

Natürlichsprachliche Interaktion beim Aufgabenlösen in Lernumgebungen

Rainer Lütticke

Bergische Universität Wuppertal, Forschungszentrum für Mikrostrukturtechnik – fmt
FernUniversität in Hagen, Intelligente Informations- und Kommunikationssysteme

Zusammenfassung

Für einen speziellen Aufgabentyp in Lernumgebungen, nämlich die natürlichsprachliche Beantwortung einer Problemstellung, haben wir einen neuen Ansatz für eine Analysetechnik entwickelt. Bei dieser Technik handelt es sich um ein tiefensemantisches Verfahren, das auf dem Einsatz semantischer Netze beruht und das MulitNet-Paradigma verwendet. Dadurch gelangt das System zu einem großen Sprachverstehen, so dass detailliertere natürlichsprachliche Rückmeldungen bzgl. Bewertung und Fehlerhinweisen gegeben werden können als in Systemen, die statistische oder „flache“ semantische Verfahren einsetzen. Als erstes Einsatzszenario dieses neuen Verfahrens wird die Analyse natürlichsprachlicher Reformulierungen semantischer Netze beschrieben.

1 Methoden der automatischen Textanalyse

Es ist seit langem bekannt, dass sich durch eigene Aktivitäten von Lernenden mentale Operationen initiieren lassen, die zu besseren Lernerfolgen führen. Ziel einer Lernumgebung sollte es daher sein, die Lernenden zu eigenen Handlungen zu motivieren. Um eine hohe Akzeptanz bei den Nutzern zu erreichen, müssen Interaktionen mit dem Lernsystem auch wirklich hilfreich und einfach durchzuführen sein. Besonders geeignet sind Interaktionen, die natürlichsprachige Kommunikation einschließen, da so einerseits Lernende direkt angesprochen werden und andererseits Lernende auf natürliche und sehr einfache Weise das System ansprechen können.

Für eine Analyse natürlichsprachlicher Eingaben eines Nutzers sind verschiedene Techniken einsetzbar (s.a. Valenti et al. 2003): Abgleich von Mustern, Analyse der Satzstruktur und des Inhalts auf Basis von vorab bewerteten Texten und Trainingsalgorithmen, latente semantische Analyse (LSA), Verwendung flacher semantischer Strukturen zur Bedeutungsdarstellung, die mit einfachen logischen Prädikaten und elementaren semantischen Rollen arbeiten, und Systeme mit großem Sprachverstehen, die auf einer ausgebauten semantischen Theorie basieren. Die ersten drei der genannten Analysetechniken nutzen vorwiegend Prozesse der automatischen Sprachverarbeitung, die auf statistischen Verfahren beruhen. Sie zeichnen

sich zwar durch Robustheit aus, stützen sich aber nicht auf eine semantische Struktur der kommunizierten Spracheinheiten. Die vierte genannte Technik ist Feature-Struktur-orientiert und bezieht bereits semantische Elemente im linguistischen Sinne ein, beruht aber noch nicht auf einem wirklichen Sprachverstehen und bleibt zu stark der Wortebene verhaftet. Diesen vier Verfahren mangelt es an Mitteln zur strukturierten Darstellung der Semantik eines Satzes, so dass sie kaum entwicklungsfähig sind. Hinsichtlich der Tiefe der Einbeziehung semantischer Strukturen werden sie auch als „flache“ Verfahren bezeichnet. Um ein großes Sprachverstehen von Seiten des Systems zu erreichen, müssen sogenannte tiefensemantische Verfahren, die zuletzt genannte Technik einsetzen, verwendet werden. Ihnen muss eine vollausgearbeitete Semantik-Theorie zu Grunde liegen und für sie müssen syntaktisch-semantische Analyseverfahren sowie entsprechende Software-Werkzeuge zur Verfügung stehen. Daher haben wir damit begonnen ein solches „tiefes“ Verfahren, das das MultiNet-Paradigma¹ (Helbig 2001) verwendet und für das eine automatische Wortklassen-gesteuerte funktionale Analyse für das Deutsche existiert (Hartrumpf 2003), in der Lernumgebung VILAB² (Lütticke et al. 2005) für die Zwecke der intelligenten Textanalyse einzusetzen

2 Analyse von textuellen Antworten in VILAB

Die Möglichkeit der freien Antwort, die dann automatisch bewertet wird, ist sehr gut dazu geeignet, das Wissen und Verständnis von Lernenden zu vertiefen (Valenti et al. 2003). Daher haben wir als erstes Einsatzszenario der natürlichsprachlichen Interaktion in VILAB mit Hilfe von Techniken, die auf MultiNet beruhen, die Analyse von textuellen Antworten von Lernenden auf spezielle Problemstellungen gewählt. Bei diesen Problemstellungen handelt es sich um die Umschreibung eines formalsprachlichen Ausdrucks, nämlich die Reformulierung semantischer Netze, die auf dem MultiNet-Paradigma beruhen, in natürliche Sprache. Der Lernende wird dabei im Rahmen einer Aufgabenstellung in VILAB dazu aufgefordert, seine Antwort in ein spezielles Textfeld zu schreiben. Nachdem dann die Antwort mit Hilfe des Parsers automatisch in MultiNet transformiert wurde, wird ein Modul zur Analyse seiner Antwort aufgerufen. Dieses Modul vergleicht durch logische Inferenzen die in MultiNet transformierte Antwort des Lerners mit der Musterantwort. Übereinstimmungen oder partielle Abweichungen können so abgeleitet werden. Daraus werden dann konkrete Hinweise für Verbesserungen in adäquater natürlichsprachlicher Form mit Hilfe von Transformationsregeln generiert, die schließlich als Feedback an den Lerner kommuniziert werden. Daraufhin kann der Lerner seine Antwort verbessern und einer erneuten Analyse unterziehen. Dieser

¹ Der Kern der MultiNet-Repräsentation ist ein semantisches Netz, das formal einen gerichteten Graph darstellt. Knoten des Graphs repräsentieren bestimmte Entitäten des Diskursgebietes und Kanten drücken semantische Relationen zwischen den Knoten aus. Jeder mögliche Begriff der realen Welt ist wählbar als Netzknoten. Innere Knoten repräsentieren komplexe Begriffe. Es gibt einen festgelegten Satz von ungefähr 110 Relationen, die die Kanten markieren. Jede Relation hat ihre vordefinierte Bedeutung (z.B. (AGT e a): a ist Handelnder in einem Geschehen e). Weitere kompliziertere MultiNet-Konzepte sind verschiedene Attribute der Knoten und Kanten.

² VILAB: Virtuelles Informatik-Labor der FernUniversität in Hagen. In dieser Internet-basierten Lernumgebung werden Studierenden Aufgaben mit verschiedenen Inhalten aus der Informatik angeboten. Die Studierenden können mit komplexen Software-Werkzeugen experimentieren und diese zur Lösung von Aufgaben einsetzen. Während der Aufgabenbearbeitung und der Lösungsfindung werden die Studierenden durch eine interaktive tutorielle Komponente unterstützt. Homepage des Labors und Guided Tour: <http://pi7.fernuni-hagen.de/vilab/>

Interaktionszyklus kann so lange durchgeführt werden, bis die richtige Antwort gefunden wird. Während der Aufgabenbearbeitung werden Daten (Anzahl der Versuche, Abstand zur Musterlösung, Verringerung des Abstandes zur Musterlösung) im Nutzermodell gespeichert und fließen in das Feedback ein (Lütticke et al. 2005). Die Kommunikationsstrategie ist dabei, den Lernenden gerade so viele Hinweise zu geben, dass eine bei der Aufgabenlösung aufgetretene Schwierigkeit behoben werden kann, aber nicht so viele, dass die Lösung trivial wird. Im Gegensatz zu flachen Verfahren sind dadurch detaillierte und didaktisch aussagekräftige Rückmeldungen aus der Fehleranalyse studentischer Antworten möglich (Beispiel in Lütticke 2006).

Während einzelne Teile des Analysemechanismus schon seit einigen Jahren in der Lehre eingesetzt werden, soll das komplette Modul zur Antwortanalyse erstmals im Rahmen eines Online-Praktikums Ende 2006 eingesetzt werden. In nachfolgenden Schritten soll dann an den bekannten sowie durch Evaluation festgestellten Nachteilen und Begrenzungen des Verfahrens gearbeitet werden (z.B. größere Robustheit bei der Transformation, Transformation mehrerer zusammenhängender Sätze, Einbeziehung mehrerer Musterantworten, linguistische Erweiterung der Muster- und Lernerantwort) (s.a. Lütticke 2006).

Literaturverzeichnis

- Hartrumpf, S. (2003): Hybrid Disambiguation in Natural Language Analysis. Osnabrück: Der Andere Verlag.
- Helbig, H. (2001): Die semantische Struktur natürlicher Sprache: Wissensrepräsentation mit MultiNet. Springer.
- Lütticke, R. (2006): Using Semantic Networks for Assessment of Learners' Answers. In: Koper, R.; Kinshuk (Hrsg.): The 6th IEEE International Conference on Advanced Learning Technologies (ICALT) – Advanced Technologies for Life-Long Learning, in press.
- Lütticke, R.; Helbig, H.; Eichhorn, C. (2005): Das Virtuelle Informatik-Labor VILAB – Konzeption, technische Realisierung und Einsatz in der Lehre, Vol. 321, Informatik-Berichte, FernUniversität in Hagen.
- Valenti, S.; Neri, F.; Cucchiarelli, A. (2003): An overview of current research on automated essay grading. Journal of Information Technology Education, Vol. 2, S. 319-330.

Kontaktinformationen

Dr. Rainer Lütticke, rainer.luetticke@fernuni-hagen.de