

Reasoning about and with Protocols

Eric Pacuit

July 18, 2007

ILLC, University of Amsterdam

staff.science.uva.nl/~epacuit

epacuit@science.uva.nl

Joint work (in progress!) with Johan van Benthem

Introduction and Motivation

We are interested in logics for reasoning about rational agents interacting in some (game-theoretic) situation.

Introduction and Motivation

We are interested in logics for reasoning about rational agents interacting in some (game-theoretic) situation.

Motivation

- Social Software
 - R. Parikh. *Social Software. Synthese* **132** (2002).
 - Refine and test our intuitions about interactive situations
 - Verify properties of social procedures
 - *Refine existing social procedures or suggest new ones*
-

Motivating Example

Ann would like Bob to attend her talk; however, she only wants Bob to attend if he is interested in the subject of her talk, not because he is just being polite.

Motivating Example

Ann would like Bob to attend her talk; however, she only wants Bob to attend if he is interested in the subject of her talk, not because he is just being polite.

There is a very simple procedure to solve Ann's problem: *have a (trusted) friend tell Bob the time and subject of her talk.*

Motivating Example

Ann would like Bob to attend her talk; however, she only wants Bob to attend if he is interested in the subject of her talk, not because he is just being polite.

There is a very simple procedure to solve Ann's problem: *have a (trusted) friend tell Bob the time and subject of her talk.*

Is this procedure correct?

Yes, but why?

Test Bob's information: the procedure forces that

Bob *does not* know that Ann knows that he knows about the talk.



Yes, but why?

Test Bob's information: the procedure forces that

Bob *does not* know that Ann knows that he knows about the talk.

Check Bob's preferences: does Bob *want* to be there?

Yes, but why?

Test Bob's information: the procedure forces that

Bob *does not* know that Ann knows that he knows about the talk.

Check Bob's preferences: does Bob *want* to be there?

Are Ann and Bob part of the same research group? If Ann is a leader, then a direct request from her would be interpreted as a command.

Plan

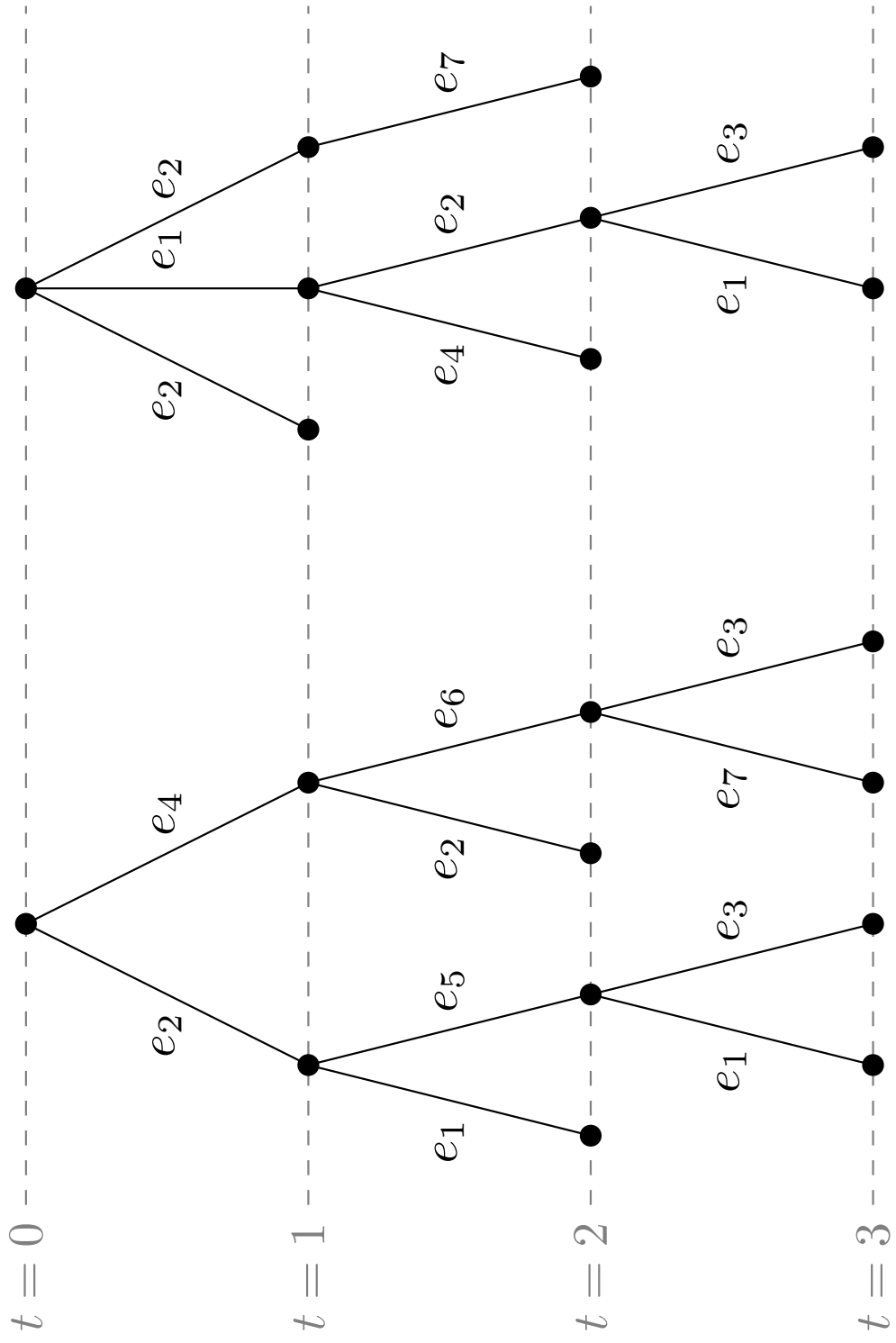
- Introduction and Motivating Example
 - Formalizing the Example
 - Reasoning *inside* a protocol
 - Issues
 - What is a protocol?
 - Observational uncertainty vs. future uncertainty.
 - What do the agents know about the protocol?
 - Concluding Remarks
-

Epistemic Temporal Logic

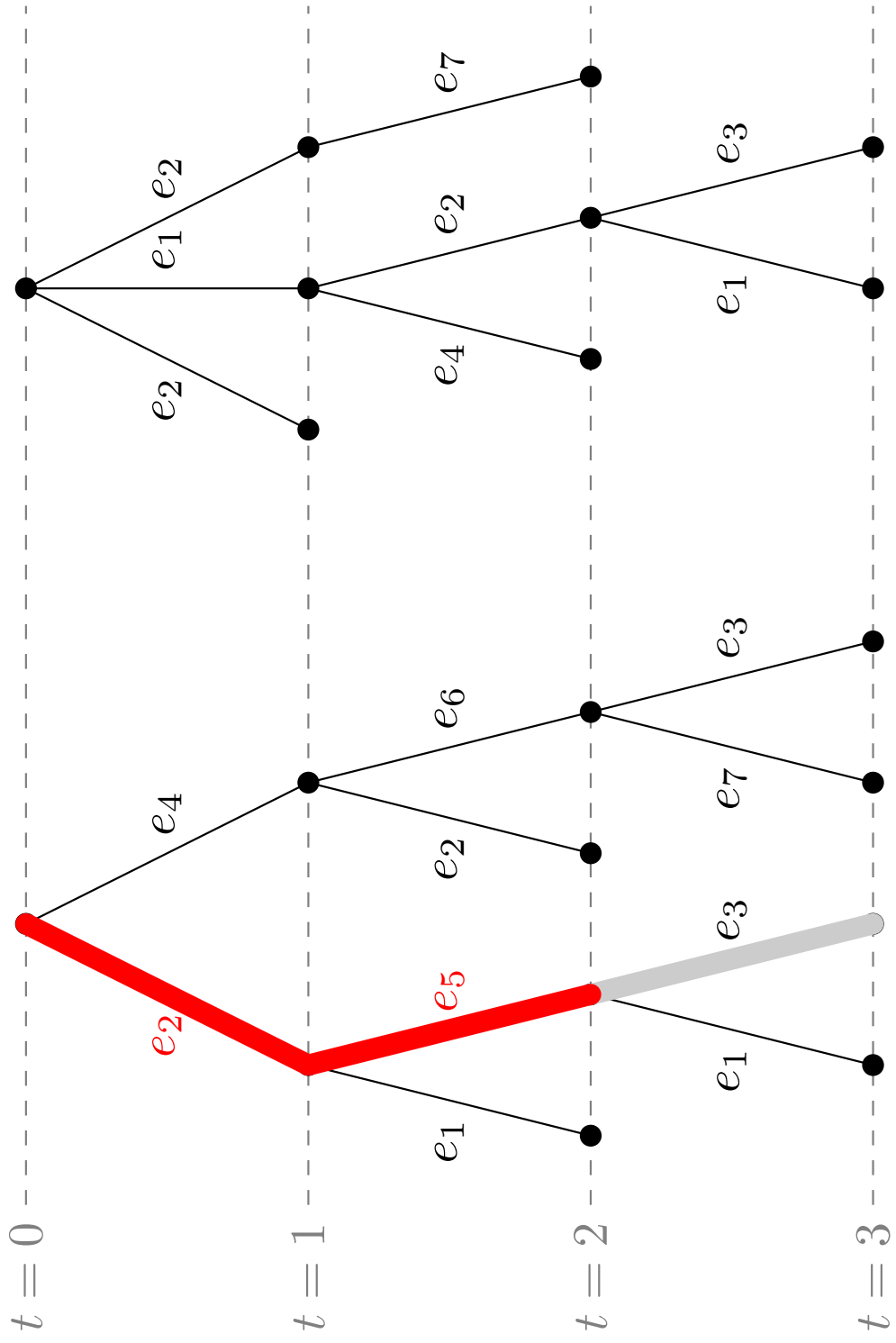
R. Parikh and R. Ramanujam. *A Knowledge Based Semantics of Messages*. *Journal of Logic, Language and Information*, 12: 453 – 467, 2003. (1985).

FHMV. Reasoning about Knowledge. MIT Press, 1995.

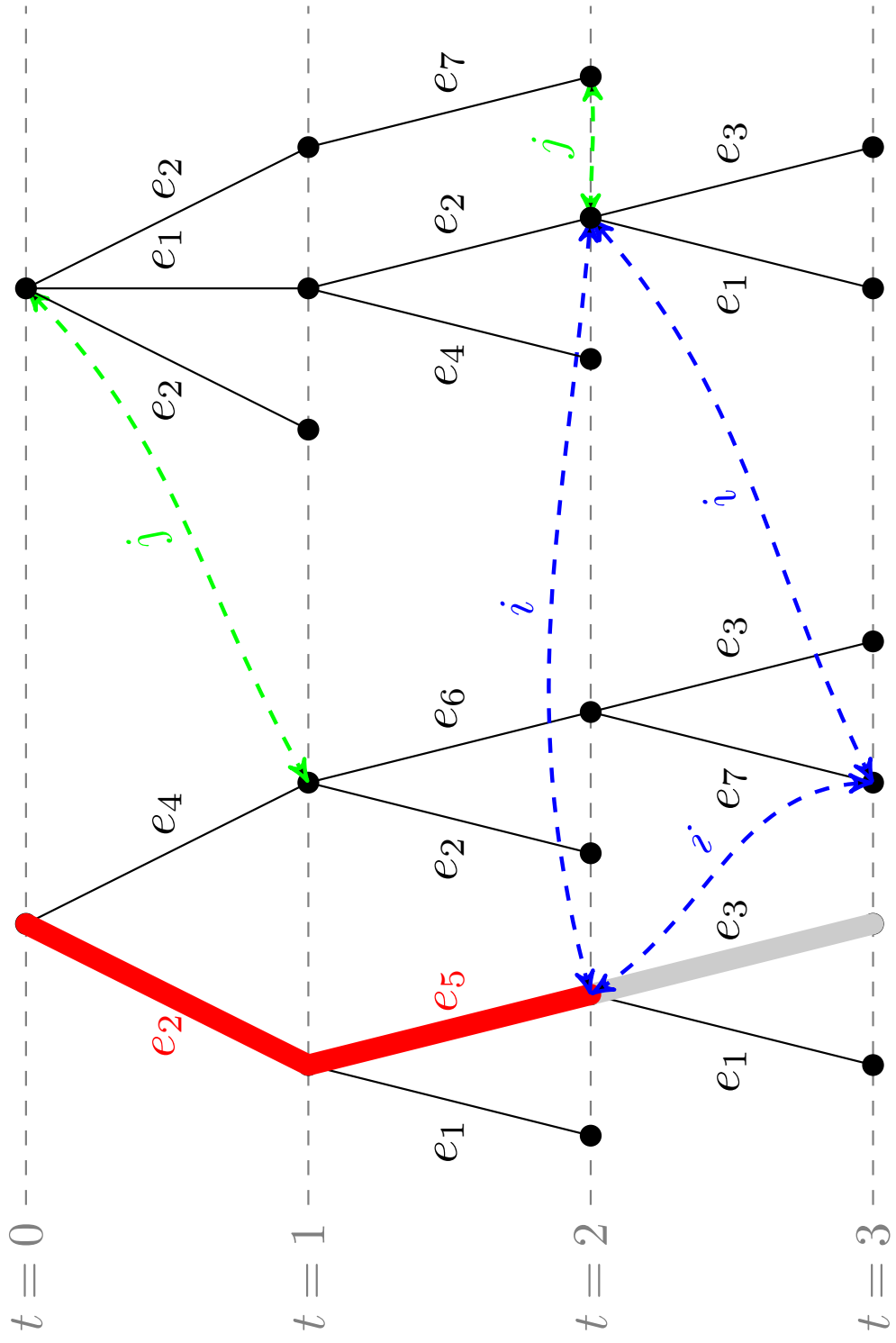
The 'Playground'



The 'Playground'



The 'Playground': Uncertainty



The 'Playground': Notation

Let Σ be a set of events.

The ‘Playground’: Notation

Let Σ be a set of events.

History: Any sequence of events (finite or infinite). I.e., any element of $\Sigma^* \cup \Sigma^\omega$.

The ‘Playground’: Notation

Let Σ be a set of **events**.

History: Any sequence of events (finite or infinite). I.e., any element of $\Sigma^* \cup \Sigma^\omega$.

Protocol: Any set of histories $\mathcal{H} \subseteq \Sigma^* \cup \Sigma^\omega$ is called a **protocol** provided \mathcal{H} is closed under finite prefixes.

The ‘Playground’: Notation

Let Σ be a set of events.

History: Any sequence of events (finite or infinite). I.e., any element of $\Sigma^* \cup \Sigma^\omega$.

Protocol: Any set of histories $\mathcal{H} \subseteq \Sigma^* \cup \Sigma^\omega$ is called a **protocol** provided \mathcal{H} is closed under finite prefixes.

ETL Frame: A tuple $\langle \mathcal{H}, \{\sim_i\}_{i \in A} \rangle$ where \mathcal{H} is a protocol, and for each $i \in A$, \sim_i is a *relation on the set of finite strings in \mathcal{H}* .

Formal Language(s)

- $P\varphi$ (φ is true *sometime* in the past),
 - $F\varphi$ (φ is true *sometime* in the future),
 - $Y\varphi$ (φ is true at *the* previous moment),
 - $N\varphi$ (φ is true at *the* next moment),
 - $N_e\varphi$ (φ is true after event e)
 - $K_i\varphi$ (agent i knows φ) and
 - $C_B\varphi$ (the group $B \subseteq A$ commonly knows φ).
-

Models

An **ETL model** is a structure $\langle \mathcal{H}, \{\sim_i\}_{i \in A}, V \rangle$ where $\langle \mathcal{H}, \{\sim_i\}_{i \in A} \rangle$ is an ETL frame and

$V : At \rightarrow 2^{\text{finite}(\mathcal{H})}$ is a valuation function.

Formulas are interpreted at pairs H, t :

$$H, t \models \varphi$$

Truth in a Model

- $H, t \models P\varphi$ iff there exists $t' \leq t$ such that $H, t' \models \varphi$
- $H, t \models F\varphi$ iff there exists $t' \geq t$ such that $H, t' \models \varphi$
- $H, t \models N\varphi$ iff $H, t + 1 \models \varphi$
- $H, t \models Y\varphi$ iff $t > 1$ and $H, t - 1 \models \varphi$
- $H, t \models K_i\varphi$ iff for each $H' \in \mathcal{H}$ and $m \geq 0$ if $H_t \sim_i H'_m$ then $H', m \models \varphi$
- $H, t \models C\varphi$ iff for each $H' \in \mathcal{H}$ and $m \geq 0$ if $H_t \sim_* H'_m$ then $H', m \models \varphi$.

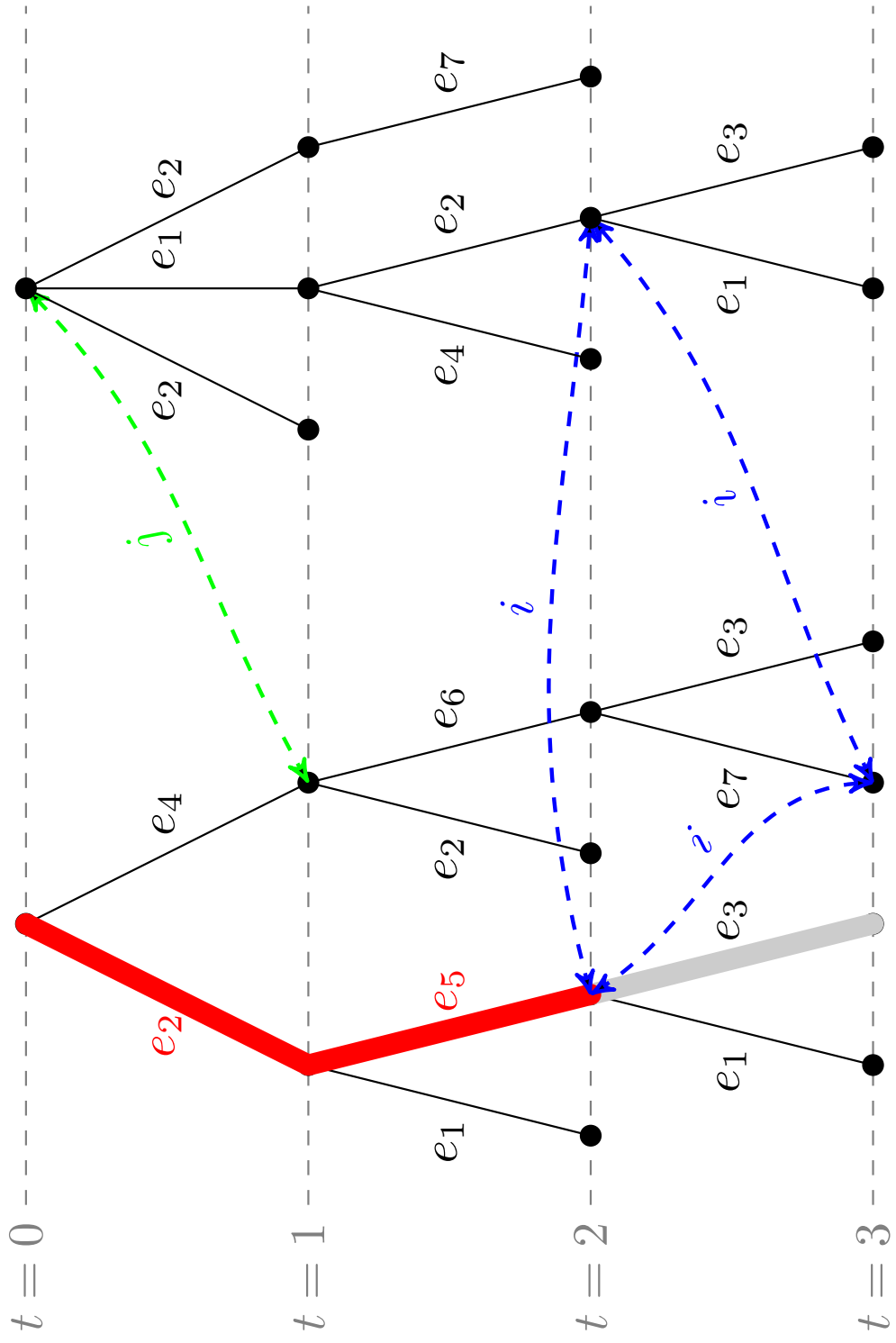
where \sim_* is the reflexive transitive closure of the union of the \sim_i .

Truth in a Model

- $H, t \models P\varphi$ iff there exists $t' \leq t$ such that $H, t' \models \varphi$
- $H, t \models F\varphi$ iff there exists $t' \geq t$ such that $H, t' \models \varphi$
- $H, t \models N\varphi$ iff $H, t + 1 \models \varphi$
- $H, t \models Y\varphi$ iff $t > 1$ and $H, t - 1 \models \varphi$
- $H, t \models K_i\varphi$ iff for each $H' \in \mathcal{H}$ and $m \geq 0$ if $H_t \sim_i H'_m$ then $H', m \models \varphi$
- $H, t \models C\varphi$ iff for each $H' \in \mathcal{H}$ and $m \geq 0$ if $H_t \sim_* H'_m$ then $H', m \models \varphi$.

where \sim_* is the reflexive transitive closure of the union of the \sim_i .

The 'Playground': Uncertainty



Digression: Defining Uncertainty

Digression: Defining Uncertainty

1. Histories or Runs?
 - Let L be a set of **local states**.
 - A **run** $r \in \mathcal{R}$ is a function $r : \mathbb{N} \rightarrow L^{n+1}$
 - Agent i cannot distinguish two points if it is in the same state in both: $(r, t) \sim_i (r', t')$ iff $r(t)_i = r'(t')_i$.
-

Digression: Defining Uncertainty

1. Histories or Runs?
 - Let L be a set of **local states**.
 - A **run** $r \in \mathcal{R}$ is a function $r : \mathbb{N} \rightarrow L^{n+1}$
 - Agent i cannot distinguish two points if it is in the same state in both: $(r, t) \sim_i (r', t')$ iff $r(t)_i = r'(t')_i$.

2. **Local View Functions**: Suppose that $\Sigma = \cup_{i \in A} E_i$.

For each $i \in A$ define $\lambda_i : \text{FinPre}(\mathcal{H}) \rightarrow E_i^*$

For each **finite** $H \in \mathcal{H}$,

$$H \sim_i H' \text{ iff } \lambda_i(H) = \lambda_i(H')$$



Digression: Defining Uncertainty

Characterization Theorem(s) Each approach is “equivalent”.

EP. *Some Comments on History Based Structures*. Journal of Applied Logic, 2007.

Parameters of the Models

1. Expressivity of the formal language. Does the language include a common knowledge operator? A future operator? Both? Both? Branching time or linear time?



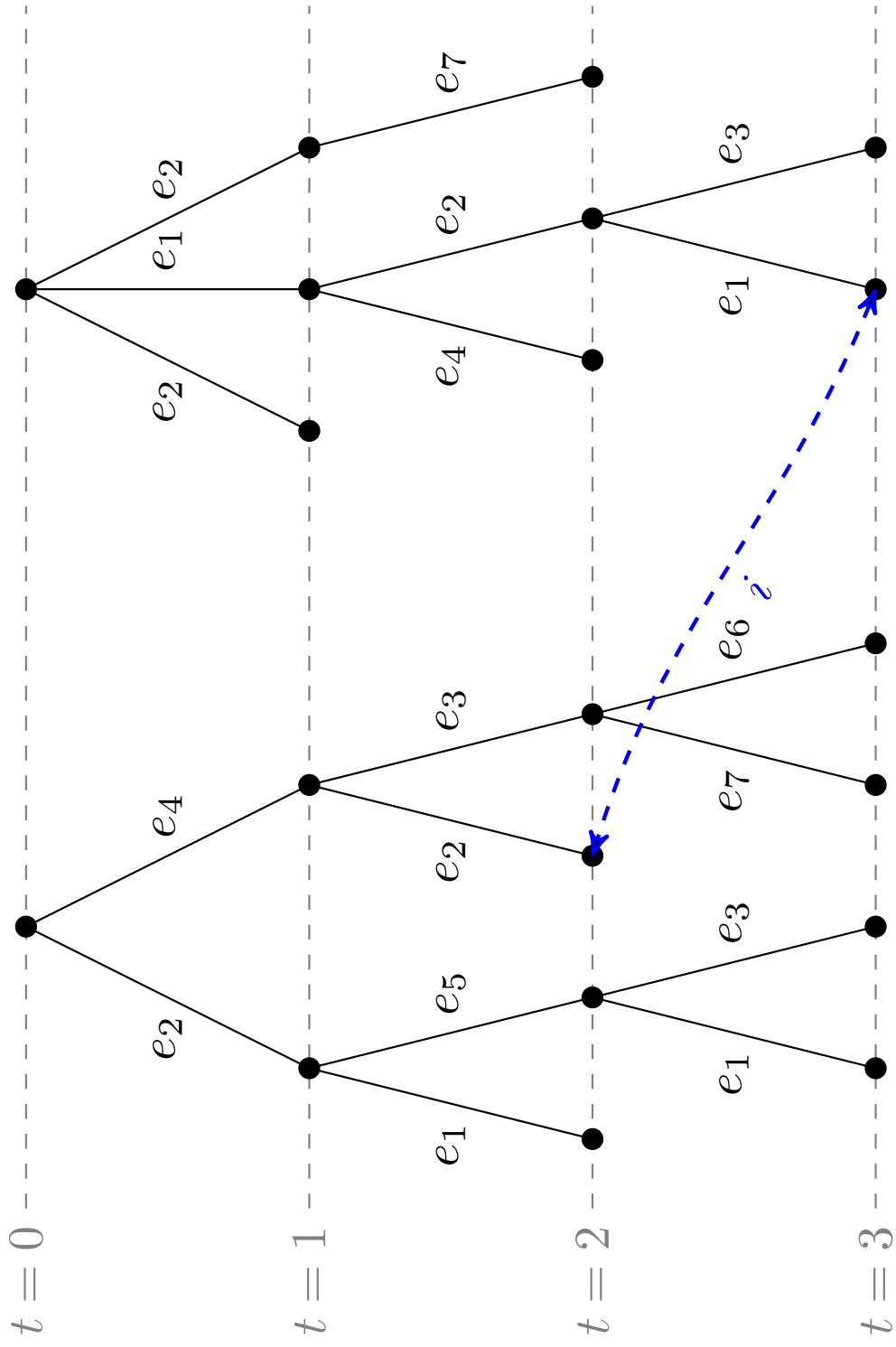
Parameters of the Models

1. Expressivity of the formal language. Does the language include a common knowledge operator? A future operator? Both? Branching time or linear time?
 2. Structural conditions on the underlying event structure. Do we restrict to protocol frames (finitely branching trees)? Finitely branching forests? Or, arbitrary ETL frames?
-

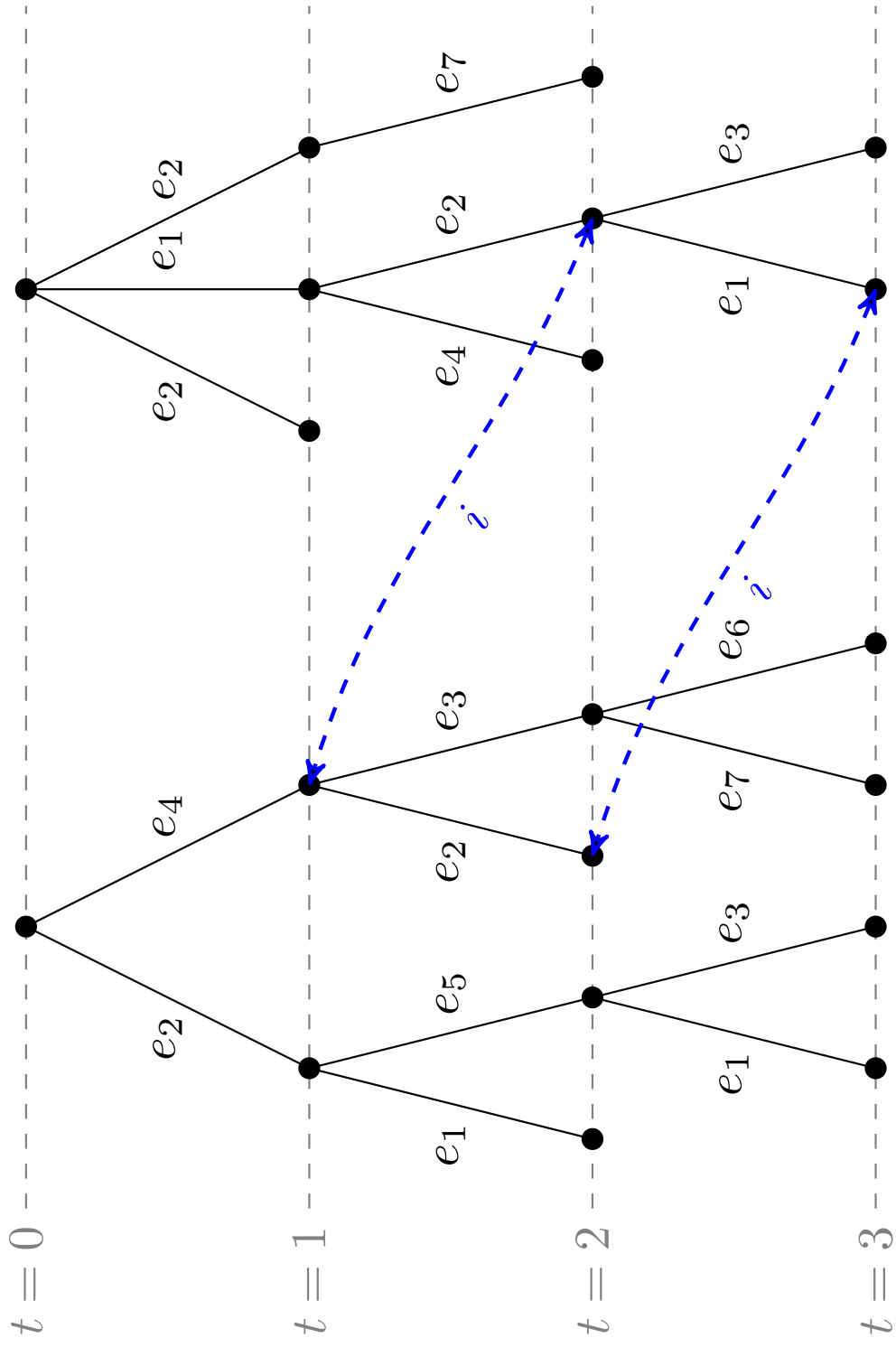
Parameters of the Models

1. Expressivity of the formal language. Does the language include a common knowledge operator? A future operator? Both? Branching time or linear time?
 2. Structural conditions on the underlying event structure. Do we restrict to protocol frames (finitely branching trees)? Finitely branching forests? Or, arbitrary ETL frames?
 3. Conditions on the reasoning abilities of the agents. Do the agents satisfy perfect recall? No miracles? Synchronization?
-

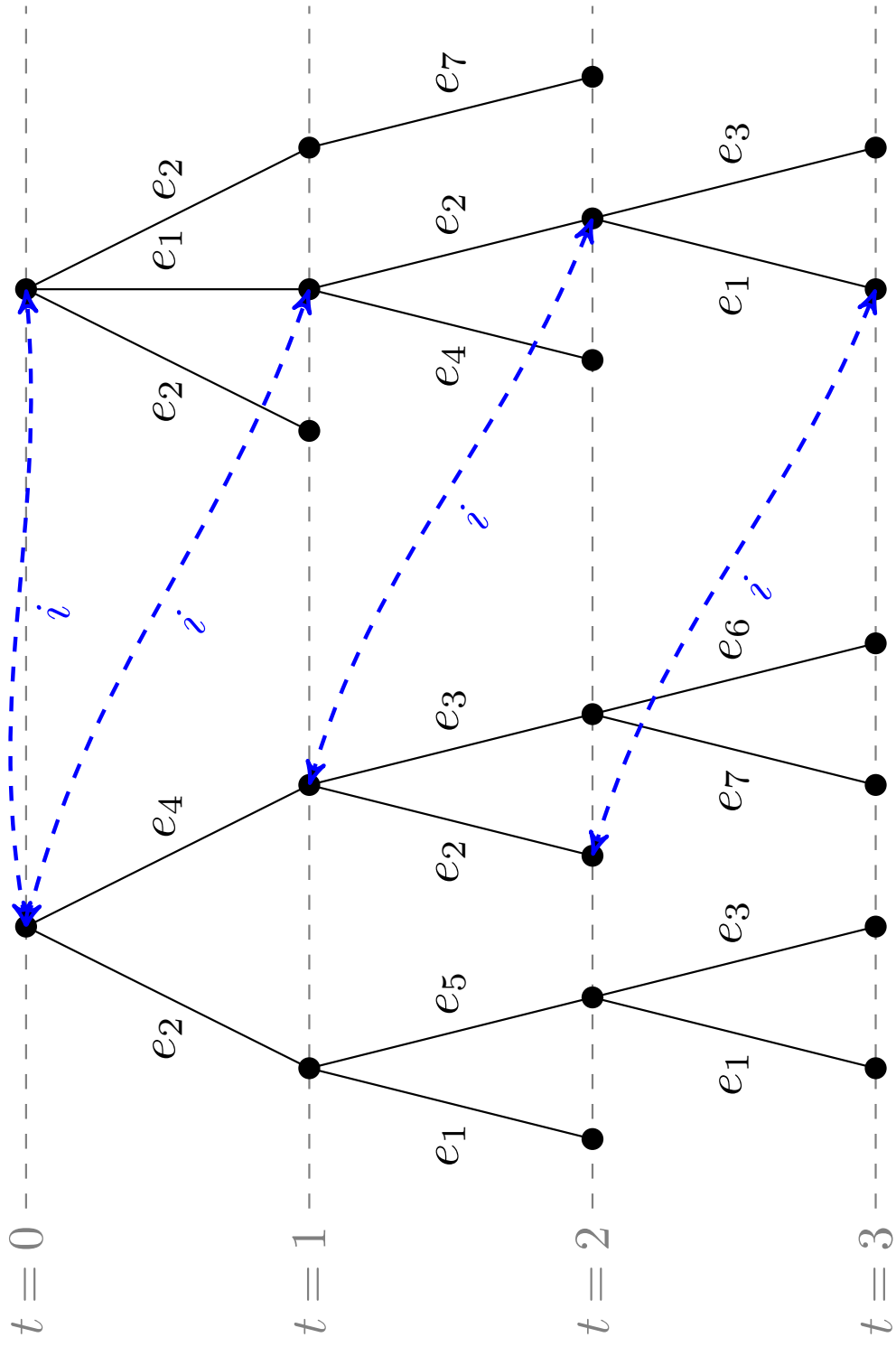
Perfect Recall



Perfect Recall



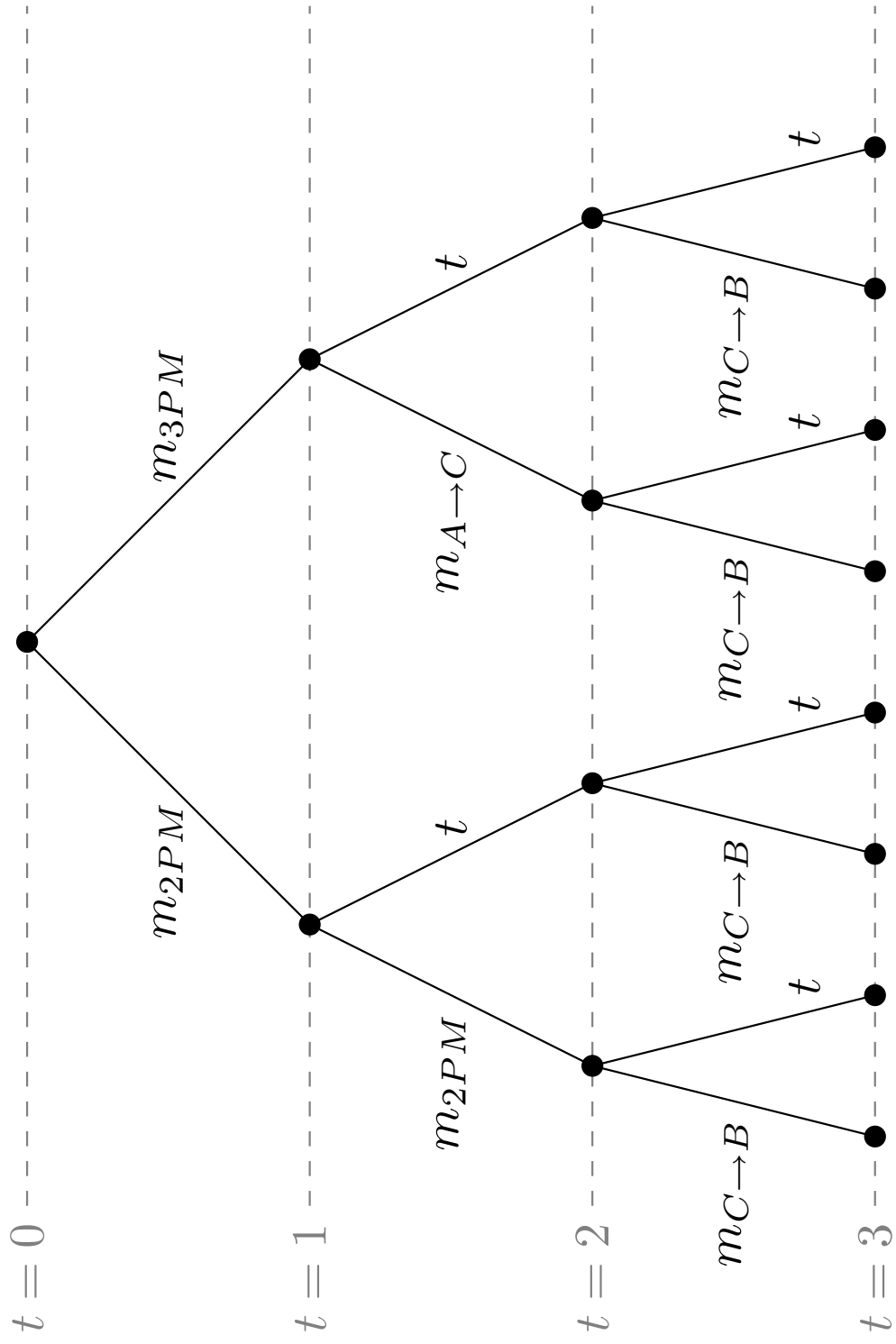
Perfect Recall

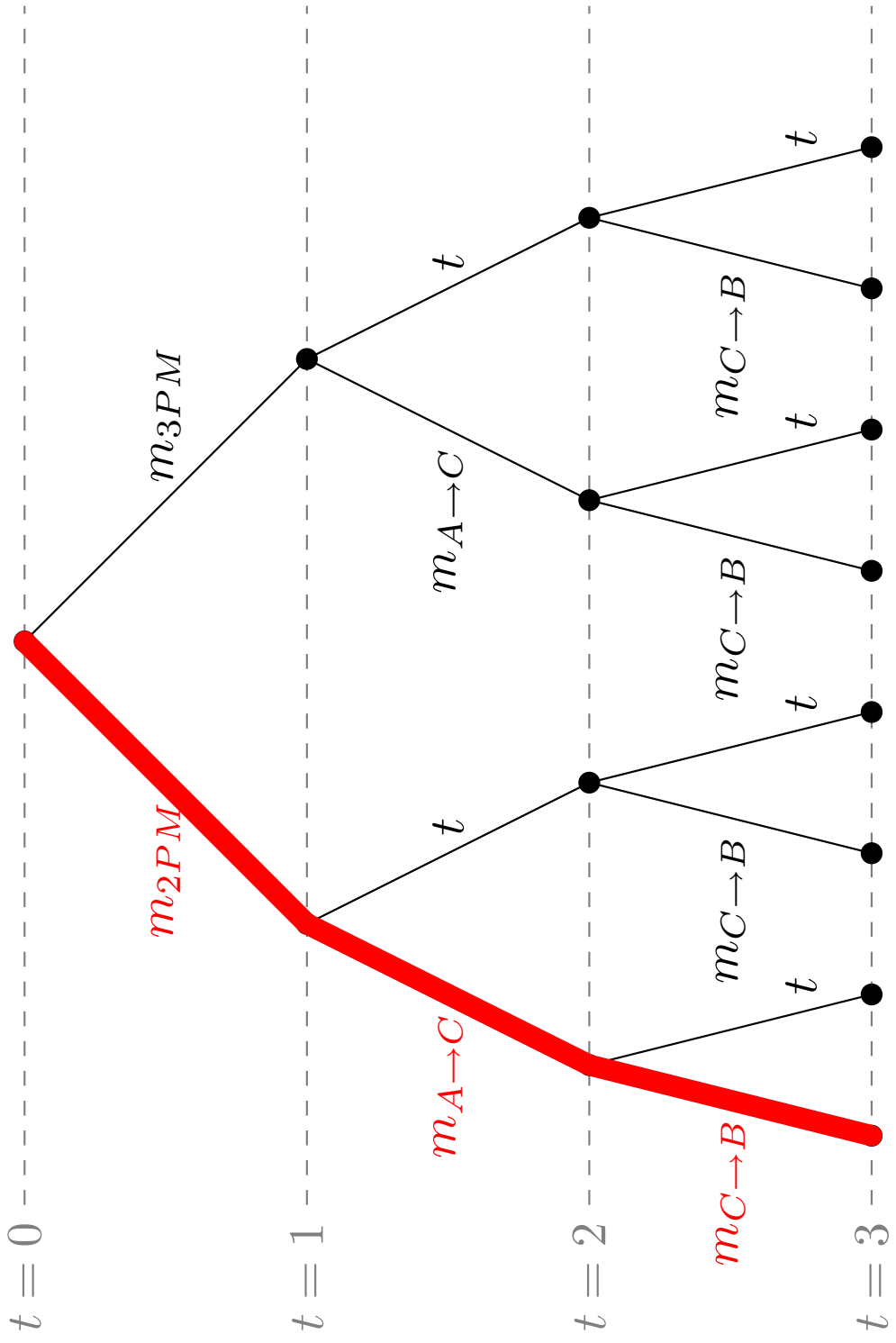


Returning to the Example

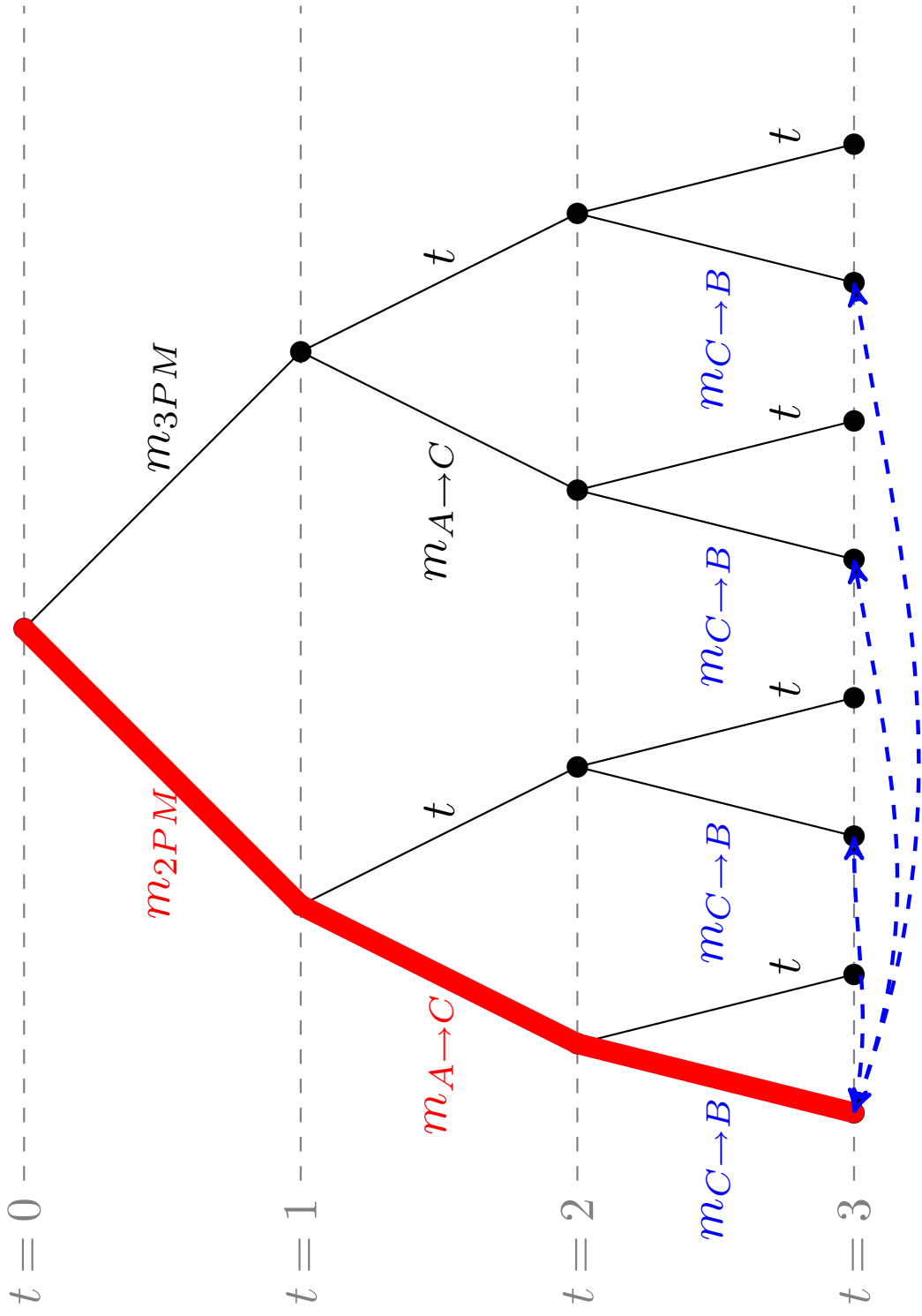
Test Bob's information: the procedure forces that

Bob *does not* know that Ann knows that he knows about the talk.

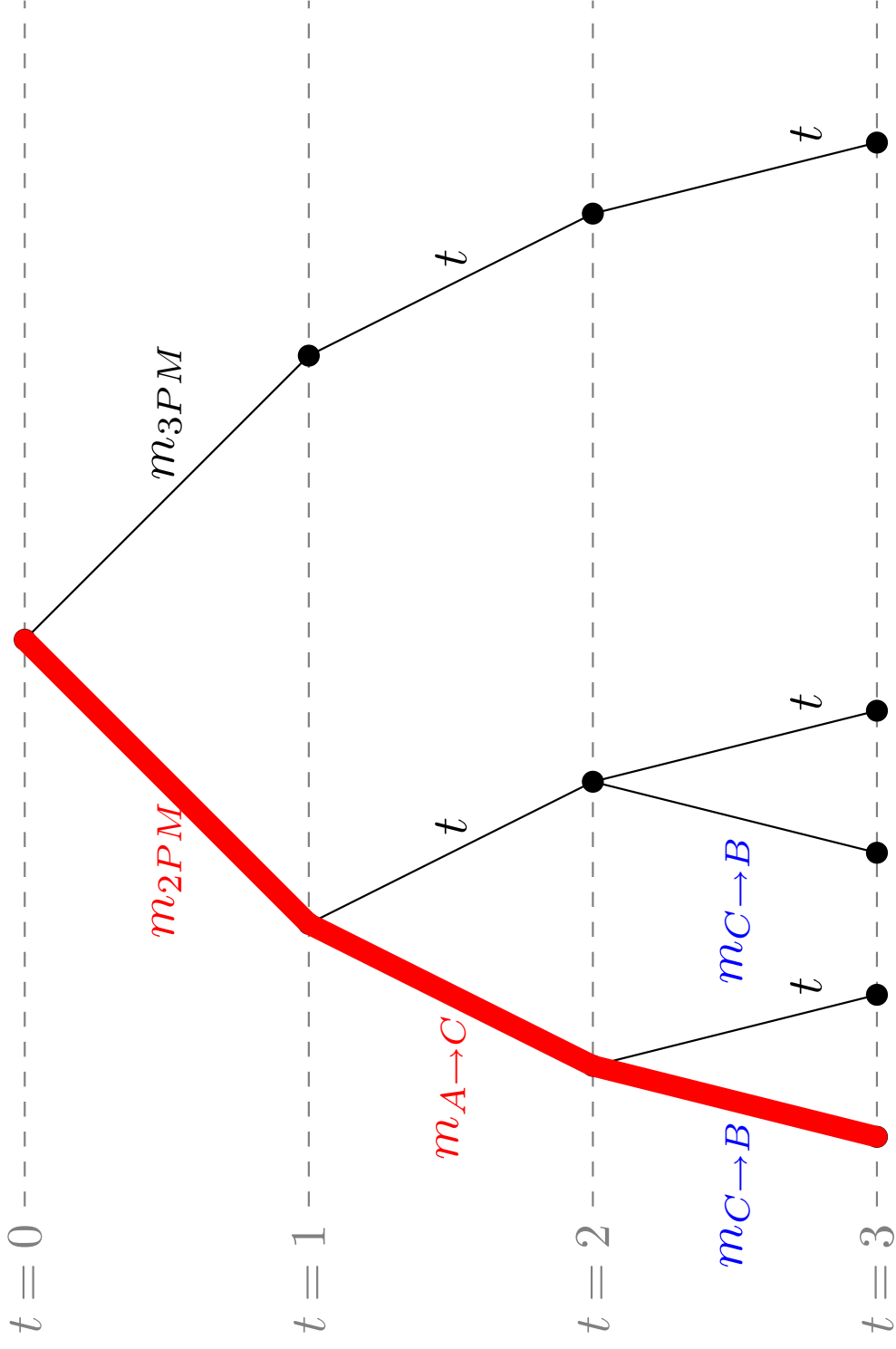




$H, 3 \models \varphi$

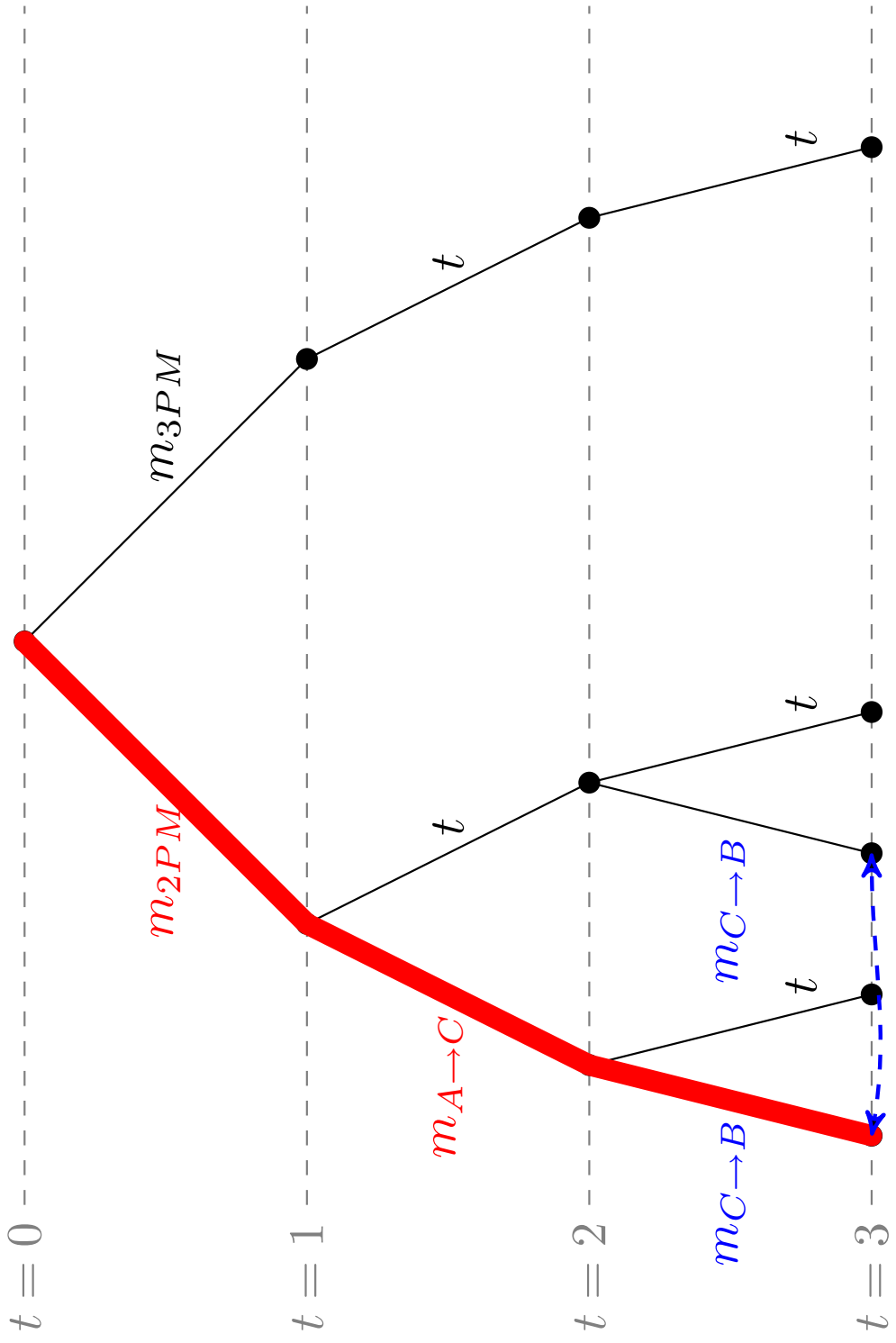


Bob's uncertainty: $H, 3 \models \neg K_B P_{2PM}$



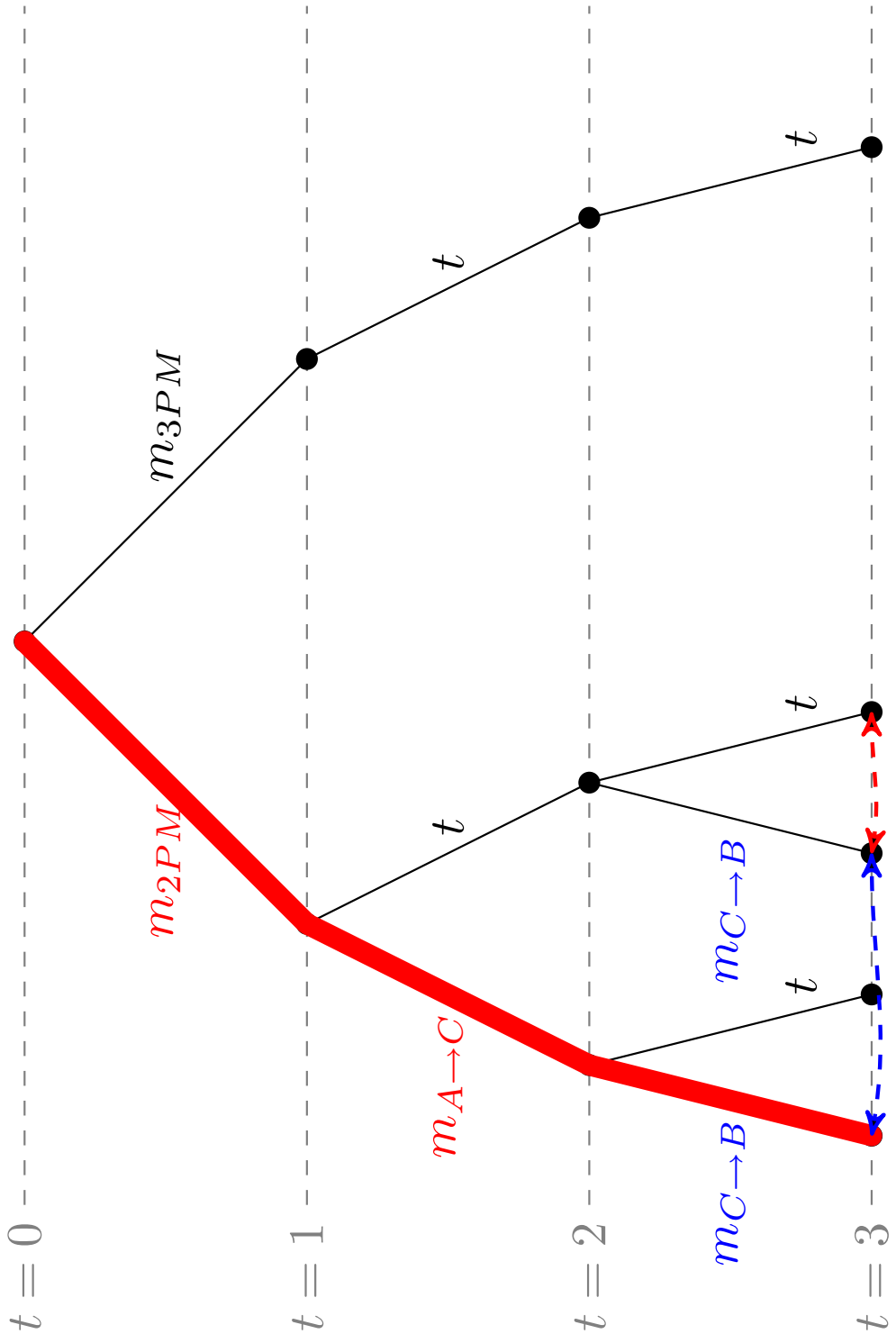
Bob's uncertainty + 'Protocol information':

$$H, 3 \models \neg K_B K_A K_B P_{2PM}$$



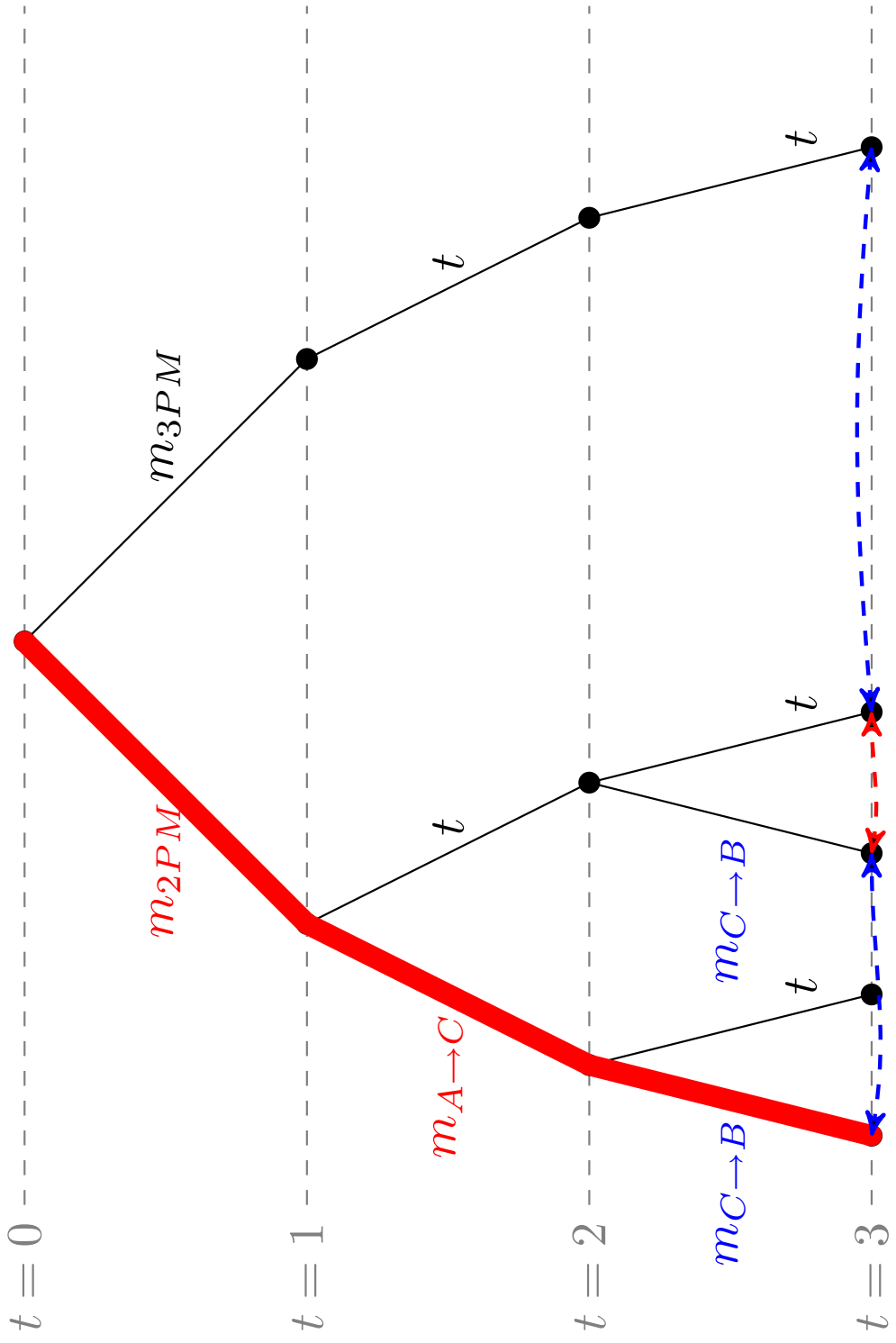
Bob's uncertainty + 'Protocol information':

$$H, 3 \models \neg K_B K_A K_B P_{2PM}$$



Bob's uncertainty + 'Protocol information':

$$H, 3 \models \neg K_B K_A K_B P_{2PM}$$



Bob's uncertainty + 'Protocol information':

$$H, 3 \models \neg K_B K_A K_B P_{2PM}$$

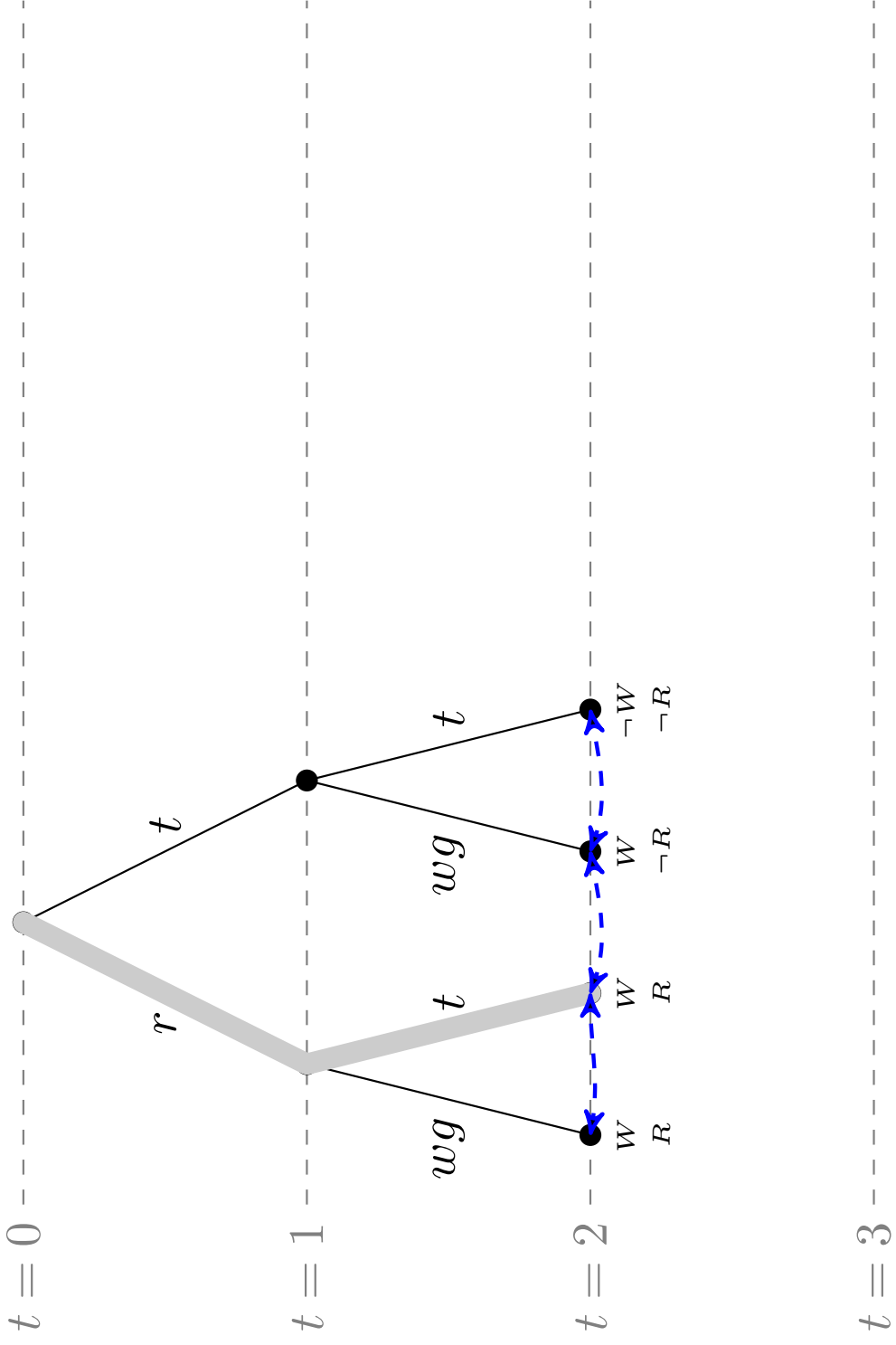
Can we be more precise about

- where the protocol comes from?
 - what the agents know about the protocol?
-

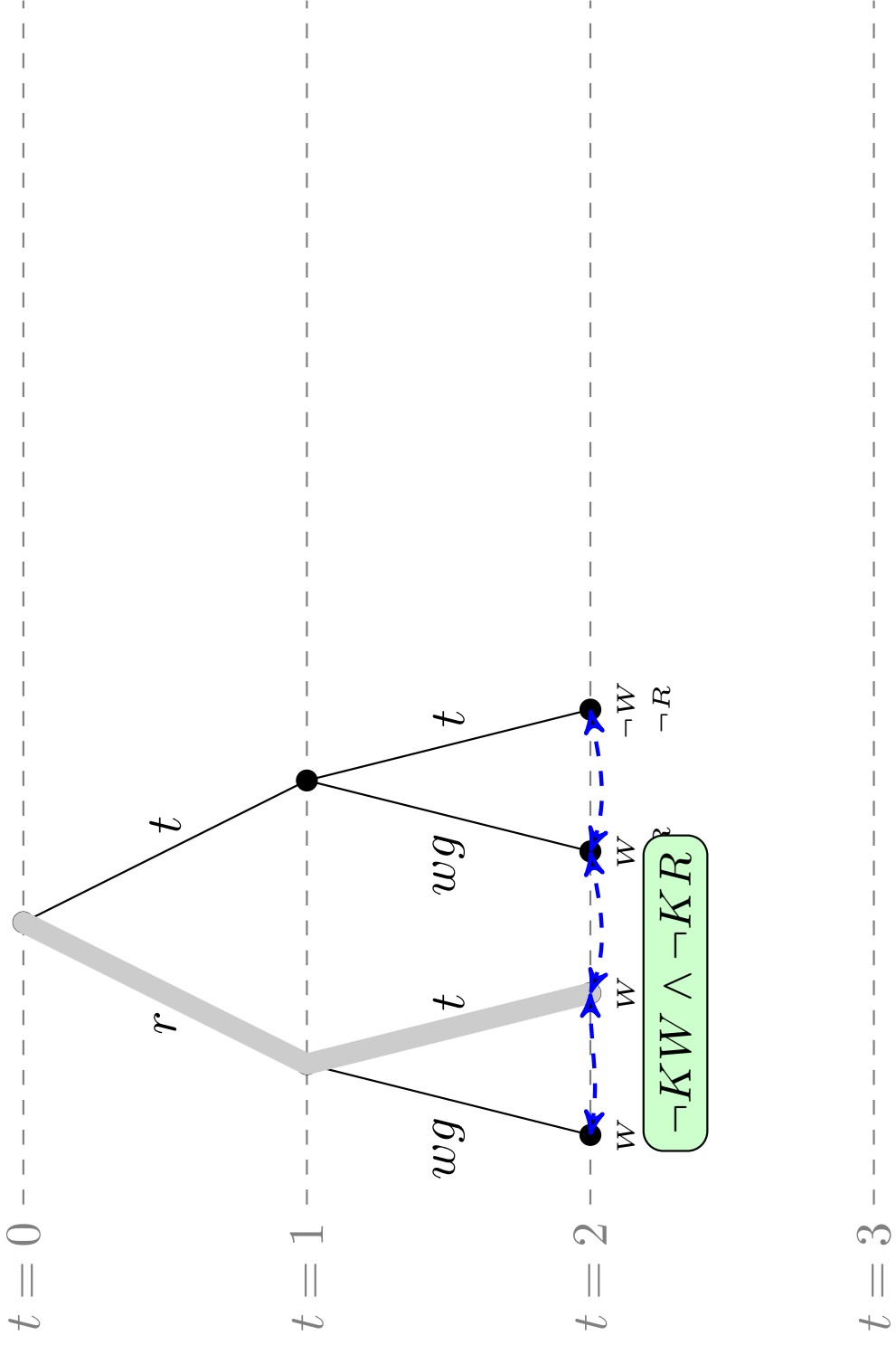
Plan

- Introduction and Motivating Example
 - Formalizing the Example
 - Reasoning *inside* a protocol
 - Issues
 - What is a protocol?
 - Observational uncertainty vs. future uncertainty.
 - What do the agents know about the protocol?
 - Concluding Remarks
-

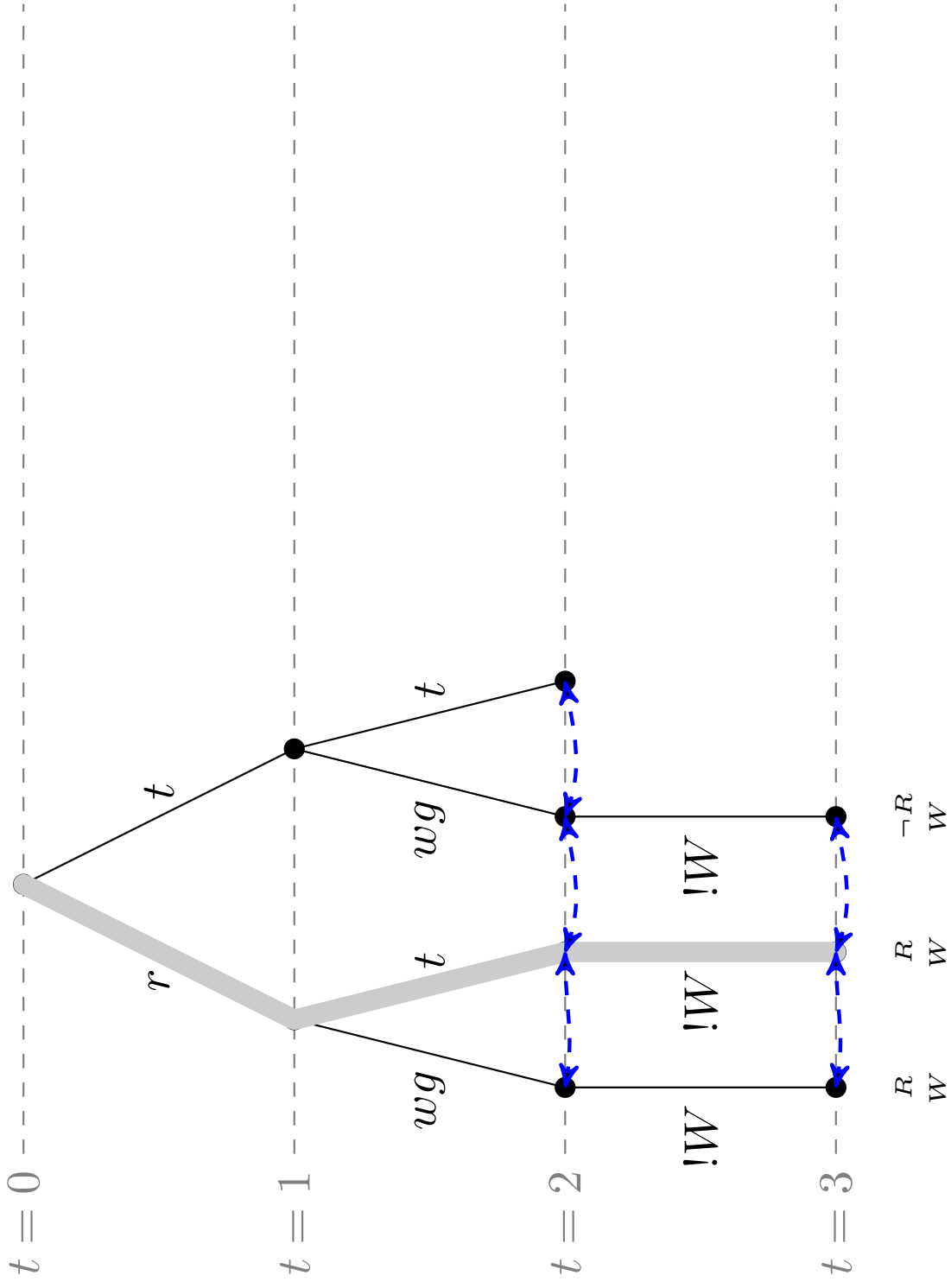
Learning From the Protocol



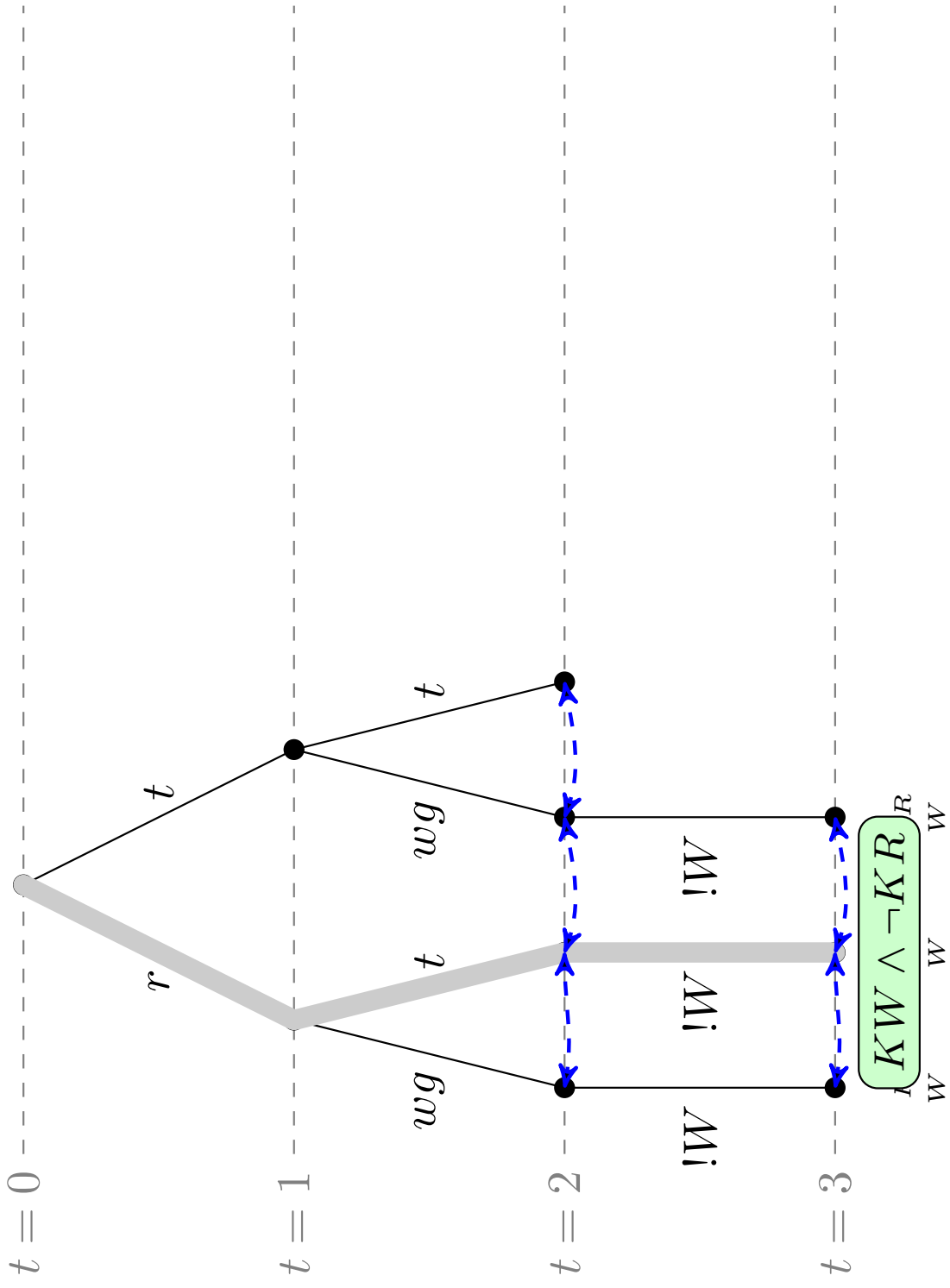
Learning From the Protocol



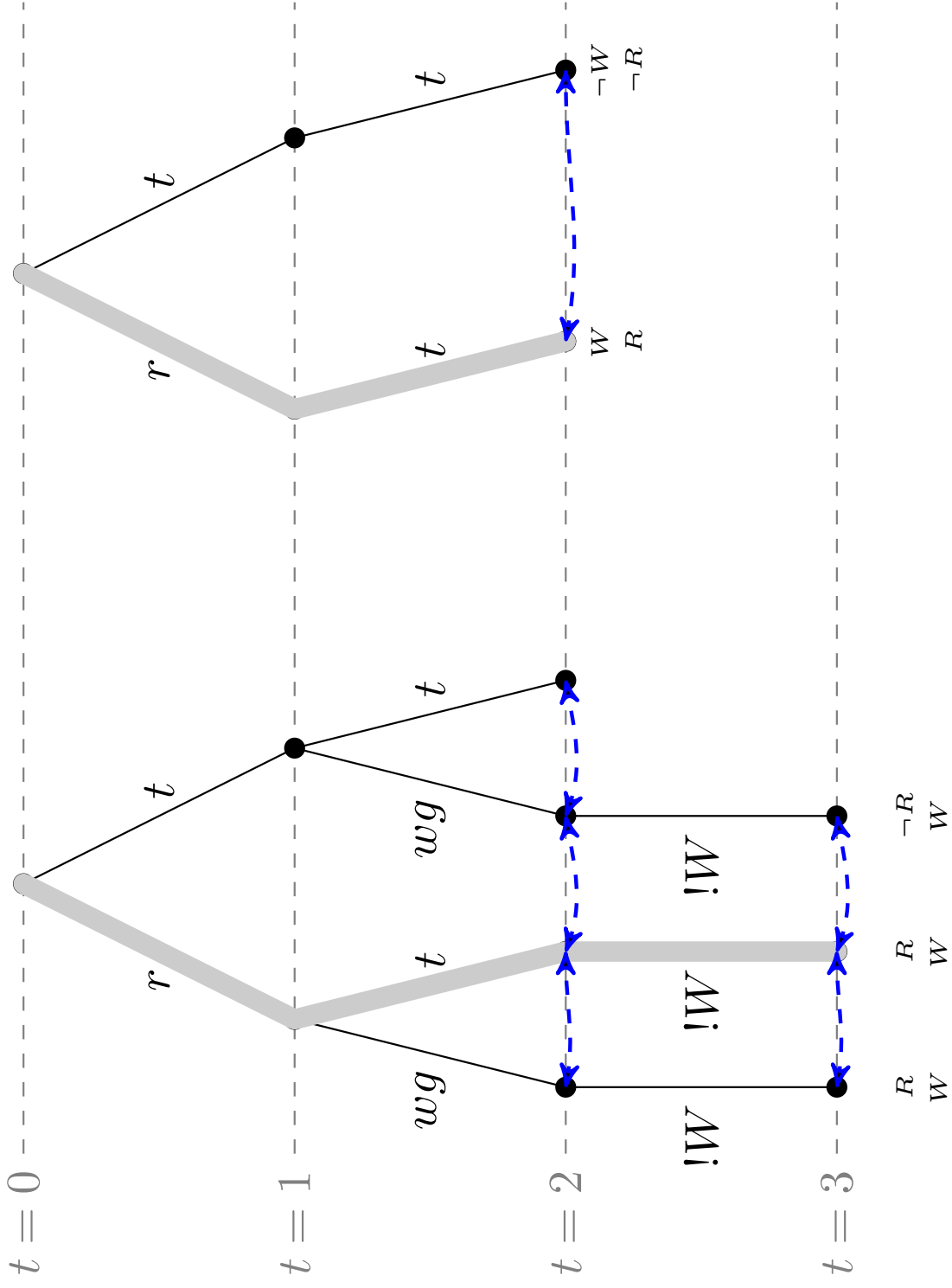
Learning From the Protocol



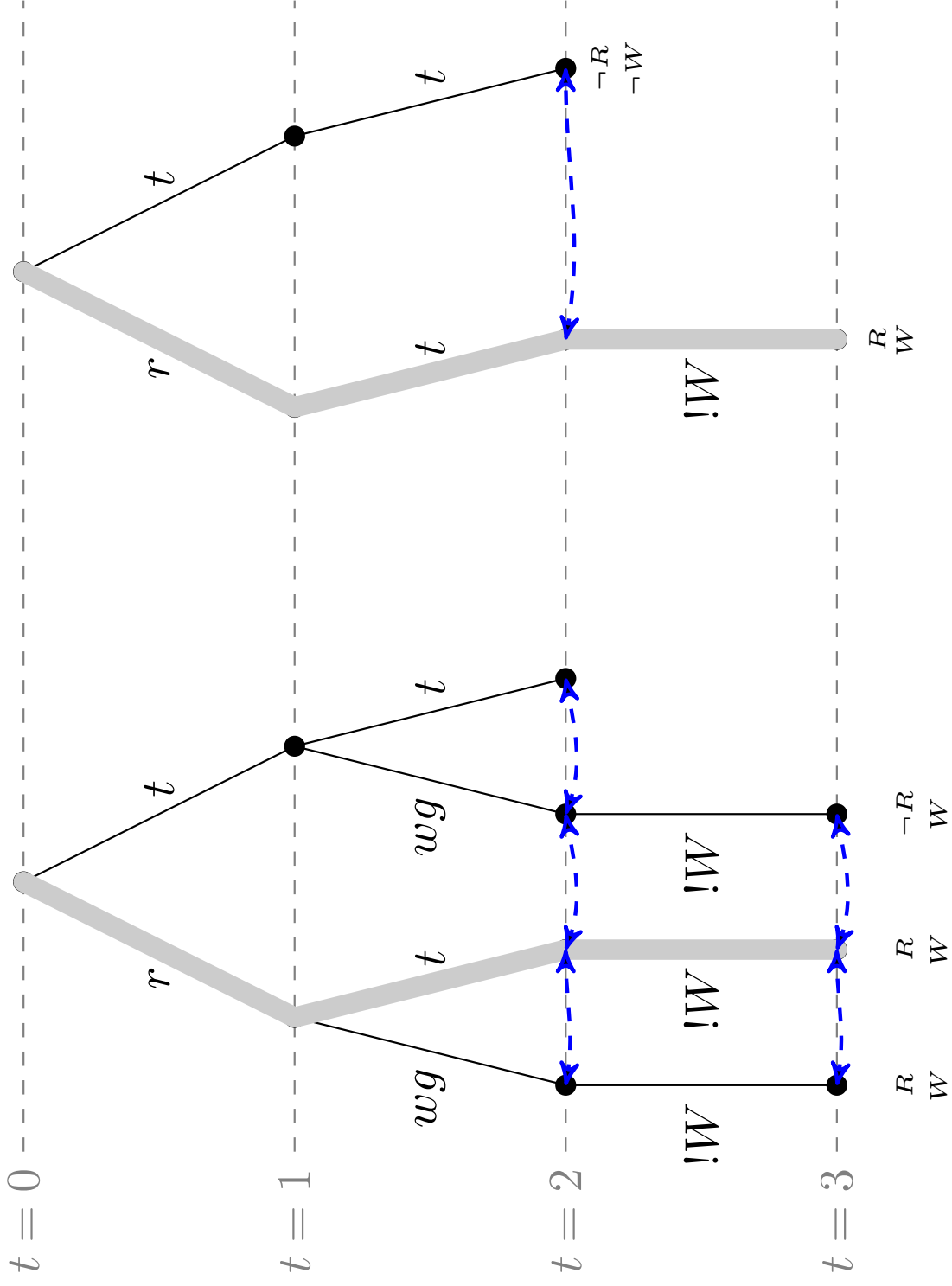
Learning From the Protocol



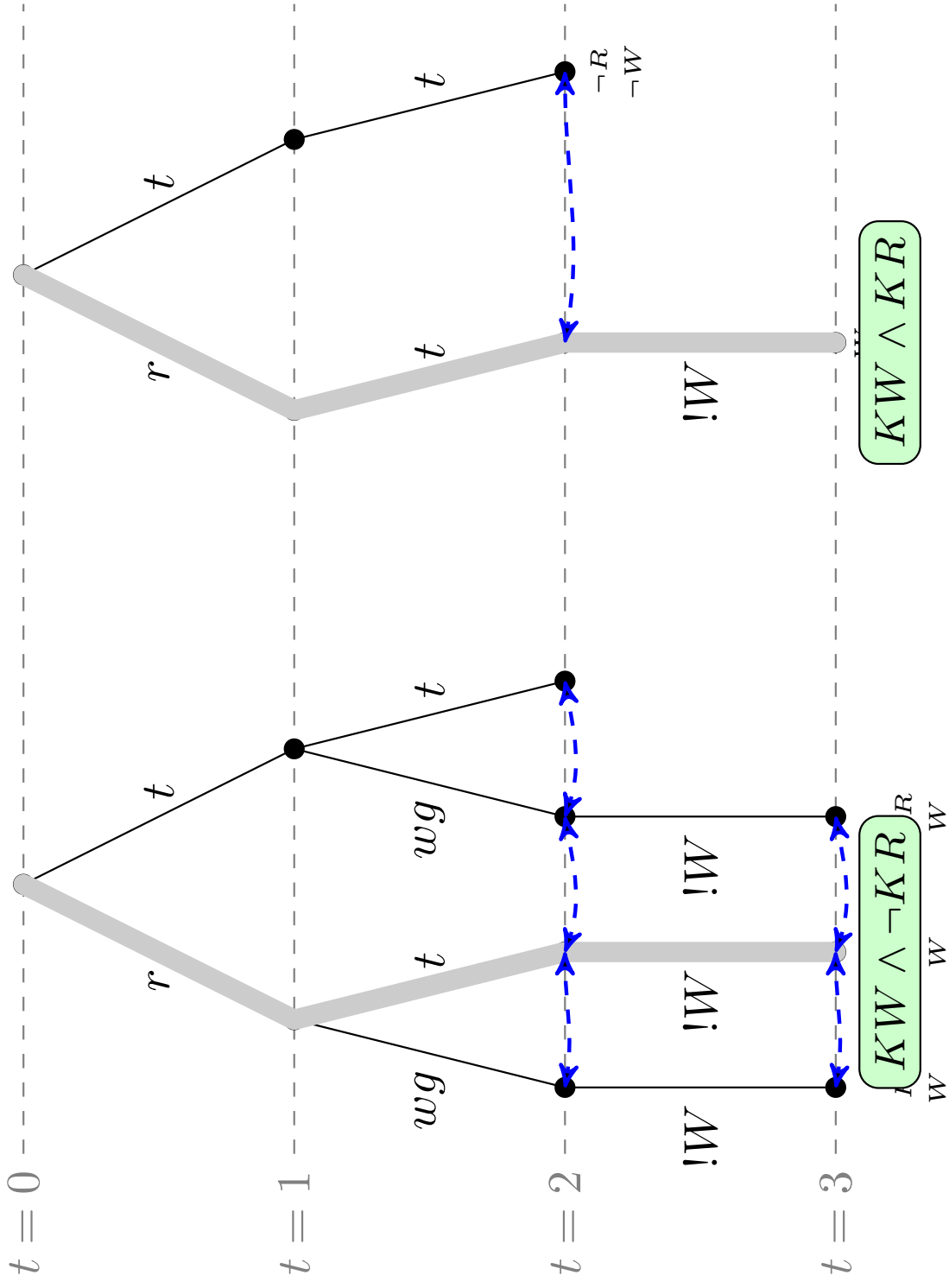
Learning From the Protocol



Learning From the Protocol



Learning From the Protocol



Plan

- Introduction and Motivating Example
 - Formalizing the Example
 - Reasoning *inside* a protocol
 - Issues
 - What is a protocol?
 - Observational uncertainty vs. future uncertainty.
 - What do the agents know about the protocol?
 - Concluding Remarks
-

What is a Protocol?

Given the full tree T of events, a **protocol** is *any* subtree of T .

What is a Protocol?

Given the full tree T of events, a **protocol** is *any* subtree of T .

- A protocol is the set of histories compatible with some **process**, i.e., it is the “unwinding” of a (multi-agent) state transition system.



What is a Protocol?

Given the full tree T of events, a **protocol** is *any* subtree of T .

- A protocol is the set of histories compatible with some **process**, i.e., it is the “unwinding” of a (multi-agent) state transition system.
 - A protocol is the set of histories satisfying some **property**:
 - Physical properties: every message is eventually answered, no message is received before it is sent
 - Agent types: agent i is the **type** of agent who always lies, agent j is the **type** who always tells the truth
-

What is a Protocol?

Given the full tree T of events, a **protocol** is *any* subtree of T .

- A protocol is the set of histories compatible with some **process**, i.e., it is the “unwinding” of a (multi-agent) state transition system.
 - A protocol is the set of histories satisfying some **property**:
 - Physical properties: every message is eventually answered, no message is received before it is sent
 - Agent types: agent i is the **type** of agent who always lies, agent j is the **type** who always tells the truth
 - A protocol is the set of histories of an extensive game consistent with a **strategy profile**.
-

Defining a Protocol

Defining a Protocol

1. What formal language should we use to define the protocol?
 2. What models do we have in mind?
-

Defining a Protocol

1. What formal language should we use to define the protocol?
2. What models do we have in mind?

Given a formula φ , two ways to think about defining a protocol:

Set of histories: the set of histories P in the full event tree T such that $h \in P$ iff $h \models \varphi$

Set of models: the set $\text{Mod}(\varphi)$ (the set of models of φ)

Plan

- Introduction and Motivating Example
 - Formalizing the Example
 - Reasoning *inside* a protocol
 - Issues
 - What is a protocol?
 - Observational uncertainty vs. future uncertainty.
 - What do the agents know about the protocol?
 - Concluding Remarks
-

Backward-Looking Uncertainty

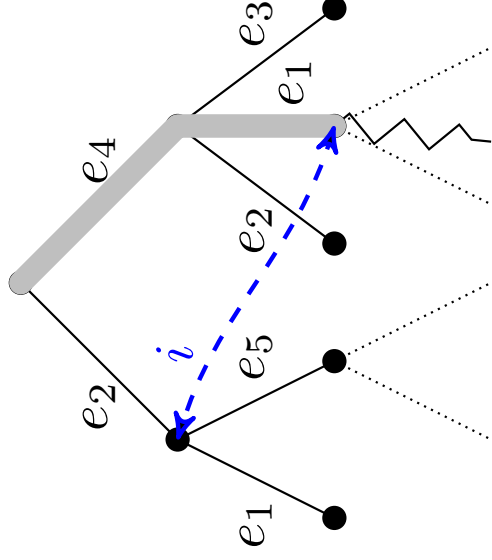
Given two finite histories h and h' ,

$h \sim_i h'$ means *given the events i has observed, h and h' are indistinguishable*

Backward-Looking Uncertainty

Given two finite histories h and h' ,

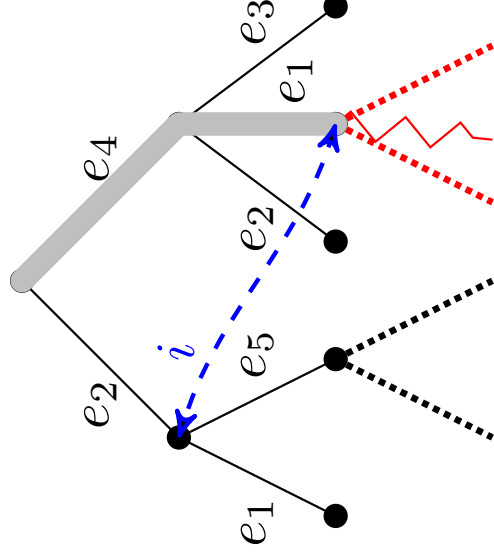
$h \sim_i h'$ means *given the events i has observed, h and h' are indistinguishable*



Forward-Looking Uncertainty

Given two maximal histories H and H' ,

agent i may be uncertain which of the two will be the final outcome.



Questions

1. What are the properties of “forward-looking” uncertainty? (eg., it is not symmetric)

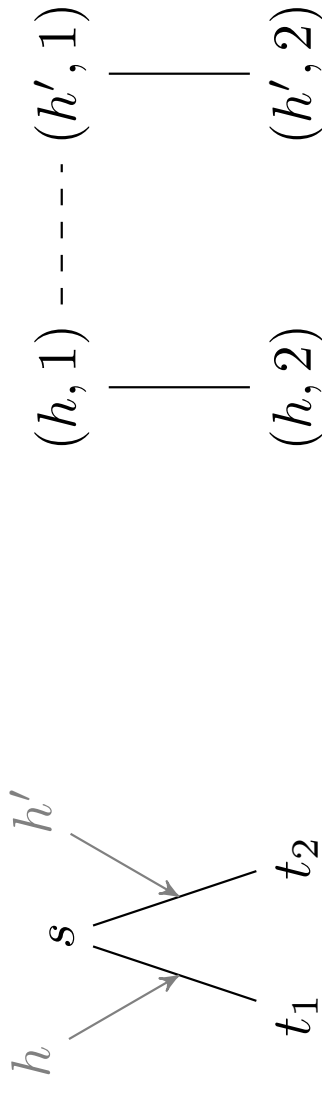


Questions

1. What are the properties of “forward-looking” uncertainty? (eg., it is not symmetric)
 2. What is the relationship between the two notions of uncertainty?
-

Questions

1. What are the properties of “forward-looking” uncertainty? (eg., it is not symmetric)
2. What is the relationship between the two notions of uncertainty?



Plan

- Introduction and Motivating Example
 - Formalizing the Example
 - Reasoning *inside* a protocol
 - Issues
 - What is a protocol?
 - Observational uncertainty vs. future uncertainty.
 - What do the agents know about the protocol?
 - Concluding Remarks
-

Uncertainty about the Protocol

Agents use the protocol to interpret events, but what if the agents are uncertain about which protocol they are in?

Uncertainty about the Protocol

Agents use the protocol to interpret events, but what if the agents are uncertain about which protocol they are in?

If the protocol *is* the model, then this is the same question as “what do the agents know about the model we (the modelers) are using to represent them”, and this is a separate issue.

Uncertainty about the Protocol

Agents use the protocol to interpret events, but what if the agents are uncertain about which protocol they are in?

If the protocol *is* the model, then this is the same question as “what do the agents know about the model we (the modelers) are using to represent them”, and this is a separate issue.

Can uncertainty about the protocol be represented by observational uncertainty?

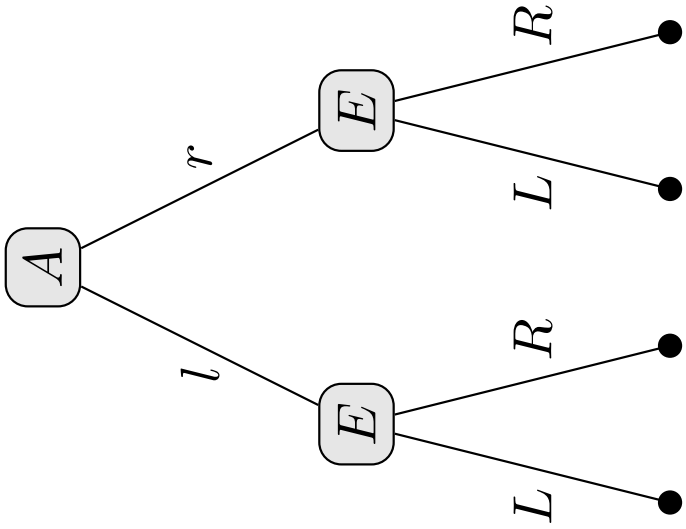
Uncertainty about the Protocol

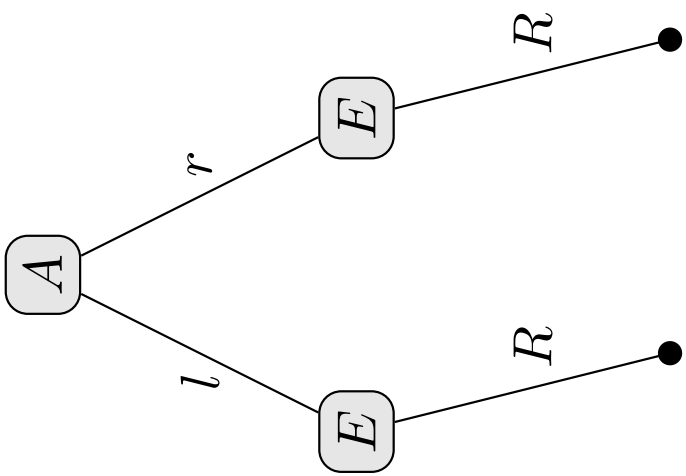
Agents use the protocol to interpret events, but what if the agents are uncertain about which protocol they are in?

If the protocol *is* the model, then this is the same question as “what do the agents know about the model we (the modelers) are using to represent them”, and this is a separate issue.

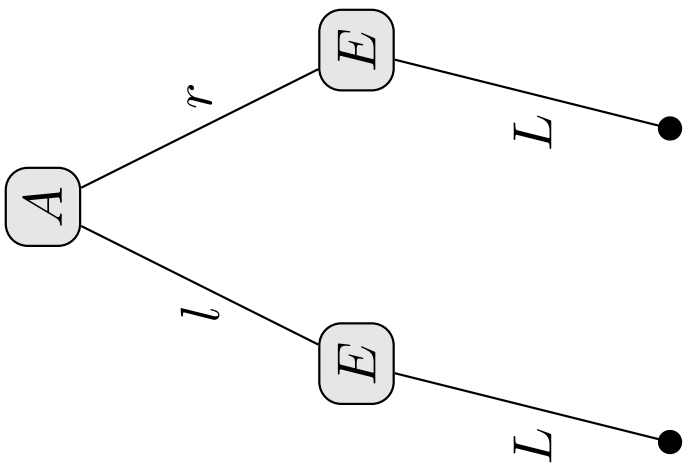
Can uncertainty about the protocol be represented by observational uncertainty?

Example: Bob is either in a protocol where Ann is a truth teller or a protocol where Ann is a liar, if Ann tells Bob something that Bob knows to be true, then Bob knows he is in the first protocol.





$\sim A$



Plan

- Introduction and Motivating Example
 - Formalizing the Example
 - Reasoning *inside* a protocol
 - Issues
 - What is a protocol?
 - Observational uncertainty vs. future uncertainty.
 - What do the agents know about the protocol?
 - Concluding Remarks
-

Thank You!
