

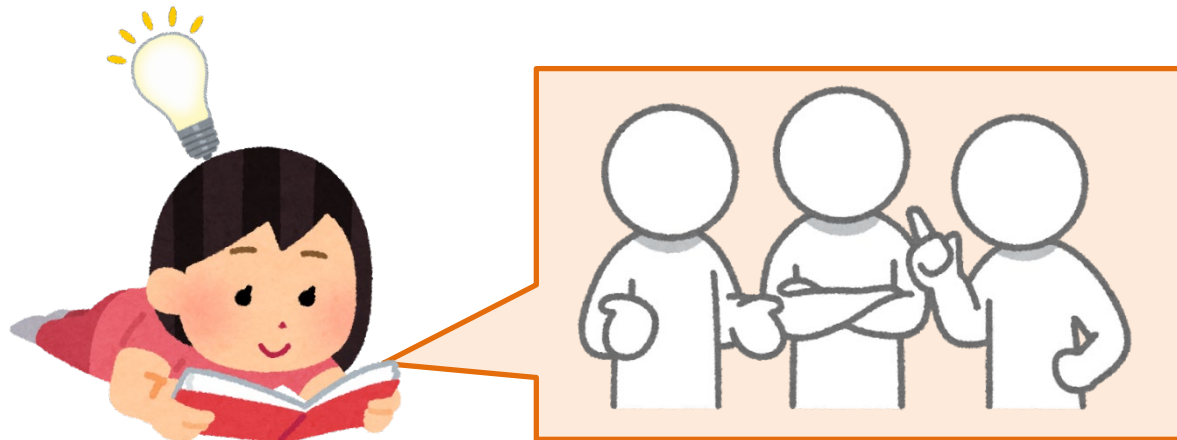
(アイデア部門作品3)
**誰が犯人か？を自動的に議論する
マルチエージェントアーキテクチャ**

白松研 feat. 59 (張翔宇, 白松俊, 神谷晃, 渡辺賢)

名古屋工業大学 工学部情報工学科

キーアイデア

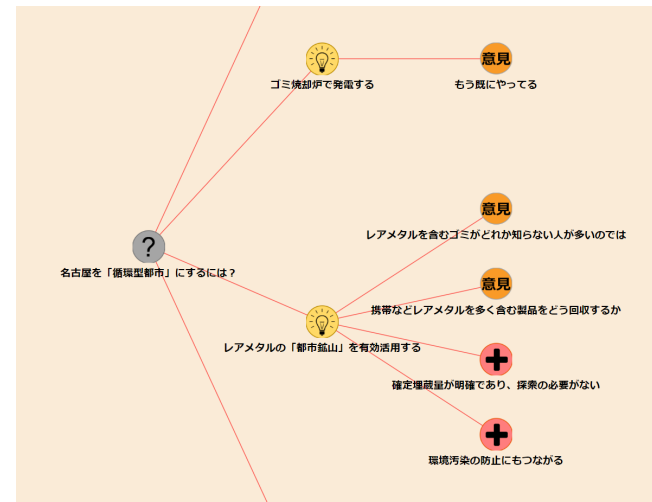
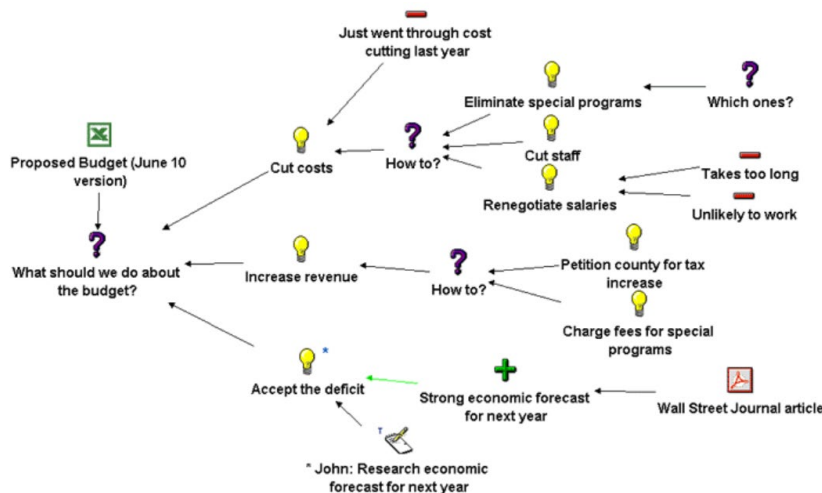
- 誰が犯人かを複数エージェントで議論することで、**説明性**を持つ推理AIを実現
- なぜ？
 - 推理小説には、犯人を言い当てる探偵だけでなく、間違った推理を披露する助手役 (サイドキック) も重要
 - それぞれの仮説を検証するような議論の場面を挿入することで、読者に**事件背景や推理の根拠を理解**させる働き



どう着想したか

これまでの議論支援研究の経緯から着想

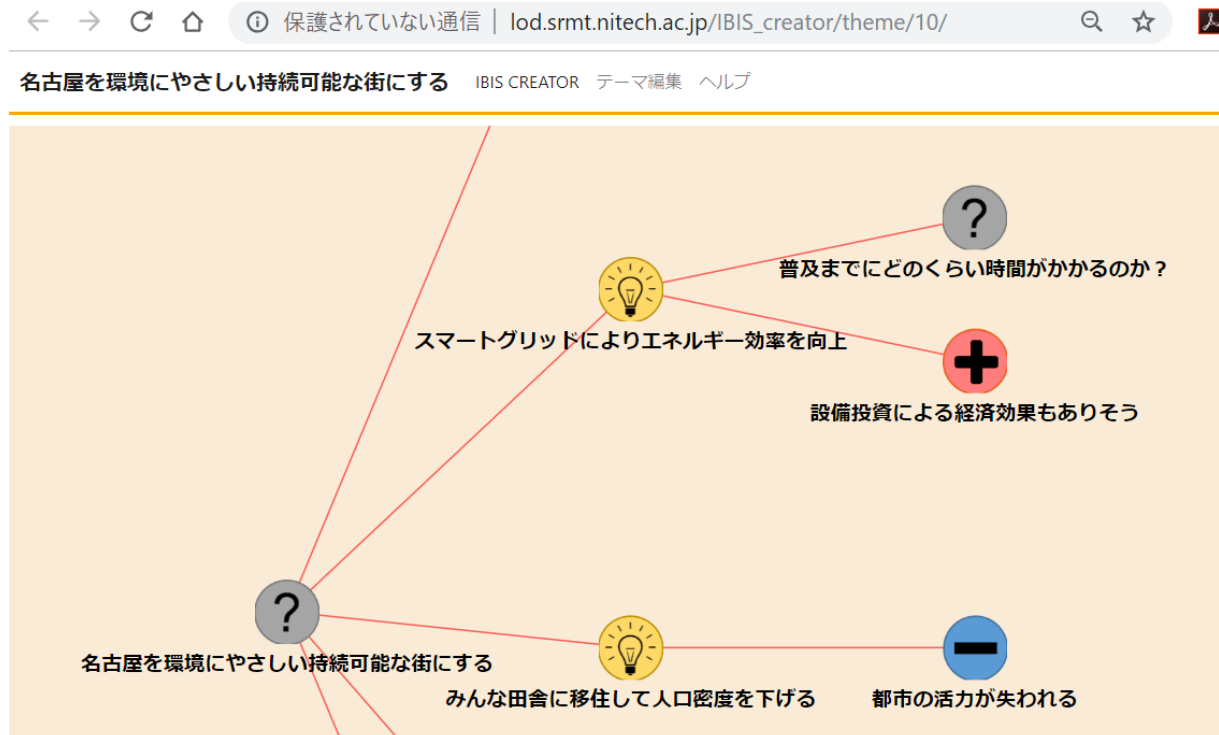
- 議論を通じた合意形成支援の研究 [白松+ 18]
 - 議論内容をIBIS (Issue-based Information System) で理解
 - 課題, アイデア, アイデアの利点／欠点をノードとする木構造



- 「犯人が誰か」という仮説をアイデアノードにすれば,
IBIS構造は仮説の根拠を構造化するのに最適なのは？

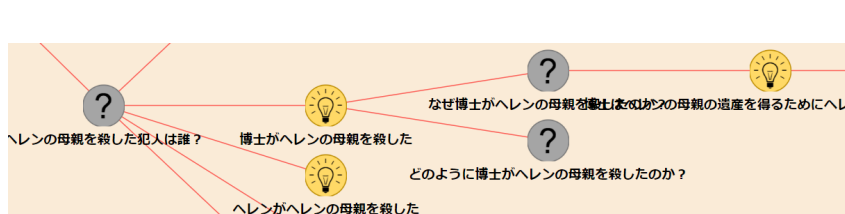
IBIS構造を扱うツール

- IBIS CREATOR http://lod.srmt.nitech.ac.jp/IBIS_creator/
 - IBIS構造を作成・編集できるWebアプリ（白松研 神谷が開発）
 - 作ったIBIS構想はLinked Open Data化される
 - 簡単なWeb APIを備え，外部システムから操作可能



推理のための議論手順の概要

- 「誰が犯人か」という仮説を担当する議論エージェント
 - ナレッジグラフ中の登場人物 x ごとに「 x が犯人」という仮説を立て、その仮説の説明を試みる議論エージェントを割り当てる
- 仮説の詳細を質問するファシリテータエージェント
 - 「 x はどのように殺したのか？」のような質問をIBIS構造に追加
- 各議論エージェントはその質問の答えとなる仮説を生成
 - 生成された仮説は、IBIS構造に追加されると同時に、
 - 各議論エージェントが保持するナレッジグラフを補完

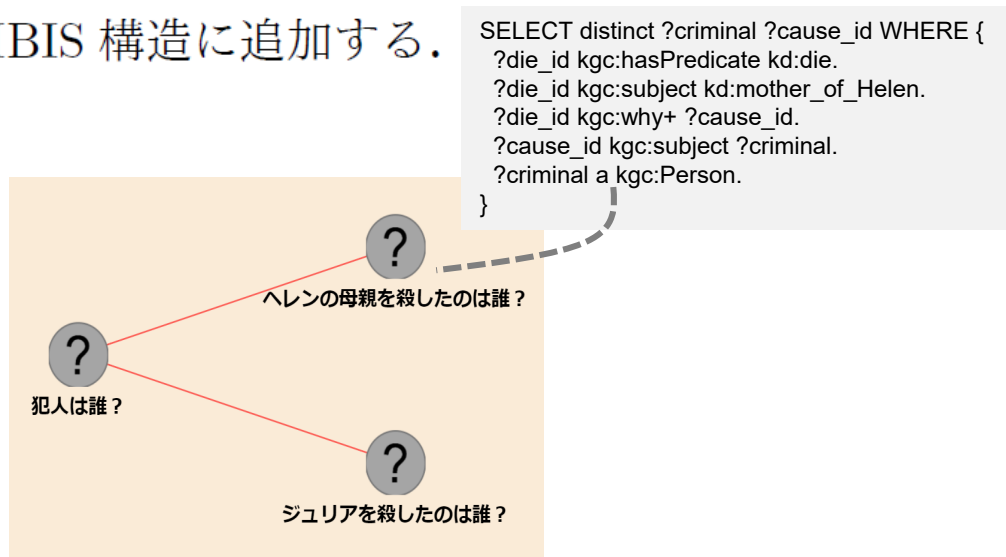


- ファシリテータエージェントが説明の整合性を評価
 - 各議論エージェントが生成した説明の整合性を評価し、最終的に評価の高かった説明を出力

議論手順の詳細 (1/4)

10ステップの手順

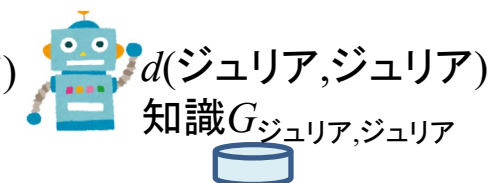
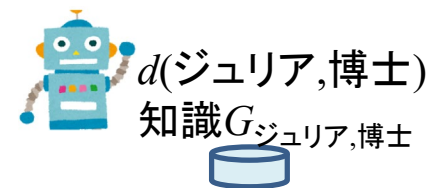
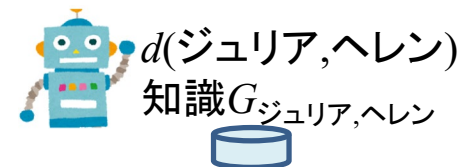
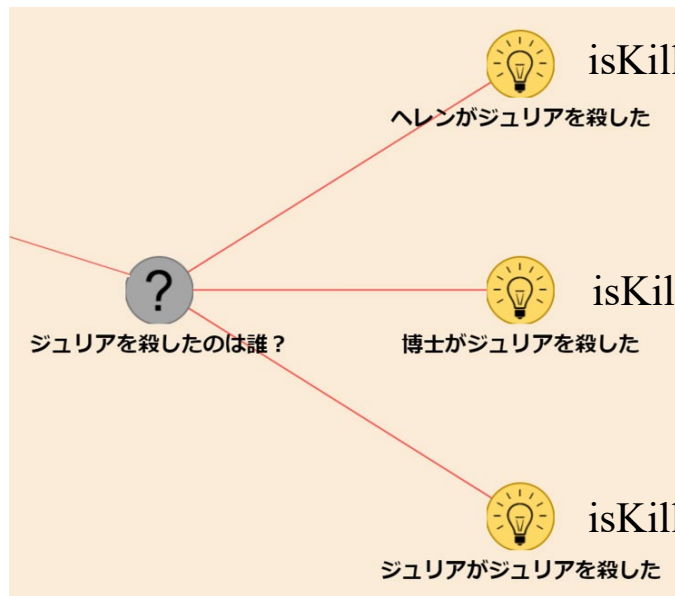
1. あらかじめ、ナレッジグラフ G に必要な常識的知識を追加しておく.
2. ナレッジグラフ G 中で死亡した犠牲者 v ごとに、犯人を問う Issue ノードを IBIS 構造に追加する.



3. 「犯人 x が犠牲者 v を殺した」という複数の仮説 $\text{isKilledBy}(v, x)$ を生成. これは死亡者 v の集合と登場人物 x の集合の直積（つまり組み合わせ）から生成される. その仮説 $\text{isKilledBy}(v, x)$ を IBIS 構造のアイデアノードとして追加する.

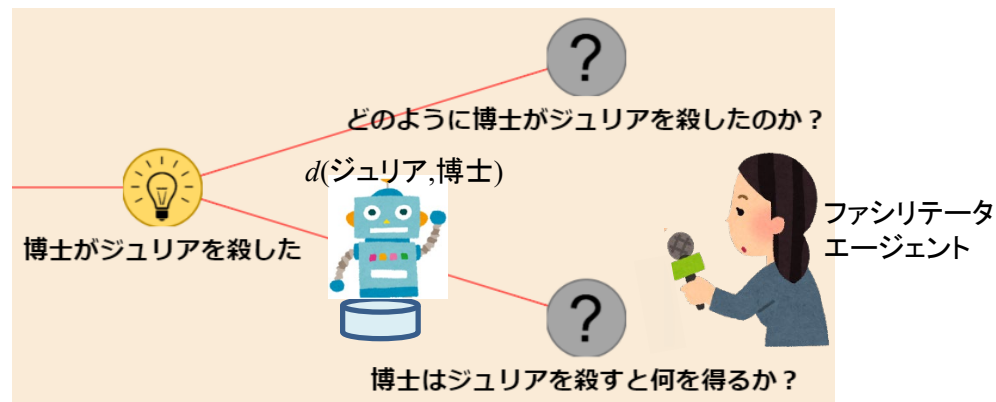
議論手順の詳細 (2/4)

4. 各仮説 $\text{isKilledBy}(v, x)$ ごとに、その仮説を導出する説明を試みる
議論エージェント $d(v, x)$ を割り当てる．また、議論エージェント
 $d(v, x)$ の知識 $G_{v,x}$ に、上記 1. で常識を補完したナレッジグラフ G
をコピーした上で仮説 $\text{isKilledBy}(v, x)$ を追加する．

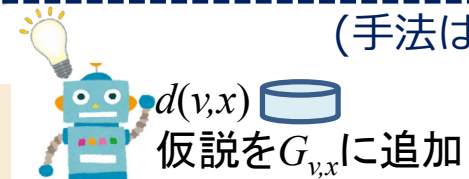
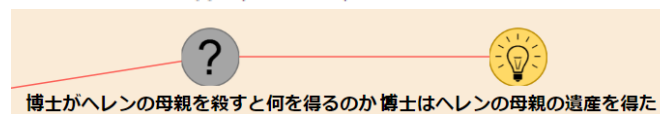


議論手順の詳細 (3/4)




5. 1つのファシリテータエージェントが、すべての仮説に「どのように x が v を殺したのか?」「 x は v を殺すと何を得るか?」を問いかける (IBIS 構造に Issue ノードを追加する)

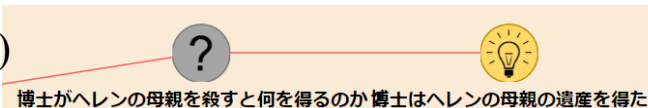
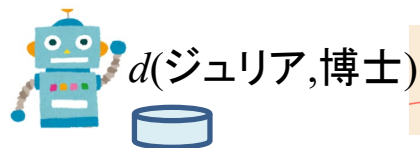


6. 議論エージェント $d(v, x)$ はそれぞれ、「どのように殺したか」に対する仮説 $how(v, x)$, 「殺すと何を得るか」に対する仮説 $gain(v, x)$ をそれぞれ仮説推論などにより生成を試み、生成できた場合は追加する. その過程で, 仮説 $how(v, x)$, $gain(v, x)$ を支持する説明を生成し, $G_{v,x}$ および IBIS 構造に追加する. (手法は未検討)



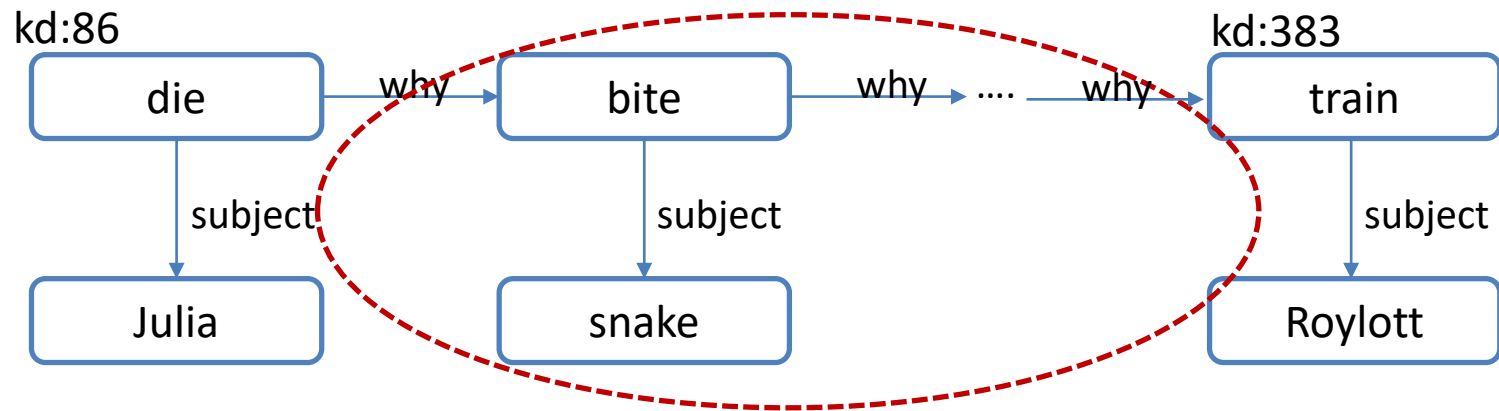
議論手順の詳細 (4/4)

7. 議論エージェント $d(v, x)$ はそれぞれ, x 以外を犯人とみなす仮説 $\text{isKilledBy}(v', x')$ の反証を試みる. すなわち, 仮説 $\text{how}(v', x')$, $\text{gain}(v', x')$ に対する 反対意見の生成を試み, 生成できた場合は $G_{v', x'}$ および IBIS 構造に追加する.(手法は未検討)
-   **反論**を $G_{v', x'}$ に追加 
8. ファシリテータエージェントは, 仮説 $\text{isKilledBy}(v, x)$ に対応する $G_{v, x}$ について, 整合性スコア $\text{cnsstcy}(v, x)$ を評価する.
(定義は未検討)
9. ファシリテータエージェントは, 被害者 v ごとに整合性スコア $\text{cnsstcy}(v, x)$ を最大化する犯人候補 $x_v = \arg \max_x \{\text{cnsstcy}(v, x)\}$ を真犯人として選択する.
10. 仮説が選ばれた議論エージェント $d(v, x_v)$ は, 被害者 v ごとに真犯人が x_v であることを説明するため, 上記 6. で生成された仮説 $\text{isKilledBy}(v, x_v)$ を支持する説明を出力する.



想定する仮説生成例

- 例えば「博士はどのようにジュリアを殺したか？」に関する仮説は、以下のようなkgc:whyの連鎖の構造で生成



- このような仮説でナレッジグラフが補完されていれば、以下のようなSPARQLクエリで犯人 x と殺人行為 $cause_id$ を検索可能

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX kgc: <http://kgc.knowledge-graph.jp/ontology/kgc.owl#>
PREFIX kd: <http://kgc.knowledge-graph.jp/data/SpeckledBand/>
SELECT distinct ?x ?cause_id WHERE {
  ?die_id kgc:hasPredicate kd:die.
  ?die_id kgc:subject kd:Julia.
  ?die_id kgc:why+ ?cause_id.
  ?cause_id kgc:subject ?x.
  ?x rdf:type kgc:Person.
}
```

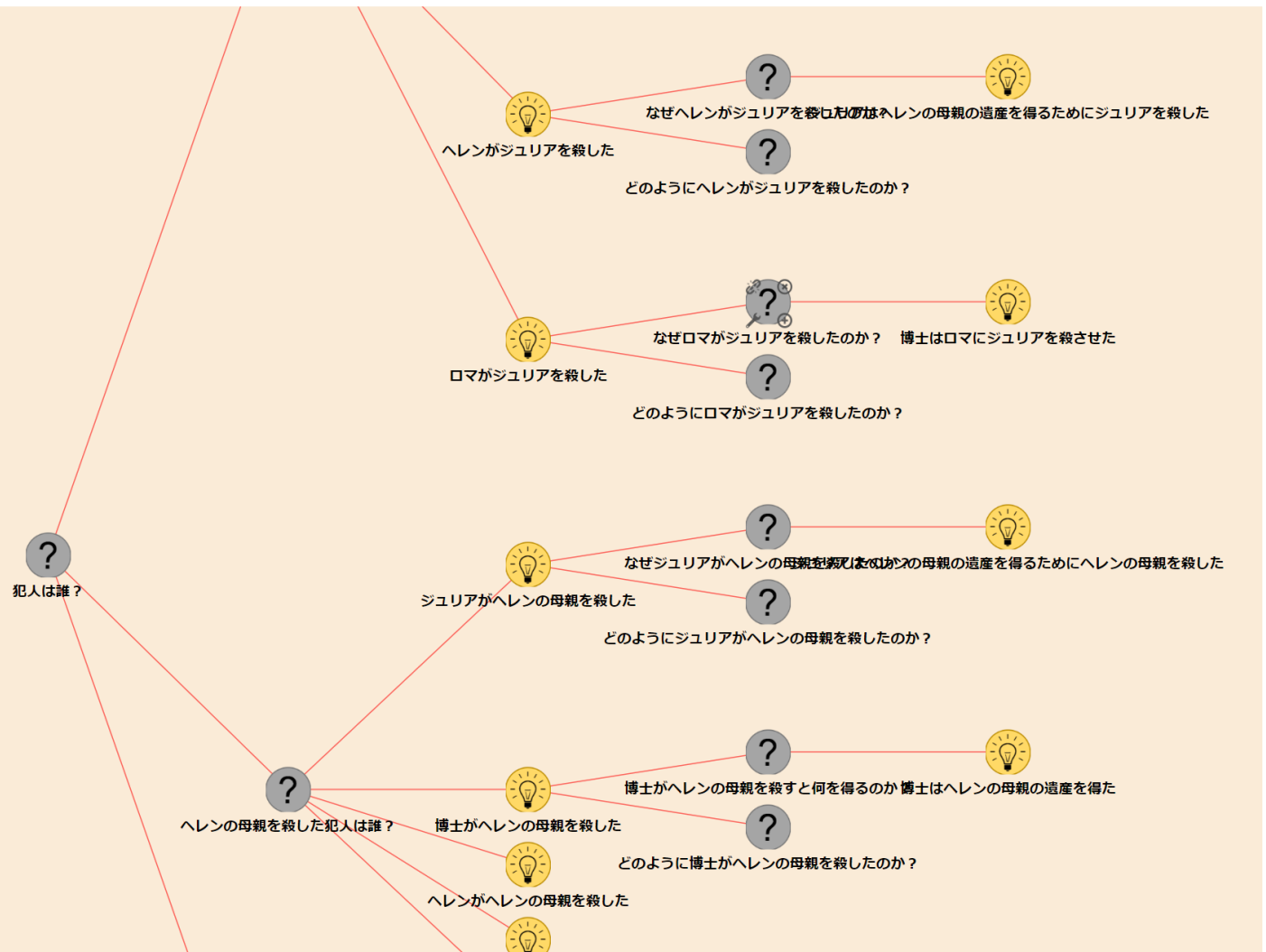
生成すべきIBIS構造を手動で検討中...

http://lod.srmt.nitech.ac.jp/IBIS_creator/theme/22/

← → ↺ ⌂ ⓘ 保護されていない通信 | lod.srmt.nitech.ac.jp/IBIS_creator/theme/22/

🔍 ☆ 📄 🖨

「まだらの紐」の犯人は誰か IBIS CREATOR テーマ編集 ヘルプ



問題点ともう1つの目論見

- 「全犠牲者を殺した犯人が同一人物」と仮定したくない
 - 😊 仮定は少ないほど良いし、同一人物とは限らない
 - 😞 しかしそのせいで、全犠牲者に関する仮説間の整合性をどう検証するかが課題に

9. ファシリテータエージェントは、被害者 v ごとに整合性スコア $\text{cnsstcy}(v, x)$ を最大化する犯人候補 $x_v = \arg \max_x \{\text{cnsstcy}(v, x)\}$ を真犯人として選択する.

犠牲者 v ごとに全仮説の整合性を最大化
(他の犠牲者に関する仮説との整合性は見えない)

- 目論見：議論支援研究にも役立つ知見が得られないか？
 - これまでの研究ではファシリテータエージェントのみに注力
 - 仮説生成手法が確立できれば、人間の議論に交じって
意見の根拠をサジェストする議論エージェントが作れるかも？