Content Analysis of Conclusions of Empirical Studies in Software Engineering

Bachelors Thesis – Closing Presentation

Lars Spitzlay

Table of Contents

- Goals

- Related Work

- Content Analysis

- The process

- Data quality

- Results/Discussion

- Conclusion

Goals

- Evaluate the quality of empirical studies in SE

- Are the conclusions „good"?
  - Do they contain appropriate generalization?

- Provide valuable insights for future lectures/tutorials on
  - Structure of the tutorials
  - Code system used for the Content Analysis

- Analysis of gathered data
  - What makes a sentence easy or hard to code?

Related Work

- Fair amount of studies about the quality of empirical studies in SE

- Focus on strengths/weaknesses, providing hints and roadmaps

- Most evaluations agree: quality is not good enough

- Perry *et al.* (2000). "Empirical Studies of Software Engineering: A Roadmap"

- Zannier *et al.* (2007). "On the success of empirical studies in the international conference on software engineering"

Content analysis – a quick overview

- Take or create a coding system
  - i.e. by open coding
- Search for all occurrences of each code in your material
- Count those occurrences
- This process should be conducted at least twice, by independent reviewers
- Calculate Inter-reviewer agreement

Lecture/Tutorial

- Develope a unified code system

- Code the conclusion + relevant parts of empirical papers from ICSE 2021

- Data validation, clustering and code correlation

Data quality

- Agreement between different reviewers limited (about 30%)

=> Makes it hard to make strong statements about anything

=> Can we improve data quality and salvage the data somehow?

Data quality improvements

- Only consider conclusion

- Merge some codes with low agreement that fit together

- Improve data validation/cleansing

=> Even after several improvements, agreement is still below 50%

Reasons for poor data quality

- High workload

- A significant amount of human error
  - Failure to abide by some coding-rules

- Disagreement on code usage
  - Generalization vs. Repeat research process
  - Generalization vs. Further work

- Coding is hard!

- High agreement is most often reached when certain keywords are present

Conclusion

- Provide more time for the students

- Reduce the number of studies used
  - Reduces stress, improves quality of the data collected

- Improve the conciseness of the code system

- Have a stricter guide to coding
  - If a code is used, certain other code sub-trees must be covered as well

Thanks for your attention!

- Any questions?

Content Analysis of Conclusions of Empirical Studies in Software Engineering