Enhancing Argumentation Structure Parsing with Deep Probabilistic Answer Set Programming

Jonas Rodrigues Lima Gonçalves
Advisee of Prof. Denis Deratani Mauá

Denis Deratani Mauá
Advisor

April 20, 2025

1 Introduction

Argumentation Mining (AM) is a significant area within Natural Language Processing (NLP) focused on extracting argumentative structures from text. Stab and Gurevych have contributed significantly to this field by proposing a pipeline approach for parsing argumentation structures in persuasive essays. Their method involves segmenting the extraction process into several sub-steps, including identifying argument components, classifying their types, and identifying argumentative relations. A crucial final step in their pipeline utilizes Integer Linear Programming (ILP) to enforce coherence between the outputs of these sub-steps and generate a globally consistent argumentation graph.

While the ILP approach offers a way to integrate local predictions into a global structure, it can face limitations in terms of capturing complex probabilistic dependencies and being seamlessly integrated with end-to-end differentiable learning. Recent advancements in neural-symbolic reasoning, particularly in Deep Probabilistic Logic Programming frameworks like dPASP, offer promising alternatives. dPASP extends Answer Set Programming (ASP) with neural predicates, probabilistic choices, and tight integration with deep learning frameworks like PyTorch, enabling end-to-end training through automatic differentiation.

This thesis project initially aimed to investigate the potential of replacing the integer programming-based coherence enforcement step in Stab and Gurevych's argumentation parsing pipeline with dPASP. However, following discussions, the project will now adopt a phased approach. The first phase will focus on implementing the coherence enforcement using Scallop, a more efficient system, in collaboration with ongoing work by Vinícius. This will involve training the necessary models for argument component identification, node classification, and edge classification using the Stab and Gurevych dataset.

The second phase, contingent on the progress of related work on dPASP 2.0, will explore a comparison with a dPASP-based implementation utilizing circuit-based inference. This phase aims to leverage advancements in dPASP, potentially incorporating additional constraints related to abstract argumentation theory.

The overarching goal is to contribute to the development of more effective neural-symbolic approaches for argumentation mining and to evaluate the capabilities of next-generation dPASP systems.

2 Objectives and Goals

The primary objective of this thesis project is to **design, implement, and evaluate a logic-based module for enforcing coherence in an argumentation structure parsing pipeline, drawing inspiration from the work of Stab and Gurevych**. This will be done in two phases, initially focusing on Scallop and subsequently exploring dPASP.

The specific goals of this project include:

- **Understanding the Existing Pipeline:** Thoroughly analyze the argumentation structure parsing pipeline proposed by Stab and Gurevych, with a specific focus on the role and implementation of the integer linear programming step. This includes understanding the input to this step (outputs of the component identification, classification, and relation identification models) and its output (the coherent argumentation graph).
- **Implementing a Scallop-based Coherence Model (Phase 1):** Develop a logic program in Scallop that can take the outputs of the preceding pipeline stages (or their representations) as input and enforce coherence rules for the argument structure. This will involve collaborating with Vinícius and leveraging his work on Scallop plugins.
- **Training Necessary Models (Phase 1):** Train the machine learning models required for argument component identification, node classification, and edge classification using the Stab and Gurevych dataset. This step is crucial for providing input to the Scallop-based coherence module.
- **Integrating the Scallop Module (Phase 1):** Implement the integration of the Scallop module into the existing argumentation parsing pipeline, potentially using publicly available resources. This will involve handling the interface between the machine learning classifiers and the Scallop framework.
- **Evaluating the Scallop-enhanced Pipeline (Phase 1):** Evaluate the performance of the modified argumentation parsing pipeline, comparing it to the original approach that uses integer programming. This evaluation will likely involve using the argumentation structure annotated corpus and standard evaluation metrics for argumentation mining.

- **Conceptualizing a dPASP-based Coherence Model (Phase 2 - Tentative):** Explore the development of a probabilistic logic program in dPASP that can take the probabilistic outputs of the preceding pipeline stages (or their representations) as input and reason about the coherent structure of the argument. This may involve defining neural predicates.
- **Implementing and Evaluating the dPASP Module (Phase 2 - Tentative):** Implement the dPASP module and compare its performance with the Scallop-based approach, potentially leveraging advancements in dPASP 2.0, such as compilation techniques using SDDs and decision sdDNNFs.
- **Analyzing the Results:** Analyze the strengths and weaknesses of both the Scallop and (if implemented) dPASP-based approaches, identifying potential improvements and discussing the implications for the field of argumentation mining and neural-symbolic reasoning.

3 Planning Major Activities

The project will be conducted through the following major activities:

- **Phase 1: Scallop Implementation and Evaluation (Months 1-5):**
  - **Month 1:** In-depth study of Stab and Gurevych's work and familiarization with the argumentation structure annotated corpus.
  - **Month 2:** Comprehensive review of the Scallop framework and collaboration with Vinícius on the plugin development.
  - **Months 3-4:** Training the BERT-based models for argument component identification, node classification, and edge classification.
  - **Month 5:** Implementation and integration of the Scallop-based coherence module and evaluation of the enhanced pipeline.
- **Phase 2: dPASP Exploration and Comparison (Months 6-7 - Tentative):**
  - **Month 6:** Study of dPASP 2.0 advancements (compilation with SDDs and decision sdDNNFs) and conceptualization of a dPASP-based coherence model.
  - **Month 7:** Implementation of the dPASP module and comparative evaluation with the Scallop-based approach (if dPASP 2.0 is sufficiently mature).
- **Final Phase: Analysis and Thesis Writing (Months 7-8):**
  - In-depth analysis of the results from both phases.
  - Discussion of the findings and their implications.
  - Writing and finalizing the thesis document.

4 Methodology

This project will primarily employ a logic-based approach for the initial implementation using Scallop, with a potential secondary exploration of a neural-symbolic approach using dPASP. The core methodology will involve:

- **Replication and Adaptation:** Starting with the well-defined pipeline of Stab and Gurevych, the project will focus on replicating the overall structure while specifically replacing the ILP-based coherence enforcement with logic-based reasoning in Scallop.
- **Logic Programming with Scallop:** The coherence of the extracted argument components and relations will be modeled using Scallop, a highly efficient deductive reasoning system. Logic rules will be defined to capture the constraints and desired properties of a coherent argumentation structure, such as acyclicity and the single sink property of the argument graph.
- **Integration with Machine Learning Models:** The Scallop module will be integrated with the output of pre-trained (or trained as part of this project) machine learning models for argument component identification, classification, and relation identification.
- **Neural-Symbolic Exploration with dPASP (Tentative):** If time and the maturity of dPASP 2.0 allow, a parallel or subsequent implementation using dPASP will be explored. This would involve defining probabilistic logic programs, potentially with neural predicates, to model coherence.
- **Experimental Evaluation:** The performance of the Scallop-enhanced pipeline will be rigorously evaluated on the annotated corpus of persuasive essays by Stab and Gurevych, using standard evaluation metrics. If a dPASP-based approach is implemented, its performance will also be evaluated and compared.

5 Expected Outcomes and Contributions

This thesis project is expected to yield the following outcomes and contributions:

- **A Scallop-based module for enforcing coherence in argumentation structure parsing.** This module will demonstrate the application of efficient logic programming to a challenging NLP task.
- **An evaluation of the effectiveness of Scallop compared to integer programming for ensuring coherence in argumentation parsing.** This comparison will provide insights into the strengths and weaknesses of logic-based approaches for this task.

- **Potential improvements in the accuracy and efficiency of argumentation structure parsing.** By leveraging the efficiency of Scallop, the project may lead to a more scalable approach to coherence enforcement.
- **A deeper understanding of the integration of machine learning and logical reasoning for argumentation mining.** The project will contribute to the field by exploring a practical application of logic programming for coherence enforcement.
- **A documented implementation and evaluation of the proposed approach**, which could serve as a foundation for future research in this area, potentially including a comparison with dPASP.

6 Timeline

Time Period	Major Activities
Month 1	Literature Review and Background Study
Month 2	Scallop Framework Study and Collaboration
Months 3-4	Training Machine Learning Models
Month 5	Scallop Module Implementation and Evaluation
Month 6	dPASP Study and Conceptualization (Tentative)
Month 7	dPASP Implementation and Comparison (Tentative)
Months 7-8	Analysis and Thesis Writing

7 Required Resources

- **Computational Resources:** Access to adequate computational resources, including GPUs, for training and evaluating deep learning models.
- **Software Resources:** Python programming environment, PyTorch for deep learning, and the Scallop framework. Access to additional libraries for NLP tasks. If Phase 2 is pursued, access to the dPASP framework.
- **Data Resources:** The annotated corpus of persuasive essays created by Stab and Gurevych.
- **Academic Resources:** Access to relevant scientific literature.

8 Risks and Mitigation Strategies

- **Risk:** Technical challenges in integrating the Scallop module with the machine learning pipeline. **Mitigation:** Start with a clear interface definition and modular design. Thoroughly test the integration at each step.

- **Risk:** The Scallop-based approach may not perform as well as the original ILP approach. **Mitigation:** Focus the evaluation on understanding the specific strengths and weaknesses of the logic-based approach.

- **Risk:** Time constraints may prevent the full exploration of a dPASP-based comparison. **Mitigation:** Prioritize the Scallop implementation and evaluation. The dPASP phase will be contingent on the progress and available time.

- **Risk:** Difficulty in training accurate machine learning models for the initial stages of the pipeline. **Mitigation:** Leverage pre-trained BERT models and focus on fine-tuning them on the Stab and Gurevych dataset.

9 Conclusion

This thesis project will investigate the application of logic programming, specifically using the Scallop framework, for enhancing argumentation structure parsing. By replacing the integer linear programming step in Stab and Gurevych's pipeline with a Scallop-based approach, the project aims to leverage the efficiency and declarative power of logic for ensuring coherence. A secondary exploration of dPASP may be conducted depending on the progress of related work. The expected outcomes include a novel implementation, empirical evaluation, and insights into the integration of machine learning and logical reasoning for this challenging NLP task. The results of this project could potentially lead to more efficient and effective approaches to argumentation mining.