From the reconstruction of enthymemes to structured datasets. Introducing YargMiner, an AI-built pipeline for extracting AMT Y-structures from the literature.*

Andrea Rocci[1,*], Carlo Torniai[2], Rudi Palmieri[3], and Costanza Lucchini[1]

[1] *Università della Svizzera Italiana, via G. Buffi 13 6900 Lugano, Switzerland*
[2] *OWITH.ai, Barcelona, Spain*
[3] *University of Liverpool, Brownlow Hill, Liverpool L69 7ZXS, United Kingdom*

Abstract
The Argumentum Model of Topics (AMT) is a theory of argument schemes which offers also an analytic tool for reconstruction and diagramming of the inferential configuration of enthymemes in context. Over the past twenty years, the AMT has been deployed in numerous qualitative studies in different domains. In order to advance the AMT approach and formulate empirical generalizations, we need to overcome the challenge of pooling results from a quite dispersed body of research. To this purpose, the present short paper proposes YargMiner, an AI-built pipeline for automatically extracting AMT analyses from existing literature. YargMiner extracts the components of the AMT diagrams (known as Y-structures) together with their key contextual information to produce a simple tabular dataset. We tested the pipeline on a small sample of AMT literature containing analyses of arguments in the finance domain. The resulting dataset is used to illustrate how the proposed literature mining method can be used to study the empirical coverage of the inventory of loci proposed by AMT and to systematize the formulation of maxims associated with each locus.

Keywords
argument schemes, AMT, loci, maxims, literature mining, AI built pipeline

1. Introduction

Theories of argument schemes, such as the Argumentum Model of Topics (AMT) [13, 15] or Douglas Walton's approach to the schemes [4, 20], fulfil the double function of guiding the reconstruction of enthymemes and of scaffolding the evaluation of arguments.

One key challenge for these theories is the need to empirically assess proposed *inventories*, *taxonomies* and *scheme components* in terms of their coverage and explanatory power against a broad sample of argumentative discourse from different contexts and genres. Meeting this challenge with a corpus annotation approach may prove problematic: annotation of argument schemes requires expert annotators, is cognitively demanding and time consuming (cf. [2]). In fact, as observed by Rocci and Lucchini [18], a pure *annotation approach* is put under strain by

---

CMNA'26: The 26th International Workshop on Computational Models of Natural Argument, Sept 14, Barcelona, Spain
* Corresponding author.
andrea.rocci@usi.ch (A. Rocci); carlo@owith.ai (C. Torniai); Rudi.Palmieri@liverpool.ac.uk (R. Palmieri); costanza.lucchini@usi.ch (C. Lucchini)
0000-0001-7156-5402 (A. Rocci); 0000-0002-3734-1859 (C. Torniai); 0000-0002-5122-3058 (R. Palmieri); 0009-0005-5901-8306 (C. Lucchini)

argument schemes, as they rather require a *reconstruction approach* where the reduction of enthymematic arguments to a canonical form can involve the extensive *addition* of implicit material and a significant *reconfiguration* of the observable linguistic form. This style of analysis makes comparison of results, generalization over cases and the incremental advance of research more challenging.

Corpus annotation projects inspired by AMT ([6], [19]) have been limited in scope and, more importantly, limited to using its taxonomy of semantic relations (*loci*) as labels (e.g. *from part and whole*, *from efficient cause*), leaving out the theory's most interesting feature, i.e. a diagrammatic layout of a single argumentative inference (called *Y-structure*).

This paper presents a method for automatically extracting AMT Y-structure diagrams from the literature, along with their corresponding text excerpts and associated macrostructural analyses.

2. The AMT analysis of argumentative inference: the Y-structure

To briefly illustrate how the AMT analysis of argument schemes works, let us consider the following enthymeme:

Standpoint: 1. The price of crude oil will remain high for several months
Argument: 1.1 The war in Iran is still there.

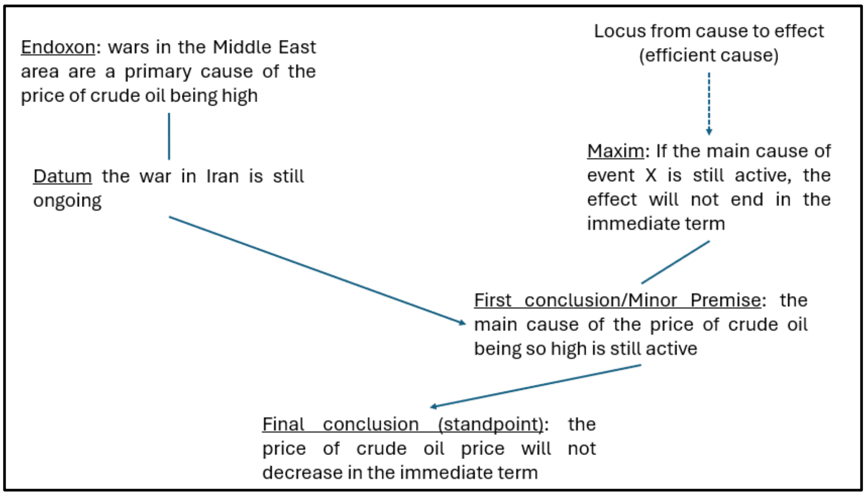

Figure 1: Diagrammatic representation (Y-structure) of the inferential structure of the enthymeme according to an AMT reconstruction.

Figure 1 shows the so-called *Y-structure*, which diagrams the inferential configuration of an enthymematic argument based on the AMT. The Y-structure diagram integrates two distinct types of premises: (1) the *topical component*, including *locus* and *maxim*, indicates the inference warrant underlying the argument-standpoint (premise-conclusion) connection; (2) the *material component*, which includes the *endoxon* and the *datum*, respectively referring to generalizations, background knowledge and shared assumptions shared in the relevant argumentative context and to situation-related facts, statements, information. The two components are integrated by a *logical form* consisting of two interlocked syllogisms so that the conclusion of the *material syllogism* becomes the minor premise of the *topical syllogism.*

Following AMT, each *locus* relation is associated with several maxims. For example, the locus from efficient cause appearing in the example entails several other maxims, such as ‘if the efficient cause of x has ceased, then x will cease to occur’. Two methods for the discovery of maxims have been proposed [15]. The first is based on the semantic-conceptual analysis of the locus relation. The second method is “heuristic” and consists in discovering new maxims through the analysis of “concrete cases of argumentation”. However, while the activity of concrete analysis has been intense, the lack of systematic collection and comparison of the maxims used by analysts in dispersed qualitative analyses has hindered the development of AMT.

3. Methods

We developed YargMiner[2], which comprises two complementary components: (1) an extraction-and-annotation pipeline that ingests academic PDFs and, through a sequence of loosely coupled stages, locates candidate figures, classifies those that are Y-structures, and extracts their components together with the source excerpt and contextual information they refer to, storing every output in a shared database; (2) an annotation-review web application that reads from the same store and presents each extracted Y-structure alongside its original figure, allowing a domain expert to assess annotation quality, correct individual fields, and promote reviewed annotations to a gold-standard reference, all without writing code.

The system is organized as four loosely coupled Python modules that share a single SQLite store and can be invoked independently or chained through one command.

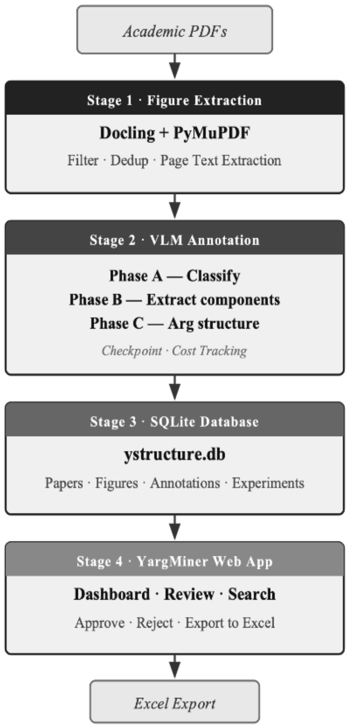

Figure 2: End-to-end data flow across the four pipeline stages.

3.1. Pipeline

Stage 1: figure extraction. Both Docling [1] and PyMuPDF run on every PDF. Docling detects figure regions from page-layout analysis and resolves captions by following internal cross-reference pointers, giving reliable caption–figure associations even when the two are far apart; PyMuPDF extracts embedded image objects directly and matches them to captions by proximity. The union of the two outputs is cleaned by a template-asset filter, which removes images of identical pixel size that recur on three or more pages without a caption (logos, headers, journal chrome), and by perceptual deduplication. The extractor also stores the full text of every page for use in Stage 2

Stage 2: three-phase VLM annotation. Each retained figure is sent to Claude Sonnet 4.5 (Anthropic) with the paper title, page number, caption and up to

[2] Details of the implementation can be found at: https://carlotorniai.github.io/YargminerNotes

200 characters of surrounding text. **Phase A** classifies the figure as *simple* (one Y-structure), *complex* (multiple linked Y-structures) or "not a Y-structure"; the system prompt specifies the six canonical components, the connector types, the expected number of boxes, multilingual label variants and explicit negative examples (IAT diagrams, attack/support graphs, flowcharts, photographs). **Phase B**, runs only on simple and complex Y-structures, asks the model to read the text in each labelled region and emit a Pydantic-validated record holding the six components, a *links* field for complex diagrams, the verbatim excerpt of the source discourse (drawn from a window of five preceding and one following page of stored text), and contextual metadata (topical domain, genre, language, modality). **Phase C** searches the same window for *diagrams* or *textual* representations of *argumentation structure*, which typically precede the Y-structure. It first scans the page text and falls back, only when nothing is found, to rendering candidate pages as PNG images. Cross-contamination between adjacent Y-structures is prevented by instructing the model to match the candidate argumentation structure against the extracted Datum/Final Conclusion, and an explicit anti-fabrication rule forbids reconstructing notation that is not literally present in the paper.

Stage 3: SQLite store. All outputs flow into one database file. Each run is recorded as an experiment; annotations carry a review-status field (pending, approved, needs-revision, rejected).

Stage 4: review web application. YargMiner is a local FastAPI + Jinja2 + htmx application [12]. A dashboard summarizes the corpus, the paper-detail view shows every extracted figure with its classification tag, and a split-pane review page places editable annotation fields beside the original figure so the expert can approve, reject or flag each Y-structure.

3.2. Dataset

We assembled a first evaluation corpus from eight scholarly sources that apply AMT to arguments in text and dialogue in the business and financial communication domains. The sources include journal articles [9–11], chapters from books [7, 8, 16, 17], and a thesis [3] co-authored by some of the authors of this study. Together the sources provide an evaluation set of 86 Y-structures, established as a gold reference by manual review.

4. Evaluation

One of the authors compared the pipeline's output against the 86 Y-structures expected to be recognized, in order to evaluate *detection* and then used YargMiner Web App to review the correctness of each of the fields extracted against the original diagram to evaluate *annotation quality*. Against the gold set of 86 Y-structures, the pipeline correctly extracted 85, missing a single one, and produced no false positives: every figure it retained is a genuine Y-structure. This gives a broad detection precision of 100.0% and a recall of 98.8% (85/86), for an F1 of 99.4%.

Turning from detection to annotation quality, of the 85 correctly extracted Y-structures 74 were approved with no changes at all and 11 required the revision of one or more annotation fields. Counting only the untouched records as correct gives a strict accuracy of 86.0% (74/86); counting the revisable records as usable gives a broad accuracy of 98.8% (85/86).

The review of the annotation revealed that the six core Y-structure components (Endoxon, Datum, First Conclusion, Locus, Maxim and Final Conclusion) were almost always recognized correctly. The majority of the corrections made to the imperfect annotations concern failures in correctly retrieving the argumentation structure associated with the examples and the original text of the example being analyzed.

5. Exploratory analysis of the mini-dataset of Y-structures

We conducted an exploratory analysis of the 74 Y-structures that YargMiner annotated correctly. While the dataset is very small, the advantages of systematically reviewing Y-structures in the AMT literature can already be gleaned.

A quick review of the *locus* field in the 74 entries shows no less than 40 different labels are used, none of them used more than 7 times. This reveals, first of all, a lack of notational consistency, even among a small sample of a close-knit community of users of the theory. Some of the notational inconsistencies are trivial (capitalization, use of articles, presence or absence of the word *locus*) and can be quickly normalized away in a revised version of the pipeline. Some other aspects of variation in naming *loci* are, however, less trivial. For instance, *locus from final cause to action* (instead of simply *from final cause*), incorporates the direction of the locus in the label. Other labels append the specification of a *sub-locus*, such as *common opinion, expert opinion, position to know, from horse's mouth, expert advice* as sub-loci of the *locus from authority.* Consistently with Rigotti's [14] examination of the rich ontology of action, the *locus from the final/ instrumental cause* presents an extremely rich variety of sub-loci highlighting the nodes of the ontology of action that are leveraged in the inferential relation (*from knowledge to action*, *from goal to action*, *from intention to side effects*, etc.). A qualitative review of the proposed sub-loci, especially for very productive loci such as *final/ instrumental cause* and *authority* appears to be a particularly worthwhile endeavor for advancing the theory. Another interesting case is represented by composite labels combining two loci, such as *locus from final / instrumental cause (+ termination and setting up).* To the authors' knowledge, this is common practice in AMT analyses: many forms of practical reasoning are analyzed as combinations of *final/ instrumental cause* with extrinsic loci such as *alternatives* or *termination and setting up.* Subsequent versions of YargMiner will need to properly represent sub-loci and combinations of loci so that the standardization of labels is achieved without information loss and the study of maxims associated with sub-loci and combinations is possible.

Table 1
Frequency of loci in the "Finance" mini-dataset (74 accepted Y-structures). Terminological variants and sub-loci are grouped together.

Locus	Frequency
From final / instrumental cause	28
From authority (incl. testimony)	14
From reliable procedure of knowledge acquisition	8
From analogy	7
From ontological implication(s)	6
From oppositions	4
From alternatives	3
From efficient cause	2
From definition (genus and species)	1
From parts and whole	1

Table 1 displays the distribution of loci in the 74 Y-structures once labels are standardized and sub-loci and second loci are discounted. Of the 20 *loci* proposed by Rigotti and Greco [15], 9 occur in the dataset, plus one additional locus (*From reliable procedure of knowledge acquisition*), closely related with authority, which was later proposed by Miecznikowski [5]. The three categories of *intrinsic*, *extrinsic* and *complex* loci are all represented, but the distribution is highly skewed in favor of two *loci*: *final/ instrumental cause* and *authority* (with the closely related *reliable procedure*). This perfectly matches Palmieri's qualitative observations [8] about finance as an arena of *practical decision* characterized by knowledge asymmetries and a *high reliance on experts* and partially opaque models and procedures.

As for maxims, a vast new landscape opens up. While Rigotti and Greco [15] also allowed for maxims without a conditional form (e.g. maxims formulated as an exclusive disjunction), all 74 maxims in the sample are formulated as conditionals (74 contain *if*, 70 start with *if*, 24 contain an *if…then* sequence). When we examine the maxims of the most frequent loci, strikingly different pictures emerge. The maxims associated with the *locus from authority* are remarkably similar, and neatly map onto the sub-loci (expert opinion, testimony, etc.). Different formulations referring to the same sub-locus are mere terminological variants. With *final/ instrumental cause* the picture is quite different: *goal to action* maxims with the structure *If action A is a (necessary) means to reach a goal G, then A should be undertaken* are common (10), but the other 18 occurrences of the locus display a strikingly diverse array of maxims, inviting the analyst to go back to conceptual and semantic analysis to put order in the set.

6. Implications and next steps

The implementation and evaluation of YargMiner on a small corpus of AMT literature demonstrated the feasibility of the proposed method of literature mining for the creation of datasets of Y-structures, while the exploratory analysis of the mini-dataset obtained was sufficient to suggest the interest of such datasets for the theoretical development of AMT and an assessment of its empirical coverage of different domains of argument and discourse genres.

In particular, in the area of the definition of sub-loci and the formulation of maxims, the dataset allows to create a virtuous circle between the two avenues of theoretical advancement envisaged by Rigotti and Greco [15], moving back from the analysis of concrete cases to conceptual-semantic analysis via comparative analysis. At the very least, the examination of extensive datasets of Y-structures will allow AMT to refine its metalanguage and identify gaps in theorizing.

The next step in the development of YargMiner will involve a *refinement of its annotation output*, optimizing it for the analysis of sub-loci and composite locus Y-structures as well as *improvements to the reviewer's interface*. We plan to apply the pipeline to a representative sample of AMT literature with a source count of about 100. This will likely produce a dataset of Y-structures ranging between 500 and 1000 entries. Such a dataset could be used then not only for systematic review and theory development, but also for the testing and training of LLM-based pipelines for the automated reconstruction of Y-structure analyses from argumentative discourse units.

Declaration on Generative AI

During the preparation of this work, the authors used Claude Sonnet 4.5 (Anthropic) in order to: generate code for the technical pipeline described in the Methods section and draft accompanying technical notes documenting its implementation. The authors reviewed, summarized and adapted the resulting description of the technical implementation in the Methods section, and take full responsibility for the content of the publication.

References

[1] C. Auer, M. Lysak, A. Nassar, M. Dolfi, N. Livathinos, P. Vagenas, C. Berrospi Ramis, M. Omenetti, F. Lindlbauer, K. Dinkla, L. Mishra, Y. Kim, S. Gupta, R. Teixeira de Lima, V. Weber, L. Morin, I. Meijer, V. Kuropiatnyk, P. W. J. Staar, Docling technical report, arXiv:2408.09869 (2024). doi:10.48550/arXiv.2408.09869.

[2] E. Bezou-Vrakatseli, O. Cocarascu, S. Modgil, EthiX: A dataset for argument scheme classification in ethical debates, in: U. Endriss, F. S. Melo, K. Bach, A. Bugarín-Diz, J. M. Alonso-Moral, S. Barro, F. Heintz (Eds.), ECAI 2024: 27th European Conference on Artificial Intelligence, volume 392 of Frontiers in Artificial Intelligence and Applications, IOS Press, 2024, pp. 3628-3635. doi:10.3233/FAIA240919.

[3] C. Lucchini, Evidentials as argumentative indicators in the argumentative activity type of earnings conference calls, Ph.D. thesis, Università della Svizzera italiana, Lugano, 2026.

[4] F. Macagno, D. Walton, Classifying the patterns of natural arguments, Philosophy & Rhetoric 48 (2015) 26-53. doi:10.5325/philrhet.48.1.0026.

[5] J. Miecznikowski, At the juncture between evidentiality and argumentation: Evidential verb complementation, Journal of Argumentation in Context 9 (2020) 42-68. doi:10.1075/jaic.00007.mie.

[6] E. Musi, T. Alhindi, M. Stede, L. Kriese, S. Muresan, A. Rocci, A multi-layer annotated corpus of argumentative text: From argument schemes to discourse relations, in: N. Calzolari, K. Choukri, C. Cieri, T. Declerck, S. Goggi, K. Hasida, H. Isahara, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, S. Piperidis, T. Tokunaga (Eds.), Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018), European Language Resources Association, Miyazaki, Japan, 2018, pp. 1629-1636. URL: https://aclanthology.org/L18-1258/.

[7] R. Palmieri, Corporate Argumentation in Takeover Bids, volume 8 of Argumentation in Context, John Benjamins, Amsterdam/Philadelphia, 2014. doi:10.1075/aic.8.

[8] R. Palmieri, The role of argumentation in financial communication and investor relations, in: A. V. Laskin (Ed.), The Handbook of Financial Communication and Investor Relations, John Wiley & Sons, Chichester, UK, 2017, pp. 45-60. doi:10.1002/9781119240822.ch4.

[9] R. Palmieri, Reasons for trust: The (counter-)argumentative dynamics of image-repair strategies, Journal of Pragmatics 240 (2025) 142-163. doi:10.1016/j.pragma.2025.02.014.

[10] R. Palmieri, S. Mazzali-Lurati, Practical reasoning in corporate communication with multiple audiences, Journal of Argumentation in Context 6 (2017) 167-192. doi:10.1075/jaic.6.2.03pal.

[11] R. Palmieri, E. Musi, Trust-repair strategies in crisis rhetorical (sub-)arenas: An argumentative perspective, International Journal of Strategic Communication 14 (2020) 272-293. doi:10.1080/1553118X.2020.1805452.

[12] S. Ramírez, FastAPI: Modern, fast (high-performance) web framework for building APIs with Python, 2018. URL: https://fastapi.tiangolo.com/.

[13] E. Rigotti, Relevance of context-bound loci to topical potential in the argumentation stage, Argumentation 20 (2006) 519-540. doi:10.1007/s10503-007-9034-2.

[14] E. Rigotti, Locus a causa finali, Analisi Linguistica e Letteraria 16 (2009) 559–576.

[15] E. Rigotti, S. Greco, Inference in Argumentation: A Topics-Based Approach to Argument Schemes, Springer International Publishing, Cham, 2019. doi:10.1007/978-3-030-04568-5.

[16] E. Rigotti, R. Palmieri, Analyzing and evaluating complex argumentation in an economic-financial context, in: C. Reed, C. W. Tindale (Eds.), Dialectics, Dialogue and Argumentation: An Examination of Douglas Walton's Theories of Reasoning and Argument, College Publications, London, 2010, pp. 85-99.

[17] A. Rocci, Modality in Argumentation: A Semantic Investigation of the Role of Modalities in the Structure of Arguments with an Application to Italian Modal Expressions, volume 29 of Argumentation Library, Springer, Dordrecht, 2017. doi:10.1007/978-94-024-1063-1.

[18] A. Rocci, C. Lucchini, Diagramming the enthymematic structure of counterarguments: An introduction to IAMT diagrams, Argumentation et Analyse du Discours 34 (2025). doi:10.4000/13q0e.

[19] N. Sheikh Asadi, A. Rocci, Revealing argumentative patterns underlying the genre-specific moves in research article introductions: A corpus-based investigation of applied linguistics journals, Journal of Argumentation in Context, online first (2026). doi:10.1075/jaic.26006.she.

[20] D. Walton, C. Reed, F. Macagno, Argumentation Schemes, Cambridge University Press, Cambridge, 2008. doi:10.1017/CBO9780511802034.