A Justification Logic with Common Knowledge

Samuel Bucheli

Institute of Computer Science and Applied Mathematics
University of Bern

Explicit Paradigms in Logic and Computer Science
June 5, 2012

Justification Logic (Artemov 95)

modal logic	**justification logic**
$\Box A$	$[t]A$
A is known	t is evidence for A
A is provable	t is a proof for A

S4 and LP

S4 **axioms & rules**

$\Box(A \to B) \to (\Box A \to \Box B)$
$\Box A \to A$
$\Box A \to \Box\Box A$

modus ponens

$$\frac{A}{\Box A}$$

LP **axioms & rules**

$[t](A \to B) \to ([s]A \to [t \cdot s]B)$
$[t]A \to A$
$[t]A \to [!t][t]A$
$[t]A \vee [s]A \to [t + s]A$

modus ponens

$$\frac{}{[c]A}$$

where c is a constant
and A an axiom

Common Knowledge

Common Knowledge of A

everybody knows A and
everybody knows that everybody knows A and
everybody knows that everybody knows that everybody knows A
and
...

Another perspective

Common knowledge of A is the greatest fixed point of

$$\lambda X.\text{everybody knows } A \text{ and everybody knows } X.$$

Common Knowledge

Common Knowledge of A

everybody knows A and
everybody knows that everybody knows A and
everybody knows that everybody knows that everybody knows A
and
...

Another perspective

Common knowledge of A is the greatest fixed point of

$$\lambda X.\text{everybody knows } A \text{ and everybody knows } X.$$

The system H_R (Fagin, Halpern, Moses, Vardi 95)

- All propositional tautologies
- $\Box_i(A \to B) \to (\Box_i A \to \Box_i B)$ (K)
- $\Box_i A \to A$ (T)
- $\Box_i A \to \Box_i \Box_i A$ (4)
- $CA \to E(A \wedge CA)$ (Co-CI)

and rules

$$\frac{A \quad A \to B}{B}\ (\text{MP}) \qquad \frac{A}{\Box_i A}\ (\text{Nec}) \qquad \frac{B \to E(A \wedge B)}{B \to CA}\ (\text{I-R})$$

The system H_{Ax} (Meyer, van der Hoek 95)

- All propositional tautologies
- $\Box_i(A \to B) \to (\Box_i A \to \Box_i B)$ (K)
- $\Box_i A \to A$ (T)
- $\Box_i A \to \Box_i \Box_i A$ (4)
- $CA \to E(A \wedge CA)$ (Co-Cl)
- $C(A \to EA) \to (EA \to CA)$ (I-Ax)

and rules

$$\frac{A \quad A \to B}{B} \text{ (MP)} \qquad \frac{A}{\Box_i A} \text{ (Nec)} \qquad \frac{B \to E(A \wedge B)}{B \to CA} \text{ (I-R)}$$

The system H_{Ax} (Meyer, van der Hoek 95)

- All propositional tautologies
- $\Box_i(A \to B) \to (\Box_i A \to \Box_i B)$ (K)
- $\Box_i A \to A$ (T)
- $\Box_i A \to \Box_i \Box_i A$ (4)
- $CA \to E(A \wedge CA)$ (Co-Cl)
- $C(A \to EA) \to (EA \to CA)$ (I-Ax)
- $C(A \to B) \to (CA \to CB)$ (C-K)

and rules

$$\frac{A \quad A \to B}{B} \text{ (MP)} \qquad \frac{A}{\Box_i A} \text{ (Nec)} \qquad \frac{B \to E(A \wedge B)}{B \to CA} \text{ (I-R)}$$

The system H_{Ax} (Meyer, van der Hoek 95)

- All propositional tautologies
- $\Box_i(A \to B) \to (\Box_i A \to \Box_i B)$ (K)
- $\Box_i A \to A$ (T)
- $\Box_i A \to \Box_i \Box_i A$ (4)
- $CA \to E(A \wedge CA)$ (Co-Cl)
- $C(A \to EA) \to (EA \to CA)$ (I-Ax)
- $C(A \to B) \to (CA \to CB)$ (C-K)

and rules

$$\frac{A \quad A \to B}{B} \text{ (MP)} \qquad \frac{A}{\Box_i A} \text{ (Nec)} \qquad \frac{B \to E(A \wedge B)}{B \to CA} \text{ (I-R)}$$

$$\frac{A}{CA} \text{ (C-Nec)}$$

A Justification Logic for Common Knowledge: General Idea

- take a copy of a justification logic for each agent
- add evidence terms for mutual and common knowledge
- operation to tuple individual evidence
- operation to access tuple
- operation to generate "infinite list"
- operations to split list into head and tail

A Justification Logic for Common Knowledge: General Idea

- take a copy of a justification logic for each agent
- add evidence terms for mutual and common knowledge
- operation to tuple individual evidence
- operation to access tuple
- operation to generate "infinite list"
- operations to split list into head and tail

A Justification Logic for Common Knowledge: General Idea

- take a copy of a justification logic for each agent
- add evidence terms for mutual and common knowledge
- operation to tuple individual evidence

$$\Box_1 A \wedge \cdots \wedge \Box_h A \rightarrow \mathsf{E}A$$

- operation to access tuple

$$\mathsf{E}A \rightarrow \Box_i A$$

- operation to generate "infinite list"
- operations to split list into head and tail

A Justification Logic for Common Knowledge: General Idea

- take a copy of a justification logic for each agent
- add evidence terms for mutual and common knowledge
- operation to tuple individual evidence

$$[t_1]_1 A \wedge \cdots \wedge [t_h]_h A \rightarrow [\langle t_1, \ldots, t_h \rangle]_{\mathsf{E}} A$$

- operation to access tuple

$$\mathsf{E} A \rightarrow \Box_i A$$

- operation to generate "infinite list"
- operations to split list into head and tail

A Justification Logic for Common Knowledge: General Idea

- take a copy of a justification logic for each agent
- add evidence terms for mutual and common knowledge
- operation to tuple individual evidence

$$[t_1]_1 A \wedge \cdots \wedge [t_h]_h A \rightarrow [\langle t_1, \ldots, t_h \rangle]_{\mathsf{E}} A$$

- operation to access tuple

$$[t]_{\mathsf{E}} A \rightarrow [\pi_i t]_i A$$

- operation to generate "infinite list"
- operations to split list into head and tail

A Justification Logic for Common Knowledge: General Idea

- take a copy of a justification logic for each agent
- add evidence terms for mutual and common knowledge
- operation to tuple individual evidence

$$[t_1]_1 A \wedge \cdots \wedge [t_h]_h A \rightarrow [\langle t_1, \ldots, t_h \rangle]_{\mathsf{E}} A$$

- operation to access tuple

$$[t]_{\mathsf{E}} A \rightarrow [\pi_i t]_i A$$

- operation to generate "infinite list"

$$\mathsf{C}(A \rightarrow \mathsf{E}A) \rightarrow (\mathsf{E}A \rightarrow \mathsf{C}A)$$

- operations to split list into head and tail

$$\mathsf{C}A \rightarrow \mathsf{E}A \ , \qquad \mathsf{C}A \rightarrow \mathsf{EC}A$$

A Justification Logic for Common Knowledge: General Idea

- take a copy of a justification logic for each agent
- add evidence terms for mutual and common knowledge
- operation to tuple individual evidence

$$[t_1]_1 A \wedge \cdots \wedge [t_h]_h A \rightarrow [\langle t_1, \ldots, t_h \rangle]_{\mathsf{E}} A$$

- operation to access tuple

$$[t]_{\mathsf{E}} A \rightarrow [\pi_i t]_i A$$

- operation to generate "infinite list"

$$[t]_{\mathsf{C}}(A \rightarrow [r]_{\mathsf{E}} A) \rightarrow ([s]_{\mathsf{E}} A \rightarrow [\mathsf{ind}(t, s)]_{\mathsf{C}} A)$$

- operations to split list into head and tail

$$\mathsf{C}A \rightarrow \mathsf{E}A \ , \qquad \mathsf{C}A \rightarrow \mathsf{E}\mathsf{C}A$$

A Justification Logic for Common Knowledge: General Idea

- take a copy of a justification logic for each agent
- add evidence terms for mutual and common knowledge
- operation to tuple individual evidence

$$[t_1]_1 A \wedge \cdots \wedge [t_h]_h A \rightarrow [\langle t_1, \ldots, t_h \rangle]_{\mathsf{E}} A$$

- operation to access tuple

$$[t]_{\mathsf{E}} A \rightarrow [\pi_i t]_i A$$

- operation to generate "infinite list"

$$[t]_{\mathsf{C}}(A \rightarrow [r]_{\mathsf{E}} A) \rightarrow ([s]_{\mathsf{E}} A \rightarrow [\mathsf{ind}(t, s)]_{\mathsf{C}} A)$$

- operations to split list into head and tail

$$[t]_{\mathsf{C}} A \rightarrow [\mathsf{head}(t)]_{\mathsf{E}} A \ , \qquad [t]_{\mathsf{C}} A \rightarrow [\mathsf{tail}(t)]_{\mathsf{E}} [t]_{\mathsf{C}} A$$

Language

Let $\circledast \in \{1, \ldots, h, \mathsf{E}, \mathsf{C}\}$

$\mathsf{Cons}_{\circledast}$ a countable set of proof constants for $\circledast$

$\mathsf{Var}_{\circledast}$ a countable set of proof variables for $\circledast$

Prop a countable set of propositional variables.

Operations on Terms

- application and sum (for agents and C)
- positive inspection (for agents)
- tupling and projection (relates individual and mutual knowledge)
- induction and co-closure (relates mutual and common knowledge)

Formulae

$A ::= P_j \mid \bot \mid A \to A \mid [t]_{\circledast} A$ where $t \in \mathsf{Tm}_{\circledast}$

$\mathrm{LP}_h^{\mathsf{C}}$: A Justification Logic with Common Knowledge

Axioms

1. all propositional tautologies
2. $[t]_*(A \to B) \to ([s]_* A \to [t \cdot s]_* B)$ (application)
3. $[t]_* A \vee [s]_* A \to [t + s]_* A$ (sum)
4. $[t]_i A \to A$ (reflexivity)
5. $[t]_i A \to [!t]_i\, [t]_i A$ (inspection)
6. $[t_1]_1 A \wedge \cdots \wedge [t_h]_h A \to [\langle t_1, \ldots, t_h \rangle]_{\mathsf{E}} A$ (tupling)
7. $[t]_{\mathsf{E}} A \to [\pi_i t]_i A$ (projection)
8. $[t]_{\mathsf{C}} A \to [\mathrm{head}(t)]_{\mathsf{E}} A, \quad [t]_{\mathsf{C}} A \to [\mathrm{tail}(t)]_{\mathsf{E}}\, [t]_{\mathsf{C}} A$ (co-closure)
9. $[t]_{\mathsf{C}}(A \to [r]_{\mathsf{E}} A) \to ([s]_{\mathsf{E}} A \to [\mathrm{ind}(t, s)]_{\mathsf{C}} A)$ (induction)

where $* \in \{1, \ldots, h, \mathsf{C}\}$.

Rules *modus ponens* and *axiom necessitation*

$$\frac{A \quad A \to B}{B}\ (\mathrm{MP})\ , \qquad \frac{}{[c]_{\mathsf{C}} D}\ (\mathrm{AN})\ ,$$

where $c \in \mathrm{Cons}_{\mathsf{C}}$ and D is an axiom from the list above.

Constructive Necessitation and Induction Rule

Theorem

If $\vdash A$, *then, for each* $\circledast$ *there exists a ground term* $t_{\circledast}$ *such that*

$$\vdash [t]_{\circledast} A \ .$$

Corollary

If $\vdash B \to [s]_{\mathsf{E}}(A \wedge B)$, *then there exists a ground term* t *such that*

$$\vdash B \to [t]_{\mathsf{C}} A \ .$$

Definable Operations

Lemma

The following operations can be defined

- $[t]_{\mathsf{E}}(A \to B) \to ([s]_{\mathsf{E}} \to [t \cdot_{\mathsf{E}} s]_{\mathsf{E}} B)$,
- $[t]_{\mathsf{E}} A \vee [s]_{\mathsf{E}} A \to [t +_{\mathsf{E}} s]_{\mathsf{E}} A$,
- $[t]_{\mathsf{C}} A \to [\downarrow_i]_i A$,
- $[t]_{\mathsf{C}} A \to [\uparrow_i]_i [t]_{\mathsf{C}} A$,
- $[t]_{\mathsf{C}} A \to [!_{\mathsf{C}} t]_{\mathsf{C}} [t]_{\mathsf{C}} A$,
- $[t]_{\mathsf{C}} A \to [\Leftarrow t]_{\mathsf{C}} [\mathsf{head}(t)]_{\mathsf{E}} A$.

Epistemic Models

An epistemic model is a quadruple $\mathcal{M} = (W, R, \mathcal{E}, \nu)$, where (W, R, ν) is a Kripke model, i.e.,

- W is a non-empty set called possible worlds.
- $R\colon \{1, \ldots, h\} \to \mathcal{P}(W \times W)$ assigns a transitive and reflexive accessibility relation to each agent.
- $\nu\colon \mathsf{Prop} \to \mathcal{P}(W)$ is a truth valuation.

and

- $\mathcal{E}\colon W \times \mathsf{Tm} \to \mathcal{P}\left(\mathsf{Fm}_{\mathsf{LP}_h^{\mathsf{C}}}\right)$ is an evidence function determining the formulae evidenced by a term at a world.

Furthermore

- $R_{\mathsf{E}} := \bigcup_{i=1}^{h} R_i$,
- $R_{\mathsf{C}} := \bigcup_{n=1}^{\infty} (R_{\mathsf{E}})^n$, i.e. R_{C} is the transitive closure of R_{E}.

Satisfaction Relation

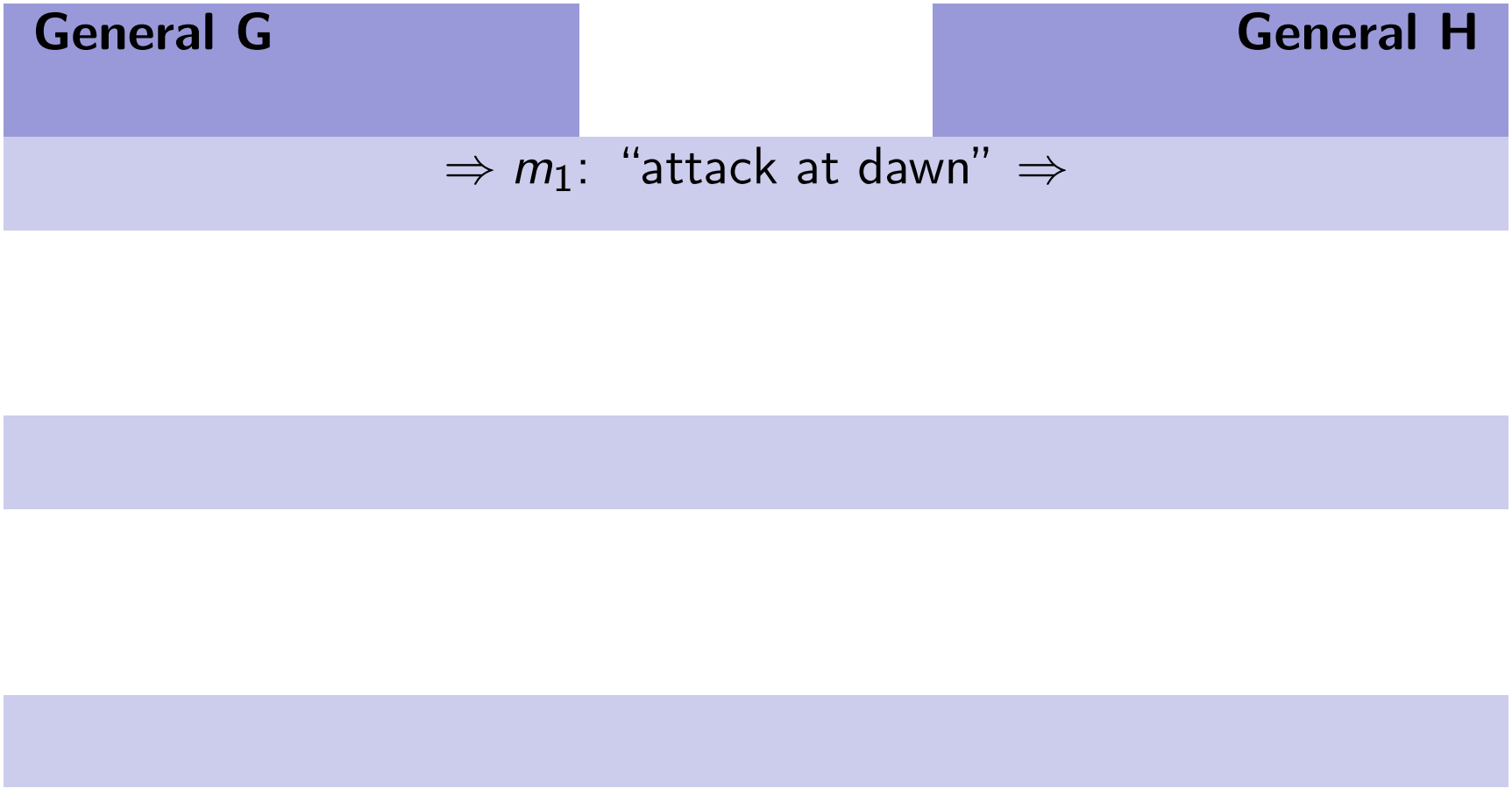
- $\mathcal{M}, w \Vdash P$ if and only if $w \in \nu(P)$;
- $\Vdash$ behaves classically with respect to propositional connectives;
- $\mathcal{M}, w \Vdash [t]_{\circledast} A$ if and only if
  1) $A \in \mathcal{E}_{\circledast}(w, t)$ and
  2) $\mathcal{M}, v \Vdash A$ for all $v \in W$ with $(w, v) \in R_{\circledast}$.

$[t]A$ if and only if

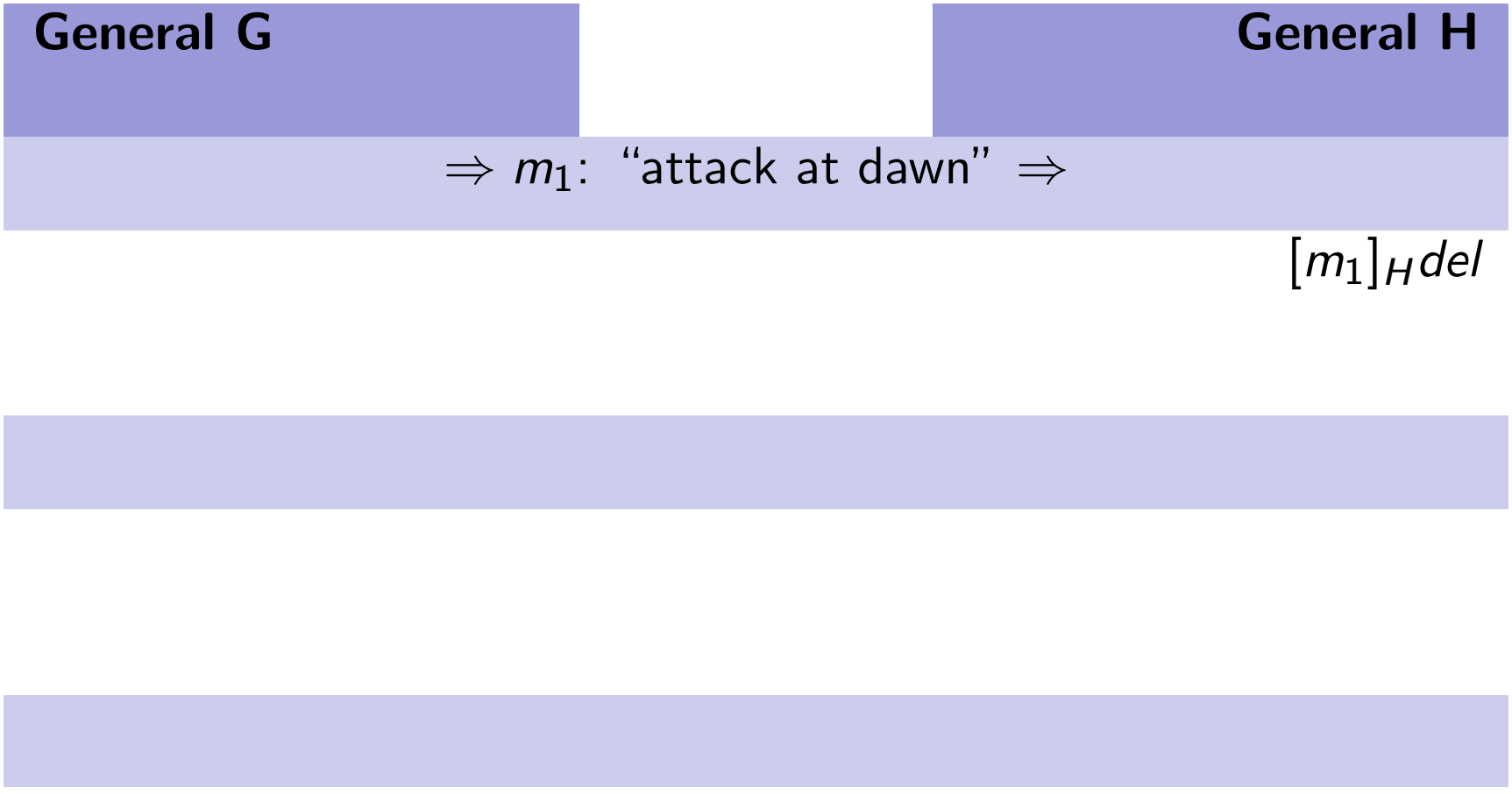
A is true in all possible worlds and t is admissible evidence for A.

Coordinated Attack

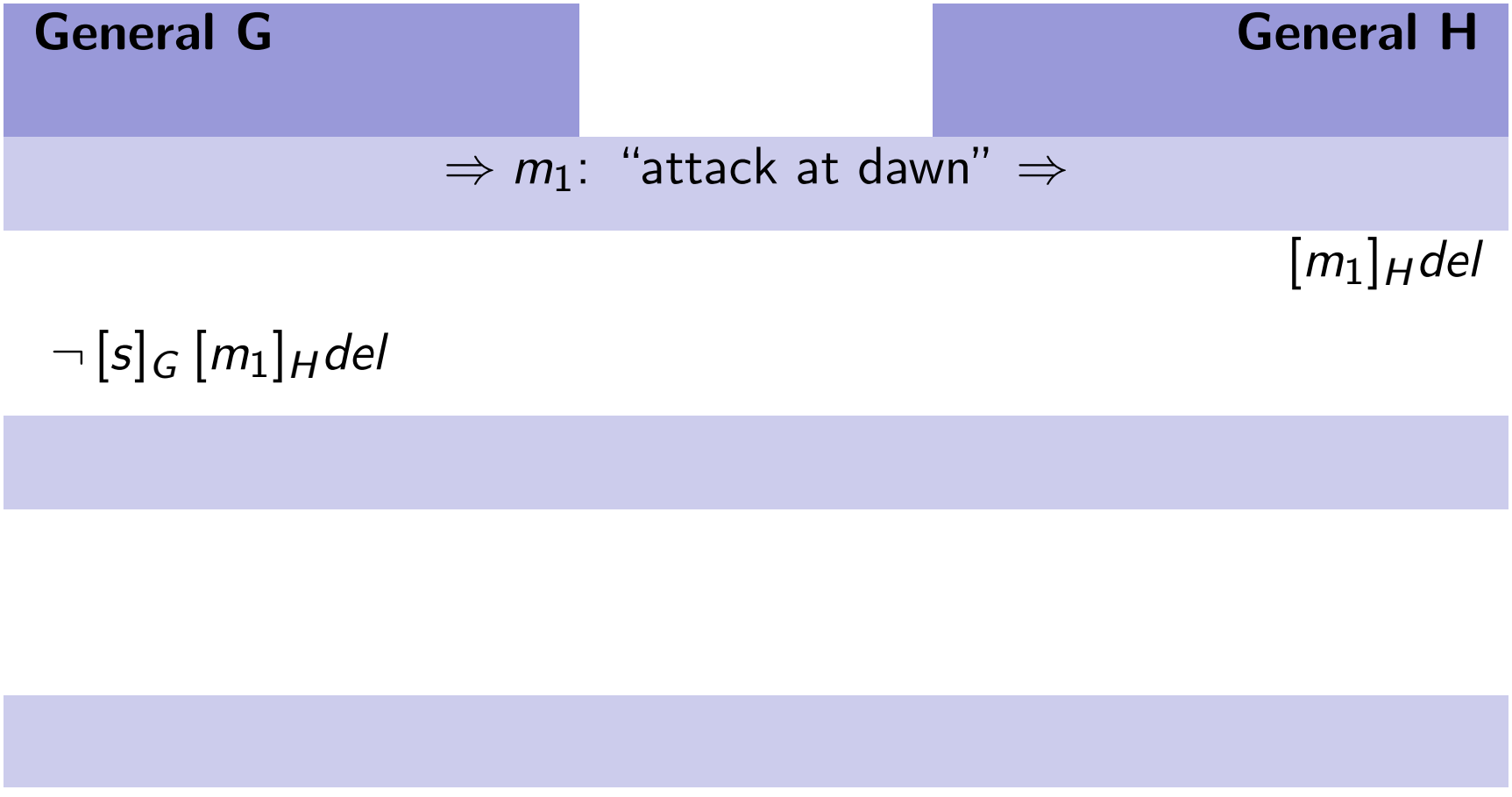
Coordinated Attack



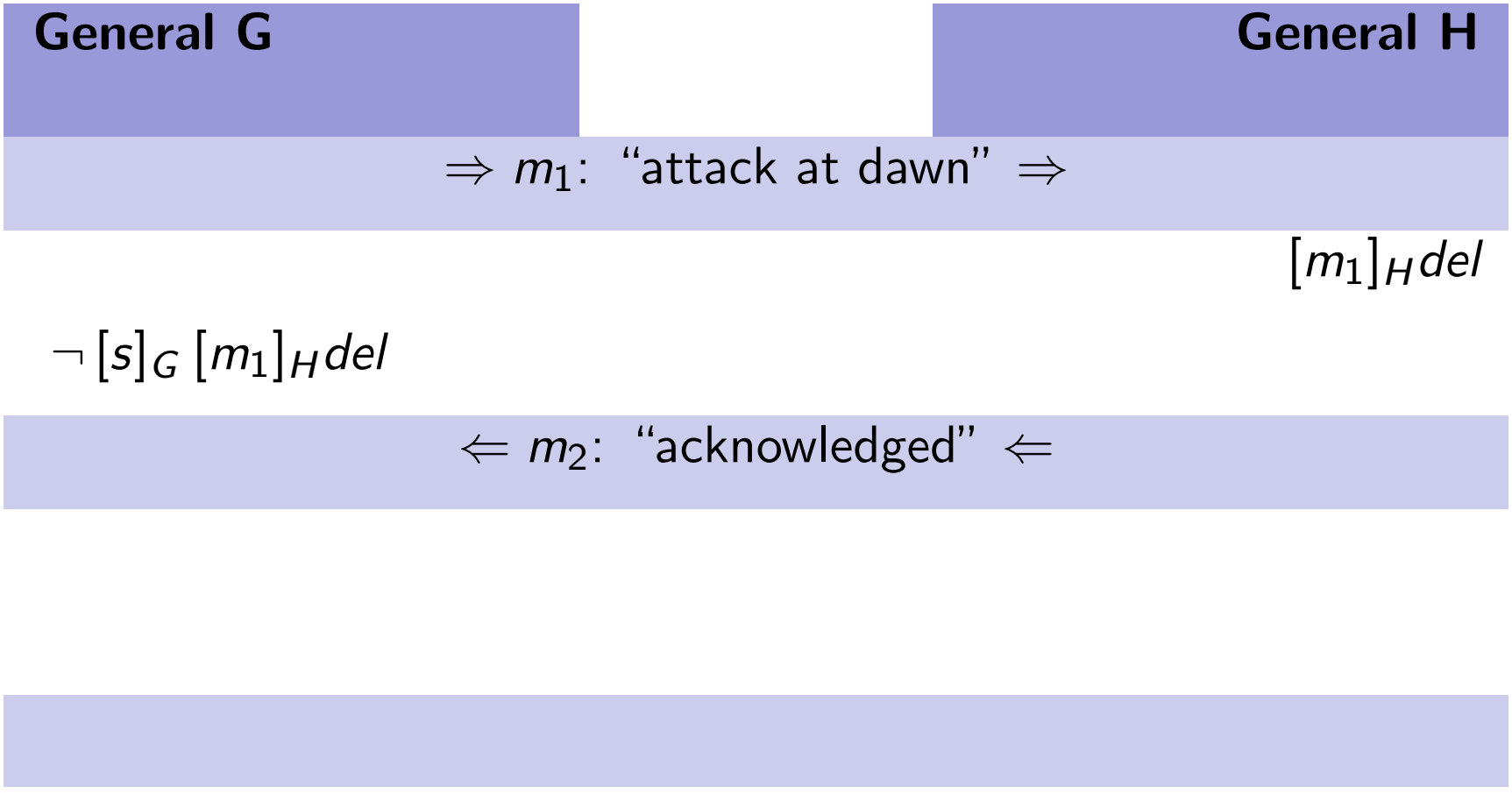
Coordinated Attack



Coordinated Attack



Coordinated Attack



Coordinated Attack

General G | **General H**

$\Rightarrow m_1$: "attack at dawn" $\Rightarrow$

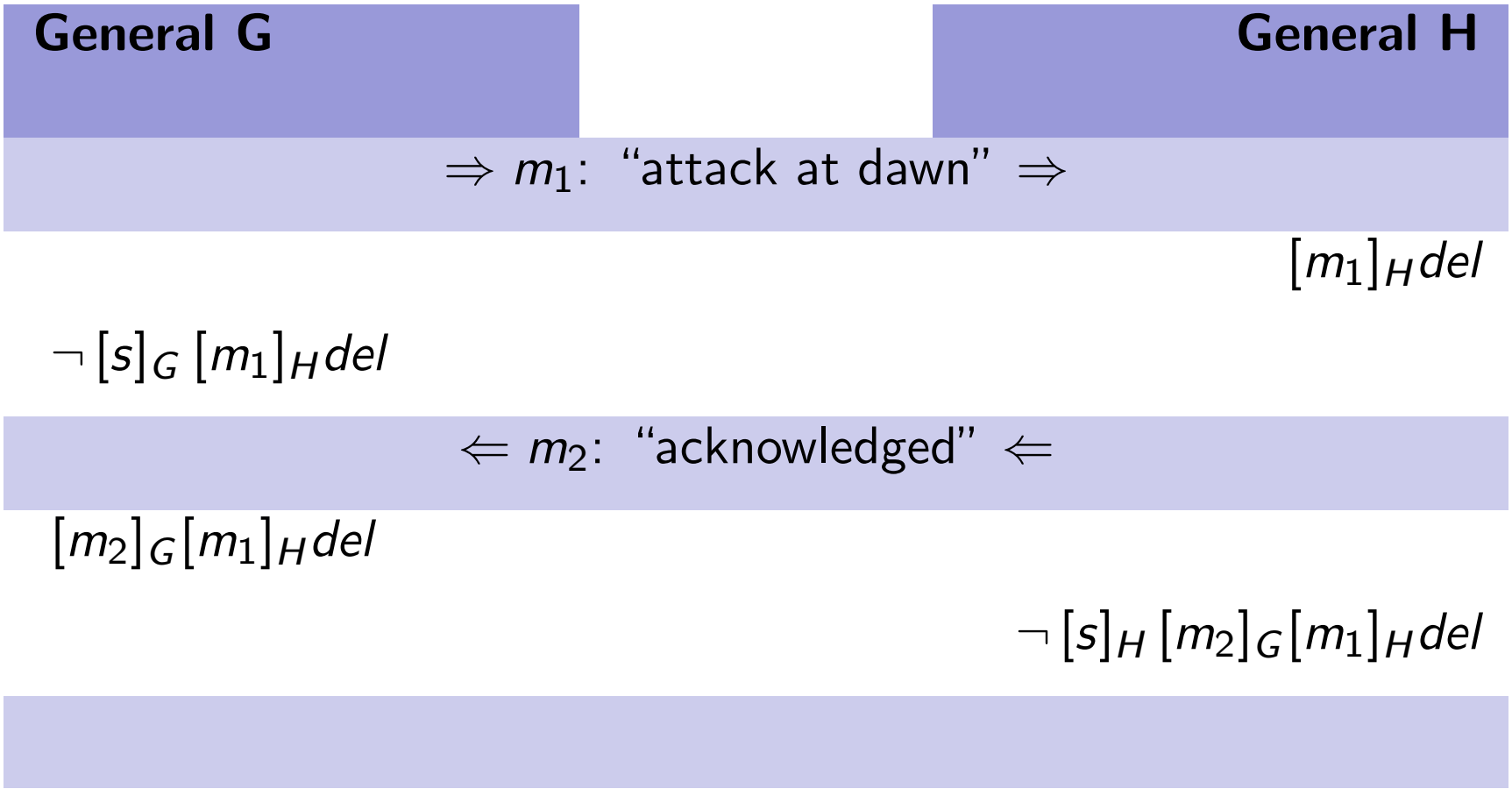
$[m_1]_H del$

$\neg [s]_G [m_1]_H del$

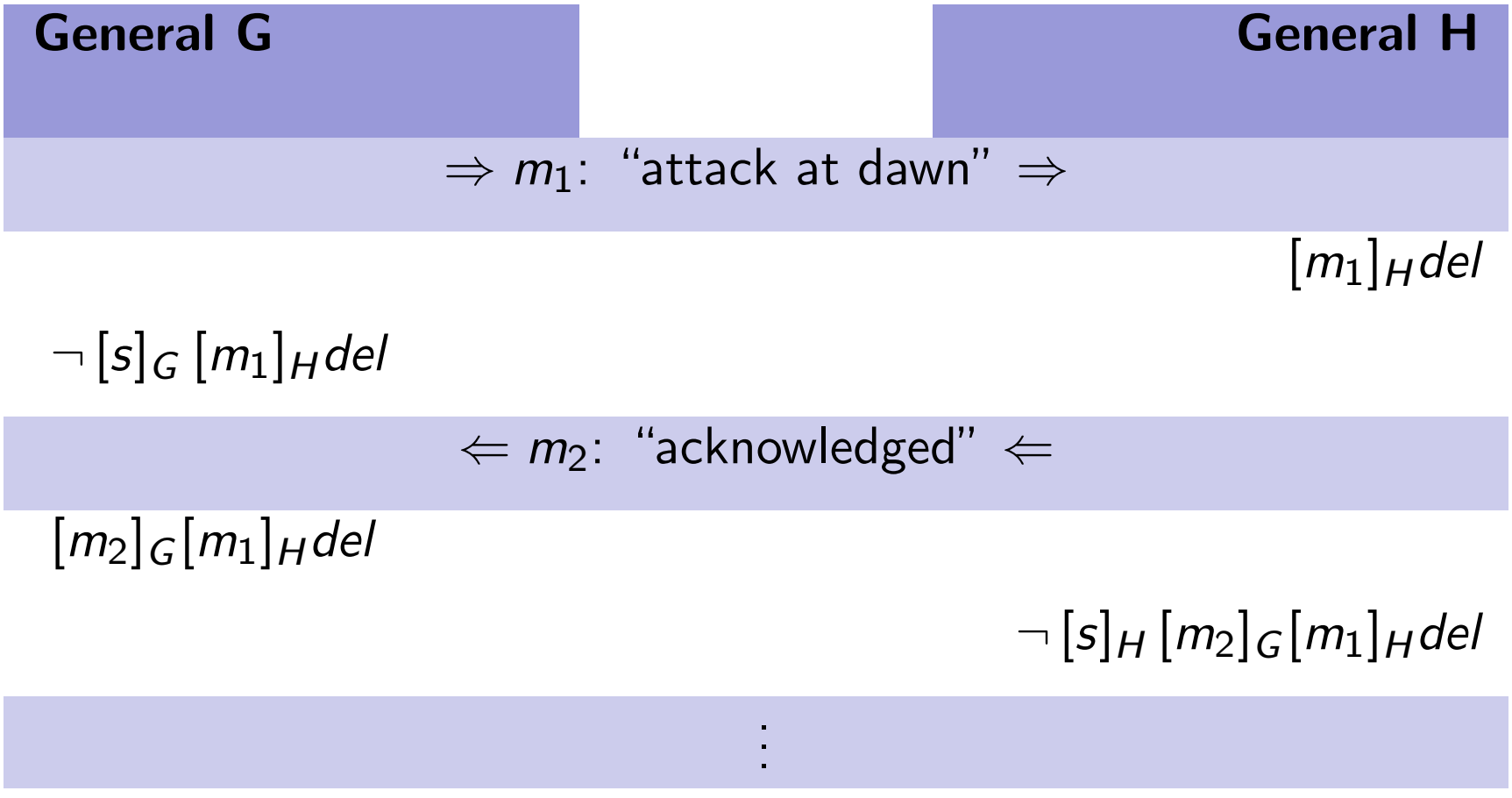
$\Leftarrow m_2$: "acknowledged" $\Leftarrow$

$[m_2]_G [m_1]_H del$

Coordinated Attack



Coordinated Attack



Soundness, Completeness and Decidability

Theorem

$\mathsf{LP}_h^{\mathsf{C}}$ *is sound and complete with respect to its class of epistemic models.*

Corollary

$\mathsf{LP}_h^{\mathsf{C}}$ *is complete with respect to the class of singleton models.*

Theorem

$\mathsf{LP}_h^{\mathsf{C}}$ *is decidable.*

Soundness, Completeness and Decidability

Theorem

$\mathsf{LP}_h^{\mathsf{C}}$ *is sound and complete with respect to its class of epistemic models.*

Corollary

$\mathsf{LP}_h^{\mathsf{C}}$ *is complete with respect to the class of singleton models.*

Theorem

$\mathsf{LP}_h^{\mathsf{C}}$ *is decidable.*

Soundness, Completeness and Decidability

Theorem

$\mathsf{LP}_h^{\mathsf{C}}$ *is sound and complete with respect to its class of epistemic models.*

Corollary

$\mathsf{LP}_h^{\mathsf{C}}$ *is complete with respect to the class of singleton models.*

Theorem

$\mathsf{LP}_h^{\mathsf{C}}$ *is decidable.*

Conservativity

Theorem

$\mathsf{LP}_h^{\mathsf{C}}$ *is a conservative extensions of its corresponding multi-agent justification logic (without common knowledge).*

Corollary

$\mathsf{LP}_2^{\mathsf{C}}$ *is a conservative extension of Yavorskaya's minimal bimodal explicit evidence logic* LP_2.

Realization

Lemma
If $\mathsf{LP}_h^{\mathsf{C}} \vdash A$, *then* $\mathsf{S4} \vdash A^\circ$.

Question

What about the other direction?

Realization

Lemma
If $\mathsf{LP}_h^{\mathsf{C}} \vdash A$, *then* $\mathsf{S4} \vdash A^\circ$.

Question

What about the other direction?

Fitting's Semantical Realization Technique

Observation

Fitting's technique relies only on a certain form of the canonical model, the Truth Lemma, and Constructive Necessitation.

Idea

If all possible realizations of a formula fail in a world of the canonical model, then also the forgetful projection of this formula must fail in the same world.

Fitting's Semantical Realization Technique

Observation

Fitting's technique relies only on a certain form of the canonical model, the Truth Lemma, and Constructive Necessitation.

Idea

If all possible realizations of a formula fail in a world of the canonical model, then also the forgetful projection of this formula must fail in the same world.

Conclusions

- common knowledge is a natural addition to multi-agent justification logics
- alternative possibilities in analyzing epistemic scenarios
- major open problem: realization
- another interesting question: complexity
- long-term goal: dynamic epistemic logics
- more model theoretic tools
- justification logics without corresponding (normal) modal logics?