AF-Xray: Visual Explanation and Resolution of Ambiguity in Legal Argumentation Frameworks

Yilin Xia
University of Illinois, Urbana-Champaign
yilinx2@illinois.edu

Heng Zheng
University of Illinois, Urbana-Champaign
zhenghz@illinois.edu

Shawn Bowers
Gonzaga University, Spokane, WA
bowers@gonzaga.edu

Bertram Ludäscher
University of Illinois, Urbana-Champaign
ludaesch@illinois.edu

Abstract

Argumentation frameworks (AFs) provide formal approaches for legal reasoning, but identifying sources of ambiguity and explaining argument acceptance remains challenging for non-experts. We present AF-XRAY, an open-source toolkit for exploring, analyzing, and visualizing abstract AFs in legal reasoning. AF-XRAY introduces: (i) layered visualizations based on game-theoretic argument length revealing well-founded derivation structures; (ii) classification of attack edges by semantic roles (primary, secondary, blunders); (iii) overlay visualizations of alternative 2-valued solutions on ambiguous 3-valued grounded semantics; and (iv) identification of critical attack sets whose suspension resolves undecided arguments. Through systematic generation of critical attack sets, AF-XRAY transforms ambiguous scenarios into grounded solutions, enabling users to pinpoint specific causes of ambiguity and explore alternative resolutions. We use real-world legal cases (e.g., Wild Animals as modeled by Bench-Capon) to show that our tool supports teleological legal reasoning by revealing how different assumptions lead to different justified conclusions.

CCS Concepts

• **Computing methodologies → Knowledge representation and reasoning**; • **Applied computing → Law, social and behavioral sciences**; • **Human-centered computing** → *Visualization*; *Human computer interaction (HCI)*; • **Theory of computation** → Logic.

Keywords

Argumentation Frameworks, Legal Reasoning, Interactive Visualization, Explainable AI

ACM Reference Format:
Yilin Xia, Heng Zheng, Shawn Bowers, and Bertram Ludäscher. 2025. AF-Xray: Visual Explanation and Resolution of Ambiguity in Legal Argumentation Frameworks. In *Twentieth International Conference on Artificial Intelligence and Law (ICAIL 2025), June 16–20, 2025, Chicago, IL, USA*. ACM, New York, NY, USA, 3 pages.

ACM ISBN 979-8-4007-1939-4

1 Introduction

Abstract argumentation frameworks [8] offer well-established, formal approaches for representing and reasoning about case law [2]. Given an argument x in an argumentation framework (AF), it is easy to determine the status of x under skeptical reasoning, i.e., whether x is *accepted* (IN), *defeated* (OUT), or *undecided* (UNDEC). In case of the latter, the AF is *ambiguous*: it has a 3-valued grounded semantics S_0, and some conflicts may require additional assumptions or choices to be made to resolve these ambiguities. *Value-based* and *Extended* AFs have been used in legal reasoning to resolve and justify the acceptance in such scenarios [3]. These approaches help users explain choices among alternative resolutions by discounting (or ignoring) certain attack edges, e.g., based on social value preferences. For AF non-experts, however, it can be difficult to pinpoint the specific reasons (i.e., *critical attacks*) causing an ambiguity, and to visualize an AF's semantics in a way that all parties understand.

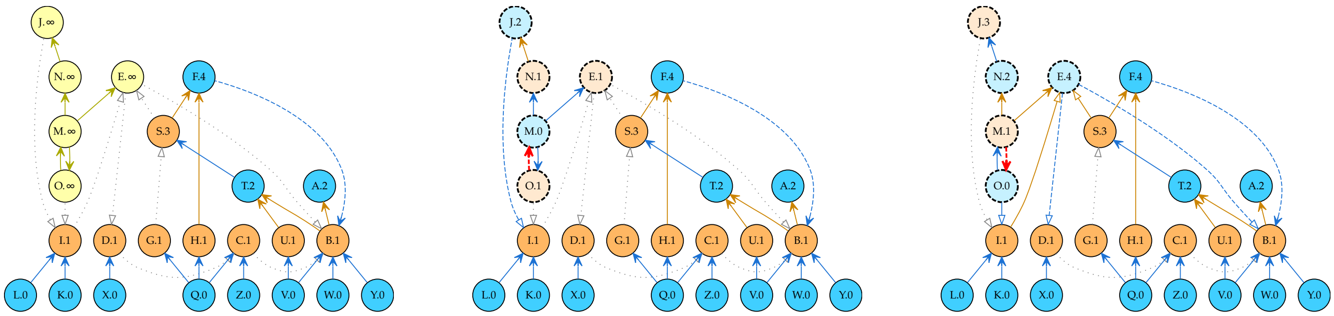
We present AF-Xray[1], a novel platform for exploring, analyzing, and visualizing AF solutions, which builds upon the state-of-the-art open source PyArg system [9]. Xray "looks deeper" into the structure of AFs and provides new analysis and visualization components for explaining the acceptance of arguments under *skeptical* reasoning, and for identifying critical attacks, whose suspension *resolves* undecided arguments under *credulous* reasoning. It adds:

(*i*) A novel *layered* AF visualization, based on the game-theoretic *length*[2] (or *remoteness* [10]) of nodes [5]; (*ii*) a novel *classification of attack edges* derived from their game-theoretic type [5]; (*iii*) the ability to switch between alternate 2-valued solutions $S_1, \ldots, S_n$ of an AF (visualized as *overlays* on the ambiguous, 3-valued S_0); and (*iv*) the identification and display of *critical attacks* in Δ_i for each solution S_i, where $\Delta_i = \{\Delta_{i,1}, \ldots, \Delta_{i,n_i}\}$ are n_i critical attack sets $\Delta_{i,j}$ for S_i: Temporarily suspending the attacks in $\Delta_{i,j}$ yields a 2-valued grounded solution $S'_{i,j}$. Together, the suspension of $\Delta_{i,j}$ and the resolution $S'_{i,j}$ explain the choices for S_i.

2 AF-Xray in Action

In Xray, similar to PyArg, users input AFs as graphs $G = (V, E)$ of arguments V and attack edges $E \subseteq V \times V$, either via a web interface or file upload. The input graph is then visualized. Users pick a semantics (e.g., *grounded*, *stable*, *preferred*) and select one of the possible *solutions* (labelings). Arguments are colored according

[1] AF-Xray: **A**rgumentation **F**ramework e**X**planation, **R**easoning, and **A**nal**Y**sis [12]
[2] … which is closely related to *min-max numberings* of *strongly admissible sets* [7]



(a) Layered *skeptical* (grounded) solution S_0 (b) Overlay *credulous* (stable) resolution $S'_{1,1}$ (c) Overlay *credulous* (stable) resolution $S'_{2,1}$

Figure 1: AF-Xray visualizations of the *Wild Animals* cases [1]: (a) The ambiguous (3-valued) grounded solution S_0 uses the *length* of nodes: e.g., F.4 requires no more than four discussion rounds to prove that F is IN. Distinct *edge types* are used to account for their semantic roles [5]. The *overlays* in (b) and (c) represent alternative resolutions $S'_{1,1}$ and $S'_{2,1}$: The UNDEC nodes E, J, M, N, O in (a) have been *decided* (M is IN and O is OUT in S_1; and vice versa in S_2). These choices are explained by *critical attacks* (red edges) $\Delta_1 = \{\Delta_{1,1} : \{O \to M\}\}$ and $\Delta_2 = \{\Delta_{2,1} : \{M \to O\}\}$, i.e., minimal sets of (temporarily) *suspended* edges: when suspensions are applied, 2-valued grounded solutions $S'_{1,1}$, $S'_{2,1}$ are obtained for S_1 and S_2.

to their status: IN (blue), OUT (orange), and UNDEC (yellow). The following highlight some of Xray's features.

Layered Visualizations. Fig. 1a shows the *Wild Animals* legal example using Xray's layered visualization. The layering is based on the *length* of argument nodes, which can be computed alongside the grounded labeling S_0 [5, 6, 11]. In S_0, an argument x that is OUT has an IN-labeled attacker; x is IN if every attacker of x is OUT; and UNDEC if x is neither IN nor OUT in S_0. In the layered visualization, the bottom layer consists of arguments that are trivially labeled IN because they have no attackers (length = 0); the next layer consists of OUT arguments (length = 1) that are defeated by length-0 attackers, etc. UNDEC arguments result from unfounded attack-chains (length = "∞"), and are displayed outside the layering. Arguments that justify an IN or OUT-labeled argument x are located at layers below x: e.g., while F.4 in Fig. 1a attacks B.1, the defeat of B.1 is known (due to V.0, W.0, and Y.0) *before* F.4's label is determined. The layering makes the *well-founded* (and thus "self-explanatory") derivation structure of the grounded semantics explicit.

Visualizing Attack Types. Xray visualizes attacks according to their role in determining argument labels [5]. Successful (blue) attacks are classified as either *primary* (solid blue) or *secondary* (dashed blue). Secondary attacks point to arguments with smaller lengths, e.g., F's attack on B, whose defeat was established in a lower layer. Dotted gray edges are "*blunders*", i.e., an edge type which is irrelevant for the acceptance status (*provenance*) of arguments [5]. A minimal explanation of an argument excludes secondary attacks and blunders, so they are de-emphasized in the visualization.

Resolving Ambiguity. To analyze and disambiguate the UNDEC portion of a 3-valued grounded solution, a less skeptical 2-valued semantics (e.g., stable or preferred) can be employed by Xray to enumerate these alternative solutions. Each solution represents a choice for resolving the (direct or indirect) circular conflicts that created the ambiguities (UNDEC nodes in Fig. 1a) in the first place. The two solutions in Fig. 1b & 1c are depicted as hybrid *overlays* of the 3-valued grounded solution S_0 (with UNDEC nodes) and the respective 2-valued stable solution S_i (without UNDEC nodes): The colors (IN/OUT-labels) of the stable solutions S_i are visualized "on top of" the grounded solution S_0, i.e., they share the same layered visualization, but now with UNDEC arguments colored according to their (newly resolved) acceptance status in S'_i. In such overlays, lighter colors and dashed outlines mark the original UNDEC subgraph.

Explaining Credulous Solutions in Xray. The grounded solution S_0 of an AF (Fig. 1a) is self-explanatory: IN, OUT, and UNDEC arguments are justified by their well-founded derivation and the *length*[3] used to rank nodes in the layered visualization [6, 13]. The explanatory structure of credulous (e.g., stable) solutions is more complex, however. It consists of a well-founded part (blue/orange nodes in Figure 1) and an ambiguous part (yellow nodes in Fig. 1a). A large number of alternative 2-valued solutions S_i usually "hide" in the ambiguous parts of S_0. In Xray, these choices can be explained via sets of *critical attacks* $\Delta_{i,j}$. If we choose to suspend these minimal sets of edges (e.g., via temporary deletions), every previously UNDEC argument x will be either IN or OUT, and for the chosen suspension $\Delta_{i,j}$, there is now a well-founded derivation of x. In this way, Xray allows the user to pinpoint critical attacks and arguments to support a desired outcome within the confines of the initial grounded solution. This approach facilitates new use cases for legal reasoning that complement earlier approaches such as Value-based and Extended AFs [3]. Whereas the latter assume that users already know which edges to attack, Xray systematically generates all such sets of critical edges, thus providing a deeper semantic analysis than any state-of-the-art system we are aware of.

Demonstration Overview. The demonstration will illustrate: (1) loading an AF with legal annotations of abstract arguments; (2) the layered visualization of the grounded solution S_0, observing the well-founded derivations of arguments (Fig. 1a); (3) exploration of different stable solutions S_i and their overlays $S'_{i,j}$, observing critical attack sets that explain the choices (suspensions) made (Fig. 1b, 1c) as part of the resolution; and (4) exporting the desired (re)solutions for future use. Legal argument annotations (hovering over a node displays its annotation; clicking on it navigates to a page with details) are used to discuss a real-world example: We

[3] The node length can be computed as a by-product of computing the well-founded model via the *alternating fixpoint procedure* [11].

study the mutual attack between two arguments: M (*mere pursuit is not enough*) and O (*bodily seizure is not necessary*), which directly reflects opposing arguments in *Pierson v. Post* [1]. Users can toggle between stable solutions S_1 and S_2 and view the critical attack sets $\Delta_1 = \{\Delta_{1,1}\}$ and $\Delta_2 = \{\Delta_{2,1}\}$ explaining each legal *possible world*. This supports the teleological structure of legal reasoning: Different assumptions lead to different legally justified conclusions, e.g., depending on which social values are prioritized [4].

References

[1] T. Bench-Capon. 2002. Representation of Case Law as an Argumentation Framework. In *15th Annual Conferece on Legal Knowledge and Information Systems (JURIX 2022)* (London, UK, December 16–17). IOS Press, Amsterdam, 103–112.

[2] T. Bench-Capon. 2020. Before and after Dung: Argumentation in AI and Law. *Argument & Computation.* 11, 1–2 (2020), 221–238.

[3] T. Bench-Capon and S. Modgil. 2009. Case law in extended argumentation frameworks. In *12th International Conference on Artificial Intelligence and Law (ICAIL 2009)* (Barcelona, Spain, June 8–12). Association for Computing Machinery, New York, NY, 118–127.

[4] D. Berman and C. Hafner. 1993. Representing Teleological Structure in Case-Based Legal Reasoning: The Missing Link. In *4th International Conference on Artificial Intelligence and Law (ICAIL 1993)* (Amsterdam, the Netherlands, June 15–18). Association for Computing Machinery, New York, NY, 50–59.

[5] S. Bowers, Y. Xia, and B. Ludäscher. 2024. The Skeptic's Argumentation Game or: Well-Founded Explanations for Mere Mortals. In *5th International Workshop on Systems and Algorithms for Formal Argumentation (SAFA 2024)* (Hagen, Germany, September 17). CEUR-WS, Aachen, 104–118.

[6] S. Bowers, Y. Xia, and B. Ludäscher. 2024. On the Structure of Game Provenance and its Applications. In *16th International Workshop on Theory and Practice of Provenance (TaPP 2024)* (Vienna, Austria, July 12). 602–609.

[7] M. Caminada and P. Dunne. 2019. Strong admissibility revisited: Theory and applications. *Argument & Computation* 10, 3 (2019), 277–300.

[8] P.M. Dung. 1995. On the Acceptability of Arguments and Its Fundamental Role in Nonmonotonic Reasoning, Logic Programming and n-Person Games. *AI* 77, 2 (1995), 321–357.

[9] D. Odekerken, A.M. Borg, and M. Berthold. 2023. Demonstrating PyArg 2.0. In *7th Workshop on Advances in Argumentation in Artificial Intelligence Advances in Argumentation in AI (AI^3 2023)* (Rome, Italy, November 9). CEUR-WS, Aachen.

[10] C. Smith. 1966. Graphs and Composite Games. *J. Comb. Theory* 1, 1 (1966), 51–81.

[11] A. Van Gelder. 1993. The Alternating Fixpoint of Logic Programs with Negation. *J. Comput. System Sci.* 47, 1 (1993), 185–221.

[12] Y. Xia, S. Bowers, and B. Ludäscher. 2025. AF-Xray: **A**rgumentation **F**ramework e**X**planation, **R**easoning, and **A**nal**Y**sis. https://github.com/idaks/xray.

[13] Y. Xia, D. Odekerken, S. Bowers, and B. Ludäscher. 2024. Layered Visualization of Argumentation Frameworks. In *10th International Conference on Computational Models of Argument (COMMA 2024)* (Hagen, Germany, September 18–20). IOS Press, Amsterdam, 373–374.