Persuasion by Inches and Non-Local Concessions

Bertram Ludäscher[1], Shawn Bowers[2]

[1]University of Illinois, Urbana-Champaign, USA
[2]Gonzaga University, Spokane, USA

Abstract
We model persuasion dialogues as movements between solutions (*complete labelings*) of a given argumentation framework. Each party begins at a maximally committed position (a *preferred labeling*), and the goal is to bring the other party to their own position. Moving from one position to another requires choosing a path through the solution space. We measure the difficulty (cost) of such a path not by the sum of label changes, but by the largest single step of a path: the *persuasion incline*. As with sequential-request techniques (e.g., "foot-in-the-door"), breaking a large disagreement into smaller stages can reduce the largest concession. Sometimes, however, a small concession on a *disputed* argument can unsettle arguments the parties already *agreed* on, creating a "ripple effect" that can make even the best path more costly. We ask a natural-argument question: do human arguers notice when concessions carry these indirect costs, and do effective persuasion strategies account for them?

Keywords
abstract argumentation, persuasion, complete labelings, argumentation motifs, paths of least resistance

1. Planning Dinner

Ada and Ben are choosing dinner. The chef offers one of **F**ish, **M**eat, or **P**asta, with either **W**hite or **R**ed wine; the mains are mutually exclusive, as are the wines, and white wine only goes with fish. The argumentation framework on the left in Fig. 1 has been used to study counterfactual explanations [AGPT24]. It has four menu solutions (preferred labelings): {F+W, F+R, M+R, P+R}. Ada wants M+R, Ben wants P+R, so they agree about the wine (red: yes, white: no) and fish (no), but differ on meat vs. pasta. As the two menu choices differ on only two labels, one might expect a short transition. Yet neither can move to the other's position without, at some point, putting fish back on the table: changing the selected main leaves all three mains undecided at the intermediate waypoint. Their agreement about fish rests on the very issue they dispute, so part of the cost is paid on an argument they do *not* disagree about.

2. From Positions to Paths

An *argumentation framework* (AF) is a finite set of arguments with a binary attack relation [Dun95]. A *complete labeling* assigns each argument IN (accepted), OUT (rejected), or UNDEC (judgment suspended) such that an argument is IN iff all its attackers are OUT, and OUT iff at least one attacker is IN [CG09]. The *information order* "$\sqsubseteq$" organizes complete labelings by commitment: $L \sqsubseteq L'$ when every argument labeled IN or OUT in L has the same label in L'. The *grounded labeling* is the least element and the *preferred labelings* are the maximal elements under $\sqsubseteq$ (Fig. 1, right).

The *distance* between labelings L and L' counts the arguments they label differently. In the weighted *routing graph*, complete labelings (our *solutions*) are the nodes; they are adjacent when one is directly above or below the other with respect to $\sqsubseteq$. Each edge is weighted by the distance between its endpoints. Each path from P_0 to some P_n in the routing graph has a largest step, its *steepest incline*. The *persuasion incline* $\mathrm{inc}(P_0, P_n)$ is the smallest of these steepest inclines over all paths from P_0 to P_n.

Fig. 1 (right) is the routing graph for the dinner menu example (blue=IN, orange=OUT, yellow=UNDEC; $\mathsf{P}_1, \ldots, \mathsf{P}_4$ are menus, $\mathsf{C}_1, \ldots, \mathsf{C}_3$ complete labelings with UNDEC menu items). Each edge is a retraction or a commitment, depending on the direction it is followed. Because edges only connect adjacent nodes

CMNA'26: 26th Workshop on Computational Models of Natural Argument, Sept. 14–15, 2026, Barcelona (at COMMA'26)
ludaesch@illinois.edu (B. Ludäscher); bowers@gonzaga.edu (S. Bowers)

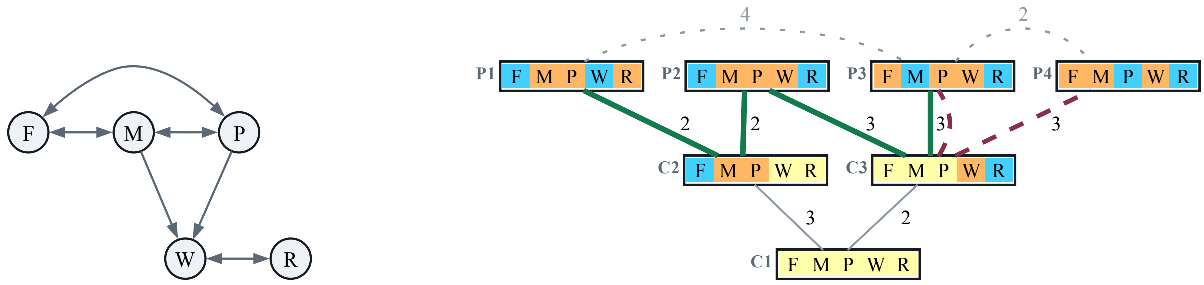

Figure 1: Menu AF, its complete labelings, two paths (*staging* solid **green**, *ripple* dashed **red**), and their endpoint distances (dotted).

in $\sqsubseteq$, IN never changes directly to OUT. Distance [BCPR12, DM18] alone does not capture the difficulty of persuasion, because a longer path can have a gentler steepest incline than a short one.

3. Staging Pays (Sometimes)

Consider another scenario: Ada now prefers F+W, i.e., fish with white wine (P_1); Ben prefers M+R, the meat with the red wine (P_3). The two positions differ on four labels. The green path P_1-C_2-P_2-C_3-P_3, in Fig. 1 first changes the wine while fish remains the main dish, then changes the main dish while red wine remains fixed. The largest steps along this path change only three labels. The path also includes P_2 (fish with red wine), which can serve as a compromise: the incline from P_1 to P_2 is 2, and from P_2 to P_3 is 3. Therefore, breaking a large disagreement into stages can reduce the largest step, similar to sequential-request techniques (cf. *foot-in-the-door* [FF66, Bur99]).

4. Concessions Are Not Local

Somewhat surprisingly (at first sight), the step cost can depend on arguments outside the disagreement. Assume Ada wants M+R (P_3) and Ben wants P+R (P_4). They assign opposite labels to meat and pasta but agree on the remaining arguments. To move from P_3 toward P_4, Ada must first retract her commitment to meat. But M attacks F, and F is OUT at P_3 only because M is IN. When meat is retracted, fish must also change from OUT to UNDEC. At the complete labeling C_3, all three main dishes are UNDEC, so the step from P_3 to C_3 changes three labels, not two. The step from C_3 to P_4 also changes three labels. The incline between P_3 and P_4 is therefore 3, even though the two positions differ on only two labels. The attack structure among the main dishes means that they cannot be changed independently. We call this situation a *ripple effect* (or *ripple* for short): fish changes labels during the transition even though both parties agree to reject it at their preferred positions.

This example shows that the persuasion incline is not determined by the disputed arguments alone. It can also depend on arguments outside the disagreement whose labels rely on disputed commitments. Retraction cascades are familiar from the early days of reason maintenance [Doy79, dK86]. Here they cause the incline to diverge from distance in a way that is not visible from the two positions alone.

5. Modeling Choices

The persuasion incline involves three modeling choices. *Allowed Moves:* a step must retract labels to UNDEC or commit UNDEC labels, but never flip IN and OUT directly. Alternatively, allowing a direct flip would let a party change disputed arguments directly without having to suspend them first (in this case, the ripple would disappear). However, these atomic flips may not be valid or acceptable in practice, and moves through intermediate concessions (with UNDEC labels) may be necessary. *Cost:* the weight of a step is the number of labels that change. Here, all changes count equally, whether retractions or commitments. *Aggregation:* the incline is the largest step, not the sum: the sum measures total effort, and with direct moves allowed it collapses to endpoint distance [BCPR12]; the largest single step is what a party refuses. Among paths with equal incline, path length and total cost can be used as tie-breakers.

6. Argumentation Motifs and Open Questions

Staging and the ripple effect are examples of *argumentation motifs*: patterns in an AF's attack structure that may affect the cost of persuasion paths. Here are two more: In *ballast*, two disputed arguments $A \leftrightarrow B$ both attack the same n additional arguments, hence rejected by both parties. The disagreement remains at two arguments, but the incline rises to $n+2$: each additional argument must change labels during the transition, just as fish does in the dinner example. In the *ridge* motif, both parties reject an argument whose only two attackers are disputed. Any path through a labeling where both attackers are UNDEC forces that argument to also become UNDEC (the same ripple mechanism). A path through a third preferred labeling can keep one attacker committed, avoiding the ripple and reducing the incline. An interactive demo [BL26] lets the reader explore persuasion incline examples and motifs.

Our model describes sequences of positions, not the dialogue acts that produce them. Any dialogue represented by such a sequence must respect the dependencies encoded in the AF [Pra06, HHP23, BC03]. This work suggests three questions for the CMNA community: (1) Do people prefer paths with a lower persuasion incline, even if the exchange is longer? (2) Can ripple effects be observed in debate or negotiation? (3) Can such a path be turned into natural argumentative moves?

Declaration on Generative AI. During the preparation of this work, the authors used AI to draft and reword portions of the text, to check grammar, and to verify examples. After each tool's use, the authors reviewed, edited, and revised all content and take full responsibility for the publication's content.

References

[AGPT24] Gianvincenzo Alfano, Sergio Greco, Francesco Parisi, and Irina Trubitsyna. Counterfactual and Semifactual Explanations in Abstract Argumentation: Formal Foundations, Complexity and Computation. In *KR*, pp. 14–26, 2024.

[BC03] Trevor J. M. Bench-Capon. Persuasion in practical argument using value-based argumentation frameworks. *Journal of Logic and Computation*, 13(3):429–448, 2003.

[BCPR12] Richard Booth, Martin Caminada, Mikołaj Podlaszewski, and Iyad Rahwan. Quantifying disagreement in argument-based reasoning. In *AAMAS*, pp. 493–500, 2012.

[BL26] Shawn Bowers and Bertram Ludäscher. The persuasion incline explorer. Interactive demo, https://idaks.github.io/tpi-demo/, 2026.

[Bur99] Jerry M. Burger. The foot-in-the-door compliance procedure: A multiple-process analysis and review. *Personality and Social Psychology Review*, 3(4):303–325, 1999.

[CG09] Martin W. A. Caminada and Dov M. Gabbay. A logical account of formal argumentation. *Studia Logica*, 93(2–3):109–145, 2009.

[dK86] Johan de Kleer. An assumption-based TMS. *Artificial Intelligence*, 28(2):127–162, 1986.

[DM18] Sylvie Doutre and Jean-Guy Mailly. Constraints and changes: A survey of abstract argumentation dynamics. *Argument & Computation*, 9(3):223–248, 2018.

[Doy79] Jon Doyle. A truth maintenance system. *Artificial Intelligence*, 12(3):231–272, 1979.

[Dun95] Phan Minh Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *AI*, 77(2):321–357, 1995.

[FF66] Jonathan L. Freedman and Scott C. Fraser. Compliance without pressure: The foot-in-the-door technique. *Journal of Personality and Social Psychology*, 4(2):195–202, 1966.

[HHP23] Emmanuel Hadoux, Anthony Hunter, and Sylwia Polberg. Strategic argumentation dialogues for persuasion: Framework and experiments based on modelling the beliefs and concerns of the persuadee. *Argument & Computation*, 14(2):109–161, 2023.

[Pra06] Henry Prakken. Formal systems for persuasion dialogue. *The Knowledge Engineering Review*, 21(2):163–188, 2006.