



UNIVERSITÄT ZU LÜBECK  
INSTITUT FÜR  
THEORETISCHE INFORMATIK

Können semantische Ähnlichkeiten von Wörtern die  
Schlussfolgerungen des gesunden Menschenverstands verbessern? Eine  
Fallstudie mit Prover E und SUMO.

*Can semantic similarities of words enhance common sense reasoning? A case  
study with prover E and SUMO.*

## **Bachelorarbeit**

verfasst am

**Institut für Software Engineering und Programming Languages**

im Rahmen des Studiengangs

**Informatik**

der Universität zu Lübeck

vorgelegt von

**Julian Britz**

ausgegeben und betreut von

**Prof. Dr. Diedrich Wolter**

mit Unterstützung von

**Moritz Bayerkuhnlein**

Lübeck, den 06. Juli 2025

### Eidesstattliche Erklärung

*Ich erkläre hiermit an Eides statt, dass ich diese Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.*

---

Julian Britz

Zusammenfassung  
Zusammenfassung.

Abstract  
Abstract.

# Inhaltsverzeichnis

1	Einleitung	1
2	Verwandte Arbeit	2
3	Behauptung	3
4	Vorwissen	5
4.1	Commen sense reasoning	5
4.2	Word Embeddings	5
4.3	Grammatiken	5
4.4	Theorembeweiser	5
5	Selektionsstrategien	7
6	Parallelen zwischen Sprachverarbeitung und Theorembeweisern	8
7	Experimente	9
8	Weiterführende Arbeit	22
9	Quellenverzeichnis	23

# 1

## Einleitung

- Bedeutung der logischen Schlussfolgerung im Bereich KI und der natürlichen Sprachverarbeitung
- Typische Ansätze
- Potenzial von semantischen Informationen zur Verbesserung der Auswahl von Axiomen
- Beschreibung von E als ein effektiver Theorembeweiser für die Aussagenlogik
- Adimen-SUMO als komplexe Wissensbasis für die Simulation und Bewertung von Schlussfolgerungsstrategien
- Ziel und Beitrag der Arbeit

# 2

## Verwandte Arbeit

- ALR12. ´Alvez, Javier, Paqui Lucio and German Rigau: Adimen-SUMO: Reengineering an Ontology for First-Order Reasoning. *Int. J. Semantic Web Inf. Syst.*, 8(4):80–116, 2012.
- Syntaktische Selektionsstrategien:
  - HV11. Hoder, Krystof and Andrei Voronkov: Sine Qua Non for Large Theory Reasoning. In Bjørner, Nikolaj S. and Viorica Sofronie-Stokkermans (editors): *Automated Deduction - CADE- 23 - 23rd International Conference on Automated Deduction*, Wroclaw, Poland, July 31 - August 5, 2011. *Proceedings*, volume 6803 of *Lecture Notes in Computer Science*, pages 299–314. Springer, 2011.
  - Meng, Jia and Lawrence C. Paulson: Lightweight relevance filter- ing for machine-generated resolution problems. *J. Appl. Log.*, 7(1):41– 57, 2009.
  - Roederer, Alex, Yury Puzis and Geoff Sutcliffe: Divvy: An ATP Meta-system Based on Axiom Relevance Ordering. In Schmidt, Renate A. (editor): *Automated Deduction - CADE-22, 22nd Interna- tional Conference on Automated Deduction*, Montreal, Canada, August 2-7, 2009. *Proceedings*, volume 5663 of *Lecture Notes in Computer Sci- ence*, pages 157–162. Springer, 2009.
- SInE: Roederer, Alex, Yury Puzis and Geoff Sutcliffe: Divvy: An ATP Meta-system Based on Axiom Relevance Ordering. In Schmidt, Renate A. (editor): *Automated Deduction - CADE-22, 22nd Interna- tional Conference on Automated Deduction*, Montreal, Canada, August 2-7, 2009. *Proceedings*, volume 5663 of *Lecture Notes in Computer Sci- ence*, pages 157–162. Springer, 2009.
- Semantisch: Sutcliffe, Geoff and Yury Puzis: SRASS - A Semantic Relevance Axiom Selection System. In Pfenning, Frank (editor): *Automated Deduction - CADE-21, 21st International Conference on Automated Deduction*, Bremen, Germany, July 17-20, 2007, *Proceedings*, volume 4603 of *Lecture Notes in Computer Science*, pages 295–310. Springer, 2007.

# 3

## Behauptung

Die semantische Nähe von Wörtern kann das Common Sense Reasoning unterstützen. Wieso? Grobe Beschreibung. Wie ist eine Grammatik strukturiert? Wie ist der Zusammenhang zwischen Logik und natürlicher Sprache?

Proover berücksichtigt nur die Signaturen und bringt solche in Verbindung, die in einer logischen Kette sind.

Wenn nun aber ein Begriff in dieser Kette nicht vorkommt, aber für den Beweis relevant ist, dann fehlt dies.

the semantic structure of ontologies often mirrors descriptions and categorizations found in natural language.

Ontologies might integrate or be constructed using existing language resources such as thesauri, taxonomies, or lexical databases (e.g., WordNet), which are themselves typically developed from natural language data.

Shared Semantic Goals:

Both natural language and ontologies aim to represent and convey meaning. While the methods and structures differ, they both operate within the realm of semantics—understanding and relating concepts and entities.

Complementary Roles:

Natural language is often the starting point for building ontologies, helping to identify and articulate the necessary concepts and properties. Ontologies, in turn, can support natural language processing by providing a structured, semantic backbone for interpreting text.

The integration of natural language semantic similarity measures into the axiom selection process of automated theorem provers significantly enhances their performance in commonsense reasoning tasks, bridging the gap between human-like understanding and machine logic through improved alignment of logical axioms with linguistic context.”

Behauptung: Logische Axiome tragen eine semantische Bedeutung, da sie einen Bezug zur realen Welt haben: Semantische Bedeutung: Auch wenn Axiome formal und symbolisch dargestellt werden, tragen sie eine semantische Bedeutung, da sie konzeptionelle Inhalte definieren, die mit bestimmten Bedeutungen in der realen Welt verbunden sind. Zum Beispiel könnte ein Axiom in einer Ontologie besagen, dass alle Menschen sterblich

### 3 Behauptung

sind, was eine direkte semantische Verbindung zu unserem Wissen über die Realität hat.  
Wie relevant ist die gröÙe der Dimension und die max. sequenz-length



# 4

## Vorwissen

### 4.1 Common sense reasoning

- Erklärung.
- Was ist das?
- Wofür braucht man es? Einsatzfelder
- Wie ist der aktuelle Stand?
- Herausforderungen

### 4.2 Word Embeddings

- Definition
- Erstellung
- Eigenschaften
- Anwendung
- Limitierung und Herausforderungen

### 4.3 Grammatiken

- Erklärung. Was sind Grammatiken und welche gibt es?
- Aufbau und Struktur
- Wofür benötigt man sie?
- SUMO

### 4.4 Theorembeweiser

- Erklärung. Was sind Theorembeweiser und welche gibt es?
- Funktionalität - wie werden Beweise gefunden - refutation
- Wofür braucht man sie?
- Grenzen und Herausforderungen

- Prover E
  - Wie arbeitet E? Wie geht E vor?
  - Satauto mode

# 5

## Selektionsstrategien

- Syntaktisch
- Semantisch
- Kombination - Union: wie übersetzt

# 6

## Parallelen zwischen Sprachverarbeitung und Theorembeweisern

- Ziel: finde Worte/ Axiome, die zum Kontext passen, um eine gewisse Aussage zu erzielen
- Mehrdeutigkeiten werden mittels Resolution gelöst
- Both represent knowledge and meaning
- Ontologies like Adimen-SUMO capture entities, relationships, and properties formally, analogous to how natural language uses words, sentences, and grammar to convey meaning.
- Just as natural language has grammatical and syntactical rules, Adimen-SUMO is built upon logical syntax and rules. However, logical syntax is rigid and unambiguous, while natural language often allows for more flexibility and nuance
- Difference: While natural languages do have grammar and syntax, they are more flexible and context-dependent, allowing for ambiguity and multiple interpretations, unlike the rigid, unambiguous nature of formal logic.
- In natural language, much of the meaning is implicit, derived from context, tone, and prior knowledge. Logical systems make everything explicit, ensuring that every inference is clear and derivable from stated axioms

# 7

## Experimente

- Whitebox Truth Test: wie werden diese erzeugt?  
[https://www.researchgate.net/publication/317230162\\_Automatic\\_White-Box\\_Testing\\_of\\_First-Order\\_Logic\\_Ontologies](https://www.researchgate.net/publication/317230162_Automatic_White-Box_Testing_of_First-Order_Logic_Ontologies)  
Falsity tests created according to Adimen Sumo -> negate them to get falsity test ->  
Prover does refutation

## – Standard

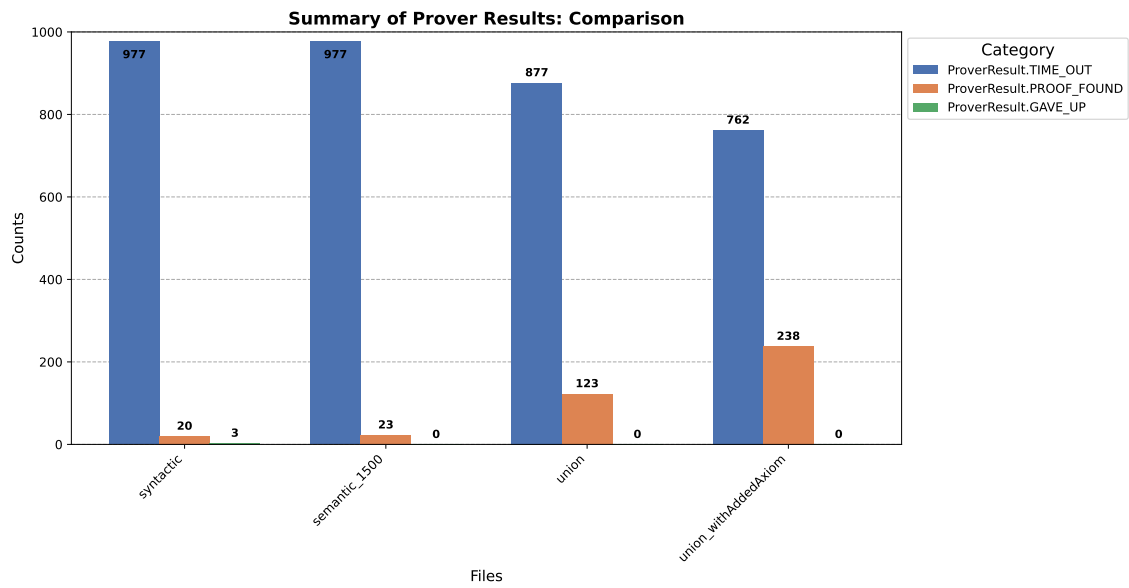


Abbildung 7.1: Summary of Prover Results: Standard Mode

## – Satauto

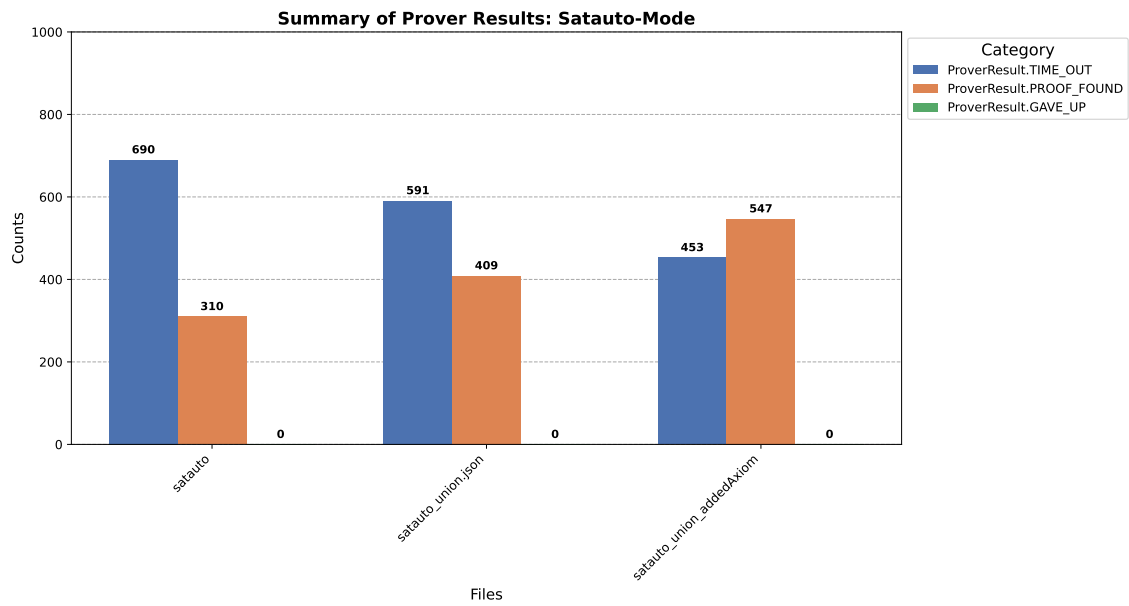


Abbildung 7.2: Summary of Prover Results: Satauto Mode

– Auto

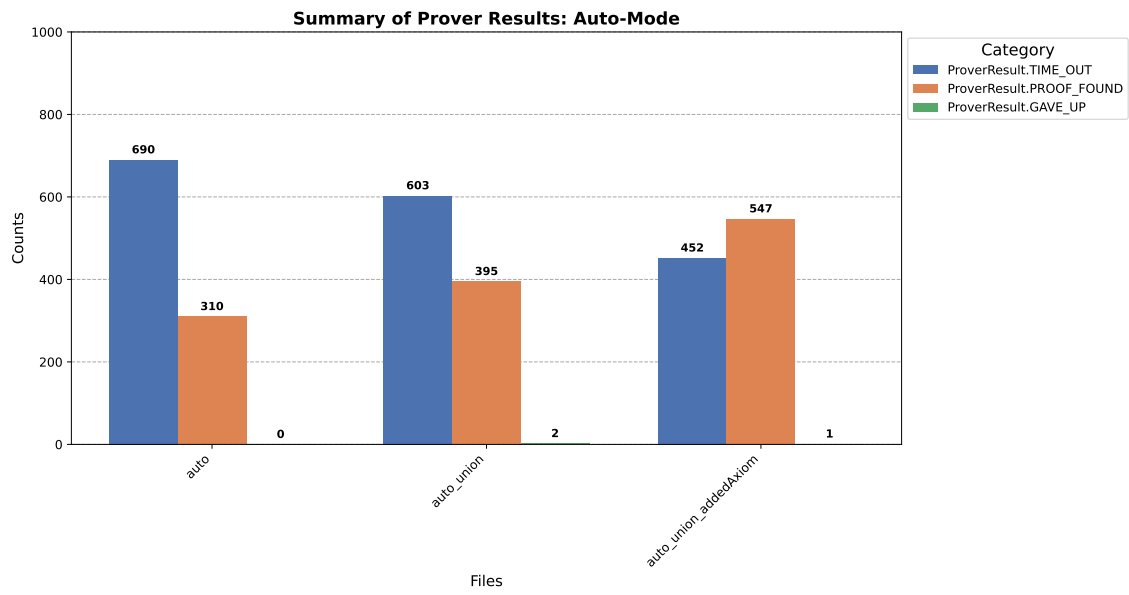


Abbildung 7.3: Summary of Prover Results: Auto Mode



## 7 Experimente

### – Modelle

- all-MiniLM-L6-v2
- mpnet-base-v2

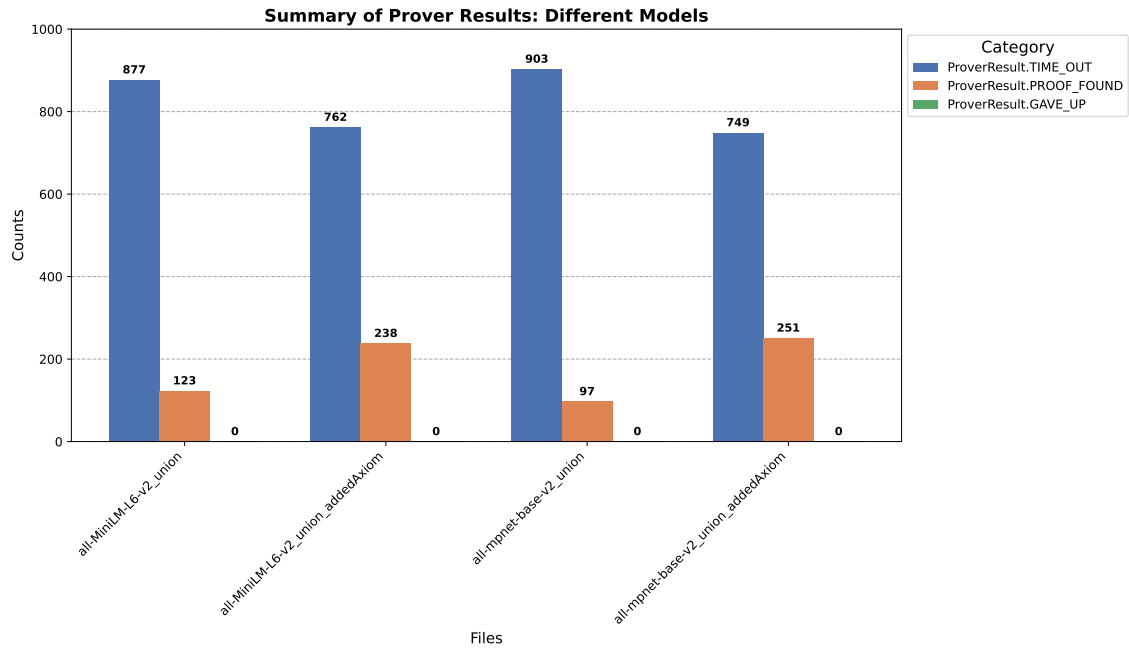


Abbildung 7.4: Summary of Prover Results: Model Mode

## – Vampire

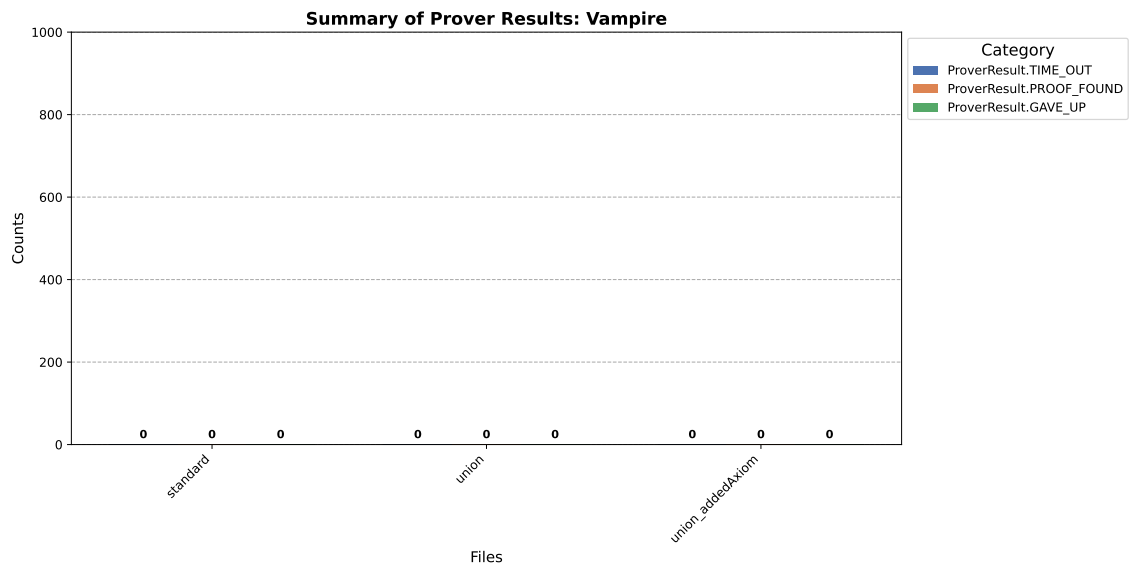


Abbildung 7.5: Summary of Prover Results: Vampire Mode

## 7 Experimente

- Durchschnittliche Kosinus-Ähnlichkeit der Axiome des Beweises und der Conjecture
- Statistiken
  - Mean variable count

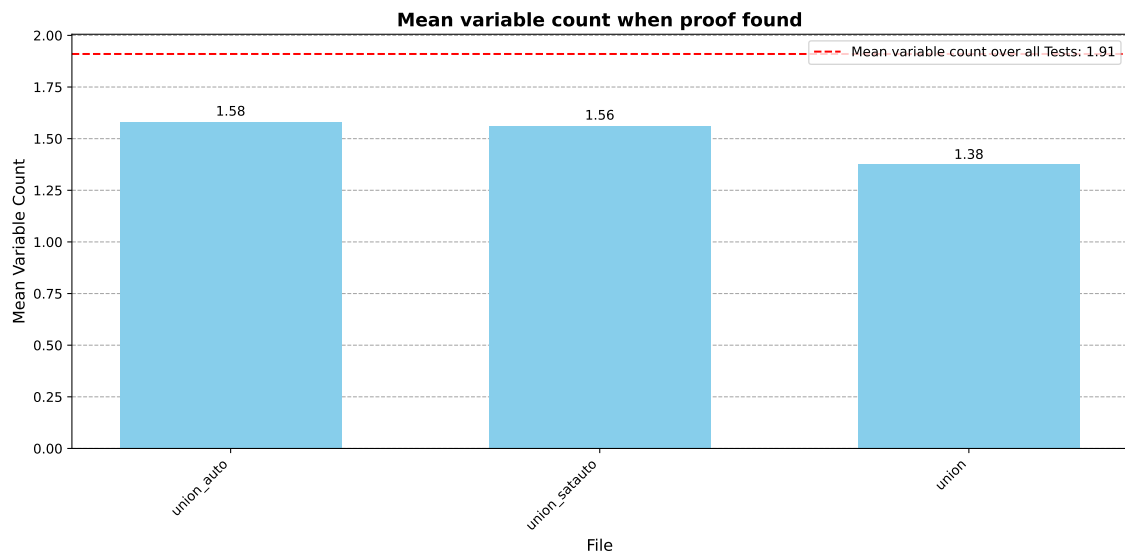


Abbildung 7.6: Summary of Prover Results: Variable Count

## 7 Experimente

### – Count signs

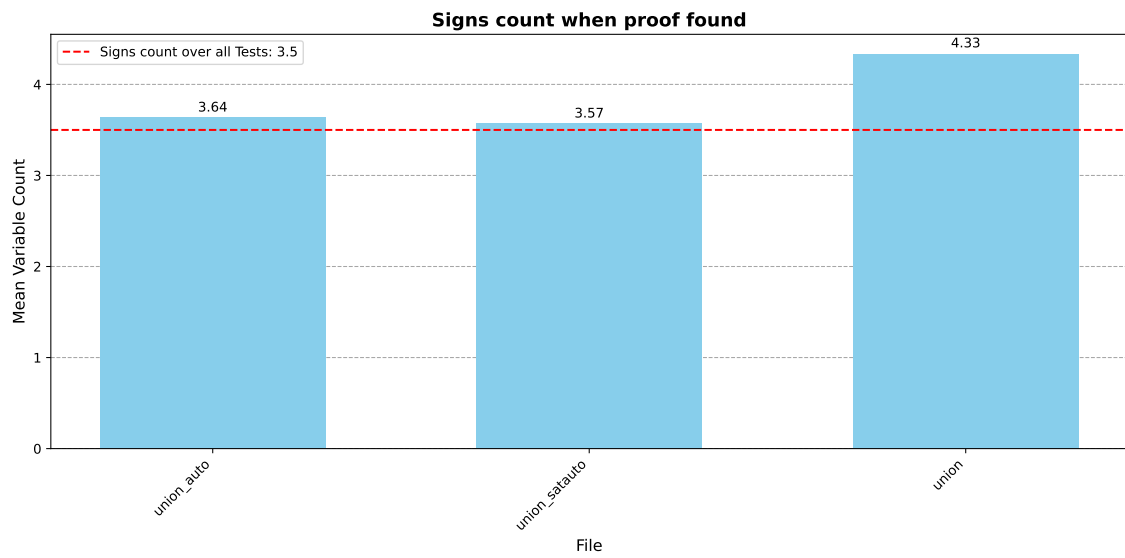


Abbildung 7.7: Summary of Prover Results: Signs Count

## – Character Count

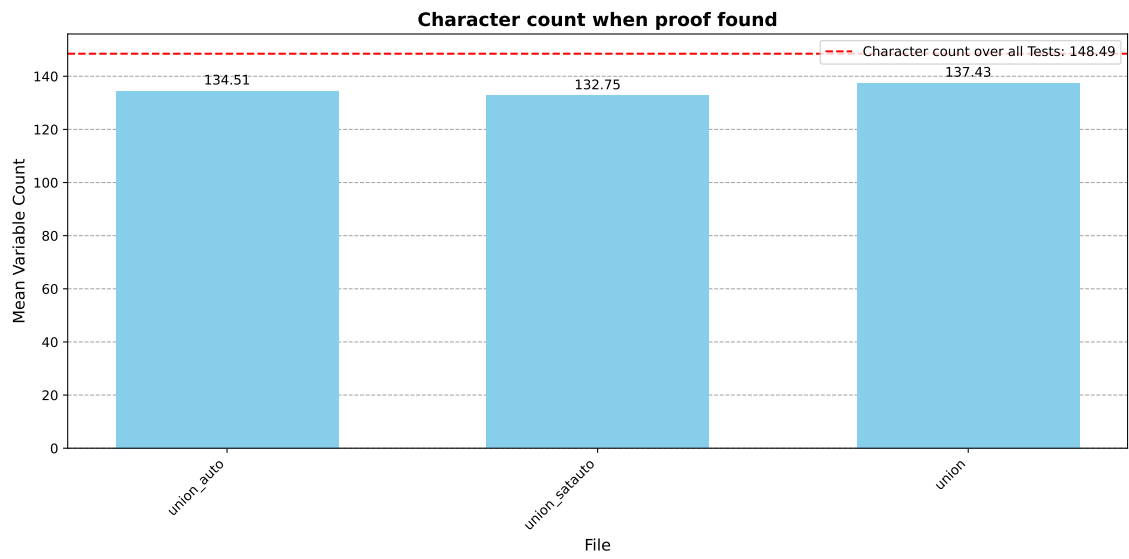


Abbildung 7.8: Summary of Prover Results: Character Count

## 7 Experimente

### – Proofs found in first named

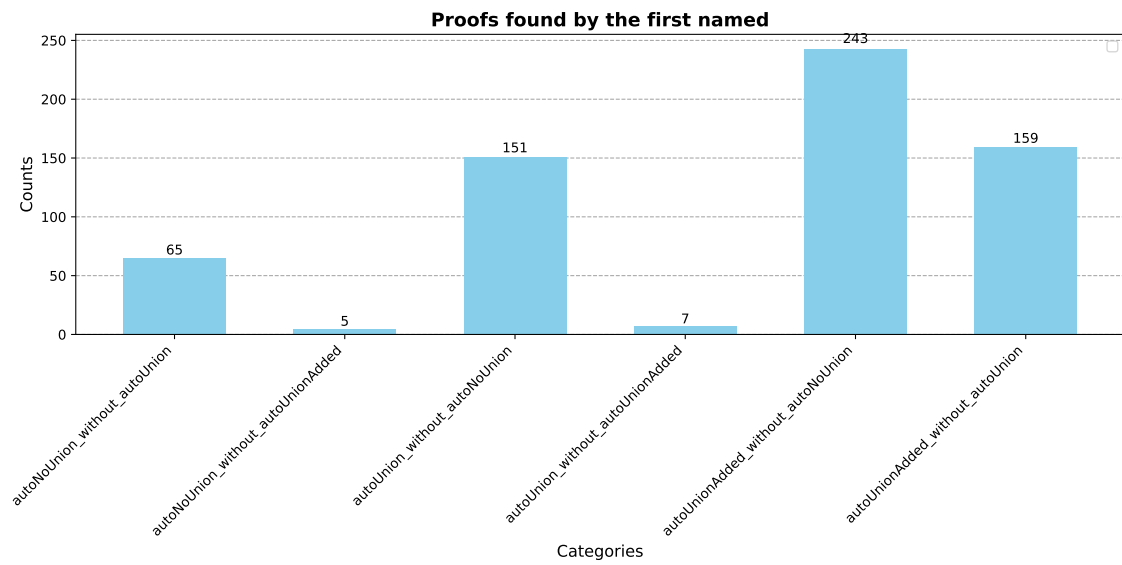


Abbildung 7.9: Summary of Prover Results: Proofs found

- Time to find proof

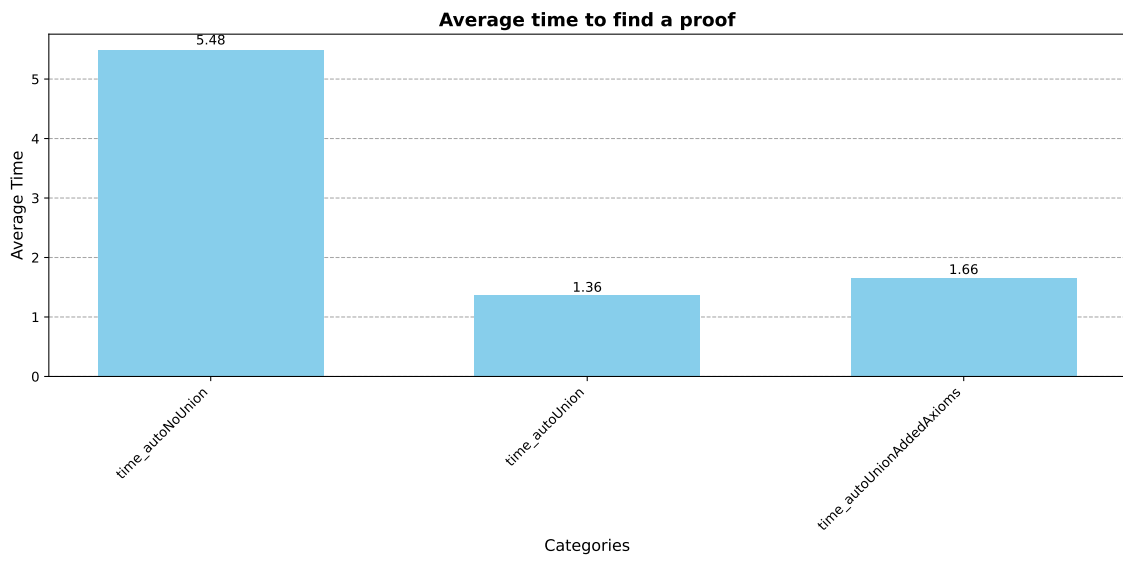


Abbildung 7.10: Summary of Prover Results: Proofs found time

## 7 Experimente

- Time to find common proof

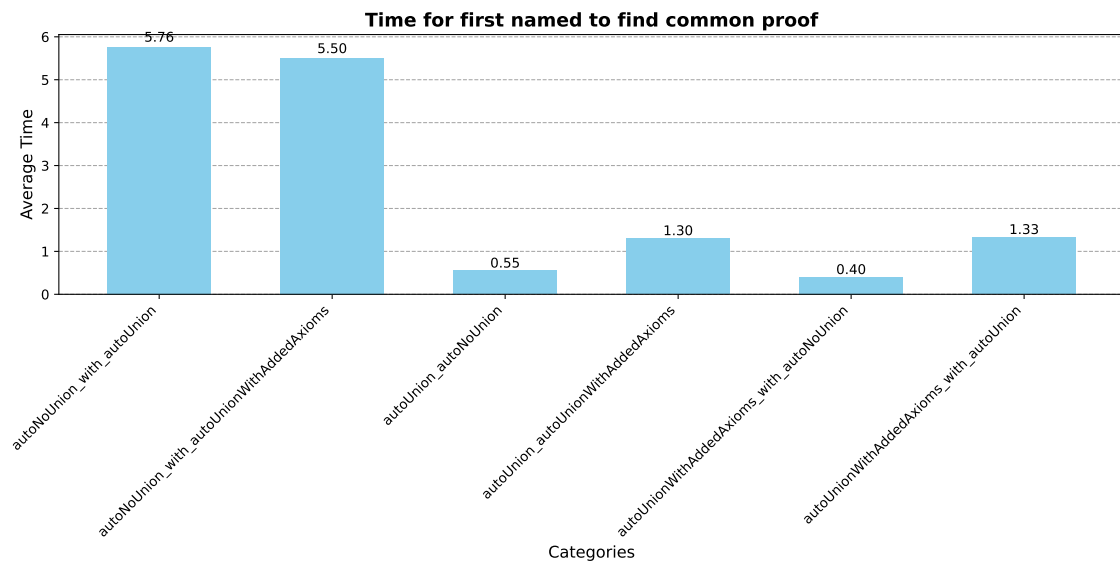


Abbildung 7.11: Summary of Prover Results: Common proofs found time



- Conclusion
- Conclusion

# 8

## Weiterführende Arbeit

- Wahl der richtigen Axiome durch neuronales Netz

# 9

## Quellenverzeichnis