Implementing Explanation-Based Argumentation using *Answer Set Programming*

Giovanni Sileno g.sileno@uva.nl
Alexander Boer, Tom van Engers

5 May 2014, ArgMAS presentation

Leibniz Center for Law
University of Amsterdam

Background

Argumentation

- Argumentation is traditionally seen in terms of *attack* and *support* relationships between *claims* brought by participants in a conversation.

Leibniz Center for Law

Argumentation

- Argumentation is traditionally seen in terms of *attack* and *support* relationships between *claims* brought by participants in a conversation.
- Argumentation seems to operate at a *meta-level* in respect to the content of arguments.

Leibniz Center for Law

Formal Argumentation

- Formal argumentation frameworks essentially target this *meta-level*

Formal Argumentation

- Formal argumentation frameworks essentially target this *meta-level*

- An *Argumentation framework* (AF) [Dung] consists of :
  - a set of arguments
  - attack relations between arguments

Leibniz Center for Law

Formal Argumentation

- To interpret/evaluate an AF we need a semantics.

- For instance, *extension-based semantics* classify sub-sets of arguments collectively *acceptable* in *extensions*:

  → the *justification state* of argument is defined in terms of memberships to extensions (*skeptically/credulously justified)*

Application of AFs

- Considering the whole process of application of argumentation theories, we recognize three steps:
  - Observation
  - Modeling/Reduction to AF
  - Analysis of AF

observer

modeler

analyst

traditional focus of formal argumentation

Leibniz Center for Law

Inside/Outside of *Argument Systems*

- In general, the extraction of attack relations may be problematic.

Inside/Outside of *Argument Systems*

- In general, the extraction of attack relations may be problematic.

- Trivial case: a claim is explicitly directed against another claim (*syntaxic* definition of attack).

Leibniz Center for Law

Inside/Outside of *Argument Systems*

- In general, the extraction of attack relations may be problematic.

- In a more general case, however, modelers have to use some background knowledge and underlying knowledge processing to identify the attacks.

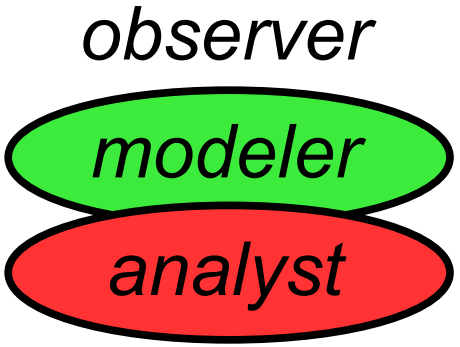
Inside/Outside of *Argument Systems*

- Usual solution: to integrate in the modeling phase default/defeasible reasoning.

- e.g. *assumption-based argumentation* (ABA)
  - Argument: *conclusion* ← *assumptions*
  - Attack to an argument holds if the “contrary” of its assumptions can be proved, or of its conclusion (*rebuttal*).

Inside/Outside of *Argument Systems*

- In practice in ABA the stress is on the *support* relation, expressed via defeasible rules, and used to extract the correspondendent AF.
  - Observation
  - Modeling/Reduction to AF
  - Analysis of AF



*(Part of) modeling is integrated, but still concerned by the **meta-level**!*

Leibniz Center for Law

The Puzzle

An interesting puzzle by Pollock

- John Pollock presents in in “Reasoning and probability”, *Law, Probability, Risk* (2007) a lucid analysis about the difficulties in reproducing certain intuitive properties with current formal argumentation theories.

An interesting puzzle by Pollock

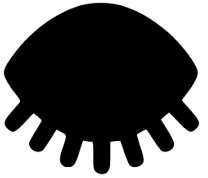
A) Jones says that the gunman had a moustache.

Leibniz Center for Law

An interesting puzzle by Pollock

A) Jones says that the gunman had a moustache.

B) Paul says that Jones was looking the other way and did not see what happened.

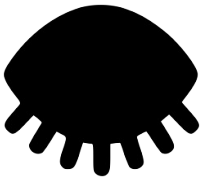


Leibniz Center for Law

An interesting puzzle by Pollock

A) Jones says that the gunman had a moustache.

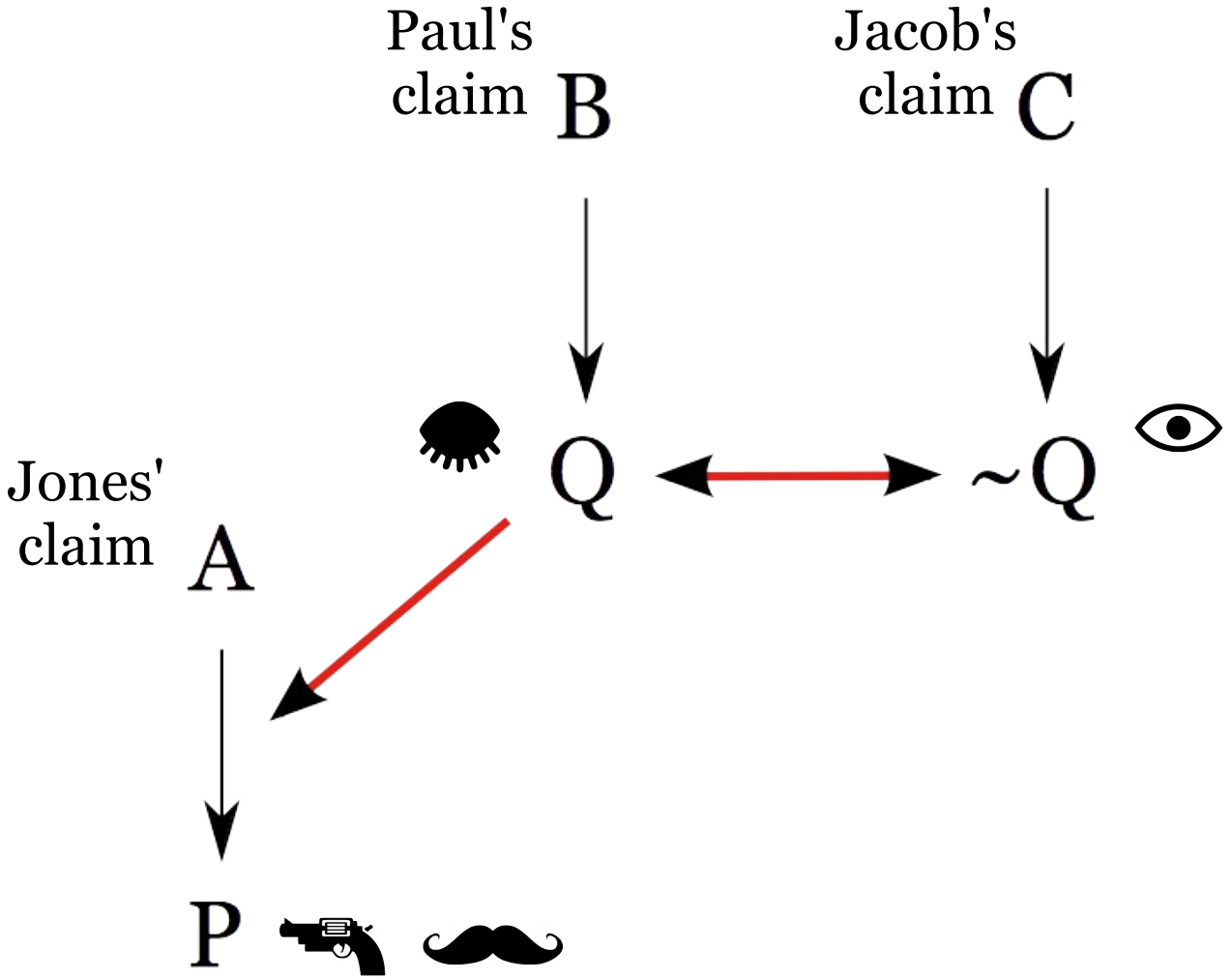
B) Paul says that Jones was looking the other way and did not see what happened.



C) Jacob says that Jones was watching carefully and had a clear view of the gunman.

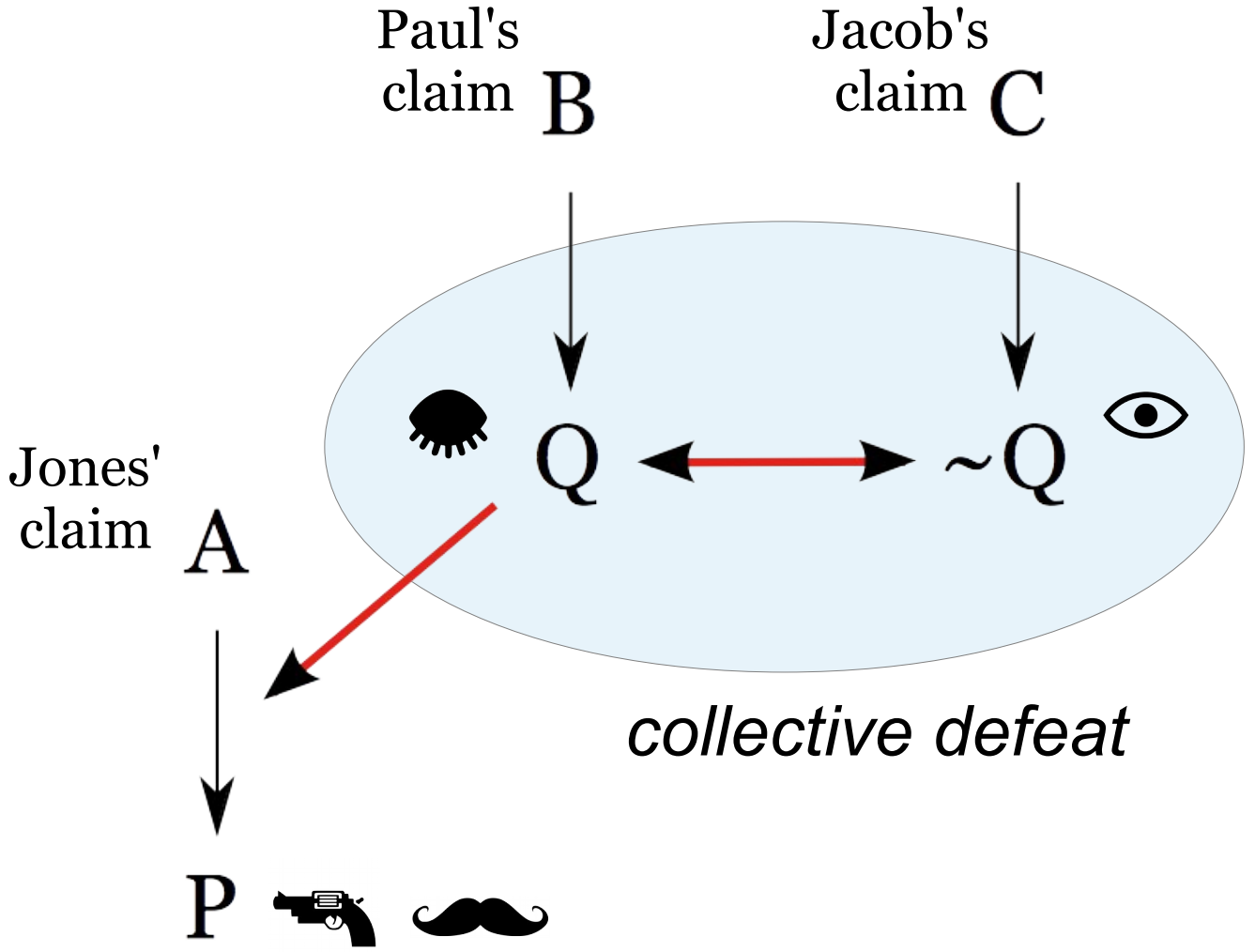
Leibniz Center for Law

Argumentation scheme of the puzzle

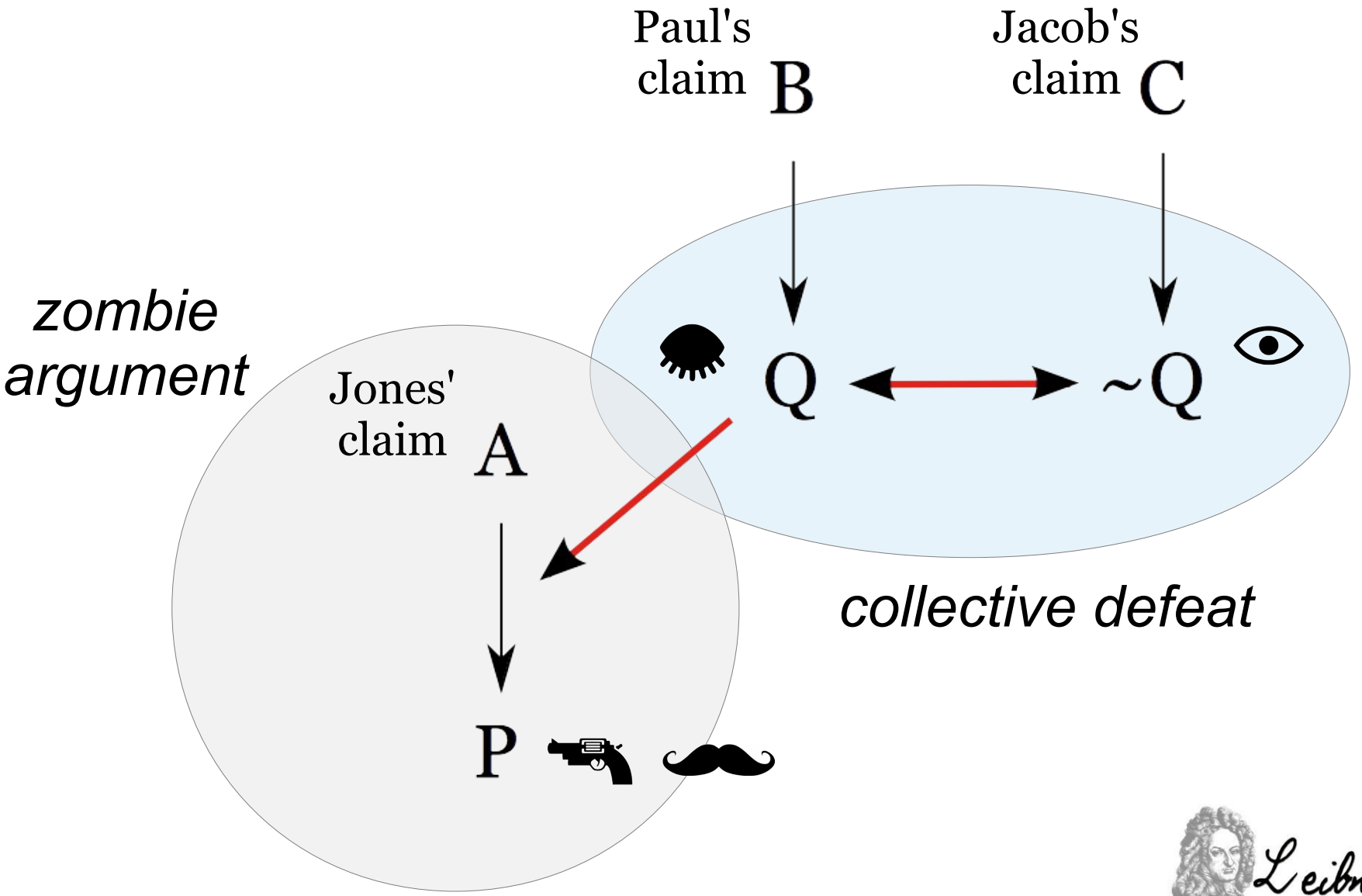


Leibniz Center for Law

Argumentation scheme of the puzzle

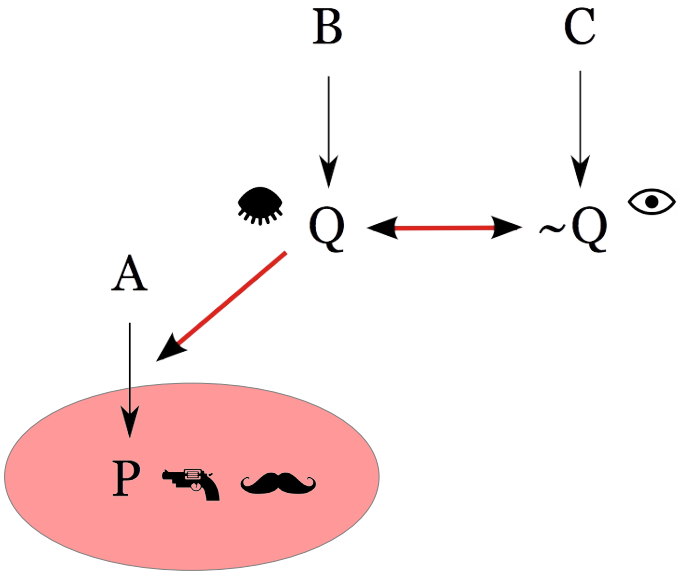


Argumentation scheme of the puzzle



Targeting intuitive properties

1. we should not believe to Jones' claim (i.e. the zombie argument) carelessly

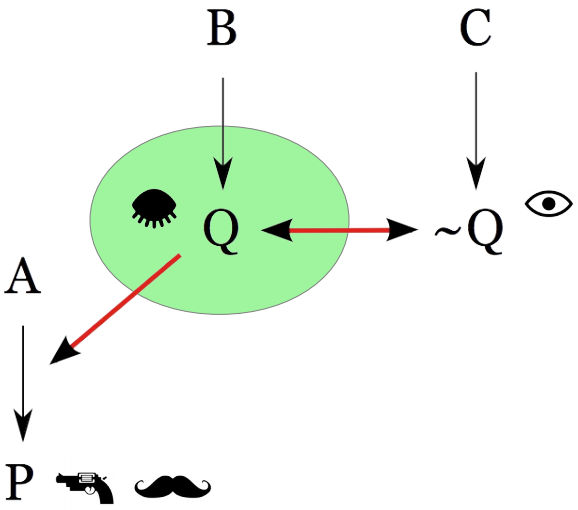


Leibniz Center for Law

Targeting intuitive properties

1. we should not believe to Jones' claim (i.e. the zombie argument) carelessly

2. if we assume Paul *more trustworthy* than Jacob, Paul's claim should be *justified* but to a *lesser degree*



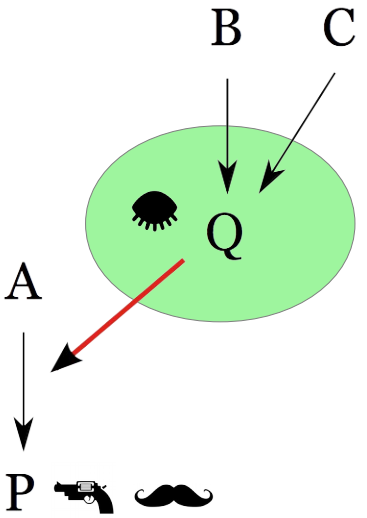
Leibniz Center for Law

Targeting intuitive properties

1. we should not believe to Jones' claim (i.e. the zombie argument) carelessly

2. if we assume Paul *more trustworthy* than Jacob, Paul's claim should be *justified* but to a *lesser degree*

3. if Jacob had confirmed Paul's claim, its *degree of justification* should have increased



Leibniz Center for Law

Pollock's puzzle

- Underlying problems:
  - zombie arguments
  - (relative) judgments of trustworthiness/reliability
  - ...
  - **how to approach justification?**

- Pollock proposed a highly elaborate preliminary solution based on *probable probabilities*.
- We propose a different solution, based on *explanation-based argumentation*.

Shift of perspective

Explanation-Based Argumentation

- Argumentation can be seen as a *dialectical process*, in which parties produce and receive messages.
- Argumentation does not concern only the matter of debate (e.g. a case, or *story*), but also the *meta-story* about about the construction of such story.

Leibniz Center for Law

EBA: observations

The Trial of Bill Burn under Martin's Act [1838]

- The sequence of collected messages consists in the **observation**.
- Sometimes the observation is collected by a third-party adjudicator, entitled to interpret the case from a neutral position.

Leibniz Center for Law

EBA: explanations

- Given a disputed case, an *explanation* is a possible scenario which is compatible
  - with the content of the messages, *and*
  - with the generation process of the messages.

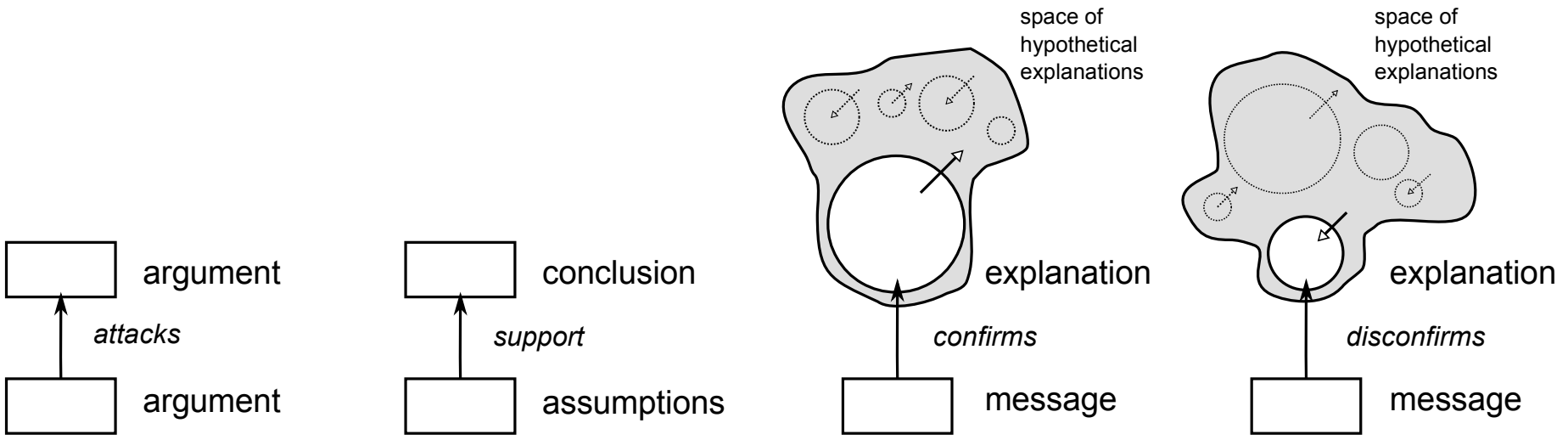
In general, the nature of such scenarios is of a *multi-representational* model, integrating physical, mental, institutional and abstract domains.

EBA: explanations

- Given a disputed case, an *explanation* is a possible scenario which is compatible
  - with the content of the messages, *and*
  - with the generation process of the messages.

- An explanation is **valid** if it reproduces the observation.
- Several explanations may be valid, i.e. fitting the same observation. Their competition is matter of *justificatio*n.

EBA: space of explanations



- Instead of being a static entity, the space of (hypothetical) explanations changes because of
  - the incremental nature of the observation (introducing new factors and constraints),
  - changes in strengths of epistemic commitment.

Leibniz Center for Law

Explanation-based Argumentation

- Referring to these ingredients, we propose the following operationalization, based on three steps.

Explanation-based Argumentation

1. Generation
   - Relevant factors, related to the observation, are grounded into *scenarios*

Explanation-based Argumentation

1. Generation
   - Relevant factors, related to the observation, are grounded into *scenarios*
2. Deletion
   - Impossible scenarios are removed, leaving a set of *hypothetical explanations*

Explanation-based Argumentation

1. Generation
   - Relevant factors, related to the observation, are grounded into *scenarios*
2. Deletion
   - Impossible scenarios are removed, leaving a set of *hypothetical explanations*

Operational assumption: effective capacity of generating adequate scenarios

Explanation-based Argumentation

1. Generation
   - Relevant factors, related to the observation, are grounded into *scenarios*
2. Deletion
   - Impossible scenarios are removed, leaving a set of *hypothetical explanations*
   - Hypothetical explanations fitting the observation select the *explanations*

Explanation-based Argumentation

1. Generation
   - Relevant factors, related to the observation, are grounded into *scenarios*
2. Deletion
   - Impossible scenarios are removed, leaving a set of *hypothetical explanations*
   - Hypothetical explanations fitting the observation select the *explanations*

Informational assumption: an observation either fits an explanation or it doesn't.

Explanation-based Argumentation

1. Generation
   - Relevant factors, related to the observation, are grounded into *scenarios*
2. Deletion
   - Impossible scenarios are removed, leaving a set of *hypothetical explanations*
   - Hypothetical explanations fitting the observation select the *explanations*
3. Justification
   - The relative position of explanations depends on the strengths of epistemic commitment

Explanation-based Argumentation

- Argumentation frameworks based on defeasible reasoning insist on the *inferential aspect* of the problem, rather than the *selection* of an adequate search space.
- The selection of (hypothetical) explanations hides already a certain commitment.
- Hypothetical explanations can be associated to a certain likelihood (*prior*).
- After some relevant message, the likelihood, i.e. the "strength" of explanations should change (*posterior*).

EBA: evaluation of explanations

- Bayesian probability
  - Subjective interpretation: probability counts as a measure of the *strength of belief*.
  - $L(E|O) = P(O|E)$

EBA: evaluation of explanations

- A relative ordinal judgment can be evaluated calculating the confirmation value for each explanation E (taken from Tentori, 2007):

$$c(O|E) = \frac{P(O|E) - P(O|\neg E)}{P(O|E) + P(O|\neg E)}$$

EBA: evaluation of explanations

- A relative ordinal judgment can be evaluated calculating the confirmation value for each explanation E (taken from Tentori, 2007):

$$c(O|E) = \frac{P(O|E) - P(O|\neg E)}{P(O|E) + P(O|\neg E)}$$

Well-known explanatory space assumption:
$P(E_1) + P(E_2) + .. + P(E_n) \sim 1$

Leibniz Center for Law

Implementation

Implementation of EBA in ASP

- *Answer set programming* is a declarative programming paradigm based on the *stable-model semantic*, oriented towards difficult (NP-hard) search problems.
  - In ASP, similarly to Prolog, the programmer models a problem in terms of rules and facts, instead of specifying an algorithm. The resulting code is given as input to a solver, which returns multiple *answer sets* or *stable models* satisfying the problem.

- We take advantage of the search capabilities of ASP solvers, in order to effectively perform the *generation* and *deletion* steps at once.

Implementation of EBA in ASP

- An ASP program related to an explanation-based argumentation consists of 3 parts:
  1. *allocation choices*, grounding all permutations of relevant factors,
  2. *world properties* and *ground facts*, modeling shared assumptions,
  3. *observation*, modeling the collected messages.

Implementation of EBA in ASP

- An ASP program related to an explanation-based argumentation consists of 3 parts:
  1. *allocation choices*, grounding all permutations of relevant factors,
  2. *world properties* and *ground facts*, modeling shared assumptions,
  3. *observation*, modeling the collected messages.

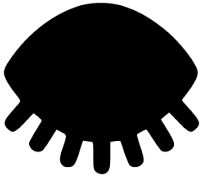
- The ASP solver gives as output hypothetical explanations (with 1+2) and explanations (1+2+3).
  - Assigning a prior probability to hyp. explanations, and analysing the final explanations we calculate the confirmation values.

Leibniz Center for Law

Relevant factors?

A) Jones says that the gunman had a moustache.

B) Paul says that Jones was looking the other way and did not see what happened.



C) Jacob says that Jones was watching carefully and had a clear view of the gunman.

Leibniz Center for Law

Relevant factors for assertion

- what an agent says may hold or not
- an agent may be *reliable* or not
- when he is reliable, what he says is what it holds.

Relevant factors for assertion

- what an agent says may hold or not
- an agent may be reliable or not
- when he is reliable, what he says is what it holds.
- e.g. *Paul says Jones was not seeing the gunman.* Writing “Paul is reliable” as **`paul`** and “Jones was seeing” as **`eye`**, we have:

```
1{eye, -eye}1.
1{paul, -paul}1.
-eye :- paul.
```

Leibniz Center for Law

Implementation of the puzzle in ASP

- An ASP program related to an explanation-based argumentation consists of 3 parts:
  1. *allocation choices*, grounding all permutations of relevant factors:

```
1{moustache, -moustache}1.
1{eye, -eye}1.
1{jones, -jones}1.
1{paul, -paul}1.
1{jacob, -jacob}1.
```

Leibniz Center for Law

Implementation of the puzzle in ASP

- An ASP program related to an explanation-based argumentation consists of 3 parts:
  1. *allocation choices*,
  2. *world properties* and *ground facts*, modeling shared assumptions:

```
eye :- jones.
```

Implementation of the puzzle in ASP

- An ASP program related to an explanation-based argumentation consists of 3 parts:
  1. *allocation choices*,
  2. *world properties* and *ground facts*,
  3. *observation*, modeling the collected messages:

```
moustache :- jones.
-eye :- paul.
eye :- jacob.
```

Leibniz Center for Law

Results

- We model the puzzle incrementally, so as to analyze the impact of each new message.

#	O_1	O_2	O_3
relevant factors	3	4	5
scenarios	8	16	32
hypothetical explanations	6	12	24
explanations	5	7	10

Prior probabilities

- How to calculate the prior probabilities?
- As we know all relevant factors characterizing the explanations, assuming that they are independent in the allocation phase we have:

  $P(E_i) = P(f_1) * P(f_2) * \ldots * P(f_n)$

- A neutral perspective is obtained assuming $P(f_i) = 0.5$
- As the inclusion of world properties and ground facts descrease the number of explanations, a normalization phase is required.

Evaluation vs targeted properties

	E_U (Jacob attacks Paul)											
	P	O_1	O_2	O_3	P	O_1	O_2	O_3	P	O_1	O_2	O_3
moustache	.50	U	U	U	.50	U	U	U	.50	T	U	T
eye	.50	U	U	U	.50	U	F	F	.50	T	U	T
jones	.50	U	U	U	.50	U	F	F	**.55**	T	U	T
paul	.50	-	U	U	**.55**	-	T	T	**.55**	-	U	F
jacob	50	-	-	U	.50	-	-	F	**.55**	-	-	T
$c(O\|E_U)$		.27	.43	.55		.27	.45	.58		.31	.48	.61

- we should not believe to Jones' claim carelessly

  → explanations in which the gunman has the moustache or not are confirmed to the same degree

Leibniz Center for Law

Evaluation vs targeted properties

	E_U (Jacob attacks Paul)											
	P	O_1	O_2	O_3	P	O_1	O_2	O_3	P	O_1	O_2	O_3
moustache	.50	U	U	U	.50	U	U	U	.50	T	U	T
eye	.50	U	U	U	.50	U	F	F	.50	T	U	T
jones	.50	U	U	U	.50	U	F	F	**.55**	T	U	T
paul	.50	-	U	U	**.55**	-	T	T	**.55**	-	U	F
jacob	.50	-	-	U	50	-	-	F	**.55**	-	-	T
$c(O\|E_U)$		.27	.43	.55		.27	.45	.58		.31	.48	.61

- if we assume Paul more trustworthy than Jacob, Paul's claim should be justified but to a lesser degree

  $\longrightarrow$ the explanation in which Paul tells the truth is more confirmed than the others.

Leibniz Center for Law

Evaluation vs targeted properties

	E_U (Jacob supports Paul)							
	P	O_1	O_2	O_3	P	O_1	O_2	O_3
moustache	.50	U	U	U	.50	T	U	U
eye	.50	U	U	U	.50	T	U	F
jones	.50	U	U	U	**.55**	T	U	F
paul	.50	-	U	U	**.55**	-	U	T
jacob	.50	-	-	U	**.55**	-	-	T
$c(O\|E_U)$		.27	.43	.51		.31	.48	.57

- if Jacob had confirmed Paul's claim, its degree of justification should have increased.

  → explanations where they both say the truth are confirmed as much as explanations in which they both lie.

Leibniz Center for Law

Extraction of attack/support

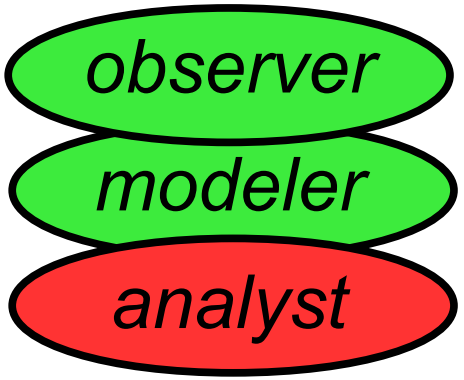
		$c_{max}(f)/c_{max}(-f)$					
		O_1		O_2		O_3	
	P	pre	post	pre	post	pre	post
moustache	.50	.19/.19	.31/.30	.18/.18	.48/.48	.17/.17	.61/.60
eye	.50	.19/.18	.31/.30	.18/.17	.48/.48	.17/.16	.61/.60
jones	.55	.19/.18	.31/.30	.18/.17	.48/.48	.17/.16	.61/.60
paul	.55			.18/.17	.48/.48	.17/.16	.60/.61
jacob	.55					.17/.16	.61/.60

	O_1	O_2	O_3
moustache	$+/+$	$=/-$	$+/+$
eye	$=/+$	$-/-$	$=/+$
jones	$=/+$	$-/-$	$=/+$
paul		$-/=$	$-/-$
jacob			$=/+$

- For each observation, we can refer to two dimensions of change:
  - post O_i − pre O_i
  - post O_i − post O_{i-1}

Conclusion

- With EBA we stress the sharing of a *deep-model* of the domain, a model for the *observation* and the explicitation of strength of commitments for the justification.
  - (Modeling) the observation
  - (Modeling) the deep model
  - Extracting (justified) explanations / AF



Leibniz Center for Law

Conclusion

- We have validated a slightly "deeper model" of reasoning, using Pollock's puzzle with EBA.
- Advantages:
  - defines justification operationally
  - handles neutral prior probability
- Disadvantages:
  - increased overload for the deep-modeling
  - explosion of explanations

Further research

- Investigate other definitions of confirmation values
- Propose an analytical definition for attack/support relations
- Integrate agent-role models into ASP
- Integrate EBA in agent architectures for diagnoser agents
- Integrate Bayesian networks for prior probabilities