

# Interaktionsrahmen in Multiagenten-Systemen

Michael Rovatsos

Lehrstuhl für Theoretische Informatik und Grundlagen der Künstlichen Intelligenz  
Fakultät für Informatik, Technische Universität München

25. Mai 2004

## Motivation

- ▶ herkömmliche Ansätze zur Spezifikation von Interaktions- und Kommunikationsformen in Multiagenten-Systemen (MAS):

## Motivation

- ▶ herkömmliche Ansätze zur Spezifikation von Interaktions- und Kommunikationsformen in Multiagenten-Systemen (MAS):
  - ▶ Agentenkommunikationssprachen (KQML/KIF, FIPA-ACL)

## Motivation

- ▶ herkömmliche Ansätze zur Spezifikation von Interaktions- und Kommunikationsformen in Multiagenten-Systemen (MAS):
  - ▶ Agentenkommunikationssprachen (KQML/KIF, FIPA-ACL)
  - ▶ Interaktionsprotokolle (CNP, Auktionen, etc.)

# Motivation

- ▶ herkömmliche Ansätze zur Spezifikation von Interaktions- und Kommunikationsformen in Multiagenten-Systemen (MAS):
  - ▶ Agentenkommunikationssprachen (KQML/KIF, FIPA-ACL)
  - ▶ Interaktionsprotokolle (CNP, Auktionen, etc.)
  - ▶ normative Vorgaben bei deren Verwendung

# Motivation

- ▶ herkömmliche Ansätze zur Spezifikation von Interaktions- und Kommunikationsformen in Multiagenten-Systemen (MAS):
  - ▶ Agentenkommunikationssprachen (KQML/KIF, FIPA-ACL)
  - ▶ Interaktionsprotokolle (CNP, Auktionen, etc.)
  - ▶ normative Vorgaben bei deren Verwendung
- ▶ stößt an Grenzen bei **offenen** MAS:

# Motivation

- ▶ herkömmliche Ansätze zur Spezifikation von Interaktions- und Kommunikationsformen in Multiagenten-Systemen (MAS):
  - ▶ Agentenkommunikationssprachen (KQML/KIF, FIPA-ACL)
  - ▶ Interaktionsprotokolle (CNP, Auktionen, etc.)
  - ▶ normative Vorgaben bei deren Verwendung
- ▶ stößt an Grenzen bei **offenen** MAS:
  - ▶ mitgliedsoffene MAS (wechselnde Populationen)

# Motivation

- ▶ herkömmliche Ansätze zur Spezifikation von Interaktions- und Kommunikationsformen in Multiagenten-Systemen (MAS):
  - ▶ Agentenkommunikationssprachen (KQML/KIF, FIPA-ACL)
  - ▶ Interaktionsprotokolle (CNP, Auktionen, etc.)
  - ▶ normative Vorgaben bei deren Verwendung
- ▶ stößt an Grenzen bei **offenen** MAS:
  - ▶ mitgliedsoffene MAS (wechselnde Populationen)
  - ▶ potentiell egoistische oder sogar böartige Agenten



# Motivation

- ▶ herkömmliche Ansätze zur Spezifikation von Interaktions- und Kommunikationsformen in Multiagenten-Systemen (MAS):
  - ▶ Agentenkommunikationssprachen (KQML/KIF, FIPA-ACL)
  - ▶ Interaktionsprotokolle (CNP, Auktionen, etc.)
  - ▶ normative Vorgaben bei deren Verwendung
- ▶ stößt an Grenzen bei **offenen** MAS:
  - ▶ mitgliedsoffene MAS (wechselnde Populationen)
  - ▶ potentiell egoistische oder sogar böartige Agenten
  - ▶ heterogenes Agenten-Design

# Motivation

- ▶ herkömmliche Ansätze zur Spezifikation von Interaktions- und Kommunikationsformen in Multiagenten-Systemen (MAS):
  - ▶ Agentenkommunikationssprachen (KQML/KIF, FIPA-ACL)
  - ▶ Interaktionsprotokolle (CNP, Auktionen, etc.)
  - ▶ normative Vorgaben bei deren Verwendung
- ▶ stößt an Grenzen bei **offenen** MAS:
  - ▶ mitgliedsoffene MAS (wechselnde Populationen)
  - ▶ potentiell egoistische oder sogar böartige Agenten
  - ▶ heterogenes Agenten-Design
- ▶ Zentrale Frage: *Wie kann man sinnvoll und erfolgreich kommunizieren, wenn man nicht weiß, ob sich Agenten an eine gegebene Kommunikationssemantik halten werden?*

## Der sozionische Ansatz

- ▶ Soziologie beschäftigt sich mit menschlichen Gesellschaften (hochgradig offen, was Kommunikation betrifft)

## Der sozionische Ansatz

- ▶ Soziologie beschäftigt sich mit menschlichen Gesellschaften (hochgradig offen, was Kommunikation betrifft)
- ▶ Anwendbarkeit sozialtheoretischer Erkenntnisse zur Entwicklung adaptiver Kommunikationsmodelle in MAS

## Der sozionische Ansatz

- ▶ Soziologie beschäftigt sich mit menschlichen Gesellschaften (hochgradig offen, was Kommunikation betrifft)
- ▶ Anwendbarkeit sozialtheoretischer Erkenntnisse zur Entwicklung adaptiver Kommunikationsmodelle in MAS
- ▶ hier: Theorien des **Symbolischen Interaktionismus** (insbes. Mead) und der Mikrosoziologie Goffmans

## Der sozionische Ansatz

- ▶ Soziologie beschäftigt sich mit menschlichen Gesellschaften (hochgradig offen, was Kommunikation betrifft)
- ▶ Anwendbarkeit sozialtheoretischer Erkenntnisse zur Entwicklung adaptiver Kommunikationsmodelle in MAS
- ▶ hier: Theorien des **Symbolischen Interaktionismus** (insbes. Mead) und der Mikrosoziologie Goffmans
- ▶ Konzentration auf Mikro-Interaktionssituationen und auf kognitiven Umgang mit Interaktionsmustern

## Der sozionische Ansatz

- ▶ zentrale Begriffe: **Rahmen** und **Rahmung** (Goffman)

## Der sozionische Ansatz

- ▶ zentrale Begriffe: **Rahmen** und **Rahmung** (Goffman)
  - ▶ beschreiben das Interaktionswissen, dass in zwischenmenschliche Interaktionssituationen einfließt



## Der sozionische Ansatz

- ▶ zentrale Begriffe: **Rahmen** und **Rahmung** (Goffman)
  - ▶ beschreiben das Interaktionswissen, dass in zwischenmenschliche Interaktionssituationen einfließt
  - ▶ Antwort auf die Frage „*was geht hier vor?*“

## Der sozionische Ansatz

- ▶ zentrale Begriffe: **Rahmen** und **Rahmung** (Goffman)
  - ▶ beschreiben das Interaktionswissen, dass in zwischenmenschliche Interaktionssituationen einfließt
  - ▶ Antwort auf die Frage „*was geht hier vor?*“
  - ▶ machen soziale Strukturen der Agentenkognition zugänglich

## Der sozionische Ansatz

- ▶ zentrale Begriffe: **Rahmen** und **Rahmung** (Goffman)
  - ▶ beschreiben das Interaktionswissen, dass in zwischenmenschliche Interaktionssituationen einfließt
  - ▶ Antwort auf die Frage „*was geht hier vor?*“
  - ▶ machen soziale Strukturen der Agentenkognition zugänglich
  - ▶ erlauben Adaption aber auch Rekonstruktion und Veränderung von Interaktionsformen

## Der sozionische Ansatz

- ▶ zentrale Begriffe: **Rahmen** und **Rahmung** (Goffman)
  - ▶ beschreiben das Interaktionswissen, dass in zwischenmenschliche Interaktionssituationen einfließt
  - ▶ Antwort auf die Frage „*was geht hier vor?*“
  - ▶ machen soziale Strukturen der Agentenkognition zugänglich
  - ▶ erlauben Adaption aber auch Rekonstruktion und Veränderung von Interaktionsformen
- ▶ informatische Operationalisierung dieser Begriffe

## Der sozionische Ansatz

- ▶ zentrale Begriffe: **Rahmen** und **Rahmung** (Goffman)
  - ▶ beschreiben das Interaktionswissen, dass in zwischenmenschliche Interaktionssituationen einfließt
  - ▶ Antwort auf die Frage „*was geht hier vor?*“
  - ▶ machen soziale Strukturen der Agentenkognition zugänglich
  - ▶ erlauben Adaption aber auch Rekonstruktion und Veränderung von Interaktionsformen
- ▶ informatische Operationalisierung dieser Begriffe
  - ▶ Entwicklung rahmenbasierter Agentenarchitektur

## Der sozionische Ansatz

- ▶ zentrale Begriffe: **Rahmen** und **Rahmung** (Goffman)
  - ▶ beschreiben das Interaktionswissen, dass in zwischenmenschliche Interaktionssituationen einfließt
  - ▶ Antwort auf die Frage „*was geht hier vor?*“
  - ▶ machen soziale Strukturen der Agentenkognition zugänglich
  - ▶ erlauben Adaption aber auch Rekonstruktion und Veränderung von Interaktionsformen
- ▶ informatische Operationalisierung dieser Begriffe
  - ▶ Entwicklung rahmenbasierter Agentenarchitektur
  - ▶ Formalisierung entsprechender Kommunikationssemantik

## Der sozionische Ansatz

- ▶ zentrale Begriffe: **Rahmen** und **Rahmung** (Goffman)
  - ▶ beschreiben das Interaktionswissen, dass in zwischenmenschliche Interaktionssituationen einfließt
  - ▶ Antwort auf die Frage „*was geht hier vor?*“
  - ▶ machen soziale Strukturen der Agentenkognition zugänglich
  - ▶ erlauben Adaption aber auch Rekonstruktion und Veränderung von Interaktionsformen
- ▶ informatische Operationalisierung dieser Begriffe
  - ▶ Entwicklung rahmenbasierter Agentenarchitektur
  - ▶ Formalisierung entsprechender Kommunikationssemantik
  - ▶ Entwurf geeigneter Lernalgorithmen

## Die MAS-Perspektive

- ▶ Praktischer Nutzen jenseits sozionischer Ziele



## Die MAS-Perspektive

- ▶ Praktischer Nutzen jenseits sozionischer Ziele
- ▶ Erlernen der strategischen Anwendung vorgegebener Interaktionsmustern

## Die MAS-Perspektive

- ▶ Praktischer Nutzen jenseits sozionischer Ziele
- ▶ Erlernen der strategischen Anwendung vorgegebener Interaktionsmustern
- ▶ Konstruktion neuer Interaktionsformen aus bestehenden

## Die MAS-Perspektive

- ▶ Praktischer Nutzen jenseits sozionischer Ziele
- ▶ Erlernen der strategischen Anwendung vorgegebener Interaktionsmustern
- ▶ Konstruktion neuer Interaktionsformen aus bestehenden
- ▶ Kombination aus vorgegebenem Kontextwissen und empirischer Kommunikationserfahrung

## Die MAS-Perspektive

- ▶ Praktischer Nutzen jenseits sozionischer Ziele
- ▶ Erlernen der strategischen Anwendung vorgegebener Interaktionsmustern
- ▶ Konstruktion neuer Interaktionsformen aus bestehenden
- ▶ Kombination aus vorgegebenem Kontextwissen und empirischer Kommunikationserfahrung
- ▶ abstraktes Modell des „Interaktions-Reasoning“ jenseits der konkreten Formalisierung

# Gliederung

Einführung

Die Agentenarchitektur InFFrA

Überblick

Interaktionsrahmen in InFFrA

Inferenzmodell

Weitere Eigenschaften

Formale Semantik & Lernalgorithmen

Experimentelle Befunde

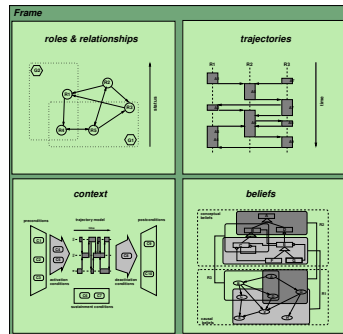
Zusammenfassung & Ausblick

## Überblick

- ▶ InFFrA = **I**nteraction **F**rames and **F**raming  
**A**rchitecture

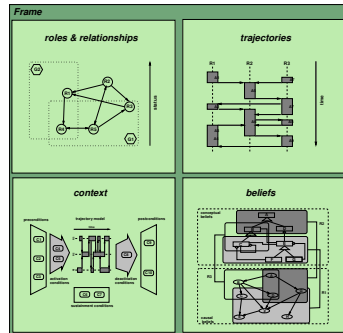
# Überblick

- ▶ InFFrA = **I**nteraction **F**rames and **F**raming **A**rchitecture
- ▶ **Rahmen** = Modell einer Klasse von Interaktionssituationen
  - ▶ Trajektorien (Handlungsverläufe)
  - ▶ Rollen & Beziehungen zwischen Teilnehmern
  - ▶ Kontextbedingungen
  - ▶ Epistemische Zustände der Akteure



# Überblick

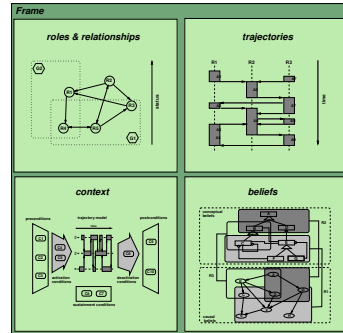
- ▶ InFFrA = **I**nteraction **F**rames and **F**raming **A**rchitecture
- ▶ **Rahmen** = Modell einer Klasse von Interaktionssituationen
  - ▶ Trajektorien (Handlungsverläufe)
  - ▶ Rollen & Beziehungen zwischen Teilnehmern
  - ▶ Kontextbedingungen
  - ▶ Epistemische Zustände der Akteure
- ▶ **Rahmung** = Prozess der Rahmenkonstruktion und -anwendung





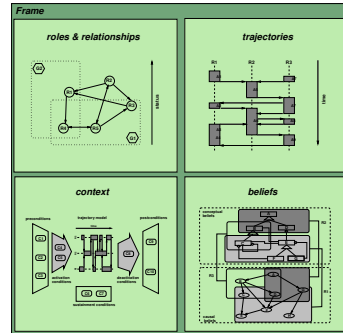
## Überblick

- ▶ InFFrA = **I**nteraction **F**rames and **F**raming Architecture
- ▶ **Rahmen** = Modell einer Klasse von Interaktionssituationen
  - ▶ Trajektorien (Handlungsverläufe)
  - ▶ Rollen & Beziehungen zwischen Teilnehmern
  - ▶ Kontextbedingungen
  - ▶ Epistemische Zustände der Akteure
- ▶ **Rahmung** = Prozess der Rahmenkonstruktion und -anwendung
- ▶ Abstraktes Modell auf konzeptioneller Ebene

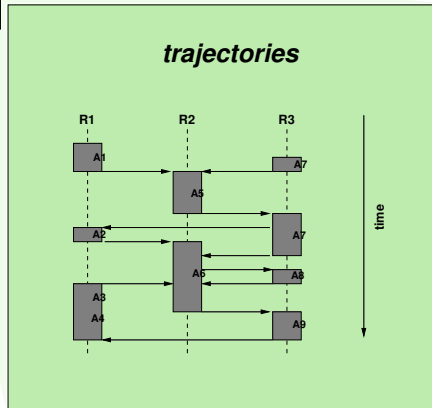
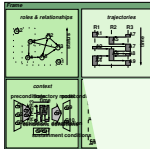


## Überblick

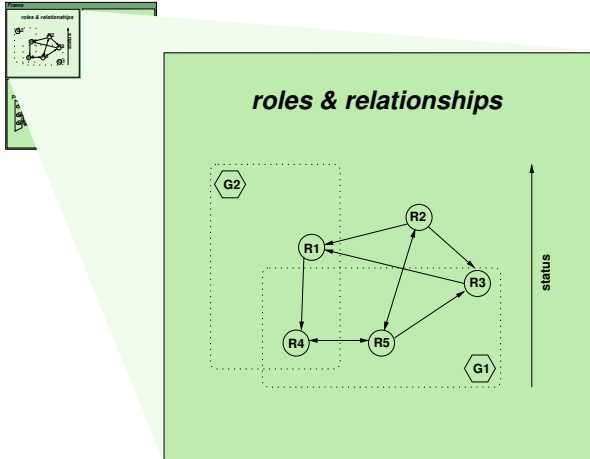
- ▶ InFFrA = **I**nteraction **F**rames and **F**raming Architecture
- ▶ **Rahmen** = Modell einer Klasse von Interaktionssituationen
  - ▶ Trajektorien (Handlungsverläufe)
  - ▶ Rollen & Beziehungen zwischen Teilnehmern
  - ▶ Kontextbedingungen
  - ▶ Epistemische Zustände der Akteure
- ▶ **Rahmung** = Prozess der Rahmenkonstruktion und -anwendung
- ▶ Abstraktes Modell auf konzeptioneller Ebene
- ▶ keine Vorgaben hinsichtlich Implementierung



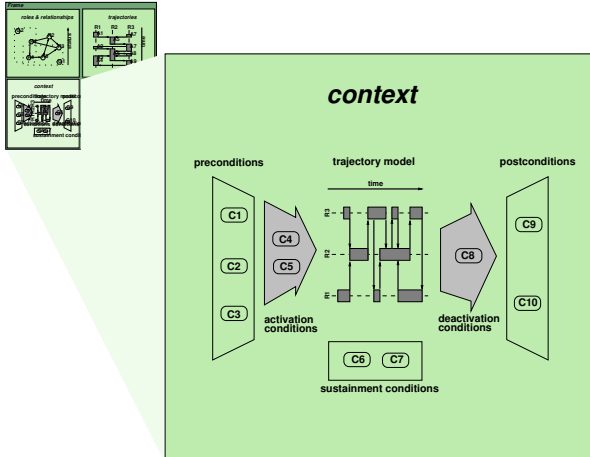
# Rahmenattribute



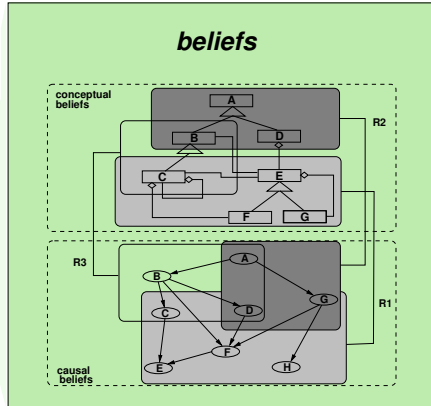
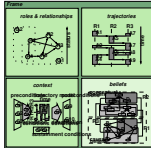
# Rahmenattribute



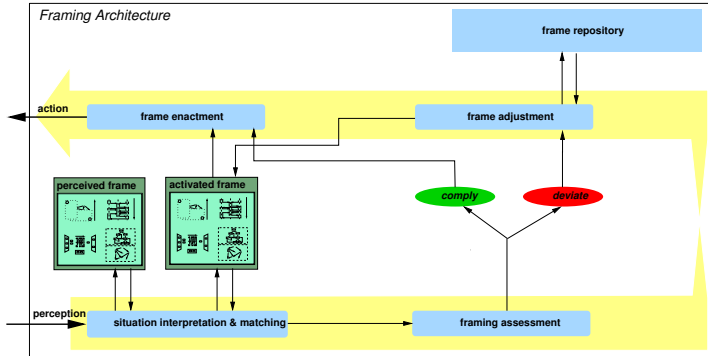
# Rahmenattribute



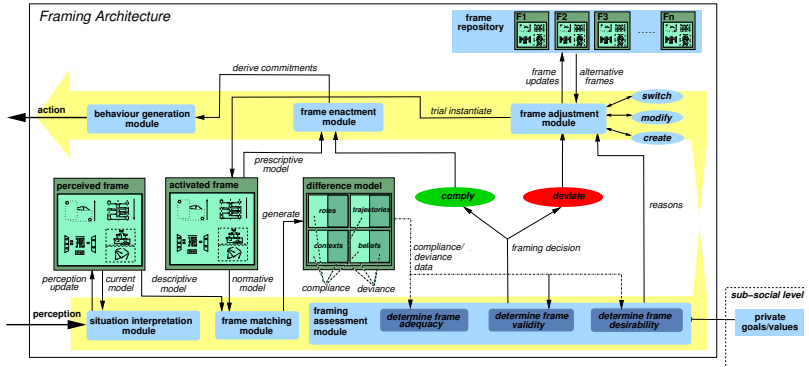
# Rahmenattribute



# Soziale Inferenz in InFFrA



# Soziale Inferenz in InFFrA





## Weitere Eigenschaften

- ▶ Meta-Attribute von Rahmen

## Weitere Eigenschaften

- ▶ Meta-Attribute von Rahmen
  - ▶ Status: Beschreibung aktuellen Rahmungszustandes

## Weitere Eigenschaften

- ▶ Meta-Attribute von Rahmen
  - ▶ Status: Beschreibung aktuellen Rahmungszustandes
  - ▶ Links: Beziehungen zwischen Rahmen (z.B. Alternativen, Vererbung, etc.)

## Weitere Eigenschaften

- ▶ Meta-Attribute von Rahmen
  - ▶ Status: Beschreibung aktuellen Rahmungszustandes
  - ▶ Links: Beziehungen zwischen Rahmen (z.B. Alternativen, Vererbung, etc.)
  - ▶ Historie: Beschreibung der Verwendung und Entwicklung eines Rahmens

## Weitere Eigenschaften

- ▶ Meta-Attribute von Rahmen
  - ▶ Status: Beschreibung aktuellen Rahmungszustandes
  - ▶ Links: Beziehungen zwischen Rahmen (z.B. Alternativen, Vererbung, etc.)
  - ▶ Historie: Beschreibung der Verwendung und Entwicklung eines Rahmens
  - ▶ „Verbreitung“: Aussagen über Wissen versch. Agenten von einem Rahmen

## Weitere Eigenschaften

- ▶ Meta-Attribute von Rahmen
  - ▶ Status: Beschreibung aktuellen Rahmungszustandes
  - ▶ Links: Beziehungen zwischen Rahmen (z.B. Alternativen, Vererbung, etc.)
  - ▶ Historie: Beschreibung der Verwendung und Entwicklung eines Rahmens
  - ▶ „Verbreitung“: Aussagen über Wissen versch. Agenten von einem Rahmen
- ▶ umfassendes Modell von Interaktionskategorien und ihrer kognitiven Verarbeitung durch Agenten

## Weitere Eigenschaften

- ▶ Meta-Attribute von Rahmen
  - ▶ Status: Beschreibung aktuellen Rahmungszustandes
  - ▶ Links: Beziehungen zwischen Rahmen (z.B. Alternativen, Vererbung, etc.)
  - ▶ Historie: Beschreibung der Verwendung und Entwicklung eines Rahmens
  - ▶ „Verbreitung“: Aussagen über Wissen versch. Agenten von einem Rahmen
- ▶ umfassendes Modell von Interaktionskategorien und ihrer kognitiven Verarbeitung durch Agenten
- ▶ dennoch keine „vollständige“ Agentenarchitektur (nur Interaktionsmanagement)

# Gliederung

Einführung

Die Agentenarchitektur InFFrA

**Formale Semantik & Lernalgorithmen**

Kommunikationssysteme/Empirische Semantik

Rahmen & Rahmensemantik

Lernalgorithmen

Experimentelle Befunde

Zusammenfassung & Ausblick



## m<sup>2</sup>inffra

- ▶ InFFrA liefert abstraktes Modell, nicht direkt implementierbar

## m<sup>2</sup>inffra

- ▶ InFFrA liefert abstraktes Modell, nicht direkt implementierbar
- ▶ Formalisierung in m<sup>2</sup>inffra („Markov<sup>2</sup>“-InFFrA)

## m<sup>2</sup>inffra

- ▶ InFFrA liefert abstraktes Modell, nicht direkt implementierbar
- ▶ Formalisierung in m<sup>2</sup>inffra („Markov<sup>2</sup>“-InFFrA)
  - ▶ Rahmentrajektorien als Sequenzen von Zweier-Konversationen (turn-taking)

## m<sup>2</sup>inffra

- ▶ InFFrA liefert abstraktes Modell, nicht direkt implementierbar
- ▶ Formalisierung in m<sup>2</sup>inffra („Markov<sup>2</sup>“-InFFrA)
  - ▶ Rahmentrajektorien als Sequenzen von Zweier-Konversationen (turn-taking)
  - ▶ Darstellung beliebiger Bedingungen mittels aussagenlogischer Ausdrücke (Rollen, Kontext, epist. Zustände)

## m<sup>2</sup>inffra

- ▶ InFFrA liefert abstraktes Modell, nicht direkt implementierbar
- ▶ Formalisierung in m<sup>2</sup>inffra („Markov<sup>2</sup>“-InFFrA)
  - ▶ Rahmentrajektorien als Sequenzen von Zweier-Konversationen (turn-taking)
  - ▶ Darstellung beliebiger Bedingungen mittels aussagenlogischer Ausdrücke (Rollen, Kontext, epist. Zustände)
  - ▶ Repräsentation früherer Instanzen durch Nachrichten-Pattern und Substitutionen

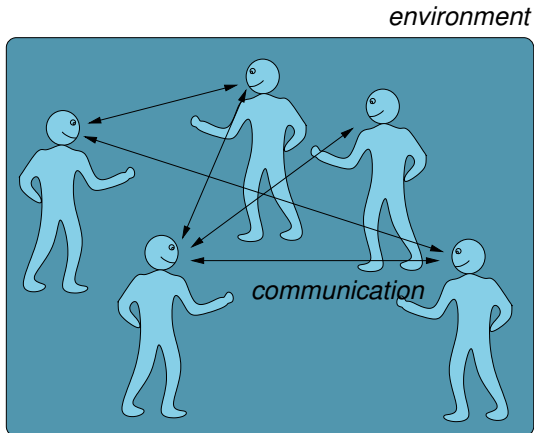
## m<sup>2</sup>inffra

- ▶ InFFrA liefert abstraktes Modell, nicht direkt implementierbar
- ▶ Formalisierung in m<sup>2</sup>inffra („Markov<sup>2</sup>“-InFFrA)
  - ▶ Rahmentrajektorien als Sequenzen von Zweier-Konversationen (turn-taking)
  - ▶ Darstellung beliebiger Bedingungen mittels aussagenlogischer Ausdrücke (Rollen, Kontext, epist. Zustände)
  - ▶ Repräsentation früherer Instanzen durch Nachrichten-Pattern und Substitutionen
- ▶ Semantik basiert auf Idee der **empirischen Semantik** („Kommunikationssysteme“)

## m<sup>2</sup>inffra

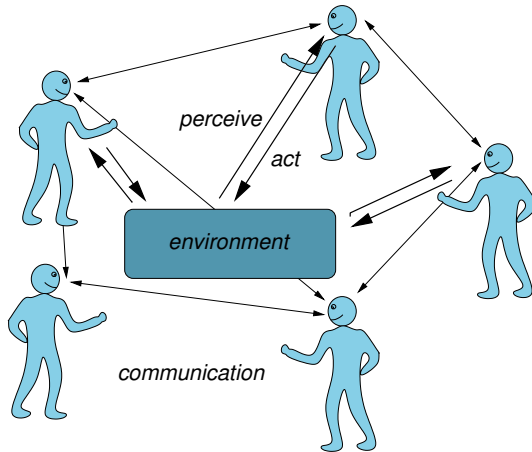
- ▶ InFFrA liefert abstraktes Modell, nicht direkt implementierbar
- ▶ Formalisierung in m<sup>2</sup>inffra („Markov<sup>2</sup>“-InFFrA)
  - ▶ Rahmentrajektorien als Sequenzen von Zweier-Konversationen (turn-taking)
  - ▶ Darstellung beliebiger Bedingungen mittels aussagenlogischer Ausdrücke (Rollen, Kontext, epist. Zustände)
  - ▶ Repräsentation früherer Instanzen durch Nachrichten-Pattern und Substitutionen
- ▶ Semantik basiert auf Idee der **empirischen Semantik** („Kommunikationssysteme“)
- ▶ Anwendung von Methoden des hierarchischen Reinforcement-Lernens

# Kommunikationssysteme

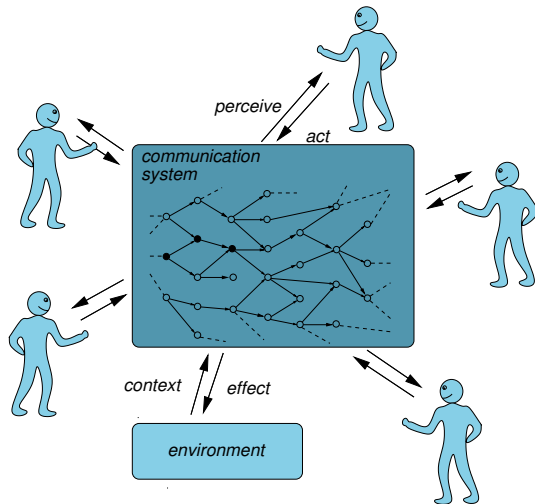




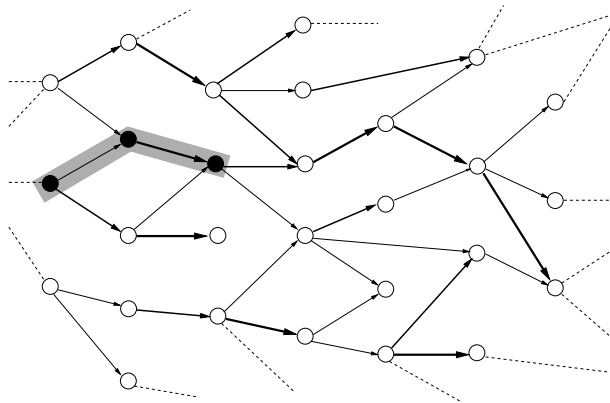
# Kommunikationssysteme



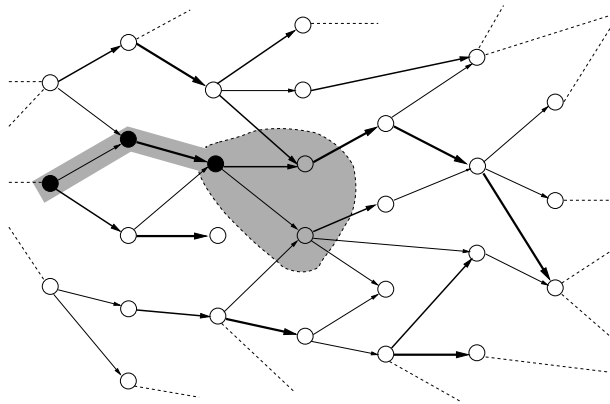
# Kommunikationssysteme



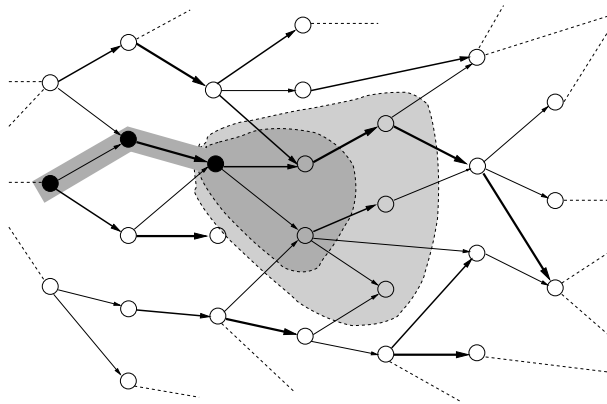
## Empirische Semantik



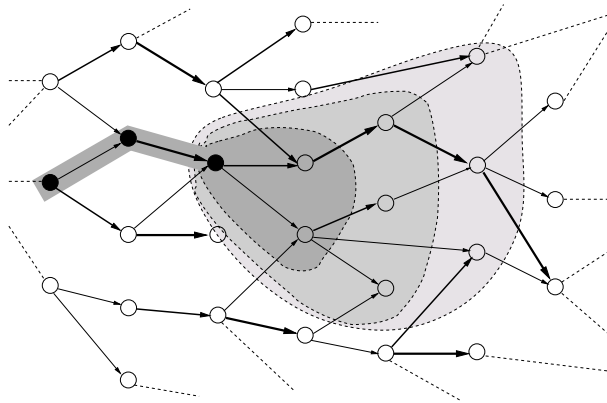
## Empirische Semantik



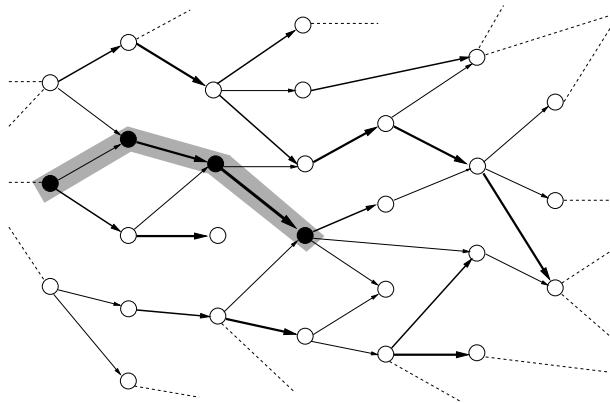
# Empirische Semantik



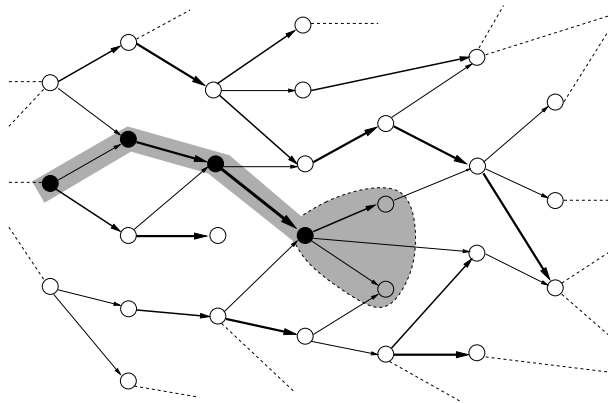
# Empirische Semantik



## Empirische Semantik

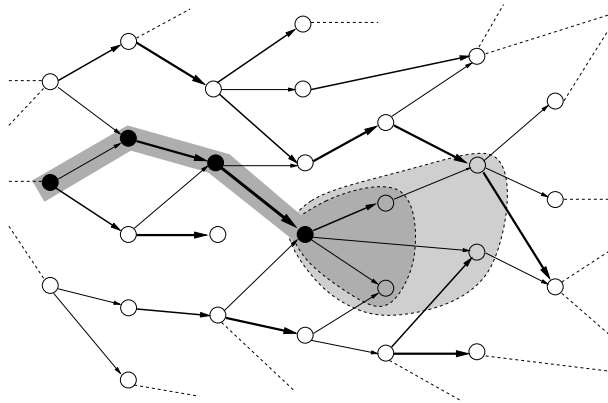


# Empirische Semantik

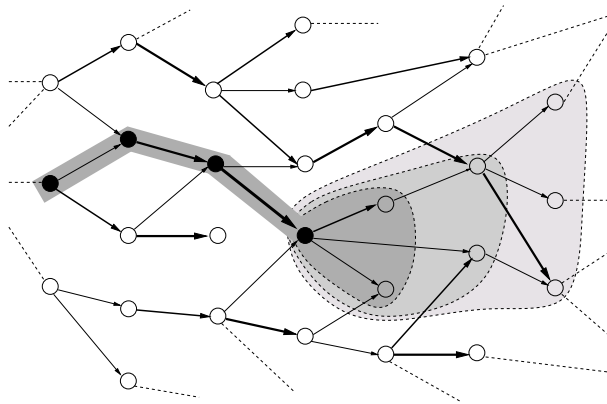




# Empirische Semantik



## Empirische Semantik



## Rahmen in $m^2$ inffra

- ▶ Rahmen  $F = (T, \Theta, C, h_T, h_\Theta)$ 
  - ▶  $T$  Sequenz von Nachrichten-Patterns mit Variablen
  - ▶  $\Theta$  Liste von Variablensubstitutionen
  - ▶  $C$  Liste logischer Klauselmengen
  - ▶  $h_T$  Trajektorienzähler,  $h_\Theta$  Substitutionszähler

## Rahmen in $m^2$ inffra

- ▶ Rahmen  $F = (T, \Theta, C, h_T, h_\Theta)$ 
  - ▶  $T$  Sequenz von Nachrichten-Patterns mit Variablen
  - ▶  $\Theta$  Liste von Variablensubstitutionen
  - ▶  $C$  Liste logischer Klauselmengen
  - ▶  $h_T$  Trajektorienzähler,  $h_\Theta$  Substitutionszähler
- ▶ Beispiel:

$$F = \left\langle \left\langle \begin{array}{l} \xrightarrow{5} \text{request}(A_1, A_2, X) \xrightarrow{3} \text{accept}(A_2, A_1, X) \\ \xrightarrow{2} \text{confirm}(A_1, A_2, X) \xrightarrow{2} \text{do}(A_2, X) \end{array} \right\rangle, \right. \\ \left. \left\langle \{ \text{self}(A_1), \text{other}(A_2), \text{can}(A_1, \text{do}(A_1, X)) \}, \right. \right. \\ \left. \left. \{ \text{agent}(A_1), \text{agent}(A_2), \text{action}(X) \} \right\rangle, \right. \\ \left\langle \xrightarrow{4} \langle [A_1/\text{agent}_1], [A_2/\text{agent}_2] \rangle, \right. \\ \left. \left. \xrightarrow{1} \langle [A_1/\text{agent}_3], [A_2/\text{agent}_1], [X/\text{deliver\_goods}] \rangle \right\rangle \right\rangle$$

## Rahmensemantik

- ▶ gegeben eine Menge von Rahmen  $\mathcal{F} = \{F_1, \dots, F_n\}$ , ein aktuelles Dialog-Präfix  $w$  und eine aktuelle Wissensbasis  $KB$

## Rahmensemantik

- ▶ gegeben eine Menge von Rahmen  $\mathcal{F} = \{F_1, \dots, F_n\}$ , ein aktuelles Dialog-Präfix  $w$  und eine aktuelle Wissensbasis  $KB$
- ▶ berechne „Anschlusswahrscheinlichkeiten“ für mögliche  $w'$

$$P(w'|w) = \sum_{F \in \mathcal{F}, ww' = T(F)\vartheta} P(\vartheta|F, w)P(F|w)$$

wobei  $\vartheta \in \Theta_{\text{poss}}(F, KB, w)$  eine „noch mögliche Substitution“ sein muss

## Rahmensemantik

- ▶ gegeben eine Menge von Rahmen  $\mathcal{F} = \{F_1, \dots, F_n\}$ , ein aktuelles Dialog-Präfix  $w$  und eine aktuelle Wissensbasis  $KB$
- ▶ berechne „Anschlusswahrscheinlichkeiten“ für mögliche  $w'$

$$P(w'|w) = \sum_{F \in \mathcal{F}, ww' = T(F)\vartheta} P(\vartheta|F, w)P(F|w)$$

wobei  $\vartheta \in \Theta_{\text{poss}}(F, KB, w)$  eine „noch mögliche Substitution“ sein muss

- ▶ Wahrscheinlichkeit von  $\vartheta$  unter  $F$  gemäß Ähnlichkeit verteilt:

$$P(\vartheta|F) \propto \sigma(\chi, F)$$

## Rahmensemantik

- ▶ Ähnlichkeit basiert auf Ähnlichkeitsmaß  $\sigma$  auf Nachrichten-Patterns, Häufigkeit und Kontextbedingungen:

$$\sigma(\vartheta, F) = \sum_{i=1}^{|\Theta(F)|} \overbrace{\sigma(T(F)\vartheta, T(F)\Theta(F)[i])}^{\text{Ähnlichkeit}} \overbrace{h_{\Theta(F)}[i]}^{\text{Häufigkeit}} \overbrace{c_i(F, \vartheta, KB)}^{\text{Relevanz}}$$



## Rahmensemantik

- ▶ Ähnlichkeit basiert auf Ähnlichkeitsmaß  $\sigma$  auf Nachrichten-Patterns, Häufigkeit und Kontextbedingungen:

$$\sigma(\vartheta, F) = \sum_{i=1}^{|\Theta(F)|} \overbrace{\sigma(T(F)\vartheta, T(F)\Theta(F)[i])}^{\text{Ähnlichkeit}} \overbrace{h_{\Theta(F)}[i]}^{\text{Häufigkeit}} \overbrace{c_i(F, \vartheta, KB)}^{\text{Relevanz}}$$

- ▶ Relevanz-Funktion z.B. durch Heranziehen gültiger Bedingungsmengen:

$$c_i(F, \vartheta, KB) = 1 \Leftrightarrow KB \models C(F)[i]\Theta(F)[i]\vartheta$$

## Rahmensemantik

- ▶ Ähnlichkeit basiert auf Ähnlichkeitsmaß  $\sigma$  auf Nachrichten-Patterns, Häufigkeit und Kontextbedingungen:

$$\sigma(\vartheta, F) = \sum_{i=1}^{|\Theta(F)|} \overbrace{\sigma(T(F)\vartheta, T(F)\Theta(F)[i])}^{\text{Ähnlichkeit}} \overbrace{h_{\Theta(F)}[i]}^{\text{Häufigkeit}} \overbrace{c_i(F, \vartheta, KB)}^{\text{Relevanz}}$$

- ▶ Relevanz-Funktion z.B. durch Heranziehen gültiger Bedingungsmengen:

$$c_i(F, \vartheta, KB) = 1 \Leftrightarrow KB \models C(F)[i]\Theta(F)[i]\vartheta$$

- ▶ nicht alle  $w'$  durch  $\mathcal{F}$  abgedeckt  $\Rightarrow$  Zuweisung kleiner Restwahrscheinlichkeit an alle nicht abgedeckten  $w'$

# Markov'sche Entscheidungsprozesse (MDPs)

- ▶ formales Modell sequentieller Entscheidungsprozesse

# Markov'sche Entscheidungsprozesse (MDPs)

- ▶ formales Modell sequentieller Entscheidungsprozesse
- ▶ Definition:
  - ▶  $\mathcal{S}$  Zustandsmenge
  - ▶  $\mathcal{A}_s$  Aktionsmenge für  $s \in \mathcal{S}$
  - ▶  $p_{ss'}^a = P(s_{t+1} | s_t = s, a_t = a)$  Zustandsübergangmodell
  - ▶  $r_s^a = E(r_{t+1} | s_t = s, a_t = a)$  (erwarteter) Reward bei Ausführung von  $a$  in  $s$

## Markov'sche Entscheidungsprozesse (MDPs)

- ▶ formales Modell sequentieller Entscheidungsprozesse
- ▶ Definition:
  - ▶  $\mathcal{S}$  Zustandsmenge
  - ▶  $\mathcal{A}_s$  Aktionsmenge für  $s \in \mathcal{S}$
  - ▶  $p_{ss'}^a = P(s_{t+1} | s_t = s, a_t = a)$  Zustandsübergangsmodell
  - ▶  $r_s^a = E(r_{t+1} | s_t = s, a_t = a)$  (erwarteter) Reward bei Ausführung von  $a$  in  $s$
- ▶ Markov-Eigenschaft:  $p_{ss'}^a$  und  $r_s^a$  hängen nur von  $s$  ab

## Markov'sche Entscheidungsprozesse (MDPs)

- ▶ formales Modell sequentieller Entscheidungsprozesse
- ▶ Definition:
  - ▶  $\mathcal{S}$  Zustandsmenge
  - ▶  $\mathcal{A}_s$  Aktionsmenge für  $s \in \mathcal{S}$
  - ▶  $p_{ss'}^a = P(s_{t+1} | s_t = s, a_t = a)$  Zustandsübergangmodell
  - ▶  $r_s^a = E(r_{t+1} | s_t = s, a_t = a)$  (erwarteter) Reward bei Ausführung von  $a$  in  $s$
- ▶ Markov-Eigenschaft:  $p_{ss'}^a$  und  $r_s^a$  hängen nur von  $s$  ab
- ▶ Agentenverhalten wird mittels (stochastischer, diskreter) sog. *Policy*  $\pi : \mathcal{S} \times \bigcup_s \mathcal{A}_s \rightarrow [0, 1]$  modelliert

## Reinforcement Learning (Verstärkungslernen)

- ▶ Ziel: Erlernen einer optimalen Policy  $\pi^*$  in einem MDP

## Reinforcement Learning (Verstärkungslernen)

- ▶ Ziel: Erlernen einer optimalen Policy  $\pi^*$  in einem MDP
- ▶ Optimalitätskriterium: Maximierung des erwarteten langfristigen Rewards,  
d.h.  $\pi^* = \arg \max_{\pi} E(\sum_{\tau=1}^{\infty} \gamma^{\tau-1} r_{t+\tau} | s_t = s, \pi)$



## Reinforcement Learning (Verstärkungslernen)

- ▶ Ziel: Erlernen einer optimalen Policy  $\pi^*$  in einem MDP
- ▶ Optimalitätskriterium: Maximierung des erwarteten langfristigen Rewards,  
d.h.  $\pi^* = \arg \max_{\pi} E(\sum_{\tau=1}^{\infty} \gamma^{\tau-1} r_{t+\tau} | s_t = s, \pi)$
- ▶ Problem: Dimensionalität der möglichen Zustands-Aktions-Paare

## Reinforcement Learning (Verstärkungslernen)

- ▶ Ziel: Erlernen einer optimalen Policy  $\pi^*$  in einem MDP
- ▶ Optimalitätskriterium: Maximierung des erwarteten langfristigen Rewards,  
d.h.  $\pi^* = \arg \max_{\pi} E(\sum_{\tau=1}^{\infty} \gamma^{\tau-1} r_{t+\tau} | s_t = s, \pi)$
- ▶ Problem: Dimensionalität der möglichen Zustands-Aktions-Paare
- ▶ mögliche Lösung: Verwendung von „Makro-Aktionen“

## Reinforcement Learning (Verstärkungslernen)

- ▶ Ziel: Erlernen einer optimalen Policy  $\pi^*$  in einem MDP
- ▶ Optimalitätskriterium: Maximierung des erwarteten langfristigen Rewards,  
d.h.  $\pi^* = \arg \max_{\pi} E(\sum_{\tau=1}^{\infty} \gamma^{\tau-1} r_{t+\tau} | s_t = s, \pi)$
- ▶ Problem: Dimensionalität der möglichen Zustands-Aktions-Paare
- ▶ mögliche Lösung: Verwendung von „Makro-Aktionen“
- ▶ deren Modellierung führt zu einem semi-MDP

## Reinforcement Learning (Verstärkungslernen)

- ▶ Ziel: Erlernen einer optimalen Policy  $\pi^*$  in einem MDP
- ▶ Optimalitätskriterium: Maximierung des erwarteten langfristigen Rewards,  
d.h.  $\pi^* = \arg \max_{\pi} E(\sum_{\tau=1}^{\infty} \gamma^{\tau-1} r_{t+\tau} | s_t = s, \pi)$
- ▶ Problem: Dimensionalität der möglichen Zustands-Aktions-Paare
- ▶ mögliche Lösung: Verwendung von „Makro-Aktionen“
- ▶ deren Modellierung führt zu einem semi-MDP
- ▶ Übergangswahrscheinlichkeiten/Rewards hängen von *Historie* der Zustandsübergänge seit Aufruf der Makro-Aktion ab

## Der Options-Ansatz

- ▶ Ansatz des **hierarchischen** RL, der gut zu  $m^2$ inffra passt

## Der Options-Ansatz

- ▶ Ansatz des **hierarchischen** RL, der gut zu  $m^2$ inffra passt
- ▶  $\mathcal{A}_s$  (des „Kern“-MDPs) wird um sog. „Optionen“ erweitert

## Der Options-Ansatz

- ▶ Ansatz des **hierarchischen** RL, der gut zu  $m^2$ inffra passt
- ▶  $\mathcal{A}_s$  (des „Kern“-MDPs) wird um sog. „Optionen“ erweitert
- ▶ *Option*  $o = (\mathcal{I}, \pi, \beta)$ 
  - $\mathcal{I} \subseteq \mathcal{S}$  Eingabemenge
  - $\pi : \mathcal{S} \times \bigcup_s \mathcal{A}_s \rightarrow [0, 1]$  Policy innerhalb  $o$
  - $\beta : \mathcal{S} \rightarrow [0, 1]$  Terminierungsbedingung

## Der Options-Ansatz

- ▶ Ansatz des **hierarchischen** RL, der gut zu  $m^2$ inffra passt
- ▶  $\mathcal{A}_s$  (des „Kern“-MDPs) wird um sog. „Optionen“ erweitert
- ▶ *Option*  $o = (\mathcal{I}, \pi, \beta)$ 
  - $\mathcal{I} \subseteq \mathcal{S}$  Eingabemenge
  - $\pi : \mathcal{S} \times \bigcup_s \mathcal{A}_s \rightarrow [0, 1]$  Policy innerhalb  $o$
  - $\beta : \mathcal{S} \rightarrow [0, 1]$  Terminierungsbedingung
- ▶  $\mathcal{I}$  und  $\beta$  bestimmen, wann  $o$  angewendet wird,  $\pi$  wird verwendet, solange  $o$  „aktiviert“ ist



## Der Options-Ansatz

- ▶ Ansatz des **hierarchischen** RL, der gut zu  $m^2$ infra passt
- ▶  $\mathcal{A}_S$  (des „Kern“-MDPs) wird um sog. „Optionen“ erweitert
- ▶ *Option*  $o = (\mathcal{I}, \pi, \beta)$ 
  - $\mathcal{I} \subseteq \mathcal{S}$  Eingabemenge
  - $\pi : \mathcal{S} \times \bigcup_S \mathcal{A}_S \rightarrow [0, 1]$  Policy innerhalb  $o$
  - $\beta : \mathcal{S} \rightarrow [0, 1]$  Terminierungsbedingung
- ▶  $\mathcal{I}$  und  $\beta$  bestimmen, wann  $o$  angewendet wird,  $\pi$  wird verwendet, solange  $o$  „aktiviert“ ist
- ▶ Preis für Komplexitätsreduktion: Konvergenz von Lernalgorithmen kann nicht mehr garantiert werden

## Q-Lernen mit Optionen

- ▶ Q-Lernen: Erlernen des „Wertes“  $Q^*(s, a)$  einer Aktion  $a$  in Zustand  $s$  unter Annahme, dass danach  $\pi^*$  verfolgt wird

## Q-Lernen mit Optionen

- ▶ Q-Lernen: Erlernen des „Wertes“  $Q^*(s, a)$  einer Aktion  $a$  in Zustand  $s$  unter Annahme, dass danach  $\pi^*$  verfolgt wird
- ▶ möglich durch Approximierung von  $Q^*$  mithilfe beobachteter Aktionen und Zustandsübergänge

## Q-Lernen mit Optionen

- ▶ Q-Lernen: Erlernen des „Wertes“  $Q^*(s, a)$  einer Aktion  $a$  in Zustand  $s$  unter Annahme, dass danach  $\pi^*$  verfolgt wird
- ▶ möglich durch Approximierung von  $Q^*$  mithilfe beobachteter Aktionen und Zustandsübergänge
- ▶ Formel für Q-Update in semi-MDP:

$$Q(s, o) \leftarrow (1 - \alpha)Q_k(s, o) + \alpha \left[ r + \gamma^T \max_{o' \in \mathcal{O}_{s'}} Q_k(s', o') \right]$$

## Q-Lernen mit Optionen

- ▶ Q-Lernen: Erlernen des „Wertes“  $Q^*(s, a)$  einer Aktion  $a$  in Zustand  $s$  unter Annahme, dass danach  $\pi^*$  verfolgt wird
- ▶ möglich durch Approximierung von  $Q^*$  mithilfe beobachteter Aktionen und Zustandsübergänge
- ▶ Formel für Q-Update in semi-MDP:

$$Q(s, o) \leftarrow (1 - \alpha)Q_k(s, o) + \alpha \left[ r + \gamma^T \max_{o' \in \mathcal{O}_{s'}} Q_k(s', o') \right]$$

- ▶ Optimale Policy ergibt sich dann als

$$\pi^*(s, a) = \arg \max_{a'} Q^*(s, a')$$

da bei geeigneter Exploration  $Q_k$  zu  $Q^*$  konvergiert

## Rahmen als Optionen

Für jedes  $F$  kann eine Option  $(\mathcal{I}_F, \pi_F, \beta_F)$  definiert werden:

$\mathcal{I}_F$  Menge  $(w, KB)$  bei der  $T(F)$  das Präfix  $w$  erfüllt und unter  $KB$  das entspr. Suffix ausführbar ist

$\pi_F$  führt die nächste eigene Aktion auf  $T(F)$  mit Wahrscheinlichkeit 1 aus (unter Maximierung des erwarteten Rewards bei mehreren möglichen Substitutionen  $\vartheta$ )  
hierzu muss „innerhalb“ von  $F$  die optimale Aktion ermittelt werden

$\beta_F$  wird durch  $T(F)$ ,  $w$  und  $KB$  bestimmt sowie durch „Desirability“-Kriterium

## Optimierung innerhalb eines Rahmens

- ▶ auf der Ebene des „unteren“ MDPs wird kein Q-Lernen verwendet, sondern Maximierung der erwarteten Utility auf der Basis von Anschlusswahrscheinlichkeiten

## Optimierung innerhalb eines Rahmens

- ▶ auf der Ebene des „unteren“ MDPs wird kein Q-Lernen verwendet, sondern Maximierung der erwarteten Utility auf der Basis von Anschlusswahrscheinlichkeiten
- ▶ Voraussetzung: Agent verfügt über Utility-Schätzung  $u(w, KB) \in \mathbb{R}$



## Optimierung innerhalb eines Rahmens

- ▶ auf der Ebene des „unteren“ MDPs wird kein Q-Lernen verwendet, sondern Maximierung der erwarteten Utility auf der Basis von Anschlusswahrscheinlichkeiten
- ▶ Voraussetzung: Agent verfügt über Utility-Schätzung  $u(w, KB) \in \mathbb{R}$
- ▶ erwartete Utility einer „eigenen“ Substitution  $\vartheta_s$ :

$$E[u(\vartheta_s, F, w, KB)] = \sum_{\vartheta_p} P(\vartheta_p | \vartheta_s, F, w) u(\text{post}(T(F), w)\vartheta_f(F, w)\vartheta_s\vartheta_p, KB)$$

## Optimierung innerhalb eines Rahmens

- ▶ Substitution des anderen Agenten wird mithilfe früherer Instanzen von  $F$  geschätzt:

$$P(\vartheta_p | \vartheta_s, F, w) = \frac{P(\vartheta_p \wedge \vartheta_s | F, w)}{P(\vartheta_s | F, w)} = \frac{P(\vartheta_f(F, w) \vartheta_s \vartheta_p | F, w)}{\sum_{\vartheta} P(\vartheta_f(F, w) \vartheta_s \vartheta | F, w)}$$

## Optimierung innerhalb eines Rahmens

- ▶ Substitution des anderen Agenten wird mithilfe früherer Instanzen von  $F$  geschätzt:

$$P(\vartheta_p | \vartheta_s, F, w) = \frac{P(\vartheta_p \wedge \vartheta_s | F, w)}{P(\vartheta_s | F, w)} = \frac{P(\vartheta_f(F, w) \vartheta_s \vartheta_p | F, w)}{\sum_{\vartheta} P(\vartheta_f(F, w) \vartheta_s \vartheta | F, w)}$$

- ▶ Auswahl optimaler Substitution/Aktion:

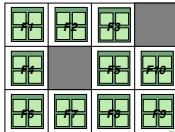
$$\vartheta^*(F, w, KB) = \arg \max_{\vartheta_s \in \Theta_{\text{poss}}(F, KB, w)} E[u(\vartheta_s, F, w, KB)]$$

$$m^*(F, w, KB) = T(F)[|w| + 1] \vartheta^*(F, w, KB)$$

# m<sup>2</sup>infra als hierarchisches RL

*framing decisions + long-term payoffs*

= *framing utility*

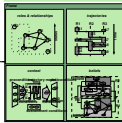


0.812	0.868	0.918	
0.762		0.611	0.534
0.705	0.655	0.611	0.388

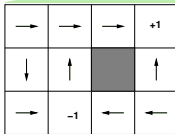
*framing*

**frame level**

**action level**



*in-frame action decisions + immediate payoffs* = *action utility*



0.455	0.686	0.874	+1
0.512	0.112		0.766
0.377	-1	0.245	0.621

## Weitere Features von $m^2$ infra

- ▶ entropiebasiertes Desirability-Kriterium
  - ▶ Balance zwischen Utility-Maximierung und Handlungssicherheit
  - ▶ Konzept der **vorübergehenden sozialen Optimalität**

## Weitere Features von $m^2$ infra

- ▶ entropiebasiertes Desirability-Kriterium
  - ▶ Balance zwischen Utility-Maximierung und Handlungssicherheit
  - ▶ Konzept der **vorübergehenden sozialen Optimalität**
- ▶ Rahmen-Konstruktion durch Kombination existierender Rahmen (Konkatenation im Sinne des Planens)

## Weitere Features von $m^2$ infra

- ▶ entropiebasiertes Desirability-Kriterium
  - ▶ Balance zwischen Utility-Maximierung und Handlungssicherheit
  - ▶ Konzept der **vorübergehenden sozialen Optimalität**
- ▶ Rahmen-Konstruktion durch Kombination existierender Rahmen (Konkatenation im Sinne des Planens)
- ▶ Rahmenmanagement: langfristige Rahmenverschmelzung nach Clustering-Kriterien
  - ▶ Bestimmung „konkretester Verallgemeinerung“ zweier Rahmen
  - ▶ Heranziehen von Isolations- und Kompatibilitätskriterien

## Weitere Features von $m^2$ infra

- ▶ entropiebasiertes Desirability-Kriterium
  - ▶ Balance zwischen Utility-Maximierung und Handlungssicherheit
  - ▶ Konzept der **vorübergehenden sozialen Optimalität**
- ▶ Rahmen-Konstruktion durch Kombination existierender Rahmen (Konkatenation im Sinne des Planens)
- ▶ Rahmenmanagement: langfristige Rahmenverschmelzung nach Clustering-Kriterien
  - ▶ Bestimmung „konkretester Verallgemeinerung“ zweier Rahmen
  - ▶ Heranziehen von Isolations- und Kompatibilitätskriterien
- ▶ Modellierung von Dialogzuständen: themenbasierter Ansatz und Generalisierung



# Gliederung

Einführung

Die Agentenarchitektur InFFrA

Formale Semantik & Lernalgorithmen

**Experimentelle Befunde**

Zusammenfassung & Ausblick

## Verhandlungsmechanismen

- ▶ vorschlags- & angebotsbasierte Methoden (CNP, Auktionen, etc.)

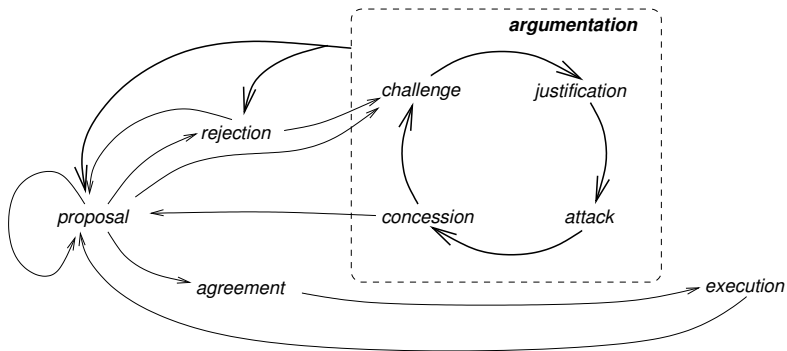
## Verhandlungsmechanismen

- ▶ vorschlags- & angebotsbasierte Methoden (CNP, Auktionen, etc.)
- ▶ argumentationsbasierte Ansätze
  - ▶ erlauben Dialog über *Gründe* für Standpunkte
  - ▶ ermöglichen damit ggs. Beeinflussung des internen Zustandes

## Verhandlungsmechanismen

- ▶ vorschlags- & angebotsbasierte Methoden (CNP, Auktionen, etc.)
- ▶ argumentationsbasierte Ansätze
  - ▶ erlauben Dialog über *Gründe* für Standpunkte
  - ▶ ermöglichen damit ggs. Beeinflussung des internen Zustandes
- ▶ **zielbasierte Verhandlung** (interest-based negotiation – IBN) (Rahwan et al.)
  - ▶ Gewinnung von Informationen über Zielstruktur und Wissen des Anderen
  - ▶ basiert auf Zyklus „Rechtfertigung-Angriff-Zugeständnis“
  - ▶ Modellierung der Zielstruktur über sog. Zielnetze

## Verandlungsmodell in IBN



# Verlinkungsszenario im WWW

LIESON v3.0 - Link Exchange Simulation (c) M. Rovatsos (TUM-AI/Cognition Group), 2001-03

g0:agent1 > modifyRating@g0:agent1,g0:agent9,1 [R1] Time: 00:01:17 Messages: 82

g0:agent2 ← 2 g0:agent7  
 g0:agent7 → 0 g0:agent1  
 g0:agent1 → 1 g0:agent2  
 g0:agent1 → 2 g0:agent5  
 g0:agent5 → -1 g0:agent4  
 g0:agent4 → 1 g0:agent0  
 g0:agent4 → -1 g0:agent9  
 g0:agent0 → 0 g0:agent8  
 g0:agent9 → 2 g0:agent6  
 g0:agent6 → 1 g0:agent3

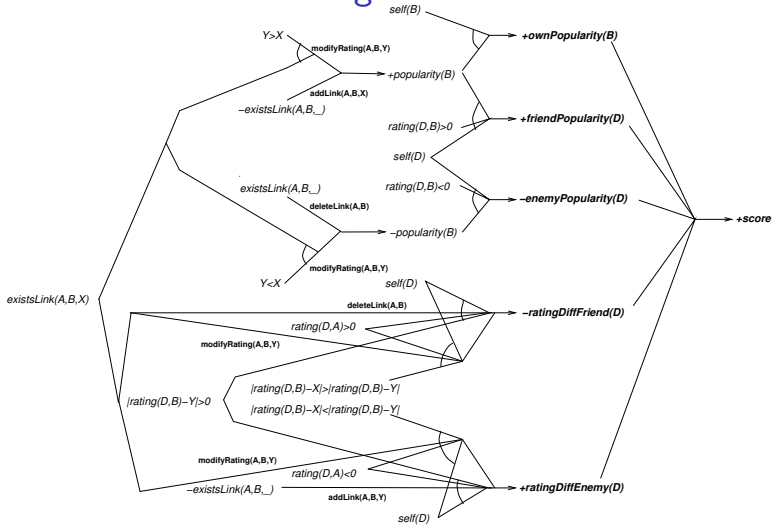
Agent popularity		
Agent	Popularity	Simple Popularity
g0:agent6	0.6111	0.2777
g0:agent5	0.1	0.1
g0:agent4	0.5908	0.1777
g0:agent3	0.5407	0.2444
g0:agent2	0.3116	0.1222
g0:agent1	0.1	0.1
g0:agent0	0.5987	0.2333
g0:agent9	0.5765	0.3555
g0:agent8	0.5101	0.1777
g0:agent7	0.6024	0.2888

```

InFFrA controller of "g0:agent1"
InFFrA messages
(00:00:44): <END ENCOUNTER>
(00:00:50): <START ENCOUNTER>
(00:00:50): initiating message request(g0:agent1,g0:agent0,modifyRating(g0:agent0,g0:agent1,0))
(00:00:50): g0:agent1 > perceived frame updated to
<frame#15/usg [1]/step:0/bind:0>
t[request(g0:agent1,g0:agent0,modifyRating(g0:agent0,g0:agent1,0))]
b[[]]
c[[]]
(00:00:51): received (as expected) accept(g0:agent0,g0:agent1,modifyRating(g0:agent0,g0:agent1,0))
(00:00:52): received accept(g0:agent0,g0:agent1,modifyRating(g0:agent0,g0:agent1,0))
(00:00:52): own turn initiated
(00:00:53): selected frame
<frame#16/usg [3, 3, 2, 2]/step:1/bind:0>
t[request(g0:agent0,g0:agent1,addLink(g0:agent1,g0:agent0,0))]
b[[]]
c[[]]
InFFrA repository
<frame#1/usg [8, 8, 8, 8]/step: -1/bind:0>
t[request(V0,V1,V2), accept(V1,V0,V2), confirm(V0,V1,V2), do(V1,V2)]
b[[]] [V0, g0:agent1, [V1, g0:agent0, [V2, addLink(g0:agent0,g0:agent1,-2)]] [V0, g0:agent1] [V1, g0:agent0]
c[[]] [other(V0)@3, number(-2)@3, other(V1)@3, can(V1,addLink(V1,V0,-2))@3, existsLink(V1,V0,-2)@4] [other(V0)@3]
<frame#29/usg [2]/step: -1/bind:0>
t[request(g0:agent0,g0:agent1,addLink(g0:agent1,g0:agent0,0))]
b[[]]
c[[]]
<frame#41/usg [1]/step: -1/bind:0>
t[request(g0:agent0,g0:agent1,addLink(g0:agent1,g0:agent0,-2))]
b[[]]
c[[]]
    
```

Debug Level 0 Start Stop Reset Redraw Script Exit

# Zielstruktur von LIESON-Agenten



# Gliederung

Einführung

Die Agentenarchitektur InFFrA

Formale Semantik & Lernalgorithmen

Experimentelle Befunde

Zusammenfassung & Ausblick