\Gamma \cup \{\neg\alpha\}$ fallisce con una strategia SLD si assume $\neg\alpha$
  (si intende una strategia SLD qualsiasi, non necessariamente *fair*)

Relazioni di inclusione
tra insiemi di clausole derivabili



Forme specifiche

Le forme di inferenza **plausibile** CWA, NAF e SLDNF sono di carattere generale

Altri sistemi logici adottano forme più **specifiche**, applicabili solo a casi particolari

- *Circumscription* (McCarthy, 1980)

  (≈ *CWA* selettiva, applicata solo a specifici predicati)

  Esempio:

  $\Pi \equiv \{\{Filosofo(socrate)\}, \{Filosofo(platone)\}, \{Mortale(felix)\}\}$

  Applicando la *circumscription* al predicato $Filosofo/1$ si avrebbe il completamento:

  $\forall x\ ((x = socrate \lor x = platone) \leftrightarrow Filosofo(x))$

  Da cui:

  $\Pi \mid\!\sim \neg Filosofo(felix)$

  Si specifica al contrario, tramite il predicato speciale $Abnormal/1$

  (≈ CWA si applica a tutti gli oggetti che non sono noti essere $Abnormal$)

  Esempio:

  $\Pi \equiv \{\{\forall x\ ((Filosofo(x) \land \neg Abnormal(x)) \rightarrow Umano(x)\}, \{Filosofo(socrate)\}, \{Filosofo(platone)\},$
  $\{Abnormal(felix)\}, \{Filosofo(felix)\}\}$

  Da cui:

  $\Pi \mid\!\sim Umano(socrate)$
  $\Pi \mid\!\sim Umano(platone)$
  ma $\Pi \not\mid\!\sim Umano(felix)$

(In realtà, la relazione tra *circumscription* e *CWA* è molto più complessa)

Regole specifiche

- *Default Logic* (Reiter, 1980)

  Si usano regole specifiche, dove le premesse (*vere* e solo *plausibili*) sono esplicitamente indicate

  $$\frac{\langle \textit{verità} \rangle : \langle \textit{plausibilità} \rangle}{\langle \textit{default} \rangle}$$ *schema informale*

  $$\frac{\alpha : \beta_1, \beta_2, \ldots, \beta_n}{\gamma}$$ *schema formale*

  Dato Γ, se $\Gamma \models \alpha$ e $\Gamma \not\models \neg\beta_1,\ \Gamma \not\models \neg\beta_2, \ldots, \Gamma \not\models \neg\beta_n$ allora $\Gamma \mid\!\sim_D \gamma$ (si assume γ per *default*)

  Vale a dire se α è vera (in Γ) e $\beta_1, \beta_2, \ldots, \beta_n$ sono *consistenti* (in Γ)

  Esempi:

  La *circumscription* di $Abnormal/1$ si esprime come

  $$\frac{true : \neg Abnormal(x)}{\neg Abnormal(x)}$$

  Dall'esempio precedente:

  $$\frac{Filosofo(x) : Umano(x)}{Umano(x)}$$

  da cui $\{Filosofo(felix)\} \mid\!\sim_D Umano(felix)$

  In alternativa:

  $$\frac{Filosofo(x) \wedge \neg Gatto(x) : Umano(x)}{Umano(x)}$$

  da cui $\{Filosofo(felix), Gatto(felix)\} \not\mid\!\sim_D Umano(felix)$

Forme di inferenza (C. S. Peirce)

Schema di inferenza

- Inferenza *deduttiva*
  - a) i fagioli di questo sacco sono bianchi
  - b) questi fagioli provengono da questo sacco

    QUINDI
  - c) questi fagioli sono bianchi

modus ponens $\dfrac{\varphi \rightarrow \psi \quad \varphi}{\psi}$ **corretto**

- Inferenza *abduttiva*
  - a) i fagioli di questo sacco sono bianchi
  - b) questi fagioli sono bianchi

    QUINDI
  - c) questi fagioli provengono da questo sacco

$\dfrac{\varphi \rightarrow \psi \quad \psi}{\varphi}$ **plausibile**

- Inferenza *induttiva*
  - a) questi fagioli provengono da questo sacco
  - b) questi fagioli sono bianchi

    QUINDI
  - c) i fagioli di questo sacco sono bianchi

$\dfrac{\varphi \quad \psi}{\varphi \rightarrow \psi}$ **plausibile**

Abduzioni come ipotesi esplicative

- La logica di base è la logica **classica**
  - E` invece diverso il **tipo** di ragionamento
    - e quindi il **tipo** di calcolo utilizzato

- In generale, in un ragionamento abduttivo:
  - Un *modello* (o descrizione astratta) rappresentato da una teoria K
  - Un insieme di *osservazioni* specifiche rappresentate da un insieme di fbf Σ
  - In generale, $K \not\models \Sigma$ (dalla teoria generale K non conseguono le osservazioni plausibili)
  - Si cerca è un'ipotesi Δ tale per cui

    $$K \cup \Delta \models \Sigma$$

    intuitivamente, Δ descrive le *ipotesi* che **spiegano** Σ

Esempio

- Modello (K)

  K_1: $batteriaScarica \rightarrow (\neg funzionanoLuci \wedge \neg funzionaAutoradio \wedge \neg motorinoGira)$

  K_2: $motorinoGuasto \rightarrow \neg motorinoGira$

  K_3: $\neg motorinoGira \rightarrow \neg macchinaParte$

  K_4: $serbatoioVuoto \rightarrow (indicatoreAZero \wedge \neg macchinaParte)$

- Osservazioni (Σ)

  Σ_1: $\neg macchinaParte$

- Possibili *ipotesi* (Δ)

  Δ_1: $batteriaScarica$ $(\{K_1, K_3\} \cup \{\Delta_1\} \models \Sigma_1)$

  Δ_2: $motorinoGuasto$ $(\{K_2, K_3\} \cup \{\Delta_2\} \models \Sigma_1)$

  Δ_3: $serbatoioVuoto$ $(\{K_4\} \cup \{\Delta_3\} \models \Sigma_1)$

Ipotesi e vincoli

- Ipotesi *plausibili*
  - Le ipotesi Δ devono essere consistenti con la teoria e le osservazioni
    - $K \cup \Delta \cup \Sigma$ deve essere soddisfacibile
  - Le ipotesi Δ devono spiegare tutte le osservazioni Σ
    - Alcune ipotesi, tuttavia, implicano anche altre osservazioni:
    - $batteriaScarica \rightarrow (\neg funzionanoLuci \wedge \neg funzionaAutoradio \wedge \neg motorinoGira)$
  - Le ipotesi Δ devono essere *minimali*
    - Non deve esistere un $\Delta^* \subset \Delta$ tale per cui $K \cup \Delta^* \vdash \Sigma$
    - Le ipotesi devono limitarsi ai soli elementi indispensabili per spiegare Σ
- Rilevanza
  - Notare che:
    $K \cup \{\neg macchinaParte\} \models \neg macchinaParte$
    - L'ipotesi è plausibile ma anche inutile: non ha valore esplicativo
  - Le ipotesi devono risalire ad elementi che abbiano un valore causale
    - La cui definizione spesso dipende dal tipo di ragionamento ...

Scelta tra ipotesi

- Ipotesi multiple possono coesistere

  Le due ipotesi $\{serbatoioVuoto\}$ e $\{batteriaScarica\}$ sono consistenti (possono coesistere)

  $K \cup \{serbatoioVuoto, batteriaScarica\} \models \neg macchinaParte$

- Altre ipotesi possono essere alternative

  Le due ipotesi $\{serbatoioVuoto, funzionaAutoradio\}$ e $\{batteriaScarica\}$ sono plausibili ma mutuamente esclusive

  $K \cup \{serbatoioVuoto, funzionaAutoradio, batteriaScarica\}$ è inconsistente

- Strategie di scelta

  Spesso è utile ridurre il numero delle ipotesi (p.es. quando si cerca un rimedio)

  Acquisizione di nuove osservazioni

  Partendo dalle ipotesi $\{serbatoioVuoto\}$, $\{batteriaScarica\}$ e $\{motorino\ Guasto\}$ l'acquisizione dei valori di *funzionaAutoradio*, *motorinoGira* e *indicatoreAZero* permette di scegliere (purchè tali fatti siano *osservabili*)

  Criteri particolari

  Scelta basata sul *costo* delle osservazioni

  Scelta basata sul *rischio* associato alle ipotesi

Abduzione e modelli

Il modello K rappresenta l'informazione che l'agente usa per spiegare le osservazioni

La scelta del tipo modello K è fondamentale per le applicazioni pratiche

- Schema-based reasoning (SBR)

    Il mondo è regolare: il comportamento atteso (p.es. dei sistemi) è descrivibile

    Le anomalie sono le differenze rispetto al comportamento atteso

  Il modello contiene la descrizione del comportamento atteso

    Il processo di ragionamento consiste nella spiegazione delle anomalie

- Case-based reasoning (CBR)

    Il mondo è regolare: problemi simili hanno spiegazioni simili

    Problemi simili tendono a ripresentarsi nel tempo

  Il modello contiene la descrizione di problemi e spiegazioni note

    Il processo di ragionamento consiste nell'adattamento per similitudine