Stit Semantics II: Knowledge and Action Types

John Horty

Introduction

1. Lindström and Segerberg: stit semantics is

   a logic of action without actions

   No author in the Anselm-Kanger-Chellas line up through Belnap has countenanced the existence of actions in logic: action talk, yes; ontology of actions, no

2. This is a bit too strong: there are already action tokens

3. Goal today: introduce knowledge and action types, leading to *labeled stit semantics*

4. Motivate by need for epistemic readings of ability and oughts

5. Outline:

Ability: causal vs epistemic

Simple combinations with knowledge fail

Labeled stit semantics

Oughts: causal vs epistemic

Review/simplification of causal oughts

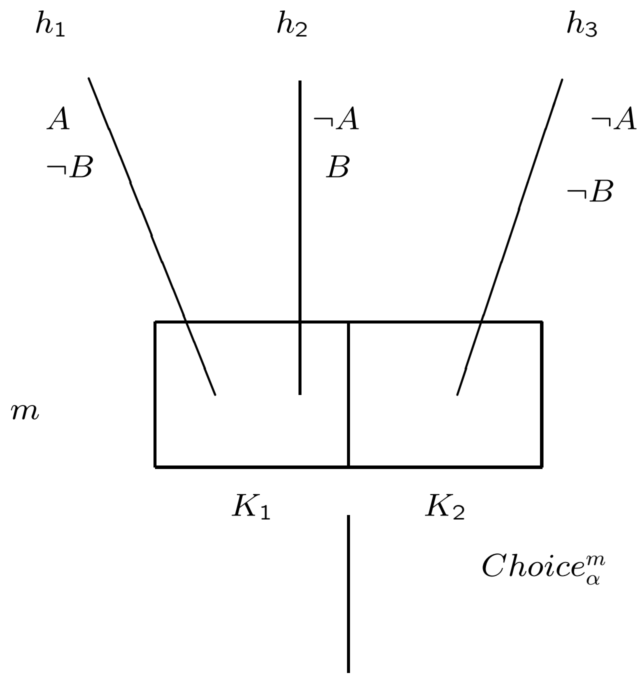
Simple combinations with knowledge fail

Ordering action types, epistemic oughts

Information sensitivity

Conditional oughts

Ability



1. Proposal:

$$\Diamond[\alpha\ cstit\text{: } A] \quad = \quad \text{It is possible that } \alpha \text{ sees to it that } A$$

$$= \quad \alpha \text{ can (has ability) see to it that } A$$

2. Validates neither

$$A \supset \Diamond[\alpha\ cstit\text{: } A]$$

$$\Diamond[\alpha\ cstit\text{: } A \vee B] \supset (\Diamond[\alpha\ cstit\text{: } A] \vee \Diamond[\alpha\ cstit\text{: } B])$$

3. "Causal" vs "epistemic" notions of ability

So let's add knowledge . . .

4. Indistinguishability relation, so that

$$m/h \sim_\alpha m'/h'$$

means that α cannot distinguish m/h from m'/h'

5. Epistemic stit models:

$$\langle \mathit{Tree}, <, \mathit{Agent}, \mathit{Choice}, \{\sim_\alpha\}_{\alpha \in \mathit{Agent}}, v \rangle$$

6. Evaluation rule: knowledge operator

   - $m/h \models \mathsf{K}_\alpha A$ iff $m'/h' \models \mathsf{K}_\alpha A$ for all m'/h' such that $m/h \sim_\alpha m'/h'$

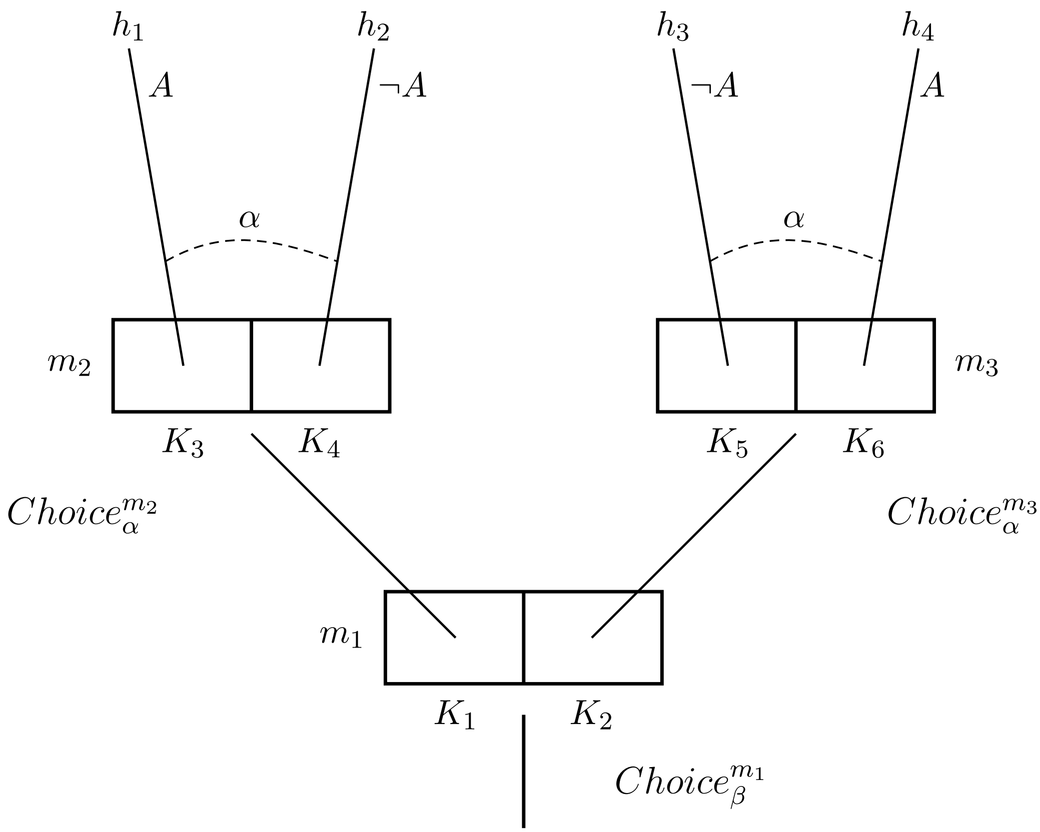
7. Natural suggestion: epistemic sense of ability should be represented by one of

$$\mathsf{K}_\alpha \Diamond [\alpha\ \mathit{stit}\text{: } A]$$

$$\Diamond \mathsf{K}_\alpha [\alpha\ \mathit{stit}\text{: } A]$$

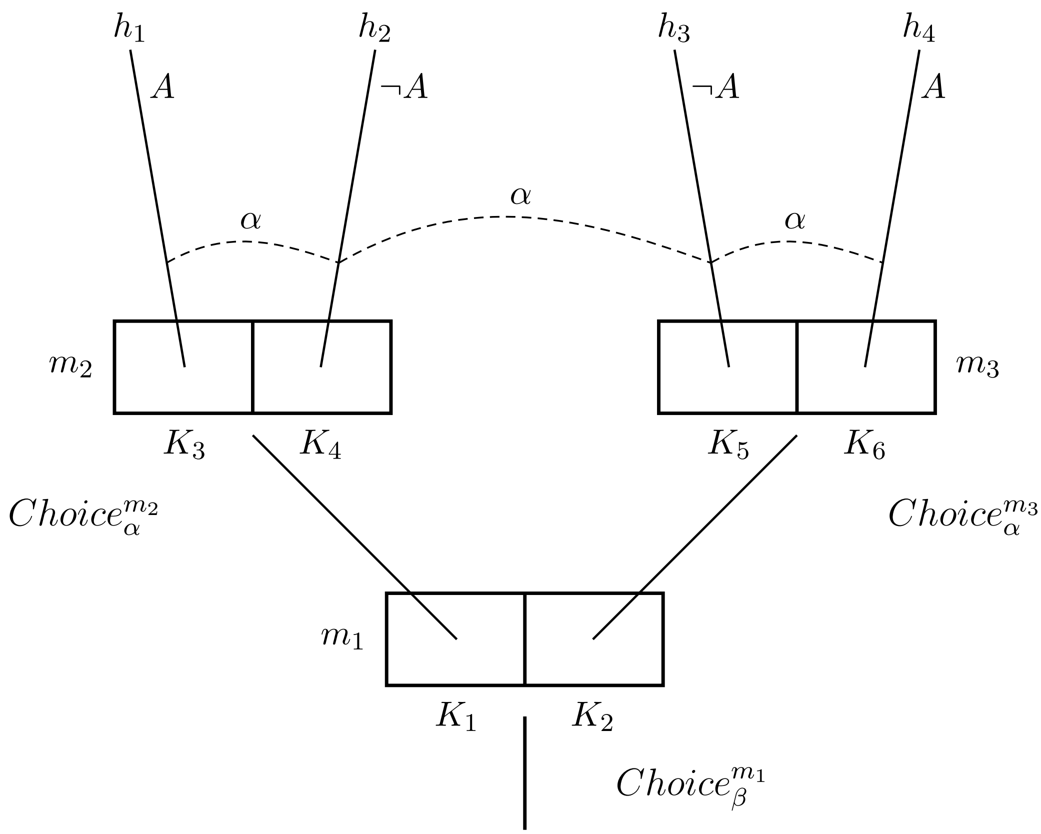
   But which one??

8. Answer: neither works, because can't distinguish this case . . .



$K_1 = \beta$ places coin heads
$K_2 = \beta$ places coin tails
$K_3 = \alpha$ bets heads
$K_4 = \alpha$ bets tails
$K_5 = \alpha$ bets heads
$K_6 = \alpha$ bets tails
$A = \alpha$ wins

Here, both $\mathsf{K}_\alpha\Diamond[\alpha\ stit\text{: } A]$ and $\Diamond\mathsf{K}_\alpha[\alpha\ stit\text{: } A]$ true from this case:



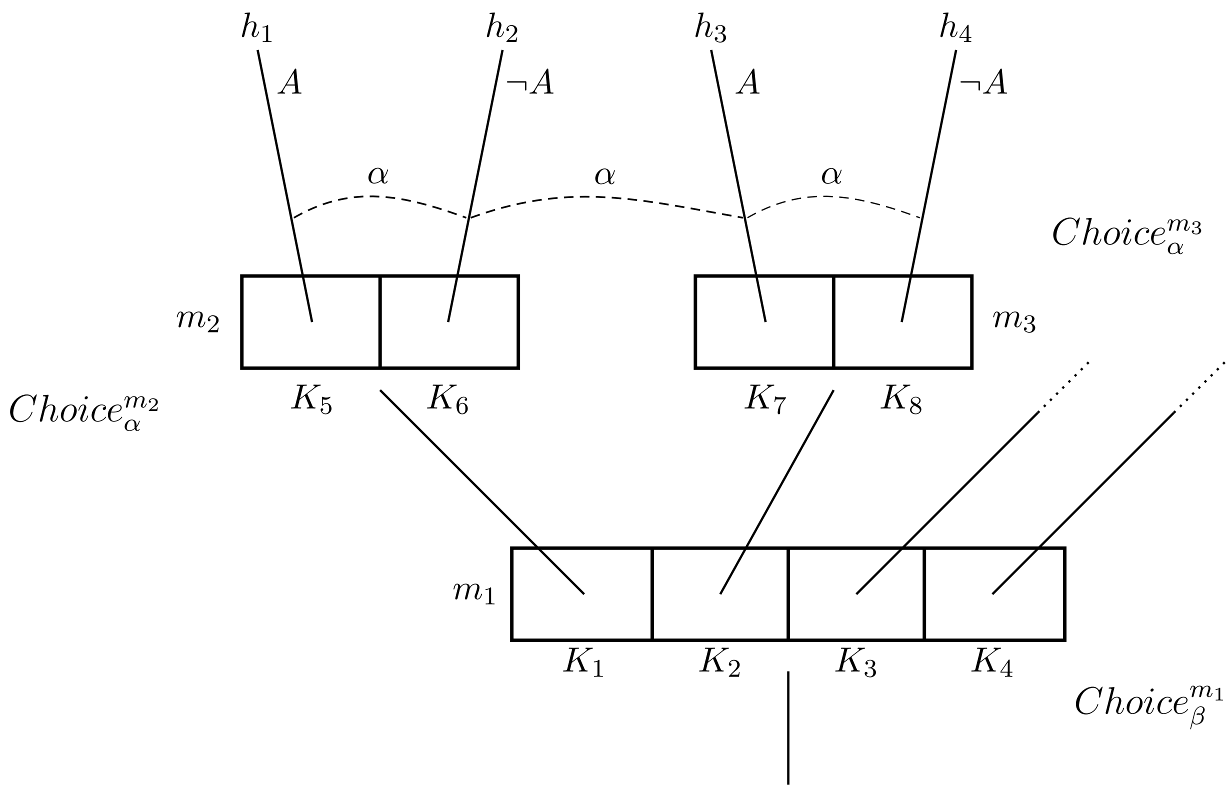
$K_1 = \beta$ places coin heads
$K_2 = \beta$ places coin tails
$K_3 = \alpha$ bets heads
$K_4 = \alpha$ bets tails
$K_5 = \alpha$ bets heads
$K_6 = \alpha$ bets tails
$A = \alpha$ wins

Here, both $\mathsf{K}_\alpha\Diamond[\alpha\ stit\colon A]$ and $\Diamond\mathsf{K}_\alpha[\alpha\ stit\colon A]$ false

9. Basic idea: $\Diamond[\alpha\ stit\colon A]$ means there is some action available to α that guarantees A

   What we want for epistemic sense is

   > there is some action available to α that which α knows to guarantee A



$K_1 = \beta$ places nickel heads up
$K_2 = \beta$ places dime heads up
$K_3 = \beta$ places nickel tails up
$K_4 = \beta$ places dime tails up
$K_5 = K_7$ α bets heads
$K_6 = K_8$ α bets tails

So this action has to be an action *type*

Labeled stit semantics

1. $Type = \{\tau_1, \tau_2, \ldots\}$ a set of *action types*

   Intuitions:

   Basic robot actions

   Agent performs a token by executing a type

   Types are repeatable

   Types (not tokens) lie within agent control

2. Partial *execution* function [ ] mapping τ to

   $$[\tau]^m_\alpha \in Choice^m_\alpha$$

   token resulting when τ is executed by α at m.

3. Label function $Label$ mapping $K \in Choice^m_\alpha$ to

   $$Label(K) \in Type$$

   This function is one-one

4. Execution/label constraints:

If $K \in Choice^m_\alpha$, then $[Label(K)]^m_\alpha = K$

If $\tau \in Type$ then $Label([\tau]^m_\alpha) = \tau$

(Note: requires $[\tau]^m_\alpha$ defined)

5. Action types available to α at moment m:

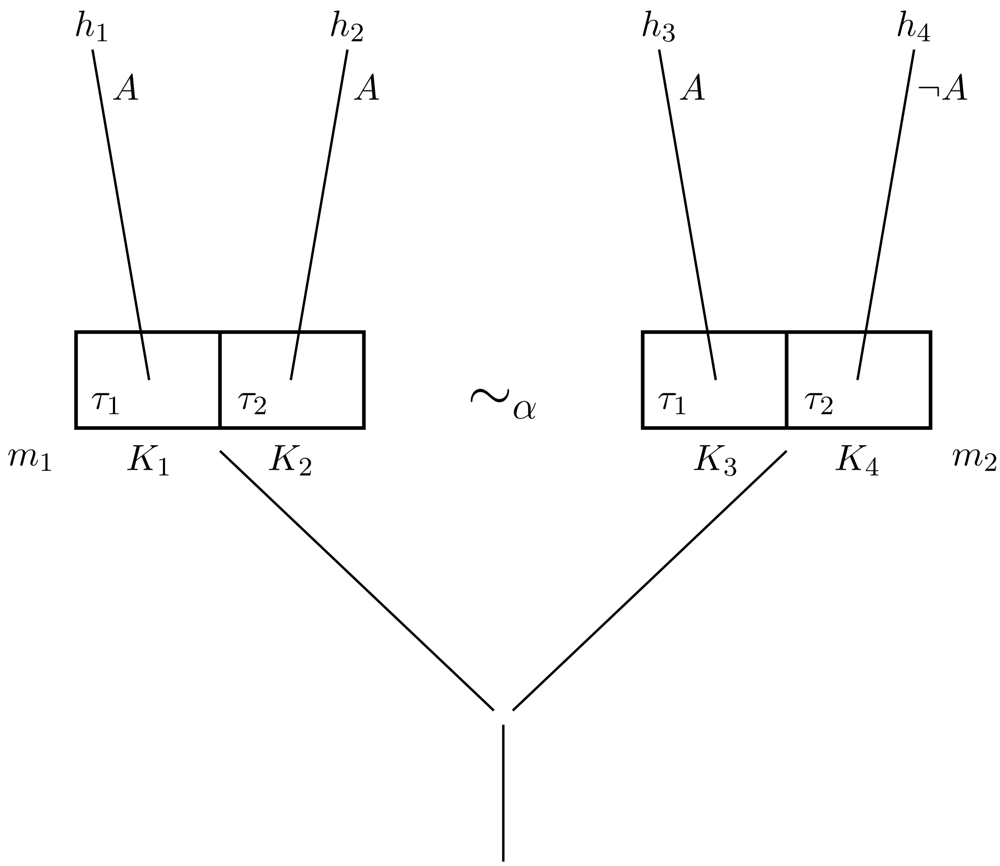
$Type^m_\alpha = \{Label(K) : K \in Choice^m_\alpha\}$

The action type executed by α at m/h is

$Type^m_\alpha(h) = Label(Choice^m_\alpha(h))$

6. Labeled stit model:

$\langle Tree, <, Agent, Choice, \{\sim_\alpha\}_{\alpha \in Agent}, Type, [\ ], Label, v\rangle$



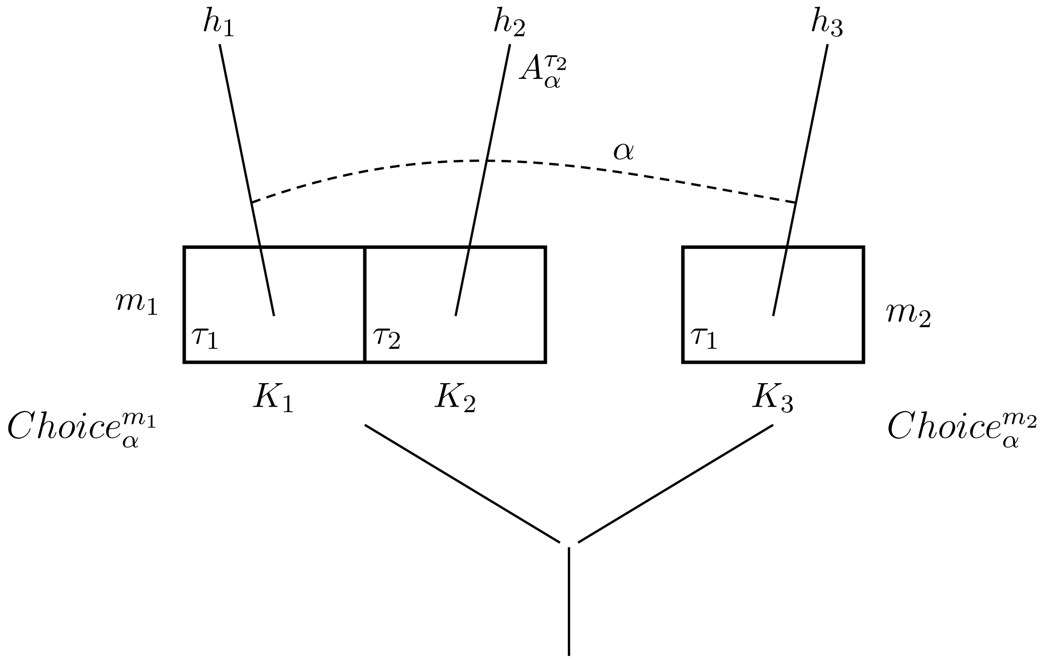
7. Evaluation rule: stit operator

- $m/h \models [\alpha\ kstit\text{: } A]$ iff $[Type^m_\alpha(h)]^{m'}_\alpha \subseteq |A|^{m'}_{\mathcal{M}}$ for all m'/h such that $m/h \sim_\alpha m'/h'$

Example:

$m_1/h_1 \models [\alpha\ kstit\text{: } A]$

$m_1/h_2 \models [\alpha\ stit\text{: } A]$, but $m_1/h_2 \not\models [\alpha\ kstit\text{: } A]$



8. This rule requires the constraint

(C1) If $m/h \sim_\alpha m'/h'$, then $Type_\alpha^m = Type_\alpha^{m'}$

Could make due with the very weird

(C2) If $m/h \sim_\alpha m'/h'$, then $[Type_\alpha^m(h)]_\alpha^{m'}$ is defined

but that's too weird. Given

- $\mathcal{M}, m/h \models A_\alpha^\tau$ iff $Type_\alpha^m(h) = \tau$.

what (C1) guarantees is the sensible

$$\Diamond A_\alpha^\tau \supset \mathsf{K}_\alpha \Diamond A_\alpha^\tau$$

9. In fact, we propose

(C4) If $m/h \sim_\alpha m'/h'$, then $m/h'' \sim_\alpha m'/h'''$ for all $h'' \in H^m$ and $h''' \in H^{m'}$

as most natural, with

$m \sim_\alpha m'$

defined as

$m/h \sim_\alpha m'/h'$ for h from H^m and h' from $H^{m'}$

10. This leads to simplified evaluation rule

    - $m/h \models [\alpha\ kstit\colon A]$ iff $[Type^m_\alpha(h)]^{m'}_\alpha \subseteq |A|^{m'}_{\mathcal{M}}$ for all m' such that $m \sim_\alpha m'$

11. Finally, return to ability:

$$\Diamond[\alpha\ stit\colon A]$$

is causal,

$$\Diamond[\alpha\ kstit\colon A]$$

is epistemic

12. Some notes on the $kstit$ logic:

S5 operator

Properly between $\mathsf{K}_\alpha + stit$ and $stit$:

$\mathsf{K}_\alpha[\alpha\ stit\colon A] \supset [\alpha\ kstit\colon A]$

$[\alpha\ kstit\colon A] \supset [\alpha\ stit\colon A]$

and converses fail

Collapses to $stit$ given

(C3) If $m/h \sim_\alpha m'/h'$ implies $m = m'$

Do you know what you're knowingly doing:

$[\alpha\ kstit\colon A] \supset \mathsf{K}_\alpha[\alpha\ kstit\colon A]$??

Ex ante and *ex interim* knowledge

$\mathsf{K}_\alpha A$ is *ex ante*

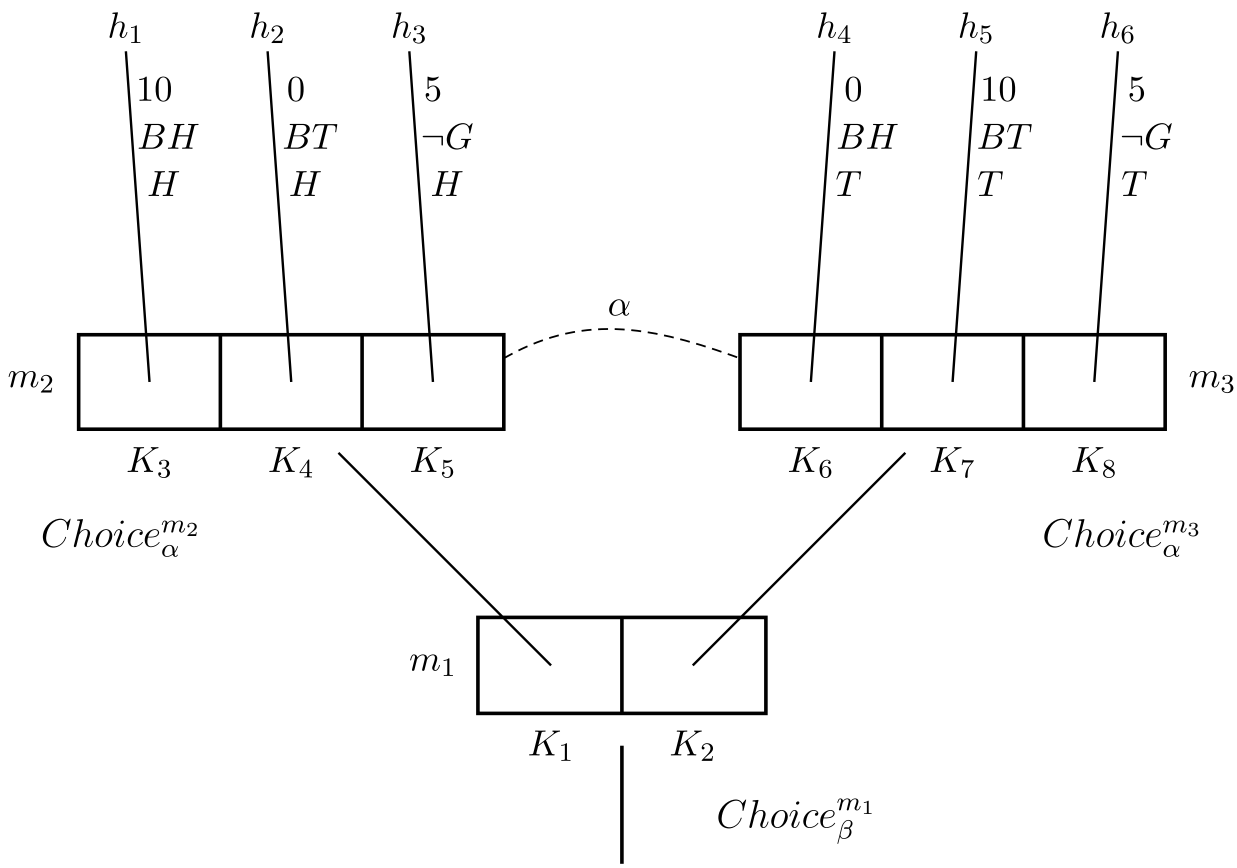
$[\alpha\ kstit\colon A]$ is *ex interim*

Relations:

$\mathsf{K}_\alpha A \supset [\alpha\ kstit\colon A]$

$\mathsf{K}_\alpha A \equiv \Box[\alpha\ kstit\colon A]$

Two kinds of oughts



$K_1 = \beta$ places coin heads up
$K_2 = \beta$ places coin tails up
$K_3 = K_6 = \alpha$ bets heads
$K_4 = K_7 = \alpha$ bets tails
$K_5 = K_8 = \alpha$ doesn't bet

Two readings of "α ought to bet heads"

causal

epistemic

Oughts, review and simplification

1. Deontic stit model:

   $$\langle \mathit{Tree}, <, \mathit{Agent}, \mathit{Choice}, \mathit{Value}, v \rangle,$$

   where $Value$ maps histories into numbers, representing values

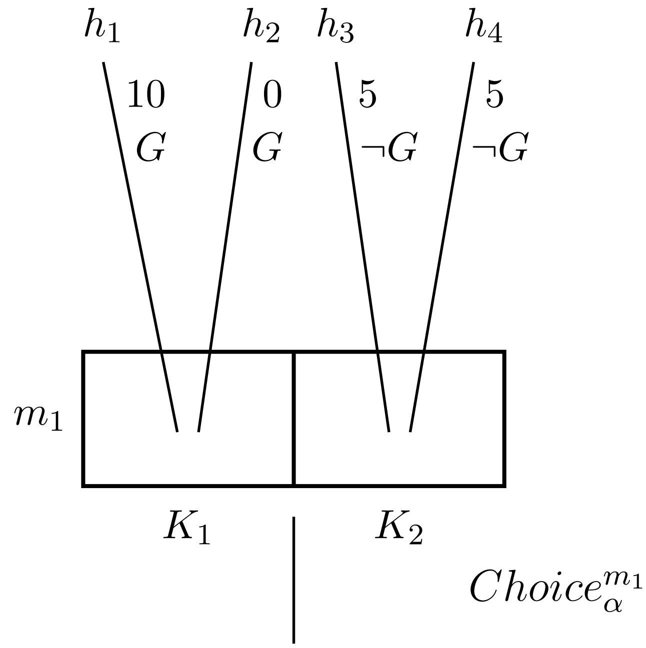
2. Evaluation rule: standard deontic operator

   - $m/h \models \bigcirc A$ iff $m/h' \models A$ for each "best" $h' \in H^m$.

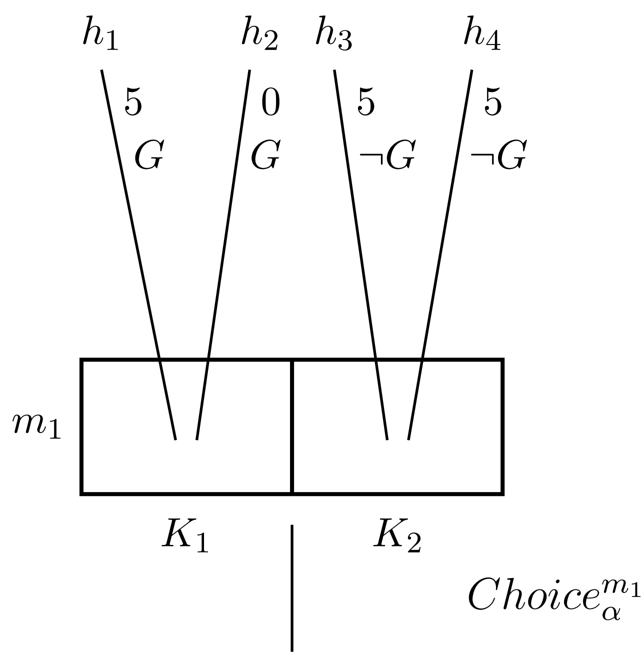
3. Meinong/Chisholm analysis:

   "S ought to bring it about that p" defined as "It ought to be that S brings it about that p"

   $$\begin{aligned} \bigcirc[\alpha\ \mathit{cstit}\text{: } A] &= \text{It ought to be that } \alpha \text{ sees to it that } A \\ &= \alpha \text{ ought to see to it that } A \end{aligned}$$



4. The gambling problem: $m_1 \models \bigcirc[\alpha\ stit\colon G]$



5. More gambling: $m_1 \not\models \bigcirc[\alpha\ stit\colon \neg G]$

6. Ordering the action *tokens*:

Where $K, K' \in Choice^m_\alpha$

$K \leq K'$ iff

For all $h \in K, h' \in K'$: $[Value(h) \leq Value(h')]$

$K < K'$ iff

$K \leq K'$ and $\neg(K' \leq K)$

Note: single-agent simplification of earlier ordering

7. Optimal action tokens:

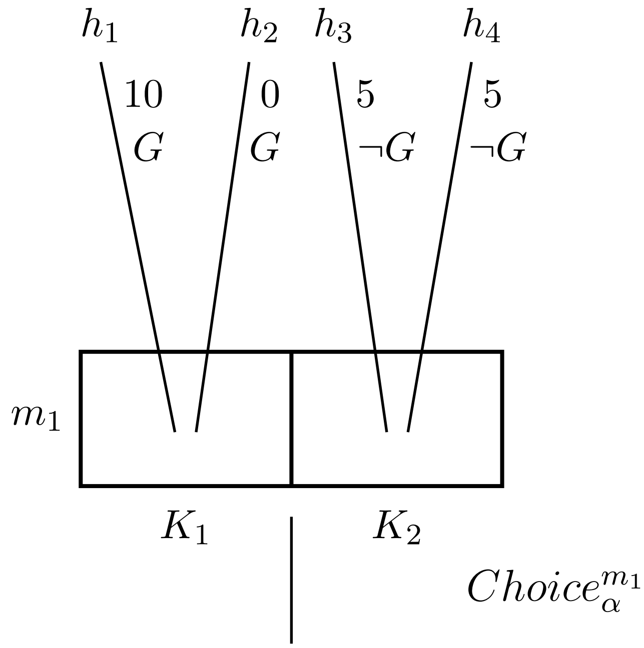
K-Optimal$^m_\alpha = \{K \in Choice^m_\alpha : \neg\exists\, K' \in Choice^m_\alpha (K < K')\}$

Note: *K-Optimal*$^m_\alpha$ here is like earlier *D-Optimal*$^m_\alpha$

8. Evaluation rule: dominance ought

- $m/h \models \bigodot[\alpha\ cstit\text{: } A]$ iff

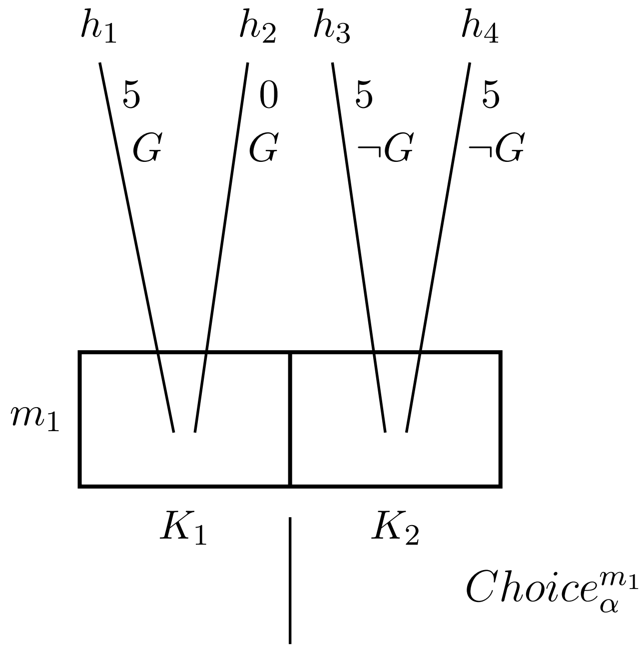
  For all $K \in$ *K-Optimal*$^m_\alpha$: $K \subseteq |A|^m$



9. The gambling problem, resolved:

$$m_1 \models \bigcirc[\alpha\ stit\colon G], \text{ but } m_1 \not\models \bigodot[\alpha\ stit\colon G]$$

$$\textit{K-Optimal}_\alpha^{m_1} = \{K_1, K_2\}$$

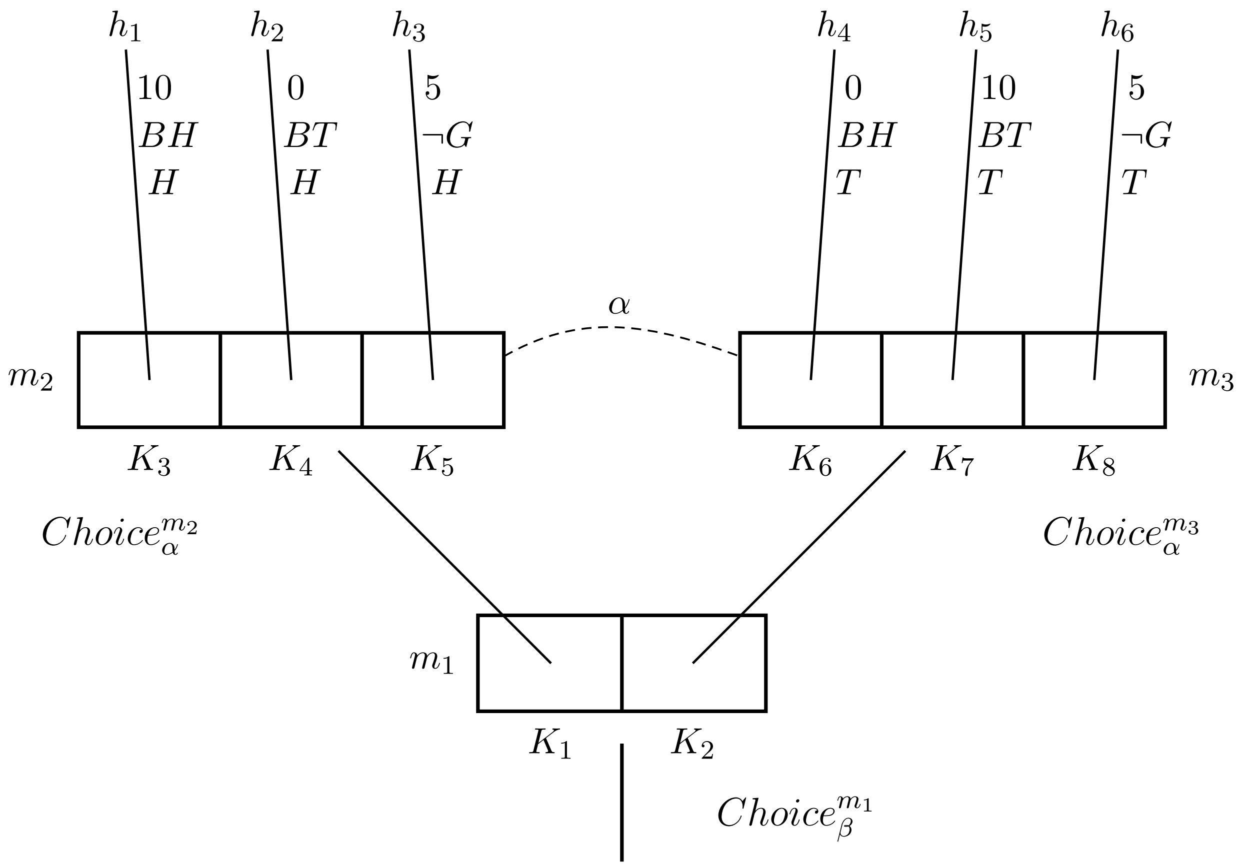


10. More gambling, resolved:

$m_1 \not\models \bigcirc[\alpha\ stit\colon \neg G]$, but $m_1 \models \odot[\alpha\ stit\colon \neg G]$

$\textit{K-Optimal}_\alpha^{m_1} = \{K_2\}$

Knowledge and oughts



1. $m_2 \models \odot[\alpha \textit{ stit}: BH]$, but is that right?

   $\textit{K-Optimal}_\alpha^{m_2} = \{K_3\}$

   Maybe it is, but agent just doesn't know it

   Criticism tied to knowledge of oughts ??

2. Indistinguishability: now an equivalence relation between *moments*, given our (C4):

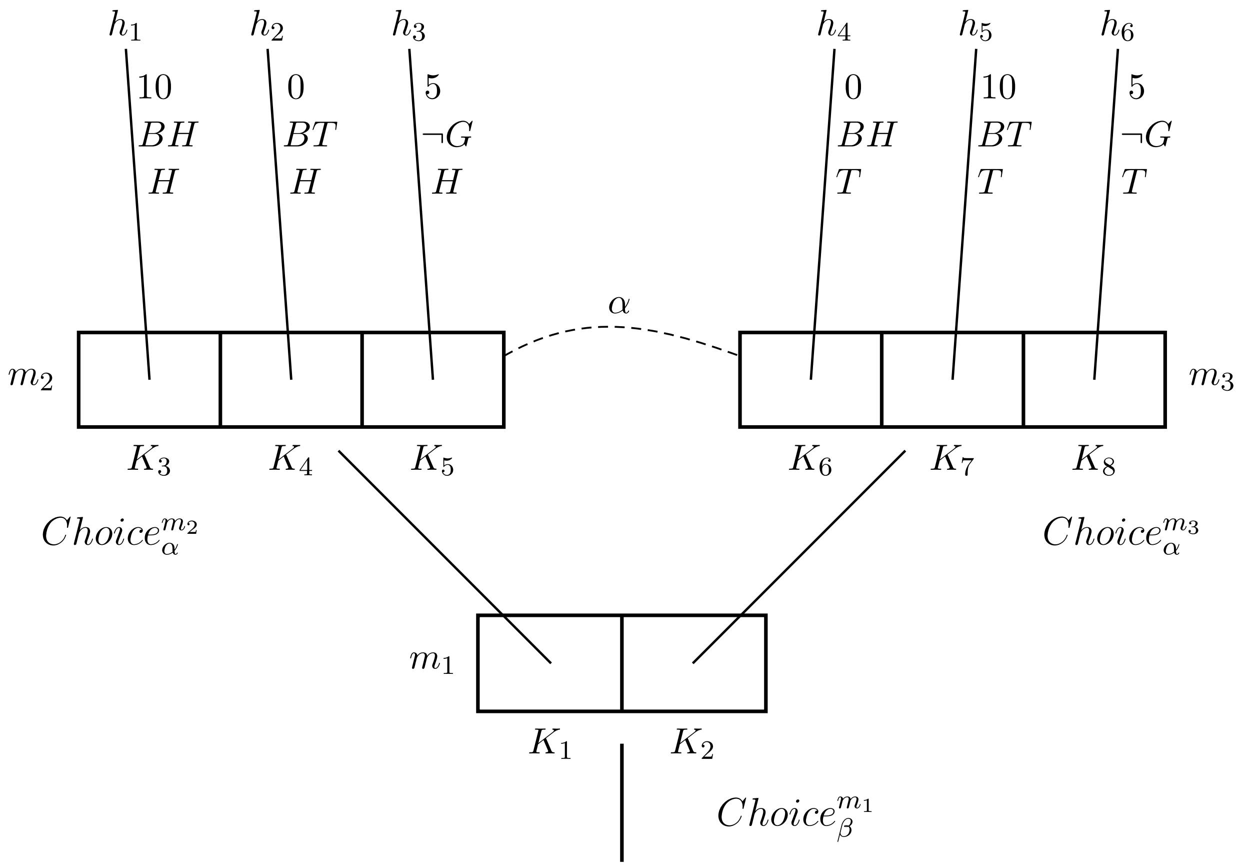
$$m \sim_\alpha m'$$

3. Epistemic deontic stit models:

$$\langle \mathit{Tree}, <, \mathit{Agent}, \mathit{Choice}, \mathit{Value}, \{\sim_\alpha\}_{\alpha \in \mathit{Agent}}, v \rangle$$

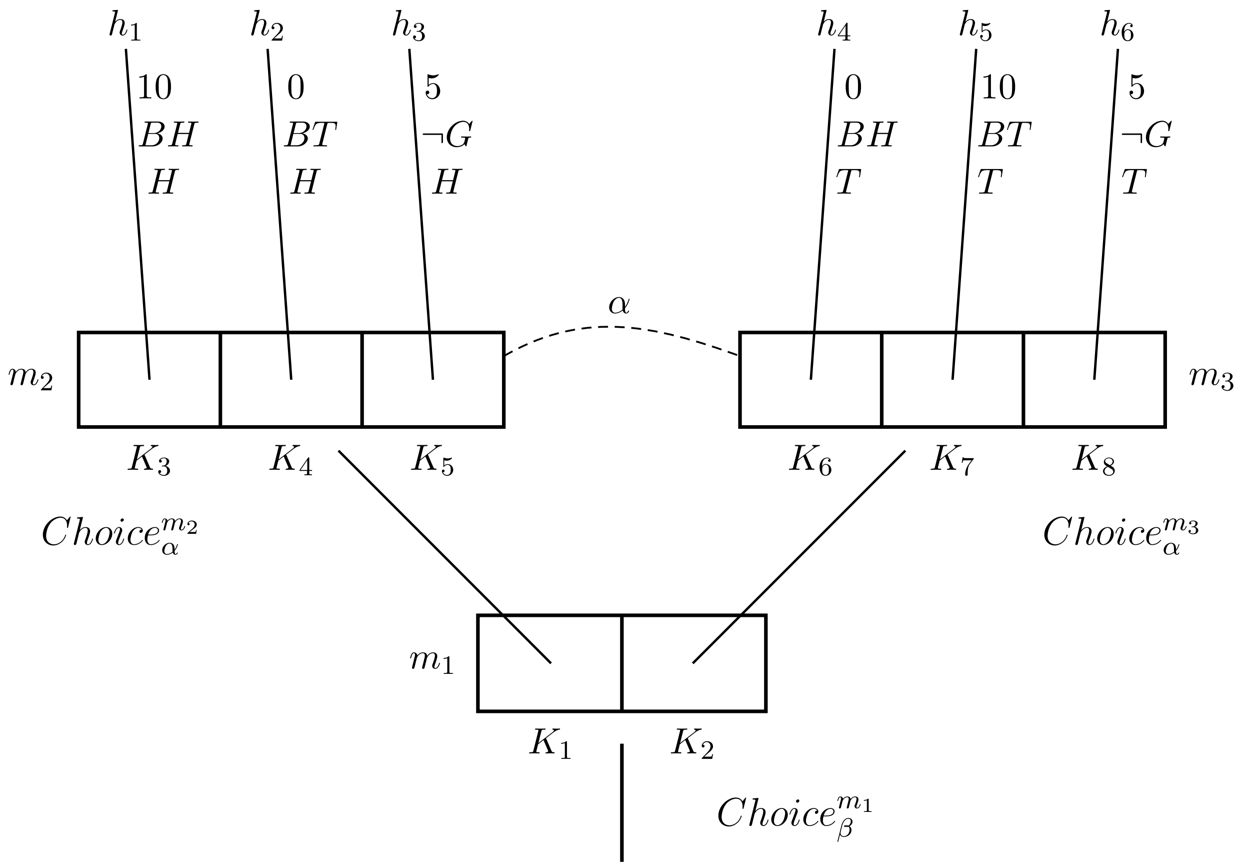
4. New evaluation rule: knowledge operator

   - $m/h \models \mathsf{K}_\alpha A$ iff $m'/h' \models A$ for all m'/h' such that $m \sim_\alpha m'$



5. $m_2 \models \odot[\alpha\ stit\colon BH]$, but $m_2 \not\models \mathsf{K}_\alpha \odot[\alpha\ stit\colon BH]$

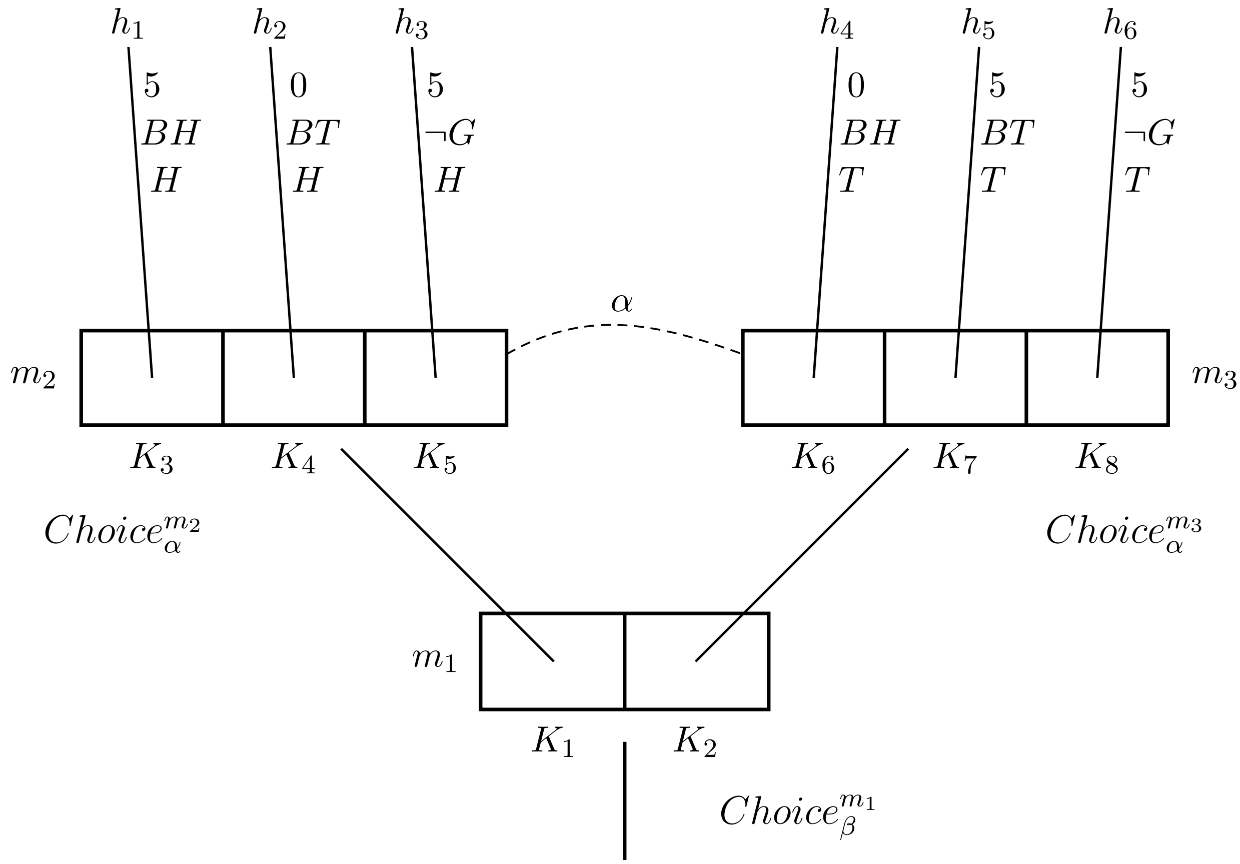
$\textit{K-Optimal}_\alpha^{m_2} = \{K_3\}$
$\textit{K-Optimal}_\alpha^{m_3} = \{K_7\}$



6. Problem #1:

   $m_2 \models \mathsf{K}_\alpha \odot[\alpha\ stit\colon G]$, but that's wrong

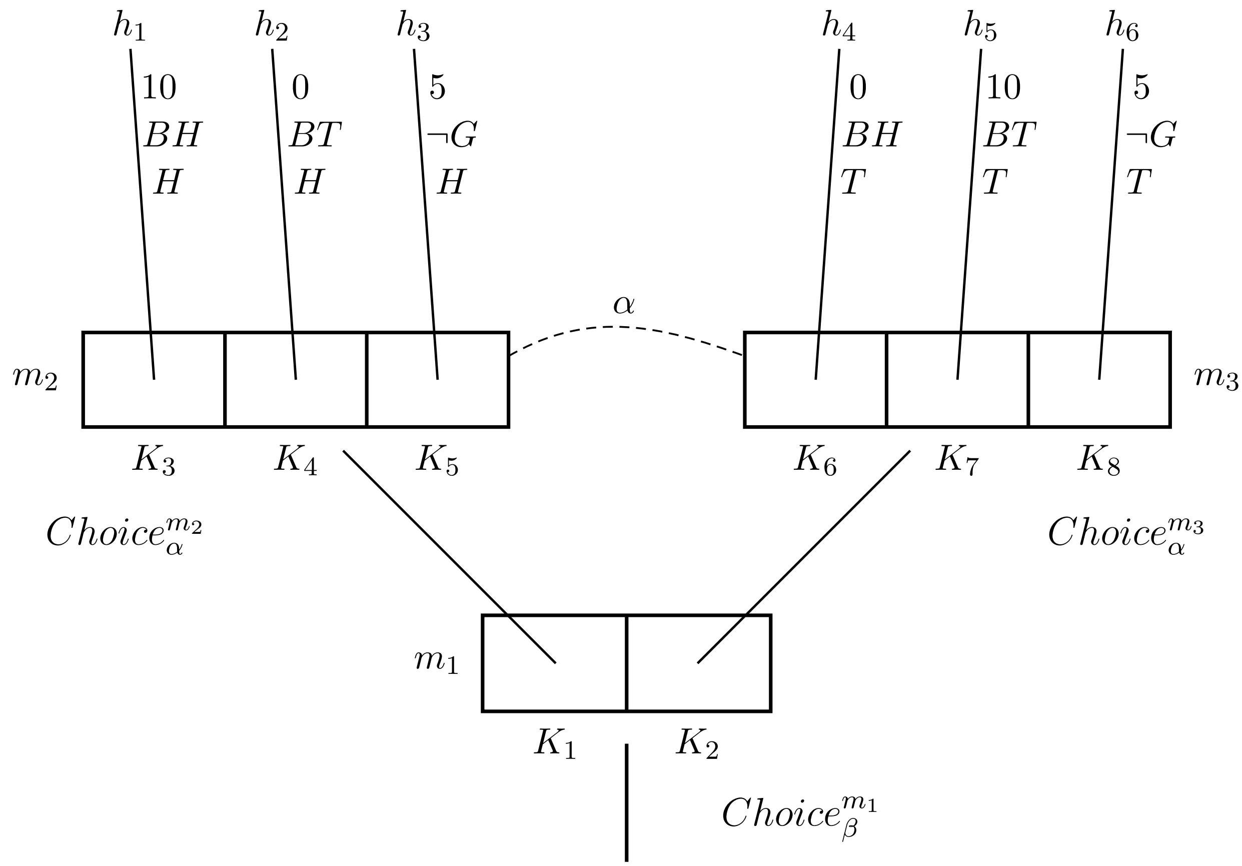
   $\textit{K-Optimal}_\alpha^{m_1} = \{K_3\}$
   $\textit{K-Optimal}_\alpha^{m_3} = \{K_7\}$



7. Problem #2:

$m_2 \not\models \mathsf{K}_\alpha \odot[\alpha\ stit\colon \neg G]$, but that's wrong too

$K\text{-}Optimal_\alpha^{m_2} = \{K_3, K_5\}$
$K\text{-}Optimal_\alpha^{m_3} = \{K_7, K_8\}$



8. Problem #3:

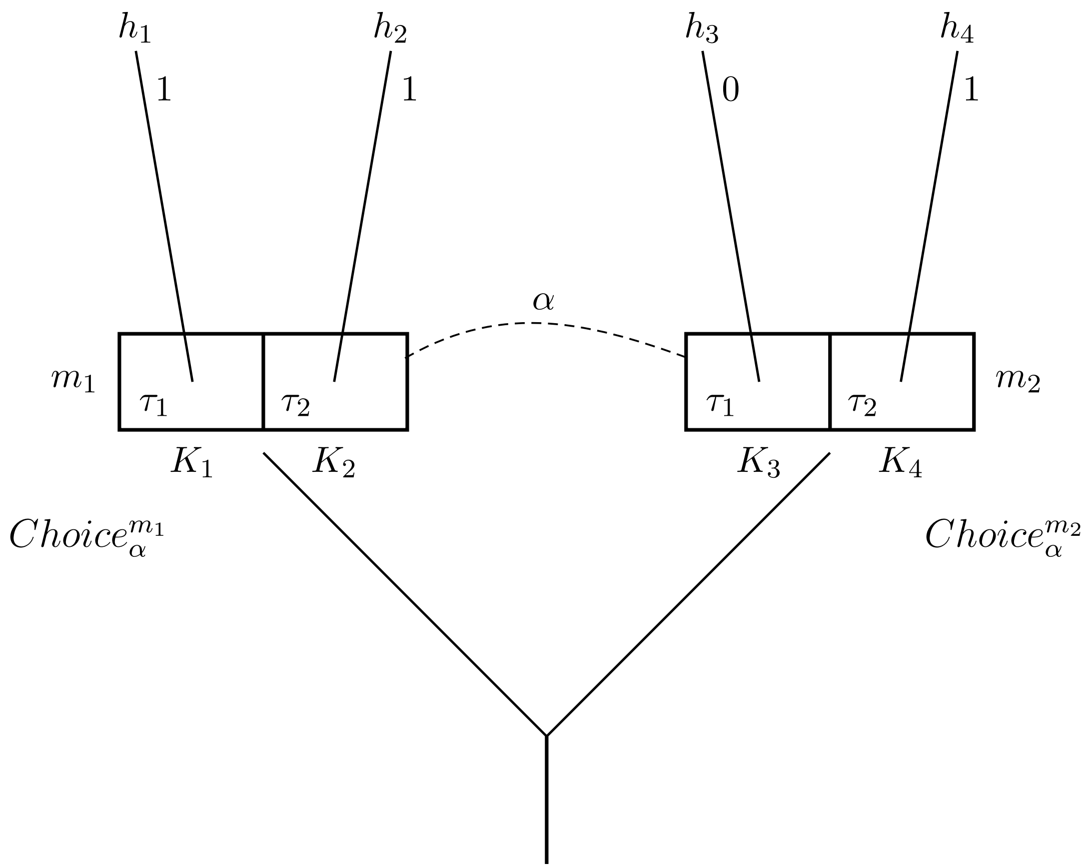
$m_2 \models \mathsf{K}_\alpha \odot[\alpha\ stit\colon W]$, but what action to take?

$\textit{K-Optimal}_\alpha^{m_2} = \{K_3\}$
$\textit{K-Optimal}_\alpha^{m_3} = \{K_7\}$

Oughts should be action guiding

Ought implies can

Epistemic oughts



1. Information set: a set I of moments subject to

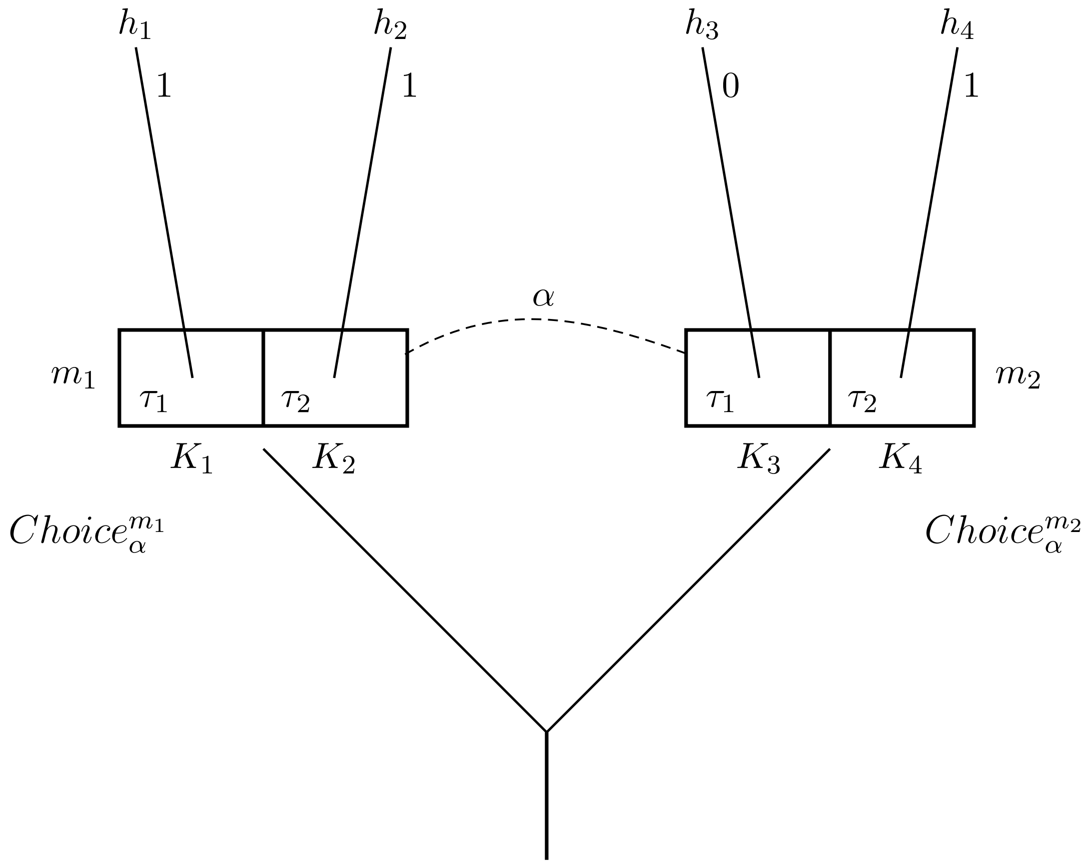
   If $m, m' \in I$, then $Type_\alpha^m = Type_\alpha^{m'}$

   In particular,

   $I_\alpha^m = \{m' : m \sim_\alpha m'\}$

   is an information set, representing information available to α at m

   Example: $I_\alpha^{m_1} = \{m_1, m_2\}$



2. Goal: rank action types, relative to I

3. One idea: take

$$[\tau]_\alpha^I = \bigcup\{[\tau]_\alpha^{m'} : m' \in I\}$$

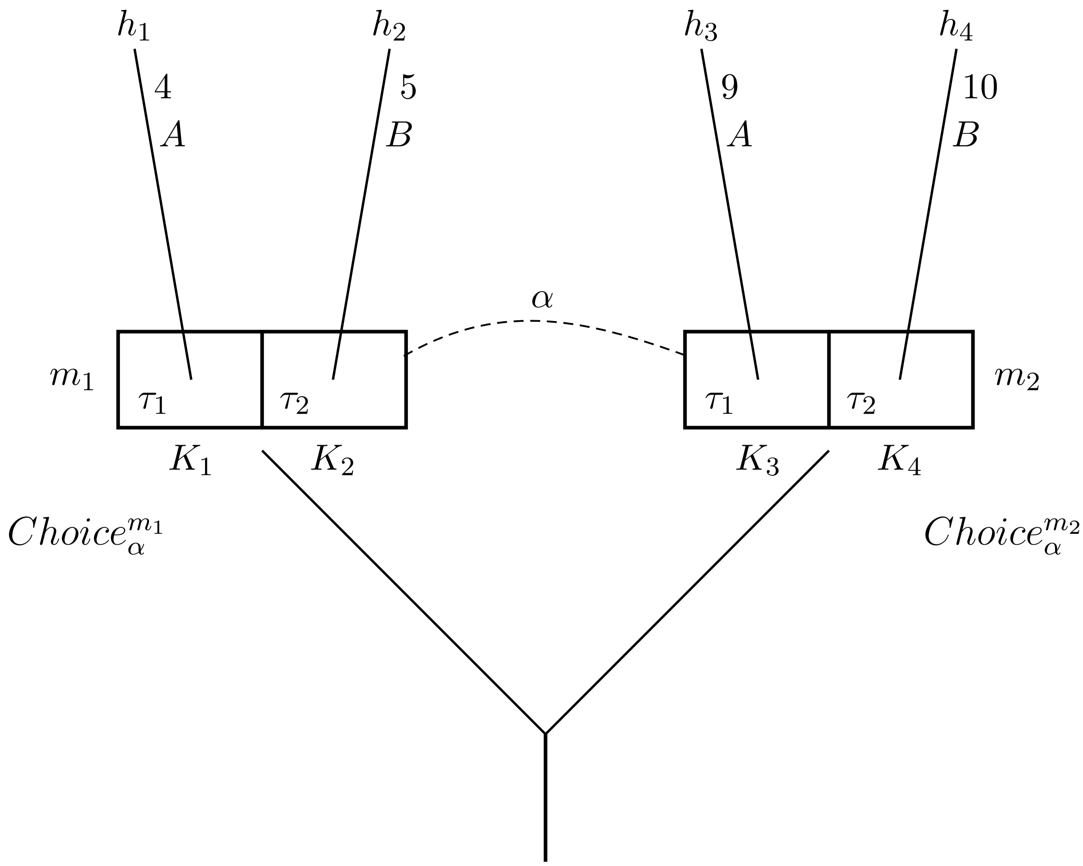
   and the define

   τ' better than τ based on I iff

$$[\tau]_\alpha^I < [\tau']_\alpha^I$$

4. Example: τ_2 better than τ_1 based on $I = \{m_1, m_2\}$, since

$$[\tau_1]_\alpha^I < [\tau_2]_\alpha^I$$

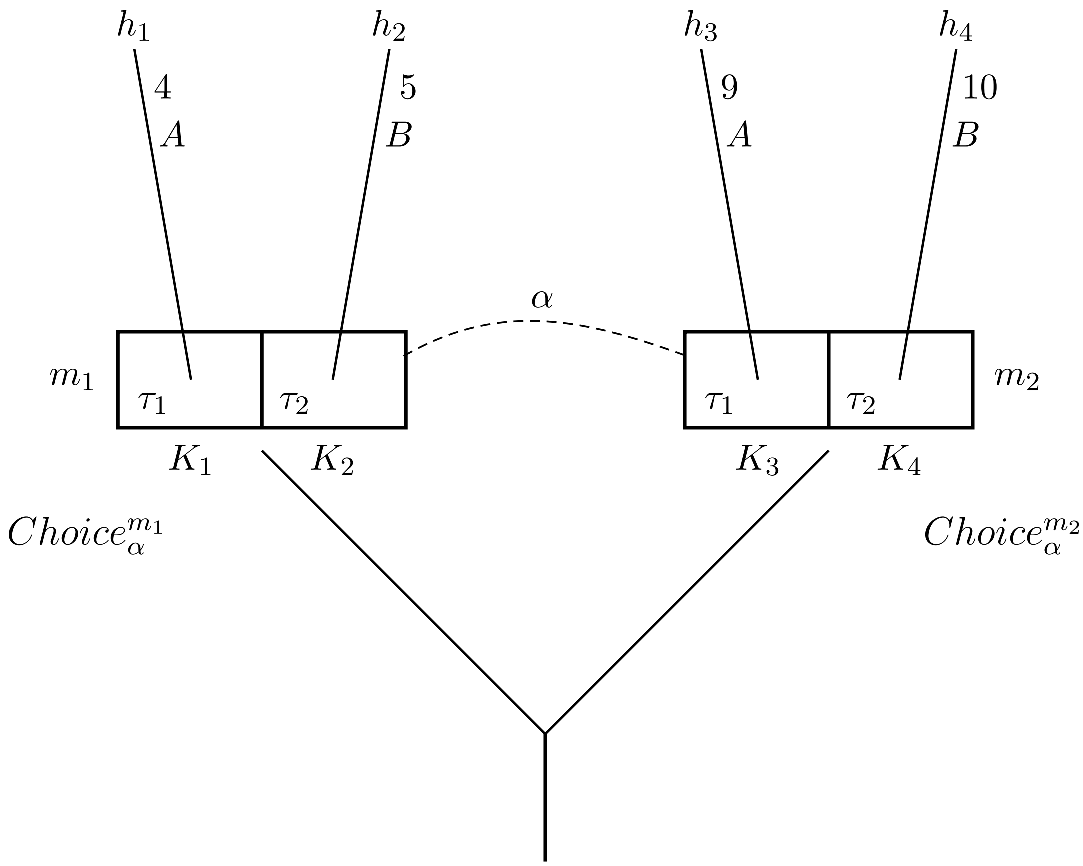


5. Problem: we do not have

$$[\tau_1]^I_\alpha < [\tau_2]^I_\alpha$$

but it seems by *sure-thing reasoning* that

τ_2 is better than τ_1



6. Instead: where $\tau, \tau' \in Type^m_\alpha$

$\tau \preceq^I_\alpha \tau'$ iff

For all $m' \in I : [\tau]^{m'}_\alpha \leq [\tau']^{m'}_\alpha$

$\tau \prec^I_\alpha \tau'$ iff

$\tau \preceq^I_\alpha \tau'$ and $\neg(\tau' \preceq^I_\alpha \tau)$

7. Optimal action types:

$$\textit{T-Optimal}^{I}_{\alpha} = \{\tau \in Type^{m}_{\alpha} : \neg\exists\tau' \in Type^{m}_{\alpha}(\tau \prec^{I}_{\alpha} \tau')\}.$$

8. Note: in this and previous definiton, I is unspecified, but I^{m}_{α} is particularly interesting

9. Labeled deontic stit model: Add $Value$ to labeled stit models, and then . . .

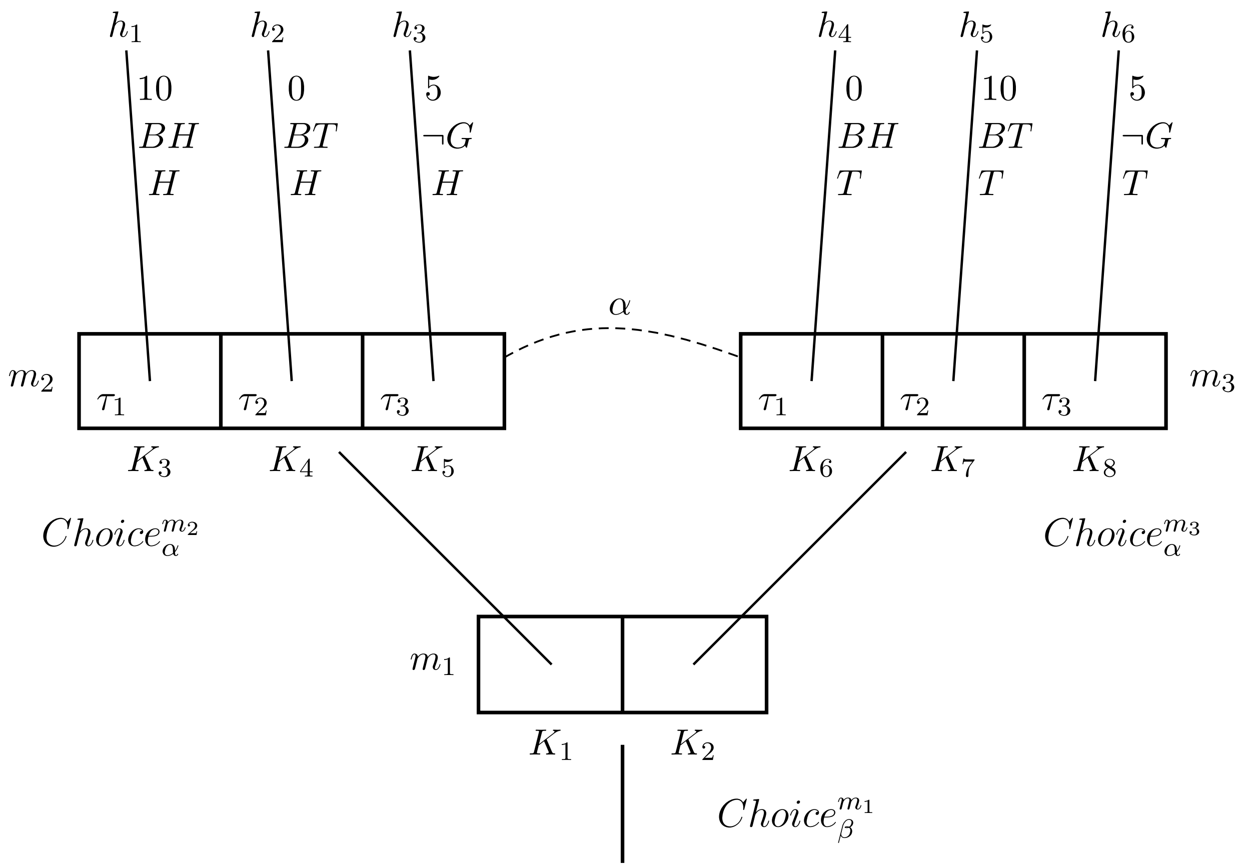
10. Evaluation rule: epistemic ought

    - $m/h \models \bigodot[\alpha\ kstit\text{: } A]$ iff

      For each $\tau \in \textit{T-Optimal}^{I^{m}_{\alpha}}_{\alpha}$:

      For each $m' \in I^{m}_{\alpha}$: $[\tau]^{m'}_{\alpha} \subseteq |A|^{m'}$

    Note: here, I is bound to I^{m}_{α}

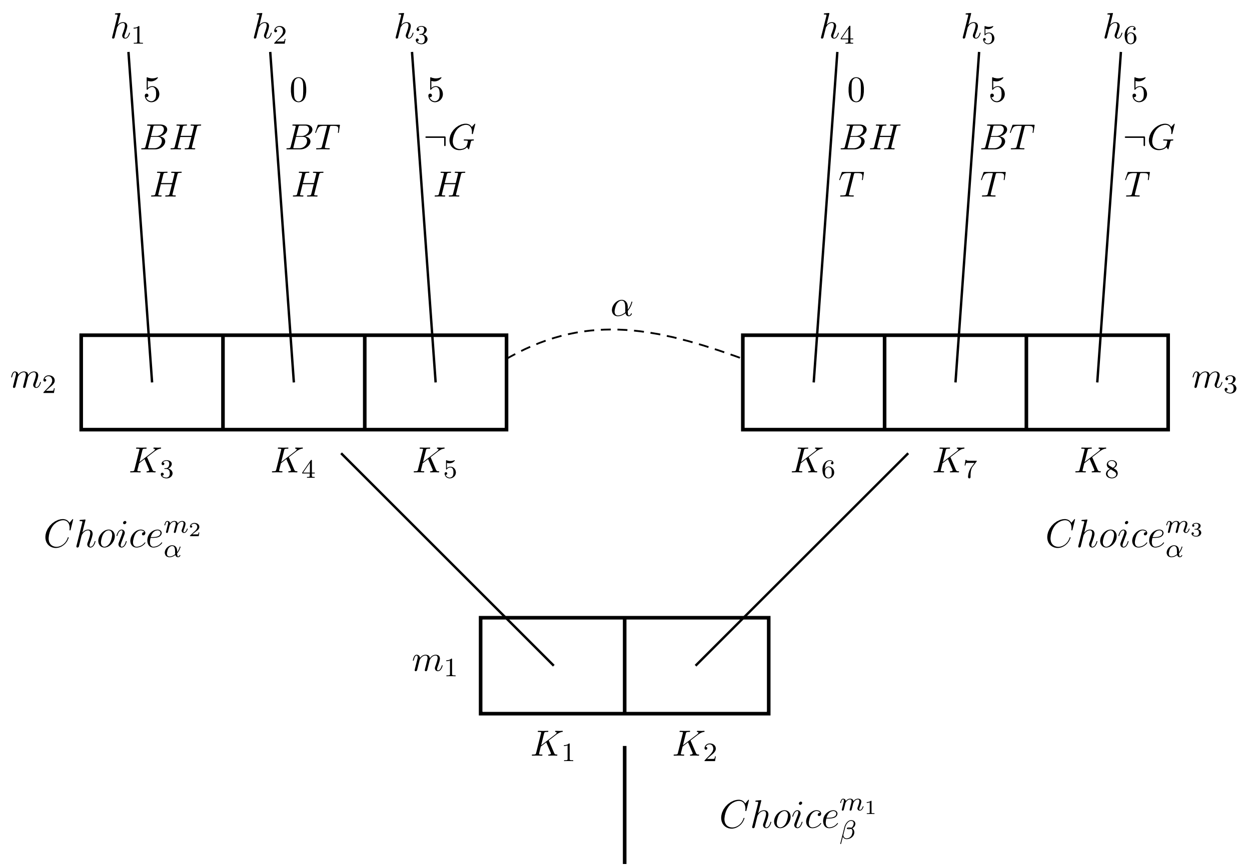


11. Problem #1:

$m_2 \models \mathsf{K}_\alpha \odot[\alpha\ stit\colon G]$, but $m_2 \not\models \odot[\alpha\ kstit\colon G]$

$\textit{K-Optimal}_\alpha^{m_2} = \{K_3\}$
$\textit{K-Optimal}_\alpha^{m_3} = \{K_7\}$

$\textit{T-Optimal}_\alpha^{I_\alpha^{m_1}} = \{\tau_1, \tau_2, \tau_3\}$

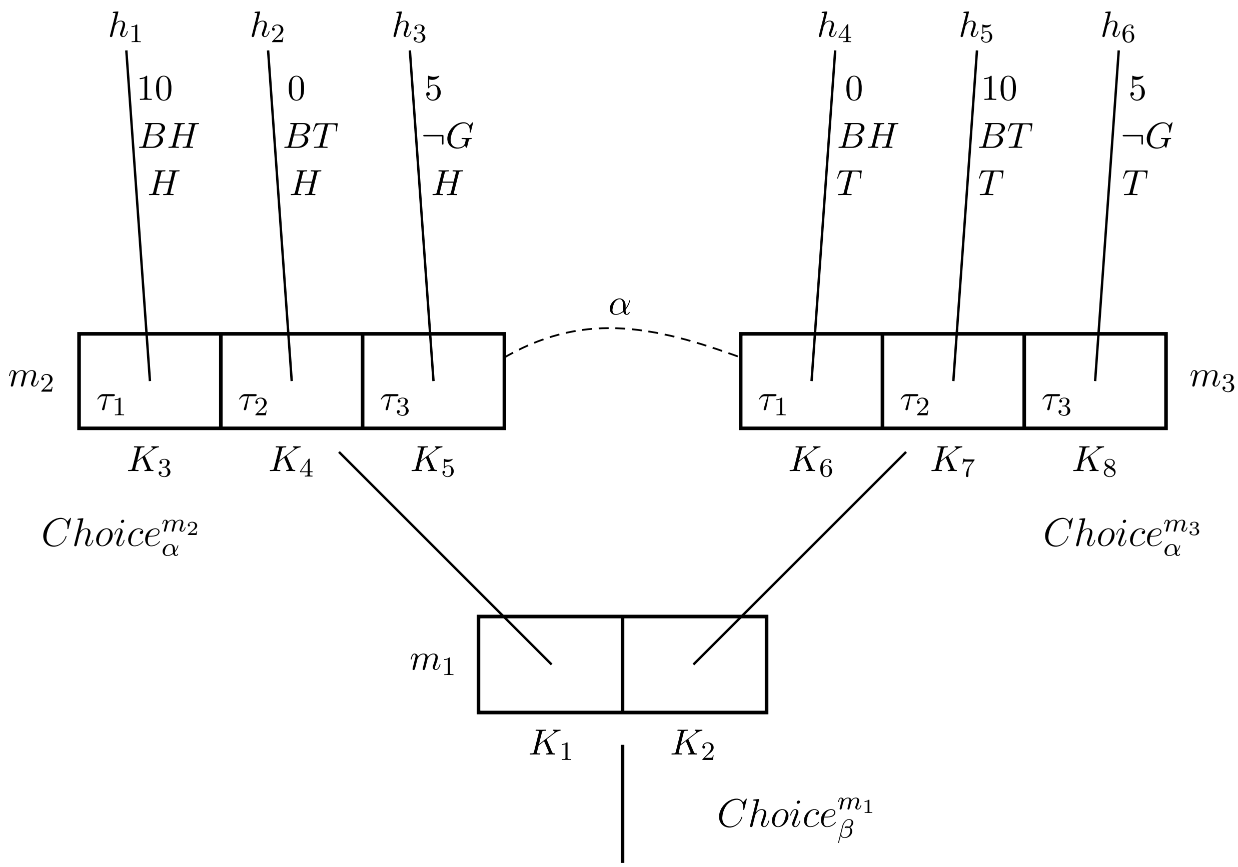


12. Problem #2:

$m_2 \not\models \mathsf{K}_\alpha \odot[\alpha\ stit\colon \neg G]$, but $m_2 \models \odot[\alpha\ kstit\colon \neg G]$

$K\text{-}Optimal_\alpha^{m_2} = \{K_3, K_5\}$
$K\text{-}Optimal_\alpha^{m_3} = \{K_7, K_8\}$

$T\text{-}Optimal_\alpha^{I_\alpha^{m_1}} = \{\tau_3\}$



13. Problem #3:

$m_2 \models \mathsf{K}_\alpha \bigodot[\alpha\ stit\colon W]$, but what action to take??

Here, do not have $m_1 \models \bigodot[\alpha\ kstit\colon W]$

$K\text{-}Optimal_\alpha^{m_2} = \{K_3\}$
$K\text{-}Optimal_\alpha^{m_3} = \{K_7\}$

$T\text{-}Optimal_\alpha^{I_\alpha^{m_1}} = \{\tau_1, \tau_2, \tau_3\}$

13. Observations on the epistemic ought:

Normal operator supporting

$$\odot[\alpha\ kstit\text{: } A] \supset \Diamond[\alpha\ kstit\text{: } A]$$

No relations between two oughts; neither

$$\odot[\alpha\ stit\text{: } A] \supset \odot[\alpha\ kstit\text{: } A]$$

$$\odot[\alpha\ kstit\text{: } A] \supset \odot[\alpha\ stit\text{: } A]$$

But everything collapses if $I^m_\alpha = \{m\}$:

$$\odot[\alpha\ stit\text{: } A] \equiv \odot[\alpha\ kstit\text{: } A],$$

since

$$\textsf{\textit{K-Optimal}}^m_\alpha = \{[\tau]^m_\alpha : \tau \in \textsf{\textit{T-Optimal}}^{I^m_\alpha}_\alpha\}$$

$$\textsf{\textit{T-Optimal}}^{I^m_\alpha}_\alpha = \{Label(K) : K \in \textsf{\textit{K-Optimal}}^m_\alpha\}$$

Finally, knowledge of epistemic oughts:

$$\odot[\alpha\ kstit\text{: } A] \supset \mathsf{K}_\alpha \odot[\alpha\ kstit\text{: } A]$$

Assessment sensitivity

1. Natural idea

   $\odot[\alpha\ stit: A] =$ "objective"

   $\odot[\alpha\ kstit: A] =$ "subjective"

2. Problem:

   Is "ought" lexically ambiguous??

3. MacFarlane's suggestion:

   Interpretation of agentive oughts depends on information at context of assessment

   Objective feel: assessment information better than agent's information

   Subjective feel: assessment information closer to agent's information

4. From epistemic to informational oughts

Epistemic:

- $m/h \models \odot[\alpha\ kstit\text{: } A]$ iff

  For each $\tau \in$ *T-Optimal*$_\alpha^{I_\alpha^m}$:

  For each $m' \in I_\alpha^m$: $[\tau]_\alpha^{m'} \subseteq |A|^{m'}$

Informational:

- $m/h/I \models \odot[\alpha\ istit\text{: } A]$ iff

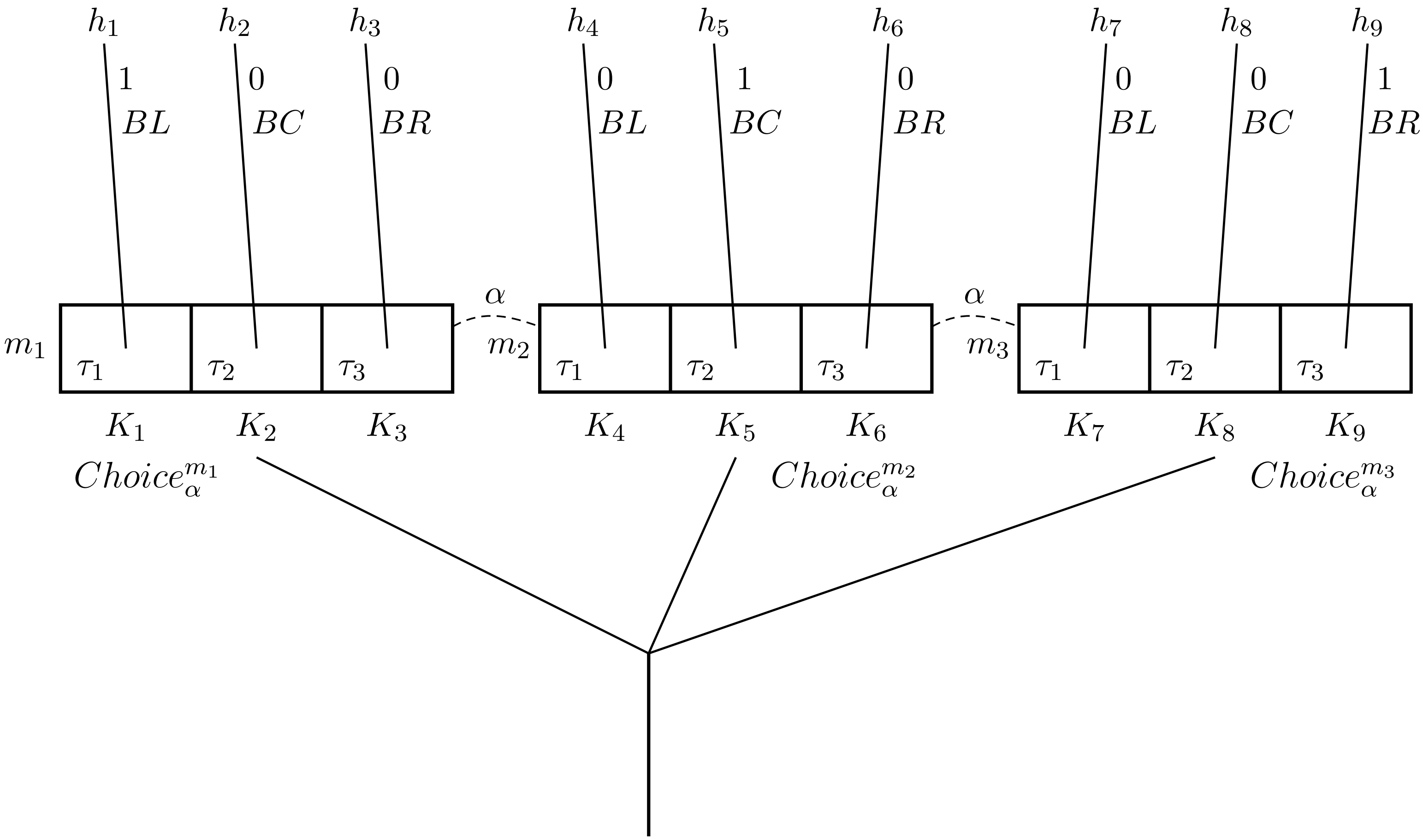
  For each $\tau \in$ *T-Optimal*$_\alpha^{I}$:

  For each $m' \in I$: $[\tau]_\alpha^{m'} \subseteq |A|^{m'}$

5. Where I_α^m = agent's knowledge and $I^* = \{m\}$:

$m/h/I^* \models \odot[\alpha\ istit\text{: } A]$ iff $m/h \models \odot[\alpha\ stit\text{: } A]$

$m/h/I_\alpha^m \models \odot[\alpha\ istit\text{: } A]$ iff $m/h \models \odot[\alpha\ kstit\text{: } A]$



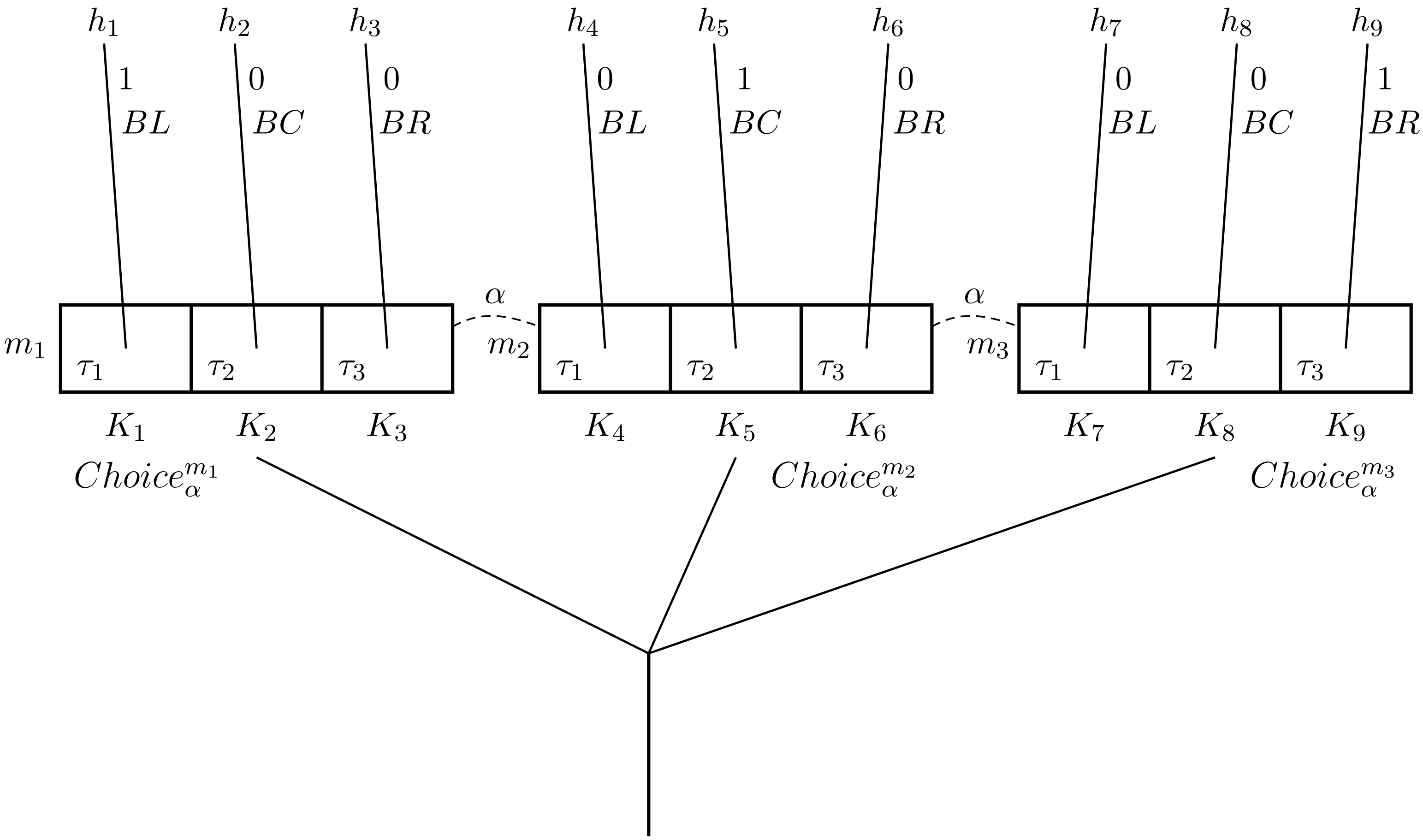
6. Example:

   Three information sets

   $$I^* = \{m_1\}$$

   $$I' = \{m_1, m_2\}$$

   $$I_\alpha^{m_1} = \{m_1, m_2, m_3\}$$

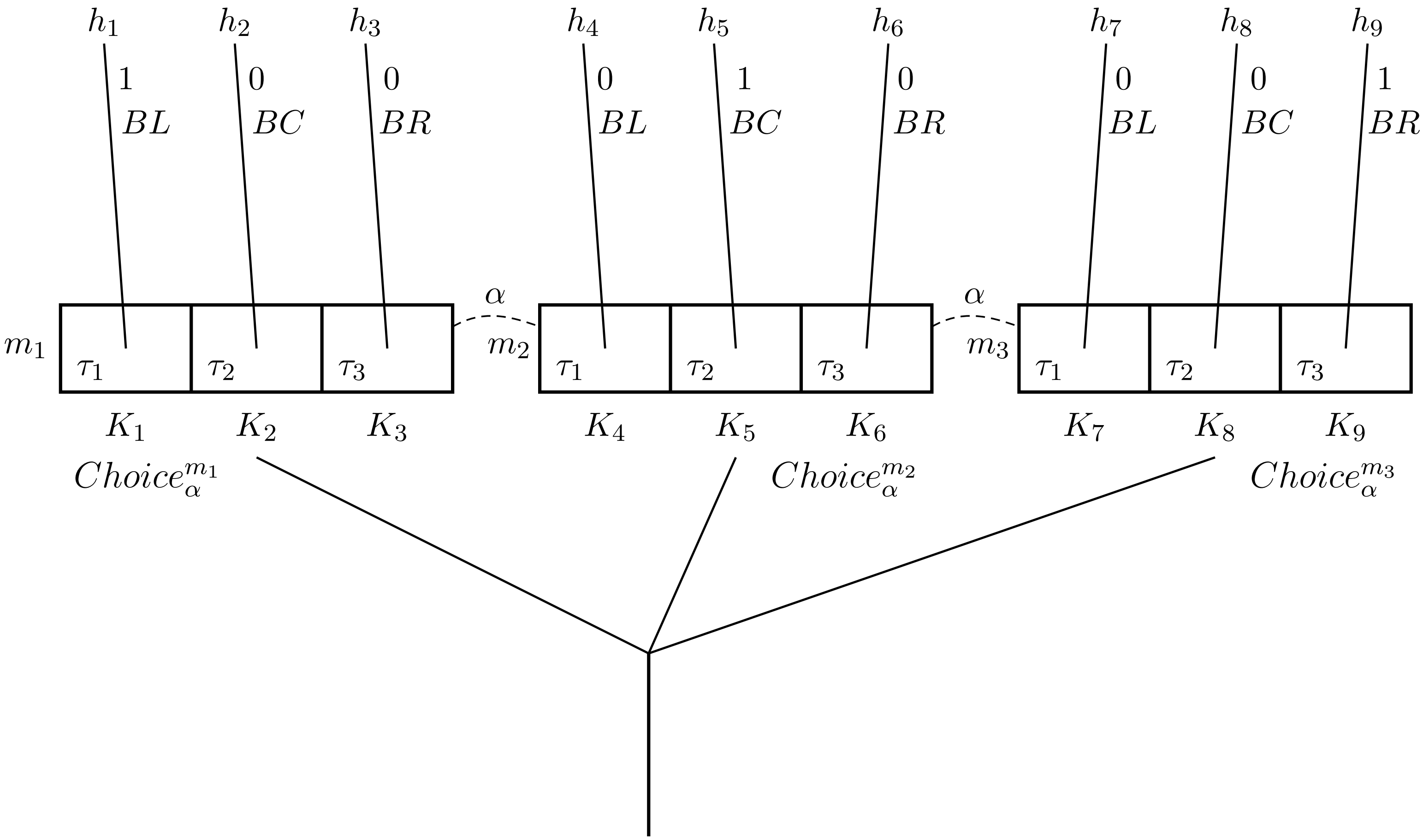


7. What is optimal?

$$T\text{-}Optimal_\alpha^{I^*} = \{\tau_1\}$$

$$T\text{-}Optimal_\alpha^{I'} = \{\tau_1, \tau_2\}$$

$$T\text{-}Optimal_\alpha^{I_\alpha^{m_1}} = \{\tau_1, \tau_2, \tau_3\}$$



8. So have

$$m_1/h_1/I^* \models \odot[\alpha\ istit\text{: } BL]$$

$$m_1/h_1/I' \models \odot[\alpha\ istit\text{: } BL \vee BC] \wedge \neg \odot[\alpha\ istit\text{: } BL]$$

$$m_1/h_1/I_\alpha^{m_1} \models \odot[\alpha\ istit\text{: } BL \vee BC \vee BR] \wedge \neg \odot[\alpha\ istit\text{: } BL \vee BC]$$

Conditional epistemic oughts

1. If A is *moment determinate* – ie, $A \equiv \Box A$ – then

   $$|A| = \{m/h : m/h \models A\}$$

   meets the constraint:

   If $m/h \in |A|$, then $m/h' \in |A|$ for each $h' \in H^m$

   so that $|A|$ can be represented as the set

   $$|A|^{\Box} = \{m : \exists h(m/h \in |A|)\}$$

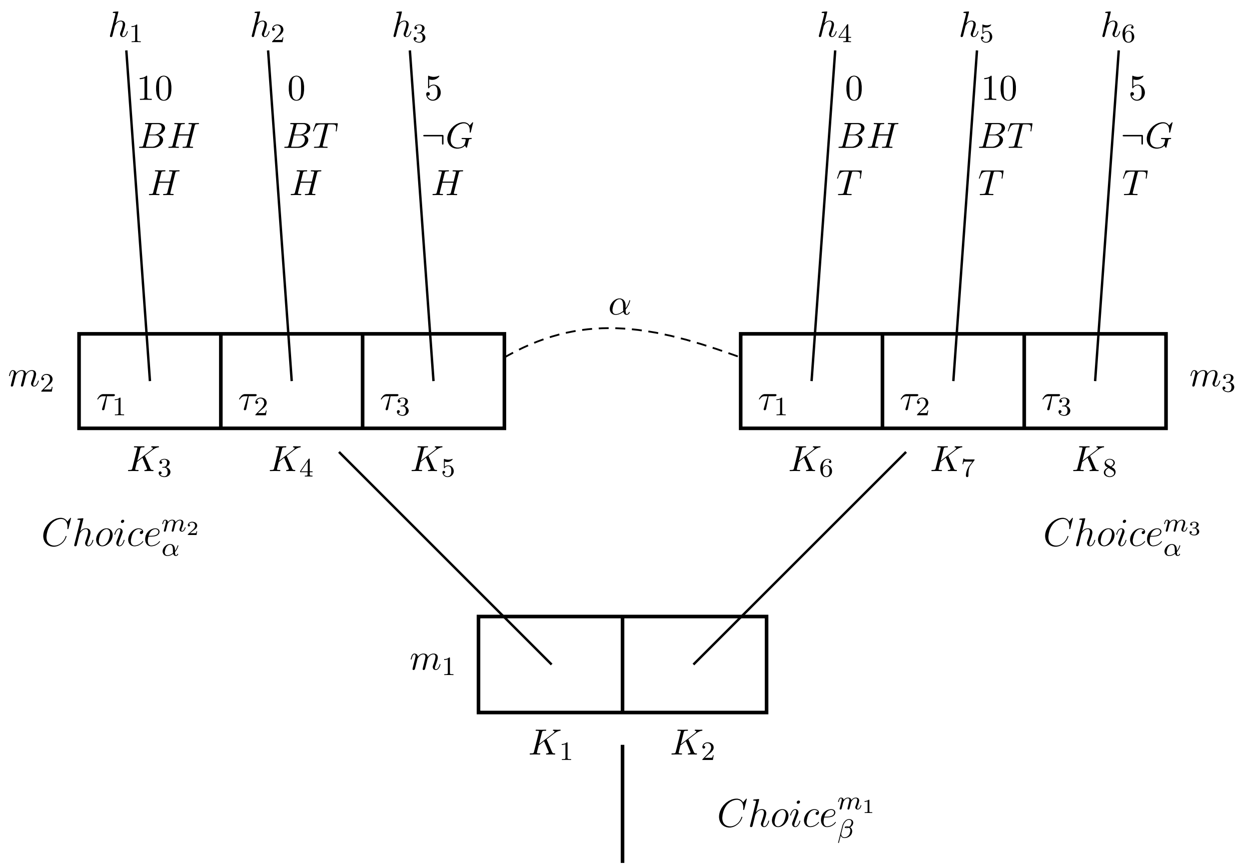
2. Local restriction: can conditionalize only on moment determinate propositions

3. Evaluation rule: conditional informational ought

   - $m/h/I \models \bigodot([\alpha\ istit\text{: } A]/B)$ iff

     For each $\tau \in$ *T-Optimal*$_{\alpha}^{I \cap |B|^{\Box}}$:

     For each $m' \in I \cap |B|^{\Box}$: $[\tau]_{\alpha}^{m'} \subseteq |A|^{m'}$

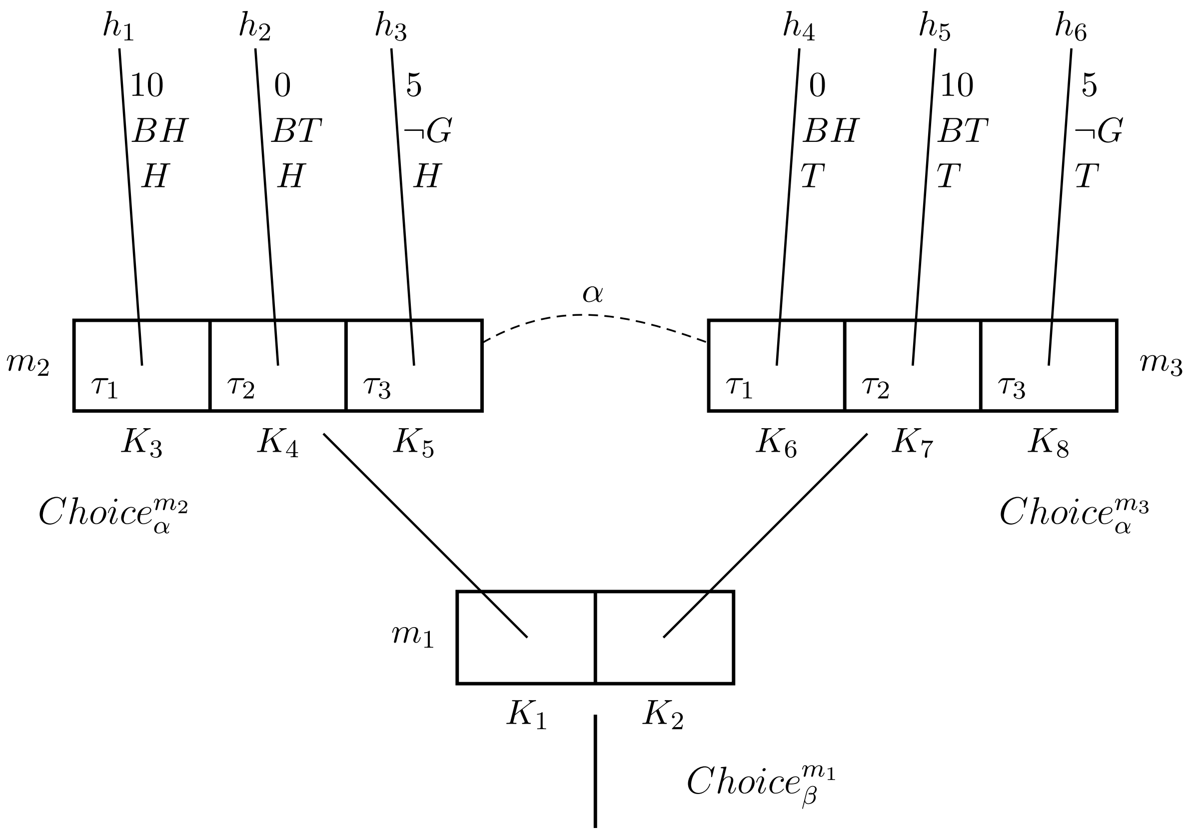


4. Modus ponens fails: at m_1 on basis of $I = \{m_1, m_2\}$, don't have

$$\frac{\begin{array}{l}\odot([\alpha \textit{ istit}\colon BH]/H) \\ H\end{array}}{\odot[\alpha \textit{ istit}\colon BH]}$$

$$\textit{T-Optimal}_\alpha^{I \cap |H|^\Box} = \{\tau_1\}$$

$$\textit{T-Optimal}_\alpha^{I} = \{\tau_1, \tau_2, \tau_3\}$$



5. Reasoning by cases fails: on basis of $I = \{m_1, m_2\}$, don't have

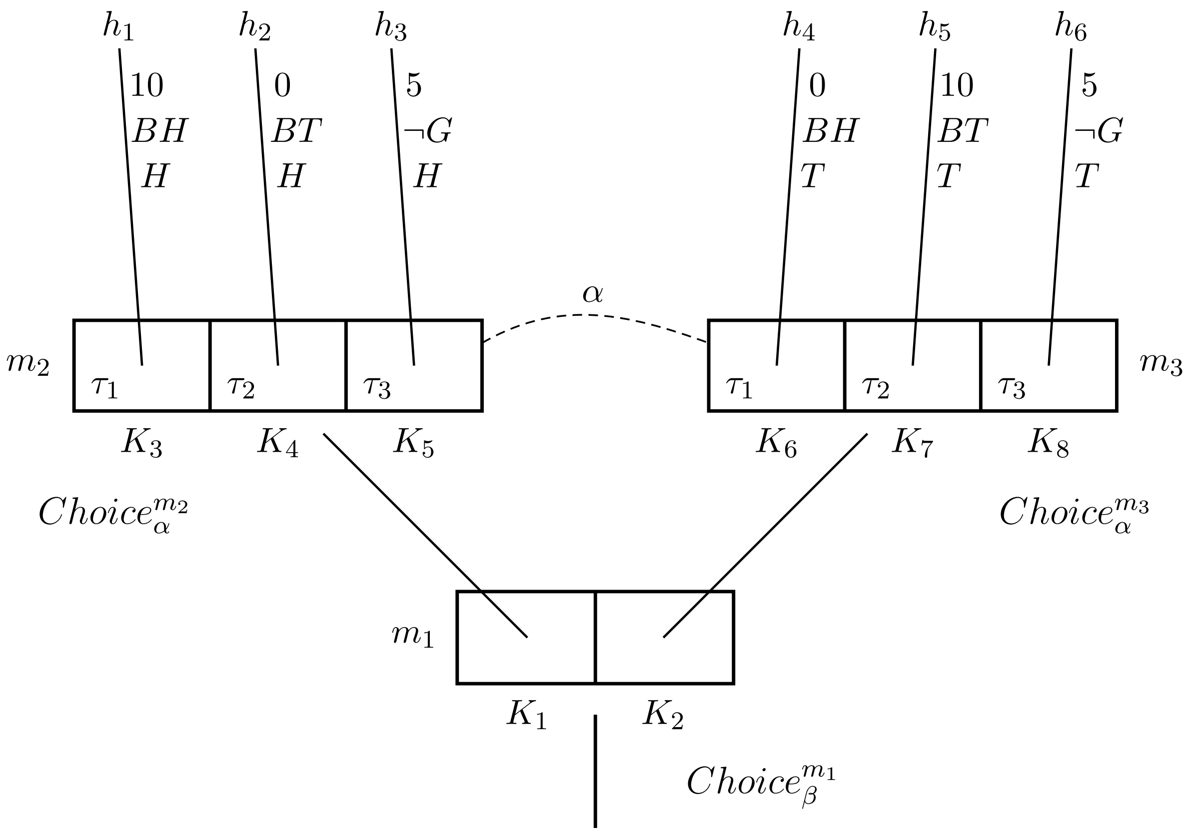
$$\frac{\begin{array}{l}\odot([\alpha\ istit\text{: } G]/H) \\ \odot([\alpha\ istit\text{: } G]/T) \\ H \vee T\end{array}}{\odot[\alpha\ istit\text{: } G]}$$

$$\textit{T-Optimal}_\alpha^{I \cap |H|^\Box} = \{\tau_1\}$$

$$\textit{T-Optimal}_\alpha^{I \cap |T|^\Box} = \{\tau_2\}$$

$$\textit{T-Optimal}_\alpha^{I} = \{\tau_1, \tau_2, \tau_3\}$$

6. Note similarity to conditional oughts in ADL

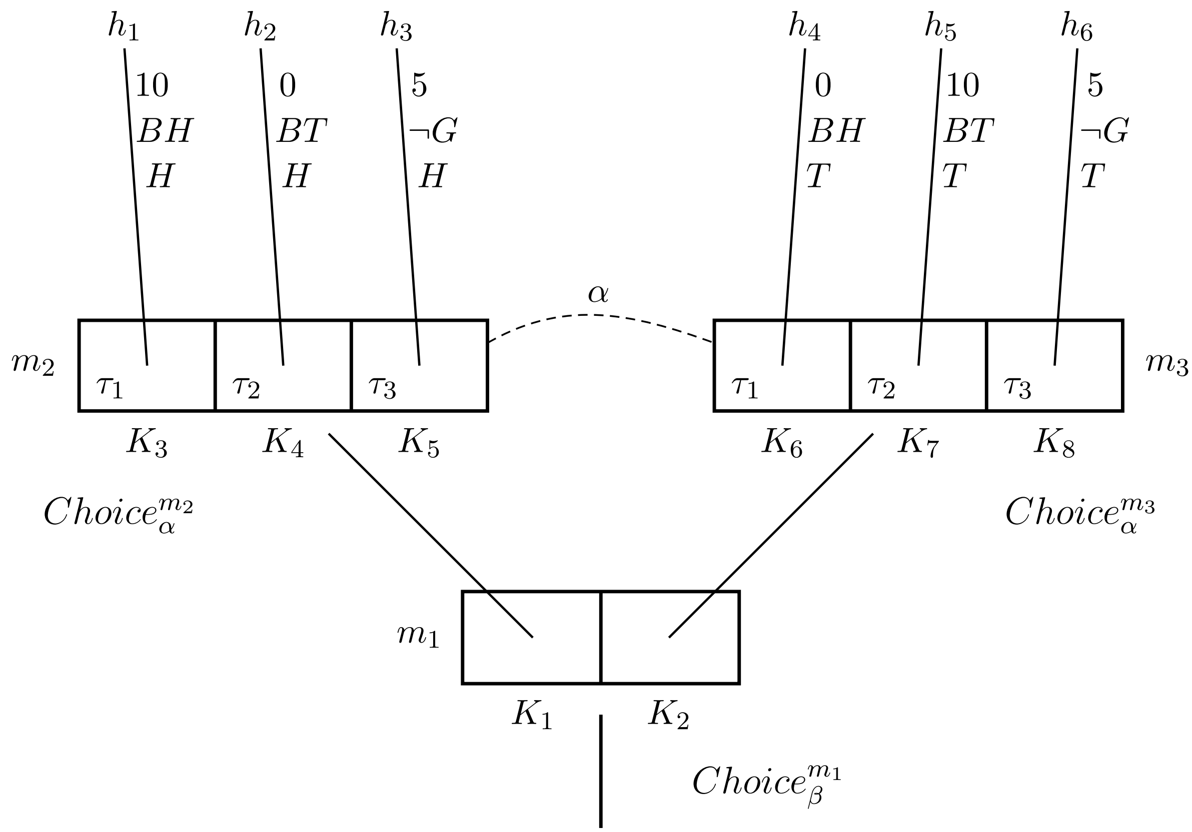


7. Note similarity to miners (Kolodny/MacFarlane):

   H = miners enter shaft A
   T = miners enter shaft B
   BH = we block shaft A
   BT = we block shaft B
   Give h_3 and h_6 value of 1

   Then on basis of $I = \{m_1, m_2\}$, don't have:

   $\odot([\alpha\ istit\text{: } BH]/H)$
   $\odot([\alpha\ istit\text{: } BT]/T)$
   $H \vee T$
   ———————————
   $\odot[\alpha\ istit\text{: } BH \vee BT]$



8. Differences too:

   Fused

   $\odot([\alpha\ istit\text{: } BH]/H)$

   but Kolodny/MacFarlane want

   [If H] $\odot[\alpha\ istit\text{: } BH]$

   Also, Kolodny/MacFarlane want

   $\odot[\alpha\ istit\text{: } \neg(BH \vee BT)]$

Today's summary

1. Talked about epistemic notion of ability

2. Introduced action types, labeled stit semantics

4. Talked about epistemic oughts

5. Theory based on ordering of actin types

6. Generalizations to informational/conditional oughts

7. Much work to be done:

    Generalize informational treatments
    Relax assumptions
    Multiple agents
    Group kstit
    Utilitarianism with knowledge