Controlling the Language of Statutes and **Regulations for Semantic Processing**

Stefan Hoefler and Alexandra Bünzli

{hoefler, buenzli}@cl.uzh.ch



Problem

Introduction
•0000

Natural language ambiguity continues to be a major obstacle to a deep semantic processing of statutes and regulations:

- lexical ambiguity
- syntactic ambiguity
- semantic ambiguity



Problem

Introduction

Natural language ambiguity continues to be a major obstacle to a deep semantic processing of statutes and regulations:

- lexical ambiguity
- syntactic ambiguity
- semantic ambiguity

Side notes

- ♠ As opposed to vagueness (open-texturedness), ambiguity is never intended in a legal text but it sometimes appears unavoidable.
- 2 Not each instance of ambiguity that is a problem for semantic processing also poses a problem for human readers.



One approach: controlling the input language

Introduction



One approach: controlling the input language

Introduction

Controlled natural languages (CNLs) restrict
the vocabulary, syntax and/or semantics available to users
in order to reduce natural language ambiguity and complexity.



- Controlled natural languages (CNLs) restrict
 the vocabulary, syntax and/or semantics available to users
 in order to reduce natural language ambiguity and complexity.
- Purposes

Introduction

- Human-oriented CNLs improve the understandability and translatability e.g. of technical texts.
- Machine-oriented CNLs ensure the processability of natural language specifications; serve as an interface to some form of logic



One approach: controlling the input language

- Controlled natural languages (CNLs) restrict
 the vocabulary, syntax and/or semantics available to users
 in order to reduce natural language ambiguity and complexity.
- Purposes

Introduction

- Human-oriented CNLs improve the understandability and translatability e.g. of technical texts.
- Machine-oriented CNLs ensure the processability of natural language specifications; serve as an interface to some form of logic
- Applications
 - 1 technical documentation (manuals)
 - 2 knowledge representation (business rules, clinical guidelines, Semantic Web)



00000

Aim

We apply the method of controlled natural language to legislative drafting to facilitate the semantic processing of statutes and regulations.



We apply the method of controlled natural language to legislative drafting to facilitate the semantic processing of statutes and regulations.

We are developing



Introduction

We apply the method of controlled natural language to legislative drafting to facilitate the semantic processing of statutes and regulations.

We are developing

Controlled Legal German (CLG)

a linguistic standard for Swiss statutes and regulations



Introduction

We apply the method of controlled natural language to legislative drafting to facilitate the semantic processing of statutes and regulations.

We are developing

- a linguistic standard for Swiss statutes and regulations
 - comprised of a set of well-defined conventions



Introduction

We apply the method of controlled natural language to legislative drafting to facilitate the semantic processing of statutes and regulations.

We are developing

- a linguistic **standard** for Swiss statutes and regulations
 - comprised of a set of well-defined conventions
 - that reduce ambiguity in legal language



Introduction

We apply the method of controlled natural language to legislative drafting to facilitate the semantic processing of statutes and regulations.

We are developing

- a linguistic standard for Swiss statutes and regulations
 - comprised of a set of well-defined conventions
 - that reduce ambiguity in legal language
 - and thus facilitate semantic processing.



Introduction

Introduction

 Historically grown domain-specific conventions: some constructions are ambiguous in full natural language but not in legislative language.



Introduction

- Historically grown domain-specific conventions: some constructions are ambiguous in full natural language but not in legislative language.
- Drafting guidelines for professional legal editors: recommend how to avoid certain types of ambiguity in statutes and regulations



- Historically grown domain-specific conventions:
 some constructions are ambiguous in full natural language but not in legislative language.
- Drafting guidelines for professional legal editors: recommend how to avoid certain types of ambiguity in statutes and regulations

Research question

Introduction

Can we **adapt and expand existing conventions** in order to facilitate the semantic processing of statutes and regulations?



Introduction

- 1 Domain characteristics
- 2 Drafting guidelines
- 3 Additional standards
- 4 State of development
- **5** Conclusion



Art. 27 Abs. 2 BGG¹

Die Veröffentlichung der Entscheide hat **grundsätzlich** in anonymisierter Form zu erfolgen.

'In principle, the decisions have to be published in anonymised form.'



Domain characteristics

Art. 27 Abs. 2 BGG¹

Die Veröffentlichung der Entscheide hat grundsätzlich in anonymisierter Form zu erfolgen.

'In principle, the decisions have to be published in anonymised form.'

The adverb grundsätzlich has two (directly opposed) meanings:



Domain characteristics

Art. 27 Abs. 2 BGG¹

Die Veröffentlichung der Entscheide hat grundsätzlich in anonymisierter Form zu erfolgen.

'In principle, the decisions have to be published in anonymised form.'

The adverb grundsätzlich has two (directly opposed) meanings:

1 'strictly', 'categorically', 'always' (no exceptions)



Domain characteristics

Art. 27 Abs. 2 BGG¹

Die Veröffentlichung der Entscheide hat grundsätzlich in anonymisierter Form zu erfolgen.

'In principle, the decisions have to be published in anonymised form.'

The adverb grundsätzlich has two (directly opposed) meanings:

- 1 'strictly', 'categorically', 'always' (no exceptions)
- 2 'generally', 'in principle', 'usually' (exceptions possible)



Domain characteristics

Art. 27 Abs. 2 BGG¹

Die Veröffentlichung der Entscheide hat **grundsätzlich** in anonymisierter Form zu erfolgen.

'In principle, the decisions have to be published in anonymised form.'

The adverb grundsätzlich has two (directly opposed) meanings:

- 'strictly', 'categorically', 'always' (no exceptions)
- 2 'generally', 'in principle', 'usually' (exceptions possible)

Convention

In statutes and regulations, *grundsätzlich* is always used in the latter sense.



§ 3 Abs. 3 UniO UZH²

Dienstleistungen sind [...] kostendeckend in Rechnung zu stellen.

'Services have to be charged so that the costs are covered.'



§ 3 Abs. 3 UniO UZH²

Dienstleistungen sind [...] kostendeckend in Rechnung zu stellen.

'Services have to be charged so that the costs are covered.'

Indefinite noun phrases have an existential and a generic usage:



§ 3 Abs. 3 UniO UZH²

Domain characteristics

Dienstleistungen sind [...] kostendeckend in Rechnung zu stellen.

'Services have to be charged so that the costs are covered.'

Indefinite noun phrases have an existential and a generic usage:



§ 3 Abs. 3 UniO UZH²

Domain characteristics

Dienstleistungen sind [...] kostendeckend in Rechnung zu stellen.

'Services have to be charged so that the costs are covered.'

Indefinite noun phrases have an existential and a generic usage:

1 $\mathcal{O} \exists^{>1} x : service(x) \land \dots$

2 $\mathcal{O} \ \forall x : service(x) \rightarrow \dots$



§ 3 Abs. 3 UniO UZH²

Domain characteristics

Dienstleistungen sind [...] kostendeckend in Rechnung zu stellen.

'Services have to be charged so that the costs are covered.'

Indefinite noun phrases have an existential and a generic usage:

1 $\mathcal{O} \exists^{>1} x : service(x) \land \dots$

2 $\mathcal{O} \ \forall x : service(x) \rightarrow \dots$

Convention

In statutes and regulations, indefinite plural noun phrases in subject position are used in the generic sense.



De-dicto vs. de-re modality

§ 2 Abs. 4 UniO UZH

Besondere Veranstaltungen können auch für eine breite Öffentlichkeit angeboten werden.

'Specific events can also be offered to a broader public.'



§ 2 Abs. 4 UniO UZH

Besondere Veranstaltungen können auch für eine breite Öffentlichkeit angeboten werden.

'Specific events can also be offered to a broader public.'

The subject can have wide scope over the modal verb or vice-versa:



§ 2 Abs. 4 UniO UZH

Besondere Veranstaltungen können auch für eine breite Öffentlichkeit angeboten werden.

'Specific events can also be offered to a broader public.'

The subject can have wide scope over the modal verb or vice-versa:

 \bullet $\exists^{>1}x$: $event(x) \land \mathcal{P} \dots$

de-re modality



§ 2 Abs. 4 UniO UZH

Besondere Veranstaltungen können auch für eine breite Öffentlichkeit angeboten werden.

'Specific events can also be offered to a broader public.'

The subject can have wide scope over the modal verb or vice-versa:

1 $\exists > 1 x : event(x) \land \mathcal{P} \dots$ de-re modality

2 $\mathcal{P} \exists^{>1} x : event(x) \land \dots$ **de-dicto modality**



De-dicto vs. de-re modality

§ 2 Abs. 4 UniO UZH

Besondere Veranstaltungen können auch für eine breite Öffentlichkeit angeboten werden.

'Specific events can also be offered to a broader public.'

The subject can have wide scope over the modal verb or vice-versa:

1 $\exists^{>1} x : event(x) \land \mathcal{P} \dots$ de-re modality

2 $\mathcal{P} \exists^{>1} x : event(x) \land \dots$ **de-dicto modality**

Convention

In statutes and regulations, the modal verb has always wide scope over the subject (de-dicto modality).



- drafts go through several editing cycles
- Central Language Services of the Federal Chancellery ensures
 - equivalence of the language versions
 - linguistic quality of the draft (including compliance with the guidelines)
- Linguistic guidelines for legislative drafting
 - issued by the Swiss Federal Office of Justice as well as by several cantonal governments
 - recommendations
 - not very systematic, not comprehensive



Anaphora resolution

Example from the guildlines of the canton of Zurich

- ¹Die Kantone können Fachhochschulen einrichten.
- ²**Sie** werden selbständig geleitet.
- ¹The cantons may establish technical universities.
- ²**They** are governered autonomously.'



Anaphora resolution

Example from the guildlines of the canton of Zurich

- ¹Die Kantone können Fachhochschulen einrichten.
- ²**Sie** werden selbständig geleitet.
- ¹The cantons may establish technical universities.
- ²**They** are governered autonomously.'

Guideline

Pronouns must only have one possible antecedent, namely the subject of the preceding sentence.



Example from the guildlines of the canton of Zurich

- ¹Die Kantone können Fachhochschulen einrichten.
- ²Die Fachhochschulen werden selbständig geleitet.
- ¹The cantons may establish technical universities.
- ²The technical universities are governered autonomously.'

Guideline

Pronouns must only have one possible antecedent, namely the subject of the preceding sentence.



§ 67 Abs. 2 UniO UZH

Ein Mitglied der Universitätsleitung führt den Vorsitz.

'A member of the executive board of the university takes the chair.'



§ 67 Abs. 2 UniO UZH

Ein Mitglied der Universitätsleitung führt den Vorsitz.

'A member of the executive board of the university takes the chair.'

 $\exists x : member(x) \land \dots$



§ 67 Abs. 2 UniO UZH

Ein Mitglied der Universitätsleitung führt den Vorsitz.

'A member of the executive board of the university takes the chair.'

 $\exists x : member(x) \land \dots$

§ 8 Abs. 7 UniO UZH

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

'A title can be revoked if the holder seriously violates the interests of the university.'



§ 67 Abs. 2 UniO UZH

Ein Mitglied der Universitätsleitung führt den Vorsitz.

'A member of the executive board of the university takes the chair.'

 $\exists x : member(x) \land \dots$

§ 8 Abs. 7 UniO UZH

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

'**A title** can be revoked if the holder seriously violates the interests of the university.'

$$\forall x : title(x) \rightarrow \dots$$



§ 3 Abs. 3 UniO UZH

Dienstleistungen sind [...] kostendeckend in Rechnung zu stellen.

'Services have to be charged so that the costs are covered.'

 $\forall x : service(x) \rightarrow \dots$



§ 3 Abs. 3 UniO UZH

Dienstleistungen sind [...] kostendeckend in Rechnung zu stellen.

'Services have to be charged so that the costs are covered.'

 $\forall x : service(x) \rightarrow \dots$

Proposed standard

In subject position, indefinite noun phrases are only used in the generic sense.



§ 67 Abs. 2 UniO UZH

Ein Mitglied der Universitätsleitung führt den Vorsitz.

'A member of the executive board of the university takes the chair.'



§ 67 Abs. 2 UniO UZH

Ein Mitglied der Universitätsleitung führt den Vorsitz.

'A member of the executive board of the university takes the chair.'

 $\forall x : member(x) \rightarrow \dots$



§ 67 Abs. 2 UniO UZH

Ein Mitglied der Universitätsleitung führt den Vorsitz.

'A member of the executive board of the university takes the chair.'

 $\forall x : member(x) \rightarrow \dots$



§ 67 Abs. 2 UniO UZH

Ein Mitglied der Universitätsleitung führt den Vorsitz.

'A member of the executive board of the university takes the chair.'

 $\forall x : member(x) \rightarrow \dots$

Rephrase (e.g. as a passive construction)

Der Vorsitz wird von einem Mitglied der Universitätsleitung geführt.

'The chair is taken by a member of the executive board of the university.'



§ 67 Abs. 2 UniO UZH

Ein Mitglied der Universitätsleitung führt den Vorsitz.

'A member of the executive board of the university takes the chair.'

 $\forall x : member(x) \rightarrow \dots$

Rephrase (e.g. as a passive construction)

Der Vorsitz wird von einem Mitglied der Universitätsleitung geführt.

'The chair is taken **by a member** of the executive board of the university.'

 $\ldots \land \exists x : member(x) \land \ldots$



§ 67 Abs. 2 UniO UZH

Ein Mitglied der Universitätsleitung führt den Vorsitz.

'A member of the executive board of the university takes the chair.'

 $\forall x : member(x) \rightarrow \dots$

Rephrase (e.g. as a passive construction)

Der Vorsitz wird von einem Mitglied der Universitätsleitung geführt.

'The chair is taken **by a member** of the executive board of the university.'

 $\ldots \land \exists x : member(x) \land \ldots$

Additional advantage:

The subject now correctly designates what the norm is about.



Plural ambiguity (distributive vs. collective reading)

Art. 12 BGG

Die Richter und Richterinnen können ihren Wohnort [...] frei wählen. '**The judges** can freely choose their place of residence [...].'

Art. 60 Abs. 2 BGG

Haben die Bundesrichter und Bundesrichterinnen den Entscheid in einer mündlichen Beratung getroffen, [....] 'If the Federal Justices have made their decision by oral deliberation, [....]'



Plural ambiguity (distributive vs. collective reading)

Art. 12 BGG

Die Richter und Richterinnen können ihren Wohnort [...] frei wählen. '**The judges** can freely choose their place of residence [...].'

Art. 60 Abs. 2 BGG

Haben die Bundesrichter und Bundesrichterinnen den Entscheid in einer mündlichen Beratung getroffen, [....] 'If the Federal Justices have made their decision by oral deliberation, [....]'

Proposed standard

Plurals are always used in the distributive sense. The collective reading is expressed with a singular term.



Plural ambiguity (distributive vs. collective reading)

Art. 12 BGG

Die Richter und Richterinnen können ihren Wohnort [...] frei wählen. '**The judges** can freely choose their place of residence [...].'

Art. 60 Abs. 2 BGG

Hat das Bundesgericht den Entscheid in einer mündlichen Beratung getroffen, [....] 'If the Federal Supreme Court has made its decision by oral deliberation, [....]'

Proposed standard

Plurals are always used in the distributive sense. The collective reading is expressed with a singular term.



Implicit anaphoric references

Art. 55 Abs. 1 AngO ETH-Bereich³

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

'Upon the birth of a child, the employee is entitled to a one-time allowance of 530 francs.'



Implicit anaphoric references

Art. 55 Abs. 1 AngO ETH-Bereich³

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

'Upon the birth of a child, the employee is entitled to a one-time allowance of 530 francs.'

 \rightarrow control in analogy to pronouns

Proposed standard

Relational nouns always refer to the subject of the main clause or, if they are part of the subject, to the subject of the immedieately preceding sentence.



State of development



The current version of Controlled Legal German comprises about two dozen rules, dealing with phenomona such as:

attachment ambiguities
 (prepositional phrases, relative clauses)



- attachment ambiguities
 (prepositional phrases, relative clauses)
- plural ambiguities
 (distributive/collective/cumulative readings)



State of development

- attachment ambiguities
 (prepositional phrases, relative clauses)
- plural ambiguities
 (distributive/collective/cumulative readings)
- scope ambiguties
 (modal verb, subject, objects, adverbials)



State of development

- attachment ambiguities
 (prepositional phrases, relative clauses)
- plural ambiguities
 (distributive/collective/cumulative readings)
- scope ambiguties
 (modal verb, subject, objects, adverbials)
- lexical ambiguities
 (articles, domain-specific function and content words)



- attachment ambiguities
 (prepositional phrases, relative clauses)
- plural ambiguities
 (distributive/collective/cumulative readings)
- scope ambiguties
 (modal verb, subject, objects, adverbials)
- lexical ambiguities
 (articles, domain-specific function and content words)
- referential ambiguities
 (pronouns, relational nouns)



State of development

- attachment ambiguities
 (prepositional phrases, relative clauses)
- plural ambiguities
 (distributive/collective/cumulative readings)
- scope ambiguties
 (modal verb, subject, objects, adverbials)
- lexical ambiguities
 (articles, domain-specific function and content words)
- referential ambiguities (pronouns, relational nouns)
- functional ambiguities
 (arising from the relatively free German word order)





• testing (processability and acceptability) and refining



- testing (processability and acceptability) and refining
- including further phenomena (e.g. underspecification)



- testing (processability and acceptability) and refining
- including further phenomena (e.g. underspecification)
- collecting re-usable linguistic building blocks for individual norm types



- testing (processability and acceptability) and refining
- including further phenomena (e.g. underspecification)
- collecting re-usable linguistic building blocks for individual norm types
- designing authoring tools



- testing (processability and acceptability) and refining
- including further phenomena (e.g. underspecification)
- collecting re-usable linguistic building blocks for individual norm types
- designing authoring tools
- composing a user guide for the standard (refining existing guidelines)



Potential and limitations I

How useful is a standard like CLG to a deep semantic processing of statutes and regulations?



Potential and limitations I

How useful is a standard like CLG to a deep semantic processing of statutes and regulations?

A standard like CLG cannot solve every problem . . .

- phenomena that cannot be controlled by a standard
- adequate representation of the content in formal logic



How useful is a standard like CLG to a deep semantic processing of statutes and regulations?

A standard like CLG cannot solve every problem ...

- phenomena that cannot be controlled by a standard
- adequate representation of the content in formal logic

. . . but it can get some major obstacles out of the way.

- instances of lexical ambiguity
- instances of syntactic ambiguity
- instances of semantic ambiguity



Potential and limitations II

Will statutes and regulations ever be written in standards like CLG?



Will statutes and regulations ever be written in standards like CLG?

Difficulties

- complexity of the drafting process (multiple parties involved)
- resistance of lawyers and politicians



Potential and limitations II

Will statutes and regulations ever be written in standards like CLG?

Difficulties

- complexity of the drafting process (multiple parties involved)
- resistance of lawyers and politicians

Preconditions

- institutionalised role of professional legal editors
- improvement also for human interpretability, translatability
- availability of useful applications (IR, expert systems, ...)

Crucial: acceptability of the standard to legal editors.



Conclusion

Acceptability

How can the acceptability of a standard like CLG be increased?



Acceptability

How can the acceptability of a standard like CLG be increased?

Usability/learnability (for professional legal editors)

- few rules that are easy to learn and use
- availability of authoring tools



How can the acceptability of a standard like CLG be increased?

Usability/learnability (for professional legal editors)

- few rules that are easy to learn and use
- availability of authoring tools

Expressivity

availability of adequate paraphrases



How can the acceptability of a standard like CLG be increased?

Usability/learnability (for professional legal editors)

- few rules that are easy to learn and use
- availability of authoring tools

Expressivity

• availability of adequate paraphrases

Naturalness/proximity to ordinary legal language

• standards must reflect existing frequency distributions



Acceptability

How can the acceptability of a standard like CLG be increased?

Usability/learnability (for professional legal editors)

- few rules that are easy to learn and use
- availability of authoring tools

Expressivity

availability of adequate paraphrases

Naturalness/proximity to ordinary legal language

standards must reflect existing frequency distributions

Required

Syntactically and semantically annotated corpora of legal texts.



We thank

- The Cogito Foundation, Switzerland
- Michael Hess, University of Zurich
- Rebekka Bratschi, Swiss Federal Chancellery
- Norbert Fuchs, University of Zurich for their support of the Collegis project.

Homepage of the Collegis project http://www.cl.uzh.ch/research/cnl/collegis.html

