Explainable Decision Support using Defeasible Logic

Structure

Eric Nakoja

Barbara Futyma

JingYang Zeng

David Pomerenke

David Schimmel

Laurens Rutten

Motivation | Background | Approach I | Approach II | Demo | Conclusion


FLIGHT
RECORDER
DO NOT OPEN


FLIGHT
RECORDER
DO NOT OPEN

INTRODUCTION

Explainable Decision Support Systems with Semantic Tableaux and Defeasible Rules

Motivation

- To investigate and make decision support systems explainable.
- To use argumentation-based approach to create explainable decision support systems.

.

PROBLEM STATEMENT

- Is it possible to create a decision support system that makes its reasoning understandable ?

Argumentation logic approach:

- Semantic tableaux approach.
- Backwards chaining approach.

RESEARCH QUESTIONS

1. How do we represent information and arguments in decision support systems?
2. How do we extract explanations for conclusions?
3. How do we solve the problem of redundant information?
4. How do we explain reasoning to the user?
5. Is an argumentation based approach limited to formally restricted domains?
6. How well does an argument-based approach scale with respect to the number of rules involved?

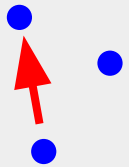
Structure

Barbara Futyma

Non-monotonic Reasoning & Defeasible Rules

- *Non-monotonic reasoning*
- *Defeasible rules*
  - *Undercutting attack*
  - *Rebutting attack*
- *Preference relation*

bird(x) ~> flies(x)
penguin(x) ~> ¬flies(x)
penguin(x) --> bird(x)



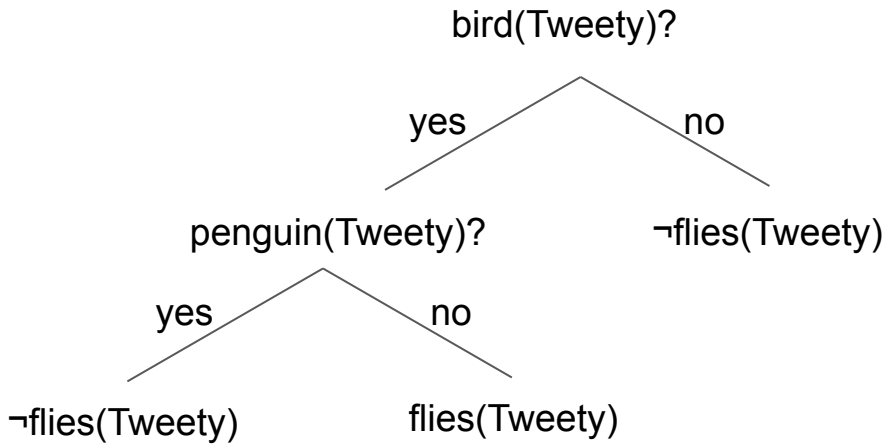
flies(Tweety)?

penguin(Tweety) **penguin(x) ~> ¬flies(x)**	¬flies(Tweety)

penguin(Tweety) penguin(x) --> bird(x) **bird(x) ~> flies(x)**	flies(Tweety)

Decision Support Systems

- *Help in the decision-making process*
  - *Binary classification*
- *User interaction*
- *Relevant questions*

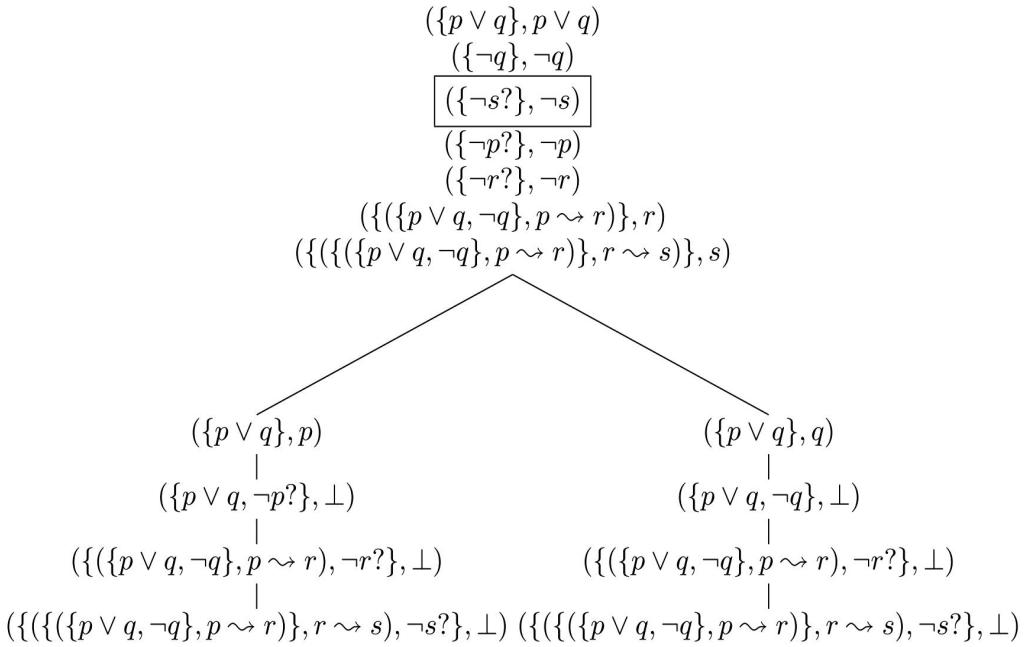


Structure

JingYang Zeng

The Argumentation Tableau

- Fact
- Tests
- Argument
- Inconsistency

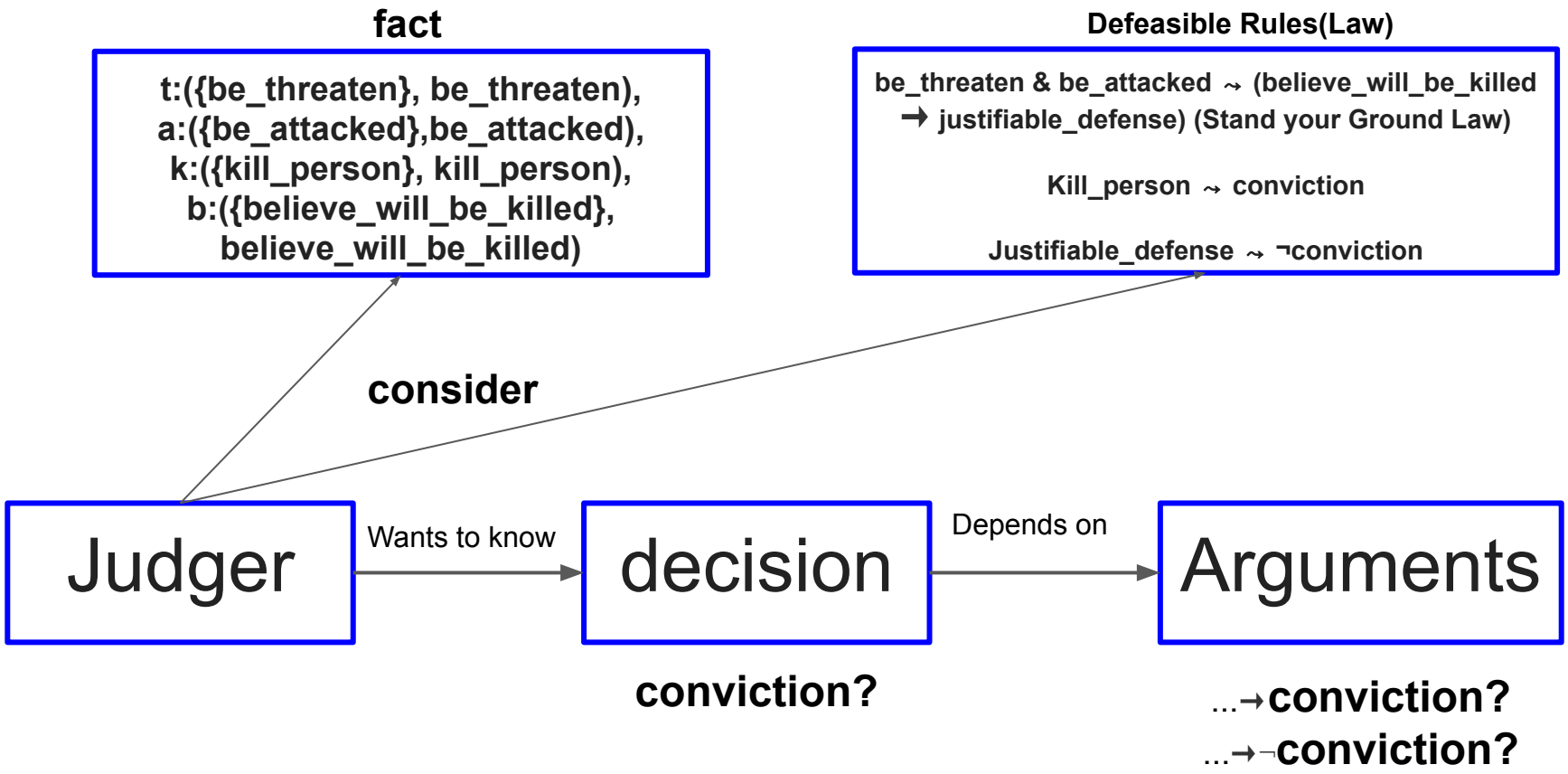


Ref: Roos, N. A Semantic Tableau Method for Argument Construction.

The court simulation - the case of Michael Drejka

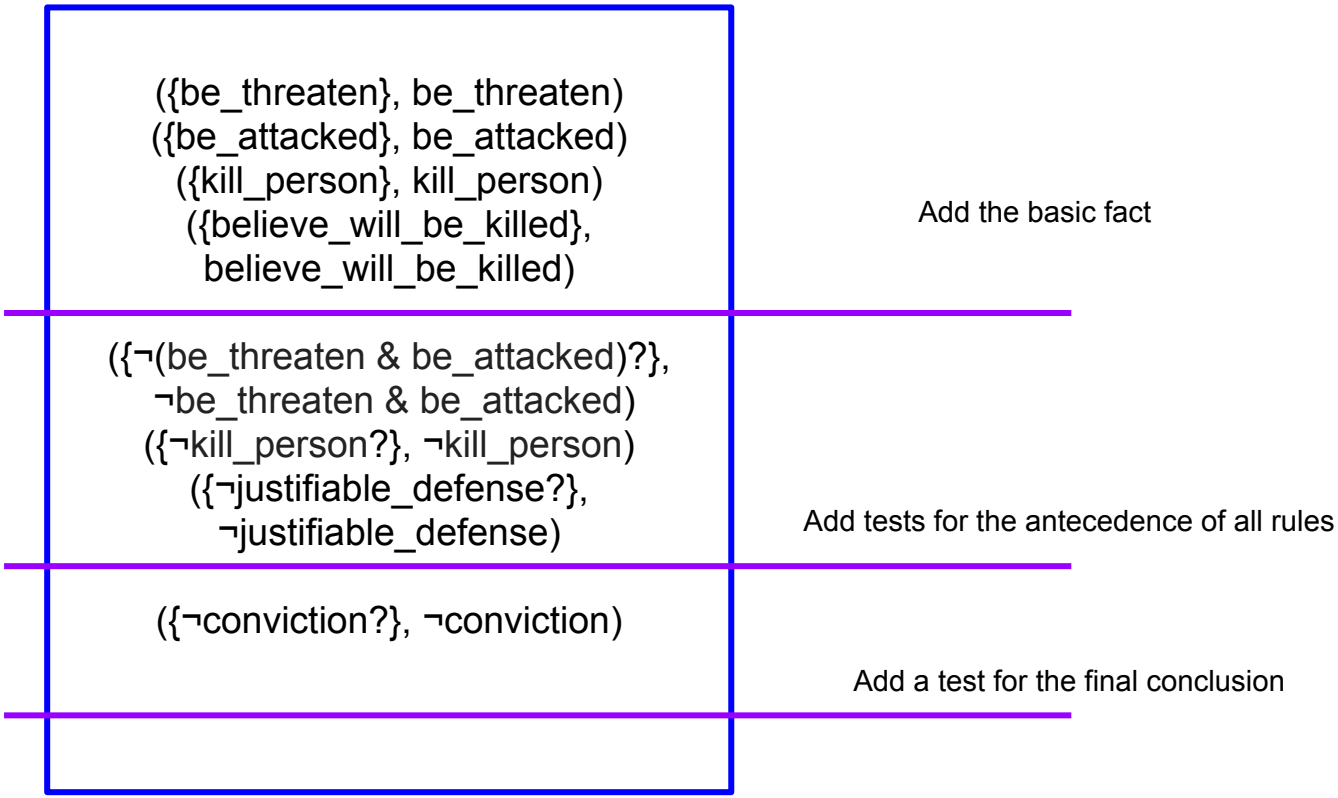
Stand your Ground Law: provides that people may use deadly force when they reasonably believe it to be necessary to defend against deadly threat.

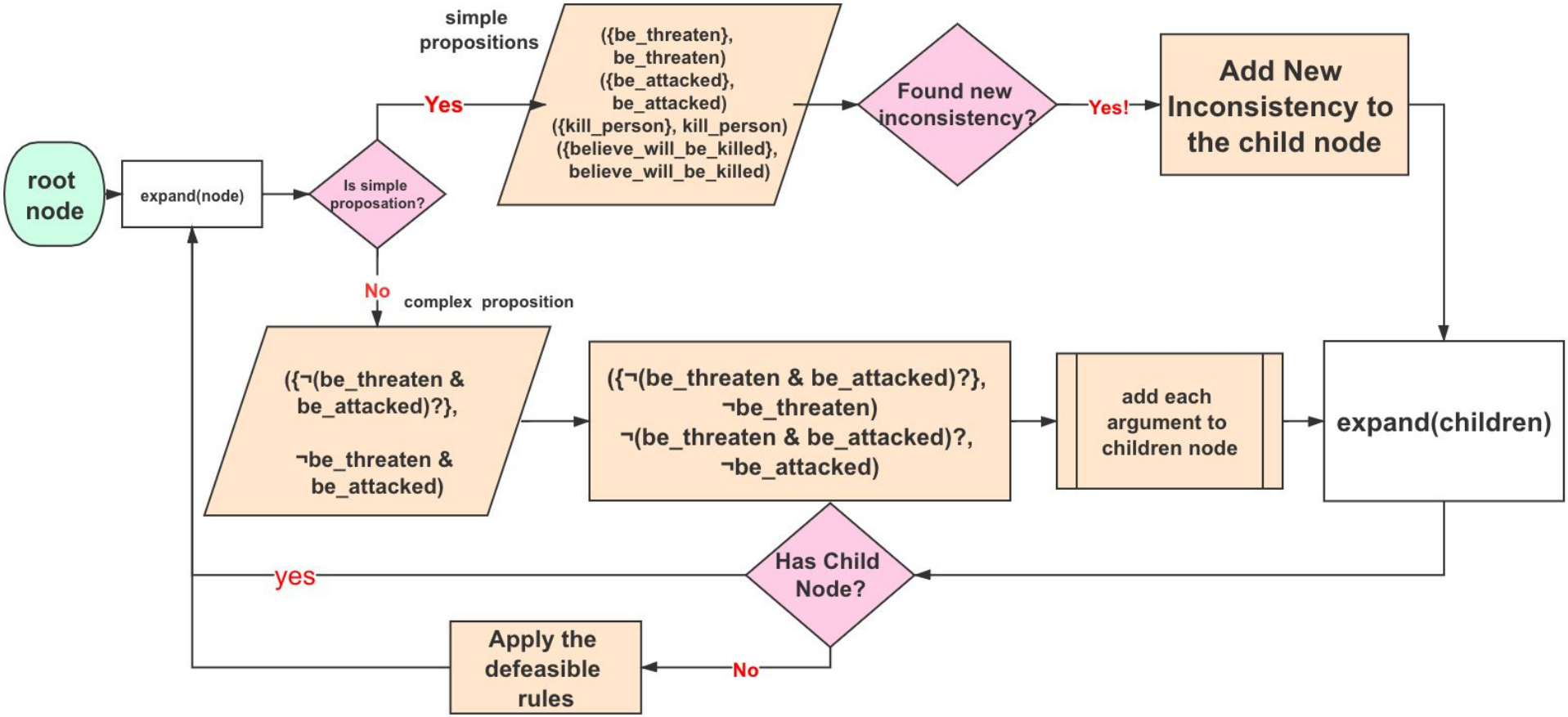
The court simulation



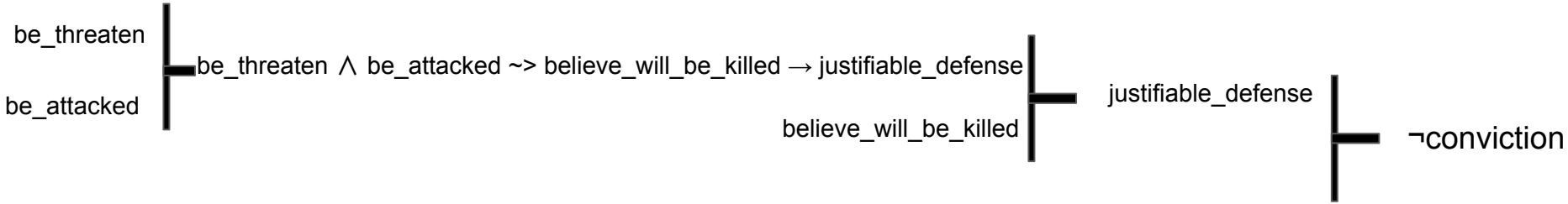
The court

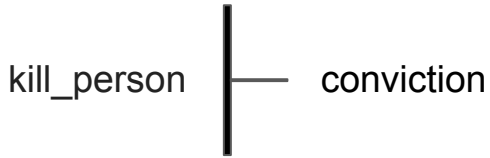
Initial root node



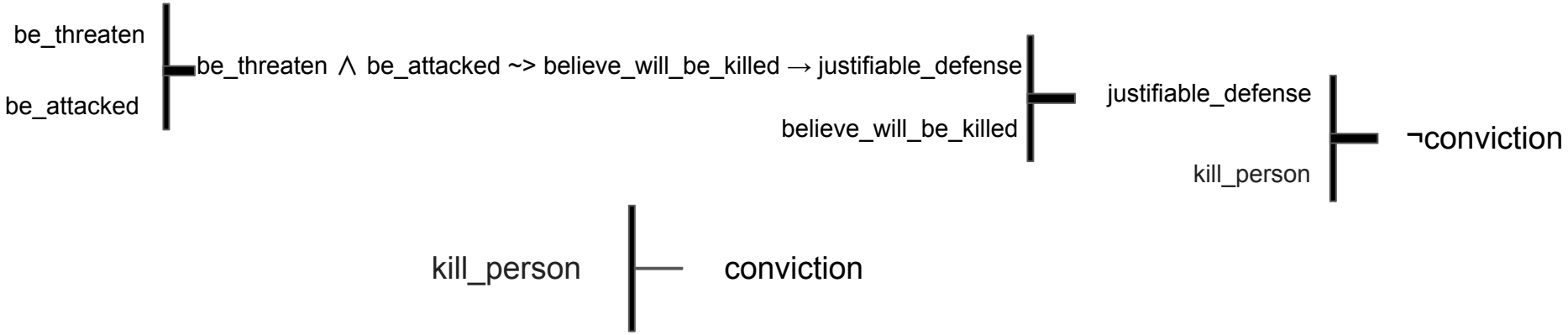

root node
expand(node)
Is simple proposation?
Yes
simple propositions
({be_threaten}, be_threaten)
({be_attacked}, be_attacked)
({kill_person}, kill_person)
({believe_will_be_killed}, believe_will_be_killed)
Found new inconsistency?
Yes!
Add New Inconsistency to the child node
No
complex proposition
({¬(be_threaten & be_attacked)?}, ¬be_threaten & be_attacked)
({¬(be_threaten & be_attacked)?}, ¬be_threaten) ¬(be_threaten & be_attacked)?, ¬be_attacked)
add each argument to children node
expand(children)
Has Child Node?
yes
No
Apply the defeasible rules

Argument Generated





Argument Generated



- If we have the priority of the defeasible rules, for example, the ‘stand your ground’ has a higher priority, then the conclusion of stronger argument is **¬conviction.**
- Add knowledge to expand new arguments by decision support system which will ask some question.

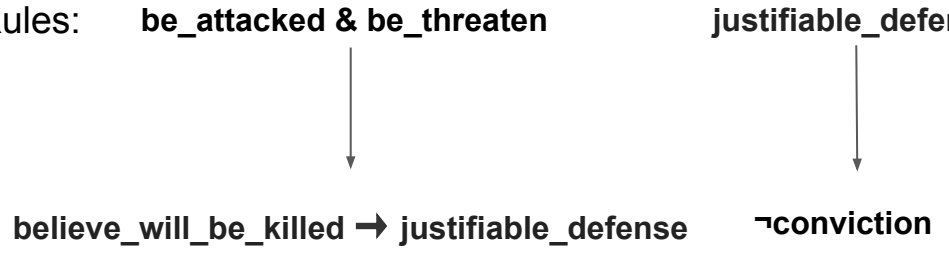
From Argumentation Tableau to Decision Support System

- Argumentation Tableau requires all information a priori → not practical
- Evaluate a Tableau as far as possible
- Ask questions about it
- Efficiency: Ask as little as possible

Question:
- Conviction?

Initial Information:
- kill person

Rules:

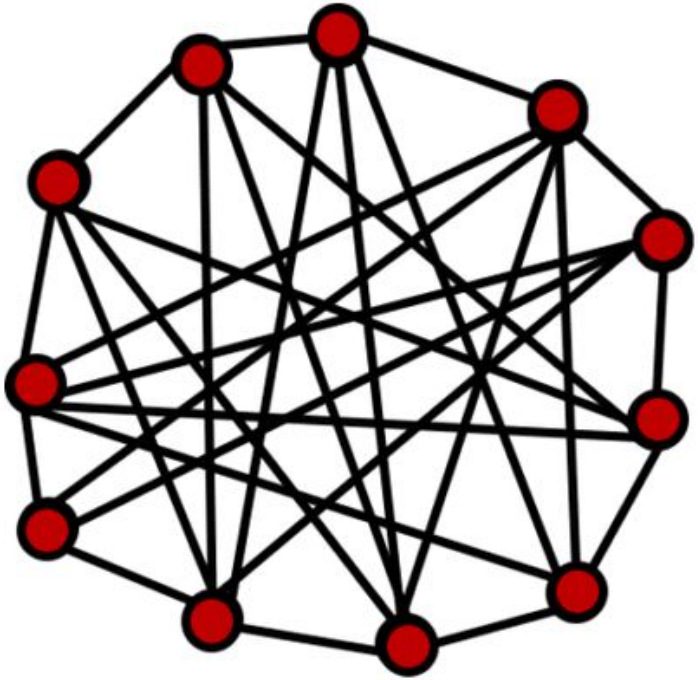


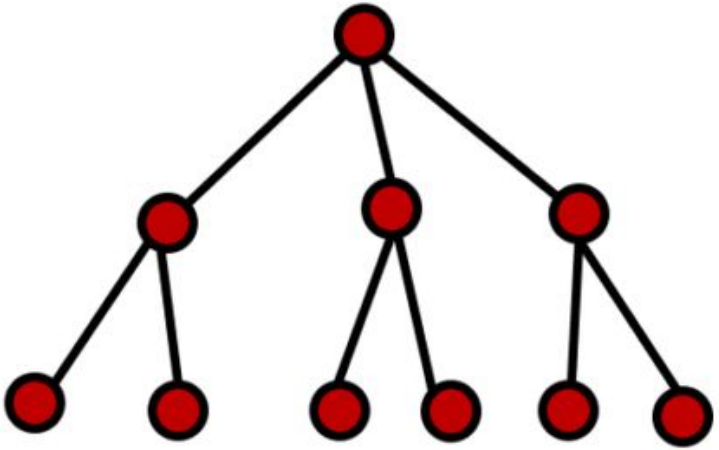
Possible questions:
- believe_will_be_killed?
- be_threaten?
- be_attacked?

Structure

David
Pomerenke

Forward-chaining vs backward-chaining





Example

is_threatened & is_attacked ->
(believes_will_be_killed -> justifiable_defense)

penguin -> not flies

Conviction?

bird -> flies

kills_person -> conviction

justifiable_defense -> not conviction

Example

not is_threatened or not is_attacked or not believes_will_be_killed or justifiable_defense

not penguin or not flies

Conviction?

not bird or flies

not kills_person or conviction

not justifiable_defense or not conviction

Example

not is_threatened or not is_attacked or not believes_will_be_killed or justifiable_defense

not kills_person or conviction

not penguin or not flies

Conviction?

not bird or flies

not justifiable_defense or not conviction

Example

not is_threatened or not is_attacked or not believes_will_be_killed or justifiable_defense

not kills_person ~~or conviction~~ 

not penguin or not flies

Conviction?

not bird or flies

not justifiable_defense ~~or not conviction~~ 

Example

not penguin or not flies

not bird or flies

not kills_person ~~or conviction~~

Conviction?

not is_threatened or not is_attacked or not believes_will_be_killed or justifiable_defense

not justifiable_defense ~~or not conviction~~

Example

not penguin or not flies

not bird or flies

not kills_person ~~or conviction~~

Conviction?

not is_threatened or not is_attacked or not believes_will_be_killed ~~or justifiable_defense~~

not justifiable_defense ~~or not conviction~~

Example

not penguin or not flies

not bird or flies

not kills_person ~~or conviction~~ +

is_threatened & is_attacked & believes_will_be_killed +

Conviction?

not is_threatened or not is_attacked or not believes_will_be_killed ~~or justifiable_defense~~ +

not is_threatened

not is_attacked -

not believes_will_be_killed

not justifiable_defense ~~or not conviction~~ -

-

Example

not penguin or not flies

not bird or flies

not kills ~~or conviction~~ +

Defeated

is_threatened & is_attacked & believes_will_be_killed +

Conviction?

not is_threatened or not is_attacked or not believes_will_be_killed ~~or justifiable_defense~~ +

not is_threatened

not is_attacked -

not believes_will_be_killed

not justifiable_defense ~~or not conviction~~ -

-

Example

is_threatened & is_attacked & believes_will_be_killed

not is_threatened

not is_attacked

not believes_will_be_killed

Asking good questions

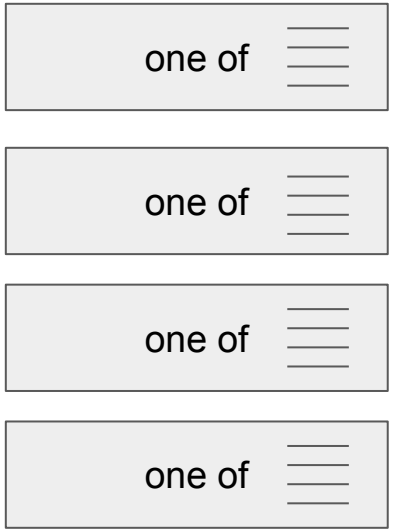
is_threatened & is_attacked & believes_will_be_killed

not is_threatened

not is_attacked

not believes_will_be_killed

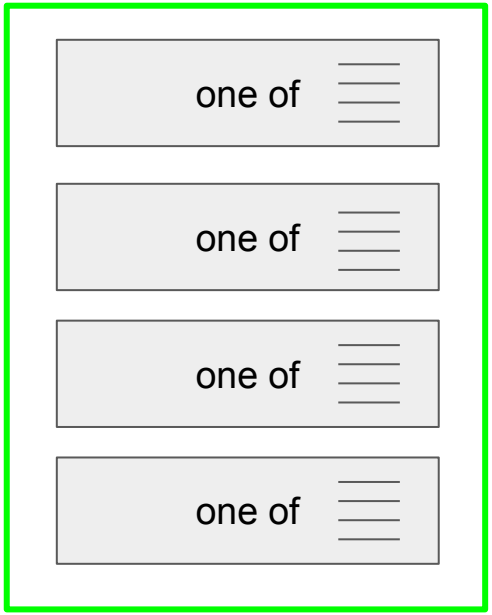
Asking good questions



Conjunctive
Normal
Form

Disjunctive
Normal
Form

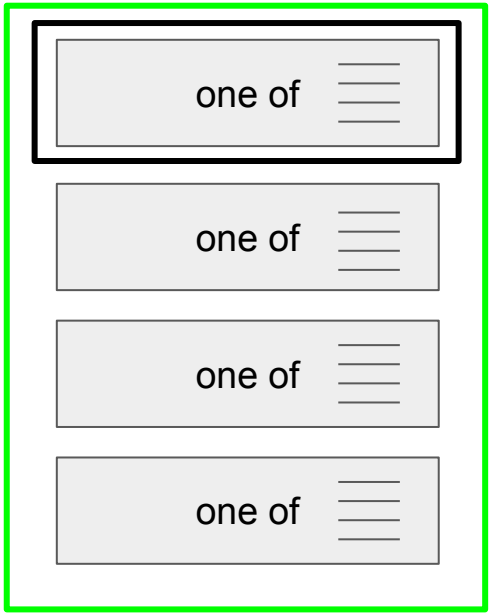
Asking good questions



Conjunctive
Normal
Form

Disjunctive
Normal
Form

Asking good questions



Conjunctive
Normal
Form

Disjunctive
Normal
Form

Structure

David
Schimmel

Demonstration

- First and second approach
- (Example from the murder case)

Discussion

- Explanations in form of arguments are shown
- Both approaches show the same argument
- For larger examples (>12 rules), performance drops drastically (making the system non interactive)
- Tableaux Method runs into problems faster
  - Prevent complete evaluation in each iteration

- Code and Examples can be found: https://github.com/explainable-reasoning

Structure

Laurens Rutten

Conclusion

- Explainable Decision Support System
  - Represent information
    - Generalized format
  - Extract explanations for conclusions
    - Arguments
  - Redundant information
    - Heuristic covering as many cases as possible
    - Checking for attacks each iteration
  - Scaling
    - 12 or more rules

Future Work

- Scalability
  - Pruning Search Space
- Predicate Logic
- Rule Mining


Thank you for listening!
Questions?

Image sources

- Front slide
  piqsels.com/en/public-domain-photo-zkxit
  public domain

Bibliography

- Roos, N. A Semantic Tableau Method for Argument Construction.
- Andone, E. C. E. A. D. (2018, July 29). *“Stand your ground” laws: Everything you need to know*. CNN. https://edition.cnn.com/2018/07/29/us/stand-your-ground-law-explainer-trnd/index.html
- T. Cremers, “Defeasible Logic as Professional Support for Regulation Analysis and Creation and Validation of the Specification of the Corresponding Services”.
- T. Bench-Capon, F. Coenen, and P. Orton, “Argument-based explanation of the british nationality act as a logic program,”Information and Communications Technology Law, vol. 2, no. 1, pp. 53–66, 1993.