

## Collegis

Alexandra Bünzli, Stefan Höfler

<http://www.c1.uzh.ch>

16. März 2010



## Allgemeine Ziele

## Hauptziel

Semantische Repräsentation für schweizerische Gesetzestexte ermöglichen

## Nebenziele

Unterstützung beim Erstellen und Redigieren von schweizerischen Gesetzestexten

- Qualität erhöhen



## Kollo 2010 – Collegis-Projekt

- 1 Einleitung
- 2 Unser Ansatz – Kurz erklärt
- 3 Lexikalische Kontrolle
- 4 Syntaktische Kontrolle
- 5 Semantische Kontrolle
- 6 Schlussfolgerungen



## Wieso will man eine semantische Repräsentation?

- Grundlage für diverse Anwendungen
  - Fragebeantwortungs- und Antwortextraktionssysteme
  - Expertensysteme
    - z.B. im Bereich der Interessenvertretung, Entscheidungsfindung, Beratung, Planung
  - Konsistenzprüfung
- ermöglicht neue Darstellungsformen
  - z.B. Flussdiagramme für einfachen Überblick
- praktisch für Übersetzungen



## Wie wollen wir das erreichen?

Grosses, altbekanntes Problem: Ambiguität

### Unsere Lösung

Sprache kontrollieren!

D.h. wir definieren eine neue Sprache:  
CONTROLLED LEGAL GERMAN (CLG)

Eröffnet neue Anwendungen: z.B. didaktische

- in Lehre und professioneller Weiterbildung
  - Bewusstmachen der Schwierigkeiten

## Der Weg dahin wird nicht einfach sein ...



## Wieso gerade Gesetzestexte?

Gesetzestexte sind speziell geeignet, kontrolliert zu werden:

- entstehen bereits in einem kontrollierten Prozess
  - administrativ
  - von der Sache her: Gesetzessprache soll nicht ambig sein
    - jedoch durchaus vage
  - Richtlinien existieren
- Akzeptanz bei Benutzern relativ gross
  - Ersteller sind froh um Unterstützung, fühlen sich teilweise unsicher
  - Sprachendienst des Bundes (Linguisten) sind offen und interessiert
    - Zusammenarbeit funktioniert

## Kontrollierte Sprache

Bis jetzt wurde kontrollierte Sprache v.a. in folgenden Bereichen eingesetzt

- in der technischen Dokumentation
- beim *requirements engineering*
- in der allgemeinen Wissensrepräsentation

Gibt zwei grössere Gruppen

- diejenigen, die für Menschen konzipiert werden
- diejenigen, die «für» Maschinen konzipiert werden

## Welche Eigenschaften soll eine kontrollierte Sprache für Gesetze haben?

- muss genügend ausdrucksstark sein
- nahe an der heutigen, redigierten Gesetzessprache, keine zu grossen Einschränkungen
- relativ einfach zu erlernen
  - wobei den Gesetzesredaktoren mehr zuzumuten ist als dem durchschnittlichen Benutzer, da sie es bereits gewohnt sind, Regeln zu beachten



## Wie funktioniert Kontrolle in CLG?

Verschiedene Ebenen:

- lexikalisch
- syntaktisch
- semantisch

Verschiedene Methoden

- lexikalische Regeln/Einschränkungen
- Konstruktionsregeln
- Interpretationsregeln
- Paraphrasen
- Unterspezifikation
- Interaktion



## Lexikalische Kontrolle in CLG

Definition von Inhaltswörtern wird dem Benutzer überlassen

- externe Terminologie, Ontologie
  - Terminologiedatenbanken existieren
  - Ontologie von Grundlagenbegriffen (in English) ebenfalls

Damit lässt man den Interpretationsspielraum weiterhin offen

- Vagheit und *open textured concepts* sind weiterhin möglich

Gewisse domänenspezifische Wörter werden aber ganz gezielt interpretiert



## Lexikalische Kontrolle: Beispiel

### Beispiel

Die Veröffentlichung der Entscheide hat grundsätzlich in anonymisierter Form zu erfolgen. (Art. 27 Abs. 2 BGG)

*grundsätzlich* hat im Deutschen zwei Bedeutungen:

- strikt
- im Prinzip

In der Domäne der Gesetze ist jedoch immer die zweite Bedeutung die richtige

- nur diese Bedeutung zulassen
- markiert *defeasibility*



## Weitere domänenspezifische Ausdrücke

- *in der Regel*
- *insbesondere*
- *sinngemäß*
- *im Rahmen von*
- etc.



## Syntaktische Ambiguität: Beispiel I

## Beispiel (Art. 25a Abs. 2 BGG)

Das Bundesgericht deckt seinen Bedarf an Gütern und Dienstleistungen im Bereich der Logistik selbständig.

Wo wird im Bereich der Logistik angeschlossen?

- *deckt*
- *Güter und Dienstleistungen*
- *Dienstleistungen*
- *Bedarf*



## Syntaktische Ambiguität: Beispiel II

## Beispiel (Art. 25a Abs. 2 BGG)

Das Bundesgericht deckt seinen Bedarf an Gütern und Dienstleistungen im Bereich der Logistik selbständig.

## Beispiel (Art. 25a Abs. 2 BGG)

Das Bundesgericht deckt seinen Bedarf an Gütern und Dienstleistungen im Bereich der Logistik selbständig.

## Beispiel (Art. 25a Abs. 2 BGG)

Das Bundesgericht deckt seinen Bedarf an Gütern und Dienstleistungen im Bereich der Logistik selbständig.



## Syntaktische Ambiguität: Beispiel III

## Beispiel (Art. 25a Abs. 2 BGG)

Das Bundesgericht deckt seinen Bedarf an Gütern und Dienstleistungen im Bereich der Logistik selbständig.



## Wie syntaktische Ambiguität kontrollieren?

## Interpretationsregel

- Eine PP wird immer an die nächste mögliche Konstituente angehängt.
- Wenn diese Konstituente eine koordinierte Phrase ist, dann wird die PP an die ganze Phrase angehängt und nicht an das letzte Konjunkt.

In unserem Beispiel also an *Güter und Dienstleistungen*

## Beispiel (Art. 25a Abs. 2 BGG)

Das Bundesgericht deckt seinen Bedarf an Gütern und Dienstleistungen  
*im Bereich der Logistik* selbständig.

## Wie erreicht man die anderen Lesarten?

Eine Möglichkeit: **Paraphrasen**

Anbindung an Verb: PP direkt nach Verb oder im Vorfeld

## Beispiel (Art. 25a Abs. 2 BGG)

Das Bundesgericht deckt *im Bereich der Logistik* seinen Bedarf an Gütern und Dienstleistungen selbständig.

## Beispiel (Art. 25a Abs. 2 BGG)

*Im Bereich der Logistik* deckt das Bundesgericht seinen Bedarf an Gütern und Dienstleistungen selbständig.



## Wie erreicht man die anderen Lesarten?

Eine Möglichkeit: **Paraphrasen**

Anbindung an einzelnes Konjunkt: Es darf nicht das letzte Konjunkt sein

## Beispiel (Art. 25a Abs. 2 BGG)

Das Bundesgericht deckt seinen Bedarf an Dienstleistungen  
*im Bereich der Logistik* und (an) Gütern selbständig.

## Paraphrase manchmal nicht möglich

## Beispiel (Art. 25a Abs. 2 BGG)

Das Bundesgericht deckt seinen Bedarf an Gütern und Dienstleistungen  
*im Bereich der Logistik* selbständig.

## Möglichkeiten

- Interaktion
- evtl. bestimmte häufige Wendungen speziell behandeln
  - z.B. *Bedarf an, Mitwirkung an, Aufsicht über ...*



## Paraphrase manchmal nicht möglich

## Beispiel (§ 9. Absatz 3 Universitätsordnung)

Voraussetzung für die Assistenzprofessur ist in der Regel die Habilitation oder eine gleichwertige wissenschaftliche Qualifikation in Forschung und Lehre.

Umstellung nicht möglich:

\*Voraussetzung für die Assistenzprofessur ist in der Regel eine **gleichwertige** wissenschaftliche Qualifikation in Forschung und Lehre oder die Habilitation.

## Paraphrase manchmal nicht möglich

## Lösung durch Ausnützung von Konventionen der Gesetzessprache

## Beispiel (§ 9. Absatz 3 Universitätsordnung)

Voraussetzung für die Assistenzprofessur ist in der Regel

- ① die Habilitation oder
- ② eine gleichwertige wissenschaftliche Qualifikation in Forschung und Lehre.

## Beispiel (§ 9. Absatz 3 Universitätsordnung)

Voraussetzung für die Assistenzprofessur ist in der Regel *a.* die Habilitation oder *b.* eine gleichwertige wissenschaftliche Qualifikation in Forschung und Lehre.

## Übersicht

- ① Einleitung
- ② Unser Ansatz – Kurz erklärt
- ③ Lexikalische Kontrolle
- ④ Syntaktische Kontrolle
- ⑤ **Semantische Kontrolle**
- ⑥ Schlussfolgerungen

## Motivation und Methoden

- **Ziel:** automatische semantische Analyse von Gesetzestexten
- **Problem:** semantische Ambiguität

Konstruktionen, die zwar nur eine syntaktische Analyse haben, aber mehrere semantische Interpretationen zulassen

- **Lösungsansatz:** Gesetzessprache so kontrollieren, dass sie keine semantische Ambiguität enthält
- **Methoden:**
  - *Konstruktionsregeln*  
den Gebrauch der ambigen Konstruktion untersagen
  - *Interpretationsregeln*  
eine Interpretation als Default festlegen, sodass die ambige Konstruktion nur in dieser Bedeutung verwendet werden darf

## Skopusambiguität: "Every man loves a woman."

## Beispiel (Art. 152 Bundesverfassung)

Jeder Rat [...] wählt eine Präsidentin oder einen Präsidenten [...].

$$\textcircled{a} \forall x : (\text{rat}(x) \rightarrow \exists y : (\text{praesidentin}(y) \wedge \text{waehlt}(x, y)))$$



$$\textcircled{b} \exists y : (\text{praesidentin}(y) \wedge \forall x : (\text{rat}(x) \rightarrow \text{waehlt}(x, y)))$$



## Überlegungen

- Interpretation gemäss Oberflächenstruktur i.A. intuitiver
- Das Subjekt bezeichnet Adressaten/Gegenstand der Norm, während die VP den eigentlichen normativen Gehalt enthält.

## Interpretationsregel

Subjekt  $\triangleright$  Objekte/Adverbialien  $\triangleright$  Verb

(Hauptsatz)



## Der Skopus des Modal-Operators

## Beispiel

Alle Radfahrer **müssen** mindestens einen Rückstrahler tragen.

$$\forall x : (\text{radfahrer}(x) \rightarrow \textcircled{a} \exists y : (\text{rueckstr}(y) \wedge \text{traegt}(x, y)))$$

$$\equiv \textcircled{b} \forall x : (\text{radfahrer}(x) \rightarrow \exists y : (\text{rueckstr}(y) \wedge \text{traegt}(x, y)))$$

## Beispiel

Mindestens eine Veranstaltung **muss** allen Personen offen stehen.

$$\exists x : (\text{veranst}(x) \wedge \textcircled{a} \forall y : (\text{pers}(y) \rightarrow \text{steht\_offen}(x, y)))$$



$$\neq \textcircled{b} \exists x : (\text{veranst}(x) \wedge \forall y : (\text{pers}(y) \rightarrow \text{steht\_offen}(x, y)))$$



## Interpretationsregel

Modalverb  $\triangleright$  Subjekt  $\triangleright$  Objekte/Adverbialien  $\triangleright$  Hauptverb

## Interpretation des unbestimmten Artikels

## Beispiel (§ 67 Abs. 2 Universitätsordnung)

Ein Mitglied der Universitätsleitung führt den Vorsitz.

$$\textcircled{a} \exists x : \text{mitglied\_unileitung}(x) \wedge \dots$$

## Beispiel (§ 8 Abs. 7 Universitätsordnung)

Ein Titel [...] kann [...] entzogen werden, wenn die Inhaberin oder der Inhaber die Interessen der Universität ernsthaft verletzt.

$$\textcircled{a} \forall x : \text{titel}(x) \rightarrow \dots$$

## Interpretationsregel

Modalverb  $\triangleright$  Subjekt<sup>v</sup>  $\triangleright$  Objekte/Adverbialien<sup>z</sup>  $\triangleright$  Hauptverb

Superskript: Interpretation indef. NPs



## Umformulierungen

## Beispiel (§ 67 Abs. 2 Universitätsordnung – umformuliert)

Mindestens ein / genau ein Mitglied der Universitätsleitung führt den Vorsitz.



$$\textcircled{a} \exists / \exists! x : (\text{mitglied\_unileitung}(x) \rightarrow \text{fuehrt\_vorsitz}(x))$$



## Beispiel (§ 67 Abs. 2 Universitätsordnung – umformuliert)

Die Forschungskommission wird von **inem Mitglied** der Universitätsleitung präsiert.

$$\exists! x : \text{forschungskommission}(x) \wedge$$

$$\textcircled{a} \exists y : (\text{mitglied\_unileitung}(y) \wedge \text{praesidiert}(y, x))$$



## Plural-Ambiguität

Da man meinen könnte, ein Pass werde von allen Staaten zusammen ausgestellt, habe ich wie folgt geändert:

## Beispiel (Art. 6 Abs. 3 Bst. e VEV)

[Von der Visumpflicht sind ausgenommen:]

e. Inhaberinnen und Inhaber eines [...] Sonderpasses, der von **denden** **einem der** in Absatz 2 genannten Staaten ausgestellt wurde;

Für den Computer ist diese Version vermutlich immer noch missverständlich (alle Inhaber haben zusammen einen Pass?)

## Interpretationsregel

Plurale werden distributiv interpretiert. Für die kollektive Lesart muss ein Singular-Begriff verwendet werden (z.B. *Inhaberschaft*).



## Unterspezifikation

## Beispiel (Art. 55 Abs. 1 AngO ETH-Bereich)

Bei der Geburt eines Kindes hat der Angestellte Anspruch auf eine einmalige Zulage von 530 Franken.

- Pro Jahr werden rund 216'000 Kinder geboren.
- Jedesmal hat der Angestellte Anspruch auf 530 Franken.
- Das macht 114'480'000 Franken pro Jahr und Angestellten.

Wie kann eine kontrollierte Gesetzessprache dieses Problem verhindern?



## Lösungsansatz 1: Konstruktionsregel

## Konstruktionsregel

Das Antezedens einer Bedingung (*bei Geburt eines Kindes*) muss mindestens einen Diskursreferenten mit dem Konsequens (*der Angestellte hat Anspruch auf ...*) gemeinsam haben.

- automatische Überprüfung nach Generierung der logischen Repräsentation
- bei Verletzung: Warnmeldung an den Benutzer

## Beispiel (Art. 55 Abs. 1 AngO ETH-Bereich – ergänzt)

Bei der Geburt eines Kindes, *gegenüber dem er elterliche Pflichten hat*, hat der Angestellte Anspruch auf eine einmalige Zulage von 530 Franken.



## Lösungsansatz 2: Interpretationsregel

*Kind* hat zwei Bedeutungen:

- ein junger Mensch: *kind(x)*
- ein direkter Nachkomme: *kind(x, y)* (relationales Substantiv)

## Interpretationsregel

Ein unterspezifiziertes Argument bezieht sich auf das Subjekt des Hauptsatzes oder das Subjekt des unmittelbar vorangehenden Satzes.

## Beispiel (Art. 55 Abs. 1 AngO ETH-Bereich)

Bei der Geburt eines Kindes hat der Angestellte Anspruch auf eine einmalige Zulage von 530 Franken.



## Übersicht

- 1 Einleitung
- 2 Unser Ansatz – Kurz erklärt
- 3 Lexikalische Kontrolle
- 4 Syntaktische Kontrolle
- 5 Semantische Kontrolle
- 6 Schlussfolgerungen



## (Vorsichtige) Schlussfolgerungen

- 1 Syntaktische und semantische Ambiguitäten lassen sich oft durch **relativ einfache sprachliche Regeln** kontrollieren.  
*Herausforderung:* Formulieren von Regeln, die natürlich wirken, die Expressivität erhalten und leicht anzuwenden sind.
- 2 Die **Konventionen der Gesetzessprache** stellen – zumindest ansatzweise – Methoden zum Kontrollieren bestimmter Ambiguitäten zur Verfügung.
- 3 Manchmal verbessern solche Regeln nicht nur die maschinelle Verarbeitbarkeit eines Gesetzestextes sondern auch seine **Verständlichkeit für den menschlichen Leser**.
- 4 Die Entwicklung und Benutzung einer kontrollierten Gesetzessprache **hilft zu hinterfragen und besser zu verstehen**, was in einer Norm eigentlich genau geregelt ist.



## Ausblick: kurz- und mittelfristige Einsatzmöglichkeiten

- 1 **Gesetzesredaktion: Sensibilisierung**  
Die Auseinandersetzung mit kontr. Sprache sensibilisiert Gesetzesredaktoren für diverse Mehrdeutigkeitsphänomene.
- 2 **Gesetzesredaktion: Empfehlungen**  
Einzelne Regeln können unmittelbar als Empfehlungen für die Gesetzesredaktion übernommen werden.
- 3 **Rechtsetzung: Konzeptualisierung von Normen**  
Neue Normen können in *natürlicher* Sprache *formal* spezifiziert werden. → Implementation diverser Tests (z.B. Konsistenz)
- 4 **Rechtsetzung: Didaktik**  
Das Zusammenspiel zwischen normativer Sprache und dem eigentlichen Inhalt von Normen kann verdeutlicht werden.

