Joachim Nitschke

Epistemische Logik

Nach dem Artikel:
Halpern, J.Y. (1995). *Reasoning about knowledge: A survey*. In: Gabbay, D., Hogger, C.J. und Robinson, J.A. (Hrsg.), Handbook of Logic in Artificial Intelligence and Logic Programing, Band 4, S. 1-34. Oxford University Press

Seminar: FGI 3 - Logik und Semantik von Programmen
Wintersemester 2009/10

natürlichsprachliches Beispiel

- Modellierung von Sachverhalten wie:

  „Dean weiß nicht, ob Nixon weiß, dass Dean weiß, dass Nixon weiß, dass McCord in O'Briens Büro in Watergate eingebrochen ist.“

Syntax

- Eine Menge von primitiven Aussagen:

  $\Phi = \{p, q, r, \dots\}$

- Modaloperatoren:

  $K_i \varphi \quad (i = 1, \dots, n)$

  „Agent i weiß φ."

- Operatoren aus der Aussagenlogik:

  $\psi \wedge \varphi, \neg \varphi$

Modellierung der Watergate-Affäre

„Dean weiß nicht, ob Nixon weiß, dass Dean weiß, dass Nixon weiß, dass McCord in O'Briens Büro in Watergate eingebrochen ist."

- Agenten: D (John Dean), N (Richard Nixon)
- Aussage p: „McCord ist in O'Briens Büro in Watergate eingebrochen."

Modellierung der Watergate-Affäre

„Dean weiß nicht, ob Nixon weiß, dass Dean weiß, dass Nixon weiß, dass McCord in O'Briens Büro in Watergate eingebrochen ist."

- Agenten: D (John Dean), N (Richard Nixon)
- Aussage p: „McCord ist in O'Briens Büro in Watergate eingebrochen."

$$\neg K_D \neg (K_N K_D K_N p) \wedge \neg K_D \neg (\neg K_N K_D K_N p)$$

Syntax (Forts.)

- Operator für „Gruppenwissen“:

  $E_G \varphi$

  „Jeder in Gruppe G weiß φ.“

- Operator für „Allgemeinwissen“:

  $C_G \varphi$

  „Jeder in Gruppe G weiß φ und jeder (in Gruppe G) weiß, dass jeder φ weiß und jeder weiß, dass jeder weiß, dass jeder φ weiß und usw.“

Semantik

- Kripkestruktur $M(S, \pi, K_1, \ldots, K_n)$

  S

  (Menge der „möglichen Welten“ oder „Zustände“)

  $\pi(s)(p) \in \{true, false\}$

  $(\text{für jedes } s \in S \text{ und jede Aussage } p)$

  (Interpretation oder „Wahrheitswertzuweisung“)

  K_i

  ((Äquivalenz-)Relation über S „Möglichkeits-“ oder „Zugänglichkeitsrelation“ von Agent i)

Semantik (Forts.)

- Relation $\models$:

$(M,s) \models p \;\; wenn \, \pi(s)(p) = true$

$(M,s) \models \neg\varphi \;\; wenn \; (M,s) \not\models \varphi$

$(M,s) \models \varphi \wedge \psi \;\; wenn \; (M,s) \models \varphi \; und \; (M,s) \models \psi$

$(M,s) \models K_i\varphi \;\; wenn \; (M,t) \models \varphi \;\; für\, alle\, t\, mit\, (s,t) \in K_i$

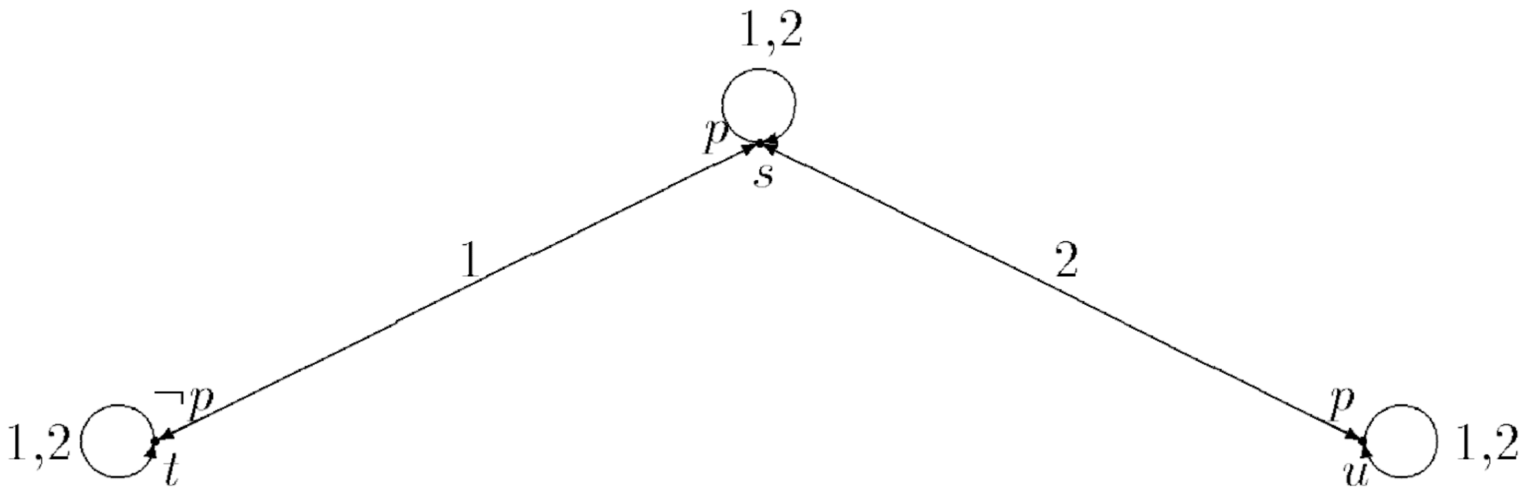
$(M,s) \models E_G\varphi \, wenn \; (M,s) \models K_i\varphi \;\; für\, alle\, i \in G$

$(M,s) \models C_G\varphi \, wenn \; (M,s) \models E_G^k\varphi \; für\, k = 1, 2, \ldots, wobei$

$E_G^1\varphi := E_G\varphi \, und \; E_G^{k+1}\varphi := E_G E_G^k\varphi$

Beispiel

- Beispielstruktur:



Kalkül

- Axiome:

$$A1: \quad Alle\ Instanzen\ aussagenlogischer\ Tautologien$$

$$A2: \quad K_i\varphi \wedge K_i(\varphi \Rightarrow \psi) \Rightarrow K_i\psi$$

$$A3: \quad K_i\varphi \Rightarrow \varphi$$

$$A4: \quad K_i\varphi \Rightarrow K_iK_i\varphi$$

$$A5: \quad \neg K_i\varphi \Rightarrow K_i\neg K_i\varphi$$

$$[(M,s) \models K_i\varphi \quad wenn\ (M,t) \models \varphi \quad für\ alle\ t\ mit\ (s,t) \in K_i]$$

Kalkül (Forts.)

- Regeln:

$$R1: \frac{\varphi, \varphi \Rightarrow \psi}{\psi} \quad (\mathit{Modus\ Ponens})$$

$$R2: \frac{\varphi}{K_i \varphi}$$

- Erweiterung für Gruppen- und Allgemeinwissen:

$$C1: \quad E_G \varphi \Leftrightarrow \wedge_{i \in G} K_i \varphi$$

$$C2: \quad C_G \varphi \Leftrightarrow E_G(\varphi \wedge C_G \varphi)$$

$$RC1: \frac{\varphi \Rightarrow E_G(\varphi \wedge \psi)}{\varphi \Rightarrow C_G \psi}$$

Multi-Agenten-System als Kripkestruktur

- System: n Agenten mit jeweils einem lokalen Zustand zu einem bestimmten Zeitpunkt
- Zustand eines Systems s (S := Menge aller Zustände):

  $$(s_e, s_1, s_2, \ldots, s_n)$$

- Funktion r (*run*) ordnet Zeitpunkten m Systemzustände zu:

  $$r(m) = (s_e, s_1, s_2, \ldots, s_n)$$

- An einem *Punkt* (r,m) ist das System im Zustand $r(m)$
- $r_i(m)$ beschreibt den lokalen Zustand von Agent i
- System wird durch eine Menge von *runs* definiert

Multi-Agenten-System als Kripkestruktur (Forts.)

- Eine Menge Φ mit primitiven Aussagen

  z.B.: „Der Wert der Variable x ist 0." oder „Das System hat einen Deadlock."

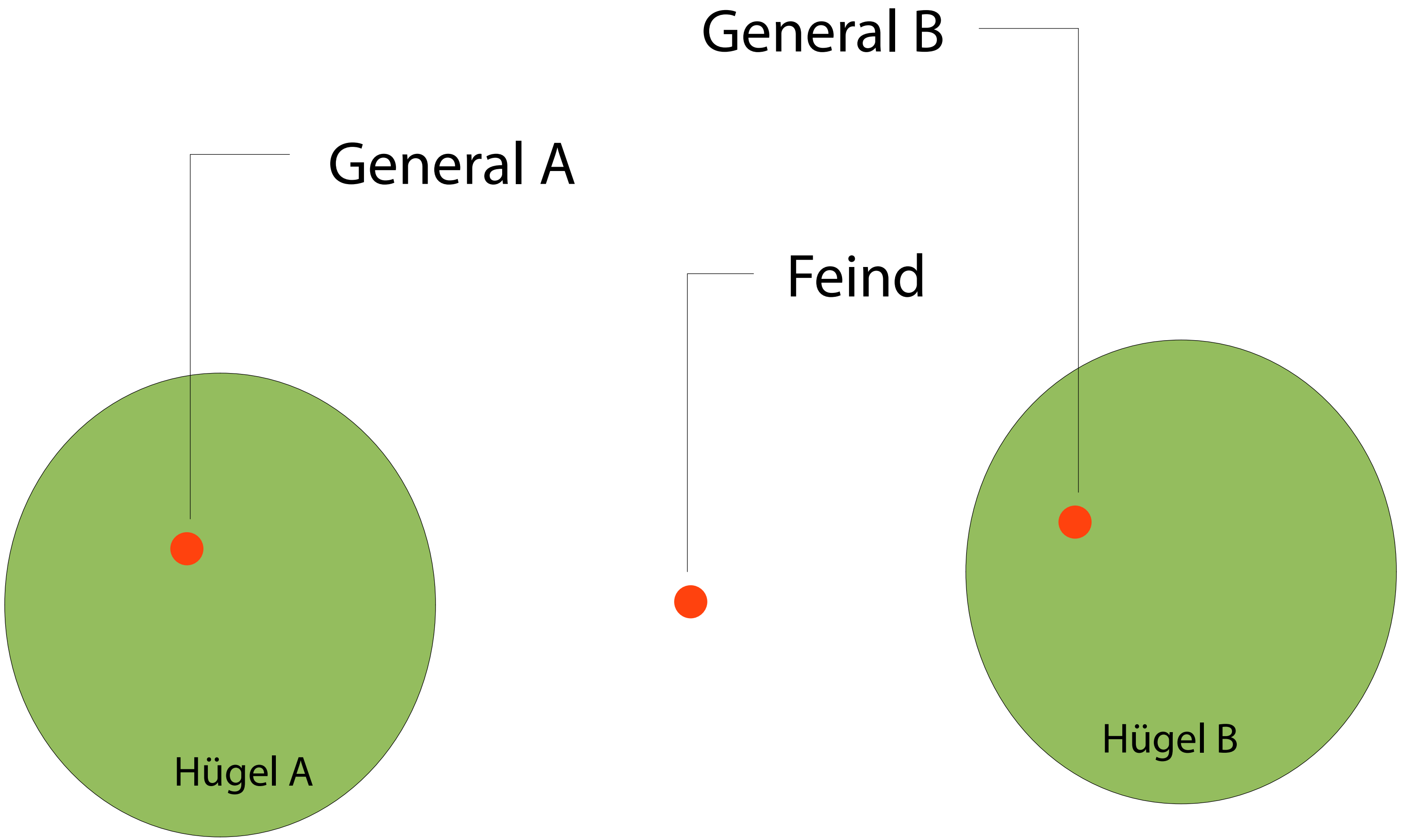
- Aussagen sind in den Systemzuständen bzw. Punkten wahr oder falsch:

  $\pi(r(m)=s)(p)\in\{true, false\}$

- Relation K_i

  $((r,m),(r',m'))\in K_i \quad wenn \ r_i(m)=r'_i(m')$

Die Byzantinischen Generäle



Die Byzantinischen Generäle: Folgerungen

- Aussage m: „Eine Nachricht mit 'Greif bei Sonnenaufgang an!' wurde von General A versendet."

  $K_B m$

  $K_A K_B m$

  $K_B K_A K_B m \dots$

- $C_G m$ bzw. $C_G(attack)$ wird niemals wahr (in einem System mit unbeschränkter Nachrichtenverzögerung)
- $attack \Rightarrow C_G(attack)$
- In einem System mit unbeschränkter Nachrichtenverzögerung greifen die Generäle niemals an.

Fragen?