

# Sistemas Multi-Agente

## Lógicas BDI

Dr. Alejandro Guerra-Hernández

**Instituto de Investigaciones en Inteligencia Artificial**  
Universidad Veracruzana

*Campus Sur, Calle Paseo Lote II, Sección Segunda No 112,  
Nuevo Xalapa, Xalapa, Ver., México 91097*

`mailto:aguerra@uv.mx`  
`https://www.uv.mx/personal/aguerra/sma`

Maestría en Inteligencia Artificial 2025



Universidad Veracruzana

# Organización

- 1 Introducción
- 2 Lógicas BDI temporalmente arborescentes



Universidad Veracruzana

# Ideas

- ▶ El estudio de agentes computacionales capaces de exhibir un comportamiento racional requiere de **formalizaciones teóricas**.
- ▶ Estas formalizaciones deben ser lo suficientemente **expresivas** para razonar acerca de:
  - ▶ Las representaciones de estos agentes en términos de **actitudes proposicionales**.
  - ▶ La ocurrencia de **eventos** y la ejecución de **acciones** en el ambiente.
  - ▶ La **comunicación** entre agentes.
- ▶ Idealmente, los formalismos usados deberían proveer una base **semántica** fuerte y una axiomatización **correcta** y **completa**.



# Aproximaciones formales a los agentes

- ▶ **Lenguajes externos a los agentes.** Metalenguajes de especificación de diseño y verificación de propiedades del comportamiento.
  - ▶ *BDI<sub>CTL</sub>* [13]
  - ▶ *Lora* [15]
  - ▶ Especificaciones en Z [3]
- ▶ **Lenguajes orientados a agentes.** Son lenguajes de especificación directamente ejecutables por el agente.
  - ▶ *Agent0* [14]
  - ▶ *Golog* [8]
  - ▶ *3APL* [7]
  - ▶ *AgentSpeak(L)* [12].



# Algunas preguntas

- ▶ ¿Cómo podemos razonar acerca del **comportamiento** de los agentes racionales?
- ▶ ¿Cómo podemos estudiar la relación entre los **estados intencionales** y dicho comportamiento?
- ▶ ¿Cómo sé que el comportamiento de mi agente es **racional**?
- ▶ ¿Qué **tipo** de formalismo podemos usar?
  - ▶ La lógica multi-modal propuesta por Cohen y Levesque [2].
  - ▶ Las lógicas *BDI* propuestas por Rao y Georgeff [11].
  - ▶ La lógica dinámica para SimpleAPL de Alechina et al. [1].
  - ▶ La lógica temporal para *AgentSpeak(L)* de Guerra-Hernández, Castro-Manzano y El-Fallah-Seghrouchni [6].



Universidad Veracruzana

# Dominios de aplicación

- ▶ ¿Cómo son los SMA que queremos implementar? Sistemas donde:
  1. El ambiente es no determinista.
  2. El sistema es no determinista.
  3. El sistema es multi-objetivo.
  4. Las acciones del sistema son dependientes del contexto e independientes del estado interno del sistema.
  5. El sistema es reactivo.
- ▶ Ej. Administración tráfico aéreo, diagnóstico de fallas en transbordador espacial, administración de redes de comunicación, asistencia en combate aéreo, minería de datos distribuida, etc.



Universidad Veracruzana

# Actuar

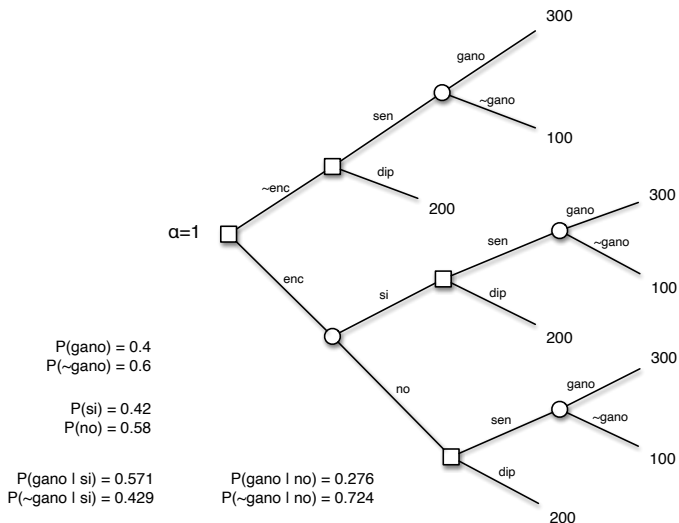
- ▶ El sistema necesita **actuar**.
- ▶ Necesita **seleccionar** las acciones o procedimientos apropiados de entre varias opciones disponibles.
- ▶ Esta selección debe permitir el cumplimiento efectivo de los **objetivos** del sistema.
- ▶ Tomando en cuenta los **recursos computacionales** disponibles y las características del **ambiente** donde el sistema está situado.



Universidad Veracruzana

# Tradicional: Teoría de decisión

□ Elección  
○ Posibilidad  
### Utilidad



Universidad Veracruzana



# Problemas

- ▶ El ambiente puede **cambiar** de manera significativa y/o no anticipada.
- ▶ Este cambio puede darse durante la función de **selección** o durante la **ejecución** de la acción seleccionada.
- ▶ Una función de selección **rápida** ayuda, pero su costo es quizás una decisión **subóptima**.



Universidad Veracruzana

# Soluciones a cambios durante ejecución

► Diferentes soluciones:

**Teoría de decisión** Reaplicar la función de selección si el ambiente cambia.

**Computación.** Nada cambia una vez iniciado el proceso de decisión.

**Intencionalidad.** Estrategias de **compromiso**.

► Las dos primeras no son aceptables dada la reactividad esperada del sistema (5).



Universidad Veracruzana

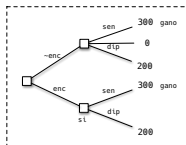
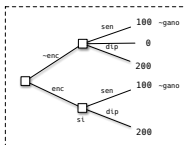
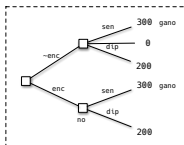
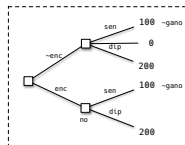
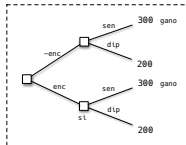
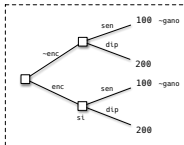
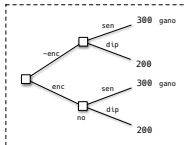
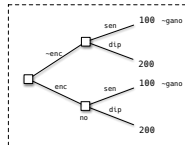
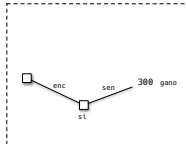
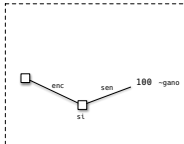
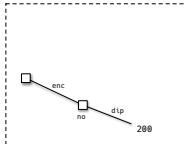
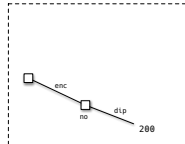
# BDI

- Creencias.** La dependencia del contexto (4) y el no determinismo del ambiente (1), exigen un componente **informativo** que toma la forma de conjuntos de expresiones lógicas, una base de datos, registros, etc.
- Deseos.** La naturaleza multi-objetivo (3) y la dependencia del contexto (4) exigen un componente de **motivación** que define las prioridades, rentabilidad o recompensa asociada con los diferentes objetivos del sistema.
- Intenciones.** La reactividad (5) y el no determinismo del ambiente (1) y del sistema (2), exigen un componente **deliberativo**, que represente el curso actual de acción elegido.



Universidad Veracruzana

# De los árboles a los mundos posibles BDI

 $w_1$   $\alpha = 0.24$ 

 $w_2$   $\alpha = 0.18$ 

 $w_3$   $\alpha = 0.16$ 

 $w_4$   $\alpha = 0.42$ 

 $w_5$ 

 $w_6$ 

 $w_7$ 

 $w_8$ 

 $w_9$ 

 $w_{10}$ 

 $w_{11}$ 

 $w_{12}$ 


Universidad Veracruzana

# Elementos del Formalismo BDI

- ▶ Intenciones como ciudadanos de primera clase: modelar diferentes tipos de **compromiso**.
- ▶ La **elección** del agente se diferencia del **resultado** de su elección, que depende del ambiente.
- ▶ Establecer las **relaciones** entre creencias, deseos e intenciones que eviten comprometerse con efectos colaterales no deseados.



# Las Lógicas multi-modales BDI

► Componentes para razonar acerca de:

**Proposiciones.** Una lógica proposicional.

**Actitudes proposicionales.** Operadores modales para razonar sobre las Creencias (B), Deseos (D) e Intenciones (I).

**Tiempo.** Una lógica temporal con futuro arborescente, pasado lineal, i.e., *CTL* y *CTL\**.

**Eventos.** Una lógica dinámica para razonar sobre la ocurrencia de **eventos**.

**Meta-expresiones.** Una lógica de primer orden para expresar propiedades de los componentes anteriores.



Universidad Veracruzana

# Lógicas de Árboles Computacionales

- ▶ Las lógicas que usaremos pueden verse como extensiones de las **lógicas de árboles computacionales**  $CTL$  y  $CTL^*$  de Emerson y Halpern [4, 5].
- ▶ La segunda es más expresiva, al permitir combinaciones arbitrarias de cuantificadores y operadores temporales.
- ▶ Estas lógicas se usan ampliamente para razonar acerca del comportamiento de **programas concurrentes**.



# Sintaxis $BDI_{CTL^*}$

- Las **fórmulas bien formadas** (fbf) de la lógica  $BDI_{CTL^*}$  son las definidas por la siguiente forma Backus Naur (BNF):

$$\phi ::= \perp \mid p \mid \neg\phi \mid \phi \wedge \phi \mid \quad (1)$$

$$BEL(\phi) \mid DES(\phi) \mid INT(\phi) \mid \quad (2)$$

$$\bigcirc\phi \mid \phi \text{ U } \phi \mid E\phi \quad (3)$$

donde  $p$  denota cualquier fórmula atómica proposicional.

- La sintaxis define fbfs **proposicionales** (1), **intencionales** (2) y **temporales** (3).



Universidad Veracruzana



# Otros operadores proposicionales

- ▶ La **disyunción**, la **implicación** y la **equivalencia material** se definen respectivamente como de costumbre:

- ▶  $\phi \vee \psi \stackrel{\text{def}}{=} \neg(\neg\phi \wedge \neg\psi).$

- ▶  $\phi \implies \psi \stackrel{\text{def}}{=} (\neg\phi \vee \psi).$

- ▶  $\phi \equiv \psi \stackrel{\text{def}}{=} ((\phi \implies \psi) \wedge (\psi \implies \phi)).$

- ▶ Los **valores de verdad** se definen en términos de la contradicción:

- ▶  $\text{false} \stackrel{\text{def}}{=} \perp$

- ▶  $\text{true} \stackrel{\text{def}}{=} \neg\perp$



# Otros operadores temporales

- ▶ Eventualmente, siempre, e inevitablemente se definen como de costumbre:

- ▶  $\Diamond\phi \stackrel{\text{def}}{=} \text{true} \cup \phi.$

- ▶  $\Box\phi \stackrel{\text{def}}{=} \neg(\Diamond\neg\phi).$

- ▶  $A\phi \stackrel{\text{def}}{=} \neg(E\neg\phi).$

- ▶ Ej. Con este lenguaje podemos expresar:

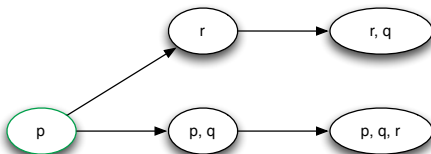
- ▶  $A\Diamond\text{BEL}(\neg\text{crisis})$  expresa que en todo futuro posible, eventualmente creeremos que no hay crisis.

- ▶  $A(\text{BEL}(\text{crisis}) \Rightarrow \neg\text{INT}(\text{viaje}))$  expresa que inevitablemente si creo que hay crisis entonces no intento hacer un viaje.



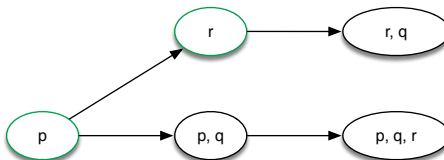
# Fórmulas de estado

- ▶ Las fbfs de estado en la lógica  $BDI_{CTL}^*$  son aquellas que son verdad en un mundo posible y un estado particular:
  1. Toda fbf proposicional (Eq. 1) es una fbf de estado.
  2. Si  $\phi$  es una fbf de estado, entonces  $BEL(\phi)$ ,  $DES(\phi)$  e  $INT(\phi)$  son fbf de estado.
  3. Si  $\phi$  es una fbf de camino, entonces  $E\phi$  (opcional) y  $A\phi$  (inevitable) son fbfs de estado.



# Fórmulas de camino

- ▶ Las fbfs de camino en la lógica  $BDI_{CTL^*}$  son aquellas que son verdaderas a lo largo de una secuencia de estados:
  1. Toda fbf de estado es una fbf de camino (de longitud 1).
  2. Si  $\phi$  y  $\psi$  son fbfs de camino, entonces  $\neg\phi$  y  $\phi \wedge \psi$  también lo son.
  3. Si  $\phi$  y  $\psi$  son fbfs de camino, entonces  $\bigcirc\phi$  y  $\phi \cup \psi$  también lo son.



# Sintaxis $BDI_{CTL}$

- ▶ La lógica restringida  $BDI_{CTL}$  se obtiene al prohibir combinaciones booleanas y **anidamiento** de los operadores temporales en las fórmulas de camino.
- ▶ Formalmente, sustituimos la definición de fbf de camino por:
  1. Si  $\phi$  y  $\psi$  son fbfs de estado, entonces  $\bigcirc\phi$  y  $\phi \cup \psi$  son fbfs de camino.
- ▶ Ej.  $\Box\Diamond crisis$ , que expresa que siempre eventualmente hay una crisis, es una fbf de camino en  $BDI_{CTL*}$ , pero no lo es en  $BDI_{CTL}$ .



# Fórmulas opcionales e inevitables

- El uso de **cuantificadores de camino** introduce otra clasificación sobre las fbf:

**O-fórmulas.** Las fórmulas opcionales son aquellas que no contienen ocurrencias de A (o de la negación de E) fuera del alcance de los operadores BEL, DES e INT.

**I-fórmulas.** Las fórmulas inevitables son aquellas que no contiene ocurrencias de E (o negaciones de A) fuera del alcance de los operadores intencionales BEL, DES, e INT.

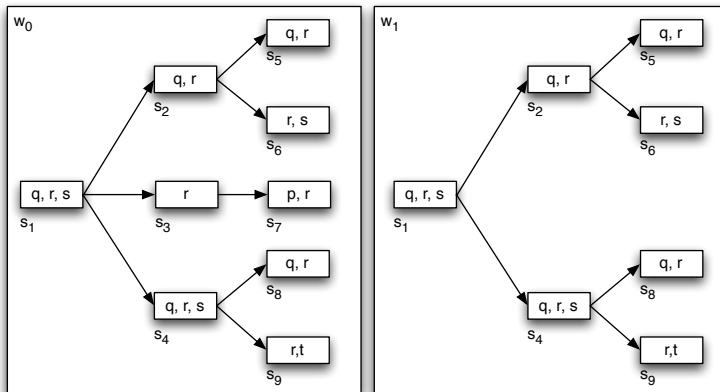


# Semántica de las lógicas $BDI_{CTL}$ y $BDI_{CTL}^*$

- ▶ Estas lógicas tienen una semántica basada en **mundos posibles**.
- ▶ Cada mundo posible se define como un **árbol** donde los nodos son conjuntos de fórmulas proposicionales, y los arcos representan transiciones en el tiempo.
- ▶ El **pasado es lineal**, pero el **futuro es arborescente**, para representar posibles cursos de acción. Una vez que el agente actúa, el pasado se vuelve lineal, debido a que solo uno de los cursos de acción posibles se ha llevado a cabo.



# Dos mundos posibles ( $w_1 \sqsubseteq w_0$ )





# Estructura de Kripke

- ▶ Se define como  $M = \langle W, \{S_w : w \in W\}, \{R_w : w \in W\}, L, \mathcal{B}, \mathcal{D}, \mathcal{I} \rangle$ , donde:
- ▶  $W$  es un conjunto de mundos posibles;
- ▶  $S_w$  es el conjunto de estados para cada mundo  $w \in W$ ;
- ▶  $R_w$  es una relación binaria total sobre los estados del mundo,  $R_w \subseteq S_w \times S_w$ ;
- ▶  $L$  es una función de asignación de verdad para las proposiciones primitivas en cada mundo  $w \in W$  en cada estado  $s \in S_w$ , p. ej.,  $L(w, s) : \phi \mapsto \{true, false\}$ ; y
- ▶  $\mathcal{B}, \mathcal{D}, \mathcal{I}$  son relaciones sobre los mundos y sus estados, p. ej.,  $\mathcal{B} \subseteq W \times S_w \times W$ .



# Semántica: fórmulas de estado

1. **Proposiciones atómicas:**  $M, w_s \models \phi$ , ssi  $\phi \in L(w, s)$
2. **Negación y Disyunción:**
  - 2.1  $M, w_s \models \neg\phi$ , ssi  $M, w_s \not\models \phi$ .
  - 2.2  $M, w_s \models \phi \vee \psi$ , ssi  $M, w_s \models \phi$  ó  $M, w_s \models \psi$ .
3. **Opcional e inevitable:**
  - 3.1  $M, w_{s_0} \models E\phi$ , ssi existe un camino  $w_{s_0}, w_{s_1}, \dots$  tal que  $M, (w_{s_0}, w_{s_1}, \dots) \models \phi$ .
  - 3.2  $M, w_{s_0} \models A\phi$ , ssi para todo camino  $w_{s_0}, w_{s_1}, \dots$ ,  $M, (w_{s_0}, w_{s_1}, \dots) \models \phi$ .
4. **Creencias, Deseos e Intenciones:**
  - 4.1  $M, w_s \models \text{BEL}(\phi)$ , ssi  $\forall v \mid (w, s, v) \in \mathcal{B}, M, v_s \models \phi$ .
  - 4.2  $M, w_s \models \text{DES}(\phi)$ , ssi  $\forall v \mid (w, s, v) \in \mathcal{D}, M, v_s \models \phi$ .
  - 4.3  $M, w_s \models \text{INT}(\phi)$ , ssi  $\forall v \mid (w, s, v) \in \mathcal{I}, M, v_s \models \phi$ .



# Semántica: fórmulas de camino

1. **Proposiciones atómicas:**  $M, (w_{s_0}, w_{s_1}, \dots) \models \phi$ , ssi  $M, w_{s_0} \models \phi$ .

2. **Negación y Disyunción:**

2.1  $M, (w_{s_0}, w_{s_1}, \dots) \models \neg\psi$ , ssi  $M, w_{s_0} \not\models \psi$ .

2.2  $M, (w_{s_0}, w_{s_1}, \dots) \models \phi \vee \psi$ , ssi  $M, (w_{s_0}, w_{s_1}, \dots) \models \phi$  ó  $M, (w_{s_0}, w_{s_1}, \dots) \models \psi$ .

3. **Next y Until:**

3.1  $M, (w_{s_0}, w_{s_1}, \dots) \models \bigcirc\psi$ , ssi  $M, (w_{s_1}, \dots) \models \psi$ .

3.2  $M, (w_{s_0}, w_{s_1}, \dots) \models \phi \cup \psi$ , ssi i)  $\exists k, k \geq 0$  t.q.  $M, (w_{s_k}, \dots) \models \psi$  y  $\forall 0 \leq j < k, M, (w_{s_j}, \dots) \models \phi$ ; ó ii)  $\forall j \geq 0, M, (w_{s_j}, \dots) \models \phi$ .



# Validez

- ▶ Una fbf se dice **válida** ssi es verdadera en todo estado, mundo y estructura.
- ▶ Una fbf se dice **satisfacible** ssi es verdadera en algún estado, mundo y estructura.
- ▶ La validez y satisfacción con respecto a una **familia de estructuras** también puede definirse:
  - ▶  $\mathcal{M}$  que requiere que  $\mathcal{R}$  sea **total**, sin imponer ninguna restricción sobre los operadores intencionales (básicamente es un sistema  $K$ ); y
  - ▶  $\mathcal{M}^{est}$ , que requiere que  $\mathcal{R}$  sea **total**;  $\mathcal{B}$  **serial, transitiva y euclidiana**; y,  $\mathcal{D}$  e  $\mathcal{I}$  sean **seriales**. Este modelo subyace en la lógica identificada como  $B^{KD45}D^{KD}I^{KD}_{CTL}$ .



# Relaciones

- ▶ Las definiciones de las relaciones, de acuerdo a Rao y Georgeff [11] son:

**Total:**  $\forall w \forall s \exists t (s, t) \in \mathcal{R}_w$ .

**Serial:**  $\forall w \forall s \exists v (w, s, v) \in \mathcal{B}$ .

**Transitiva:**  $\forall w, v, x \forall s$  Si  $(w, s, v) \in \mathcal{B} \wedge (v, s, x) \in \mathcal{B}$ , entonces  $(w, s, x) \in \mathcal{B}$ .

**Euclidiana:**  $\forall w, v, x \forall s$  Si  $(w, s, v) \in \mathcal{B} \wedge (w, s, x) \in \mathcal{B}$ , entonces  $(v, s, x) \in \mathcal{B}$ .

- ▶ **Nota.** Su relación **total** parece más una serial, ya que la definición estándar de total (o connex) sería:

$$\forall w \forall s \forall t (s, t) \in \mathcal{R}_w \vee (t, s) \in \mathcal{R}_w$$



Universidad Veracruzana

# El Axioma $K$

- ▶ En lo que sigue nos concentraremos en  $BDI_{CTL}$ .
- ▶ El axioma  $K$  establece que las creencias, deseos e intenciones son sujeto del *modus ponens*:

**BK**  $BEL(\phi) \wedge BEL(\phi \Rightarrow \psi) \Rightarrow BEL(\psi)$ ;

**DK**  $DES(\phi) \wedge DES(\phi \Rightarrow \psi) \Rightarrow DES(\psi)$ ;

**IK**  $INT(\phi) \wedge INT(\phi \Rightarrow \psi) \Rightarrow INT(\psi)$ ;

- ▶ Se adopta también la regla de **generalización**, i.e., toda fbf válida es creída, deseada e intentada:

**B-Gen** Si  $\vdash \phi$  entonces  $\vdash BEL(\phi)$ ;

**D-Gen** Si  $\vdash \phi$  entonces  $\vdash DES(\phi)$ ;

**I-Gen** Si  $\vdash \phi$  entonces  $\vdash INT(\phi)$ ;



- Se asume que las creencias son **consistentes** y hay **introspección** positiva y negativa sobre ellas:

$$BD \quad BEL(\phi) \Rightarrow \neg BEL(\neg \phi);$$

$$B4 \quad BEL(\phi) \Rightarrow BEL(BEL(\phi));$$

$$B5 \quad \neg BEL(\phi) \Rightarrow BEL(\neg BEL(\phi)).$$



$B^{KD45} D^{KD} I^{KD}$ 

- ▶ Para los deseos y las intenciones se adopta además el axioma  $D$  para expresar **consistencia** entre los deseos y las intenciones.

$$DD \text{ DES}(\phi) \Rightarrow \neg \text{DES}(\neg \phi);$$

$$ID \text{ INT}(\phi) \Rightarrow \neg \text{INT}(\neg \phi);$$

- ▶ Es consistente y completa con respecto a la familia de estructuras  $\mathcal{M}^{est}$





# Relaciones entre conjuntos de mundos

- Hay cuatro casos posibles, ejemplificados aquí con mundos creídos ( $\mathcal{B}$ ) y deseados ( $\mathcal{D}$ ):
  - $\mathcal{D} \subseteq \mathcal{B}$  El agente cree un mundo posible, pero no desea estar en él;
  - $\mathcal{B} \subseteq \mathcal{D}$  Hay ciertos mundos deseados que el agente no cree posibles;
  - $\mathcal{B} \cap \mathcal{D} \neq \emptyset$  Algunos mundos que son deseados no son creídos y viceversa, una combinación de los dos primeros casos.
  - $\mathcal{B} \cap \mathcal{D} = \emptyset$  Esta posibilidad no es significativa.



# Relaciones estructurales entre mundos

- ▶ Dados los árboles de dos mundos posibles, ejemplificados aquí por un mundo creído ( $w_b$ ) y mundo deseado ( $w_d$ ), también hay cuatro casos posibles:
  - $w_d \sqsubseteq w_b$  De todos los caminos que el agente cree poder elegir, solo algunos son deseables.
  - $w_b \sqsubseteq w_d$  De todos los caminos que el agente desea, solo algunos se creen realizables.
  - $w_b = w_d$  La combinación de los casos anteriores, todos los caminos creídos son deseados y viceversa.
  - $w_b \not\sqsubseteq w_d$  Este caso no es significativo, los mundos creídos y deseados son disjuntos.



# Combinaciones

- ▶ Combinando los casos significativos, es posible obtener nueve relaciones basadas en conjuntos de mundos posibles y sus estructuras.
- ▶ Tres de estas relaciones han sido consideradas en la literatura: **realismo** [2], **realismo fuerte** [10], y **realismo débil** [9].



# Realismo fuerte

- ▶ Se da cuando  $\mathcal{B} \subseteq \mathcal{D} \subseteq \mathcal{I}$  y cada  $w_i \sqsubseteq w_d \sqsubseteq w_b$ .
- ▶ Si el agente desea opcionalmente lograr una proposición, entonces cree que la proposición es una opción que es realizable (si la elige).
- ▶ La relación puede aplicar a mundos deseados e intentados.
- ▶ Si  $\phi$  es una fórmula-O, las condiciones anteriores se expresan con los axiomas de realismo fuerte:

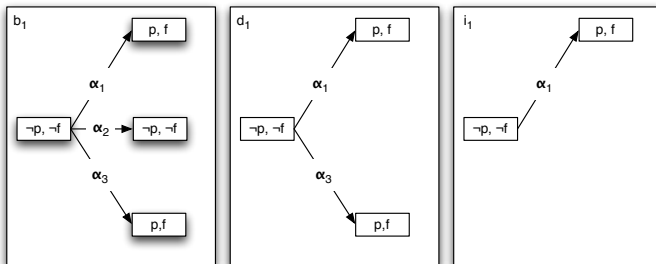
Realismo fuerte ID.  $INT(\phi) \implies DES(\phi)$ ;

Realismo fuerte DB.  $DES(\phi) \implies BEL(\phi)$ .



Universidad Veracruzana

# Ejemplo



# Realismo

- ▶ Se da cuando  $\mathcal{I} \subseteq \mathcal{D} \subseteq \mathcal{B}$  y cada  $w_b = w_d = w_i$ .
- ▶ Si un agente cree una proposición, también tiene una intención hacia esa proposición.
- ▶ Si  $\phi$  es una fórmula-O:

Realismo DI.  $DES(\phi) \implies INT(\phi)$ ;

Realismo BD.  $BEL(\phi) \implies DES(\phi)$ .



Universidad Veracruzana

# Realismo débil

- ▶ Se da cuando  $\mathcal{B} \cap \mathcal{D} \neq \emptyset$  y cada  $w_d \sqsubseteq w_b$ . Lo mismo para deseos-intenciones y creencias-intenciones.
- ▶ Un punto de equilibrio donde los agentes no desean proposiciones cuya negación es creída; no intentan proposiciones cuya negación es deseada; y no intentan proposiciones cuya negación es deseada.

Realismo débil DB.  $DES(\phi) \implies \neg BEL(\neg\phi)$ ;

Realismo débil IB.  $INT(\phi) \implies \neg BEL(\neg\phi)$ ;

Realismo débil ID.  $INT(\phi) \implies \neg DES(\neg\phi)$ .



Universidad Veracruzana

# Consistencia BI, BD y DI

- ▶ Si un agente tiene una intención, éste cree que tiene tal intención:  
 $INT(\phi) \implies BEL(INT(\phi))$
- ▶ Si un agente tiene un deseo, éste cree que tiene tal deseo:  
 $DES(\phi) \implies BEL(DES(\phi))$
- ▶ Si un agente tiene una intención, debe desear tal intención:  
 $INT(\phi) \implies DES(INT(\phi))$



Universidad Veracruzana



# Tesis de asimetría

$$AT1 \models \text{INT}(\phi) \implies \neg \text{BEL}(\neg \phi)$$

$$AT2 \not\models \text{INT}(\phi) \implies \text{BEL}(\phi)$$

$$AT3 \not\models \text{BEL}(\phi) \implies \text{INT}(\phi)$$

$$AT4 \models \text{INT}(\phi) \implies \neg \text{DES}(\neg \phi)$$

$$AT5 \not\models \text{INT}(\phi) \implies \text{DES}(\phi)$$

$$AT6 \not\models \text{DES}(\phi) \implies \text{INT}(\phi)$$

$$AT7 \models \text{DES}(\phi) \implies \neg \text{BEL}(\neg \phi)$$

$$AT8 \not\models \text{DES}(\phi) \implies \text{BEL}(\phi)$$

$$AT9 \not\models \text{BEL}(\phi) \implies \text{DES}(\phi)$$



# Asimetría en sistemas $B^{KD45}D^{KD}I^{KD}$

	AT1	AT2	AT3	AT4	AT5	AT6	AT7	AT8	AT9
Realismo fuerte	✓		✓	✓		✓	✓		✓
Realismo	✓	✓		✓	✓		✓	✓	
Realismo débil	✓	✓	✓	✓	✓	✓	✓	✓	✓



# Posposición finita

- ▶ Si un agente forma una intención, entonces en algún momento futuro la **abandona**.
- ▶ Esto se conoce como posposición finita o **compromiso finito** (*no infinite deferral*):

$$INT(\phi) \Rightarrow A\Diamond(\neg INT(\phi))$$



# Sintáxis

- ▶ Sea  $E$  un conjunto de símbolos que denotan **eventos**;
- ▶ Si  $e \in E$ , entonces  $succeeds(e)$ ,  $fails(e)$ ,  $does(e)$ ,  $succeeded(e)$ ,  $failed(e)$ , y  $done(e)$  son fbfs de estado.



# Semántica

- ▶ Incluimos en  $M$  los siguientes elementos: Sea  $E$  un conjunto de **tipos de evento** primitivos;  $SE_w : S_w \times S_w \mapsto E$  y  $FE_w : S_w \times S_w \mapsto E$  las ocurrencias con fracaso y éxito de los eventos.
  - ▶  $M, w_{s_1} \models \text{succeeded}(e)$  si y sólo si  $SE_w(s_0, s_1) = e$
  - ▶  $M, w_{s_1} \models \text{failed}(e)$  si y sólo si  $FE_w(s_0, s_1) = e$



# Axiomatización de los eventos

- ▶ Las intenciones tienen un carácter **volitivo**:

$$INT(does(e)) \implies does(e)$$

- ▶ Un agente es **consciente** de los eventos que ocurren en su medio ambiente.

$$done(e) \implies BEL(done(e))$$



# Compromiso ciego

- ▶ Este agente mantiene sus intenciones hasta que cree que las ha logrado **satisfacer**:
  - ▶  $\text{INT}(A \Diamond \phi) \implies A(\text{INT}(A \Diamond \phi) \cup \text{BEL}(\phi))$
  - ▶  $\text{INT}(A \Diamond \phi) \implies A \Diamond (\text{BEL}(\phi))$



# Compromiso racional

- ▶ El agente mantiene sus intenciones en tanto considera que siguen siendo una **opción viable**:
  - ▶  $\text{INT}(A \Diamond \phi) \implies A(\text{INT}(A \Diamond \phi) \cup (\text{BEL}(\phi) \vee \text{BEL}(E \Diamond \phi)))$
  - ▶  $\text{INT}(A \Diamond \phi) \wedge A(\text{BEL}(E \Diamond \phi) \cup \text{BEL}(\phi)) \implies A \Diamond (\text{BEL}(\phi)).$





# Compromiso emocional

- ▶ Este agente mantiene sus intenciones mientras siga **deseándolas**:
- ▶  $\text{INT}(A \Diamond \phi) \implies A(\text{INT}(A \Diamond \phi) \cup (\text{BEL}(\phi) \vee \neg \text{DES}(E \Diamond \phi)))$
- ▶  $\text{INT}(A \Diamond \phi) \wedge A(\text{DES}(E \Diamond \phi)) \cup \text{BEL}(\phi) \implies A \Diamond (\text{BEL}(\phi)).$



# Otros resultados

- ▶ Un agente **competente** sigue  $BEL(\phi) \implies \phi$ , e independientemente de su compromiso adoptado, llega  $BEL(A\Diamond\phi)$



# Referencias I

- [1] N Alechina et al. "A Logic of Agent Programs". En: *Proceedings of the Twenty-Second AAAI Conference on Artificial Intelligence, July 22-26, 2007, Vancouver, British Columbia, Canada*. New York, NY, USA: AAAI Press, 2007, págs. 795-800.
- [2] PR Cohen y HJ Levesque. "Intention is choice with commitment". En: *Artificial Intelligence* 42.3 (1990), págs. 213-261.
- [3] M d'Inverno y M Luck. *Understanding Agent Systems*. Second. Berlin, Germany New York: Springer, 2004.
- [4] A Emerson y J Halpern. "Decision Procedures and Expressiveness in The Temporal Logic of Branching Time". En: *Journal of Computer and System Sciences* 30.1 (1985), págs. 1-24.
- [5] A Emerson y J Halpern. ""Sometime" and "Not Never" Revisited: On Branching versus Linear Time Temporal Logic". En: *Journal of the Association for Computer Machinery* 33.1 (1986), págs. 151-178.
- [6] A Guerra-Hernández, JM Castro-Manzano y A El-Fallah-Seghrouchni. "CTL AgentSpeak(L): a Specification Language for Agent Programs". En: *Journal of Algorithms* 64 (2009), págs. 31-40.



Universidad Veracruzana

# Referencias II

- [7] KV Hindriks et al. "Agent Programming in 3APL". En: *Autonomous Agents and Multi-Agent Systems 2.4* (1999), págs. 357-401. ISSN: 1387-2532. URL: [http://portal.acm.org/ft\\_gateway.cfm?id=608641&type=external&coll=Portal&dl=GUIDE&CFID=90115497&CFTOKEN=29439121](http://portal.acm.org/ft_gateway.cfm?id=608641&type=external&coll=Portal&dl=GUIDE&CFID=90115497&CFTOKEN=29439121).
- [8] HJ Levesque et al. "GOLOG: A logic programming language for dynamic domains". En: *Journal of Logic Programming* 31.1997 (1997), págs. 59-84.
- [9] AS Rao y MP Georgeff. *Asymmetry Thesis and Side-Effect Problems in Linear-Time and Branching-Time Intention Logics*. Inf. téc. 13. published in proceedings of IJCAI-91. Carlton, Victoria: Australian Artificial Intelligence Institute, 1991.
- [10] AS Rao y MP Georgeff. *Modelling Rational Agents within a BDI-Architecture*. Inf. téc. 14. Carlton, Victoria: Australian Artificial Intelligence Institute, 1991.
- [11] AS Rao y MP Georgeff. "Decision Procedures for BDI Logics". En: *Journal of Logic and Computation* 8.3 (1998), págs. 293-342.
- [12] A Rao. "AgentSpeak(L): BDI Agents Speak Out in a Logical Computable Language". En: *Seventh European Workshop on Modelling Autonomous Agents in a Multi-Agent World*. Ed. por R van Hoe. Eindhoven, The Netherlands, 1996.



# Referencias III

- [13] A Rao y M Georgeff. *Formal Models and Decision Procedures for Multi-Agent Systems*. Technical Report 61. Carlton, Victoria: Australian Artificial Intelligence Institute, 1995.
- [14] Y Shoham. "AGENT0: A Simple Agent Language and Its Interpreter". En: *Proceedings of the 9th National Conference on Artificial Intelligence*. Ed. por TL Dean y K McKeown. Vol. 2. Menlo Park, CA, USA: AAAI Press, 1991, págs. 704-709.
- [15] M Wooldridge. *Reasoning about Rational Agents*. Cambridge, MA, USA: The MIT Press, 2000.



Universidad Veracruzana