FAKULTÄT FÜR !NFORMATIK
Faculty of Informatics

dbai

Conflicts in Abstract Argumentation[1]

Christof Spanring

Department of Computer Science, University of Liverpool, UK

Institute of Information Systems, TU Wien, Austria

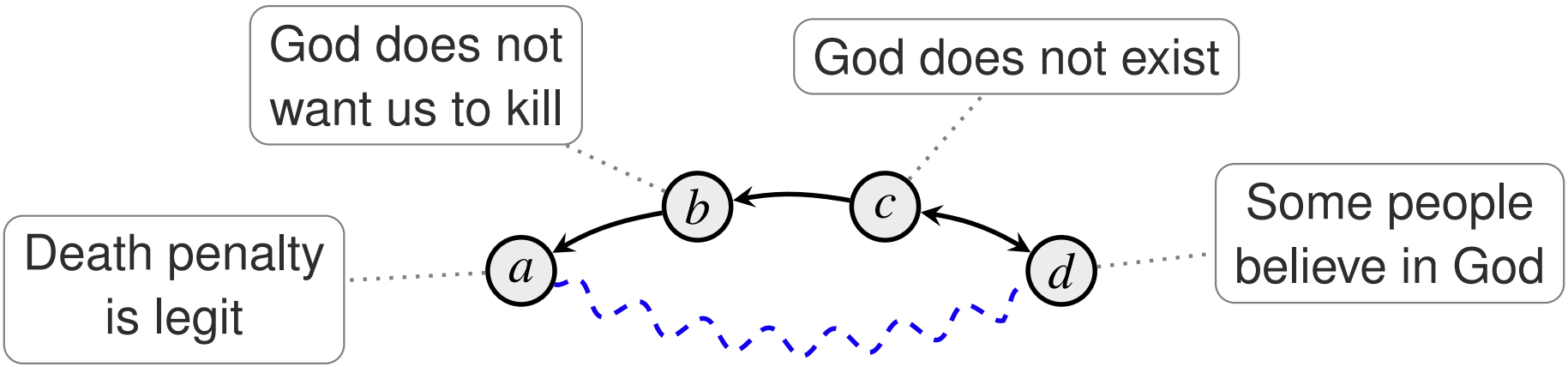
DBAI research seminar, November 10, 2016

FWF Der Wissenschaftsfonds.

UNIVERSITY OF LIVERPOOL

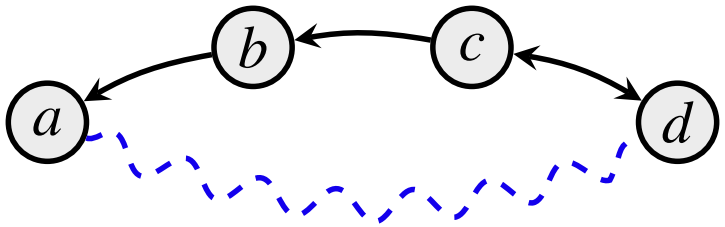
[1]This research has been supported by FWF (projects I1102 and I2854).

Argumentation

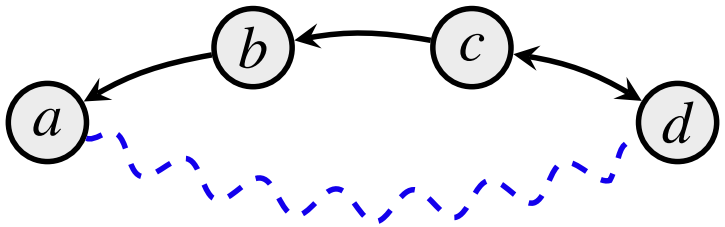


Natural Language Example, Is Death Penalty Legit?

Abstract Argumentation



Abstract Argumentation

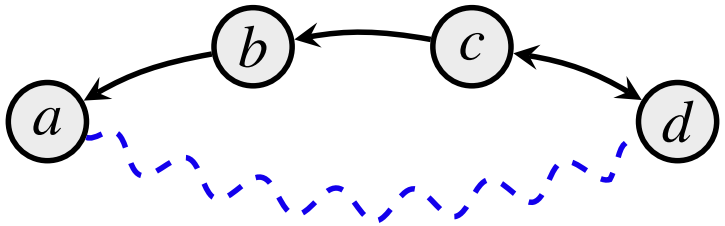


- Arguments: a, b, c, d
- Attacks: $(b, a), (c, b), (d, c), (c, d)$

Definition (Abstract Argumentation, Syntax)

Argumentation Framework (AF): $F = (A, R)$
A: set of arguments
$R \subseteq A \times A$: set of attacks

Abstract Argumentation



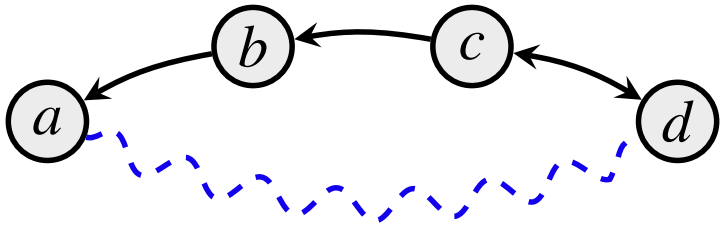
- Arguments: a, b, c, d
- Attacks: $(b, a), (c, b), (d, c), (c, d)$

- Conflicts: $[a, b], [b, c], [c, d]$

Definition (Syntactic Conflict and Compatibility)

Syntactic Conflict, $[X, Y]_F$: X attacks Y or Y attacks X
Syntactic Compatibility, $\{X, Y\}_F$: otherwise

Abstract Argumentation

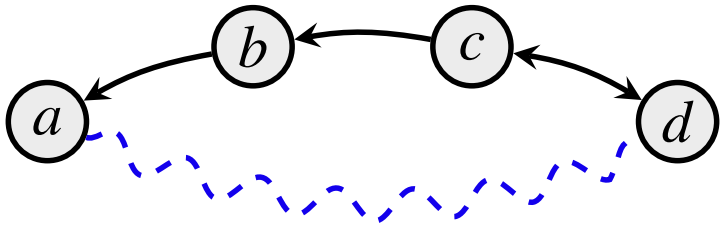


- Arguments: a, b, c, d
- Attacks: $(b, a), (c, b), (d, c), (c, d)$
- Extensions: $\{a, c\}, \{b, d\}$

Definition (Argumentation Semantics)

Conflict-freeness, $S \in \mathit{cf}(F)$: $\{S, S\}_F$
Stable Extension, $S \in \mathit{sb}(F) \subseteq \mathit{cf}(F)$: $A \setminus S = \{x \in A \mid S \textit{ attacks } x\}$

Abstract Argumentation

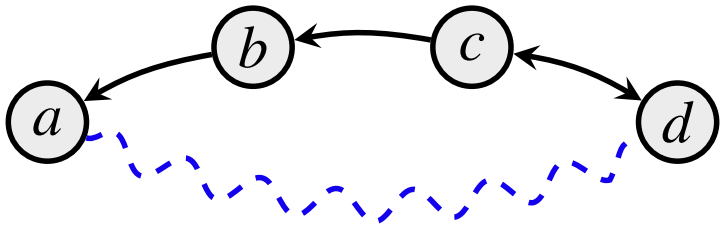


- Arguments: a, b, c, d
- Attacks: $(b, a), (c, b), (d, c), (c, d)$
- Extensions: $\{a, c\}, \{b, d\}$
- Conflicts: $[a, b], [b, c], [c, d], [a, d]$

Definition (Semantic Conflict and Compatibility)

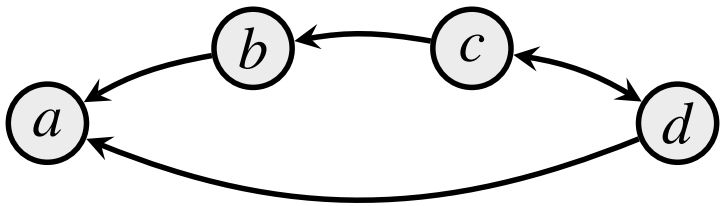
Semantic Compatibility, $\{X, Y\}_{\mathbb{S}}$: f.a. $x \in X, y \in Y$ ex. $S \in \mathbb{S}$, $\{x, y\} \subseteq S$
Semantic Conflict, $[X, Y]_{\mathbb{S}}$: otherwise

Framework Modifications



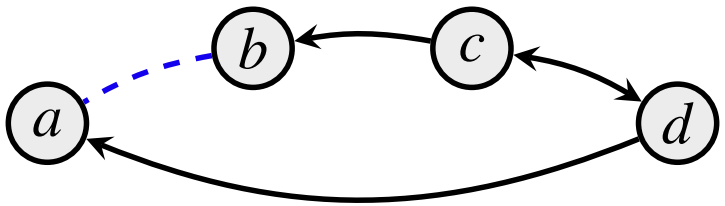
- Arguments: a, b, c, d
- Attacks: $(b, a), (c, b), (d, c), (c, d)$
- Extensions: $\{a, c\}, \{b, d\}$
- Conflicts: $[a, b], [b, c], [c, d], [a, d]$

Framework Modifications



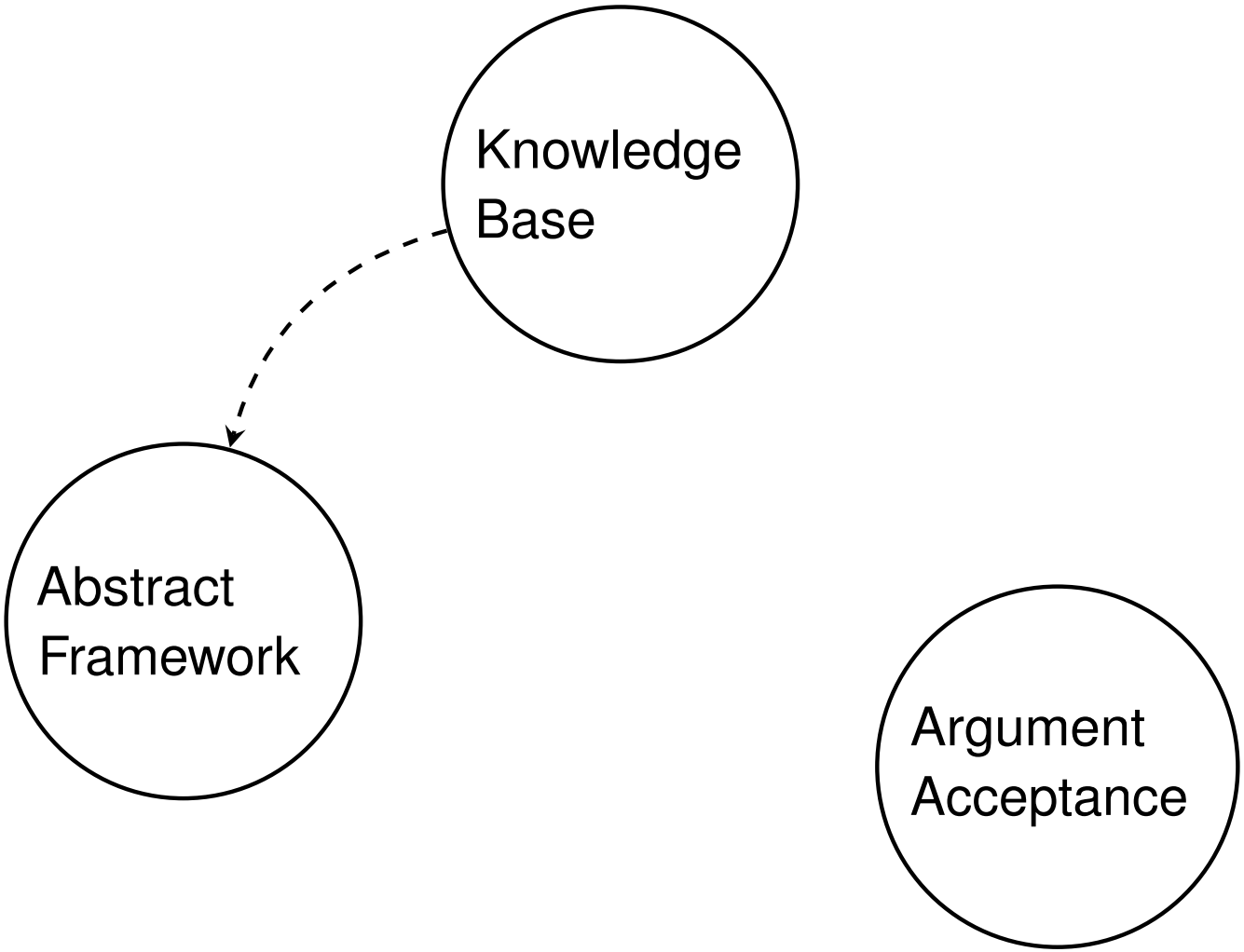
- Arguments: a, b, c, d
- Attacks: $(b, a), (c, b), (d, c), (c, d), (d, a)$
- Extensions: $\{a, c\}, \{b, d\}$
- Conflicts: $[a, b], [b, c], [c, d], [a, d]$

Framework Modifications

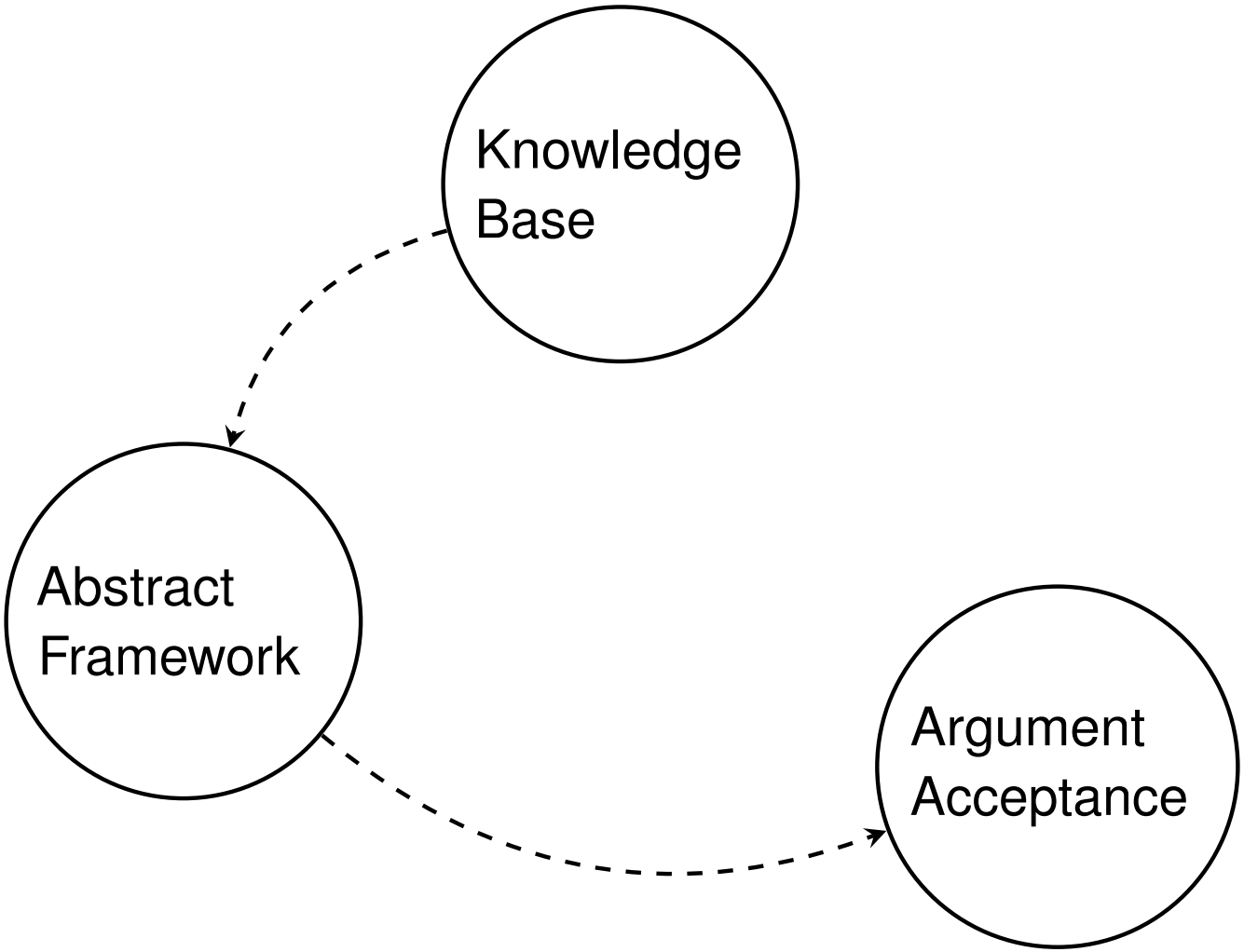


- Arguments: a, b, c, d
- Attacks: ~~(b, a)~~, (c, b), (d, c), (c, d),(d, a)
- Extensions: $\{a, c\}, \{b, d\}$
- Conflicts: $[a, b], [b, c], [c, d], [a, d]$

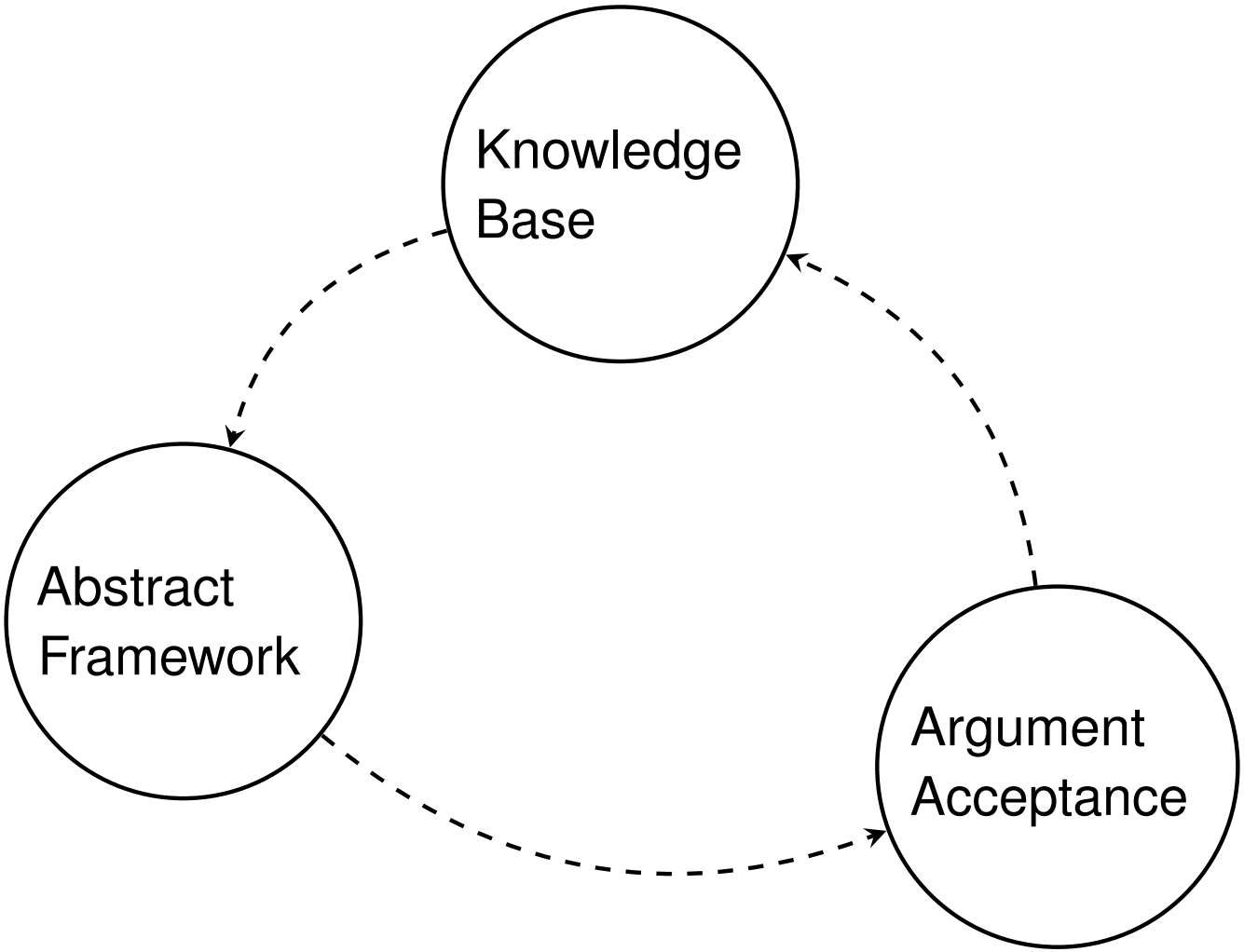
What are we doing?



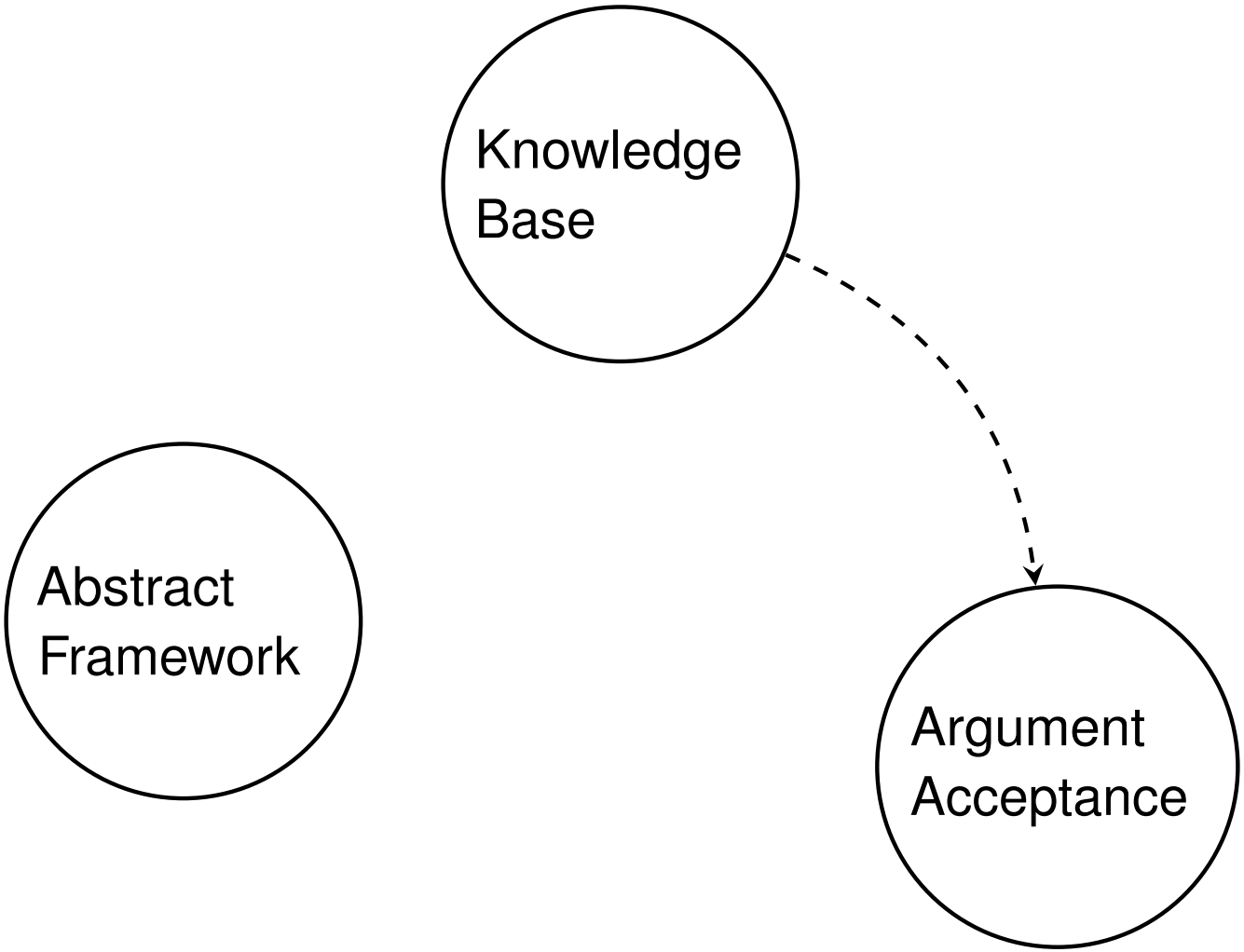
What are we doing?



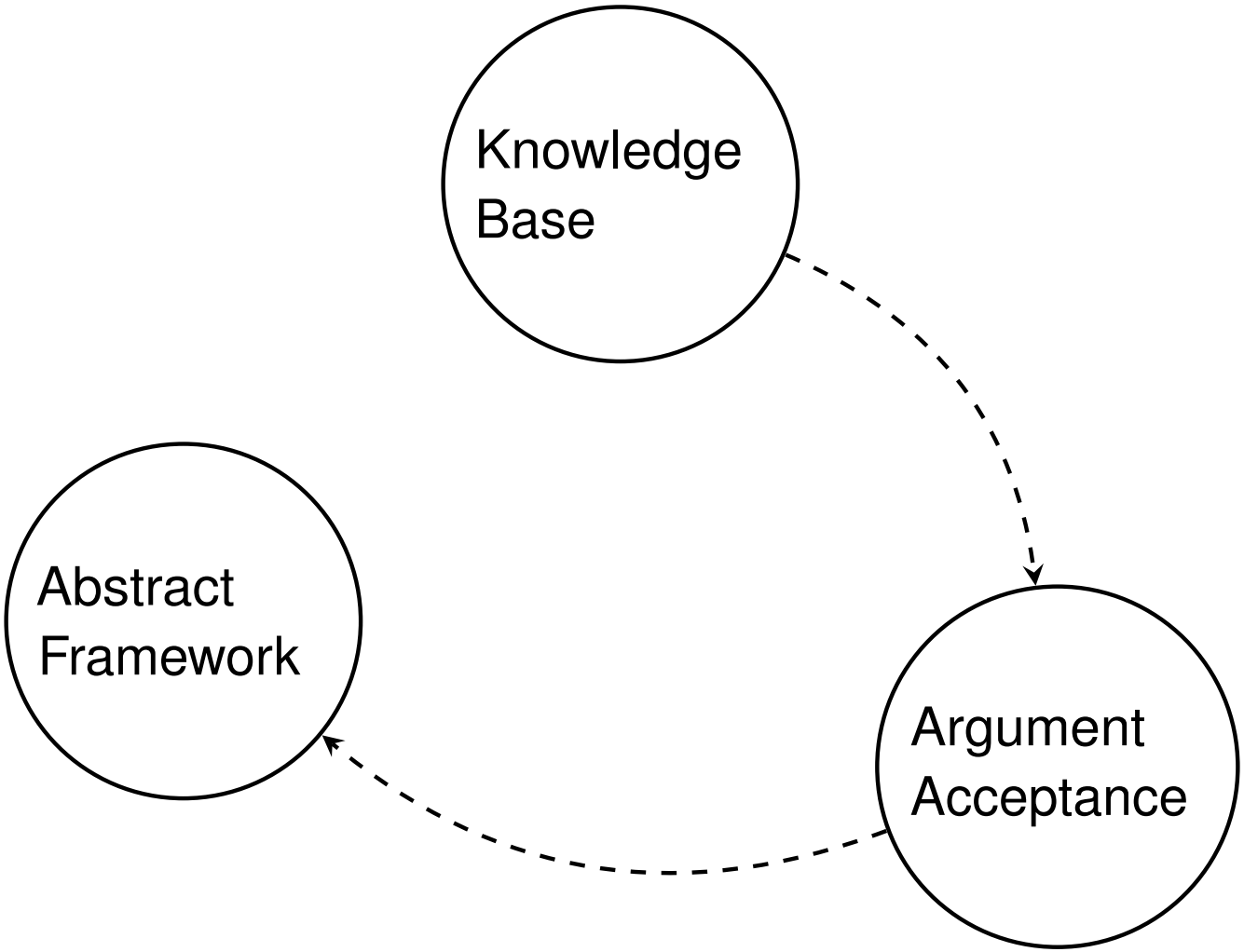
What are we doing?



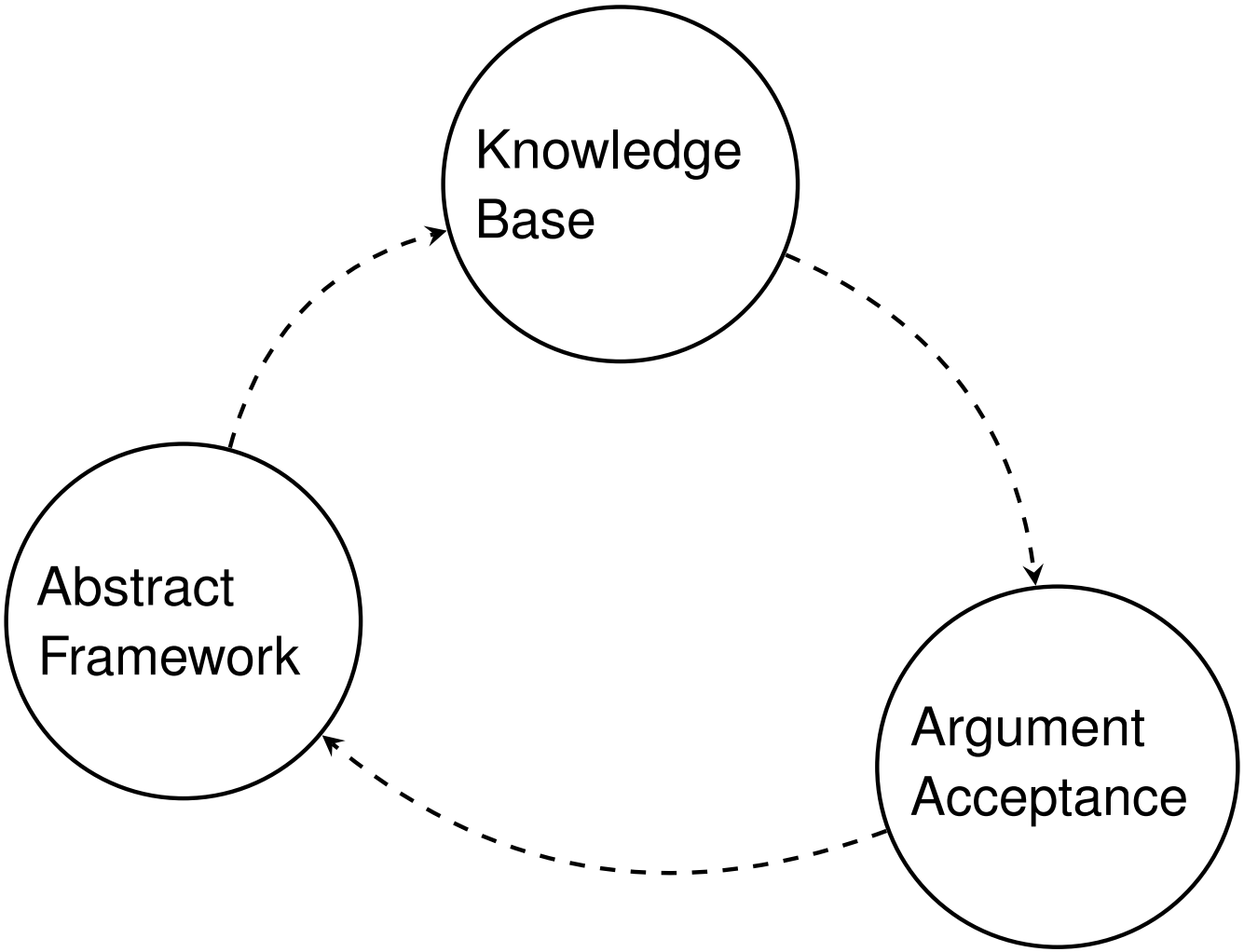
What are we doing?



What are we doing?



What are we doing?



Realizability and Conflict

Definition (Realizability)

$\mathbb{S}$ is σ-realizable if ex. AF F with $\sigma(F) = \mathbb{S}$

$\mathbb{S}$ is σ_A-realizable if ex AF $F = (A, R)$ with $\sigma(F) = \mathbb{S}$

Definition (Conflict)

A semantic conflict $[a, b]_{\mathbb{S}}$ is

- *pure* (semantic) if there is no realization F with $[a, b]_F$;
- *necessary* (syntactic) if any realization F has $[a, b]_F$;
- optional otherwise.

Levels of Conflict

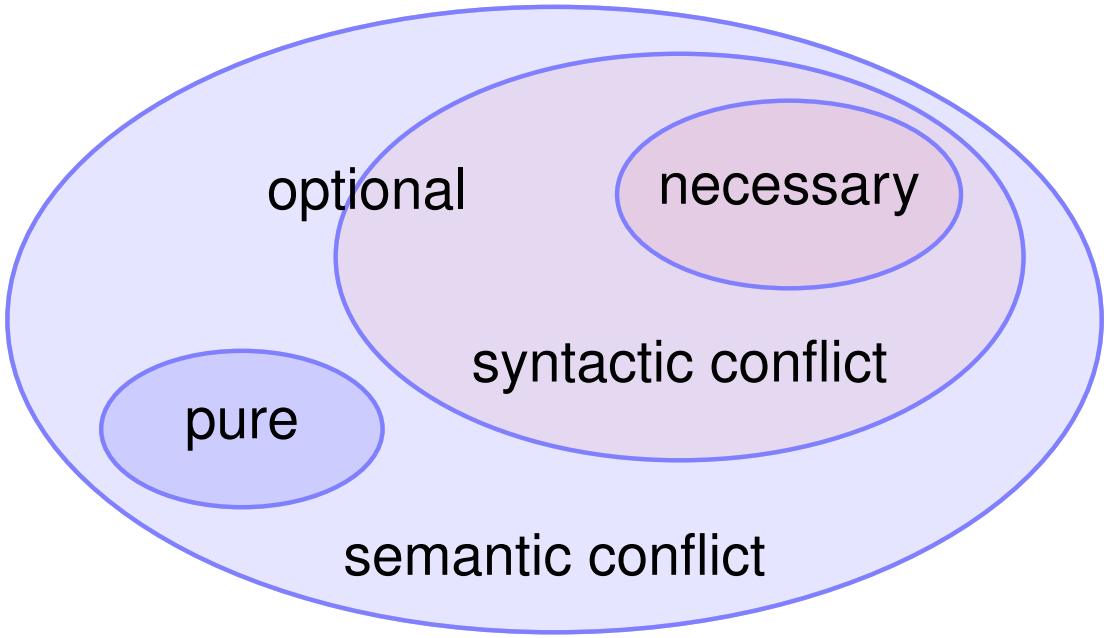
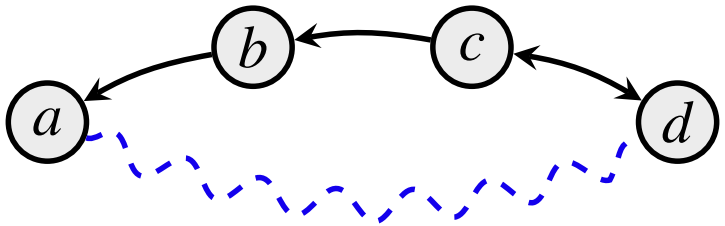


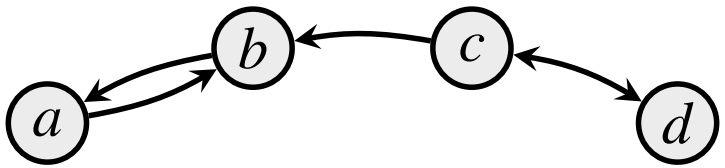
Figure : A Venn-diagram illustrating different levels of conflict.

Arbitrary Modifications



- Arguments: a, b, c, d
- Attacks: $(b, a), (c, b), (d, c), (c, d)$
- Extensions: $\{a, c\}, \{b, d\}$
- Conflicts: $[a, b], [b, c], [c, d], [a, d]$

Arbitrary Modifications



- Arguments: a, b, c, d
- Attacks: $(b, a), (c, b), (d, c), (c, d), (a, b)$
- Extensions: $\{a, c\}, \{b, d\}, \{a, d\}$
- Conflicts: $[a, b], [b, c], [c, d], [a, d]$

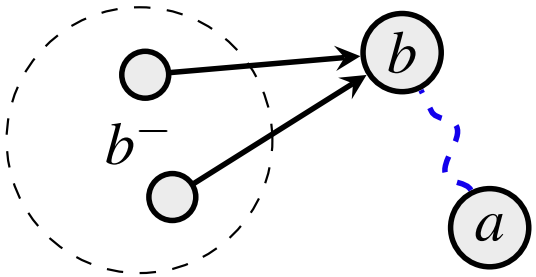
Conflict Characterizations

Theorem (Stable Conflicts)

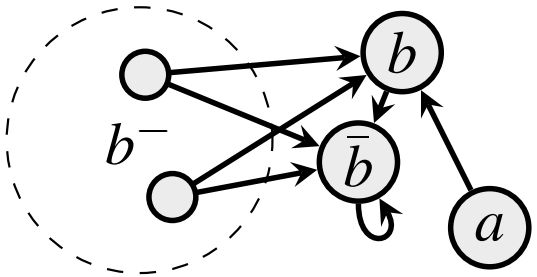
$[a, b]_{\mathbb{S}}$ *is necessary attack* $(a, b)_F$ *for each sb-realization* F *of* $\mathbb{S}$
if and only if
there is $S \in \mathbb{S}$, $a \in S$ *and* $\{b, S \setminus \{a\}\}_{\mathbb{S}}$.

All other conflicts for sb are optional.

Modifications for Stable Semantics

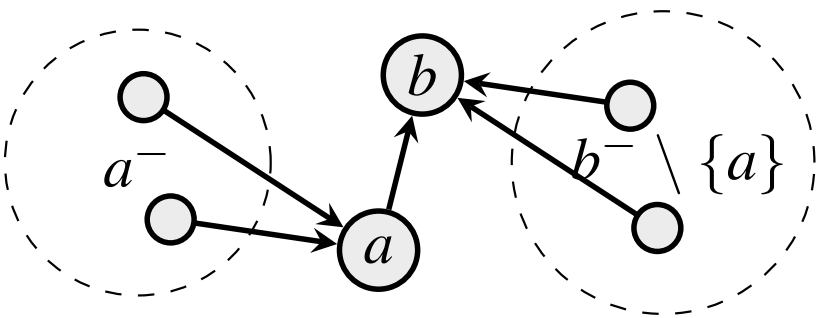


(a) Original AF, $[a, b]_{\mathbb{S}}$.

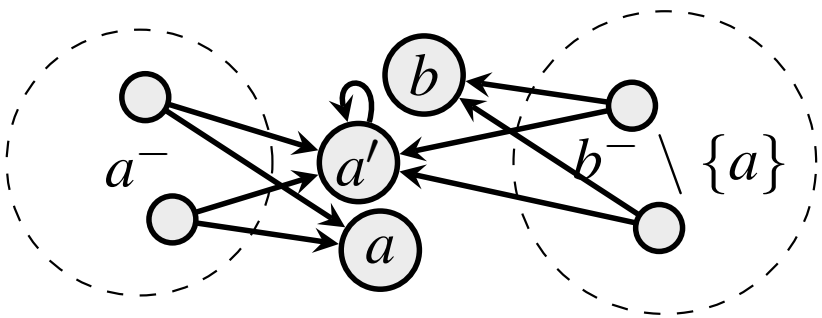


(b) Modified AF, $(a, b)_G$.

Figure : Forcing attacks for stable semantics.



(a) Original AF, $(a, b) \in R_F$.



(b) Modified AF, $(a, b) \notin R_G$.

Figure : Purging Attacks for Stable Semantics.

Illustration of Stable Modifications

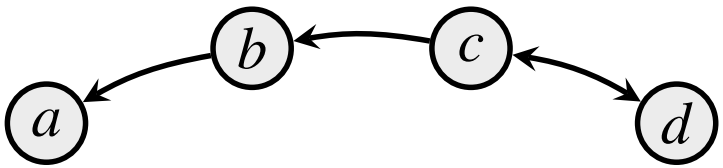
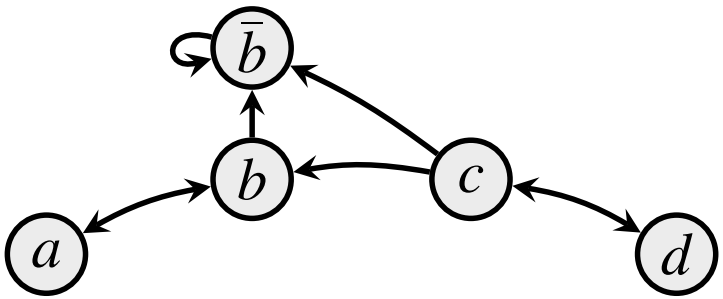
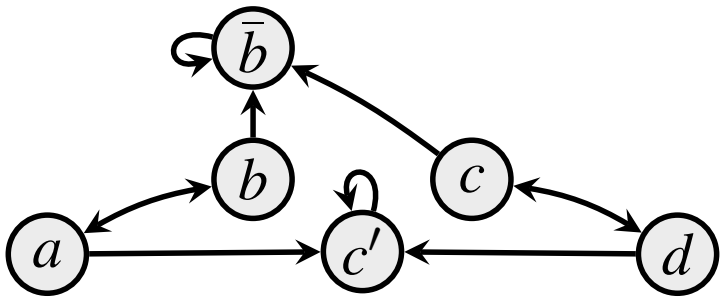


Figure : Original AF.



(a) Forcing Attack (a, b)



(b) Purging Attack (c, b)

Realizability and Conflict

Definition (Realizability)

$\mathbb{S}$ is σ-realizable if ex. AF F with $\sigma(F) = \mathbb{S}$

$\mathbb{S}$ is σ_A-realizable if ex AF $F = (A, R)$ with $\sigma(F) = \mathbb{S}$

Definition (Conflict)

A semantic conflict $[a, b]_{\mathbb{S}}$ is

- *pure* (semantic) if there is no realization F with $[a, b]_F$;
- *necessary* (syntactic) if any realization F has $[a, b]_F$;
- optional otherwise.

Realizability and Conflict

Definition (Realizability)

$\mathbb{S}$ is σ-realizable if ex. AF F with $\sigma(F) = \mathbb{S}$
$\mathbb{S}$ is σ_A-realizable if ex AF $F = (A, R)$ with $\sigma(F) = \mathbb{S}$

Definition (Conflict)

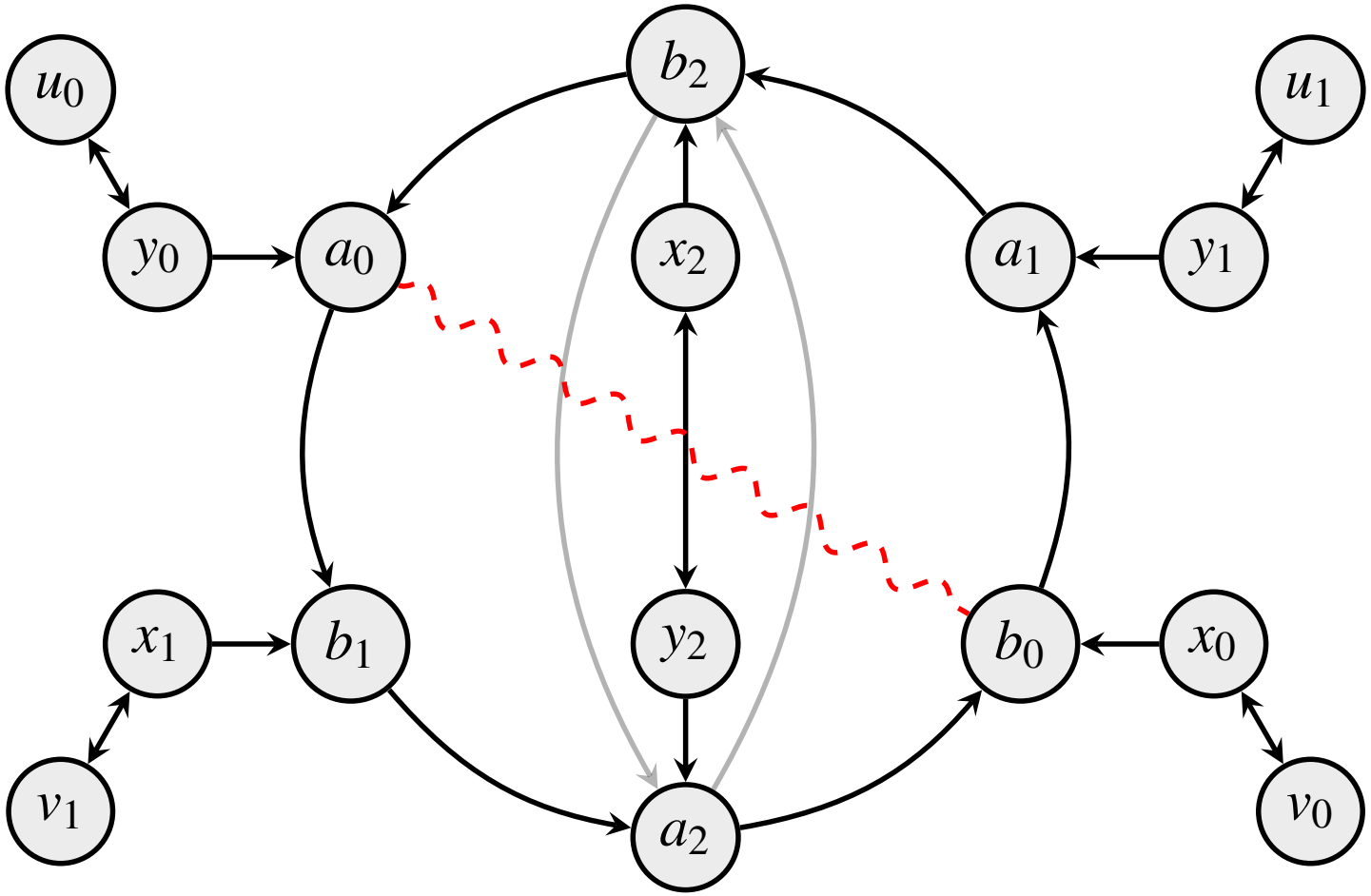
A semantic conflict $[a, b]_{\mathbb{S}}$ is

- *pure* (semantic) if there is no realization F with $[a, b]_F$;
- *necessary* (syntactic) if any realization F has $[a, b]_F$;
- optional otherwise.

Definition (Conditional Conflicts)

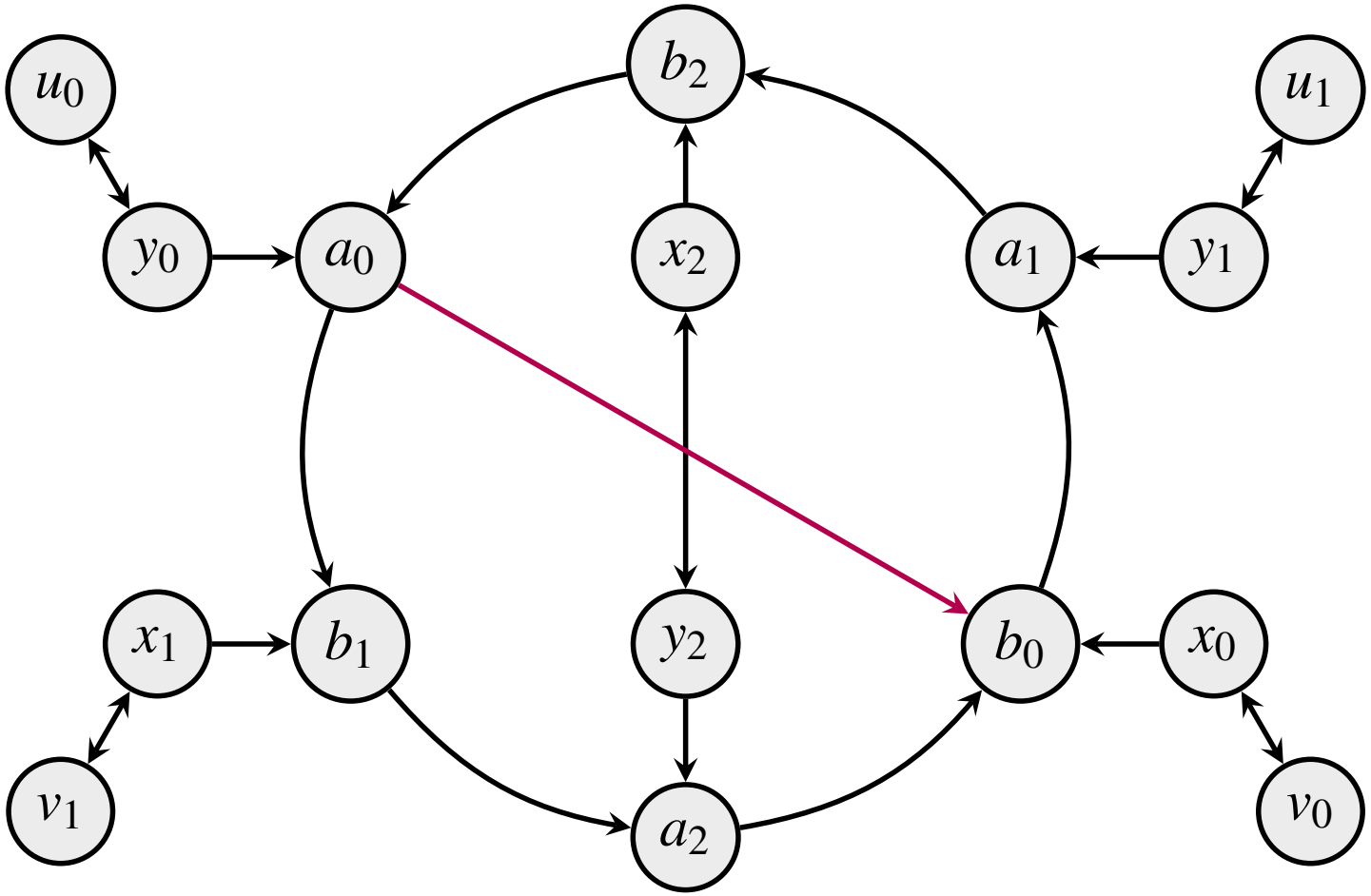
Extend pure, necessary and optional to A-realizability

A-Purity



Argument set of interest: $\{a_0, a_1, y_2, u_0, u_1, v_0, v_1\}$

A-Purity



Argument set of interest: $\{a_0, a_1, y_2, u_0, u_1, v_0, v_1\}$

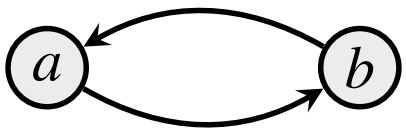
Conclusions

For Stable Semantics

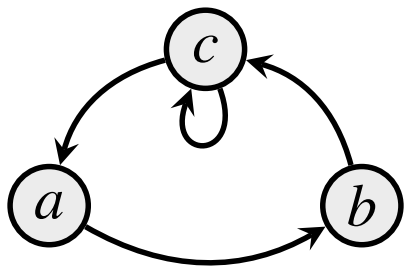
- necessary Conflicts can be directed (attacks) or undirected (symmetric attacks);
- in general there are no pure conflicts;
- A-purity however is possible;
- one could allow bigger extensions to get rid of necessary conflicts;
- manipulation only requires compatibilities.

Other Semantics

- Preferred and Semi-stable semantics have only symmetric necessary attacks $[a, b]$ where there are $S, T \in \mathbb{S}$ with $a \in S$, $b \in T$ and otherwise compatibilities $\{a, T \setminus \{b\}\}_{\mathbb{S}}$, $\{b, S \setminus \{a\}\}_{\mathbb{S}}$.
- Stage semantics has the same necessary conflicts as Stable, but without directions.
- Cf2 semantics probably has the same necessary conflicts as Stable, no necessary symmetric attacks but allows general pure conflicts.



(c) Symmetric Attack



(d) Directed Attack

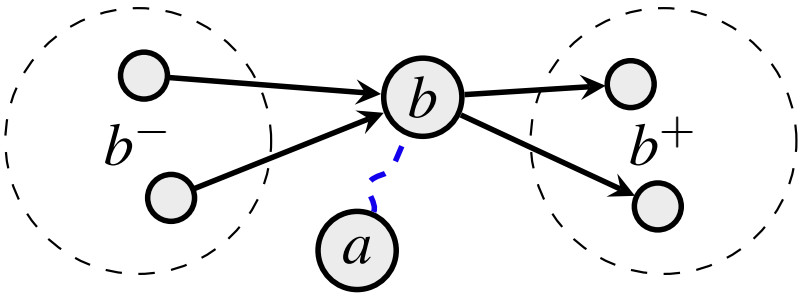
Future Work, Open Questions

- Conflicts between sets of arguments.
- Conditional Conflicts: exact characterizations for A-pure definitions, under what circumstances can A-pure conflicts arise?
- Formal definition of attack-minimal AFs
- Other semantics, labellings, . . .
- Instantiation-related questions; what does it mean to use such modifications? How can we use this knowledge to manipulate or analyse/detect manipulation?
- Other directions: Given some AF, which arguments necessarily are jointly acceptable? How can we detect semantic conflicts without computing all extensions?
- Syntactic Conflict is a *semantics*, extend approach to arbitrary pairs of semantics.

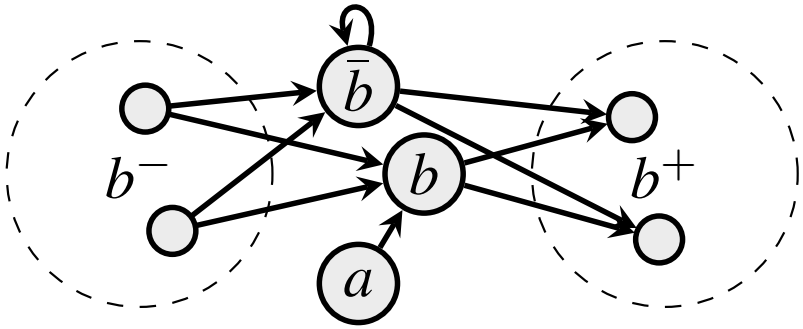
References

- Baroni, P., Caminada, M., and Giacomin, M. (2011).
  An introduction to argumentation semantics.
  Knowledge Eng. Review, 26(4):365–410.
- Dung, P. M. (1995).
  On the Acceptability of Arguments and its Fundamental Role in Nonmonotonic Reasoning, Logic Programming and n-Person Games.
  Artif. Intell., 77(2):321–358.
- Dunne, P. E., Dvořák, W., Linsbichler, T., and Woltran, S. (2015).
  Characteristics of multiple viewpoints in abstract argumentation.
  Artif. Intell., 228:153–178.
- Linsbichler, T., Spanring, C., and Woltran, S. (2015).
  The Hidden Power of Abstract Argumentation Semantics.
  The 2015 International Workshop on Theory and Applications of Formal Argument.

Preferred Modifications

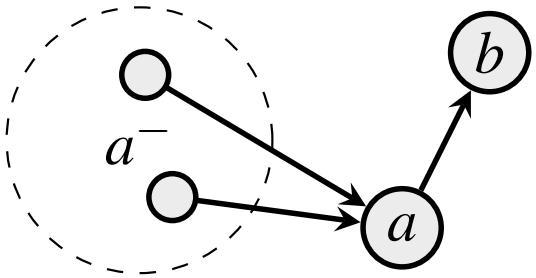


(e) Original AF, $[a, b]_{\mathbb{S}}$.

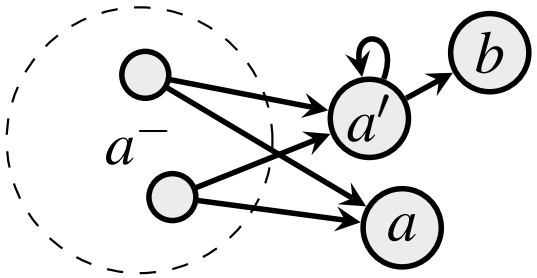


(f) Modified AF, $(a, b)_G$.

Figure : Forcing Attacks for Preferred Semantics.



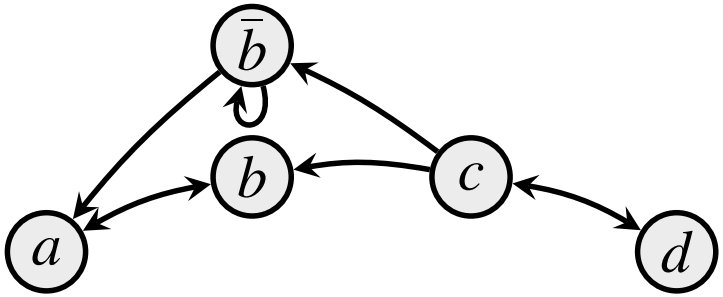
(a) Original AF, $(a, b) \in R_F$.



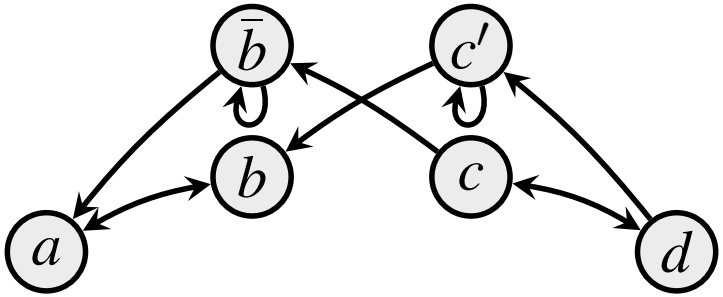
(b) Modified AF, $(a, b) \notin R_G$.

Figure : Purging Attacks for Preferred Semantics.

Illustration of Preferred Modifications.

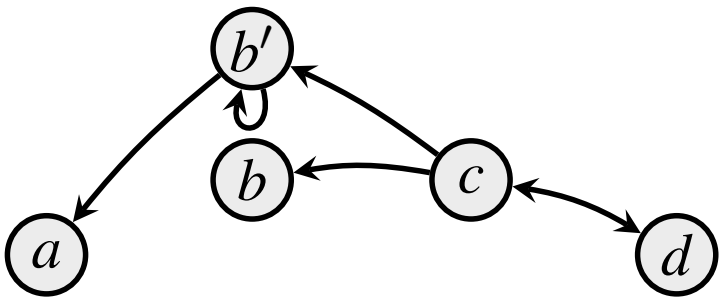


(a) Forcing Attack (a, b).

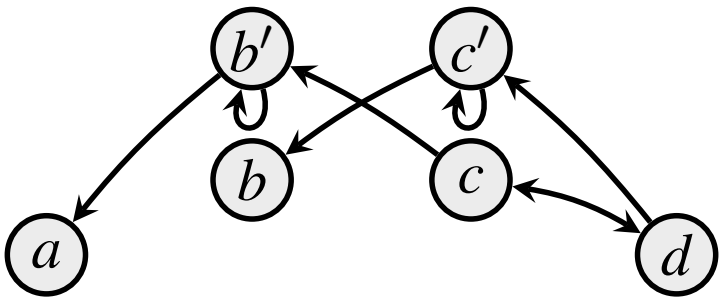


(b) Purging Attack (c, b).

Figure : Analogy to Stable Illustration.



(a) Purging Attack (a, b).



(b) Purging Attack (c, b).

Figure : For an attack-minimal AF.