King's College London

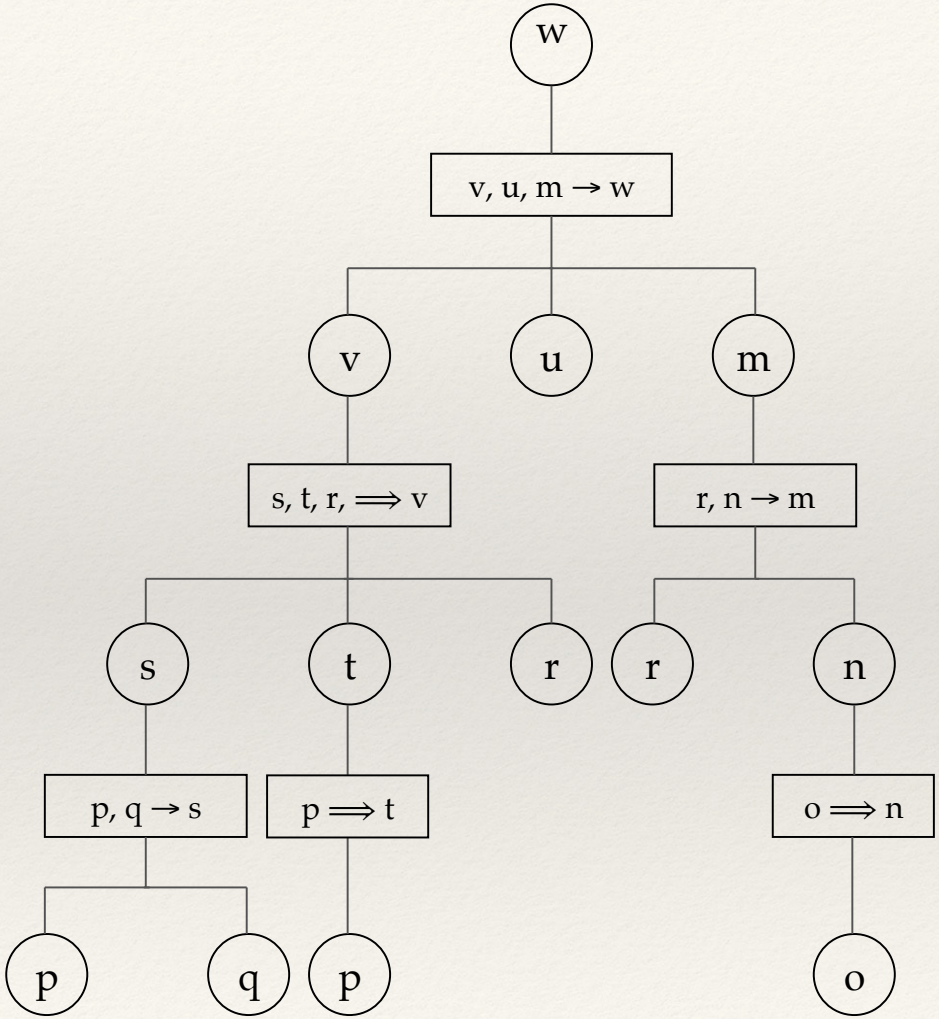
Assigning Likelihoods to Interlocutors' Beliefs and Arguments

Seyed Ali Hosseini
Sanjay Modgil
Odinaldo Rodrigues

Structured argumentation based on ASPIC+: Preliminaries

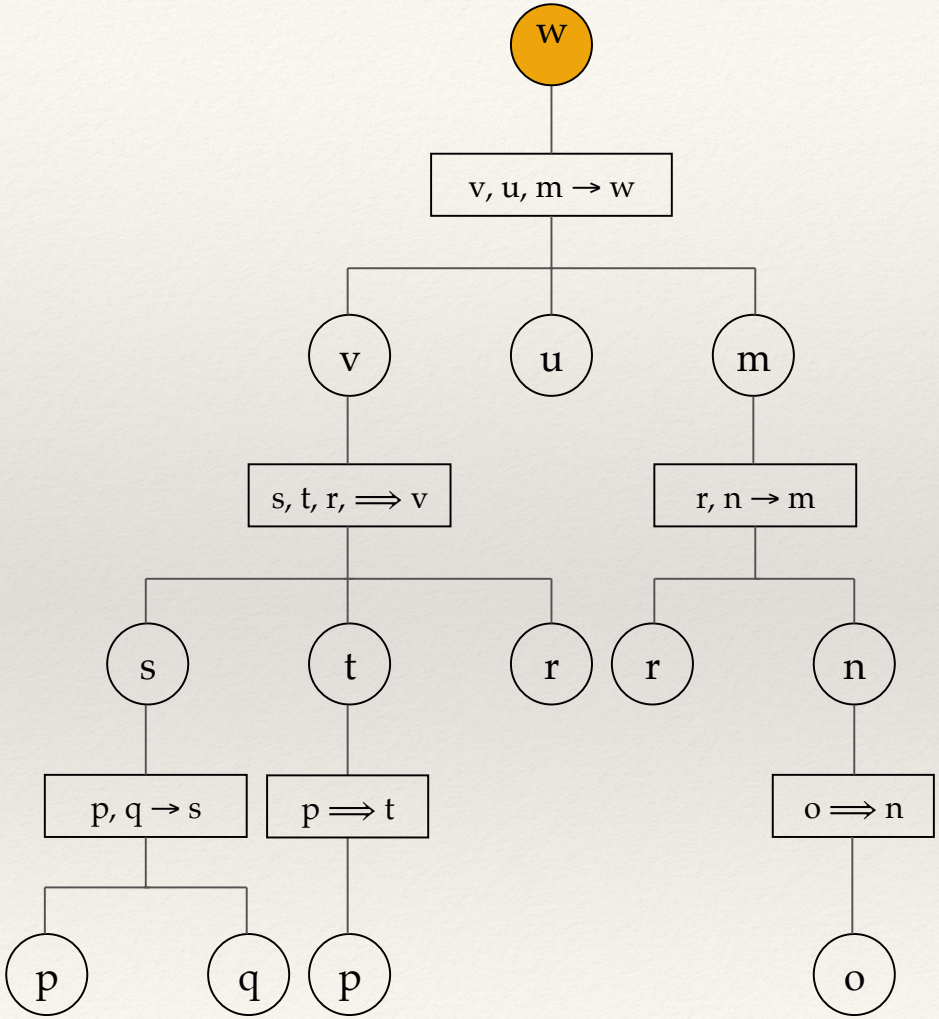
- Argumentation Theory $AT = \langle AS, K \rangle$
- Argumentation System $AS = \langle L, R, —, n \rangle$
  - L Language
  - $R = R^s \cup R^d$ Strict/Defeasible inference rules
  - $—$ Generalised negation
  - $n : R^d \rightarrow L$ Name for defeasible rules
- Knowledge-base $K = K^n \cup K^p$ Axiom and ordinary premises

Structured argumentation based on ASPIC+: Arguments



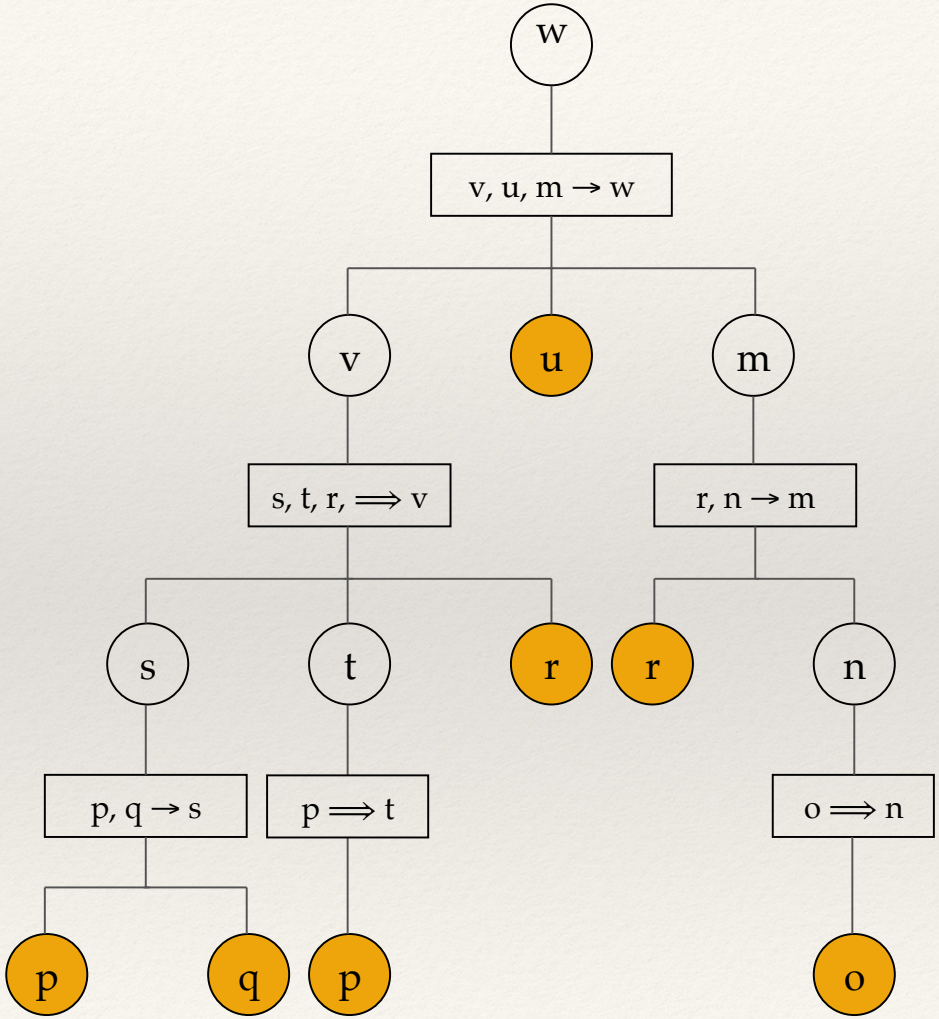
Structured argumentation based on ASPIC+: Arguments

Conclusion



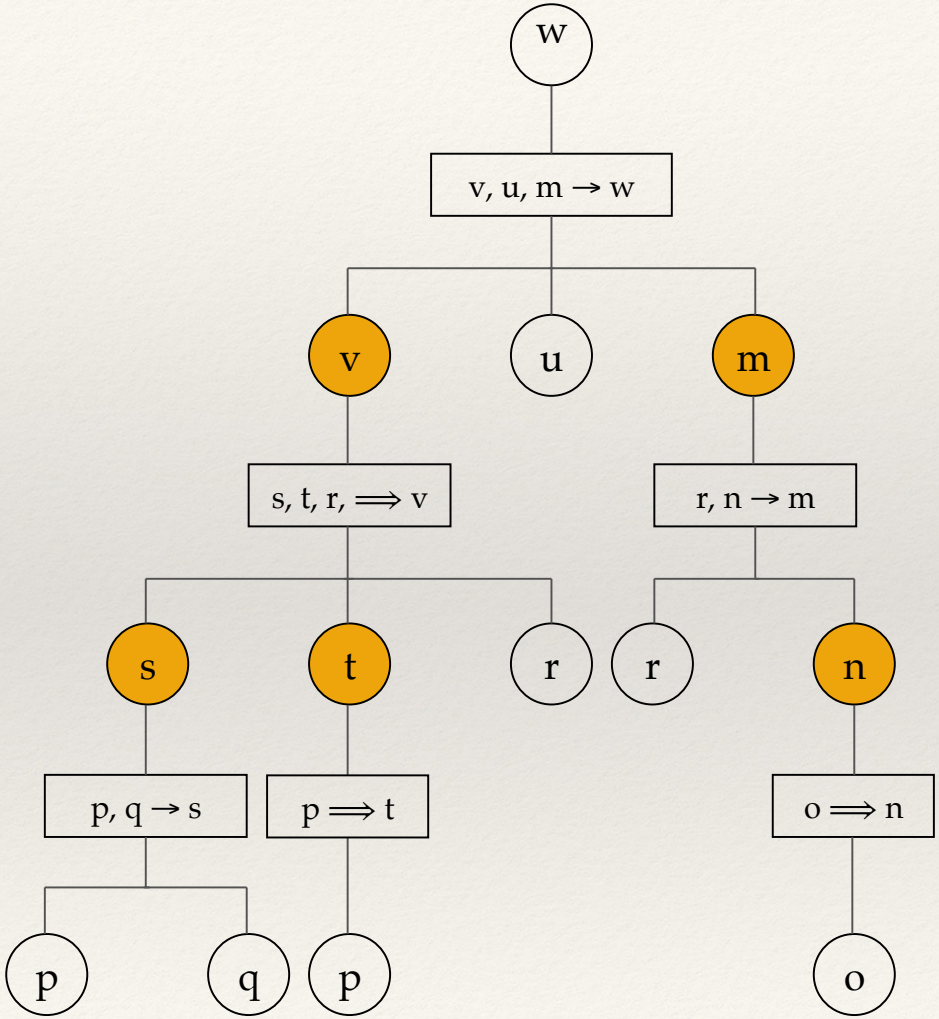
Structured argumentation based on ASPIC+: Arguments

Premises



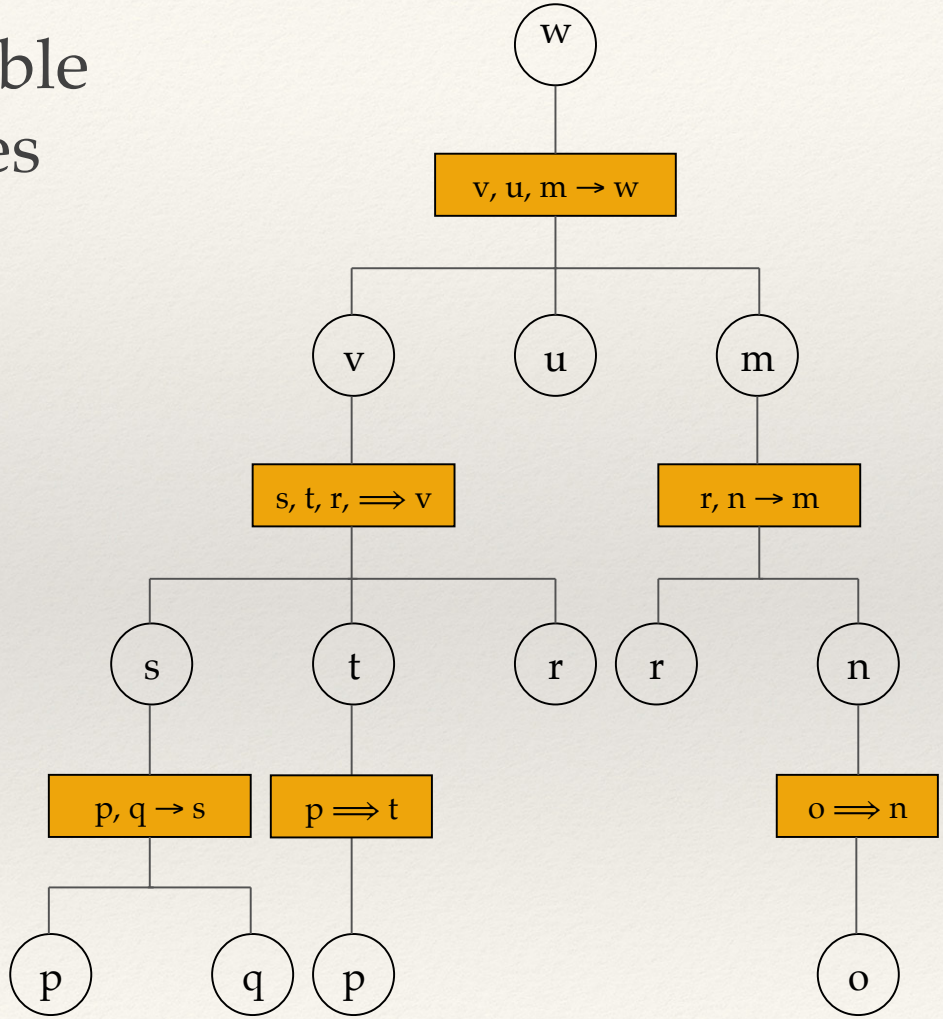
Structured argumentation based on ASPIC+: Arguments

Intermediate Conclusions

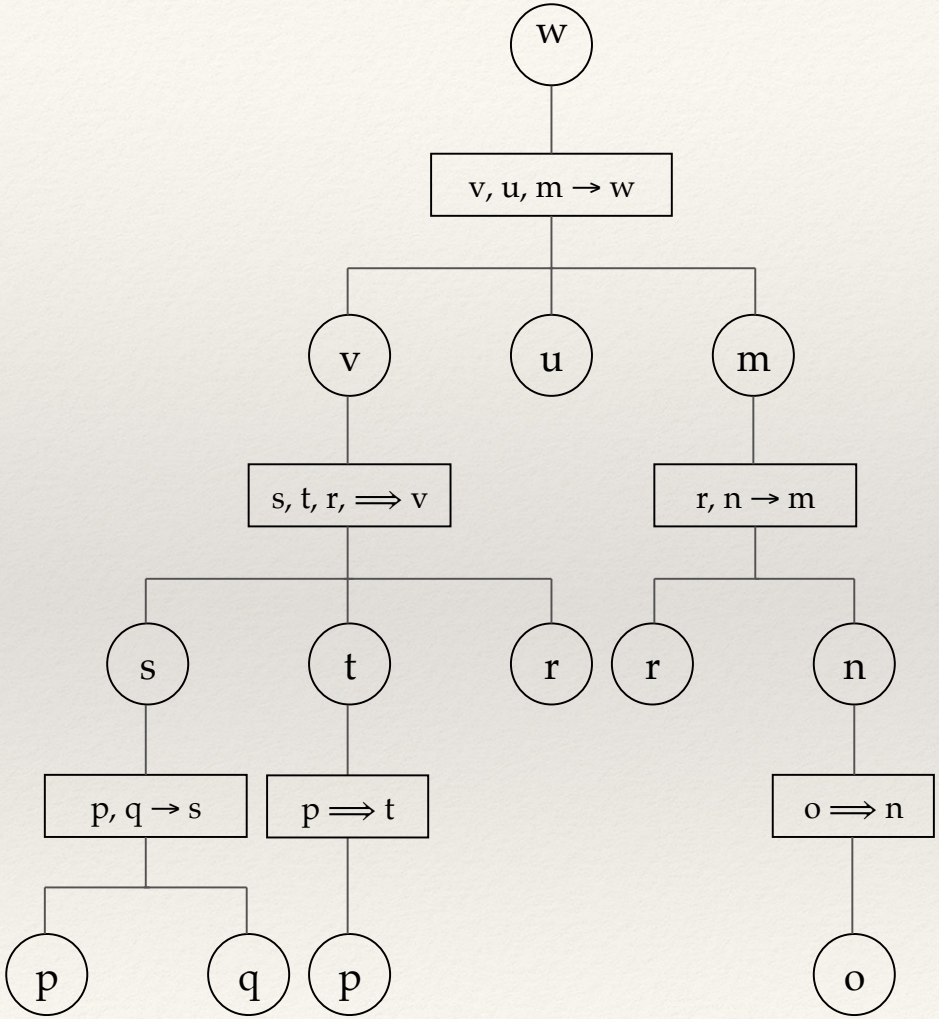


Structured argumentation based on ASPIC+: Arguments

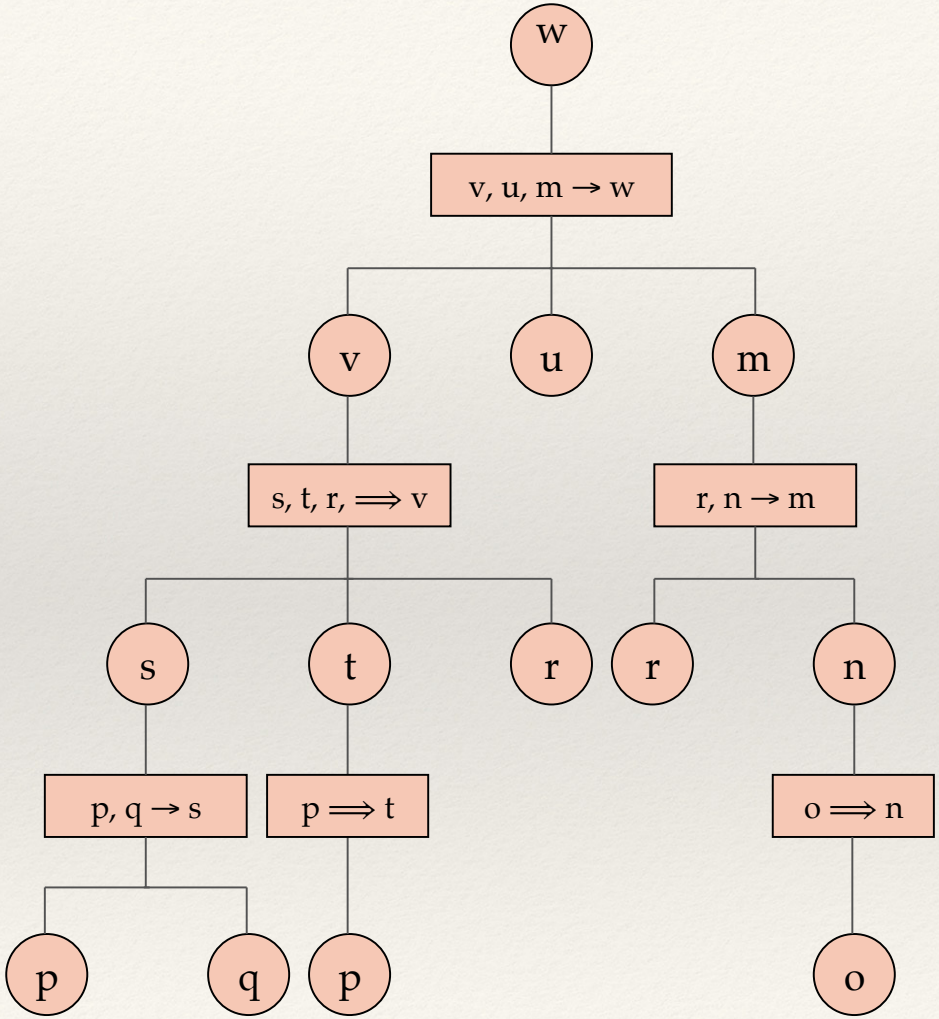
Strict/Defeasible Inference Rules



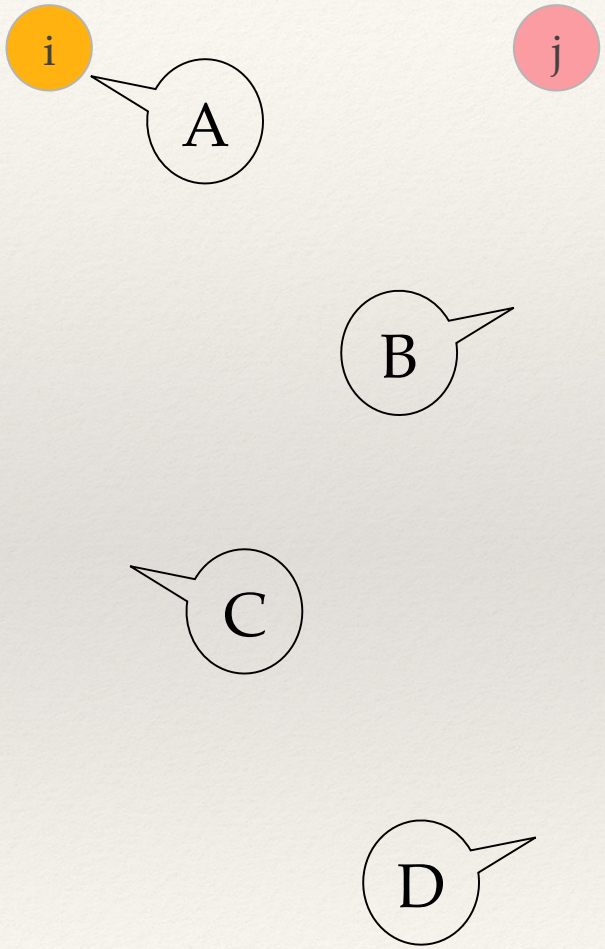
Belief



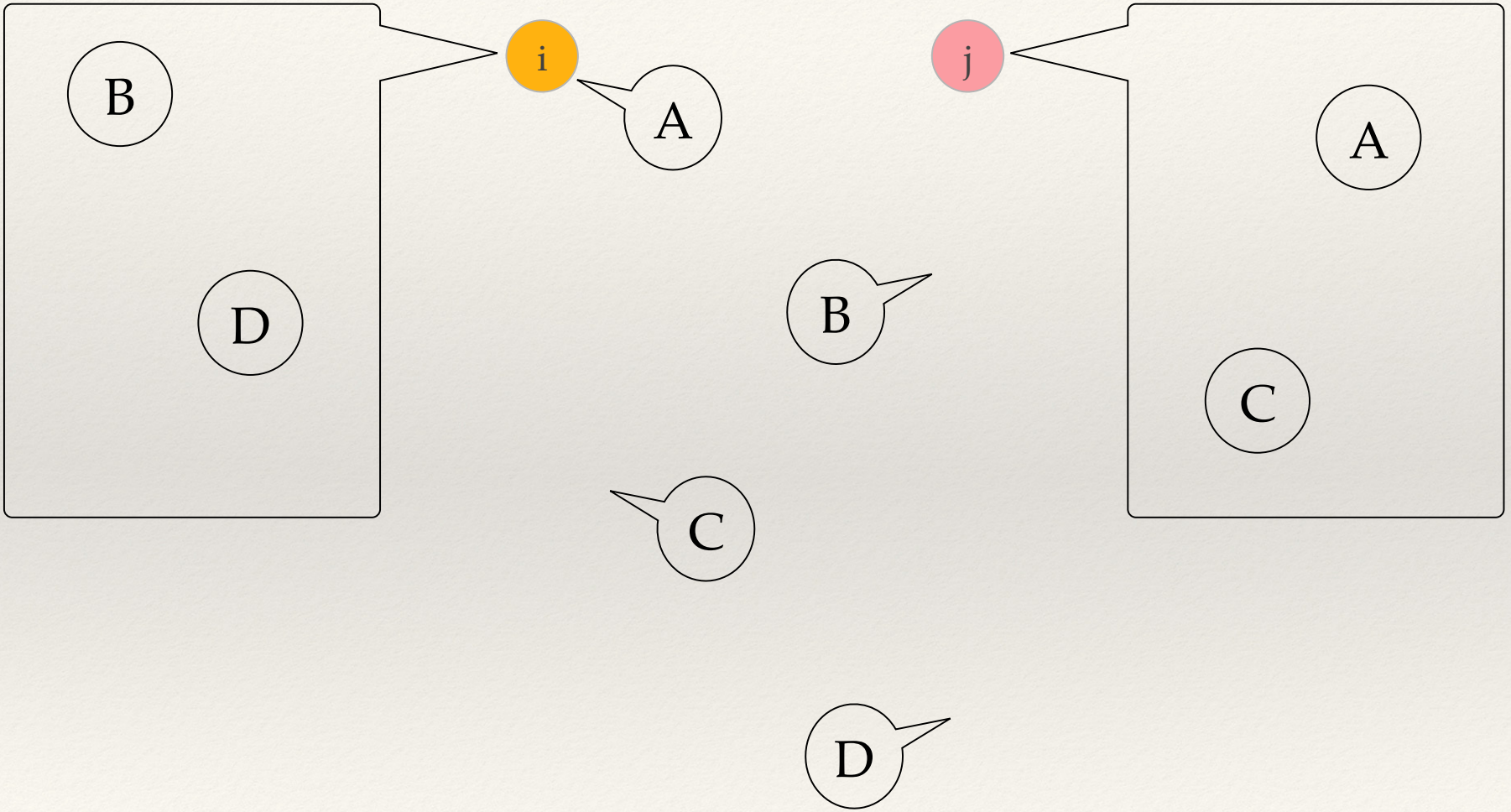
Belief



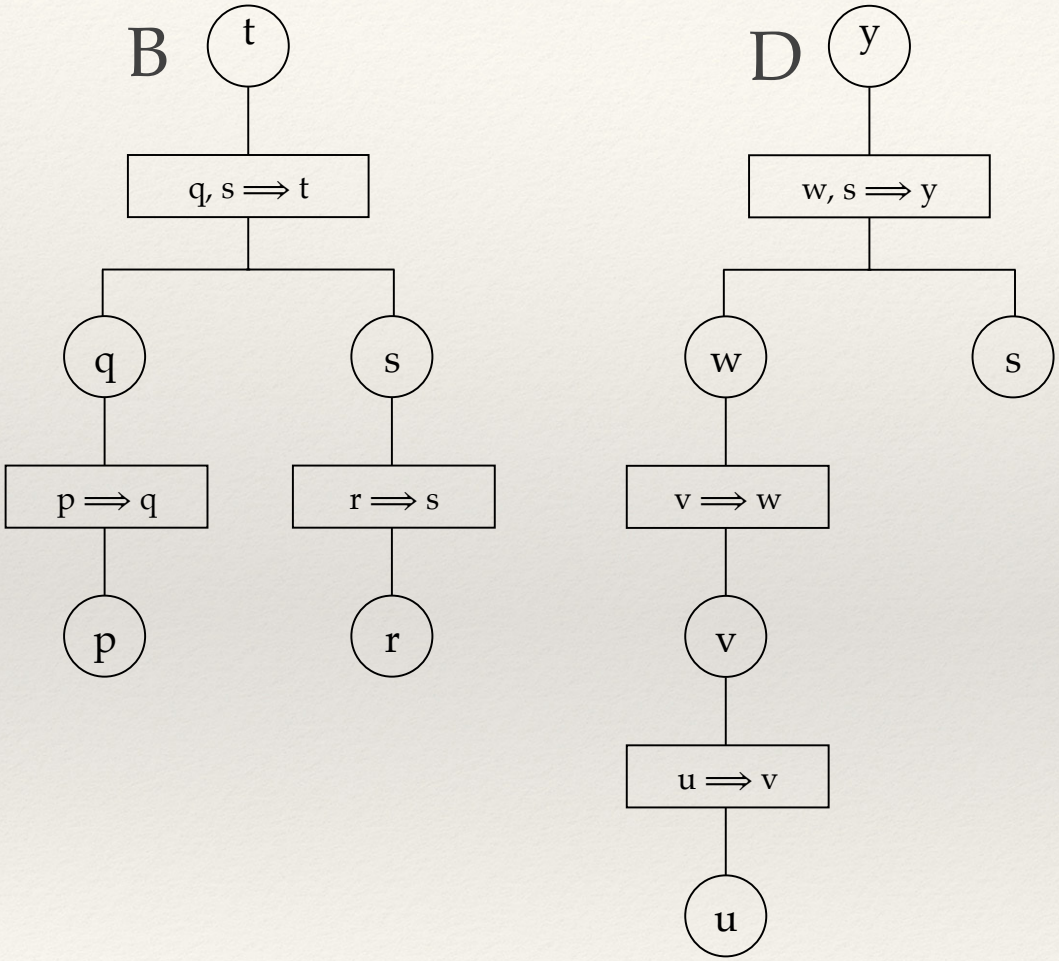
Opponent modelling using abstract arguments



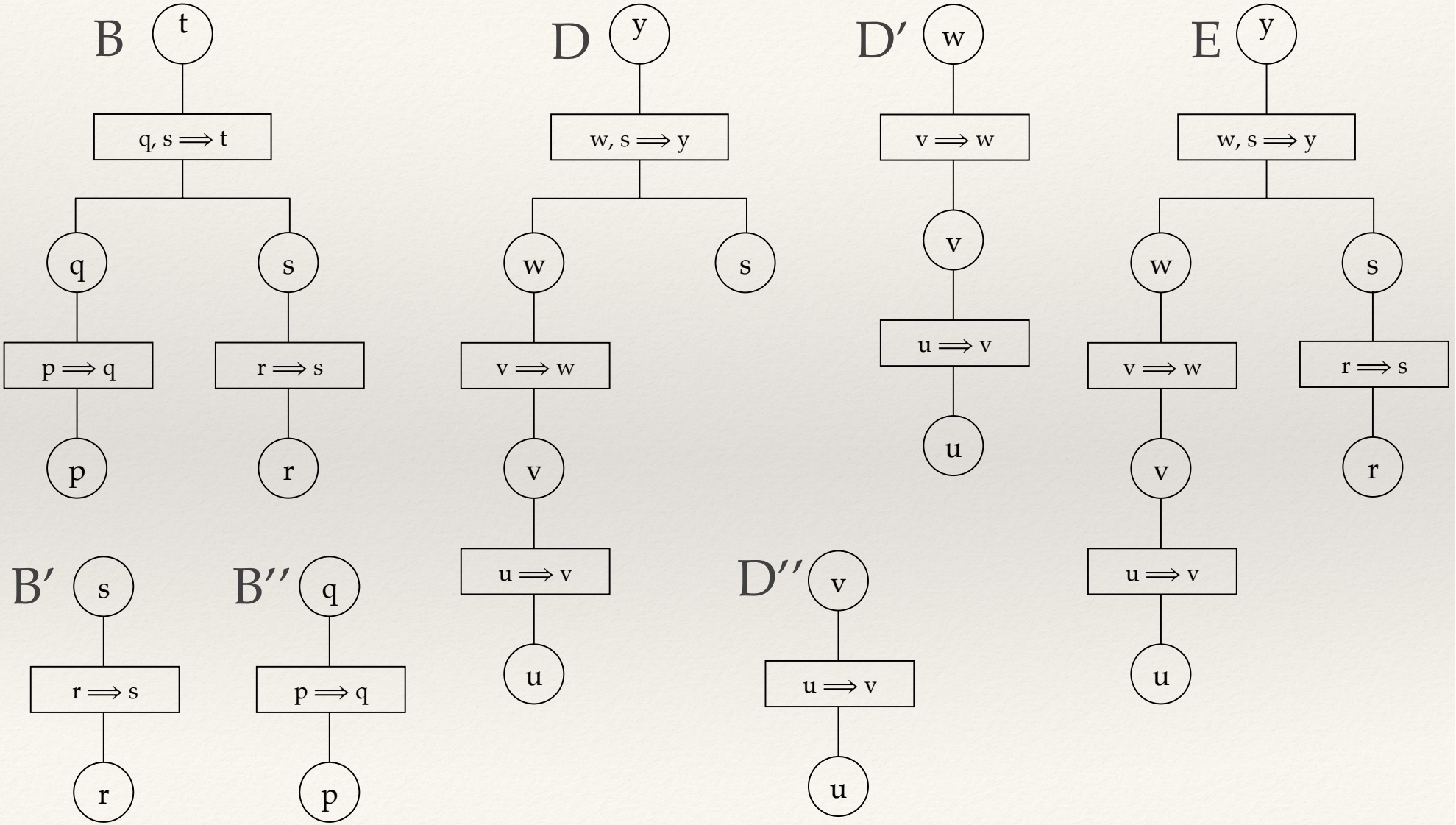
Opponent modelling using abstract arguments



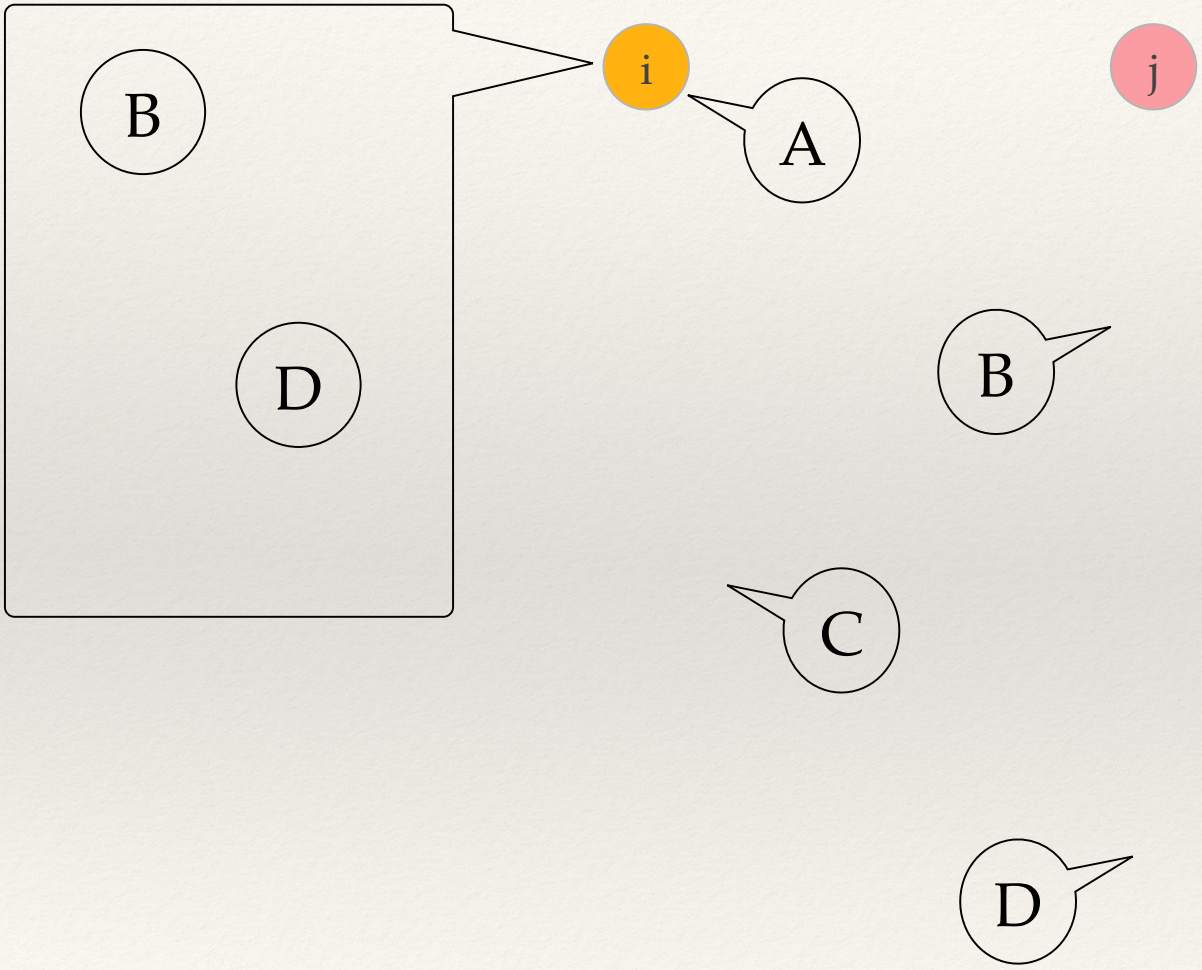
Opponent modelling: from abstract to structured arguments



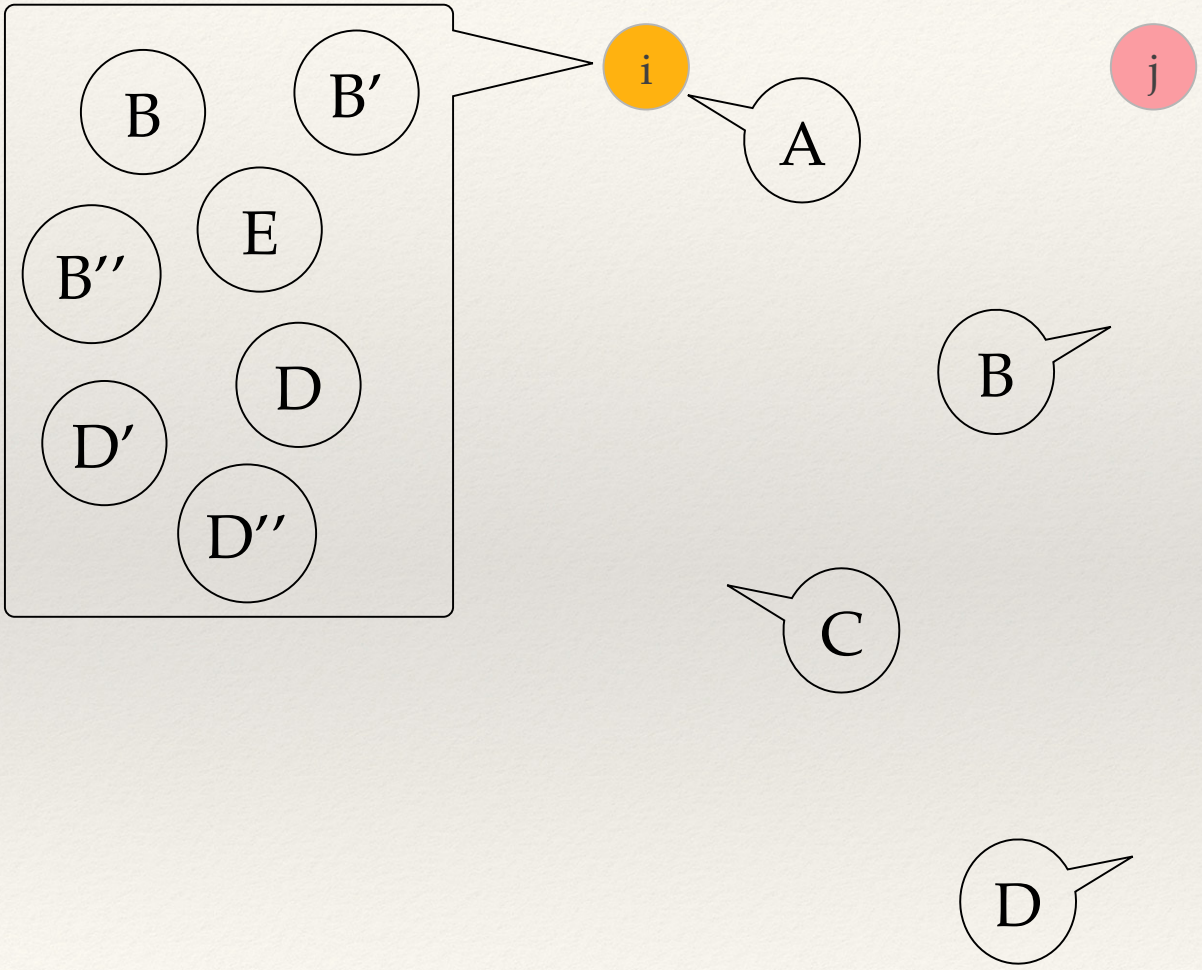
Opponent modelling: from abstract to structured arguments



Opponent modelling: from abstract to structured arguments



Opponent modelling: from abstract to structured arguments



Estimating the likelihood of a second-order belief

DEF An agent *i*'s second-order belief with regards to agent *j* is a belief that agent *i* assumes that agent *j* has.

We use two complementary methods, the first of which uses

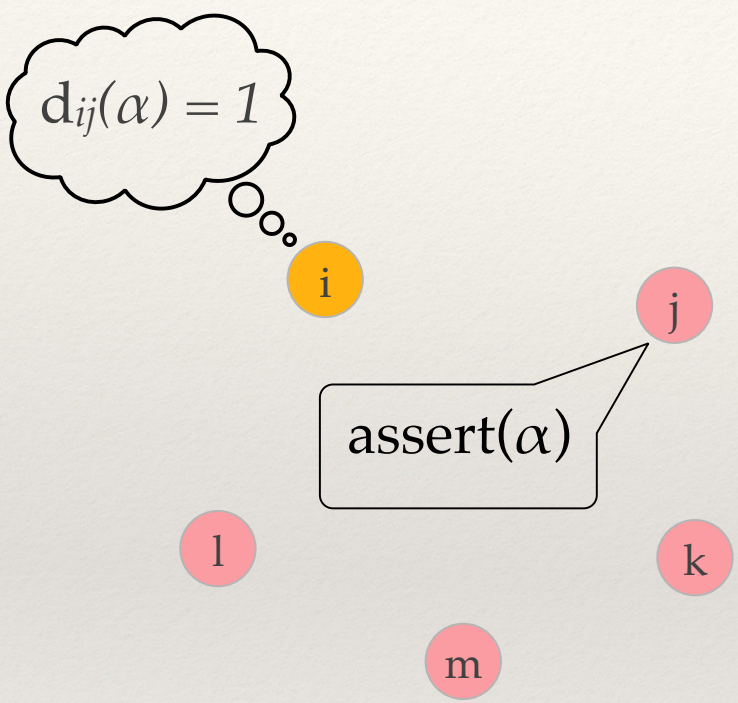
- Evidence obtained through dialogue
  - *Example:* agent *j* asserts α in a dialogue with *i*
  - *Example:* an agent *k* informs *i* that *j* believes in α
  - *Example:* an agent *i* finds out that *j* does not believe in α through a failed inquiry dialogue

Dialogical Evidence (DE)

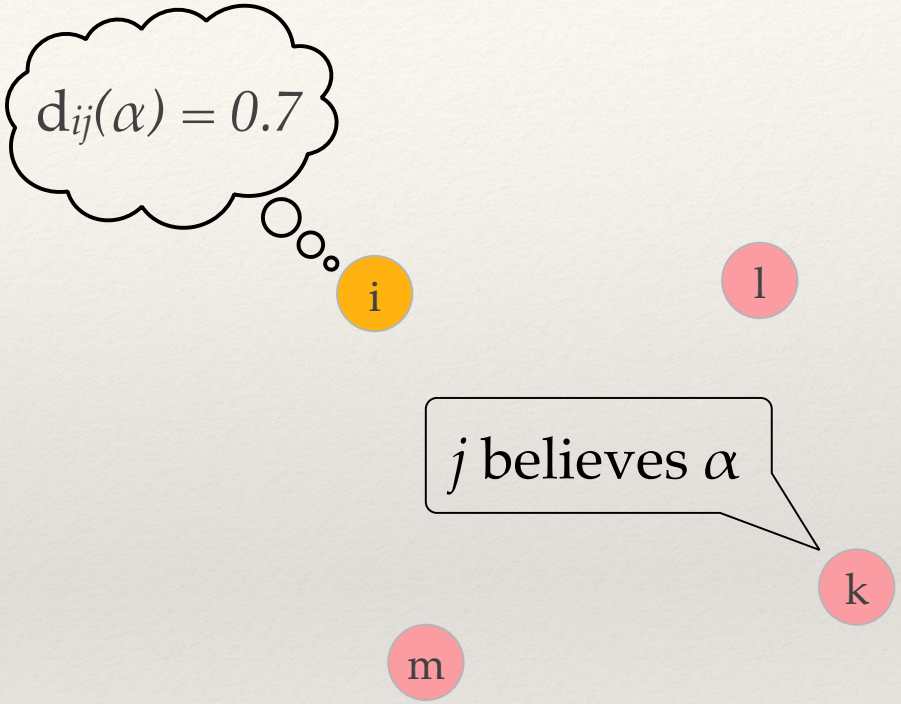
We assume the existence of some mechanism that allows an agent i to assign the likelihood that j believes α based on dialogical interactions. This is denoted by $d_{ij}(\alpha)$:

- $d_{ij}(\alpha)=0$ means i has no information about j's believing (or not) in α
- $0 < d_{ij}(\alpha) \leq 1$ means that i believes that j believes in α with degree $d_{ij}(\alpha)$ of confidence
- $d_{ij}(\alpha) = \perp$ means that i believes that j does not believe in α

Dialogical Evidence (DE)

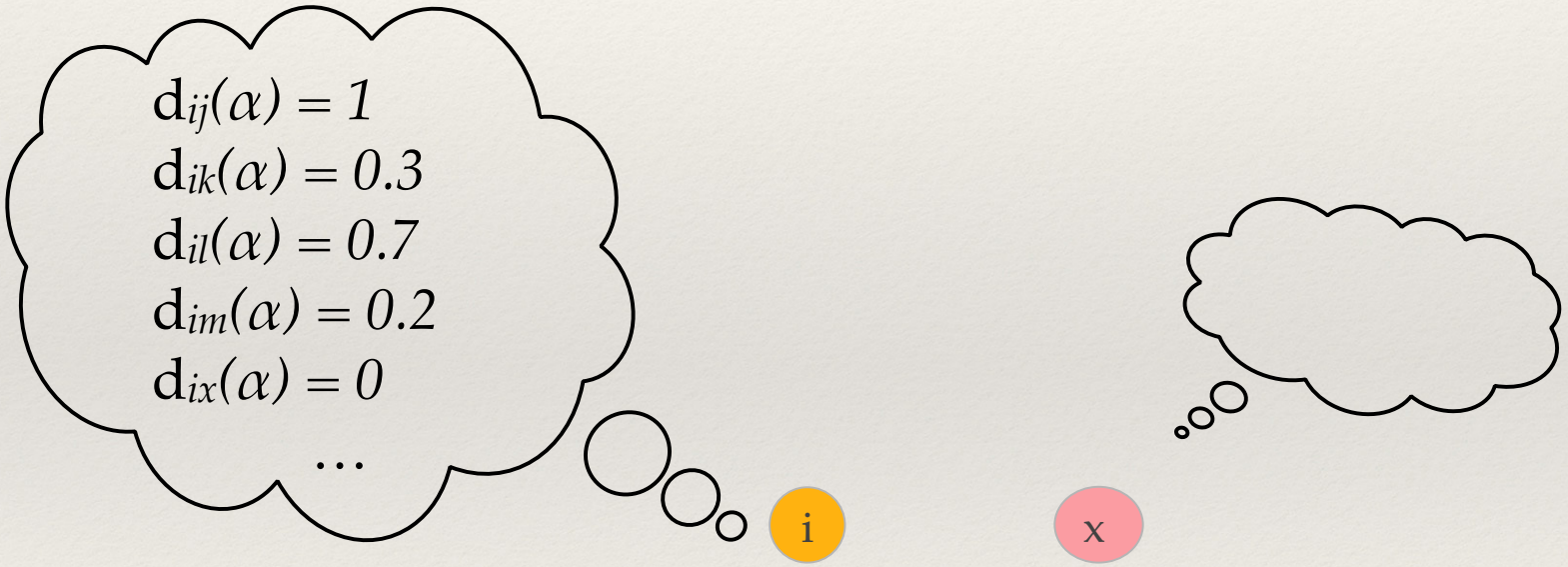


Direct Evidence



Indirect Evidence

Dialogical Evidence (DE) not always available



Estimating the likelihood of a second-order belief

Two complementary methods:

- Evidence obtained through dialogue
  - *Example:* agent j asserts α in a dialogue with i
  - *Example:* an agent k informs i that j believes in α
- Estimate the likelihood based on the agents' membership to communities

Agent Group

DEF A *group* of agents is a set of agents who share certain characteristics.

Example: agents with the same goal, history, institutional roles, etc.

Example: Lawyers, Logicians, Researchers in the field of argumentation, etc.

Agents may belong to more than one groups.

> ”Membership to a group licenses the assumption that the member shares some beliefs with the others in the group."

Agent Communities

DEF The power set of the set of groups forms the set of all communities: an *agent community* is an element of the set.

So if A and B are groups, the community {A,B} contains the agents that belong to both A and B, and so forth.

If an agent is in community {A,B}:

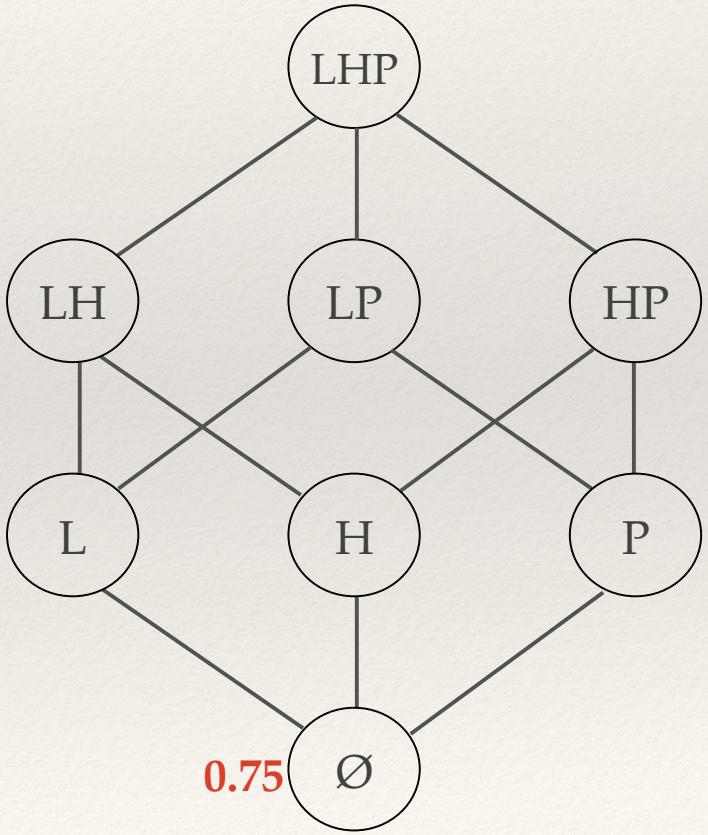
- *i* has the beliefs shared by the agents in A
- *i* has the beliefs shared by the agents in B
- plus possibly some other beliefs specific to members of both A and B only.

Ø is the most general community, to which every agent belongs.

P-scores

Since an agent knows the communities to which each agent belongs:

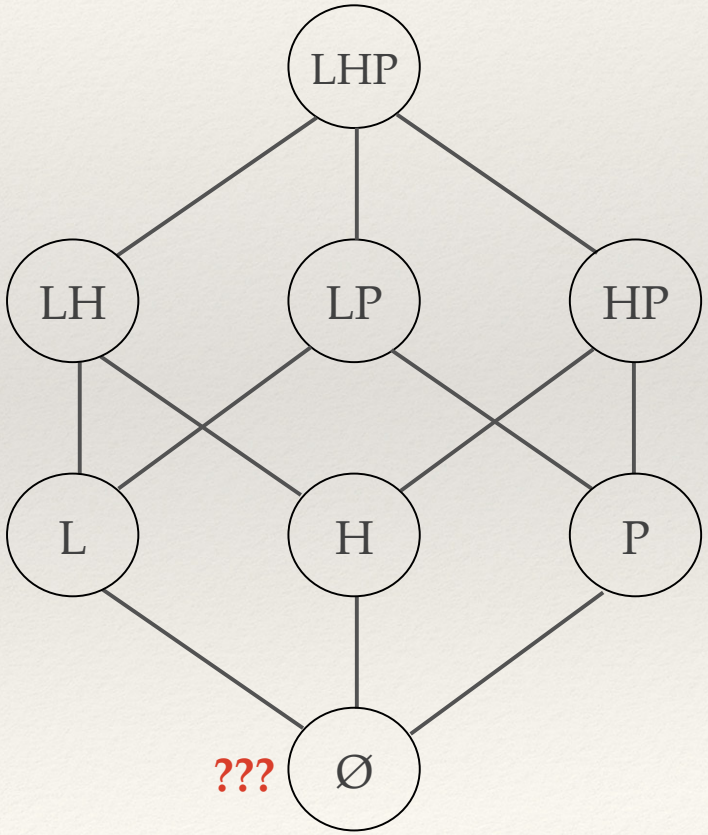
- Using the dialogical evidence gathered, an agent assigns to each community κ and belief α, a *p-score* representing the proportion of agents in that community with positive dialogical evidence.
- Assuming we have 20 agents in total and agent *i* knows through dialogical evidence that 15 agents have belief α and the remaining 5 do not to have the belief α: the p-score for α and Ø according to *i* will be 0.75



Refining this notion…

Suppose that out of the 15 agents *i* knows to believe α, 10 actually belong to community L.

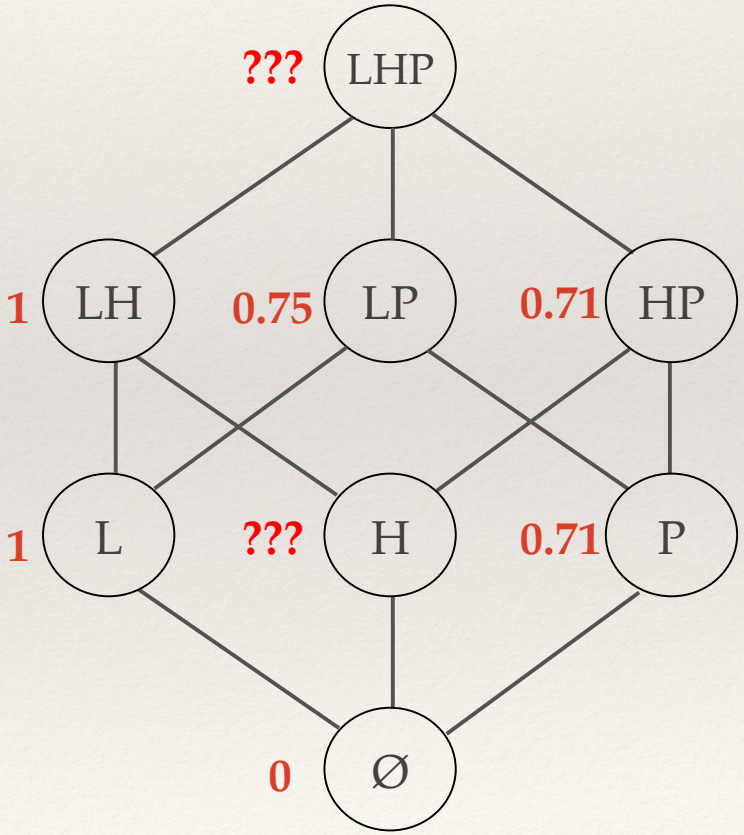
- Then in fact, the reason for believing α could be attributed to the membership to the more specific community.
- This means we then need to remove these agents from Ø's p-score.
- The systematic way of doing this is to successively remove the agents in a community with maximal p-score from the other more general communities they belong too.



We need to refine…

This will yield a distribution such as the one on the right, where some values may be missing.

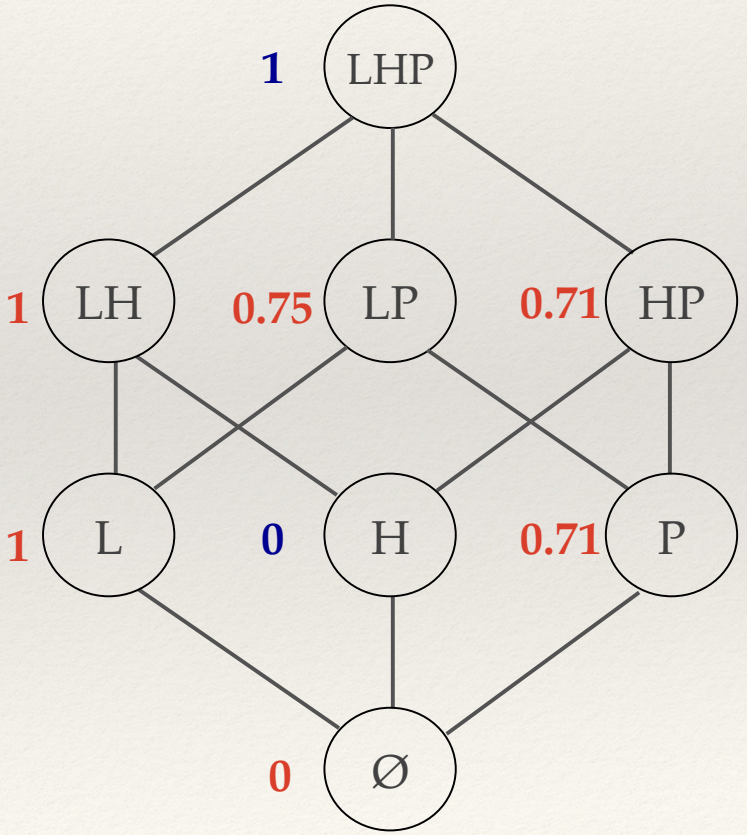
- To complete the assignment, we push the values up, giving each community missing a p-score the maximum of its children.
- So we end up with an assignment F(κ,α) giving the likelihood that a member of the community κ has the belief α.



We need to refine…

This will yield a distribution such as the one on the right, where some values may be missing.

- To complete the assignment, we push the values up, giving each community missing a p-score the maximum of its children.
- So we end up with an assignment F(κ,α) giving the likelihood that a member of the community κ has the belief α.



Community-Based Estimate (CE)

$c_{ij}(\alpha)$ denotes the likelihood *i* assigns to *j* believing α based on *j*'s membership in communities using the assignment F just defined.

DEF $c_{ij}(\alpha)=F(\kappa,\alpha)$, where κ is the most specific community agent *j* belongs to.

Combining the Estimates

DEF $u_{ij}(\alpha)$ denotes the likelihood i assigns to j believing α

- If $d_{ij}(\alpha)>0$, then $u_{ij}(\alpha)=d_{ij}(\alpha)$
- If $d_{ij}(\alpha)=0$, then $u_{ij}(\alpha)=c_{ij}(\alpha)$
- If $d_{ij}(\alpha)=\perp$, then $u_{ij}(\alpha)=0$

We can lift this likelihood to arguments via an appropriate function, e.g., min aggregating the values $u_{ij}(\alpha)$ of the argument's constituent beliefs.

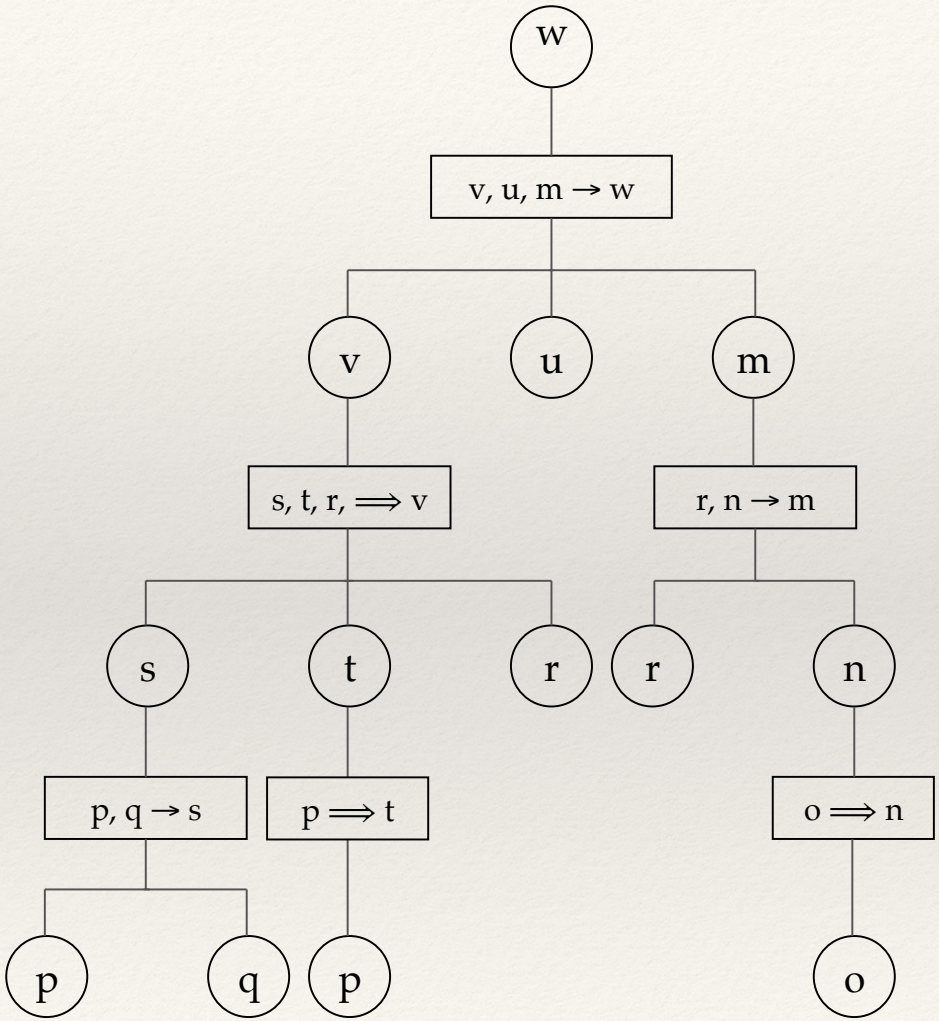
Application: Enthymeme Construction

Sending agent i:

- Uses the values $u_{ij}(\alpha)$ of each belief α in an argument.
- Removes nodes exceeding a given threshold τ
- Reconnects remaining components using ancestor-descendant relationship in the original argument

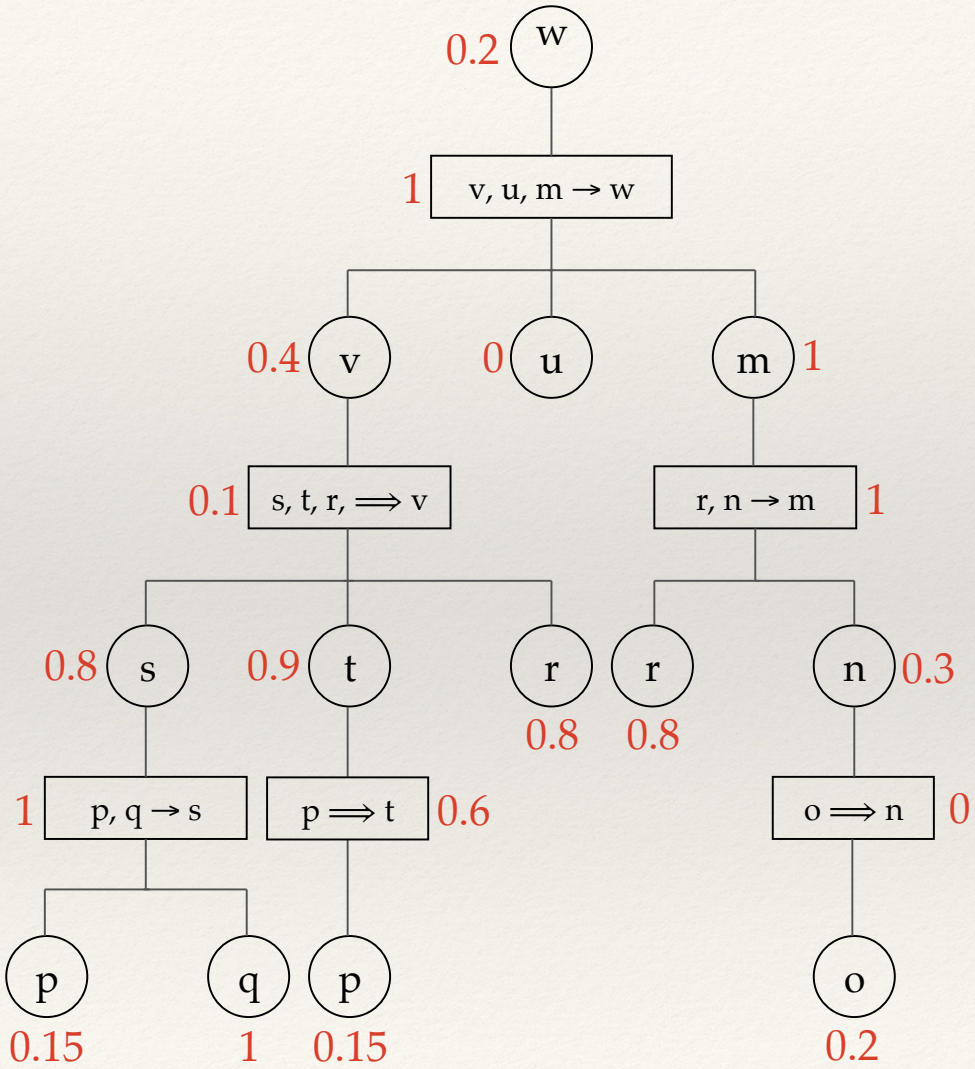
Example

Threshold $\tau = 0.7$



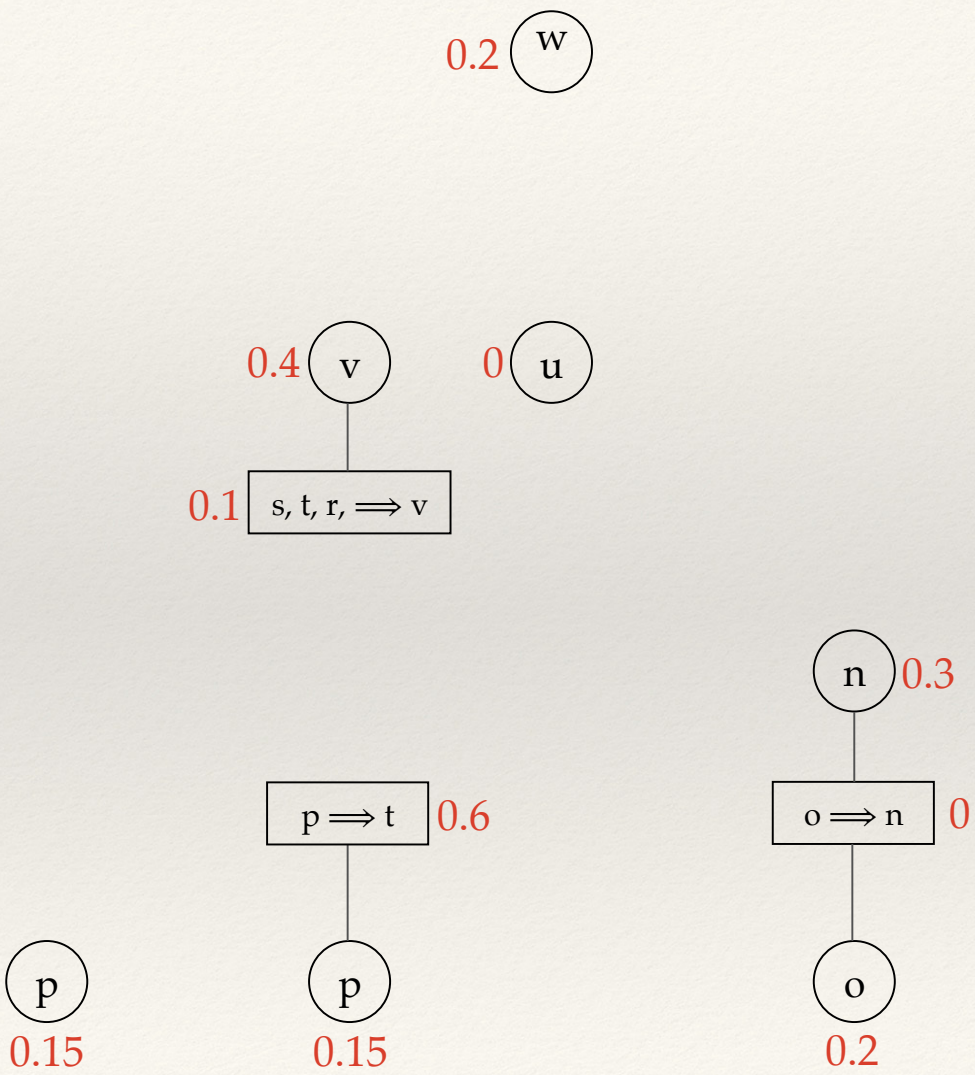
Example

Threshold $\tau = 0.7$



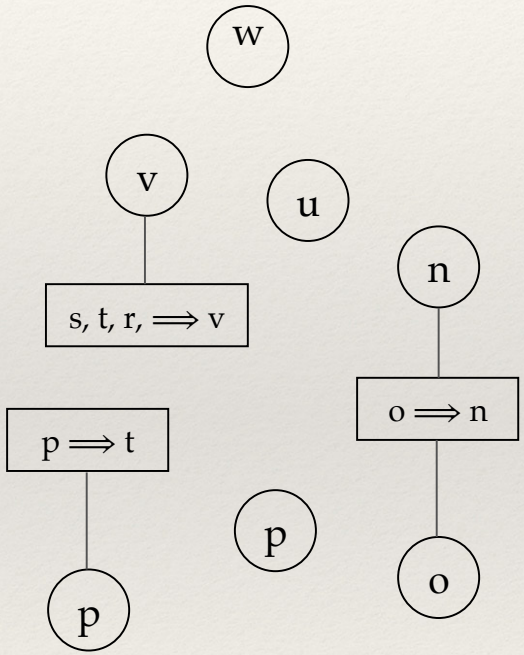
Example

Threshold $\tau = 0.7$



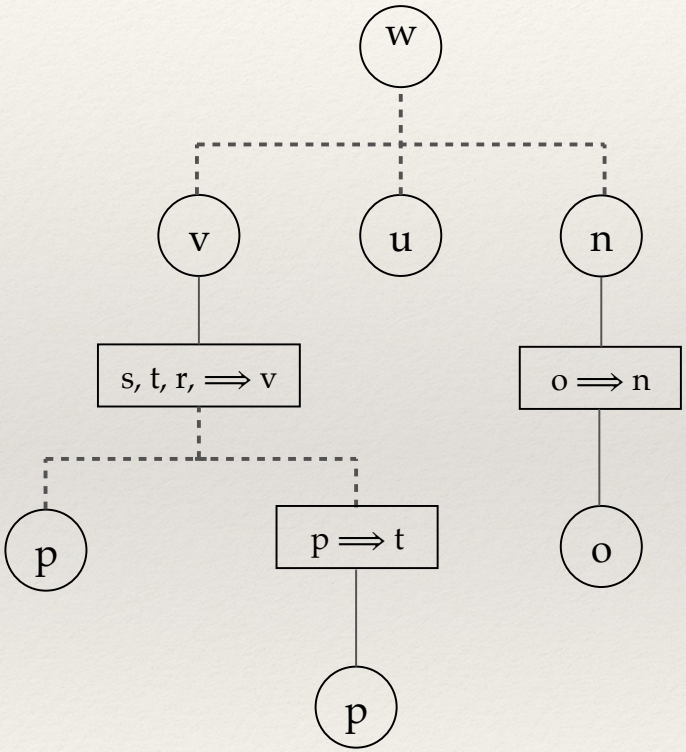
Example

Threshold $\tau = 0.7$



Example

Threshold $\tau = 0.7$



Conclusions and Future work

- We allow agents to store the full argument structure.
- Based on dialogical evidence, we record the likelihood that another agent has certain belief.
- We also use this information to estimate the likelihood that "similar" agents have some beliefs.
- Combining these two estimates, we can estimate the likelihood that an agent can construct a certain argument. In the future we would like to estimate the likelihood that an agent considers an argument to be *acceptable*.
- Our estimates can be made more sophisticated.
- Conduct some empirical evaluation using human subjects?