Argument and Story Strength - Bayesian vs. Qualitative Approaches

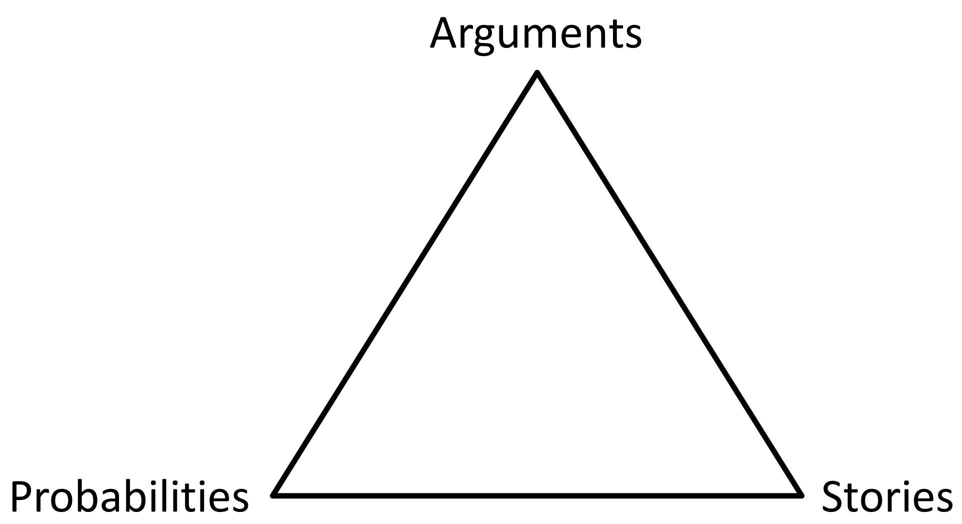
Floris Bex
Utrecht University
Tilburg University

Stories – Arguments – Probabilities

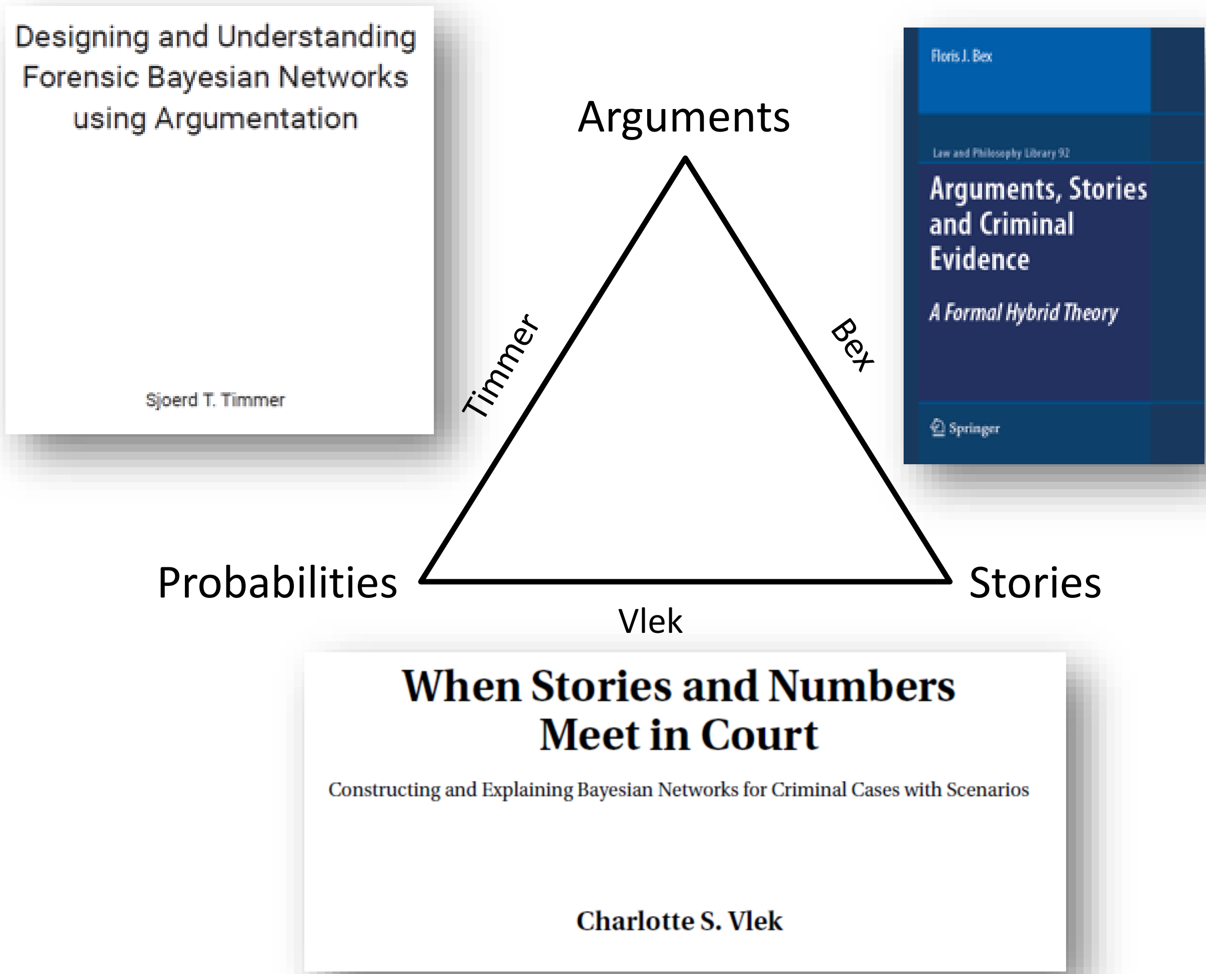
- Explanations are causally coherent sequences of events (*stories*) that explain the evidence in a case.
- Multiple explanations for different conclusions have to be proposed, analysed and compared (*argumentation*), and the “best” (most likely) one should be chosen (*probabilities*)

Stories – Arguments – Probabilities

- Stories: coherent sequences of events
- Arguments: reasons for or against a conclusion
- Probabilities: measure of likelihood that some event has occurred



Stories – Arguments – Probabilities



Stories vs. Arguments

- Stories are “holistic”
- Stories provide an overview
- Stories encapsulate causal reasoning
- Stories represent how humans order a mass of evidence

- Arguments are “atomistic”
- Arguments provide a means of detailed analysis
- Arguments encapsulate evidential reasoning
- Arguments represent how humans talk about individual evidence

Qualitative vs. Quantitative

- Probabilities allow for fine-grained degrees of uncertainty
- Probabilities allow for the correct modelling of probabilistic influences between evidence & events

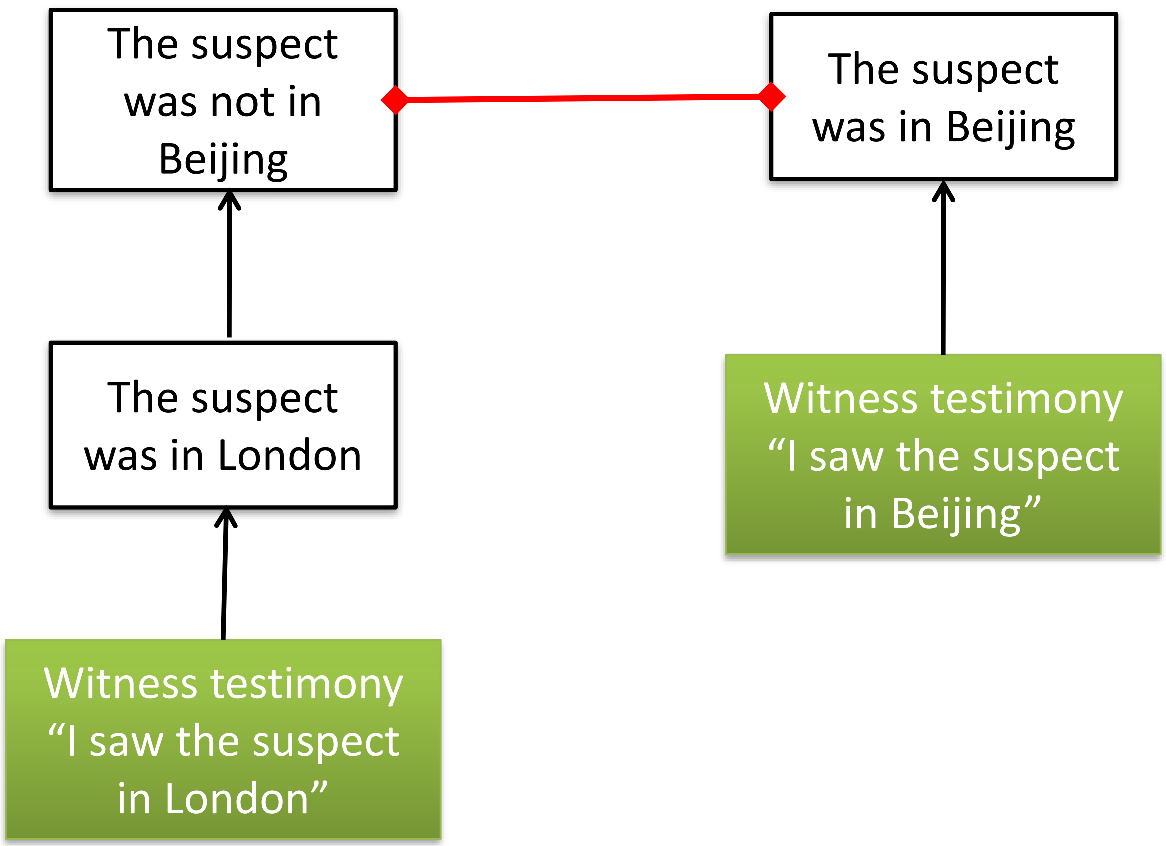
- Qualitative approaches require no precise estimates of probabilities
- Qualitative approaches are closer to how many domain experts reason

Comparing arguments & stories

- There are various pitfalls when reasoning with stories and arguments, but can we measure how good or strong a story or an argument is?

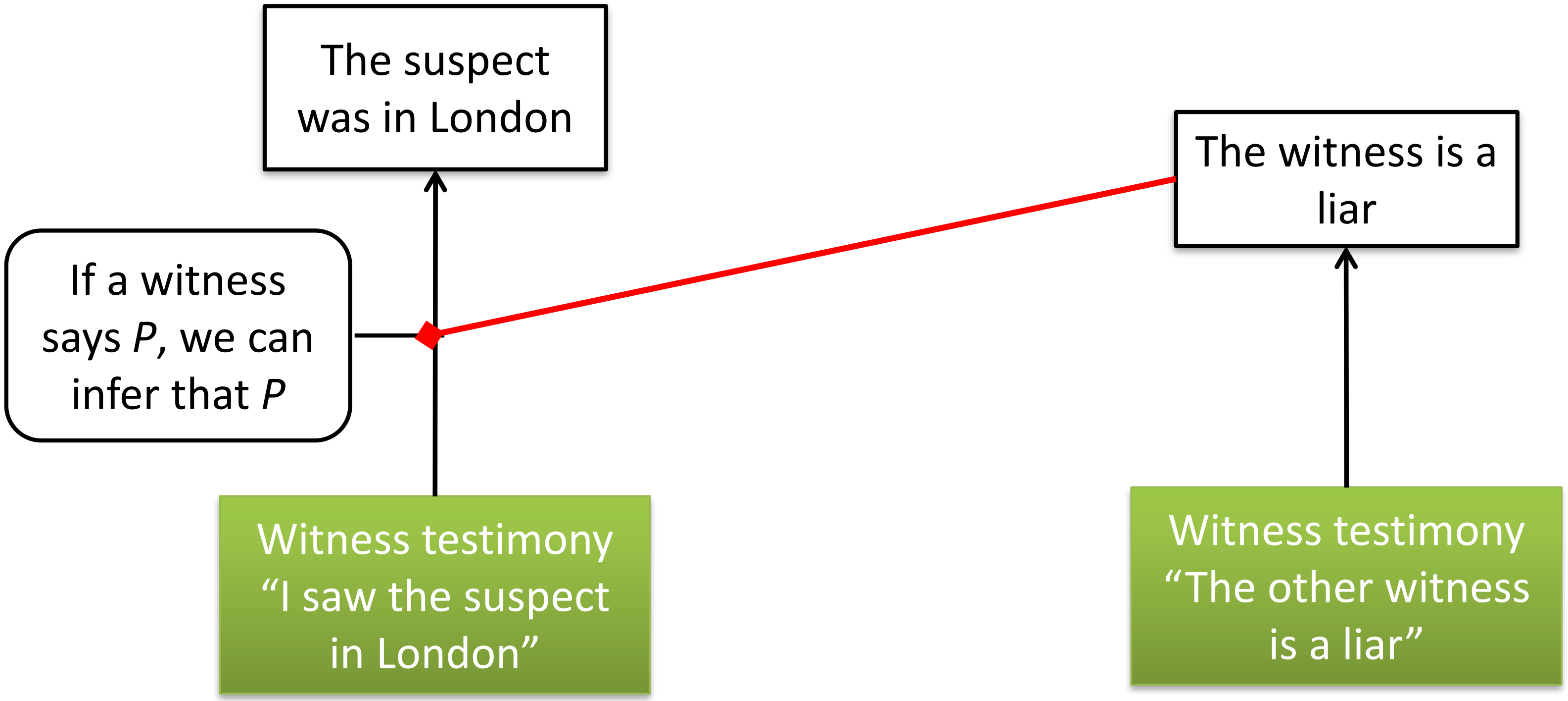
Argument Strength

- Which argument wins?



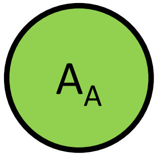
Argument Strength

- Is the attacker strong enough?



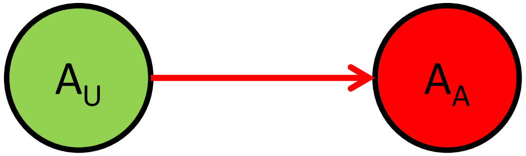
Dialectical semantics

- Dynamically assign status to arguments
  - Status may change if new arguments are put forward



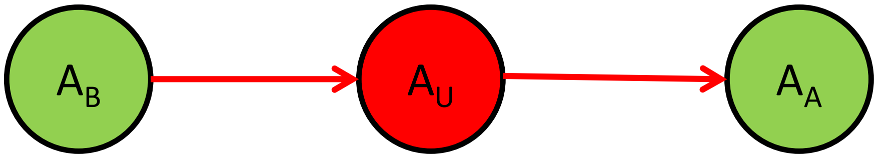
Dialectical semantics

- Dynamically assign status to arguments
  - Status may change if new arguments are put forward



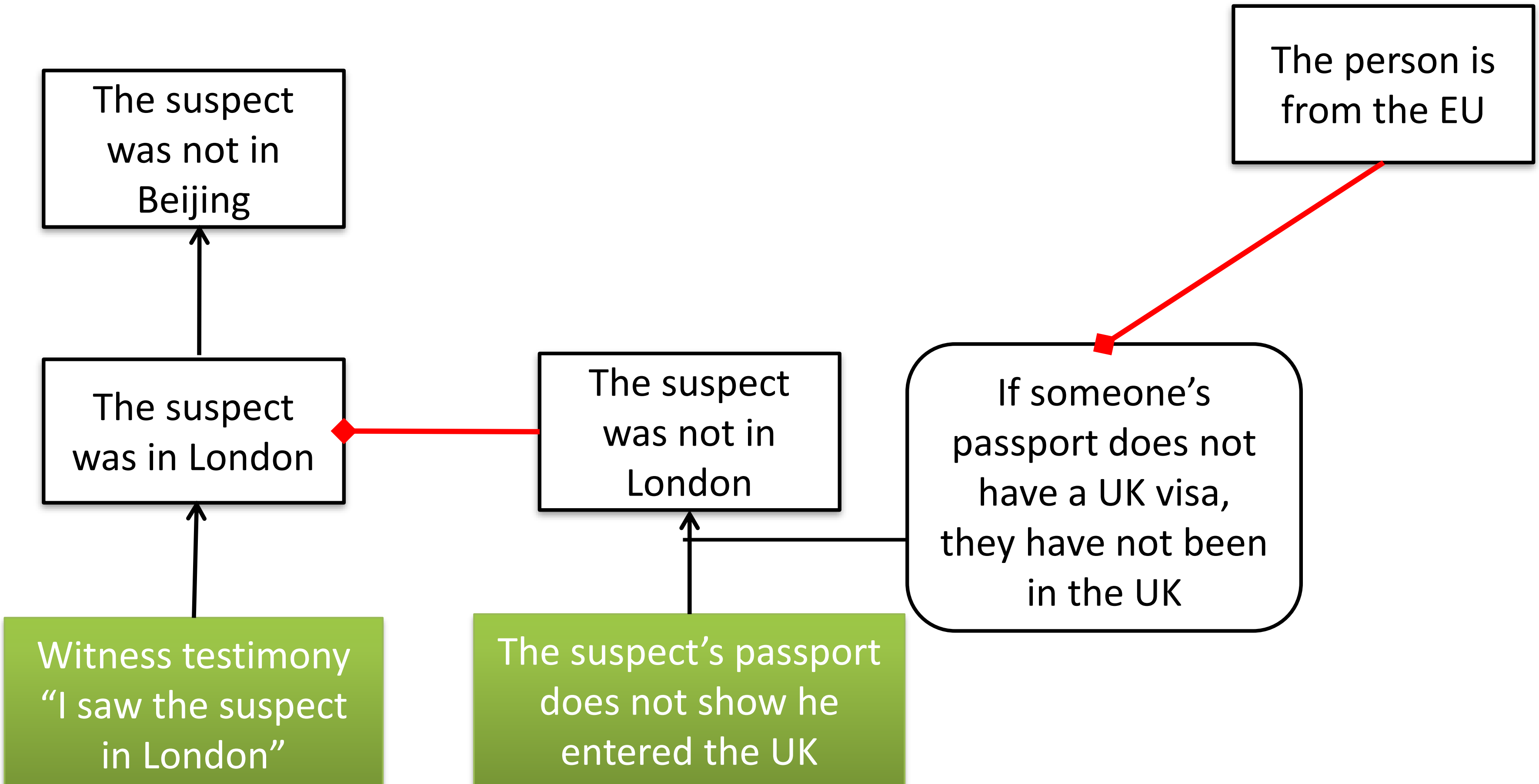
Dialectical semantics

- Keep attacking until you win!



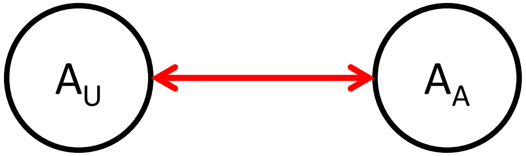
"*The one who has the last word laughs best*"

Reinstatement



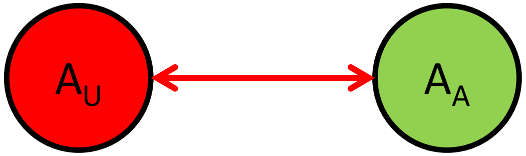
Dialectical semantics

- But how to choose between 2 arguments that attack each other?



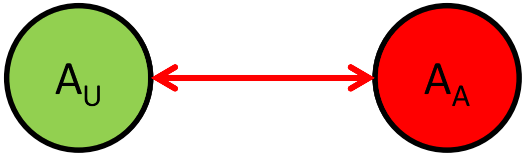
Dialectical semantics

- Strength of arguments
  - $A_U < A_A$ (A_a is preferred over A_U)



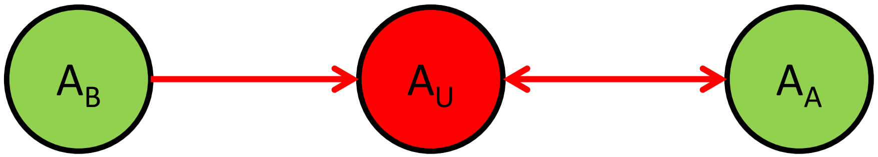
Dialectical semantics

- Strength of arguments
  - $A_U > A_A$ (A_U is preferred over A_A)



Dialectical semantics

- Keep attacking until you win!

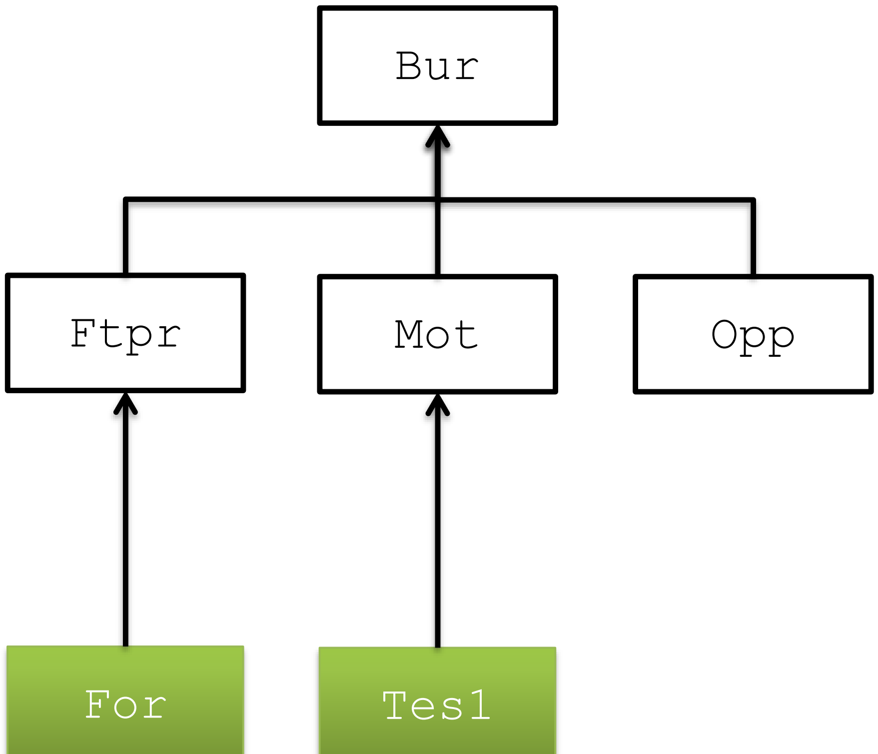


"*The one who has the last word laughs best*"

Reinstatement

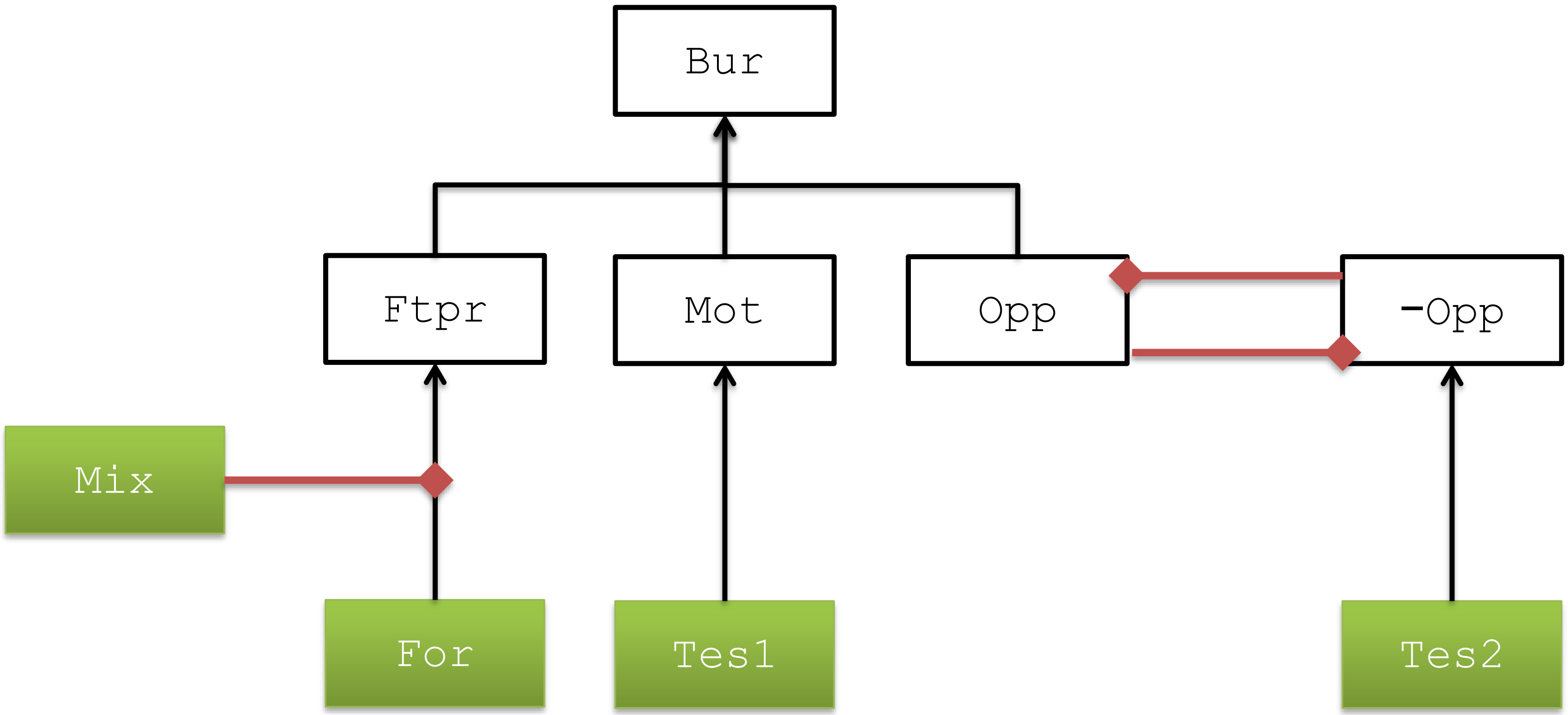
Structured arguments vs. Bayesian Networks

- The burglary (`Bur`) was committed by the suspect, because there is a footprint match (`Ftpr`) and a motive (`Mot`) backed by a report (`For`) and a testimony (`Tes1`), and the suspect has no alibi, so `Opp`.



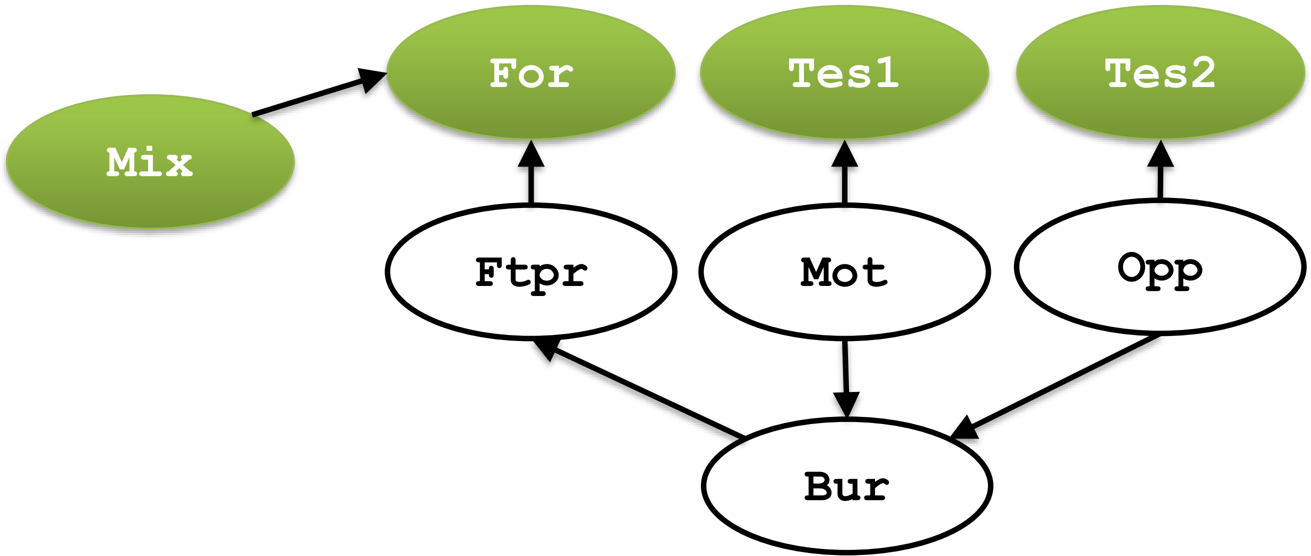
Structured arguments vs. Bayesian Networks

- However, there is evidence of a mixup in the lab (`Mix`), which means the footprint match is not really backed by evidence. Furthermore, the suspect later gave a testimony (`Tes2`) with an alibi, so ¬`Opp`.



Structured arguments vs. Bayesian Networks

- Represent joint probability distribution as DAG + CPT
- Directed Acyclic Graph
  - Nodes are variables **Bur** = [Bur, ¬Bur]
  - Arcs represent probabilistic dependencies between nodes (**Mot**, **Bur**)

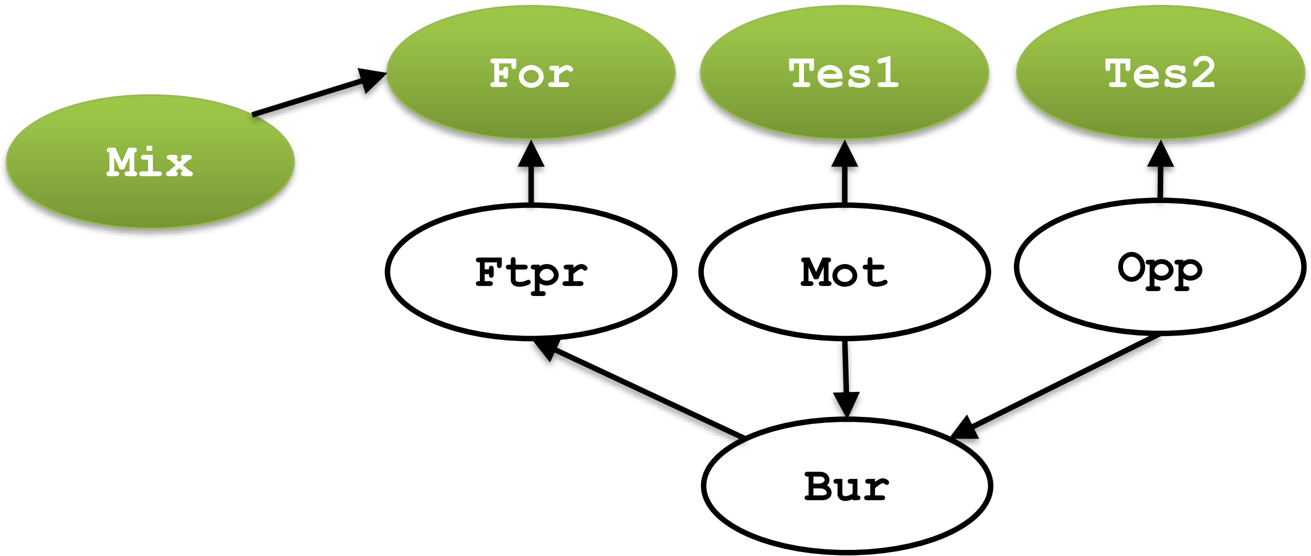


Probabilistic reasoning

- Probability of events and the links between evidence/events
- Probability of a proposition (event) being true or false
  - $P(e)$, $P(\neg e)$
  - $P(e) + P(\neg e) = 1$
- Conditional probability of e given evidence ev
  - $P(e \mid ev)$
- Probability of observed variable (evidence) = 1

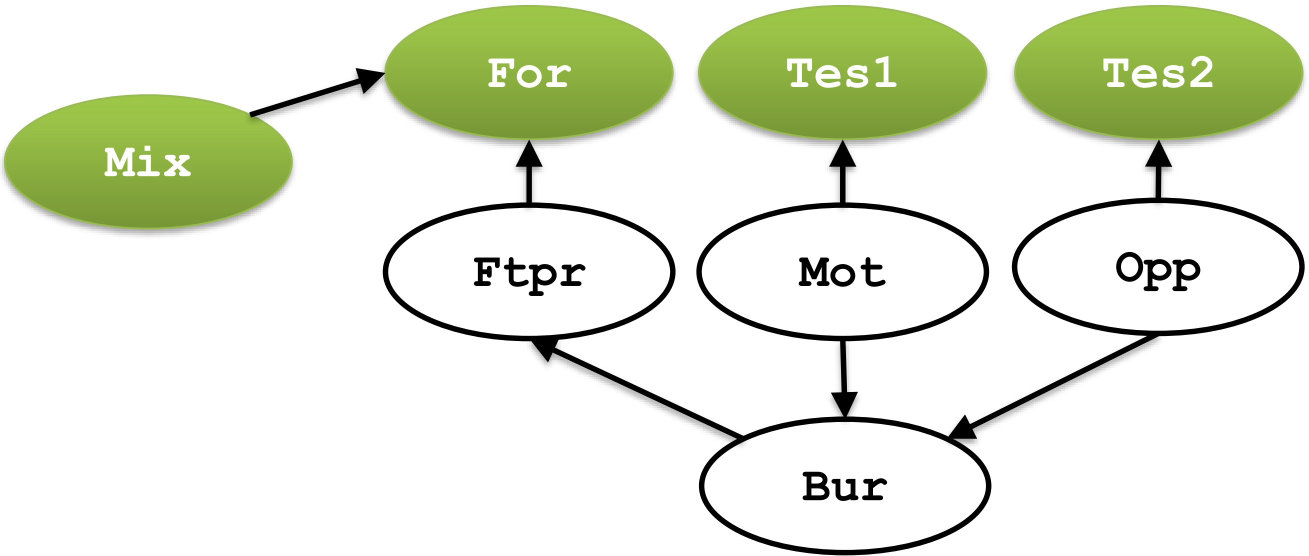
Bayesian Networks

- (Conditional) probabilities
  - Pr(Mot)=0.4; Pr(−Mot)=0.6;
  - Pr(Ftpr | Bur)=0.8; Pr(−Ftpr | Bur)=0.2
    Pr(Ftpr | −Bur)=0.01; Pr(− Ftpr | −Bur)=0.99
  - Pr(Tes1) = 1



Bayesian Networks

- Given the evidence and all the probabilities, we can precisely calculate the posterior probability of the conclusion (Bur)



Inference to the Best Explanation

- Given observations, hypothesise possible explanations
  - I have a cough – cold or flu?
  - Computer fails to start – why?
  - Body found – what happened?

- Choose the “best” explanation
  - Strongest explanation
- How to determine strength of explanations?
  - Using argumentation? Using Bayesian networks?

Formal IBE

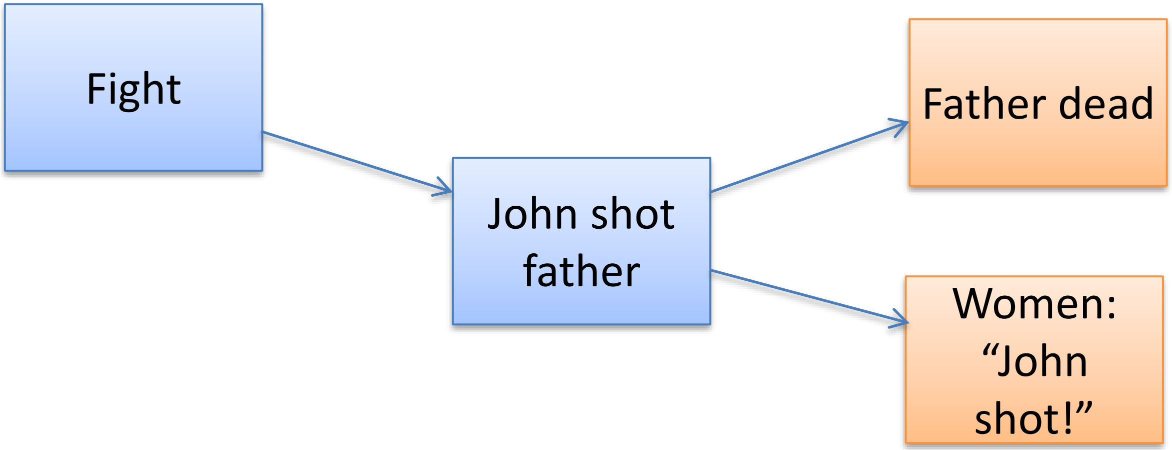
- Given a set of observations O

Father dead

Women: “John shot!”

Formal IBE

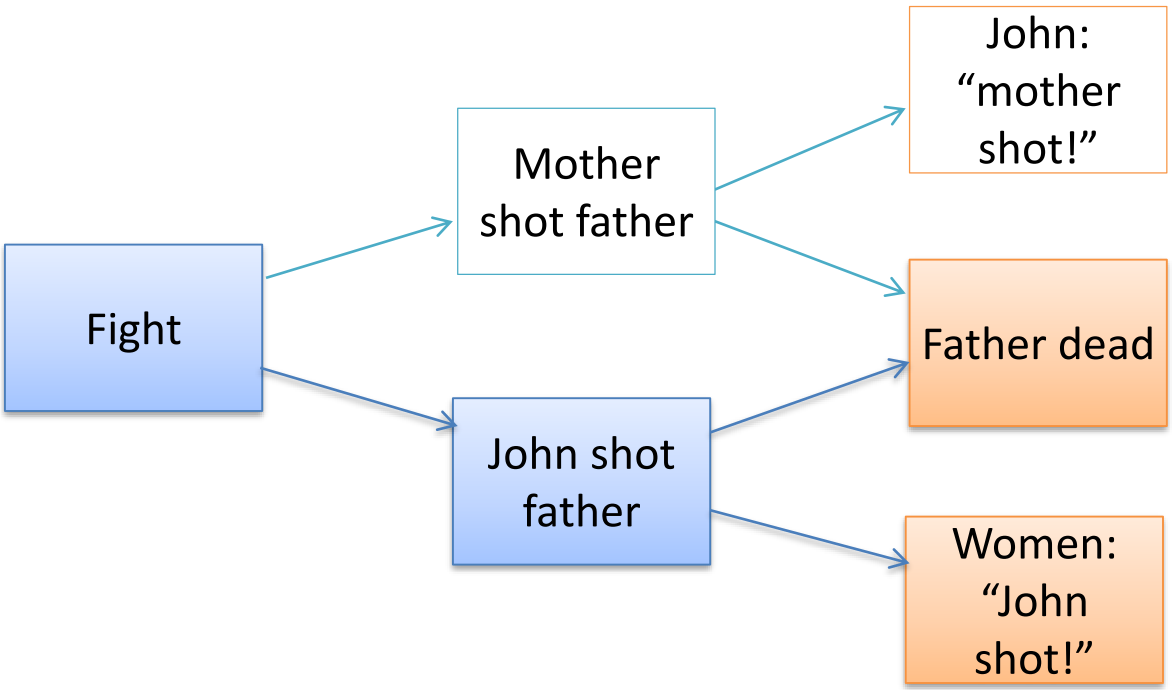
- Assume hypothesis H and rules R s.t. H,R ⊢ O



Abductive IBE – Console & Torasso, Poole

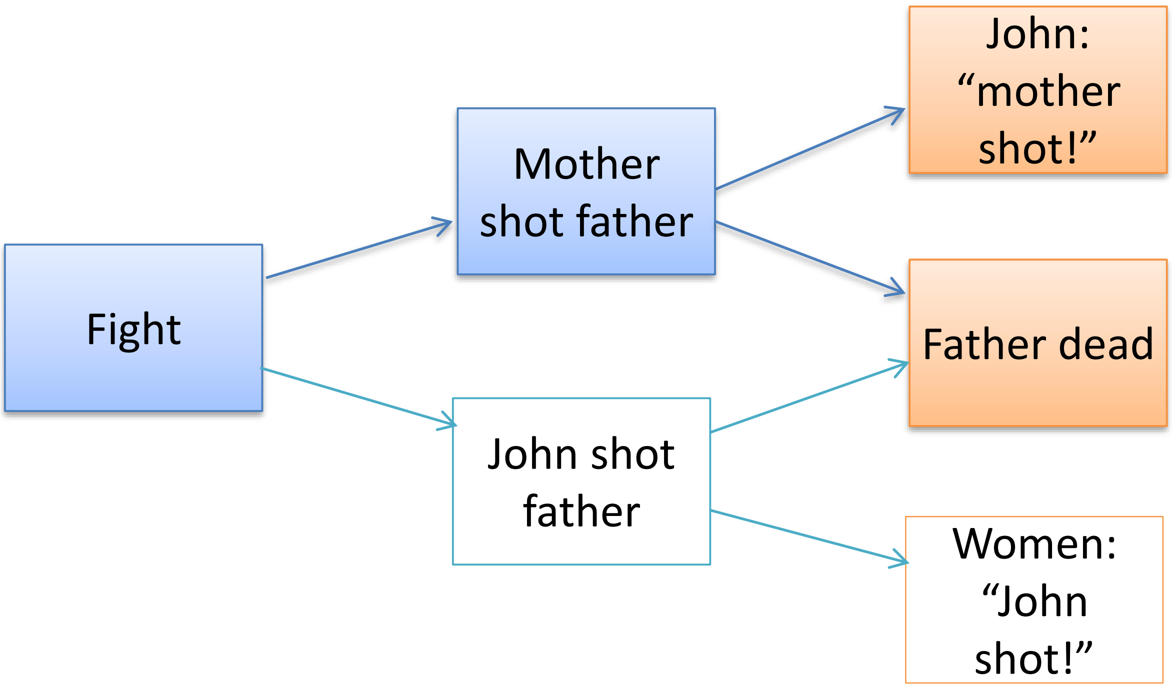
Formal IBE

- Alternative explanations



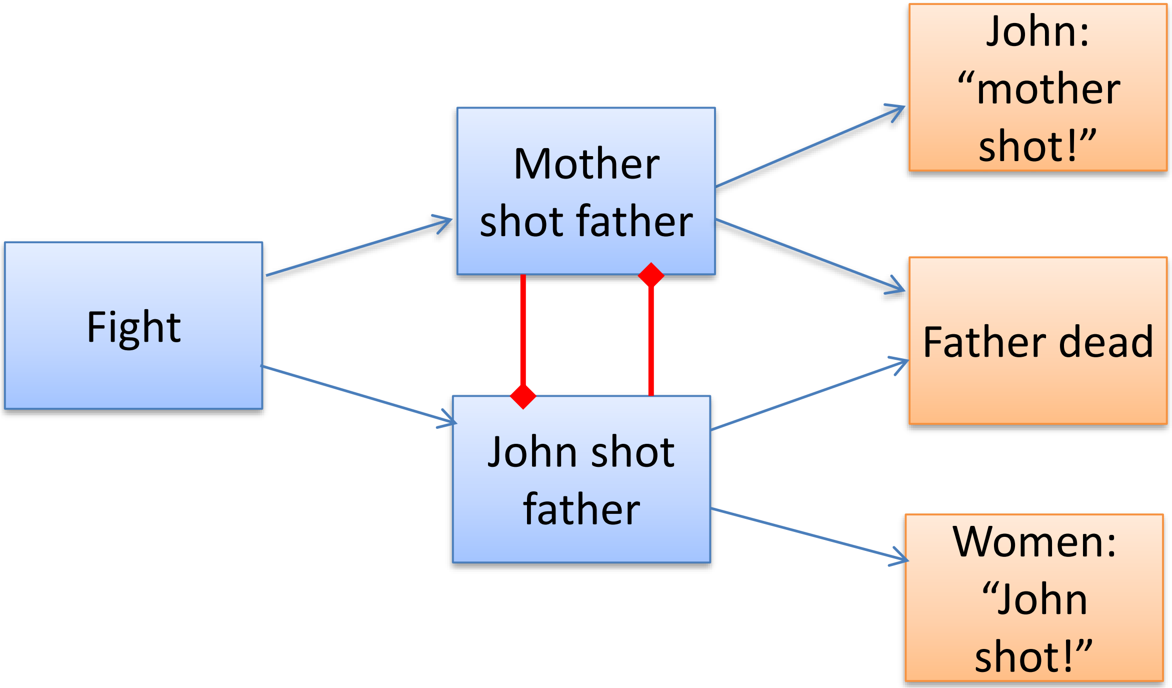
Formal IBE

- Alternative explanations



Argumentative IBE

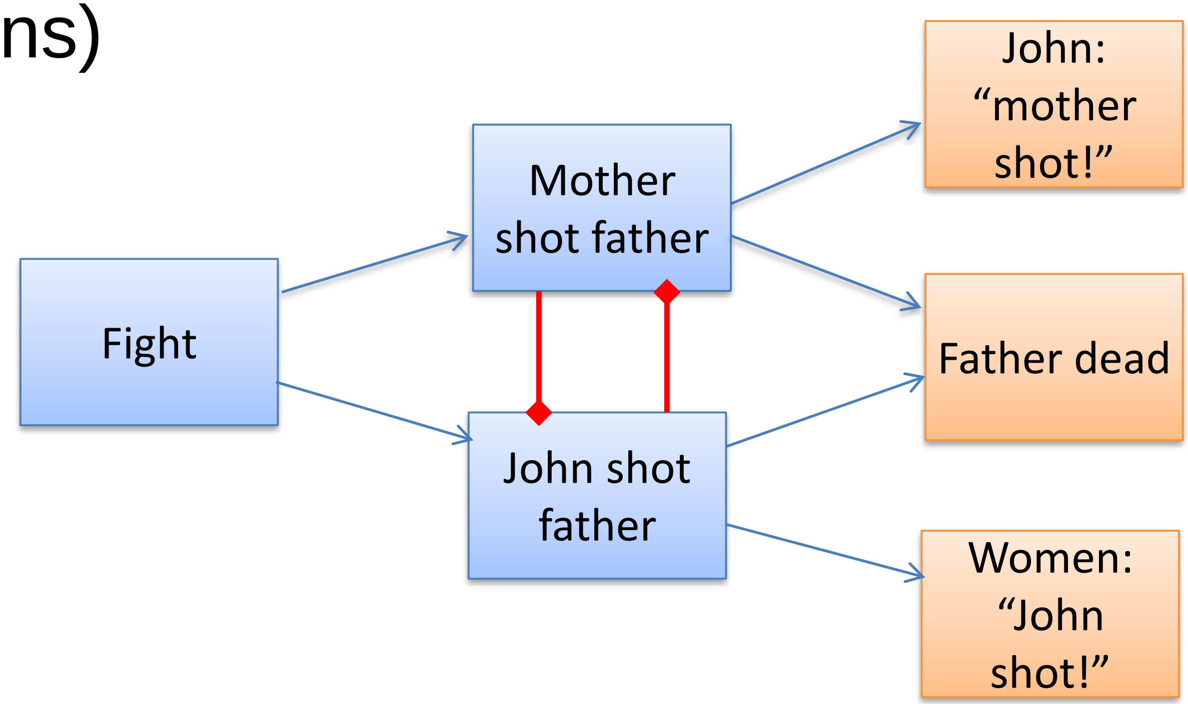
- Defeasible explanations (i.e. H, R |~ O)
- Explanations as contradictory arguments



Default Reasoning – Poole; ABA – Bondarenko et al.

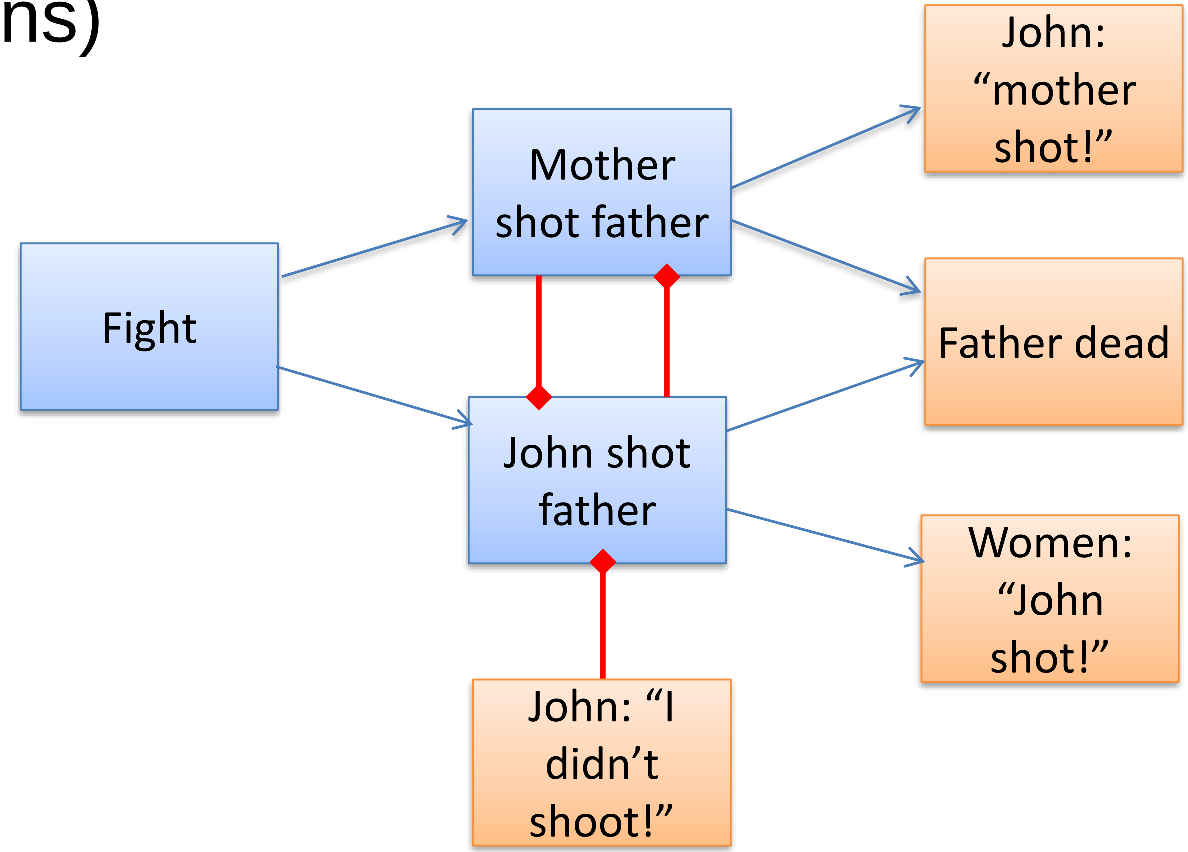
Argumentative IBE

- Explanations themselves can be attacked/supported by arguments (based on observations)



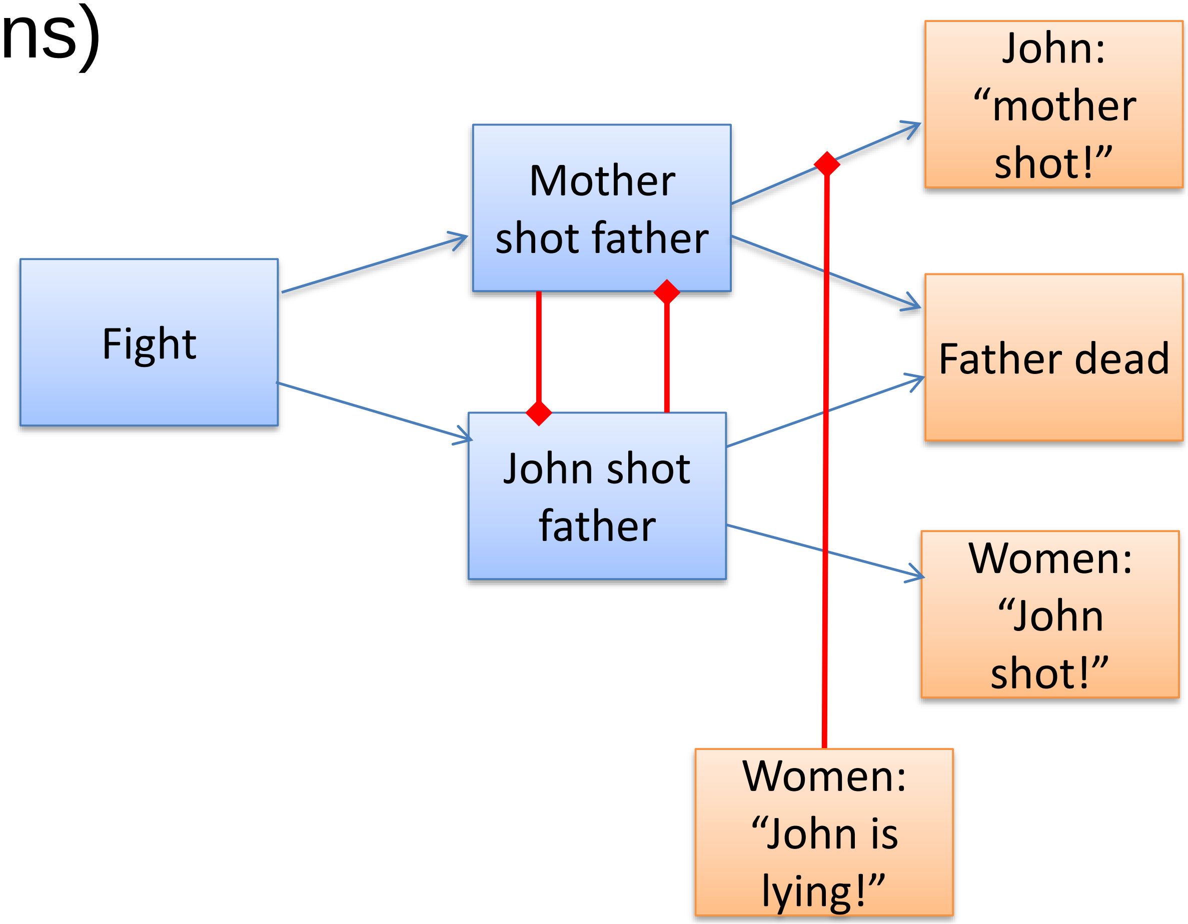
Argumentative IBE

- Explanations themselves can be attacked/supported by arguments (based on observations)



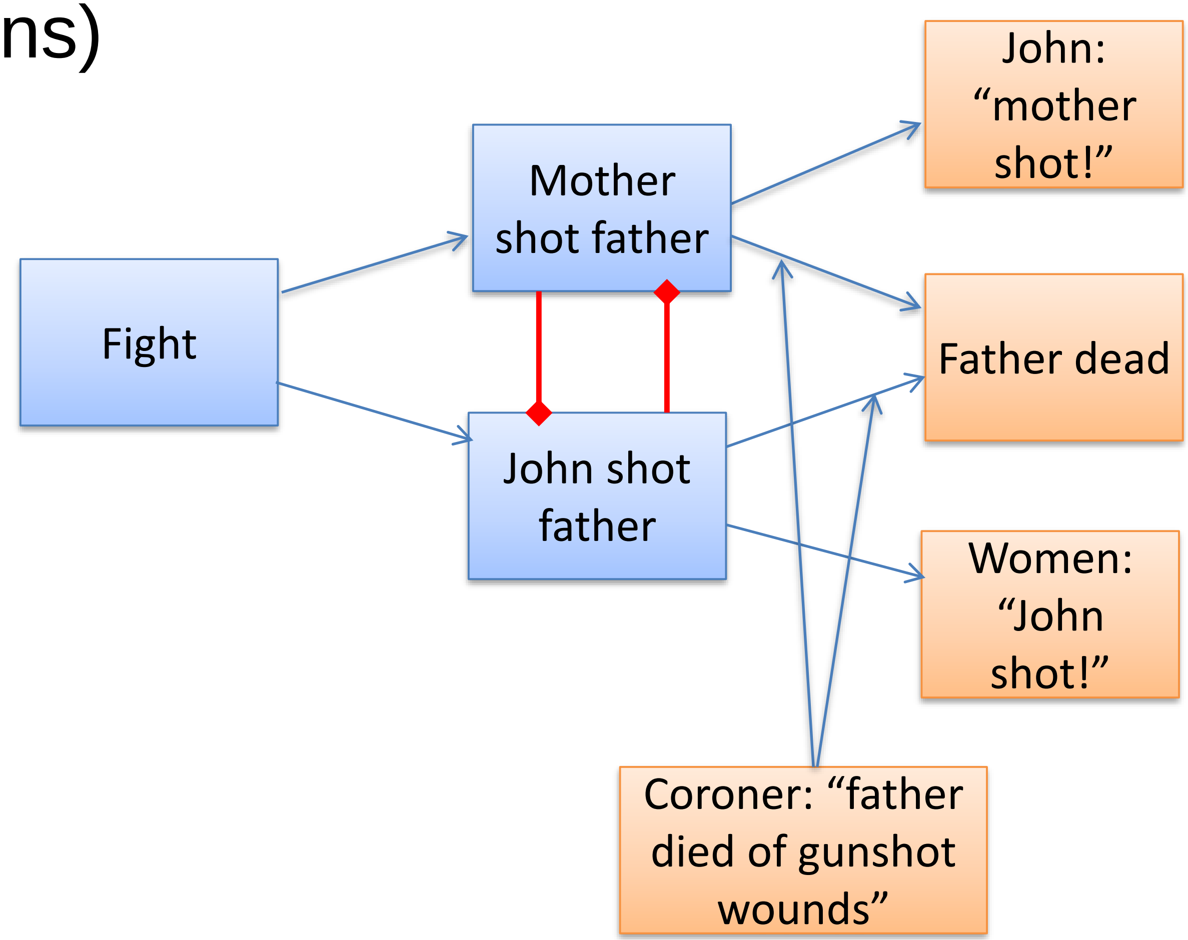
Argumentative IBE

- Explanations themselves can be attacked/supported by arguments (based on observations)



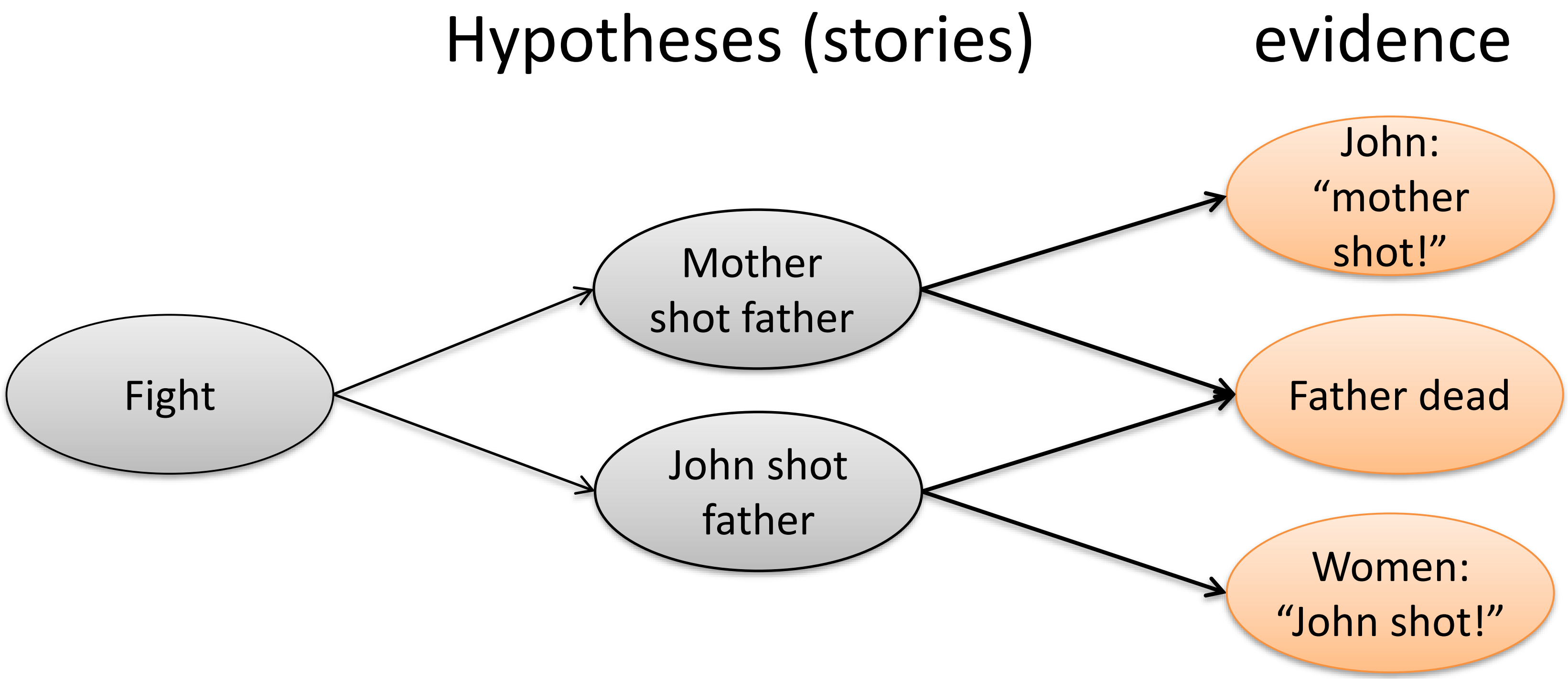
Argumentative IBE

- Explanations themselves can be attacked/supported by arguments (based on observations)



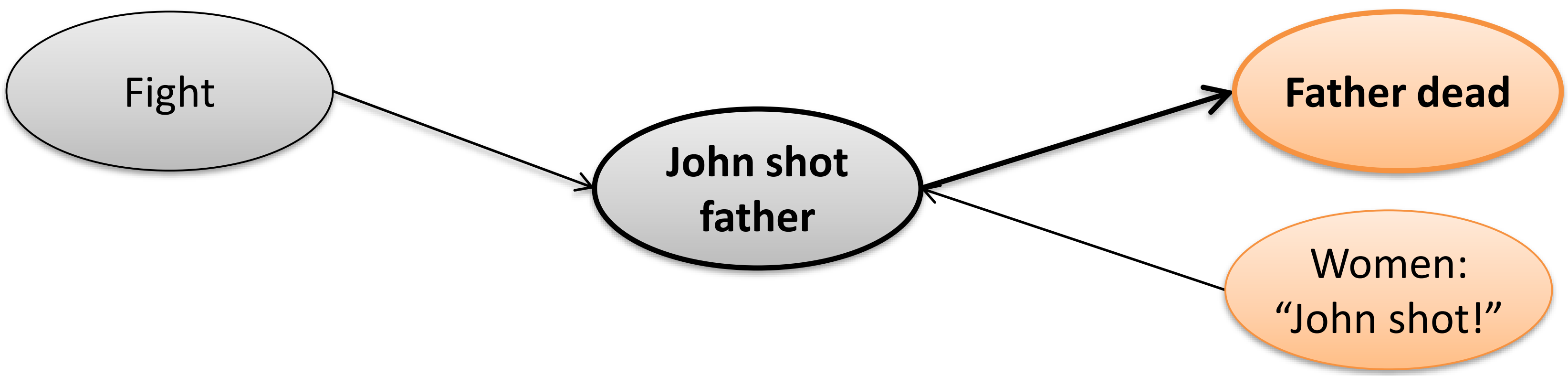
Capturing IBE
Structure

- Alternative stories



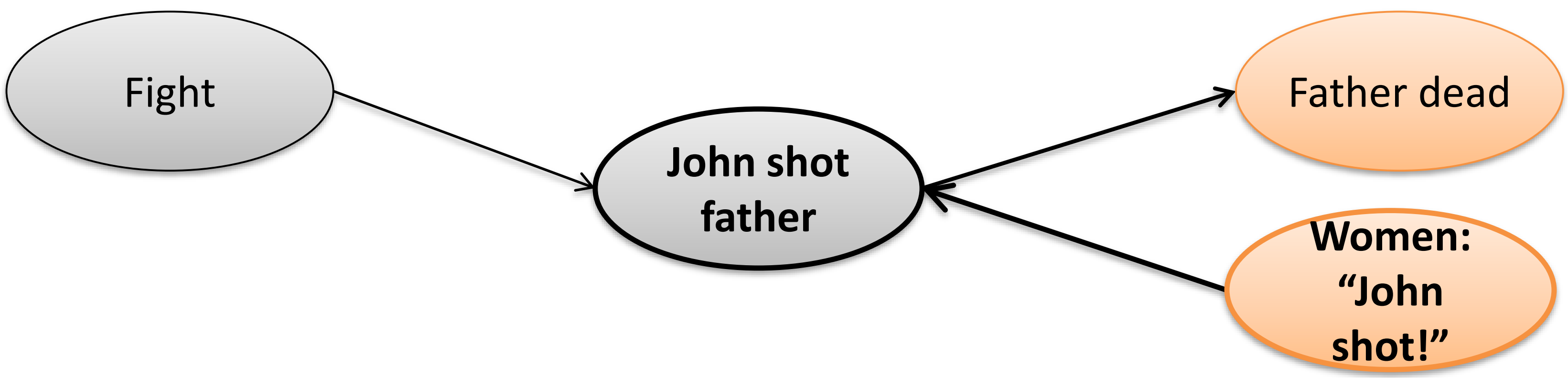
Capturing IBE
Structure

- Causal reasoning:
  – John shooting father causes father to die
    The story explains the evidence



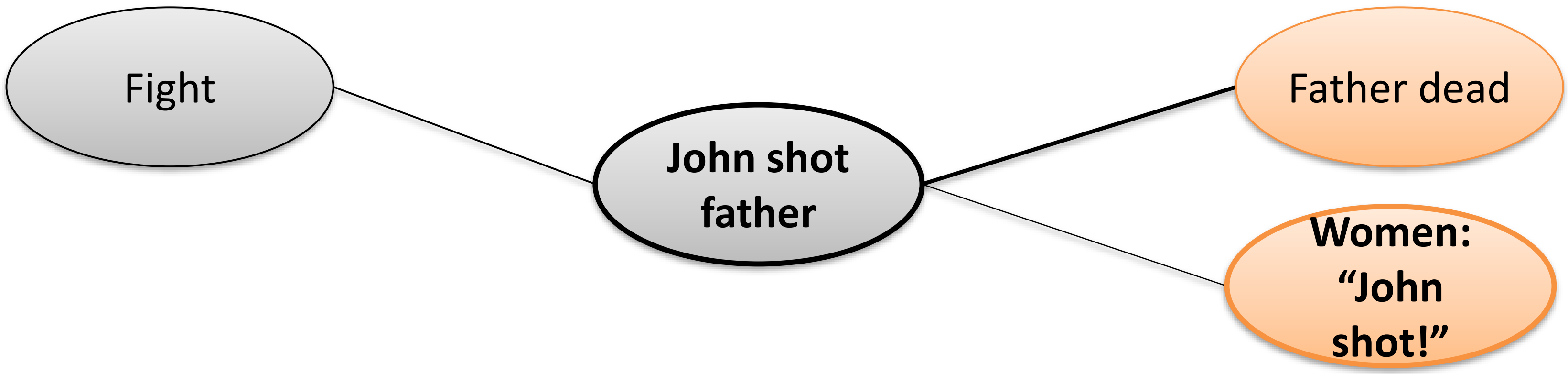
Capturing IBE
Structure

- Evidential reasoning:
  - Women saying "John shot father" <u>is evidence for</u> John shot father
    Testimony <u>supports</u> the story



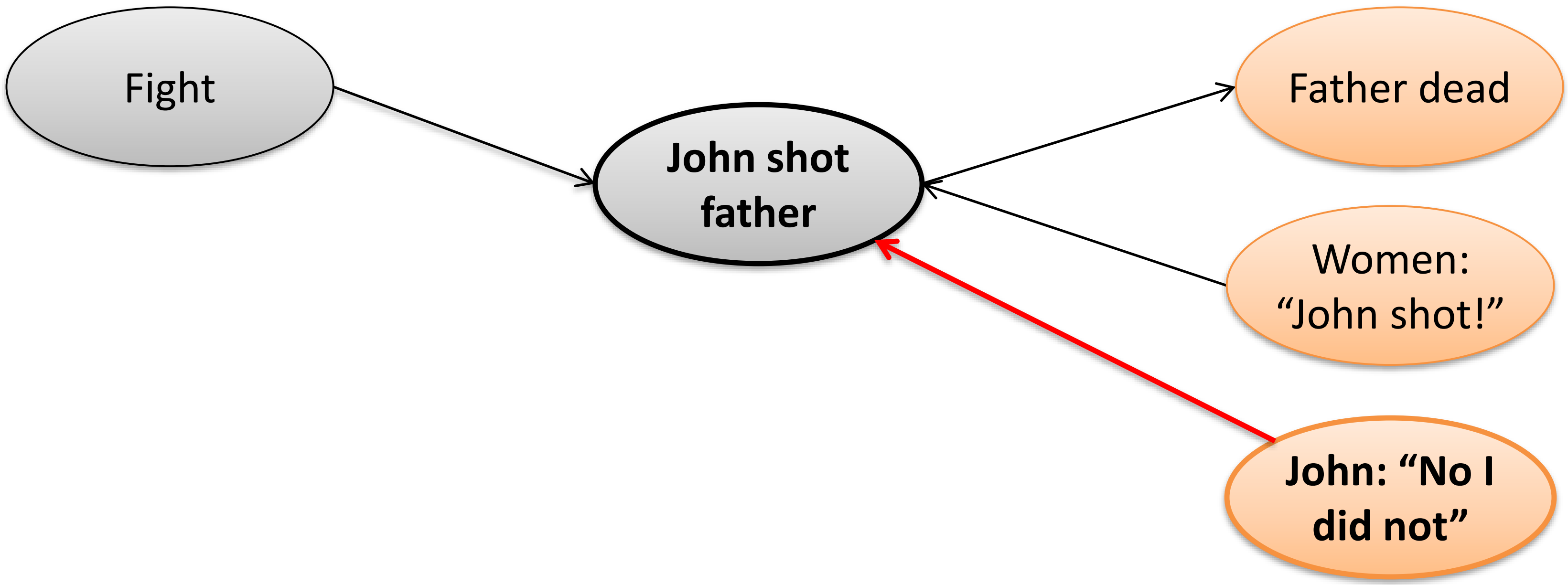
Capturing IBE
Structure

- Directions of arrows (inference) does not matter!



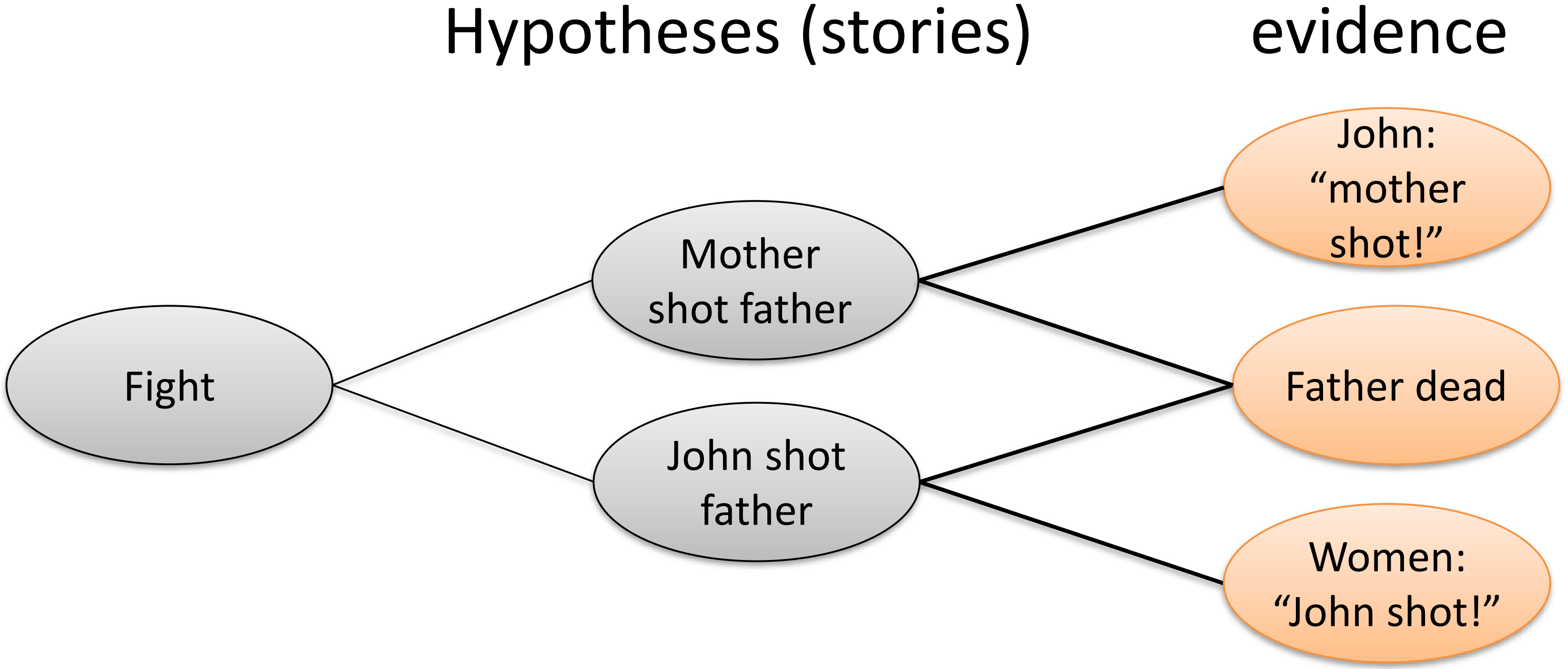
Capturing IBE
Structure

- Contradictory evidence
  – John's denial <u>attacks</u> the fact that John shot father
  The evidence <u>contradicts</u> the story



Capturing IBE
Structure

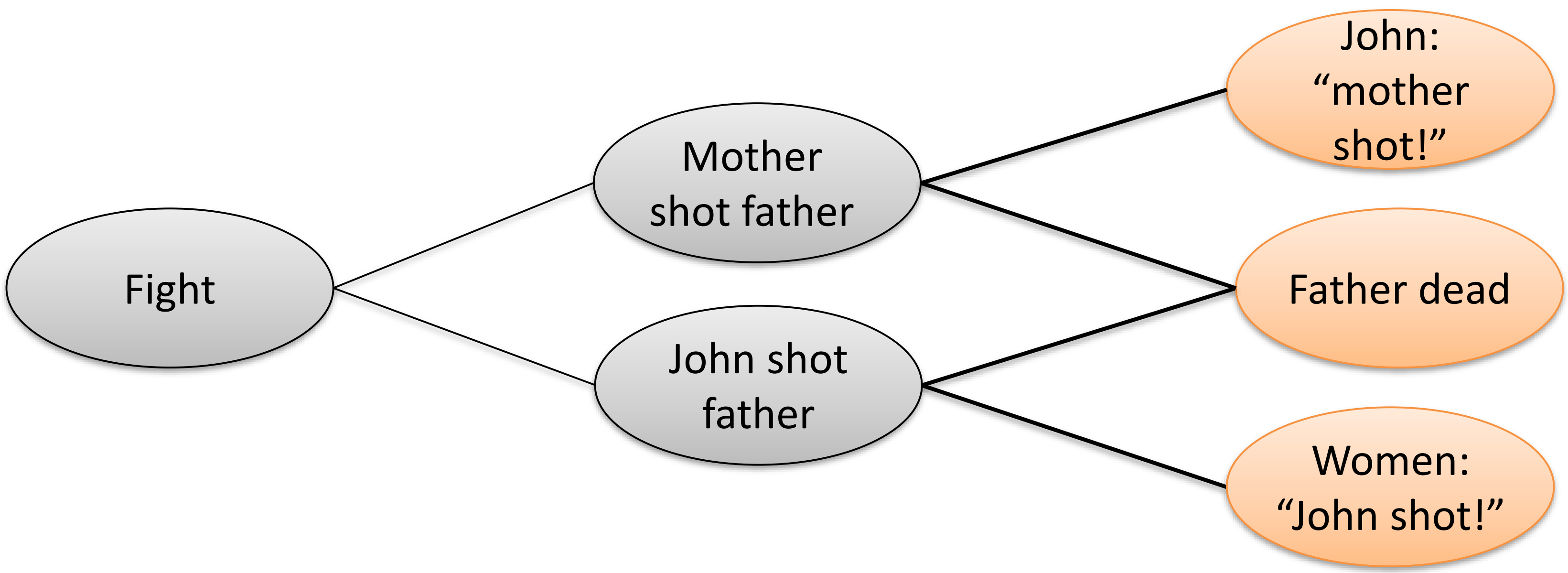
- Can be sets of (logical) propositions with support (argumentation) and causal (story) links



Capturing IBE
Structure

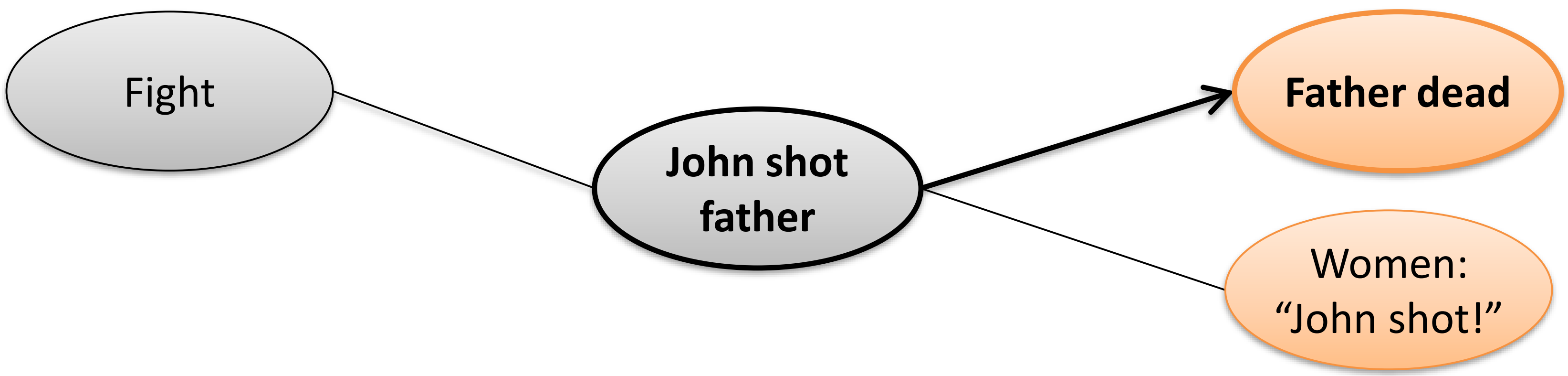
- But also a Bayesian Network where nodes represent variables and link dependencies

Hypotheses (stories) evidence



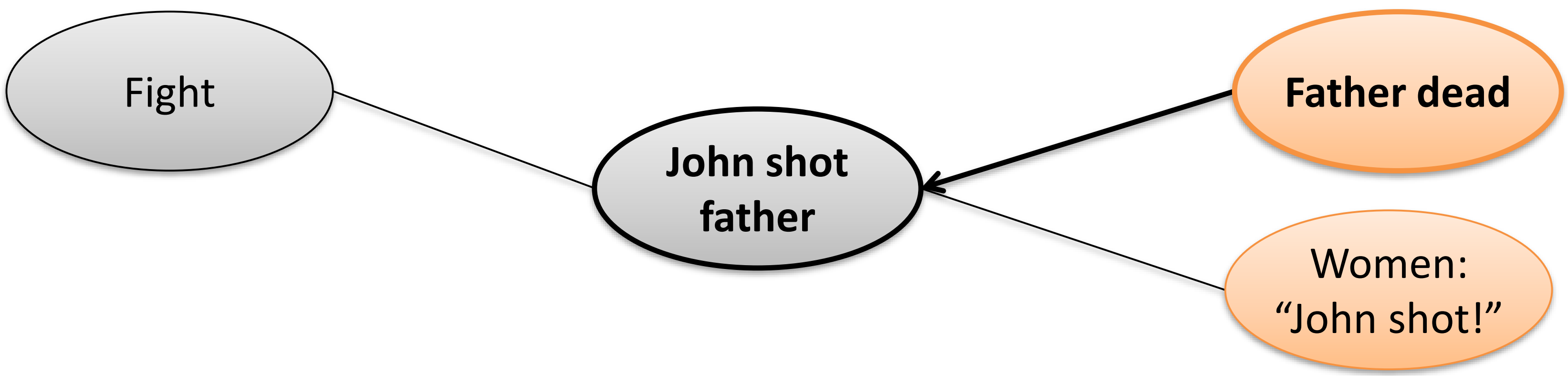
Capturing IBE
Adding probabilities

- Conditional probabilities
  - $Pr(f\_dead \mid J\_shot) + Pr(\neg f\_dead \mid J\_shot) = 1$
    $Pr(f\_dead \mid \neg J\_shot) + Pr(\neg f\_dead \mid \neg J\_shot) = 1$
- Depends on direction of arrow



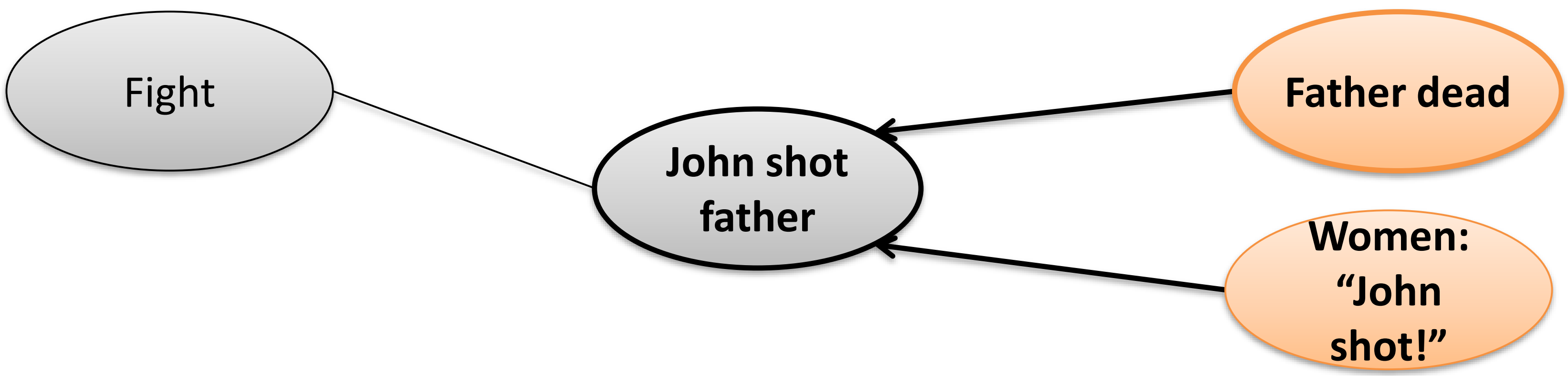
Capturing IBE
Adding probabilities

- Conditional probabilities
  - $Pr(J\_shot \mid f\_dead) + Pr(\neg J\_shot \mid f\_dead) = 1$
    $Pr(J\_shot \mid \neg f\_dead) + Pr(\neg J\_shot \mid \neg f\_dead) = 1$
- Depends on direction of arrow



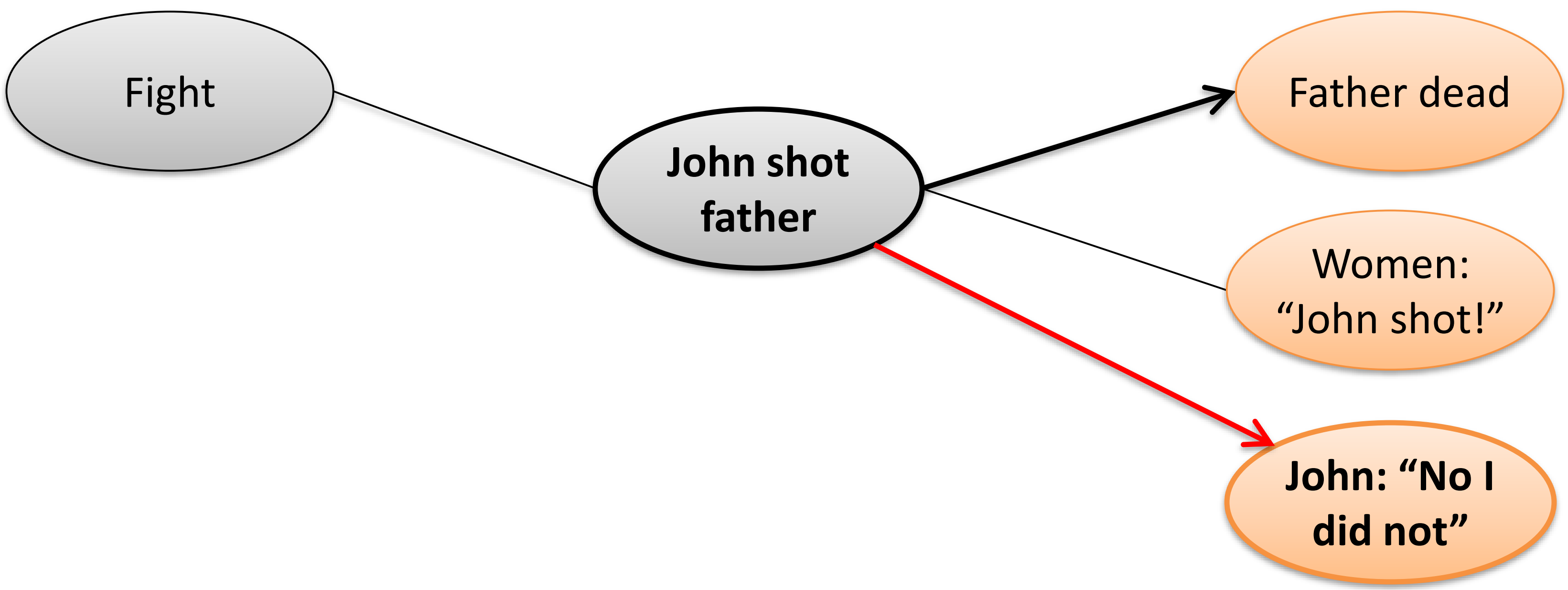
Capturing IBE
Adding probabilities

- Conditional probabilities
  - Pr(J_shot | f_dead, women) + Pr(¬J_shot | f_dead, women) = 1
  - Pr(J_shot | f_dead, ¬women) + Pr(¬J_shot | f_dead, ¬women) = 1
  - Pr(J_shot | ¬f_dead, women) + Pr(¬J_shot | ¬f_dead, women) = 1
  - Pr(J_shot | ¬f_dead, ¬women) + Pr(¬J_shot | ¬f_dead, ¬women)=1



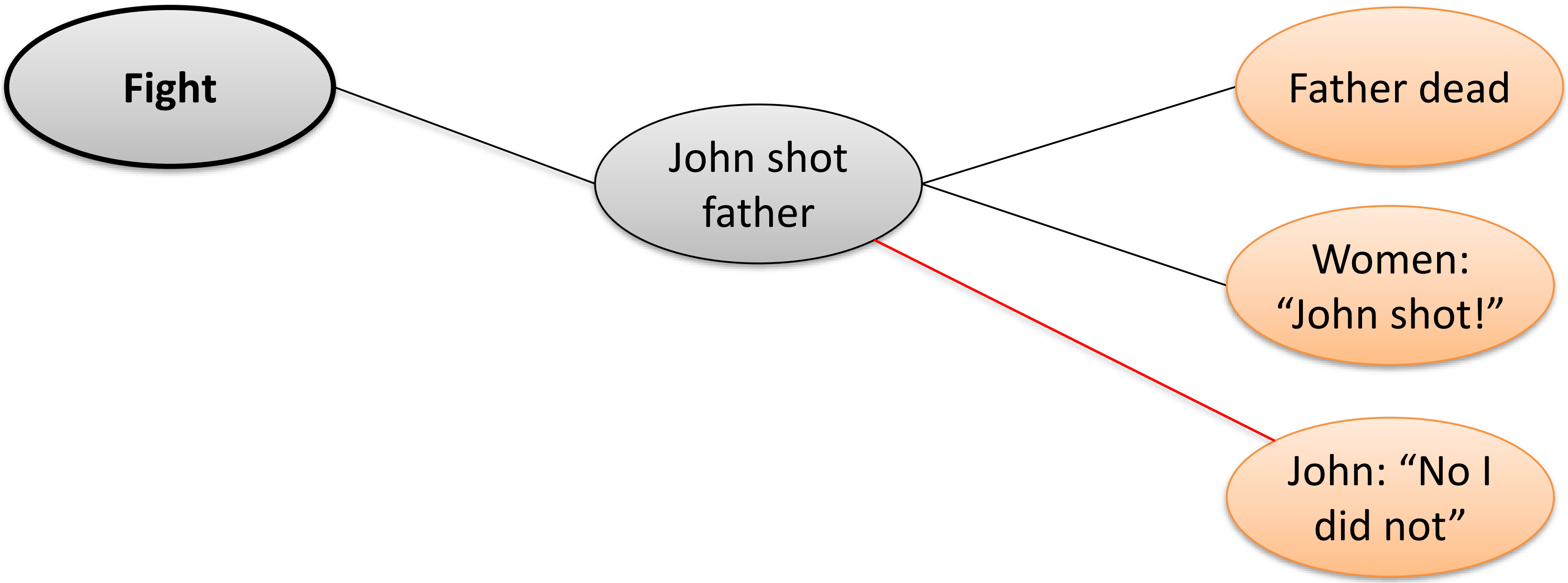
Capturing IBE
Support vs attack

- Supporting evidence
  - Pr(f_dead | J_shot) > Pr(f_dead | ¬J_shot)
- Attacking evidence
  - Pr(J_denial | ¬J_shot) > Pr(J_denial | J_shot)



Capturing IBE
Adding probabilities

- Prior probabilities
  - $\Pr(\text{Fight}) + \Pr(\neg\text{Fight}) = 1$
    The prior probability that a fight breaks out



Strength of explanations

- Stories need to be compared
  - How well do they conform to the evidence?
  - How coherent are they of themselves?
- A good/strong story is complete, plausible and conforms to much of the important evidence

Story model for juror decision making – Pennington & Hastie

Strength of explanations

- Evidential Coverage
  - Evidential Support: how much of the evidence supports the story (is explained by it)?
  - Evidential Attack: how much of the evidence attacks the story (is contradicted by it)?
- Completeness
  - Does the story mention all the relevant events we expect to see?
- Plausibility
  - Are the story and its elements plausible (irrespective of the evidence)?
- Consistency
  - Is the story consistent?

Strength of explanations

- Evidential Coverage
  - Evidential support
  - Evidential attack
- Completeness
- Plausibility
- Consistency

- Given these elements of story strength, we can
  - Reason about them (*Argumentation*)
  - Measure them (*Probabilities)*

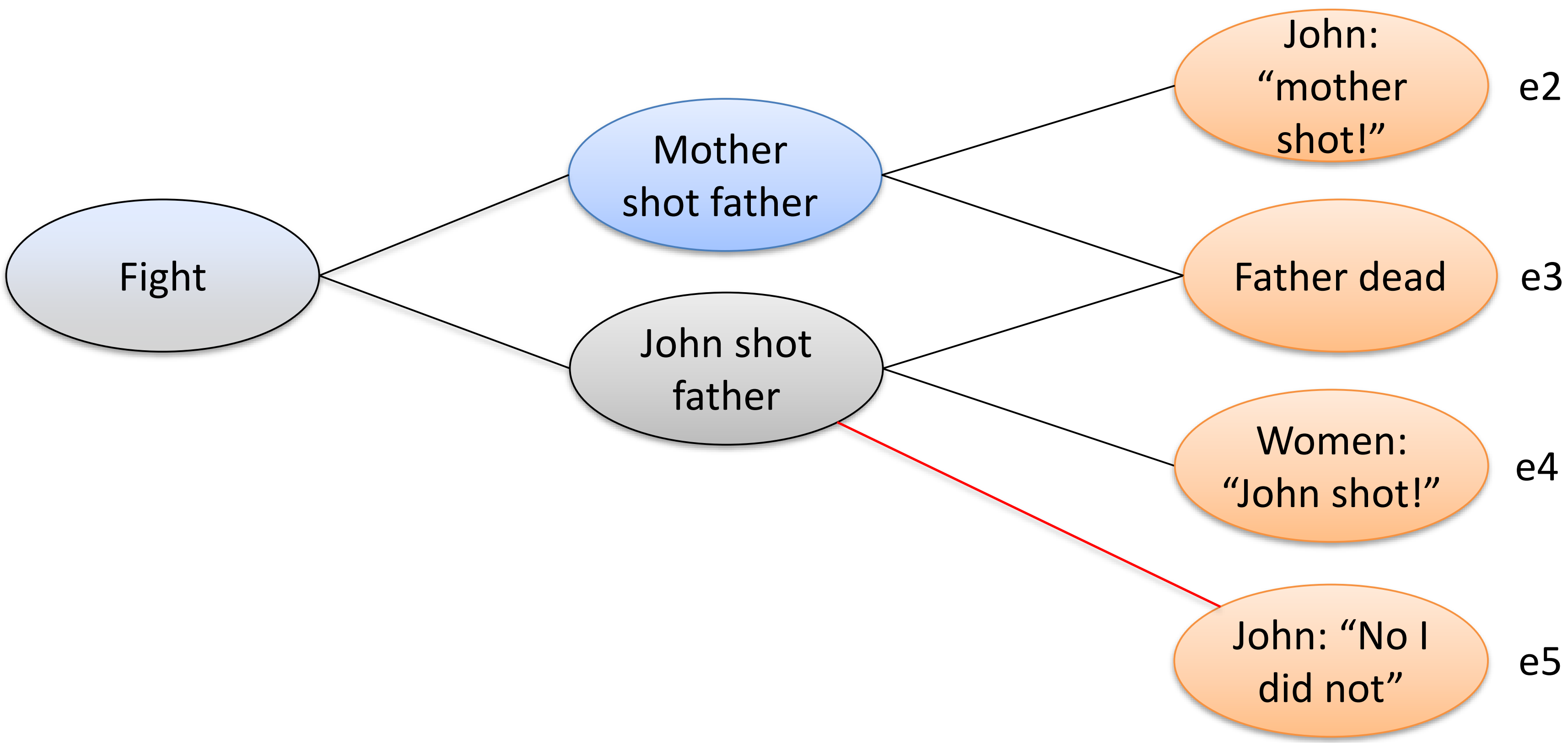
Strength of explanations
Evidence

- Reasoning with evidential support and attack

- Check which evidence directly supports or attacks a story
  - Support: the evidence supporting a story
  - Attack: the evidence attacking a story

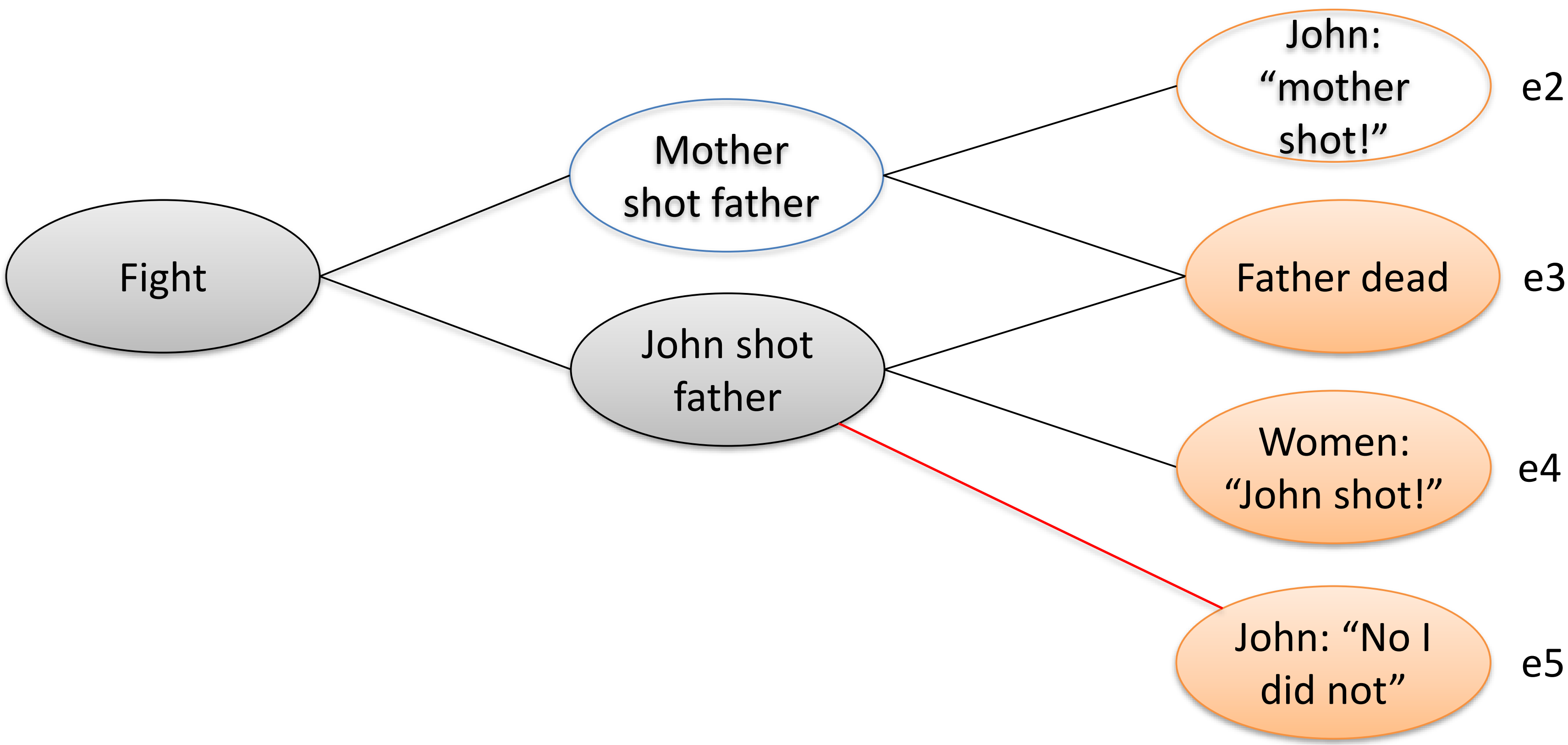
- What are the differences with other (competing) stories?
  - Which evidence does my story not (yet) explain?
  - Which attacks do I need to respond to?

Support and Attack
Reasoning



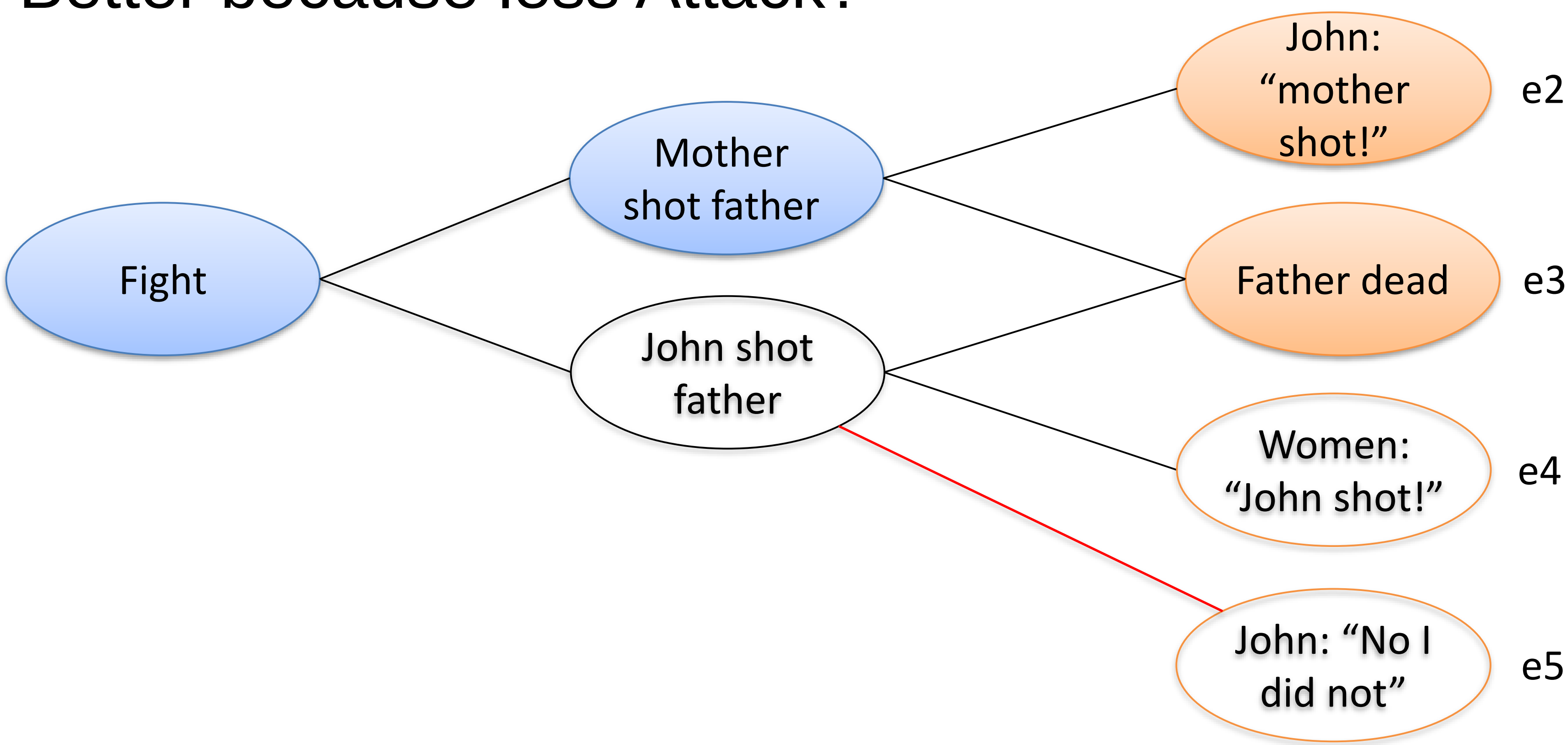
Support and Attack
Qualitative interpretations

- Support: {e3,e4}
- Attack: {e5}



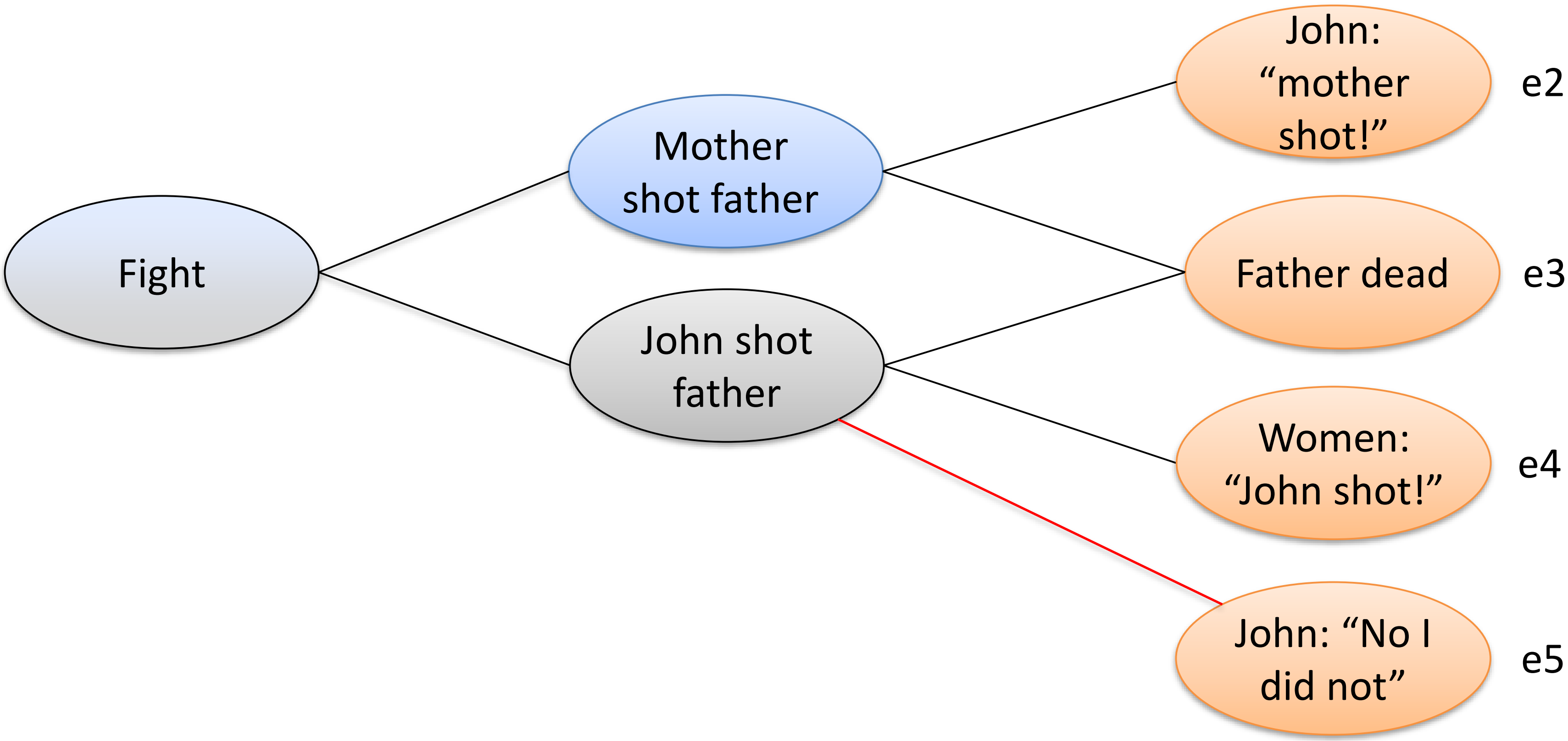
Support and Attack
Qualitative interpretations

- Support = {e2,e3}
- Attack = {}
- Better because less Attack?



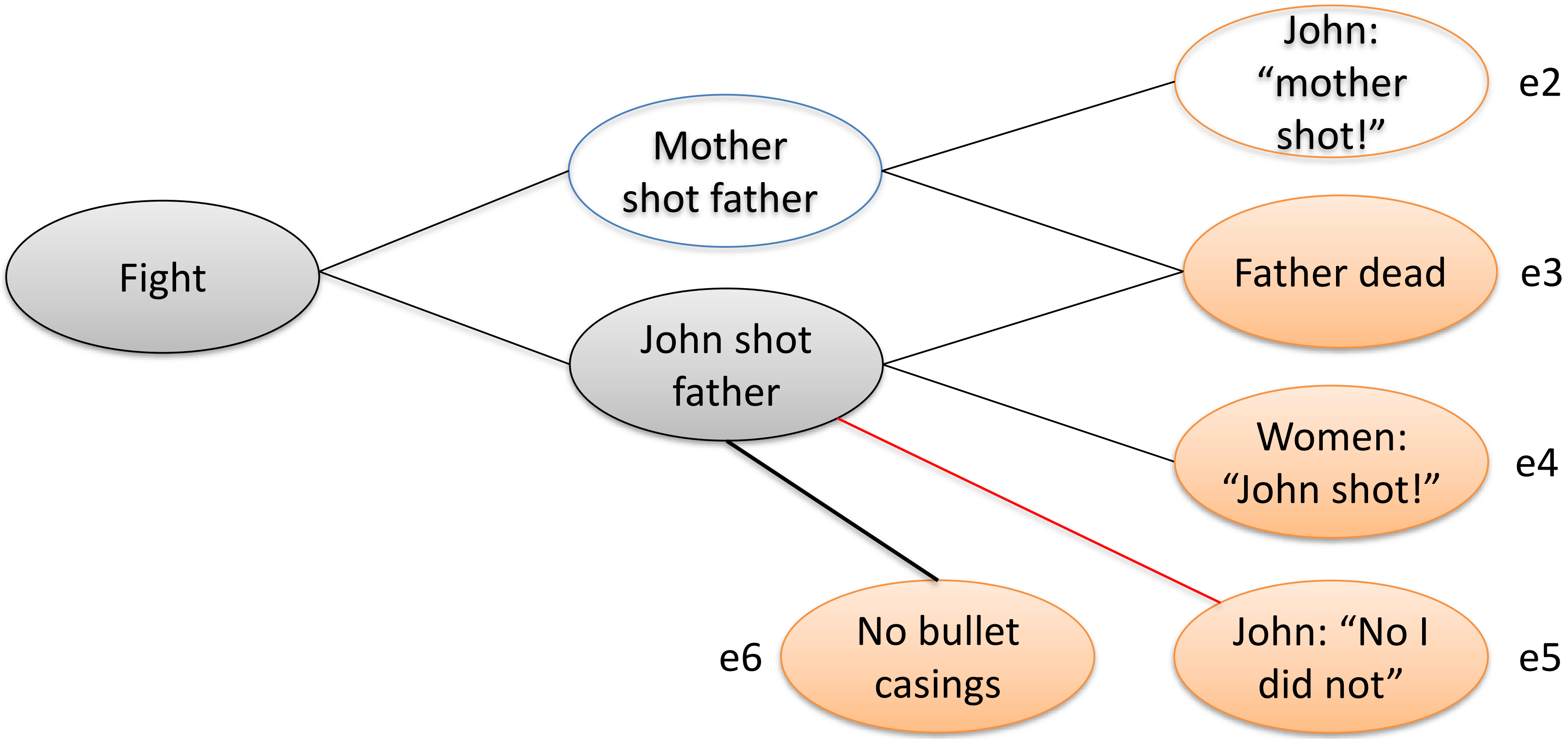
Support and Attack
Reasoning

- (J) has a support of two pieces of evidence, same as (M), and is attacked by 1 piece of evidence while (M) is not attacked.



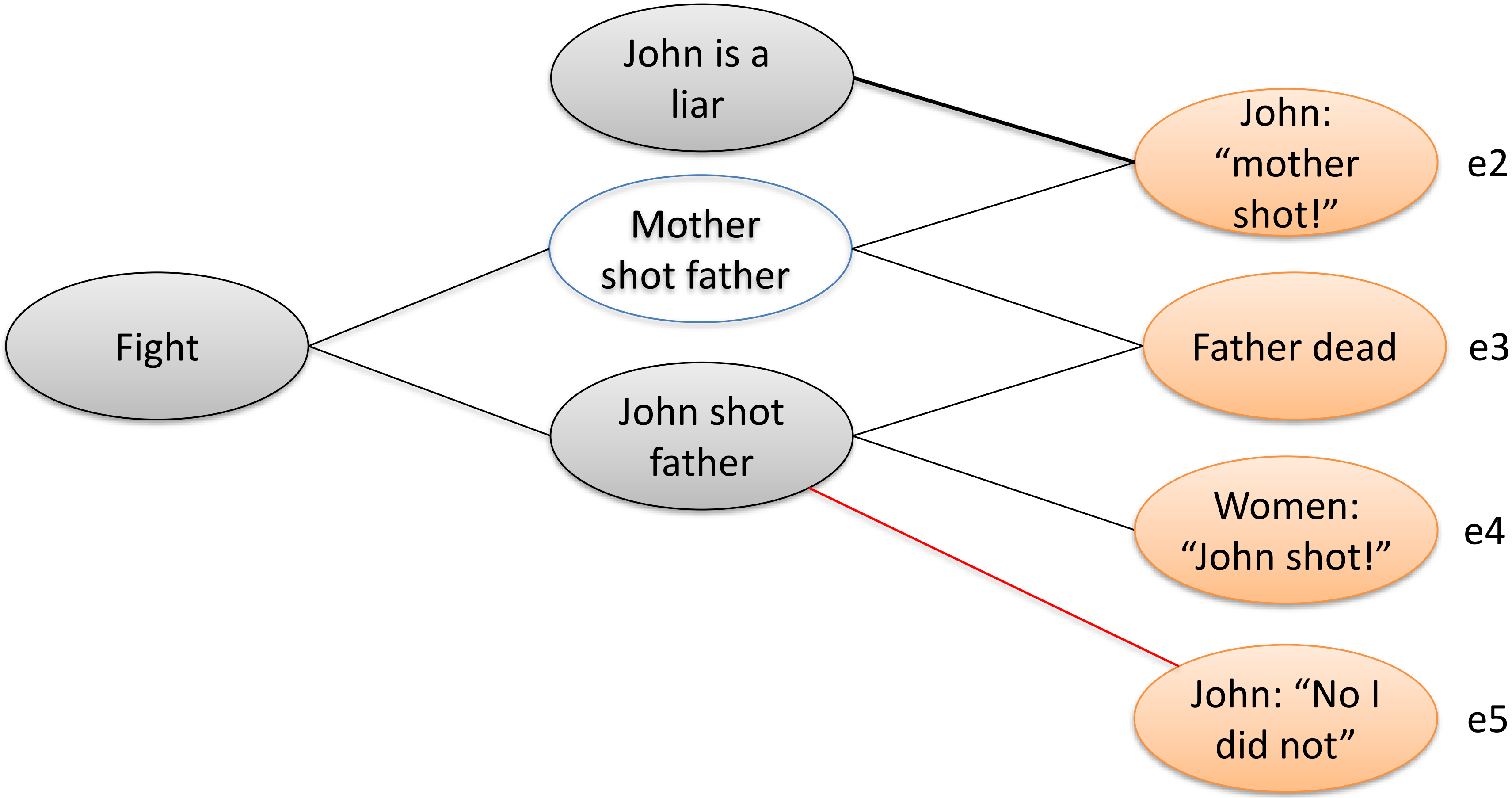
Support and Attack
Reasoning

- Find extra supporting evidence for J, increasing Support



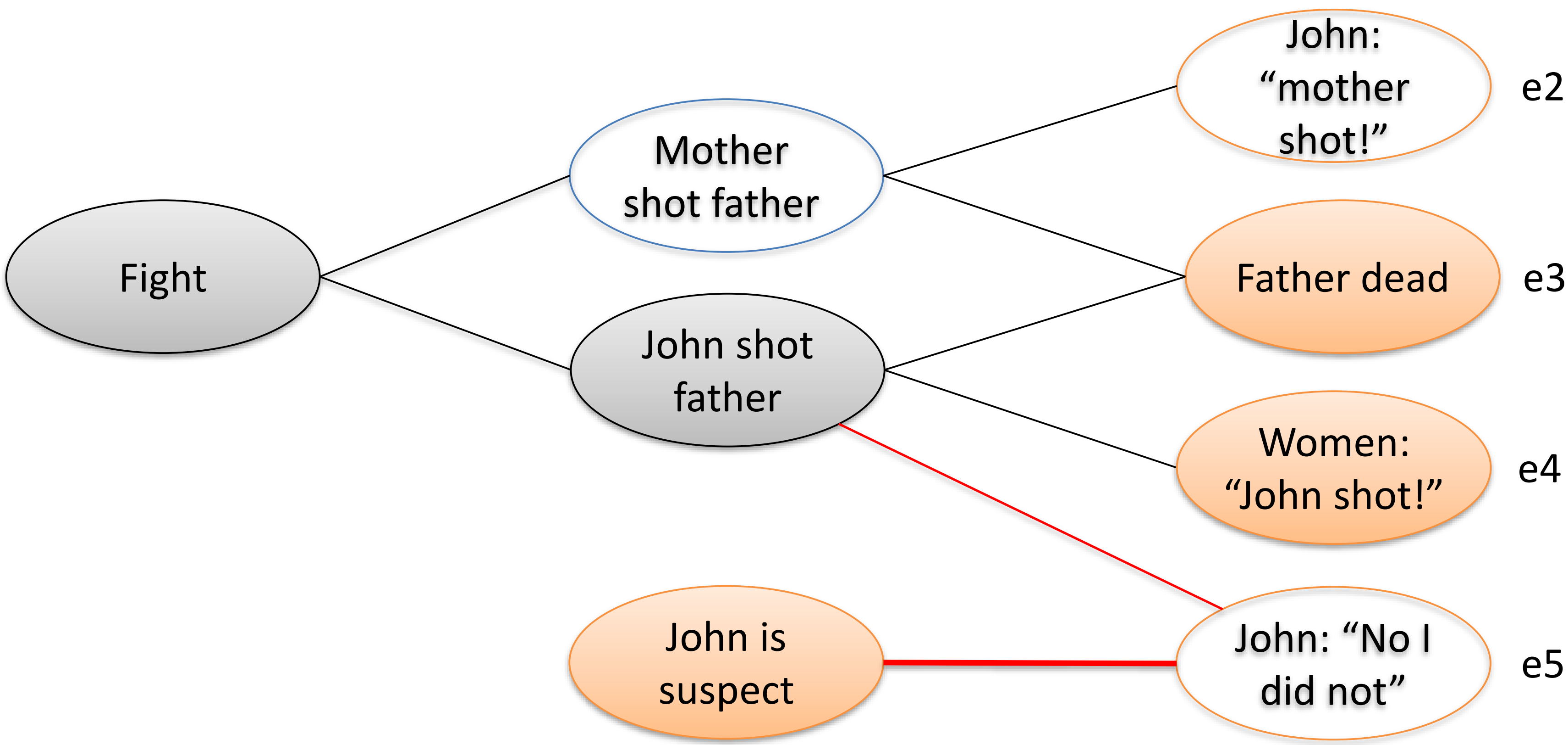
Support and Attack
Reasoning

- Explain the evidence that the other story explains by expanding your story (increasing support)



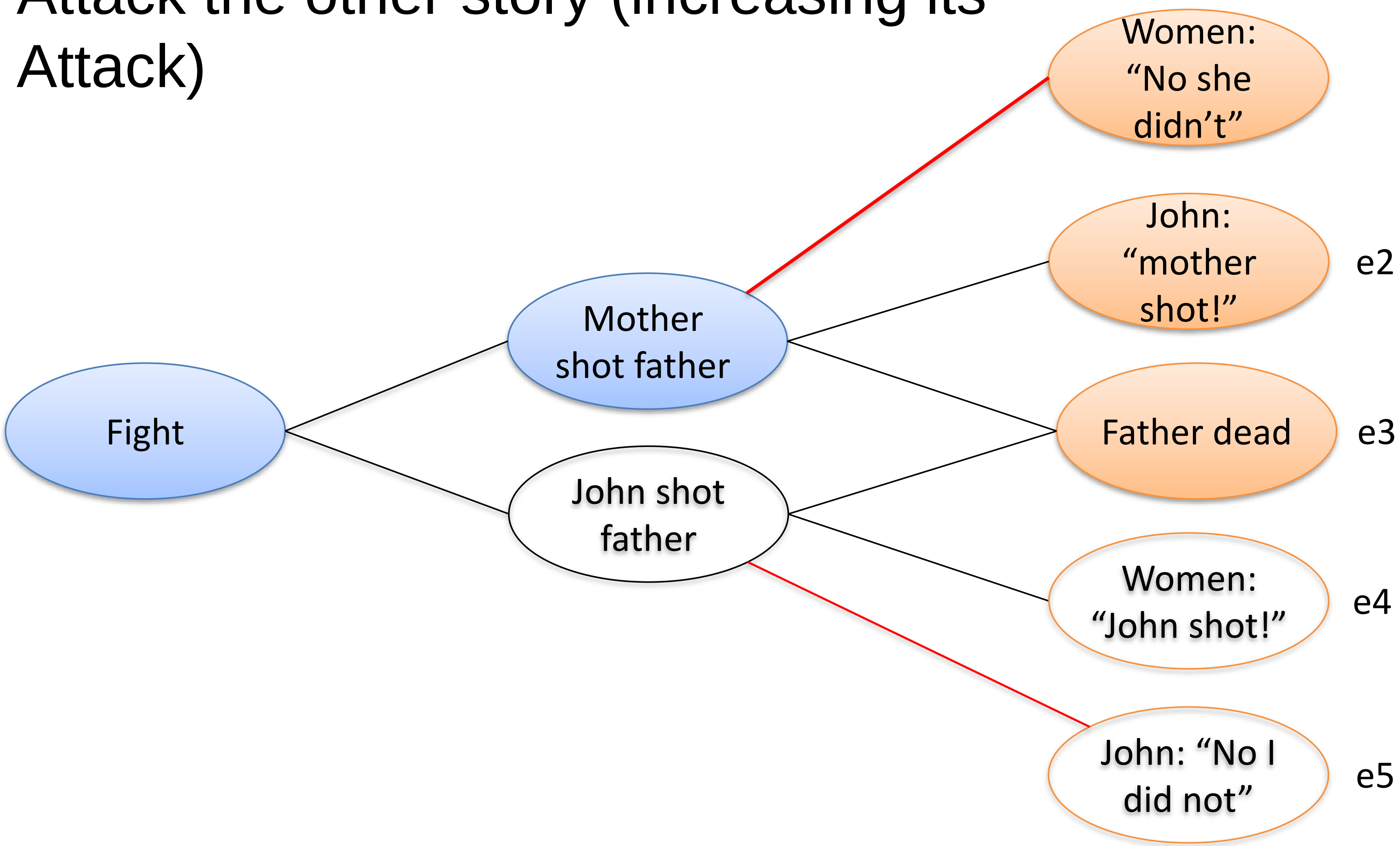
Support and Attack
Reasoning

- Attack the attacking evidence (decreasing Attack)



Support and Attack
Reasoning

- Attack the other story (increasing its Attack)



Evidence
Reasoning

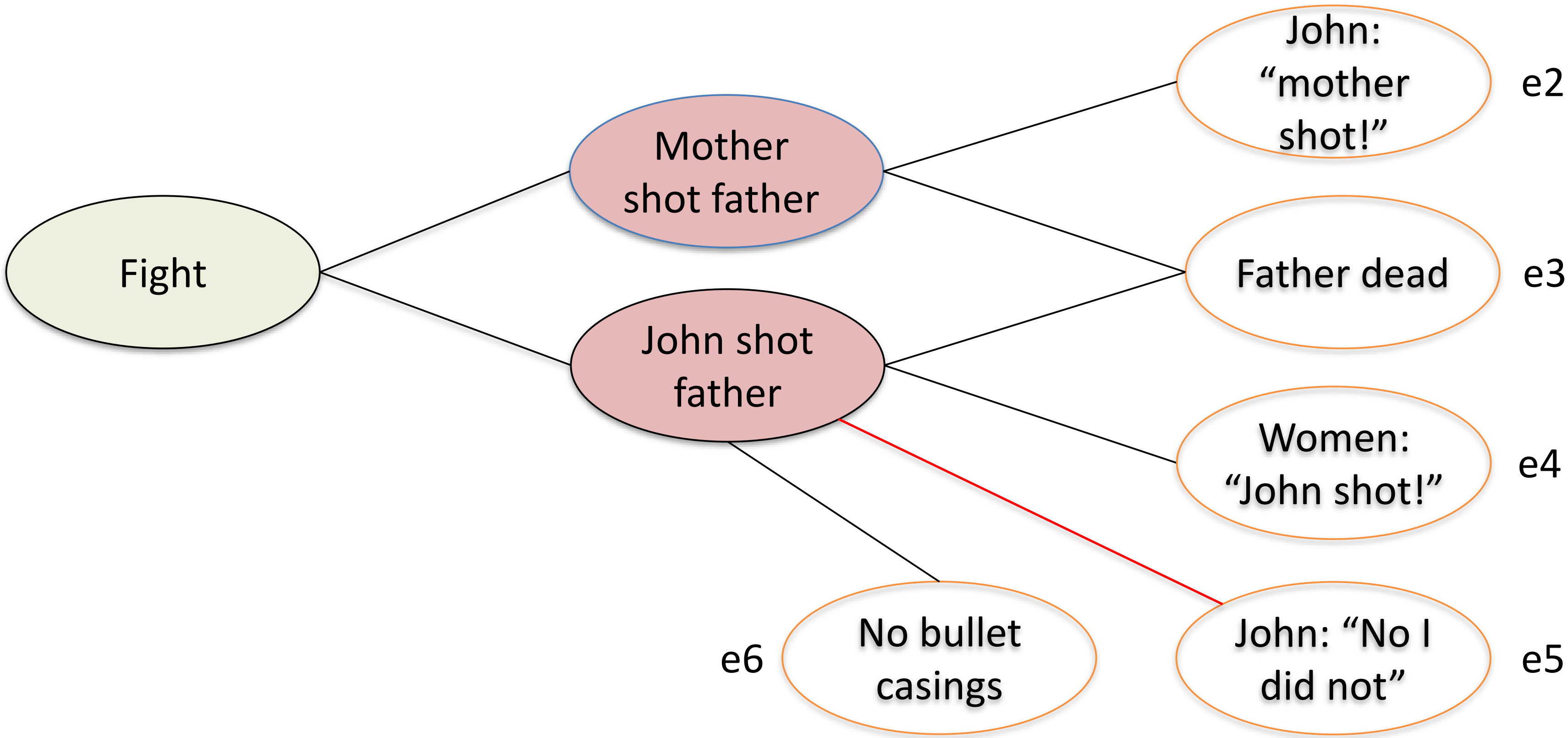
- Qualitative reasoning about the strength of stories given the evidence
- Improve your story by
  - Finding new supporting evidence
  - Expanding your story to explain more existing evidence
  - Attacking the other story
  - Attacking your attackers
- No final “decision”, but also no numbers needed

Strength of explanations
Evidence

- Measuring support and attack
- Supporting evidence
  – Pr(Story | Evidence) > Pr(Story)
- Attacking evidence
  – Pr(Story | Evidence) < Pr(Story)
- “Evidence” can be 1 piece, but also a set

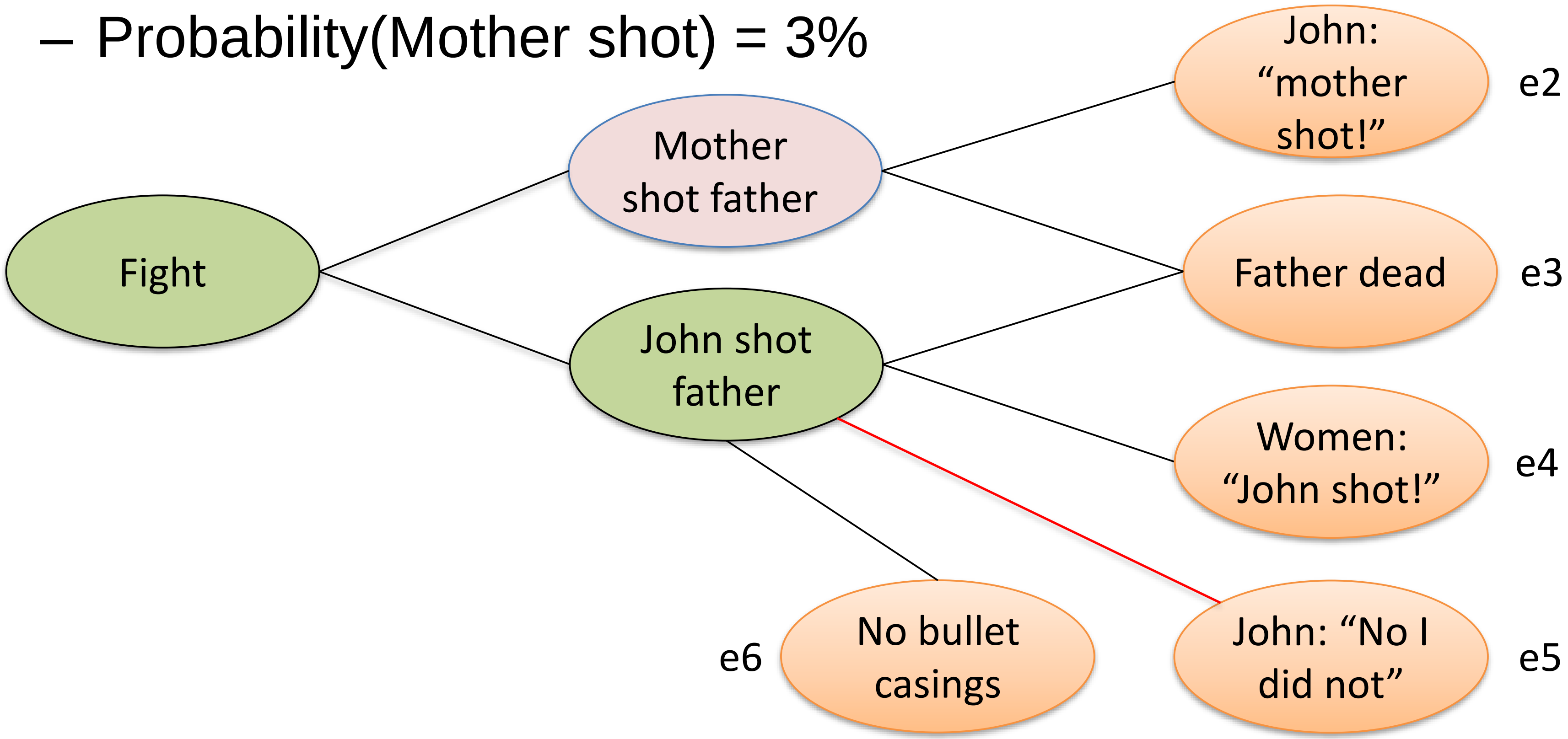
Support and Attack
Measuring

- Prior probabilities



Support and Attack
Measuring

- Compare the posterior probabilities with the priors for all stories
  - Probability(John Shot) = 95%
  - Probability(Mother shot) = 3%



Comparing Support and Attack
Measuring

- Total evidential support/attack
  - SuppAtt(M) = Pr(M | e2,…,e6) / Pr(M)
  - SuppAtt(J) = Pr(J | e2,…,e6) / Pr(J)
- SuppAtt(M) < SuppAtt (J)
  - J is more strongly supported (or less strongly attacked) by the evidence than M

- Measuring a story's conformance to the evidence
  - Aggregation, strong vs weak evidence, total influence of evidence on story
- However: all numbers have to be filled in

Strength of explanations
Completeness & plausibility

- A story is *coherent* if it conforms to our world knowledge

- World knowledge can be encoded as rules/generalizations
  - If you shoot someone they might die
- World knowledge can be encoded as story schemes
  - person *x* has a motive *m* to kill person *y*
  - person *x* kills person *y* (at time *t*) (at place *p*) (with weapon *w*)
  - person *y* is dead

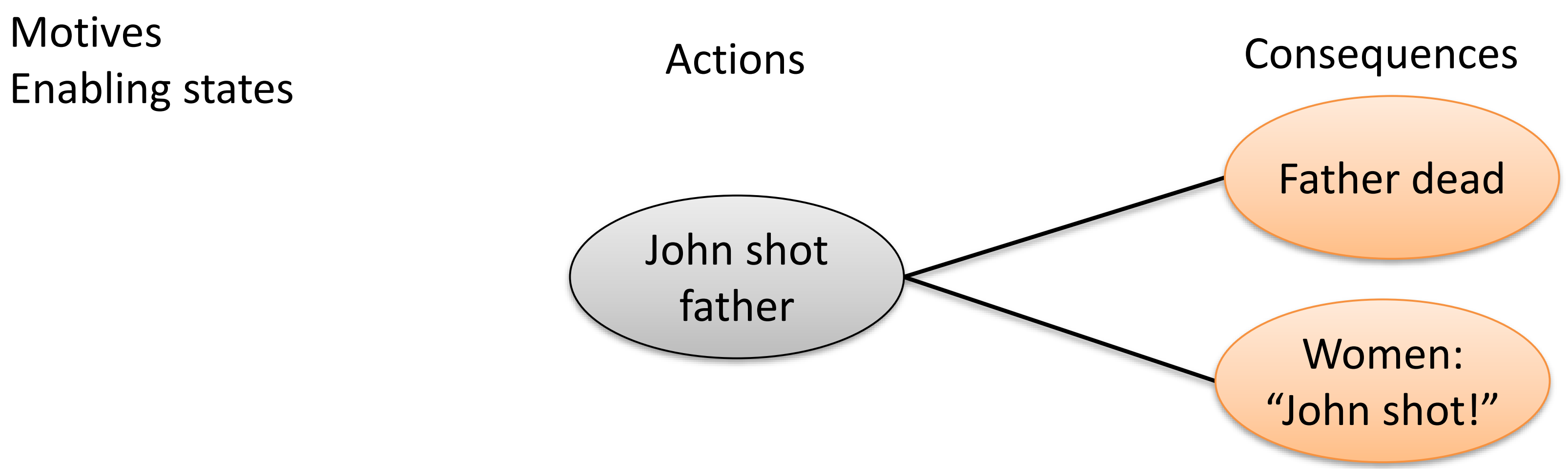
Strength of explanations
Completeness

- Completeness
  - Does the story mention all the relevant events we expect to see?

- A complete story mentions all parts of a script, and nothing more

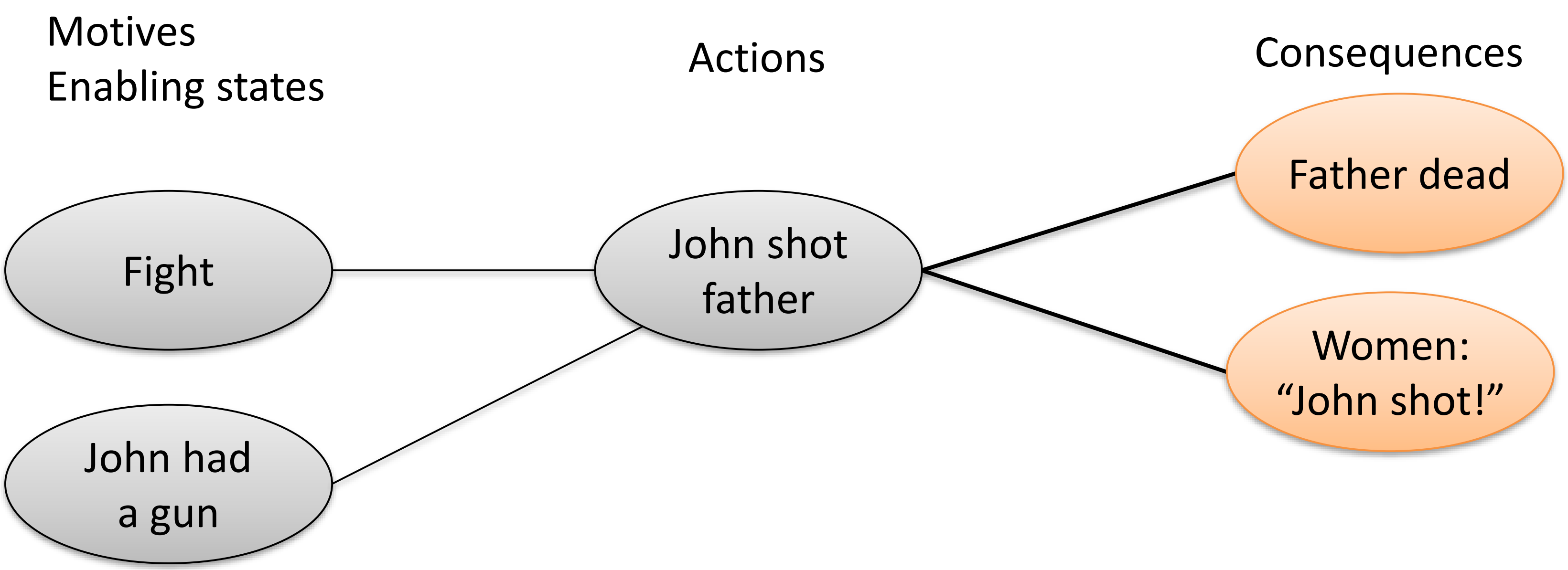
Completeness

- Missing elements



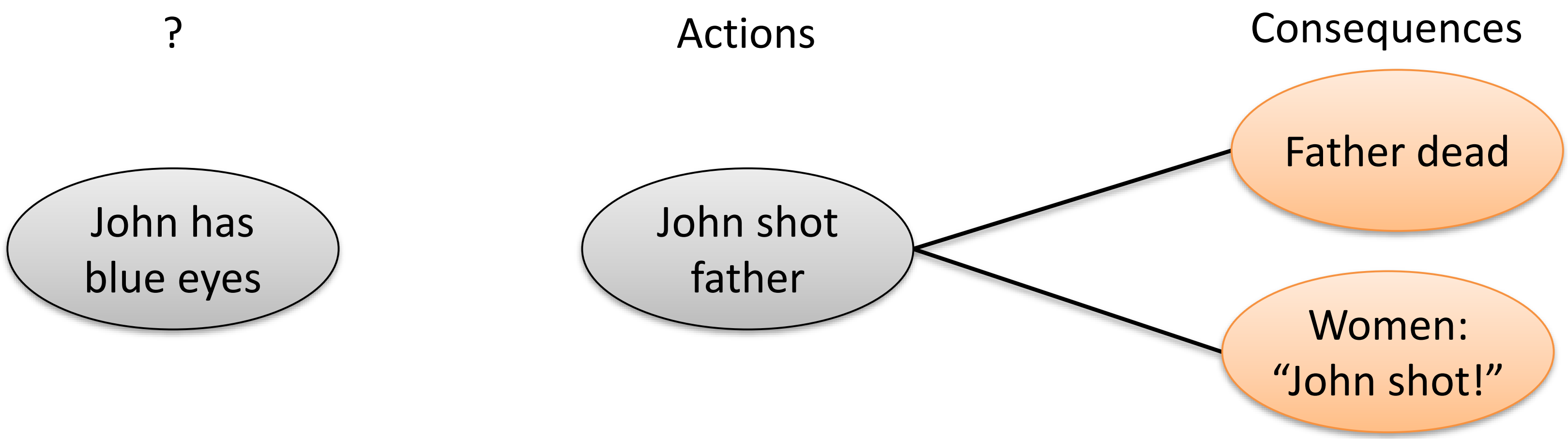
Completeness

- Add missing elements



Completeness

- Superfluous elements can be deleted

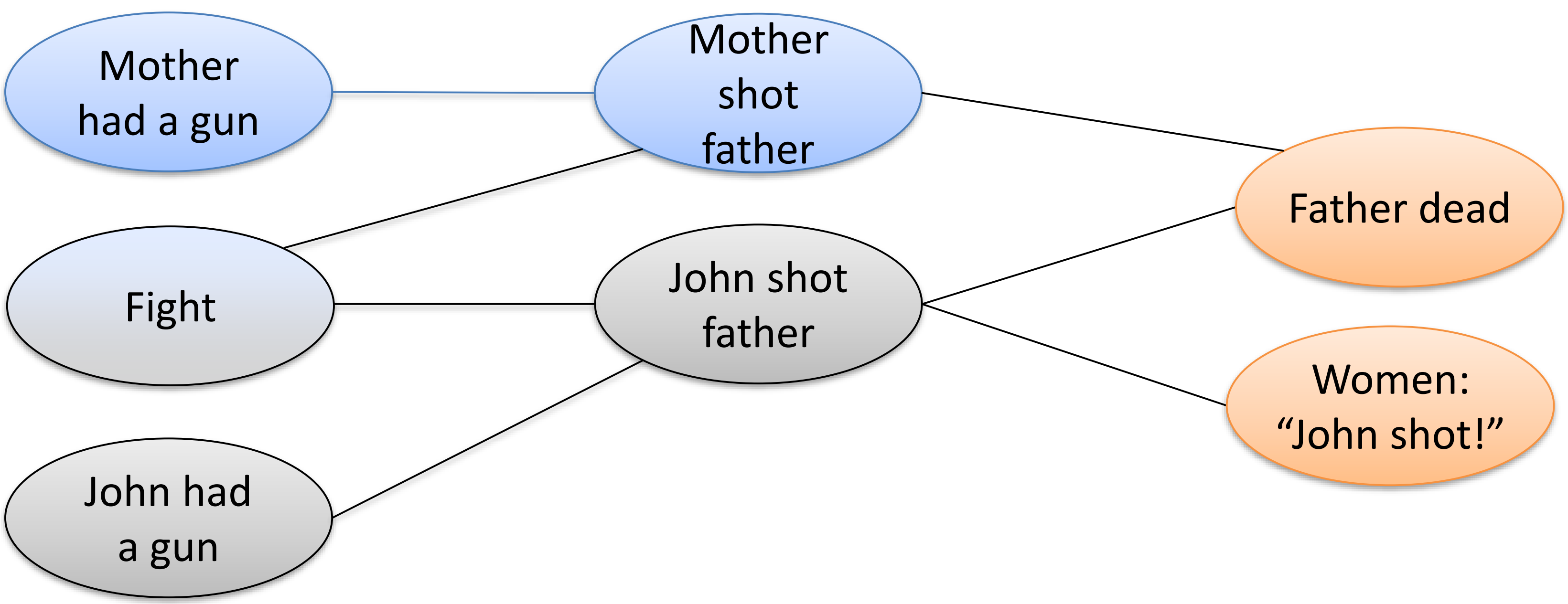


Plausibility

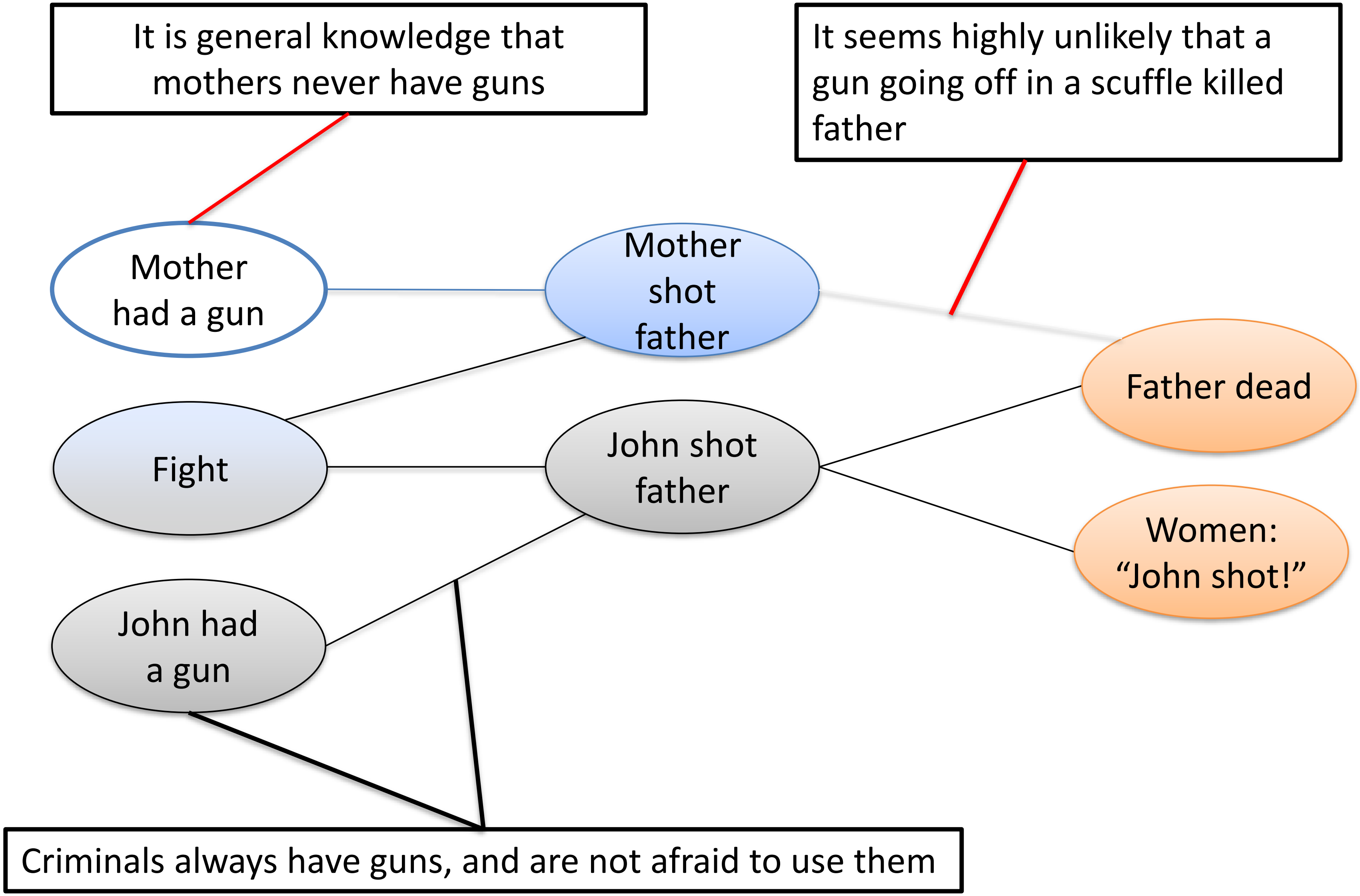
- The inherent plausibility of events and links between events
  - Mothers never carry guns
  - Criminals like John always carry guns
  - Shooting someone causes them to die
  - It is implausible that a gun going off in a scuffle would have killed father
  - Suspects always deny the charges against them

Plausibility
Reasoning

- Argue for the plausibility of your own story, and against that of the others

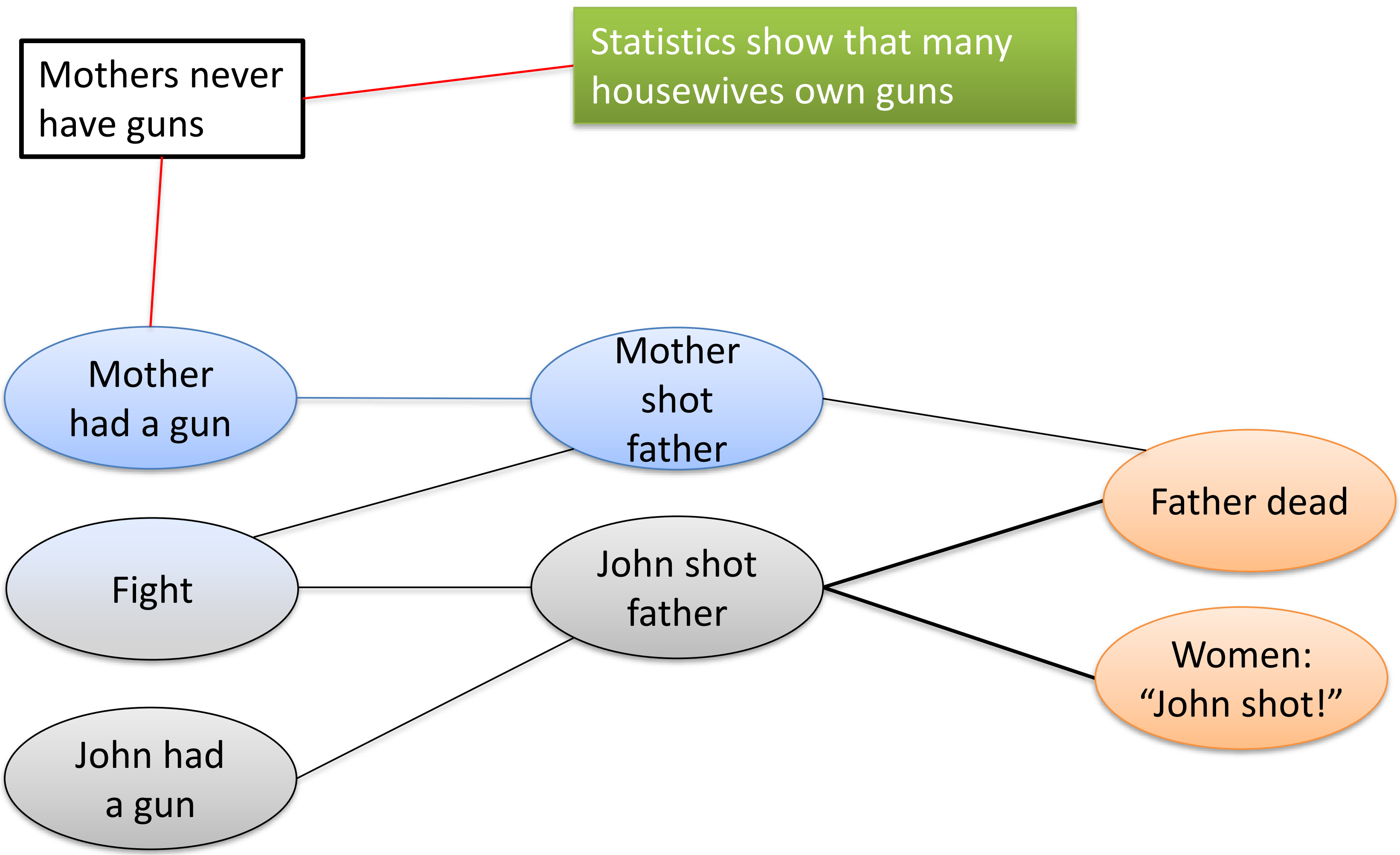


Plausibility
Reasoning



Plausibility
Reasoning

- Attacking attackers



Evidence
Reasoning

- Qualitative reasoning about the completeness & plausibility of stories

- Improve your story by
  - Completing it and deleting superfluous elements
  - Arguing that the other story is incomplete
  - Arguing for the plausibility of your story
  - Arguing against the plausibility of the other story
  - Attacking your attackers

- No final “decision”, but also no numbers needed

Plausibility
Measuring

- Plausibility can be expressed as probabilities

- Criminals always carry guns
  - Pr(Criminal_John_has_gun) = 1
- Criminals are not afraid to use guns
  - Pr(J_shot | fight, J_has_gun) > 0.5
- Mothers (almost) never carry guns
  - Pr(Mother_has_gun) = 0.001
- It is implausible that a gun going off in a scuffle would have killed father
  - Pr(f_dead | m_shot) < 0.1

- Measuring plausibility is necessary to come to a decision
- Measuring plausibility is dangerous if probabilities left implicit
  - Argue about them!

Conclusions

- Reasoning based on competing *stories*
- Stories can be *argued* about
  - How well they conform to the evidence
  - How complete & plausible they are
  - How much better than other stories they are
- *Probabilities* can be used to measure how good stories are
  - How inherently plausible the events and links are
  - How (much more or less) likely they are given the evidence

Conclusions

- Always need world knowledge
  - Possible counterarguments
  - Probabilities
- Qualitative reasoning
  - No complete probability distribution needed
- Quantitative reasoning
  - Close to statistical (machine learning) in AI