Intelligenza Artificiale II

Logiche modali e temporali

Marco Piastra

Un paradosso?

Una particolare fbf di L_P :

$(\varphi \to \psi) \lor (\psi \to \varphi)$

Si tratta di una *tautologia*:

φ	ψ	$\varphi \to \psi$	$\psi \to \varphi$	$(\varphi \to \psi) \lor (\psi \to \varphi)$
0	0	1	1	1
0	1	0	1	1
1	0	1	0	1
1	1	1	1	1

- Lettura informale:

  Una relazione di implicazione sussiste <u>comunque</u> tra due fbf φ e ψ qualsiasi di L_P in un senso o nell'altro.

Implicazione stretta

Si direbbe che la relazione di *implicazione* sia troppo 'pervasiva'

Non si possono rappresentare coppie di proposizioni
che non hanno alcuna relazione logica:

φ : "Questi fagioli sono bianchi"

ψ : "Oggi c'è lezione di IA"

- L'origine storica della logica modale

Il desiderio di rappresentare una forma di *implicazione*
per cui questo 'paradosso' non vale

Originariamente detta ***implicazione stretta*** (Lewis, 1912)

Si affianca all'implicazione classica (anche *implicazione materiale*)

Come si vedrà, la definizione dell'implicazione stretta non è univoca
(già Lewis era arrivato a definire diverse logiche modali, non una sola)

Implicazione modale

$\Box\,(A \to B)$ (si legga $\Box$ come "L'agente ritiene che")

p. es. A : "Il treno non è riportato dall'orario ferroviario", B : "Il treno non c'è"

L'obiettivo è definire delle regole di ragionamento in cui si possa distinguere le verità soggettive ("l'agente ritiene che") dalle verità oggettive.

- Intuitivamente ci si aspetta che:

$\Box\,(A \to B), \Box\,A \vdash \Box\,B$

Se l'agente ritiene che $A \to B$ e A, allora ritiene che B

$\Box\,(A \to B), A \vdash \Box\,B$

Se A è (oggettivamente) vera, allora l'agente ritiene che B

$\Box\,(A \to B), A \nvdash B$

Se la regola è soggettiva, la verità (oggettiva) di A non implica la verità di B

- Inoltre:

$\vdash \Box\,A \lor \neg\Box\,A$

L'agente può ritenere A vera oppure no …

$\nvdash \Box\,A \lor \Box\,\neg A$

… ma non è obbligato ad avere un'opinione definita su A

$\nvdash \Box\,(A \to B) \lor \Box\,(B \to A)$

In particolare, l'agente non è obbligato a ritenere vera l'implicazione in uno dei due sensi

Linguaggio e derivazione

- L_{MP} : **Logica modale proposizionale**

  Linguaggio della logica proposizionale classica + il simbolo unario $\Box$

  si legga $\Box\,\varphi$ come "è *necessario* che φ" o anche "l'agente ritiene che φ"

  ed un simbolo unario *derivato*:

  $\Diamond \;\Leftrightarrow\; \neg\Box\neg$

  si legga $\Diamond\,\varphi$ come "è *possibile* che φ" o anche "l'agente ritiene possibile che φ"

- Assiomi proposizionali

  Valgono gli schemi di assioma Ax_1, Ax_2, Ax_3 di L_P (si estende la logica classica)

  Nell'istanziazione degli assiomi, non si può separare $\Box$ da quel che segue:
  "$\Box\,\varphi$" deve essere considerato come un'unica proposizione

- Regole di inferenza

  modus ponens $\quad \varphi \to \psi, \varphi \vdash \psi$

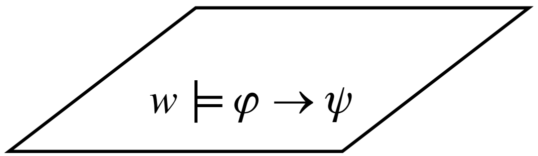
  necessitazione (Nec) $\quad \varphi \vdash \Box\,\varphi$

  Identico il concetto di *dimostrazione* (metodo *a la Hilbert*)

Semantica modale: mondi possibili ed accessibilità

- Intuitivamente, in logica classica
    - L'agente considera un solo mondo (un *mondo possibile*)
        - Le formule di una teoria descrivono i fatti noti (esempio: il mondo dei blocchi)
- Intuitivamente, in logica modale (nella declinazione *epistemica*)
    - L'agente considera una pluralità di ***mondi possibili***
        - Ciascuno di essi descrive una situazione logicamente coerente
        - Le formule di una teoria descrivono fatti noti e fatti ritenuti veri
  - L'insieme dei mondi possibili è strutturato
    - La relazione di **accessibilità** rappresenta i legame mondi ritenuti possibili

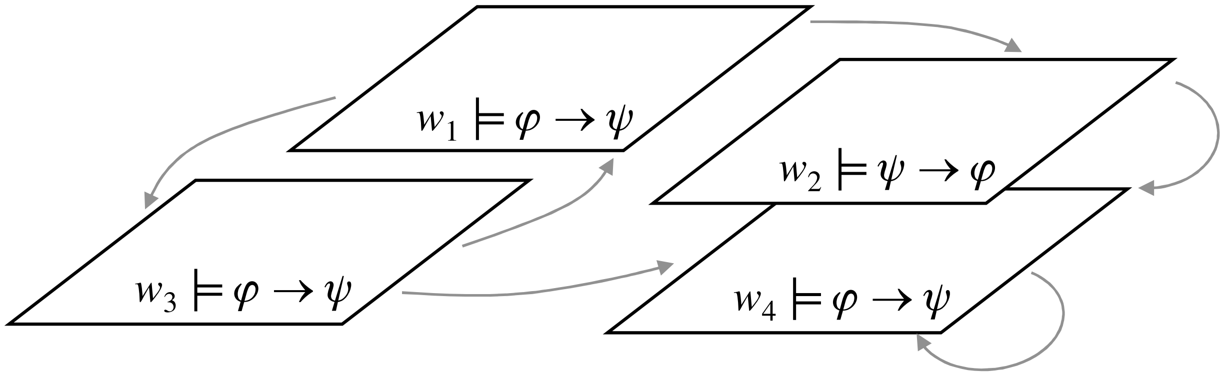
logica **classica**



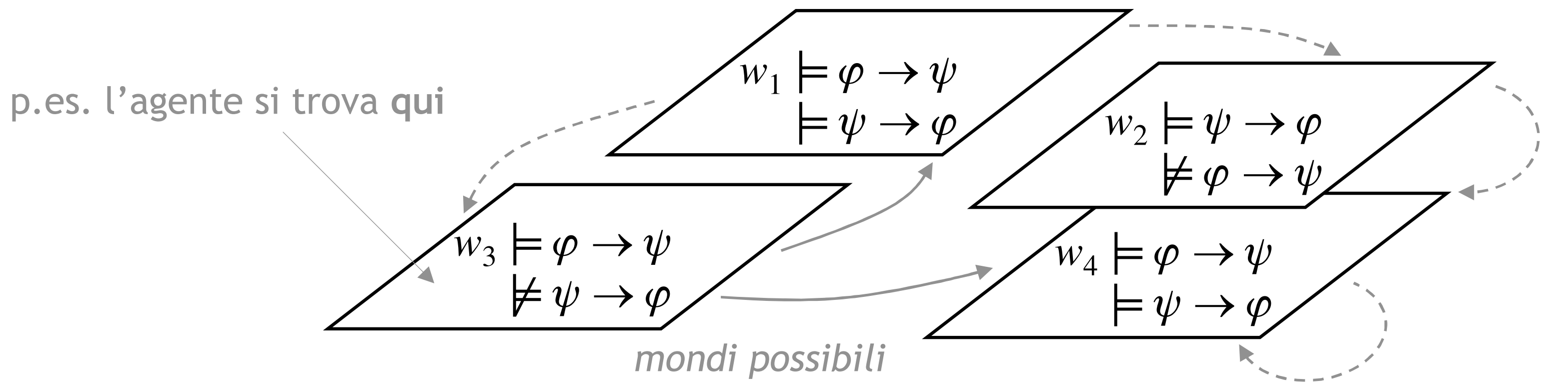
mondo possibile

logica **modale**

mondi possibili ed accessibilità



Interpretazione degli operatori modali



In ciascun mondo le formule φ non modali hanno un valore definito

- La semantica di $\Box$ è definita in riferimento a un singolo mondo

  $\Box\,\varphi$ è vera in un mondo w sse φ è vera in tutti mondi accessibili (o *possibili*) da w

  Esempio:

  Per un agente che si trova in w_3, $\Box\,(\varphi \to \psi)$ è vera, $\Box\,(\psi \to \varphi)$ è vera

  Per un agente che si trova in w_1, $\Box\,(\varphi \to \psi)$ è falsa, $\Box\,(\varphi \to \psi)$ è falsa, $\Diamond\,(\psi \to \varphi)$ è vera

Una formula modale può essere vera in tutti i mondi possibili

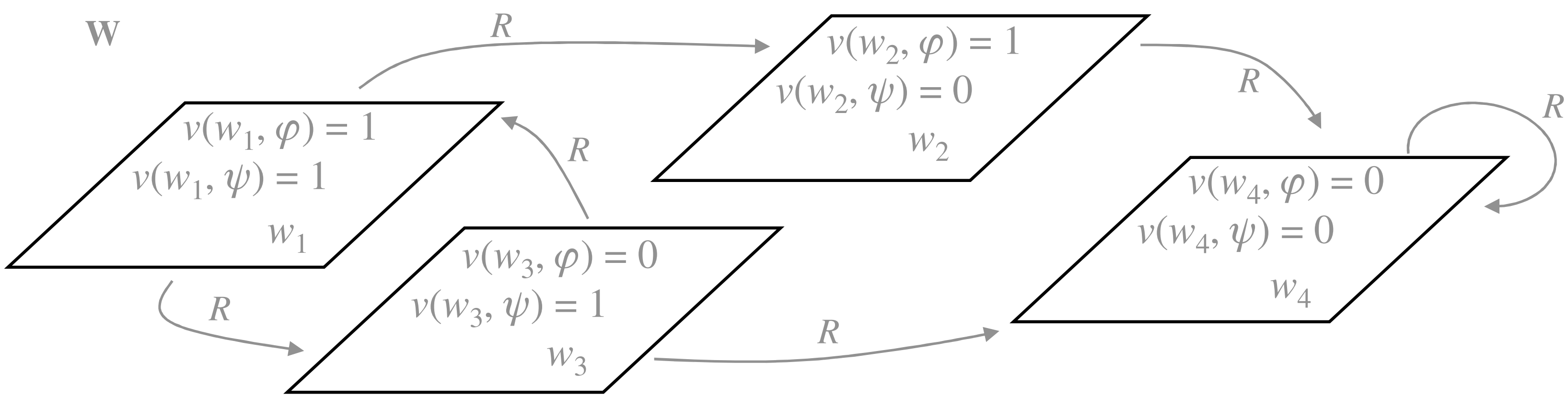
Esempio: $\Diamond\,(\psi \to \varphi)$ è vera in tutti mondi in figura

Strutture semantiche modali

- Modello come insieme strutturato di mondi possibili

  (dato un linguaggio proposizionale modale L_{MP})

  Si definisce una struttura (di Kripke) $<\mathbf{W}, R, v>$:
  - $\mathbf{W}$ è un insieme di punti detti anche 'stati' o 'mondi possibili'
  - R è una relazione binaria (in $\mathbf{W}^2$) che definisce l'accessibilità tra mondi
  - v è una funzione che assegna un valore di verità alle fbf di L_{MP} in ciascun mondo $w \in \mathbf{W}$



(*è la stessa struttura dell'esempio precedente*)

Regole semantiche

Definizione in tre passi

1) Soddisfacimento di formule non modali in un mondo w della struttura $<\mathbf{W}, R, v>$
2) Soddisfacimento di formule in un mondo w della struttura $<\mathbf{W}, R, v>$
3) Soddisfacimento nell'intera struttura

- Soddisfacimento

Una fbf *non* modale φ è soddisfatta in un mondo $w \in \mathbf{W}$ di una struttura $<\mathbf{W}, R, v>$ sse φ è vera in w

$<\mathbf{W}, R, v>, w \models \varphi$ sse $v(w, \varphi) = 1$ (secondo le regole di L_P)

fbf modali elementari, in un mondo w

$<\mathbf{W}, R, v>, w \models \Box\, \varphi$ sse $\forall w' \in \mathbf{W}, wRw'; <\mathbf{W}, R, v>, w' \models \varphi$

$<\mathbf{W}, R, v>, w \models \Diamond\, \varphi$ sse $\exists w' \in \mathbf{W}, wRw'; <\mathbf{W}, R, v>, w' \models \varphi$

Tramite le usuali regole di composizione semantica,
le definizioni si estendono alle fbf composite

Formula qualsiasi $\psi \in L_{MP}$, nell'intera struttura $<\mathbf{W}, R, v>$

$<\mathbf{W}, R, v> \models \psi$ sse $\forall w \in \mathbf{W}; <\mathbf{W}, R, v>, w \models \psi$

Logiche modali ed assiomi

Le logiche modali costituiscono una famiglia di logiche

- Logica modale **normale**

  K: $\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi)$

  (corrisponde alla possibilità di una semantica dei mondi possibili – *a la Kripke*)

- Alcuni assiomi modali

  D: $\Box\varphi \rightarrow \neg\Box\neg\varphi$ (equivale a $\Box\varphi \rightarrow \Diamond\varphi$)

  T: $\Box\varphi \rightarrow \varphi$

  4: $\Box\Box\varphi \rightarrow \Box\varphi$

  5: $\neg\Box\neg\varphi \rightarrow \Box\neg\Box\neg\varphi$ (equivale a $\Diamond\varphi \rightarrow \Box\Diamond\varphi$)

  B: $\varphi \rightarrow \Box\neg\Box\neg\varphi$ (equivale a $\varphi \rightarrow \Box\Diamond\varphi$)

- Assiomatizzazione di una (particolare) logica modale (normale)

  Gli assiomi Ax_1, Ax_2, Ax_3 di L_P

  Una combinazione di assiomi modali che include K (normalità)

  ad esempio K, D, 4 e 5

Attenzione: in logica modale NON vale il teorema di deduzione

Corrispondenze semantiche

I principali assiomi corrispondono a proprietà della relazione di accessibilità R tra i mondi possibili

In termini formali, le due proposizioni sono equivalenti:

- L'insieme Γ di fbf espresse in una determinata logica modale è soddisfacibile
- L'insieme Γ di fbf è soddisfacibile da una struttura modale in cui la R possiede le proprietà che corrispondono agli assiomi modali che caratterizzano tale logica

- Assiomi e proprietà corrispondenti di R

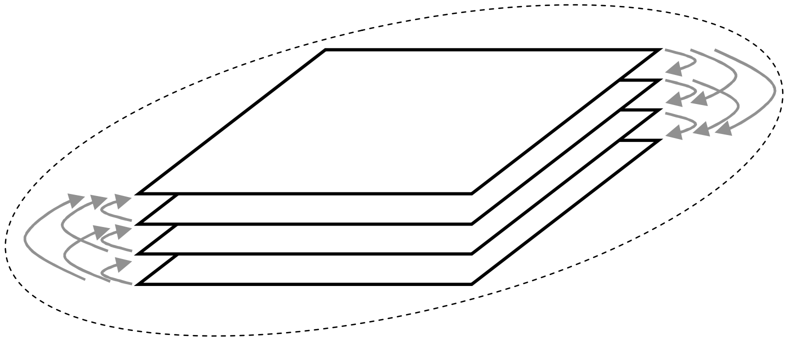
D: $\Box\varphi \to \Diamond\varphi$	corrisponde a	R *seriale*	$(\forall w\ \exists v,\ wRu)$
T: $\Box\varphi \to \varphi$	corrisponde a	R *riflessiva*	
5: $\Diamond\varphi \to \Box\Diamond\varphi$	corrisponde a	R *euclidea*	$(wRv,\ wRu \Rightarrow vRu)$
4: $\Box\Box\varphi \to \Box\varphi$	corrisponde a	R *transitiva*	
B: $\varphi \to \Box\Diamond\varphi$	corrisponde a	R *simmetrica*	

Quindi la logica KT45 (= KTB45 = KT5 = S5) corrisponde alla classe di strutture dove R è una relazione di *equivalenza*

Non tutte le proprietà della relazione R corrispondono ad un assioma modale (e.g. *irriflessività*)

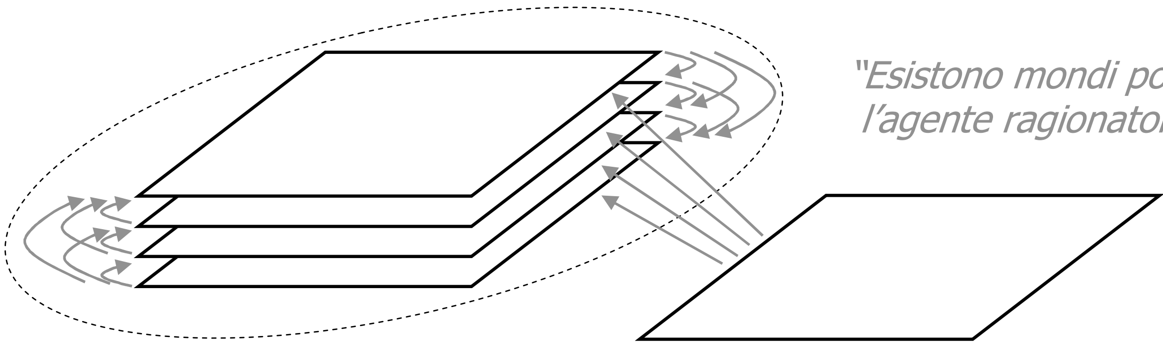
Strutture di mondi possibili

La logica KT45 (= S5) è soddisfacibile in una classe di strutture in cui tutti i mondi appartengono ad un'unica classe di equivalenza



"L'agente ragionatore ha accesso diretto a tutti i mondi possibili"

La logica KD45 è soddisfacibile invece in una classe di strutture in cui almeno un mondo rimane '*all'esterno*'



"Esistono mondi possibili a cui l'agente ragionatore non ha accesso"

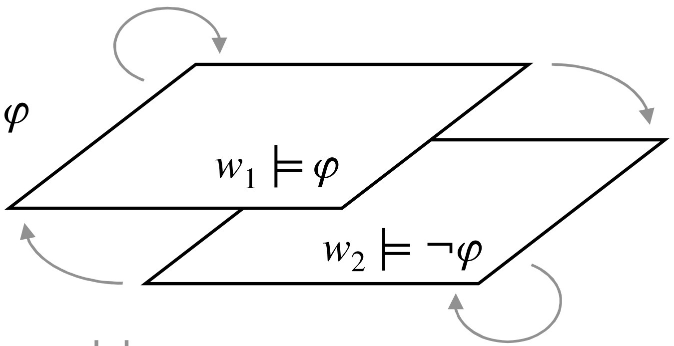
Verità e conoscenza

- L'operatore modale $\Box$ <u>non</u> è un quantificatore sui mondi possibili

  La semantica si basa infatti sulla relazione di accessibilità R

  Esempio: $\varphi \rightarrow \Box\, \varphi$ non è una fbf valida in S5

  $\not\models_{S5}\ \varphi \rightarrow \Box\, \varphi$



  La struttura è S5 ma $\varphi \rightarrow \Box\, \varphi$ non è vera in alcuno dei due mondi

    Si ricordi la regola Nec $\varphi \vdash \Box\, \varphi$:
    evidentemente il teorema di deduzione **non** vale in logica modale

  La validità di $\varphi \rightarrow \Box\, \varphi$ in S5 provocherebbe il 'collasso' della logica modale: il simbolo $\Box$ diventerebbe inutile e si avrebbe $S5 \equiv L_P$

  Che significa tutto questo?

  - In S5 il sistema delle conoscenze $\Box\, \varphi$ è certo (e progressivo) ma non si confonde con φ

    La fbf $\varphi \rightarrow \Box\, \varphi$ significherebbe "ciò che è certo è anche conosciuto"

  - La regola $\varphi \vdash \Box\, \varphi$ significa "ciò che è certo può essere derivato (o conosciuto)"

    In KD45 il sistema delle conoscenze $\Box\, \varphi$ <u>non è</u> certo (e nemmeno progressivo, o monotono)

La conoscenza dell'agente ragionatore

- Ciascuna logica modale caratterizza il sistema di conoscenze dell'agente

KT45 (= S5) è la logica della conoscenza *infallibile*

infatti vale T: $\Box \varphi \rightarrow \varphi$

KD45 è invece la logica della conoscenza *falsificabile*

infatti vale D: $\Box \varphi \rightarrow \neg\Box \neg\varphi$

L'assioma D esprime un requisito più debole, di non contraddittorietà

Gli assiomi 4 e 5 riguardano le capacità introspettive

4: $\Box\Box \varphi \rightarrow \Box \varphi$

5: $\Diamond \varphi \rightarrow \Box\Diamond \varphi$

L'agente "sa di sapere"?

Esempio: logica modale K45 e *Default Logic* (Konolige, 1988)

$(\Box \alpha \wedge \Diamond \beta_1 \wedge \Diamond \beta_2 \wedge \ldots \wedge \Diamond \beta_n) \rightarrow \Box \gamma$

(Trascrizione modale dello schema di inferenza per *default*)

La logica modale di riferimento è K45 : il ragionamento per *default* non garantisce infatti la 'consistenza' (assioma D)

(*La trascrizione modale K45 non è del tutto equivalente alla *Default Logic*)

Logiche modali e L_{PO}

- L'operatore modale □ è comunque un quantificatore

  L_{MP} corrisponde infatti ad un frammento (un sotto-insieme) di L_{PO}

  Regola di traduzione standard (*Standard Translation*)

$\mathrm{ST}_x(A) = A(x)$ (da proposizione a predicato unario, x variabile libera)

$\mathrm{ST}_x(\neg\varphi) = \neg\mathrm{ST}_x(\varphi)$

$\mathrm{ST}_x(\varphi \to \psi) = \mathrm{ST}_x(\varphi) \to \mathrm{ST}_x(\psi)$

$\mathrm{ST}_x(\Box\, \varphi) = \forall y\, (R(x,y) \to \mathrm{ST}_y(\varphi))$ (notare l'uso di y in ST_y , y vincolata)

$\mathrm{ST}_x(\Diamond\, \varphi) = \exists y\, (R(x,y) \wedge \mathrm{ST}_y(\varphi))$ (notare l'uso di y in ST_y , y vincolata)

Validità di $\mathrm{ST}_x(\varphi)$ in un mondo $w \in \mathbf{W}$: $<\mathbf{W}, v> (x{:}w) \models \mathrm{ST}_x(\varphi)$

Validità di $\mathrm{ST}_x(\varphi)$ in tutti i mondi di $\mathbf{W}$: $<\mathbf{W}, v> \models \forall x\, \mathrm{ST}_x(\varphi)$

Logiche modali

- In generale, le logiche modali
  - Sono caratterizzate dalla scelta di un particolare insieme di assiomi (e.g. KT5, KD45) a seconda del tipo di nozione informale si vuole rappresentare
  - Sono complete rispetto alla corrispondente classe di strutture
  - Sono *decidibili* (in versione proposizionale)
    Quindi anche il corrispondente frammento di L_{PO} lo è

- Inoltre
  - Non sono *vero-funzionali*:
    non esiste la possibilità di creare le tavole di verità con un numero finito di valori
  - Sono considerate *intensionali* e non *estensionali*
    in quanto il valore di verità dipende anche da un 'mondo possibile' o contesto

- Automazione

  Tipicamente, si usano metodi a tableau
  (con regole di inferenza diverse, a seconda della particolare logica modale)

2

Logiche temporali

Mondi e sequenze temporali

- Stessa struttura, diversa concettualizzazione
  - La struttura dei mondi possibili come una successione di istanti discreti
  - Ogni mondo rappresenta una descrizione completa ad un dato istante
  - La relazione di accessibilità descrive le possibili transizioni temporali
- Operatori temporali

A	(*always* A)	A è vero d'ora in poi
$\bigcirc A$	(*nexttime* A)	A sarà vero nell'istante successivo
$\Diamond A$	(*eventually* A)	A sarà vero, prima o poi

Esempi

$\Box (A \rightarrow B)$	D'ora in poi, se A allora B
$A \rightarrow \Box B$	Se A allora d'ora in poi B
$A \rightarrow \bigcirc B$	Se A allora B sarà vero nell'istante successivo
$\Diamond\Box A$	Prima o poi A sarà vero per sempre
$\Box\Diamond A$	In tutti gli istanti successivi, A sarà vero prima o poi

$$\Box((\neg passport \lor \neg ticket) \rightarrow (\bigcirc \neg board\_flight))$$

Linguaggio e assiomi

- LTL_P : logica temporale lineare (proposizionale)

  Linguaggio di L_P più i simboli unari $\Box$, $\bigcirc$, $\Diamond$

- Assiomi

  Gli schemi di assioma Ax_1, Ax_2, Ax_3 di $\mathrm{L_P}$

  Assiomi specifici:

  ltl_1: $\neg\bigcirc \varphi \leftrightarrow \bigcirc\neg \varphi$

  ltl_2: $\bigcirc (\varphi \to \psi) \to (\bigcirc \varphi \to \bigcirc \psi)$ (sussume l'assioma K)

  ltl_3: $\Box \varphi \to (\varphi \wedge \bigcirc\Box \varphi)$

- Regole di inferenza

  Modus Ponens $\varphi \to \psi, \varphi \vdash \psi$

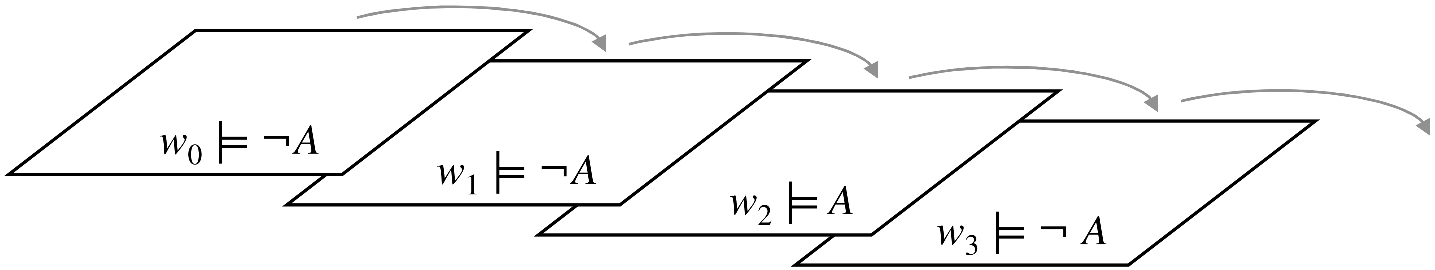
  Nex $\varphi \vdash \bigcirc \varphi$

  Ind $\varphi \to \psi, \varphi \to \bigcirc \varphi \vdash \varphi \to \Box \psi$

  Identico il concetto di *dimostrazione* (metodo '*a la Hilbert*', a partire dagli assiomi)

Strutture di mondi possibili, semantica

- La logica LTL_P è soddisfacibile in una classe di strutture in cui tutti i mondi appartengono ad un'unica sequenza infinita



- Regole semantiche:

  Data una struttura $<\mathrm{W}, w_0, v>$, dove $\mathbf{W}$ è una sequenza di mondi $\{w_i\}$, $i \in \mathrm{N}$

  $<\mathbf{W}, v>, w_i \models \Box\, \varphi$ sse $\forall\, j \geq i, <\mathbf{W}, v>, w_j \models \varphi$

  $<\mathbf{W}, v>, w_i \models \bigcirc\, \varphi$ sse $<\mathbf{W}, v>, w_{i+1} \models \varphi$

  $<\mathbf{W}, v>, w_i \models \Diamond\, \varphi$ sse $\exists\, j \geq i, <\mathbf{W}, v>, w_j \models \varphi$

Proprietà di LTL_P- Automazione

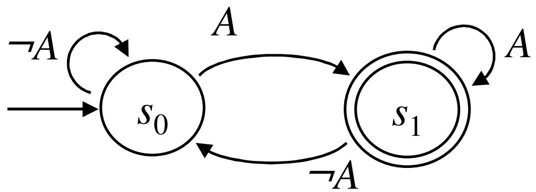
- LTL_P è completa

  $\models \varphi \;\Rightarrow\; \vdash \varphi$

  $\Gamma \models \varphi \;\Rightarrow\; \Gamma \vdash \varphi$ (se Γ è un insieme finito)

- La soddisfacibilità di una teoria (finita) in LTL_P è riconducibile ad un automa a stati finiti (automa di Büchi)

$\Box\Diamond A$



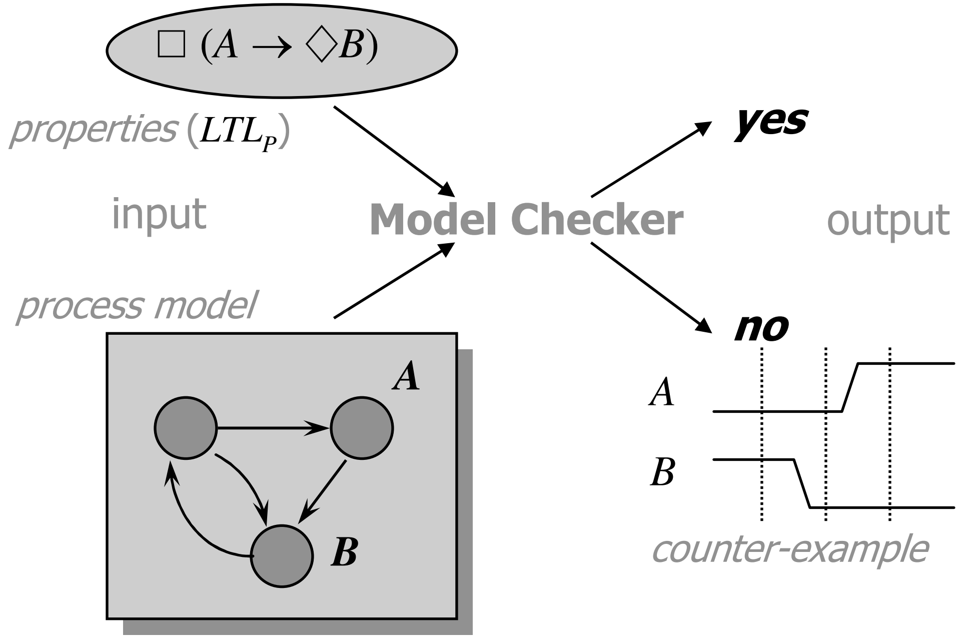
Nel diagramma, le etichette degli archi hanno solo un valore di esempio

Una fbf di LTL_P è sempre traducibile in un automa di Büchi

Un automa di Büchi riconosce sequenze *infinite*:
una sequenza $\{\omega_i\}, i \in \mathrm{N}$ è riconosciuta se produce una sequenza di stati $\{\rho_i\}$ in cui lo stato finale occorre infinite volte

Una fbf di LTL_P è **soddisfacibile** se l'insieme di sequenze riconosciute dal corrispondente automa di Büchi non è vuoto

Model Checking



- Verifica formale delle proprietà
  di un **modello di processo**
  (descritto da un automa a stati finiti)
- Sistemi effettivi (p.es. SPIN)
  - Traducono le fbf in LTL_P nel corrispondente automa di Büchi
  - Confrontano l'automa di Büchi con l'automa a stati finiti che descrive il processo
  - Producono una conferma o un contro-esempio (una sequenza che non soddisfa)

  Esempi di proprietà del modello di processo:
  - Sicurezza
    - $\Box\, \neg(Condition1 \wedge Condition2)$ (le due condizioni non si verificano mai simultaneamente)
  - Produttività (*Liveness*)
    - $\Box\, (Request \to \Diamond\, Service)$ (la richiesta sarà servita – prima o poi)
  - Buon funzionamento (*Fairness*)
    - $\Box\Diamond\, Request \to \Box\Diamond\, Service$ (il sistema continuerà a rispondere – assenza di *deadlock*)

Appendice: SPIN

- Per ulteriori approfondimenti

  http://spinroot.com/spin/whatispin.html