Identifying and Classifying Ambiguity for Regulatory Requirements

Aaron Massey
Postdoctoral Fellow
School of Interactive Computing
akmassey@gatech.edu
@akmassey

Co-Authors: Richard L. Rutledge, Annie I. Antón, and Peter Swire

27 Aug 2014
Legal Domain: Healthcare

• Health Insurance Portability and Accountability Act (HIPAA) passed in 1996
  – Regulates security and privacy for healthcare organizations
  – $25,000 fines per violation per year for non-criminal violations
  – Amended by the HITECH Act in 2009 to address data breaches and increase enforcement actions

• Recent Settlement Actions:
  – Concentra Health Services – $1.7 Million (April 2014)
  – New York and Presbyterian Hospital – $3.3 Million (May 2014)
  – Columbia University Hospital – $1.5 Million (May 2014)
Legal Ambiguity: a Critical Challenge for Requirements

- Legal texts are often **intentionally ambiguous**.
  - Example:
    “make **reasonable** efforts to limit protected health information to the minimum necessary to accomplish the intended purpose of the use” – HIPAA §164.502(b)
  - The word “**reasonable**” **appears 61 times** in HIPAA!

- Traditional approaches, such as disambiguation or removal, do not work for legal ambiguities.
  - Legal texts cannot easily be re-written
  - Legal stakeholders cannot easily be sought out for definitive clarification.
  - Requirements engineers must interpret ambiguities in legal texts!
What is ambiguity?

- **ANSI/IEEE Standard 830-1993**: a requirements specification is **unambiguous only when each requirement has a single interpretation.**

- **Definitional Concerns**:  
  - Should a statement with no clear interpretation be considered ambiguous?  
  - What constitutes a valid interpretation? Who decides?

- **No objective standard exists.**  
  - There is no “correct” identification or classification of ambiguity.  
  - We do have relative standards: Does a group agree as a whole on an interpretation?
Research Overview

- **Case study of 18 students** identifying and classifying ambiguity
  - technologists
  - policy analysts
- Using a **taxonomy** based on linguistics, software engineering, and legal understandings of ambiguity.
- Legal text: **§170.302 of the HITECH Act.**
  - 23 paragraphs (104 lines)
  - Meaningful Use Stage 1 Criteria for a certified EHR
- **Tutorial introducing the taxonomy** and study procedure
- 5 Research Questions
# A Taxonomy of Ambiguity

<table>
<thead>
<tr>
<th>Lexical</th>
<th>Syntactic</th>
<th>Semantic</th>
<th>Vagueness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incompleteness</td>
<td>Referential</td>
<td>Other</td>
<td>Unambiguous</td>
</tr>
</tbody>
</table>

© 2006-2014 Aaron Massey et al., Georgia Institute of Technology
Lexical Ambiguity

- Lexical ambiguity occurs when a word or phrase has multiple valid meanings.
- Examples:
  - **Conversational**: Melissa walked to the bank.
  - **§170.302(d)**: Enable a user to electronically record, modify, and retrieve a patient’s active medication list as well as medication history for longitudinal care.
Syntactic Ambiguity

- Syntactic ambiguity occurs when a sequence of words has multiple valid grammatical parsings.

- Examples:
  - Conversational: I saw the man with the binoculars.
  - §170.302(f): Enable a user to electronically record, modify, and retrieve a patient’s vital signs…
Semantic Ambiguity

- Semantic ambiguity occurs when a sentence has more than one interpretation based entirely on the surrounding context.

- Examples:
  - Conversational: Fred and Ethel are married.
  - §170.302(j): Enable a user to electronically compare two or more medication lists.
Vagueness

- Vagueness occurs when a term or statement admits borderline cases or relative interpretation.
- Examples:
  - **Conversational**: George is tall.
  - **§170.302(h)(3)**: Electronically attribute, associate, or link a laboratory test result to a laboratory order or patient record.
Incompleteness

- Incompleteness occurs **when a statement fails to provide enough information** to have a single clear interpretation.

- Examples:
  - **Conversational**: Combine flour, eggs, and salt to make fresh pasta.
  - **§170.302(a)(2)**: Provide certain users with the ability to adjust notifications provided for drug-drug and drug-allergy interaction checks.
Referential Ambiguity

- Referential ambiguity occurs when a word or phrase in a sentence cannot be said to have a clear reference.

- Examples:
  - Conversational: The boy told his father about the damage. He was very upset.
  - §170.302(n): For each meaningful use objective with a percentage-based measure, electronically record the numerator and denominator...
Please complete the table below using the procedure detailed in the tutorial.

Ambiguities Found in 45 CFR Subtitle A, § 170.302(a)(1)–(2)

<table>
<thead>
<tr>
<th>Lexical</th>
<th>Syntactic</th>
<th>Semantic</th>
<th>Vagueness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Incompleteness  Referential  Other  None

Software engineers should be able to build software that complies with this legal text.

Circle one: 

Agree  Disagree
Research Questions 1 to 3

1. Does the taxonomy provide adequate coverage of the ambiguities found in § 170.302?

2. Do participants agree on the number and types of ambiguities they identify in § 170.302?

3. Do participants agree on the number and types of intentional ambiguities they identify in § 170.302?
Research Questions 4 and 5

4. **Do participants agree** on whether software engineers should be able to build software that complies with each paragraph of § 170.302?

5. Does an **identified ambiguity affect** whether participants believe that software engineers should be able to build software that complies with each paragraph of § 170.302?
Q1 Measures: (1) Use of each of the first six ambiguity types and (2) minimal use of the “Other” type.

Q2 Measures: ICC for both number and type of ambiguities identified

Q3 Measures: ICC for both number and type of intentional ambiguities identified

Q4 Measures: Fleiss Kappa agreement on implementability of the paragraph.

Q5 Measures: The percentage of paragraphs deemed unimplementable that contain identified ambiguities
Q1: Taxonomy Coverage

- Participants identified on average 33.47 ambiguities for the 23 paragraphs examined.
  - 50 minutes provided for the study
  - All participants finished before time was up

- Every ambiguity type was used.
  - Least frequent: Semantic (1.59 on average)
  - Most frequent: Vagueness (9.82 on average)

- The “Other” type was less common than the least common ambiguity classification we defined (0.82 on average).

- Result: Yes, the taxonomy provides adequate coverage.
Q2: Number and Type agreement

- **Number agreement:** ICC: 0.316, indicating *fair agreement on number* *(p < 0.001)*

- **Type agreement:**
  - For 2 of the 23 paragraphs, the participants demonstrated near-universal agreement.
  - For the remaining 21, the participants demonstrated only slight agreement.
  - Overall Fleiss Kappa for type agreement: 0.0446, indicating *slight agreement on type* *(p < 0.0029)*
Q3: Intentional Number and Type

- **Number:** ICC 0.141, (p < 0.0001)
- **Type:** ICC 0.201 (p < 0.001)
- The Incompleteness category was a primary driver of type disagreement.
  - Technologists identified significantly more ambiguities of this type.
  - Removing Incompleteness, Type agreement ICC becomes 0.39, indicating fair agreement (p < 0.0001)
- **Result:** Participants agreed less on intentional ambiguities than on total ambiguities.
Ambiguities by Type and Intent

Bar chart showing the number of ambiguities identified by category, with subcategories of Technologist (unintentional and intentional), Policy Analyst (unintentional), and Referential. The chart compares different types of ambiguities such as lexical, syntactic, semantic, and vagueness.
Q4: Implementability

- **All participants:** Fleiss Kappa value of 0.0052, \( p < 0.788 \) — not statistically significant.
- **Technologists:** Fleiss Kappa value of 0.0455, \( p < 0.116 \) — not statistically significant.

- **Result:** Participants agreement on implementability was not statistically significant.
Q5: Ambiguity and Implementability

- 89% of unimplementable paragraphs contained an ambiguity
- 48% of implementable paragraphs contained an ambiguity

Result: Yes, ambiguity is more commonly identified in paragraphs deemed unimplementable.
Summary

- In 50 minutes over 104 lines of legal text our participants identified **33.47 ambiguities on average**
- **The taxonomy provided reasonable coverage:** **97.5%** of all ambiguities identified were classified as one of the six defined types
- Participants accepted paragraphs with unintentional ambiguity as implementable!
Future Work

- Participants did not exhibit strong agreement on the number and type of ambiguity.
  - 50 minute limit?
  - Better guidelines for the taxonomy?
- Additional case studies
  - More participants
  - Different legal domains
  - Does identifying and classifying ambiguity prior to other legal requirements activities improve performance?
Thank You!
Questions?

Aaron Massey
Postdoctoral Fellow
School of Interactive Computing
akmassey@gatech.edu
@akmassey

Co-Authors: Richard L. Rutledge, Annie I. Antón, and Peter Swire

27 Aug 2014