The Severity and Prevalence of Ambiguity in Software Engineering Requirements

Cristina Ribeiro
PhD Candidate
University of Waterloo

Overview

- Introduction
- Research Problem
- Proposed Solution
- Ambiguity Model
- Research Method

Introduction

- Requirements Engineering (RE) captures all requirements from all stakeholders in a Requirements Specification (RS)
- Ambiguity in a RS could cause the incorrect implementation of software
- Almost all RSs are written in natural language (NL) and NL is inherently ambiguous

- Project failure has often been attributed to ambiguities in RSs
- This attribution claim has not been conclusively empirically proven
- This claim has fuelled research in methods and tools for removing ambiguities
- The earlier an ambiguous requirement is found the less expensive it is to fix
- The paradox is that finding ambiguities is expensive and time-consuming

Research Problem

- We don't know:
- the exact the benefits of the tools & methods,
- if these tools & methods are worth the effort,
- if ambiguity is in fact costly

Identifying the Effects of RS Ambiguity

- De Bruijn reviewed 1 RS for a failed project
- He used a published taxonomy of ambiguity types
- He found far too many instances of ambiguities in the RS, more than feasible to analyze each for its effect
- He cut the number to size by taking a random sampling

- De Bruijn found only 1 not very severe defect was caused by an ambiguous requirement
- The test and development teams worked through all other ambiguities & successfully developed those requirements
- He concluded that for the RS examined the ambiguities that remained were not critical and did not contribute to the failure
- Few ambiguities affected development because conversations during requirements analysis took care of most ambiguities

Drawbacks to De Bruijn's Study

- His random sampling strategy could have missed a lot of ambiguities, some which could have caused damage
- Ambiguous requirements that are expensive to fix may be too infrequent to catch with random sampling
- This drawback could have caused him to come to an incorrect conclusion

Research Questions

- What is the severity of ambiguous requirements?
- What has the lower cost -- searching for ambiguities in the RS during RE or repairing the damage caused by undetected ambiguities in later stages of the software development life cycle?

Proposed Strategy

- Focus on ambiguities likely to be missed by stakeholders & remain after analysis
- These ambiguities are more likely to cause expensive problems, requiring fixing late in the development
- These ambiguities are the ones likely to suffer subconscious disambiguation (SD)

Subconscious Disambiguation

- occurs when an ambiguous sentence is interpreted as unambiguous, with only one possible interpretation
- My strategy is to focus on the ambiguities that people are unaware of
- People unaware of an ambiguity cannot identify that another interpretation exists

Ambiguity Types that People are not Aware of

- Demonstrative reference
- Conditional clausal reference
- Ellipsis
- Misplaced modifier
- Plural

The Model's Restrictions & Application

- Restrict focus to the English language
- Focuses on how people use language, syntactically correct or incorrect
- Restrict focus to written linguistic ambiguities in RSs
- The model will be used to identify ambiguities likely to suffer SD

Referential Ambiguity

- occurs when a personal or demonstrative pronoun can refer to more than one referent
- Example 1. Bob said to Joe that he must leave
- He could refer to either
- Meaning A. Bob
- Meaning B. Joe
- Meaning C. someone else

- Example 2. This prevents security breaches

- This could refer to
- a word
- a phrase
- a clause
- multiple sentences
- an idea

- Example 3. They kidnapped my dog
- Meaning A. They could refer to some specific plural noun previously given, such as in Three thieves came to my house
- Meaning B. They could mean an indeterminate number of people such as in Some people kidnapped my dog
- Meaning C. They could mean a person of indeterminate gender such as in Someone kidnapped my dog

Conditional Clause Reference Ambiguity

- occurs when a conditional clause can refer to more than one condition
- Example 4.

I. When a user requests a book with an available status, assign book to user.

II. When the user requests a book with a checked-out status, place a hold on the book for the user.

III. If so, increase the book's number-of-user-requests counter.

- so could refer to
- When a user requests a book with an available status
- When the user requests a book with a checked-out status
- The meaning of so or not can go back further than the convention of going to the most recent referential element

Elliptic Ambiguity

- occurs when an element has been elided, and there is more than one possible element from the discourse that the elided element could be
- Example 5. Perot knows a man richer than Trump
- Meaning A. Perot knows a man richer than Trump knows
- Meaning B. Perot knows a man richer than Trump is

Modifier Ambiguity

- occurs when a sentence contains a modifier and there is more than one possible element that the modifier could be modifying

- Correct Placement
  occurs when a modifier is placed immediately preceding what it modifies

- Common Placement
  occurs when a modifier is placed before the main verb regardless of what is intended to be modified.

- Example 6. I only nap in the afternoon
- Meaning A. The only thing I do is nap in the afternoon

- If the writer intended that I do not take naps at any time of the day other than the afternoon, the correct placement of only is
- Meaning B. I nap only in the afternoon

- There are cases where common placement is unambiguous

- Example 7. I ONLY ATE VEGETABLES

- Meaning A. I DIDN'T BUY, WASH OR COOK THE VEGETABLES, I ONLY ATE THEM

- Meaning B. I ATE ONLY VEGETABLES

- When placed at end of sentence, can refer to only the previous word

- Example 8. I NAP IN THE AFTERNOON ONLY

- Meaning A. I NAP IN ONLY THE AFTERNOON

- This is uncommon placement of a modifier

Plural Ambiguity

- occurs when a sentence contains a plural subject and or object and it's unclear whether the object or subject complement refers to a collective or a distributive interpretation
- Example 9. Two men lift a table
- Meaning A. Two men lift a single table together
- Meaning B. Each of two men lifts his own table

Distributive Sub-Entailment

- Example 10. JOHN LIFTED THREE TABLES
- Meaning A. JOHN LIFTED THREE TABLES AS A SET
- Meaning B. JOHN LIFTED EACH TABLE

- Meaning B is a distributive sub-entailment of Meaning A

Plural Amb Resolution

- To avoid plural ambiguity for a collective interpretation
- a writer should use a singular noun naming the collection
- Example 18. A GROUP OF TWO MEN LIFTS A TABLE
- Example 19. A PAIR OF MEN LIFTS A TABLE

Plural Amb Resolution

- To avoid ambiguity for a distributive interpretation
- a writer should use singular
- Example 20. EACH OF THE TWO MEN LIFTS A TABLE

Research Method

- A major company supplied three high quality RSs
- Each RS was successfully implemented for a major computer based system
- Review each RS, searching for ambiguities likely to suffer SD
  - This keeps the number of instances down to a manageable number

Research Method Con't

- Meet with RS analyst to review findings
- Will examine development histories, if available, for signs that the ambiguities found caused development problems
- The severity of these problems will be estimated

Validity Threats

- External validity
  - An empirical study limits generalizability
  - This study is of realistic size to real-life problems
  - This study is using real-life data, and the data sets are large

Construct Validity

- Assumption:
  People are not aware of the ambiguity types likely to suffer SD, and they remain in RSs after multiple inspections
- Assumption validity
  - If everyone were aware of all ambiguities likely to suffer SD, RS inspections would not be required
  - The assumption is necessary when working with large real-life data sets

Internal Validity

- Unless the issues documented clearly state that ambiguity caused the issue the cause is debatable
- A positive result is not diminished by the assumption
- A negative result is conclusive because the inspection is a good representation of what an industrial inspection would find

Thank You!

Lexical Ambiguity

- occurs when a homonymous or polysemous word occurs in a sentence and these words have multiple meanings.
- Example 21. BANK
- Meaning A. FINANCIAL INSTITUTION
- Meaning B. EDGE OF A RIVER

- Example 22. Green

- Meaning A. The colour of an object

- Meaning B. An emotion such as green with envy

- Meaning C. Youth, vitality or inexperience

Analytical Ambiguity

- occurs when the role of the elements in a phrase or sentence is ambiguous
- Example 23. The french history teacher
- Meaning A. The teacher of French history
- Meaning B. The history teacher who is French

Attachment Ambiguity

- occurs when either a prepositional phrase or relative clause can be syntactically attached to more than one part of a single sentence and renders different interpretations
- Example 24. THE POLICE SHOT THE RIOTERS WITH GUNS
- The prepositional clause WITH GUNS could be attached to
- Meaning A. THE POLICE
- Meaning B. THE RIOTERS

- Example 25. The lamp near the painting in the house that was damaged in the flood
- the relative clause was damaged in the flood could be attached to
  - Meaning A. the lamp
  - Meaning B. the painting
  - Meaning C. the house

Coordination Ambiguity

- occurs in a sentence with either more than one conjunction or a modifier and a conjunction
- Example 26. I saw Jane and Calvin and Jack saw me
- Meaning A. I saw Jane and Clavin, and Jack saw me
- Meaning B. I saw Jane, and Calvin and Jack saw me

- Example 27. YOUNG MAN AND WOMAN
- Meaning A. YOUNG MAN AND YOUNG WOMAN
- Meaning B. WOMAN AND YOUNG MAN

- Resolution: a writer should use line breaks and indentation, as in this sentence, to show the structure of coordination,
use punctuation,
use additional words, or
change the wording.

Plural Ambiguity

- Example 11. She summarized the proposals
- Meaning A. The proposals were summarized as a whole
- Meaning B. Each proposal was individually summarized

Multiple Plural Noun Phrases

- Example 12. Two men lift three tables
- Meaning A. Two men together lift three tables as a set
- Meaning B. Two men together lift three tables individually
- Meaning C. Two men individually lift three tables as a set
- Meaning D. Two men individually lift three tables individually

Quantifiers

- Example 13. Many bring their dogs
- Example 14. Few bring their dogs
- Meaning A: Many to many relationship
- Meaning B: Many to one relationship

- Example 15. Many bring their dog
- Meaning A: Many to one relationship
- Meaning B: One to one relationship

- Example 16. All lights in the room are connected to a single switch

- Meaning A: All lights in the room are connected to a single shared switch

- Meaning B: Each light in the room has its own unshared single switch

- Example 17. Every light has their switch
- Meaning A: All the lights share a single switch
- Meaning B: Each light has its own switch

Vagueness Phenomenon

- occurs when a sentence contains a subjective noun phrase and results in more than one possible interpretation
- Example 28. TALL
- Meaning A. A TALL PERSON could mean OVER TWO METERS IN HEIGHT
- Meaning B. A TALL BASKETBALL PLAYER could mean OVER 2.5 METERS IN HEIGHT
- Meaning C. A TALL JOCKEY could mean 1.5 METERS

Generality Phenomenon

- occurs when a sentence contains a general non-specific noun phrase and results in more than one interpretation

- Example 29. Cousin
- Meaning A. a female cousin
- Meaning B. a male cousin