

Bedeutung und Repräsentation: Fragebeantwortung

Wintersemester 2004/05

Christopher Habel

FB Informatik &

FB Sprach-, Literatur- & Medienwissenschaft

Universität Hamburg

1

Fragen & Antworten – Seminar WS 2004/05

Fragen & Antworten: 1. Sitzung

Themenschwerpunkte des Seminars

- Semantik & Pragmatik von Fragen
- Frage-Antwort-Dialoge
- Maschinelle Beantwortung von Fragen

Ziele des Seminars

- Erkenntnisse über Bedeutung von Fragen und die Bedeutungsrelationen zwischen Fragen und Antworten
- Nutzbarmachung dieser Erkenntnisse für das Design und die Realisierung von Frage-Antwort-Systemen
 - Beantwortung von komplexen Fragen über komplexen Wissensstrukturen
 - kooperative Fragebeantwortung
 - Frage-Antwort-Dia "loge" in ausersprachlicher Kommunikation

Fragen & Antworten – Seminar WS 2004/05

3

Fragen – Fragen & Antworten – Fragebeantwortung

Fragen

- Nicht-deklarativer Typ von Sätzen / Äußerungen
- Grammatikalisch markiert:
 - syntaktisch, morphologisch oder phonologisch

Fragen & Antworten

- Fragen und Antworten
 - stehen in Bedeutungszusammenhängen
 - betreffen unterschiedliche Rollen in Dialogen (nicht-monologischen Diskursen)

Fragebeantwortung (question answering)

- betrifft kognitive Fähigkeit
- Voraussetzung für erfolgreiche Teilnahme am F&A-Dialog

2

Semantik von Fragen

- Bedeutungstheorien behandeln überwiegend Deklarativsätze
 - insbesondere sind Fragen, Imperative und Aufforderungen meist nicht Gegenstand der Untersuchungen (→ Satzmodi)
- Formale Semantik ist meist an der Prädikatenlogik (oder einer Erweiterung) orientiert.
 - Wahrheitwertsemantik (→ für Fragen muss ein alternativer Ansatz entwickelt werden)
- Bedeutungstheorien behandeln überwiegend „monologische Texte“
 - Die Bedeutungsrelationen zwischen Fragen & Antworten sind dialogischer Natur.

Fragen & Antworten – Seminar WS 2004/05

4

Fragen & Antworten

Linguistische Klassifizierung von Fragen

Klassifizierung aus der Syntax der Fragen

- Entscheidungsfragen (*yes/no questions*)
Geht Flug LH112 von Terminal_1 ?
- Ergänzungsfragen (*wh-questions*)
An welchem Terminal wird Flug LH112 abgefertigt ?
Wann ist boarding time für Flug LH112 ?

aber auch:

- *Wie komme ich zu Terminal_1 ?*
Warum verzögert sich der Abflug von Flug LH112 ?
- Klassifizierung über die Antworten
(genauer: angemessene Antworten)

Fragen & Antworten

Klassifizierung von Antworten (1)

Entscheidungsfragen (*yes/no questions*)

- Geht Flug LH112 von Terminal_1 ?
- Flug LH112 geht von Terminal_1. ≈ JA!
Flug LH112 geht nicht von Terminal_1. ≈ NEIN!
- Menge möglicher Antworten entspricht der Menge von Wahrheitswerten.

Ergänzungsfragen (*wh-questions*)

- An welchem Terminal wird Flug LH112 abgefertigt wird?
- Flug LH112 wird am Terminal_1 abgefertigt. ≈ am Terminal_1!
- Menge möglicher Antworten entspricht einer Menge von durch die Frage spezifizierten Entitäten/Objekten.

(Diskurs-)Syntax von Fragen & Antworten

Fragen

meist kompletter Satz

- Geht Flug LH112 von Terminal_1 ?
- Kann mir einer von Ihnen sagen, ob Flug LH112 von Terminal_1 abgeht ?
- Ich wüsste gern, ob Flug LH112 von Terminal_1 abgeht .

Antworten

oft nur Phrasen oder Partikel, insbesondere in elliptischen Konstruktionen

- Geht Flug LH112 von Terminal_1 ?
JA!
- An welchem Terminal wird Flug LH112 abgefertigt?
am Terminal_1! / Terminal_1! / Terminal_1 ab 12h!

Fragen & Antworten

Klassifizierung von Antworten (2)

temporale Fragen (als Spezialfall)

- Wann ist boarding time für Flug LH112 nach München?
- Um 12:15.
- Menge möglicher Antworten ist eine Menge von Zeitentitäten.
- Eine halbe Stunde, nachdem LH111 aus München eingetroffen ist.
- Antworten können auch durch zeitliche Beziehungen zu Ereignissen spezifiziert sein.

Fragen & Antworten

Eine Frage-Antwort-Logik zu $\mathcal{P}L1$

Assertorische Ausdrücke $\mathcal{A}L_{\mathcal{P}L1}$

Geschlossene Formeln über $\mathcal{P}L1$

Interrogative Ausdrücke $IL_{\mathcal{P}L1}$

Erweiterung der Syntax von $\mathcal{P}L1$:

- Falls $\alpha \in \mathcal{A}L_{\mathcal{P}L1}$, dann $\langle ?, \alpha \rangle \in IL_{\mathcal{P}L1}$
ANTW($\langle ?, \alpha \rangle$) = $\{\alpha, \neg\alpha, \alpha \vee \neg\alpha\}$
- Falls α offene Formeln aus $\mathcal{A}L$, und x die freie Variable (bzw. Liste aller freien Variablen), dann $\langle ?x, \alpha \rangle \in IL_{\mathcal{P}L1}$
ANTW($\langle ?x, \alpha \rangle$) = $\{\alpha[x/t] \mid KB \vdash \alpha[x/t]\}$

Belnap, N. D. & Steel, T. B. (1976). *The logic of questions and answers*. New Haven: Yale Univ. Press.

Fragen & Antworten

Komplexe Fragen

Wonach wird eigentlich gefragt bei:

- *Wie komme ich zu Terminal_1 ?*
- *Warum verzögert sich der Abflug von Flug LH112 ?*
- Nach Plänen, Handlungssequenzen,...
- Nach Gründen, Erklärungen,...

Zwischenfazit

Antworten auf *wie-* bzw. *warum-Fragen*

- betreffen gegebenenfalls Strukturen (Sequenzen) von einfachen Antworten
- unterscheiden sich darin, wie sie verwendet werden

Maschinelle Fragebeantwortung:

Question-Answering Systems

Fragen an Datenbanken / Wissenbasen

- natürliche Sprache als Interface zu Datenbanken
- Semantik & Pragmatik der Fragebeantwortung ist auch für formale Anfragesprachen zu spezifizieren!

Fragen an Informationssysteme: Text- & Dokumentenretrieval

- Internet-Suchmaschinen suchen insbesondere in Textbeständen
- Fragebeantwortung im WWW

Fragen in nicht-sprachlicher Kommunikation

- Formular-Rahmen
- Graphische Frageformulierung

Struktur des Seminars:

(1) Einleitung (2) Maschinelle Fragebeantwortung 1

•25.10 & •1.11.	Semantik und Pragmatik von Fragen und Antworten Wunderlich, Dieter (1976). <i>Fragesätze und Fragen</i> . In D. Wunderlich, <i>Studien zur Sprechakttheorie</i> . (pp. 181–250). Suhrkamp: Frankfurt am Main.
•8.11.	Fragebeantwortung durch Maschinen: Phase 1 (1960 – 1990) Waltz, D. L. (1978). An English language question answering system for a large relational database. <i>Communications of the ACM</i> , 21. 526–539. Kaplan, S. Jerrold (1984). Designing a portable natural language database query system. <i>ACM Transactions on Database Systems (TODS)</i> , 9. 1–19. Copestake, A. & Spark Jones, K. (1990). Natural language Interfaces to Databases. <i>The Knowledge Engineering Review</i> , 5. 225–249.

Struktur des Seminars:

(3) Semantik und Pragmatik von Fragen und Antworten (4) Maschinelle Fragebeantwortung 2

•15.11.	Logik von Fragen und Antworten
•&	Beinap, N. D. & Steel, T. B. (1976). <i>The logic of questions and answers</i> . New Haven: Yale Univ. Press
•22.11.	
•29.11.	Semantik und Pragmatik von Fragen und Antworten
•–	Groenendijk, J. & Stokhof, M. (1997). Questions. In J. van Benthem & A. ter Meulen (eds), <i>Handbook of logic and language</i> , (pp. 1055-1124). Amsterdam/Cambridge, MA: Elsevier/MIT Press.
•13.12.	
•3.1.	Fragebeantwortung durch Maschinen: Phase 2 (nach 1990) Hirschman, Lynette & Gaizauskas, Rob (2001). Natural Language Question Answering: The View from Here. <i>Natural Language Engineering</i> , 7. 275–300

Fragen & Antworten – Seminar WS 2004/05

13

Struktur des Seminars:

(5) Temporale Fragen

•10.1	Temporale Fragen (1): Question Answering Systems
•	Pustejovsky, J., Castaño, J., Ingria, R., Sauri, R., Gaizauskas, R., Setzer, A. & Katz, G. (2003). TimeML: Robust specification of event and temporal expressions in text. In <i>Proceedings of the Fifth International Workshop on Computational Semantics (IWCS-5)</i> . Tilburg, Netherlands.
	Schilder, Frank & Habel, Christopher (2003). Temporal information extraction for temporal question answering. In <i>New Directions in Question Answering. Papers from the 2003 AAAI Spring Symposium TR SS-03-07, (Stanford, CA. March 2003)</i> . 34–44.
	Saquete, E.; Martínez-Barco, P.; Muñoz, R. & Vicedo, J.L. (2004). Splitting complex temporal questions for question answering systems. <i>Proceedings of the 42nd Annual Meeting of the Association for Computational Linguistics: ACL 2004, Barcelona, Spain</i> . 566–573.

Fragen & Antworten – Seminar WS 2004/05

14

Struktur des Seminars:

(5) Temporale Fragen (Forts.)

•17.1	Temporale Fragen (2): Linguistische Semantik
•	Nelken, R. & Francez, N. (1999). The semantics of temporal questions. <i>Proceedings of Formal Grammar 1999</i> . Ed. by. Geert-Jan M. Kruij and Richard T. Oehrlé.
•24.1	Temporale Fragen (3): Datenbanken
•	Androutsopoulos, I.; Ritchie, G.D. & Thanisch, P. (1998). Time, tense and aspect in natural language database interfaces. <i>Natural Language Engineering</i> , 4. 229–276.
•31.1	Abschlussdiskussion
•	

Fragen & Antworten – Seminar WS 2004/05

15

Scheine und Leistungsnachweise

Seminarscheine für die Fächer

- Informatik
- Germanistische Linguistik
- Allgemeine Sprachwissenschaft

Anforderungen

- Regelmässige, aktive Teilnahme am Seminar
- Vortrag
- gegebenenfalls Ausarbeitung / Hausarbeit

Fragen & Antworten – Seminar WS 2004/05

16