

Kontrollierte Sprache für juristische Informationssysteme

Stefan Höfler, Alexandra Bünzli, Michael Hess

Universität Zürich, Institut für Computerlinguistik
Binzmühlestrasse 14, 8050 Zürich, Schweiz

hoefler@cl.uzh.ch, buenzli@cl.uzh.ch, hess@cl.uzh.ch

Wissensbasierte juristische Informationssysteme haben zum Ziel, das automatische Ziehen von Schlussfolgerungen über Gesetzestexten und anderen Rechtstexten zu ermöglichen. Zu diesem Zweck muss einerseits juristisches Schlussfolgern formalisiert werden, und andererseits muss das juristische Wissen, über dem geschlussfolgert werden soll, in einer formalen logischen Repräsentation verfügbar sein. Während im letzten Jahrzehnt Fortschritte in diesem Bereich erzielt werden konnten (RISSLAND et al. 2003), bleibt ein Haupthindernis für die Entwicklung juristischer Informationssysteme bestehen: «Since the raw materials of the law are embodied in natural language – cases, statutes, regulations, etc. – the designer of a knowledge-based legal information system today must translate them, by hand, into a formal language, just to get started» (MCCARTY 2007:217). Eine manuelle Übersetzung von Gesetzestexten in eine formale Repräsentationsform ist nicht nur ausgesprochen zeitaufwändig, sondern auch fehleranfällig. Dies lässt sich insbesondere auch darauf zurückführen, dass hier zwei sehr unterschiedliche Fachrichtungen aufeinandertreffen: Entwickler von Informationssystemen verfügen üblicherweise nicht über juristisches Fachwissen, Juristen hingegen sind in formalen Methoden nicht genügend ausgebildet, um die Korrektheit von formalen Darstellungen von Gesetzestexten beurteilen zu können. Die Alternative zur manuellen Übersetzung, die maschinelle Übersetzung, ist momentan noch zu wenig präzise, um zuverlässiges automatisches Schlussfolgern über Gesetzestexten zu erlauben: Besonders die auch in Gesetzestexten weit verbreitete (wenn auch nicht unbedingt intendierte) natürlichsprachliche Mehrdeutigkeit stellt ein Problem dar.

Das Projekt Collegis, eine Kooperation von Computerlinguisten der Universität Zürich und Gesetzesredaktoren der Schweizerischen Bundeskanzlei, nimmt sich dieses Problems an, indem es eine kontrollierte Version der schweizerischen Rechtssprache entwickelt. Kontrollierte Sprachen schränken das Lexikon, die Syntax und/oder die Semantik einer natürlichen Sprache ein, um die Komplexität dieser Sprache zu verringern und das Vorkommen von Mehrdeutigkeiten zu reduzieren. Diese Methode wurde ursprünglich eingesetzt, um die Verständlichkeit und Übersetzbarkeit von Texten zu verbessern; neuerdings wird sie auch dazu verwendet, natürlichsprachliche Schnittstellen zur formalen Logik zu definieren (POOL 2006; FUCHS 2009).

Die Methode der kontrollierten natürlichen Sprache wird nun im Projekt Collegis auf die Domäne der Gesetzestexte angewendet (HÖFLER und BÜNZLI 2010). Die im Projekt entwickelte kontrollierte Sprache *Controlled Legal German* (CLG) soll es erlauben, den Inhalt von Gesetzestexten in einer Form darzustellen, die für den menschlichen Leser ohne Hintergrundwissen in formalen Methoden verständlich ist, aber zugleich über eine eindeutige formale Semantik verfügt und damit automatisch in eine formal-logische Repräsentation übersetzt werden kann. Die (manuelle) Übertragung von Gesetzestexten in CLG wird dadurch erleichtert, dass CLG der konventionellen Gesetzessprache ähnlich ist. Bei der Entwicklung von CLG wird dabei die Tatsache ausgenützt, dass Gesetzessprache selber schon stark konventionalisiert ist – dies unter anderem gerade auch im Hinblick darauf, Mehrdeutigkeiten zu vermeiden.

In unserem Vortrag stellen wir die zentralen syntaktischen und semantischen Eigenschaften von Controlled Legal German vor. Wir erklären, welche Design-Entscheidungen bei der Entwicklung von CLG getroffen werden müssen, um die diversen Formen von natürlichsprachlicher Mehrdeutigkeit zu verhindern, und wie bestehende Konventionen der schweizerischen Gesetzessprache zu diesem Zweck genutzt werden können. Abschliessend zeigen wir einige relevante Anwendungsmöglichkeiten auf.

Literatur

FUCHS, N.E., Hrsg. (2009). *Pre-Proceedings of the Workshop on Controlled Natural Language (CNL 2009)*. CEUR-WS.

HÖFLER, STEFAN und A. BÜNZLI (2010). *Controlling the Language of Statutes and Regulations for Semantic Processing*. In: *Proceedings of the LREC Workshop on the Semantic Processing of Legal Texts*. Im Erscheinen.

MCCARTY, L. T. (2007). *Deep Semantic Interpretation of Legal Texts*. In: *Proceedings of the 11th International Conference on Artificial Intelligence and Law*, S. 217–224, New York. ACM Press.

POOL, J. (2006). *Can Controlled Languages Scale to the Web?*. In: *5th International Workshop on Controlled Language Applications*.

RISSLAND, E. L., K. D. ASHLEY und R. P. LOUI (2003). *AI and Law: A Fruitful Synergy*. *Artificial Intelligence*, 150(1–2):1–15.