Intelligenza Artificiale II

Logiche modali e temporali

Marco Piastra

1

Logiche modali

Un paradosso?

- Una fbf di L_P
  $(\varphi \to \psi) \lor (\psi \to \varphi)$
- Si tratta di una *tautologia*

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

- La lettura informale è abbastanza inquietante:
  - comunque prese due fbf φ e ψ di L_P
  - una delle due è *conseguenza logica* dell'altra
    - Infatti:
      per il teorema di deduzione, $\varphi \to \psi$ equivale a $\varphi \vdash \psi$
      data la correttezza di L_P, $\varphi \vdash \psi$ equivale a $\varphi \models \psi$

Implicazione stretta

- Si direbbe che la relazione di *conseguenza logica*
  - è troppo 'pervasiva'
  - non si possono rappresentare coppie di proposizioni che non hanno alcuna relazione logica
    - "questi fagioli sono bianchi"
    - "oggi c'è lezione di IA"

- L'origine storica della logica modale (Lewis):
  - il desiderio di rappresentare una forma di implicazione per cui questo 'paradosso' non vale
    - originariamente detta *implicazione stretta*
    - che si affianca all'implicazione **classica** (anche *implicazione materiale*)
  - Già Lewis era arrivato a definire diverse logiche modali, non una sola

Implicazione modale

$\Box\,(A \rightarrow B)$ (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"

– Si intende rappresentare una regola di ragionamento che vale a prescindere delle situazioni particolari (p.es. conoscenza o meno di A)

- Intuitivamente, data la regola, ci si aspetta che:

  $\Box\,(A \rightarrow B),\ \Box\,A \vdash \Box\,B$
  Se l'agente ritiene che A, allora ritiene anche che B

  $\Box\,(A \rightarrow B),\ A \vdash \Box\,B$
  Se l'agente è certo che A, allora ritiene anche che B

  $\Box\,(A \rightarrow B),\ A \nvdash B$
  La regola non implica che la certezza di A implichi la certezza di B

- Inoltre:

  $\vdash \Box\,A \vee \neg\Box\,A$
  Si può assumere A oppure no …

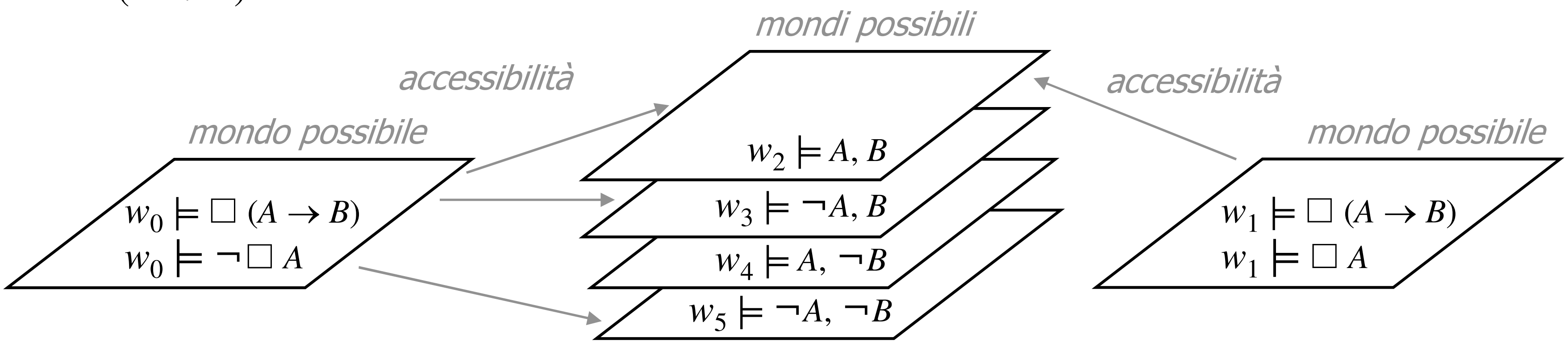
  $\nvdash \Box\,A \vee \Box\,\neg A$
  … ma non si è obbligati ad avere un'opinione

  $\nvdash \Box\,(A \rightarrow B) \vee \Box\,(B \rightarrow A)$
  Non si è obbligati a ritenere vera la regola o la regola opposta

Interpretazione intuitiva

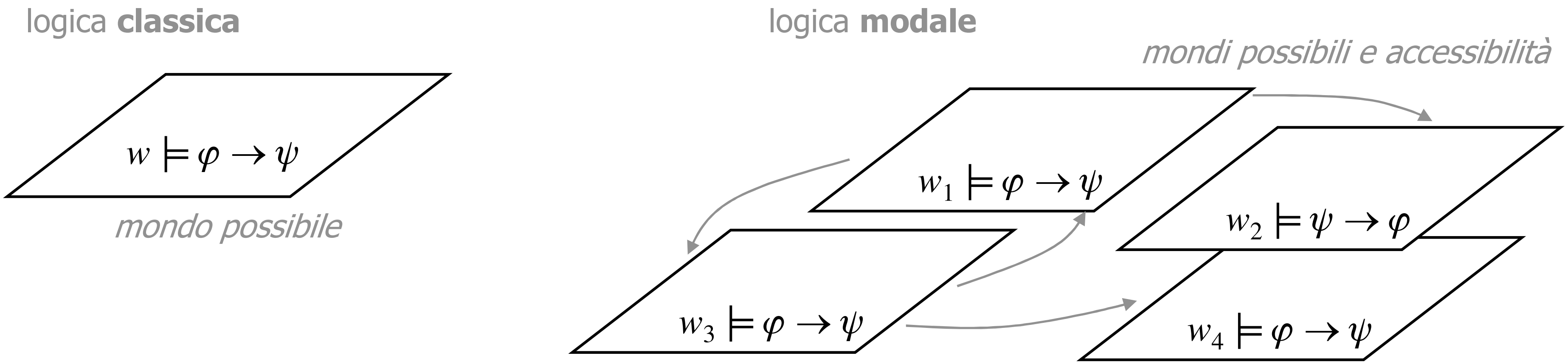
- Formule e mondi possibili
  - Premesso che:
    - L'agente adotta la regola $\Box (A \rightarrow B)$ (in generale, $\Box A$ o $\neg\Box A$)
  - Allora:
    - In alcuni casi (mondi), l'agente $\Box A$ (= viene a conoscenza di A): in tutti i mondi ritenuti possibili dall'agente, B è vero
    - In altri casi (mondi), l'agente $\neg\Box A$ (= non viene a conoscenza di A): in alcuni mondi ritenuti possibili dall'agente, B può essere vero e, in altri, falso

$\Box (A \rightarrow B)$



Non solo mondi possibili

- In logica classica
    - L'agente ha una vista univoca del mondo (un solo mondo possibile)
- In logica modale (nella declinazione *epistemica*)
    - L'agente considera una pluralità di **mondi possibili**
    - Ciascuno descrive una situazione logicamente coerente
  - L'insieme dei mondi possibili è strutturato
    - La relazione di **accessibilità** descrive i mondi
      che l'agente vede possibili a partire da ciascun mondo



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)
  - Nelle sostituzioni, però, 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)

Logiche modali ed assiomi

- Le logiche modali costituiscono una famiglia di logiche
- Logica modale **normale**

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

  (corrisponde alla possibilità di una semantica dei mondi possibili)

- Alcuni assiomi modali

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

  T: $\Box\varphi \to \varphi$

  4: $\Box\Box\varphi \to \Box\varphi$

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

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

- Assiomatizzazione di una (particolare) logica modale
  - 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

Assiomi e agente ragionatore

- Gli assiomi determinano il comportamento dell'agente ragionatore
  - Ad esempio KT45 (= S5) è la logica della conoscenza *infallibile*
    - infatti vale T: $\Box\, \varphi \to \varphi$
  - La logica KD45 è invece la logica della conoscenza *falsificabile*
    - infatti vale D: $\Box\, \varphi \to \neg\Box\, \neg\varphi$
      che esprime un requisito più debole, di semplice razionalità
  - Gli assiomi 4 e 5 riguardano le capacità introspettive

    4: $\Box\Box\, \varphi \to \Box\, \varphi$

    5: $\Diamond\, \varphi \to \Box\Diamond\, \varphi$
    - L'agente "sa di sapere"?
  - Logica modale K45 e *Default Logic* (Konolige, 1988)

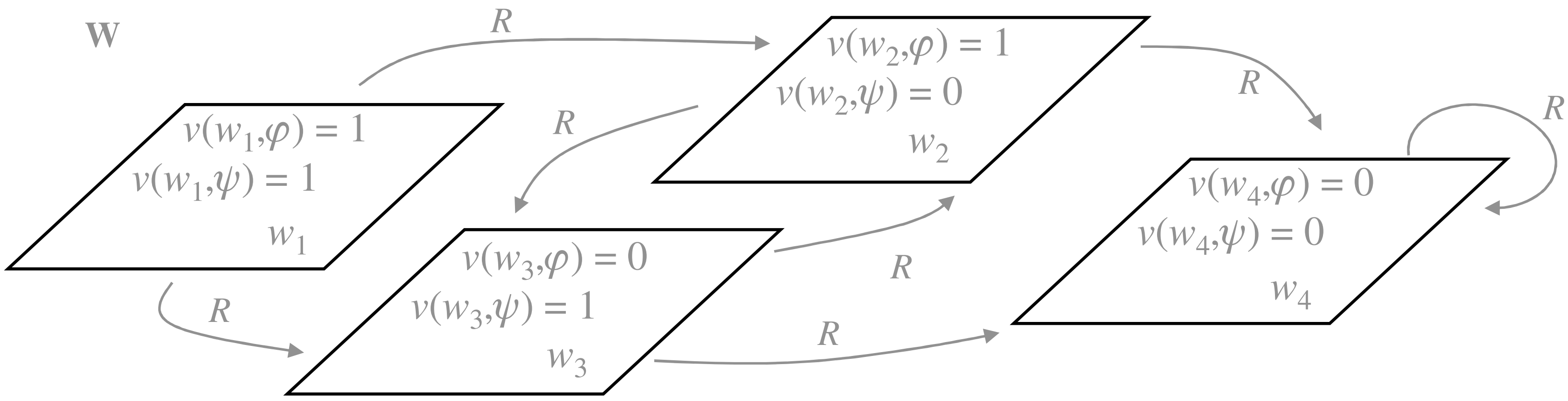
    $$(\Box\, \alpha \wedge \Diamond\, \beta_1 \wedge \Diamond\, \beta_2 \wedge \ldots \wedge \Diamond\, \beta_n) \to \Box\, \gamma$$

    (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*)

Semantica dei mondi possibili

- Strutture semantiche per la logica modale
  - (dato un linguaggio proposizionale modale L_{MP})
  - Strutture di mondi possibili $<\mathbf{W}, R, v>$ dove:
    - $\mathbf{W}$ è un insieme di punti detti anche 'stati' o 'mondi possibili'
    - R è una relazione binaria su $\mathbf{W}^2$ che definisce l'accessibilità tra mondi
    - v è una funzione che assegna in ciascun mondo $w \in \mathbf{W}$ un valore di verità in $\{0, 1\}$ alle fbf di L_{MP}



Regole semantiche

- Definizione in due passi
  1) Soddisfacimento in un mondo della struttura $<\mathbf{W}, R, v>$
  2) Soddisfacimento nell'intera struttura

- Soddisfacimento
  - una formula *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 $\mathrm{L_P}$)
  - formule 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 $\exists w' \in \mathbf{W}, wRw'; <\mathbf{W}, R, v>, w' \models \varphi$

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

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

Corrispondenze semantiche

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

- Ad esempio:

  D: $\Box\varphi \rightarrow \Diamond\varphi$ $\Leftrightarrow$ *seriale* $(\forall w\ \exists v,\ wRu)$

  T: $\Box\varphi \rightarrow \varphi$ $\Leftrightarrow$ *riflessiva*

  5: $\Diamond\varphi \rightarrow \Box\Diamond\varphi$ $\Leftrightarrow$ *euclidea* $(wRv,\ wRu \Rightarrow vRu)$

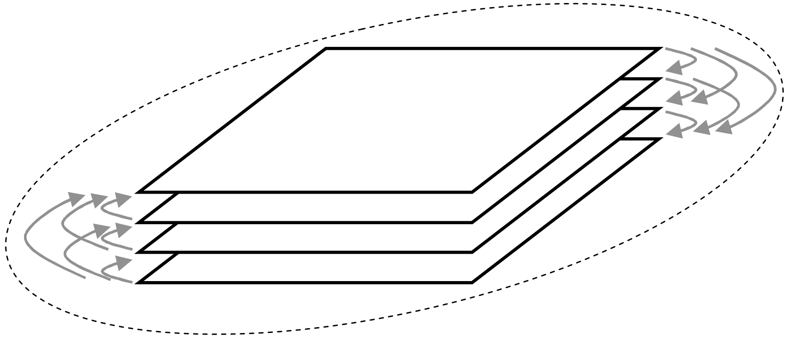
  4: $\Box\Box\varphi \rightarrow \Box\varphi$ $\Leftrightarrow$ *transitiva*

  B: $\varphi \rightarrow \Box\Diamond\varphi$ $\Leftrightarrow$ *simmetrica*

  - quindi la logica KT45 (= KTB45 = KT5 = S5) corrisponde alla classe di strutture dove R è una relazione di *equivalenza*
  - non tutte le proprietà di R corrispondono ad un assioma modale (e.g. *irriflessività*)

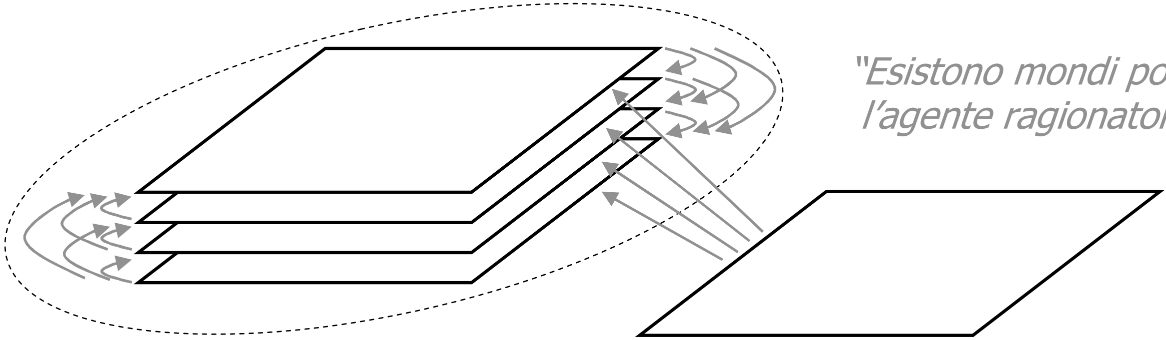
Strutture di mondi possibili

- 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"

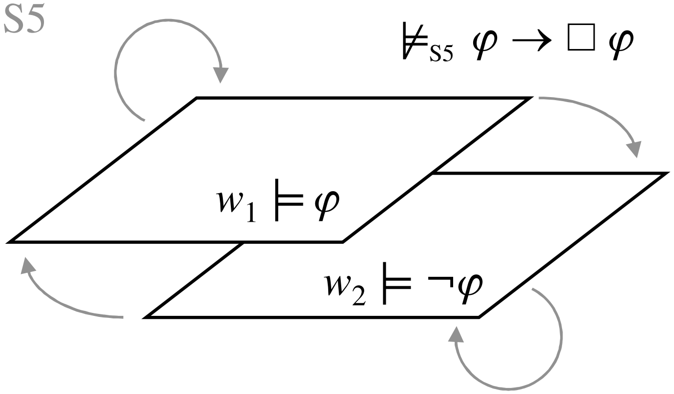
- La logica KD45 è soddisfacibile invece in una classe di strutture in cui almeno un mondo rimane 'all'esterno'



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

Specificità dell’operatore modale

- L’operatore modale $\Box$ <u>non</u> è un quantificatore sui mondi possibili
  - La semantica si basa infatti sulla relazione di accessibilità R
  - La verità delle fbf è definita in relazione a ciascun singolo mondo
    - esempio: $\varphi \rightarrow \Box\, \varphi$ non è una fbf valida in $\mathrm{S5}$



    - La struttura è $\mathrm{S5}$ 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 $\mathrm{S5}$ provocherebbe il ‘collasso’ della logica modale: il simbolo $\Box$ diventa inutile e $\mathrm{S5} \equiv \mathrm{L_P}$

Logiche modali e L_{PO}

- L'operatore modale $\Box$ è comunque un quantificatore
  - L_{MP} corrisponde infatti ad un frammento (limitato) di L_{PO}
  - Regola di traduzione standard (*Standard Translation*)
    - $ST_x(A) = A(x)$ (da proposizione a predicato unario, x variabile libera)
    - $ST_x(\neg\varphi) = \neg ST_x(\varphi)$
    - $ST_x(\varphi \rightarrow \psi) = ST_x(\varphi) \rightarrow ST_x(\psi)$
    - $ST_x(\Box\, \varphi) = \forall y\ (R(x,y) \rightarrow ST_y(\varphi))$ (notare l'uso di y in ST_y , y vincolata)
    - $ST_x(\Diamond\, \varphi) = \exists y\ (R(x,y)\ \wedge\ ST_y(\varphi))$ (notare l'uso di y in ST_y , y vincolata)

    - Validità di $ST_x(\varphi)$ in un mondo $w \in \mathbf{W}$: $<\mathbf{W}, v>\ (x{:}w) \models ST_x(\varphi)$
    - Validità di $ST_x(\varphi)$ in tutti i mondi di $\mathbf{W}$: $<\mathbf{W}, v> \models \forall x\ ST_x(\varphi)$

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 (o di struttura dei mondi possibili) si vuole rappresentare
  - sono complete rispetto alla corrispondente classe di strutture
  - sono *decidibili* (in versione proposizionale)
- Inoltre
  - non sono *vero-funzionali*, ovvero non esiste la possibilità di creare le tavole di verità con un numero finito di valori
  - sono talvolta intese come 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 può essere vista 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

$\square A$	(*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

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

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 L_P
  - Assiomi specifici:

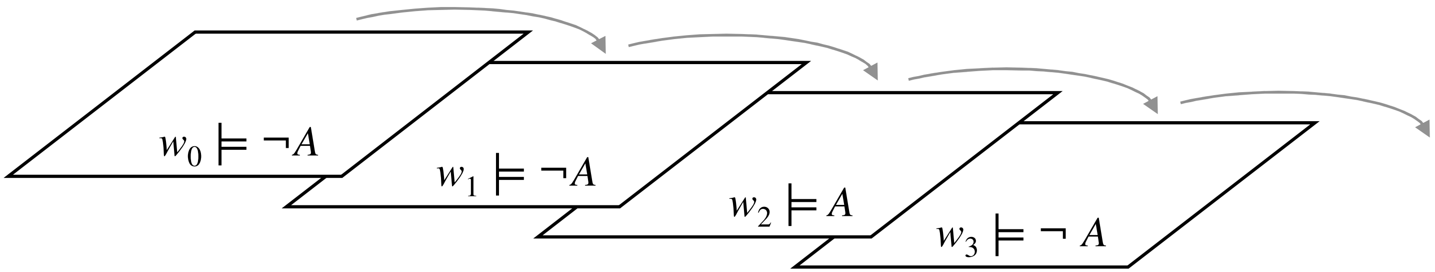
    ltl_1: $\neg \bigcirc \varphi \leftrightarrow \bigcirc \neg \varphi$

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

    ltl_3: $\Box \varphi \rightarrow (\varphi \wedge \bigcirc \Box \varphi)$
- Regole di inferenza
  - *modus ponens* $\varphi \rightarrow \psi, \varphi \vdash \psi$
  - *nex* $\varphi \vdash \bigcirc \varphi$
  - *ind* $\varphi \rightarrow \psi, \varphi \rightarrow \bigcirc \varphi \vdash \varphi \rightarrow \Box \varphi$
  - Identico il concetto di *dimostrazione* (metodo a la Hilbert)

Strutture di mondi possibili, semantica

- La logica $\mathrm{LTL_P}$ è soddisfacibile in una classe di strutture in cui tutti i mondi appartengono ad un'unica **sequenza** infinita



- Regole semantiche:
  - Data una struttura $<\mathrm{W}, v>$, 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 $\mathrm{LTL_P}$- Automazione

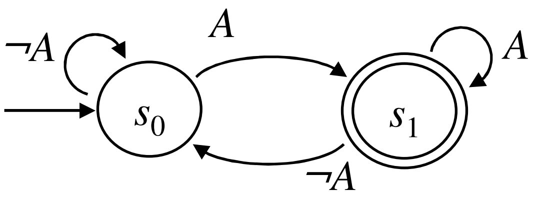
- $\mathrm{LTL_P}$ è completa

$$\models \varphi \quad \Rightarrow \quad \vdash \varphi$$

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

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

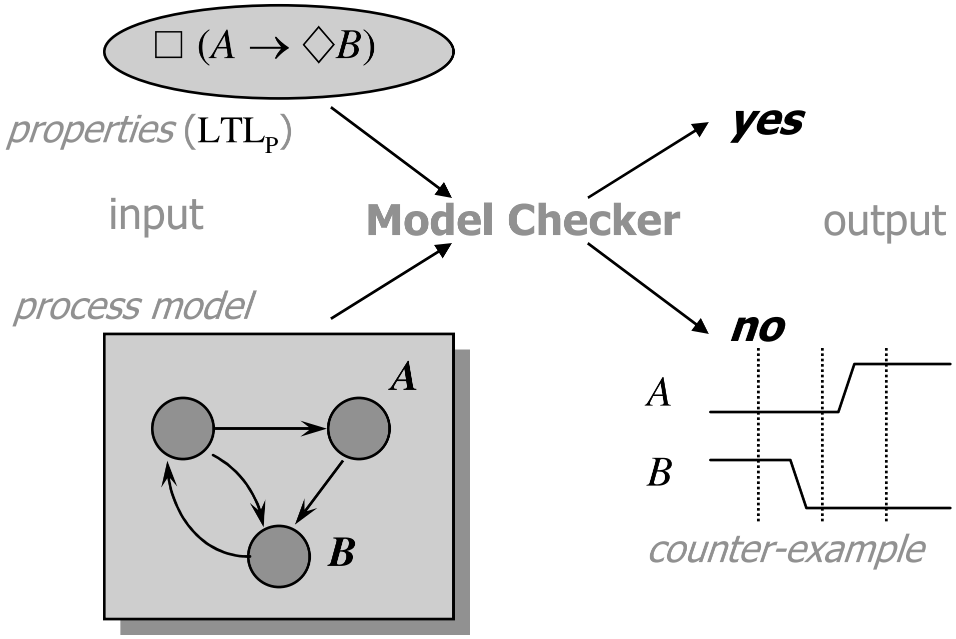
$\Box\Diamond A$



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

- Una fbf di $\mathrm{LTL_P}$ 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 $\mathrm{LTL_P}$ 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**
- 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*)
      $\Box (Request \rightarrow \Diamond Service)$ (la richiesta sarà servita – prima o poi)
    - Buon funzionamento (*Fairness*)
      $\Box\Diamond Request \rightarrow \Box\Diamond Service$ (il sistema continuerà a rispondere – assenza di *deadlock*)