Challenges with Obligations in Multiagent Deontic Logic

Challenges 2, 3 and 4

Raik Hipler and Mark Scheibner

11.03.2021

Outline

Introduction

Non-Deterministic Actions

Moral Luck

Procrastination

Summary

Introduction

Ought-to-be and Ought-to-do

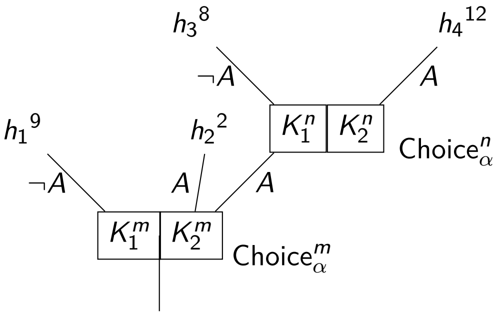
- SDL and DSDL give us frameworks for handling how things *ought to be*
- Since we are interested in the actions of agents, we need to describe what agents *ought to do*
- Simply redefining *ought to do* as *ought to be done* leads to various problems

STIT logic

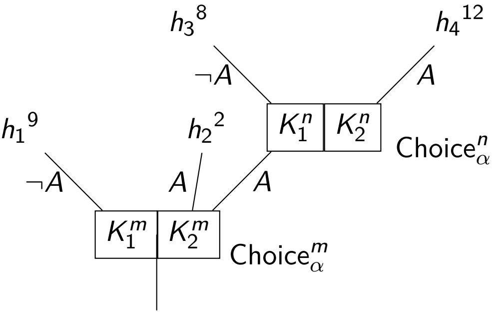
- STIT trees model possible decisions of (multiple) agents
- Inner node are sets of actions the agent has to decide between
- Each action may non-deterministically lead to different outcomes
- Leafs are histories with associated utility values

STIT logic

- STIT trees model possible decisions of (multiple) agents
- Inner node are sets of actions the agent has to decide between
- Each action may non-deterministically lead to different outcomes
- Leafs are histories with associated utility values

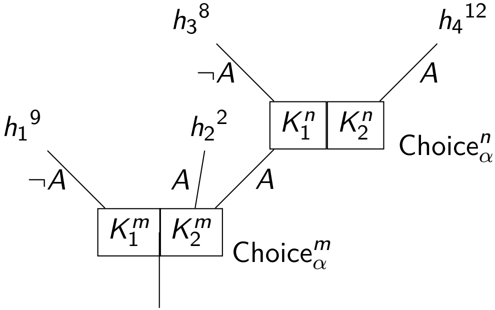


STIT logic
Semantics



- Consider moment-history pairs
- Does $m, h \models \varphi$ hold?

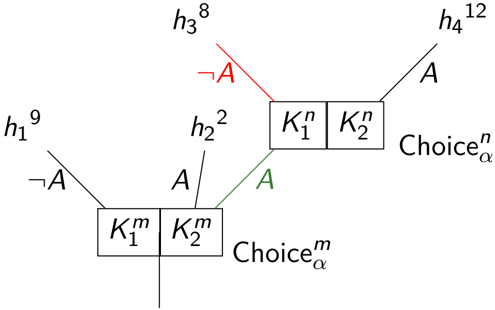
STIT logic
Semantics



Atomic propositions

- $m, h_3 \models A$?
- $n, h_3 \models A$?

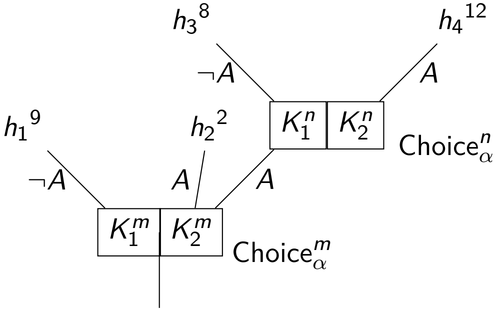
STIT logic
Semantics



Atomic propositions

- $m, h_3 \models A$
- $n, h_3 \not\models A$

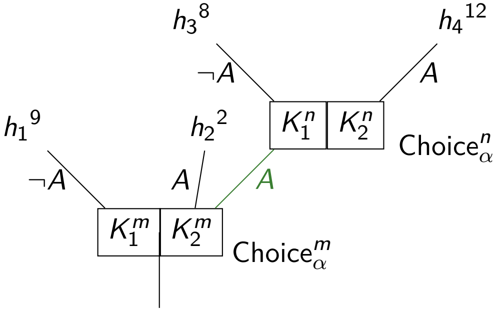
STIT logic
Semantics



Obligations

- $m, h_2 \models \bigcirc A$?
- $m, h_2 \models \bigcirc \neg A$?

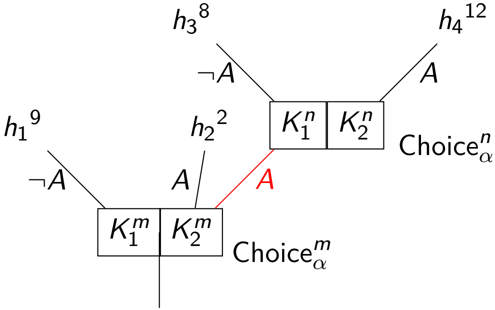
STIT logic
Semantics



Obligations

- $m, h_2 \models \bigcirc A$
- $m, h_2 \models \bigcirc \neg A$

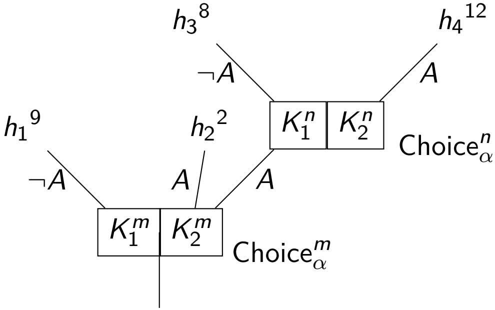
STIT logic
Semantics



Obligations

- $m, h_2 \models \bigcirc A$
- $m, h_2 \not\models \bigcirc \neg A$

STIT logic
Semantics

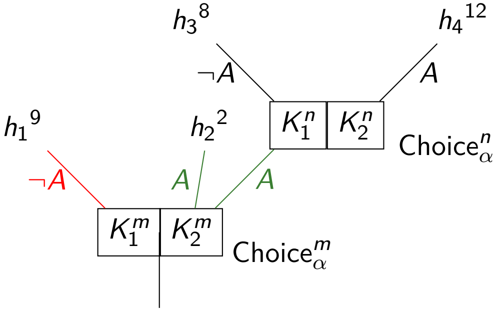


See-To-It-That

- $m, h_3 \models [\alpha\,\mathrm{cstit}\colon A]$?
- $m, h_1 \models [\alpha\,\mathrm{cstit}\colon A]$?

STIT logic

Semantics



See-To-It-That

- $m, h_3 \models [\alpha \,\text{cstit}\colon A]$
- $m, h_1 \not\models [\alpha \,\text{cstit}\colon A]$

Challenge 2: Non-Deterministic Actions

How can we deal with different possible outcomes caused by non-determinism?

Challenge 2: Non-Deterministic Actions

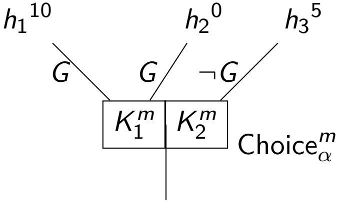
The gambling problem

- The agent is presented with two options
- If he gambles, he can either double or lose his bet
- If he does not gamble, he preserves his money

Challenge 2: Non-Deterministic Actions

The gambling problem

- The agent is presented with two options
- If he gambles, he can either double or lose his bet
- If he does not gamble, he preserves his money



How can we formulate that the agent should arrive at h_1?

Challenge 2: Non-Deterministic Actions

Optimality of non-deterministic actions

- The optimal path ends in h_1
- The agent should "ought to see to it that G"
- Intuitive formulation $\bigcirc[\alpha\, \text{cstit}\colon G]$ problematic since K_1^m is not necessarily better than K_2^m

Challenge 2: Non-Deterministic Actions

Optimality of non-deterministic actions

- The optimal path ends in h_1
- The agent should "ought to see to it that G"
- Intuitive formulation $\bigcirc[\alpha\, \text{cstit}\colon G]$ problematic since K_1^m is not necessarily better than K_2^m
- We cannot infer optimal actions from optimal histories
- An action may lead non-deterministically to worse histories

Challenge 2: Non-Deterministic Actions

More appropriate approach:

- Compare all possible histories
- K_1 is more optimal than K_2 if it has one history that is better than every history of K_2 and no worse history

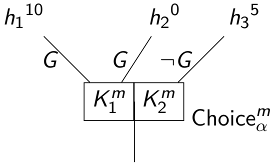
Challenge 2: Non-Deterministic Actions

More appropriate approach:

- Compare all possible histories
- K_1 is more optimal than K_2 if it has one history that is better than every history of K_2 and no worse history

New operator: $\odot[\alpha\,\text{cstit}\colon G]$

- Fulfilled if no action not leading to G is the optimal action
- Back to the gambling problem:
  Neither $\odot[\alpha\,\text{cstit}\colon G]$ nor $\odot[\alpha\,\text{cstit}\colon \neg G]$ hold for m



Challenge 3: Moral Luck

How do we deal with luck?

Challenge 3: Moral Luck

- Sometimes the consequences of our actions are outside of our control
- This may for example happen if the outcome is also dependent on someone else's actions

Challenge 3: Moral Luck

- Sometimes the consequences of our actions are outside of our control
- This may for example happen if the outcome is also dependent on someone else's actions
- How do we evaluate the formula $\odot[\alpha\,\mathsf{cstit}\colon A]$ if the outcome also depends on another agents decision?

Challenge 3: Moral Luck

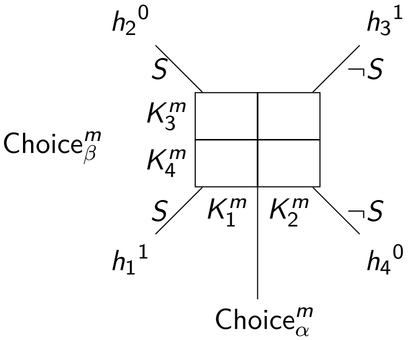
The driving example

- Two cars are driving towards each other on a small road
- The drivers (α and β) both have to decide whether to swerve to the side (S) or drive straight ($\neg S$)

Challenge 3: Moral Luck

The driving example

- Two cars are driving towards each other on a small road
- The drivers (α and β) both have to decide whether to swerve to the side (S) or drive straight ($\neg S$)
- If both decide on the same action, they will crash
- How do we deal with $\odot[\alpha\,\text{cstit}\colon S]$?



Challenge 3: Moral Luck

There are two views:

Challenge 3: Moral Luck

There are two views:

- Dominance act utilitarianism:

Challenge 3: Moral Luck

There are two views:

- Dominance act utilitarianism:
  - $\odot[\alpha \,\mathsf{cstit}\colon S]$ if all optimal outcomes guarantee S
  - In the example this is not the case, since $\neg S$ can also lead to an optimal outcome
  - The same goes for $\odot[\alpha \,\mathsf{cstit}\colon \neg S]$

Challenge 3: Moral Luck

There are two views:

- Dominance act utilitarianism:
  - $\odot[\alpha\,\mathsf{cstit}\colon S]$ if all optimal outcomes guarantee S
  - In the example this is not the case, since $\neg S$ can also lead to an optimal outcome
  - The same goes for $\odot[\alpha\,\mathsf{cstit}\colon \neg S]$
- Orthodox perspective

Challenge 3: Moral Luck

There are two views:

- Dominance act utilitarianism:
  - $\odot[\alpha\,\text{cstit}\colon S]$ if all optimal outcomes guarantee S
  - In the example this is not the case, since $\neg S$ can also lead to an optimal outcome
  - The same goes for $\odot[\alpha\,\text{cstit}\colon \neg S]$
- Orthodox perspective
  - We don't evaluate the formula at the time of m
  - We decide whether $\odot[\alpha\,\text{cstit}\colon S]$ should be true based on the outcome
  - Intuitively, we ask whether α *should have* swerved

Challenge 4: Procrastination

How do we deal with actions that will never be taken?

Challenge 4: Procrastination

- Agents can put off on following through with obligations
- Since they can do so over and over, tasks need deadlines
  - See also: Thread Starvation
- But what if agents still procrastinate?

Challenge 4: Procrastination

The story of Professor Procrastinate

- Professor Procrastinate is requested to write a review
- He is the best person available for writing the review
- He is known to procrastinate and will not actually write the review
- Should Professor Procrastinate accept the request?

Challenge 4: Procrastination
The story of Professor Procrastinate

- Professor Procrastinate is requested to write a review
- He is the best person available for writing the review
- He is known to procrastinate and will not actually write the review
- Should Professor Procrastinate accept the request?

We now run into a paradox:

Challenge 4: Procrastination

The story of Professor Procrastinate

- Professor Procrastinate is requested to write a review
- He is the best person available for writing the review
- He is known to procrastinate and will not actually write the review
- Should Professor Procrastinate accept the request?

We now run into a paradox:

- He ought to accept since him writing the review is the best scenario

Challenge 4: Procrastination

The story of Professor Procrastinate

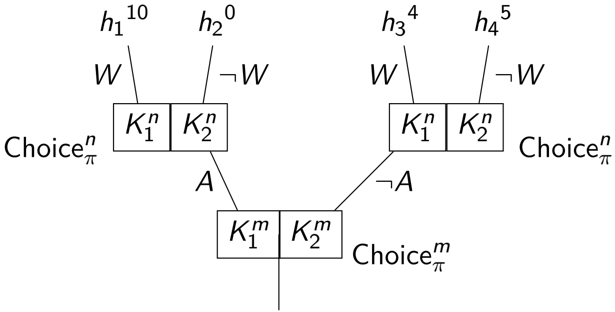
- Professor Procrastinate is requested to write a review
- He is the best person available for writing the review
- He is known to procrastinate and will not actually write the review
- Should Professor Procrastinate accept the request?

We now run into a paradox:

- He ought to accept since him writing the review is the best scenario
- He ought to decline since he will not actually write the review, which leads to the worst outcome

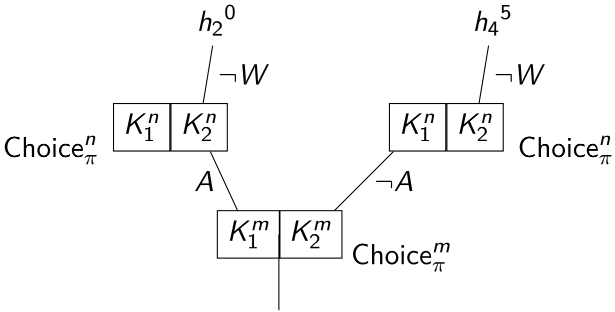
Challenge 4: Procrastination

The solution: Just prune the branch!



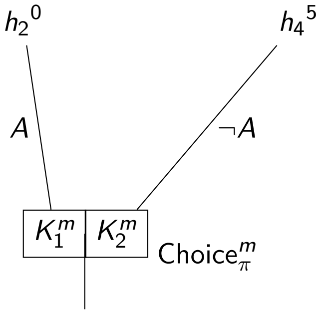
Challenge 4: Procrastination

The solution: Just prune the branch!



Challenge 4: Procrastination

The solution: Just prune the branch!



Summary

- For the gambling problem we can circumvent the problem by modifying the semantics
  - Instead of considering only the best outcome we consider every possible outcome
- We we're unable to solve the driving example
  - Instead two perspectives on how the semantics should work were given
  - Which one is to be used is a design decision
- The problem of procrastination can be solved by using additional knowledge to cut branches from the STIT-tree