

Remarks on Formalising Contracts from Natural Language

Adam Wyner, University of Aberdeen

17 March 2016

Runtime Verification COST Action
Prague, Czech Republic

Process

- Work with legal texts in natural language.
 - Use formal languages to formally represent the content – LegalRuleML (LRML) or a Controlled Natural Language (CNL).
 - Convert the formalisation to an executable rule language.
 - Draw inferences.
 - Reason with inconsistency (argumentation).
-
- Gist of talk: the methodology of translation from NL to any formal language needs to be addressed to facilitate subsequent automation and to understand the issues.

Problems

- Complexity of legal documents
 - Structure of sections and paragraphs
 - Cross references
 - Volume

Problems

- Complexity of sentences
 - Arguments (syntactic and semantic)
 - Predicates
 - Modifiers
 - Subordination
 - Classification: constitutive, prescriptive exemptions, reparations
 - Length
 - Ambiguities, e.g. scope of terminology, modifiers, exceptions
 - etc

Problems

- General conception of the relation between syntax and semantics.
 - Every man is rich
 - All x [$\text{man}(x) \rightarrow \text{rich}(x)$]
 - All (man, happy) - subset
 - Some man is rich
 - Some x [$\text{man}(x)$ and $\text{rich}(x)$]
 - Some (man, happy) - intersection
- Structural uniformity; constituent structure might matter:
 - [Every man and some woman] is rich. They are happy.

Previous work

- British Nationality Act (1981).
 - Formalised in Prolog with no methodology, no NLP tools, complex predicates of NL analysed as atomic predicates in Prolog.
 - Some follow on, but not much.
 - Oracle Policy Automation. Scopes problem. ‘just enough’ language analysis.

Rule-based Rule Extraction

Deontic rules:

- Agent and theme, which are semantic roles that must be associated with noun phrases in grammatical (subject or object) roles in the sentence. These are used to account for active-passive alternations and identify the individual's relationship to the deontic concept.
- Deontic modals and verbs.
- Main verbs.
- Exception clauses, which may appear in lists.

Conditional rules:

- Conditional sentences along with their antecedents and consequences. Antecedents may appear in lists.

Start with this. In future work, add punctuation, negation, temporal phrases, generics, tense in conditionals, references, and so on....

Sample Outputs

You may use human blood from a donor with a previous record of a reactive screening test for evidence of infection due to a communicable disease agent that is designated in paragraph a of this section, if:

- (1) At the time of donation, the donor is shown to be suitable by a requalification method; and
- (2) tests performed under paragraphs a are nonreactive.

Consequence, antecedent list structure

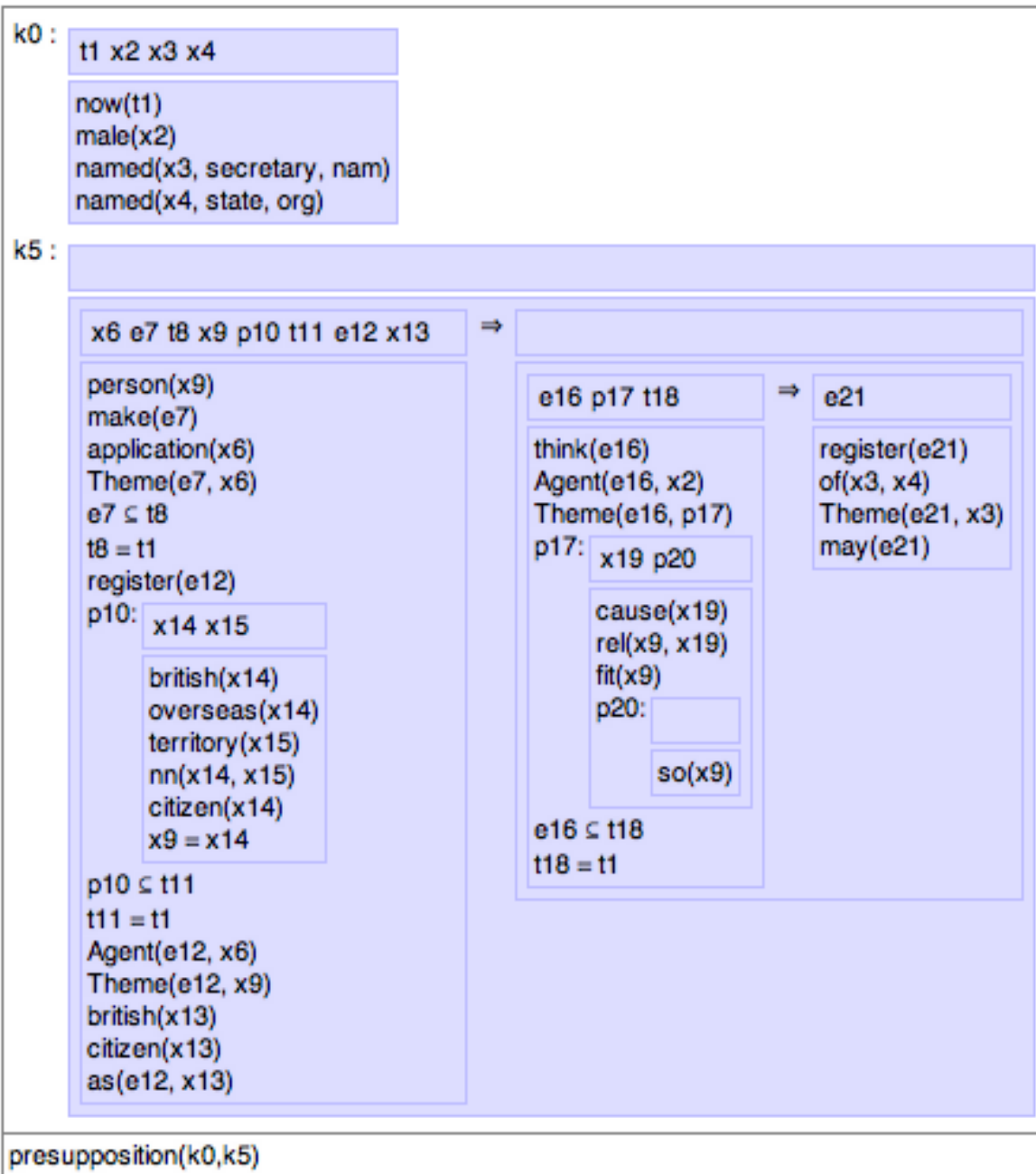
Except as specified in paragraphs c, you, an establishment that collects blood, must test each donation of human blood that is intended for use in preparing a product for evidence of infection due to the following communicable disease agents:

- (1) Human immunodeficiency virus, type 1;
- (2) Human T-lymphotropic virus, type I; and
- (3) Human T-lymphotropic virus, type II.

Exception, agent NP, deontic concept, active main verb, theme.

C&C/Boxer

- If an application is made to register as a British citizen a person who is a British overseas territories citizen, the Secretary of State may, if he thinks fit, cause the person to be so registered.



Output Analysis

First Antecedent

1. A *person* (x9) is identified as a *british overseas citizen* (see x14 in P10).
2. A *making* event (e7) has an *application* (x6) as *Theme*.
3. A *registering* event (e12) has the *application* (x6) as *Agent* and x9 as *Theme*.
4. The *registering* event (e12) is in the *as* relation with a *british citizen* (x13), which is distinct from x14.

This captures some interesting issues - the difference between different forms of predication, e.g. *as a British Citizen* and *is a British Citizen*, while leaving their exact interpretation aside. The proposition (p10) correctly predicates being a British overseas territory citizen of the person (x9) who wants to be a British citizen. There are two events - a making event and a registering event, where the theme of making (the application) is the agent of registering; this captures accurately the link between the passive and the subject of the infinitive. However, it makes an abstract noun the agent of registering, which is a conceptual ambiguity of the text. The agent of making is left unspecified, as it is in the source, though conceptually there ought to be one.

Policy statement in ACE for arguments

- p1: Every household should pay some tax for the household's garbage.
- p2: No household should pay some tax for the household's garbage.
- p3: Every household which pays some tax for the household's garbage increases an amount of the household's garbage which the household recycles.
- p4: If a household increases an amount of the household's garbage which the household recycles then the household benefits the household's society.
- p5: If a household pays a tax for the household's garbage then the tax is unfair to the household.
- p6: Every household should pay an equal portion of the sum of the tax for the household's garbage.
- p7: No household which receives a benefit which is paid by a council recycles the household's garbage.
- p8: Every household which does not receive a benefit which is paid by a council supports a household which receives a benefit which is paid by a council.
- p9: Tom says that every household which recycles the household's garbage reduces a need of a new dump which is for the garbage.

Sentences in ACE

- p10: Every household which reduces a need of a new dump benefits the household's society.
- p11: Tom is not an objective expert about recycling.
- p12: Tom owns a company that recycles some garbage.
- p13: Every person who owns a company that recycles some garbage earns some money from the garbage which is recycled.
- p14: Every supermarket creates some garbage.
- p15: Every supermarket should pay a tax for the garbage that the supermarket creates.
- p16: Every tax which is for some garbage which the supermarket creates is passed by the supermarket onto a household.
- p17: No supermarket should pay a tax for the garbage that the supermarket creates.
- p18: Tom is an objective expert about recycling.
- p19: If an objective expert says every household which recycles the household's garbage reduces a need of a new dump which is for the garbage, then every household which recycles the household's garbage reduces a need of a new dump which is for the garbage.

Previous work

- Semantics of Business Vocabulary and Rules (SBVR) on financial regulations. Ceci et al.
- Formal contracts – deontic logic + action/event logics. Maibaum/Sergot.

Text Source – Legal Instruments

Smoking, Health and Social Care (Scotland) Act 2005

The Prohibition of Smoking in Certain Premises (Scotland) Regulations 2006

The act is a general statement of principles and the law. The regulations contain further detail about how the act is implemented and interpreted.

Act

1. Offence of permitting others to smoke in no-smoking premises
 1. A person who, having the management or control of no-smoking premises, knowingly permits another to smoke there commits an offence.
 2. A person accused of an offence under this section is to be regarded as having knowingly permitted another to smoke in no-smoking premises if that person ought to have known that the other person was smoking there.
 3. It is a defence for an accused charged with an offence under this section to prove that the accused (or any employee or agent of the accused) took all reasonable precautions and exercised all due diligence not to commit the offence; or that there were no lawful and reasonably practicable means by which the accused could prevent the other person from smoking in the no-smoking premises.

Act

3. Display of warning notices in and on no-smoking premises
 1. If notices are not conspicuously displayed in, on or near no-smoking premises so as to be visible to and legible by persons in and persons approaching the premises; and stating that the premises are no-smoking premises; and that it is an offence to smoke there or knowingly to permit smoking there, the person having the management or control of the premises commits an offence.
 2. It is a defence for an accused charged with an offence under this section to prove that the accused (or any employee or agent of the accused) took all reasonable precautions and exercised all due diligence not to commit the offence.

Act

7. Powers to enter and require identification

1. An authorised officer of the appropriate council may enter and search any no-smoking premises in order to ascertain whether an offence under section 1, 2 or 3 has been or is being committed there.
2. A power under this section may be exercised, if need be, by force.

Regulation

“adult” means a person aged 16 years or over;

“private vehicle” means any car which is not a public transportation vehicle; and any other vehicle which is used primarily for the private purposes of the person who owns it or of a person having the right to use it, provided always that such right to use a vehicle does not, in relation to that vehicle, include a reference to a person whose right to use the vehicle derives only from having paid, or undertaken to pay, for the use of the vehicle and its driver for a particular journey nor does it include a reference to any public transportation vehicle;

Regulation

(a) “premises” includes any building or part of a building; any structure or part of a structure, whether moveable or otherwise; any installation on land (including the foreshore and other land intermittently covered by water), any offshore installation, and any other installation (whether floating, or resting on the seabed or the subsoil thereof, or resting on other land covered with water or the subsoil thereof); any tent, marquee or stall; and any vehicle.

(c) “substantially enclosed” means for premises other than a vehicle or part of a vehicle, having a ceiling or roof and, except for doors, windows and passageways, substantially enclosed, whether permanently or temporarily; or for premises that are a vehicle, or part of a vehicle, having a top or roof and, except for doors, windows or exits, substantially enclosed, whether permanently or temporarily, and in determining whether premises are “substantially enclosed”, no account is to be taken of openings in which there are doors, windows or other fittings that can be opened or shut;

Regulation

Schedule 1

No-smoking premises:

Restaurants, Hotels, Public toilets, Sports centres, Public telephone kiosks,

Schedule 2

Exemptions:

Residential accommodation, adult hospices, designated rooms in psychiatric hospitals and psychiatric units,

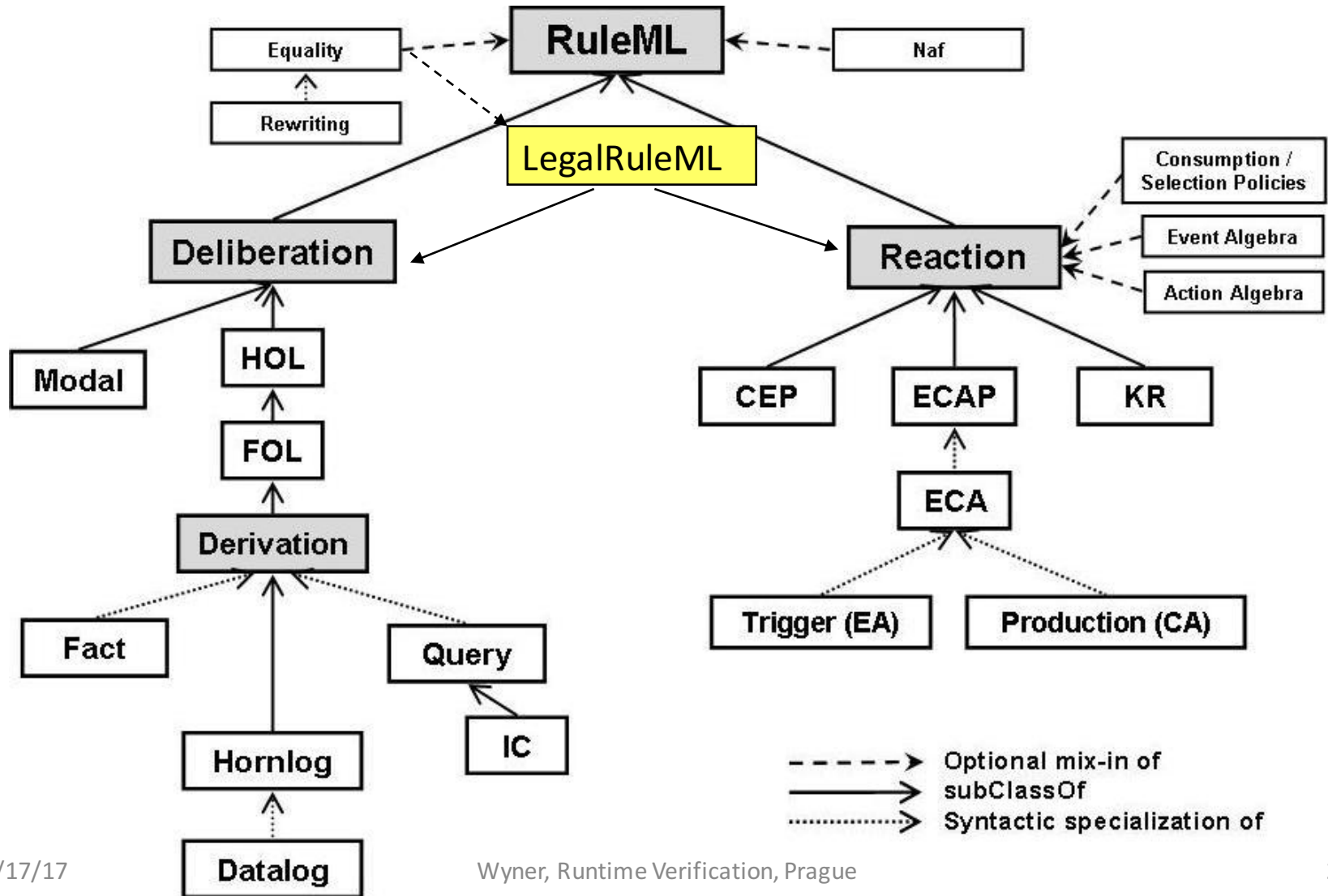
Formalisation

- In formalising legislation, the issue is not just pleasing logicians and/or computer scientists.
- The legislators, lawyers, and members of the public also have to be pleased (in some way for some reason).
- Use cases and applications that serve the interests of the non-maths people.
- The formalisation also has to suit the domain that is being represented - custom fit.

Formal language - LegalRuleML

- Motivations, Goals, Principles
- Design principles
- LegalRuleML main blocks: meta, context, rules
 - Legal Statements and References
 - Temporal Events and Temporal Situations
 - Deontic
 - Penalty and Reparation
 - Defeasibility
 - Alternatives
 - Authority, Jurisdiction
 - Actor, Figure, Roles

RuleML family of sublanguages



Telecommunications Consumer Protections Code, Australia

Date of Assent: 30 May 2012

Date of Registration: 11 July 2012

Date of Efficacy: 1 September 2012

2.1 sec2.1-v2

Complaint

sec2.1-list1-itm31-v2

means an expression of dissatisfaction made to a Supplier in relation to its Telecommunications Products or the complaints handling process itself, where a response or Resolution is explicitly or implicitly expected by the Consumer.

rule1a

rule1b

rule2

rule3

rule4

An initial call to a provider to request a service or information or to request support is not necessarily a Complaint. An initial call to report a fault or service difficulty is not a Complaint. However, if a Customer advises that they want this initial call treated as a Complaint, the Supplier will also treat this initial call as a Complaint.

rule1b<rule2

rule1b<rule3

rule3<rule4

If a Supplier is uncertain, a Supplier must ask a Customer if they wish to make a Complaint and must rely on the Customer's response.

rule5

TCP Code, Australia

```
<lrml:hasStatements key="rulebase1-v2">  
  <lrml:ConstitutiveStatement key="rule1b-v2">  
    <ruleml:if>  
      <ruleml:Atom key="rule1-atom2-v2">  
        <ruleml:Rel iri="#rule1-rel2-v2">is an expression of  
dissatisfaction made to a Supplier in relation to its Telecommunications Products  
or the complaints handling process itself, where a response or Resolution is  
explicitly or implicitly expected by the Consumer</ruleml:Rel>  
        <ruleml:Var>X</ruleml:Var>  
      </ruleml:Atom>  
    </ruleml:if>  
    <ruleml:then>  
      <ruleml:Atom key="rule1-atom1-v2">  
        <ruleml:Rel iri="#complaint-v2"/>  
        <ruleml:Var>X</ruleml:Var>  
      </ruleml:Atom>  
    </ruleml:then>  
  </lrml:ConstitutiveStatement>
```

TCP Code, Australia

```
<lrml:PrescriptiveStatement key="rule5-v2">
  <lrml:if>
    <ruleml:Atom key="rule5-atom1-v2">
      <ruleml:Rel iri="rule5-rel1-v2">is uncertain if/wishes to make a Complaint</ruleml:Rel>
      <ruleml:Var type="#supplier-v2">S</ruleml:Var>
      <ruleml:Var type="#customer-v2">C</ruleml:Var>
    </ruleml:Atom>
  </lrml:if>
  <lrml:then>
    <lrml:Obligation key="rule5-ob1-v2">
      <lrml:And key="rule5-and1-v2">
        <ruleml:Atom key="rule5-atom2-v2">
          <ruleml:Rel iri="rule5-rel2-v2">asks/if they wish to make a Complaint</ruleml:Rel>
          <ruleml:Var>S</ruleml:Var>
          <ruleml:Var>C</ruleml:Var>
        </ruleml:Atom>
        <ruleml:Atom key="rule5-atom3-v2">
          <ruleml:Rel iri="#rule5-rel3-v2">relies on the response of</ruleml:Rel>
          <ruleml:Var>S</ruleml:Var>
          <ruleml:Var>C</ruleml:Var>
        </ruleml:Atom>
      </lrml:And>
    </lrml:Obligation>
  </lrml:then>
</lrml:PrescriptiveStatement>
```

How?!

- What is the methodology to go from the source NL text to the formalisation?
- What are the tools to support this translation?

LegalRuleML Problem

- No methodology. No tools. All hand-crafted.
- Auxiliary note: while a Controlled Natural Language (fixed vocabulary, syntax, and guidance on usage) might help, it still does not say that an "accepted" sentence means what the source intends.
 - Manual checking is problematic
 - The more complex the sentences and documents, the greater the problem.

LegalRuleML Analysis Methodology

- One thing to make a formal language. Another thing to systematically translate from NL to the formal language, using some uniform methodology for analysts.
- Methodology for uniformity, transparency, comparison....
- Similar to a methodology for translating from NL to FOL, but for legal statements....

- Proposed an initial methodology, starting from a very coarse-grained analysis, then (incrementally) refining and developing tool support.
- Use feedback.

Methodology in General

- Define a coarse-grained level of analysis
 - Fix high-level interpretation
 - Pinpoint ambiguities
 - Provide support for rule mining in legal sources
 - Provide clear methodology: guidelines and tool support
- Work bottom-up towards refinement of the analysis, rather than top-down.

Segmentation

Annotation Guideline 1: annotate each sentence or each independent clause of a sentence as a paraphrase of a rule in a LegalRuleML statement and associated it with a unique identifier, e.g. KEY for some particular value:

```
<lrml:Statement key="KEY">  
  <ruleml:Rule>  
    <lrml:Paraphrase> ANNOTATED SEGMENT </lrml:Paraphrase>  
  </ruleml:Rule>  
</lrml:Statement>
```


Segmentation

Annotation Guideline 2: annotate each entitled or numbered sequence of sentences as a group of statements (`lrml:statements`) associated with a unique identifier, e.g. “b”:

```
<lrml:Statements key="b">
  <lrml:PrescriptiveStatement>
    <ruleml:Rule>
      <lrml:Paraphrase>To test for..., you must... </lrml:Paraphrase>
    </ruleml:Rule>
  </lrml:PrescriptiveStatement>
  <lrml:PrescriptiveStatement>
    <ruleml:Rule>
      <lrml:Paraphrase>You must perform tests... </lrml:Paraphrase>
    </ruleml:Rule>
  </lrml:PrescriptiveStatement>
</lrml:Statements>
```

Classification

Annotation guideline 3: If a statement contains deontic markers like *must*, *is required to*, *It is obligatory that*, *obligatorily*, etc. or terminology related to permission or prohibition, consider categorizing it as a prescriptive statement.

Annotation guideline 4: If a statement contains definition markers like *definition*, *means*, etc. consider categorizing it a constitutive statement.

Annotation guideline 5: Associate a category (e.g. prescriptive or constitutive) to each statement.

```
<lrml:ConstitutiveStatement>  
  <ruleml:Rule>  
    <lrml:Paraphrase>Complaint means an expression of dissatisfaction made  
    to a Supplier...</lrml:Paraphrase>  
  </ruleml:Rule>  
</lrml:ConstitutiveStatement>
```

Exception

Annotation guideline 6: for each statement (associated with a KEY) with an exception marker, consider adding an override statement:

```
<lrml:PrescriptiveStatement key="cli">
  <ruleml:Rule>
    <lrml:Paraphrase>You must test donations ...; </lrml:Paraphrase>
  </ruleml:Rule>
</lrml:PrescriptiveStatement>

<lrml:PrescriptiveStatement key="clibis">
  <ruleml:Rule>
    <lrml:Paraphrase>if..., you may perform such testing only on... </lrml:Paraphrase>
  </ruleml:Rule>
</lrml:PrescriptiveStatement>

<lrml:OverrideStatement>
  <lrml:Override over="#clibis" under="#cli"/>
</lrml:OverrideStatement>
```

Future work

- Collaborating with Scottish Parliamentary Counsel Office (authors of legislation/regulation) on specific texts to provide detailed examples, identify problems, and refine the methodology
- Consider quality assurance, e.g. consistency of annotations as well as principled variant interpretations
- Develop interactive NLP tools to support the work to move from text to LegalRuleML
- Extend the guidelines to further LegalRuleML components
- See if there are ways to augment LegalRuleML (implicit rules)

Thanks!

Questions

azwyner@abdn.ac.uk