Towards Smart Cities at the Mental Level via Tentacular AI (TAI) Agents

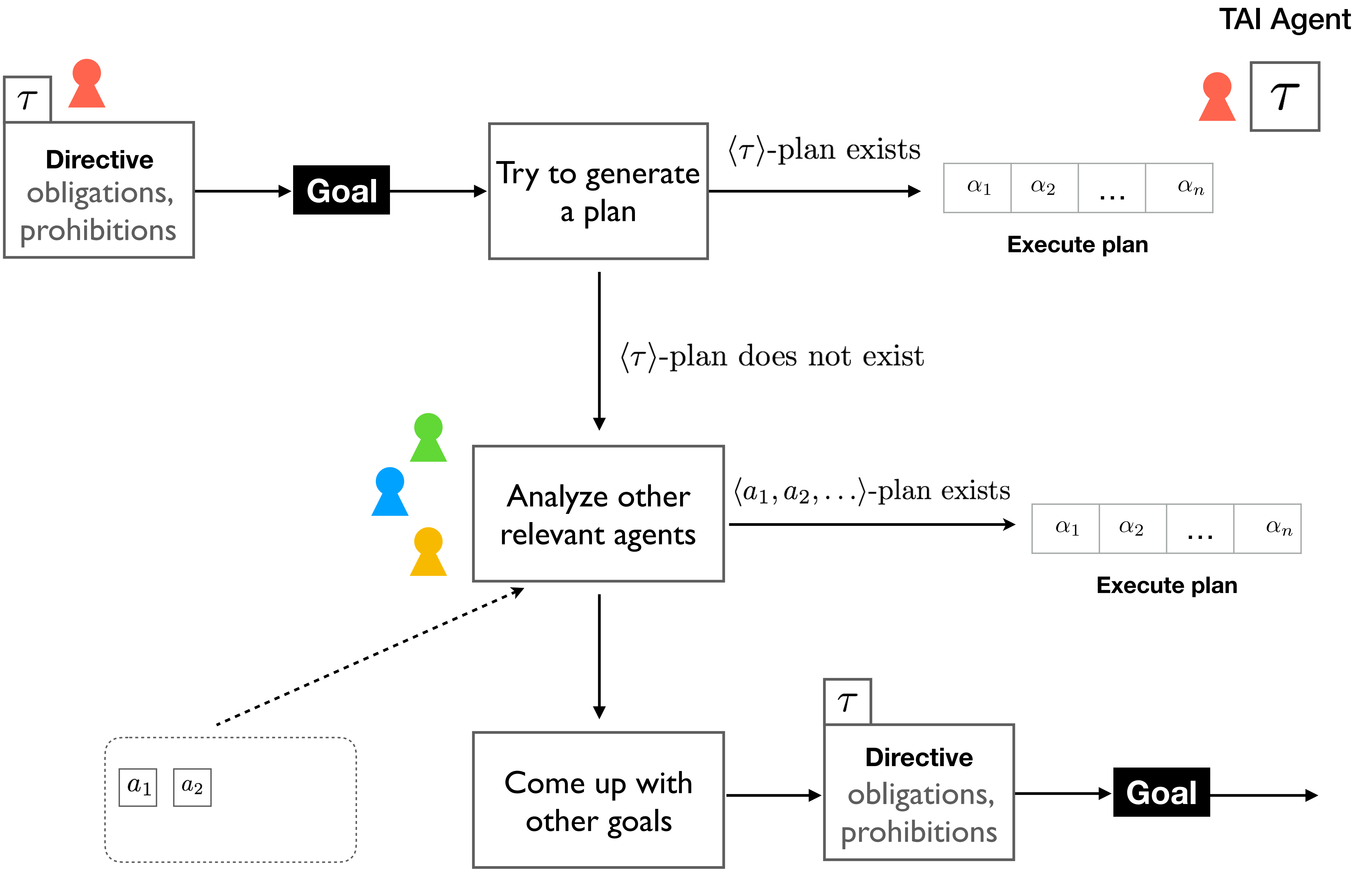
Selmer Bringsjord[1]
Naveen Sundar Govindarajulu[1]
Mike Giancola[1]
Paul Mayol[1]
Atriya Sen[1]
Biplav Srivastava[2]
Kartik Talamadupula[2]

1 RAIR Rensselaer AI and Reasoning Lab

2 IBM

Smart Cities Workshop — ICRES
July 30, 2019

Sponsored by Office of Naval Research Science & Technology, AFOSR Air Force Office of Scientific Research, Rensselaer


TAI Agent
τ
τ
Directive
obligations,
prohibitions
Goal
Try to generate
a plan
$\langle \tau \rangle$-plan exists
α_1 α_2 ... α_n
Execute plan
$\langle \tau \rangle$-plan does not exist
Analyze other
relevant agents
$\langle a_1, a_2, \ldots \rangle$-plan exists
α_1 α_2 ... α_n
Execute plan
a_1 a_2
Come up with
other goals
τ
Directive
obligations,
prohibitions
Goal

A TAI agent must be …

1. capable of **problem-solving** via planning, reasoning, learning, communicating;

2. capable of solving at least important instances of problems that are at and/or "above" **Turing-unsolvable problems**;

3. able to supply justification, **explanation**, and certification of supplied solutions, how they are arrived at, and that these solutions are safe/ethical;

4. capable of ***theory-of-mind*-level** reasoning, planning, learning, and communicating;

5. capable of **creativity**, minimally to the level of so-called "MacGyveresque", or ***m***-*creativity*; and

6. in possession of "**tentacular**" power wielded throughout **I/IoT**, edge computing, cyberspace, etc.

Formal Background

- Deontic Cognitive Event Calculus $\mathcal{DCEC}$
  - First Order Multi-Operator Modal Logic
  - Well-Defined Syntax & Inference Schemata
  - Based on Natural Deduction

Sort	Description
Agent	Human and non-human actors.
Time	The Time type stands for time in the domain. E.g. simple, such as t_i, or complex, such as $birthday(son(jack))$.
Event	Used for events in the domain.
ActionType	Action types are abstract actions. They are instantiated at particular times by actors. Example: eating.
Action	A subtype of Event for events that occur as actions by agents.
Fluent	Used for representing states of the world in the event calculus.

Syntax

$$S ::= \textsf{Agent} \mid \textsf{ActionType} \mid \textsf{Action} \sqsubseteq \textsf{Event} \mid \textsf{Moment} \mid \textsf{Fluent}$$

$$f ::= \begin{cases} action : \textsf{Agent} \times \textsf{ActionType} \rightarrow \textsf{Action} \\ initially : \textsf{Fluent} \rightarrow \textsf{Formula} \\ holds : \textsf{Fluent} \times \textsf{Moment} \rightarrow \textsf{Formula} \\ happens : \textsf{Event} \times \textsf{Moment} \rightarrow \textsf{Formula} \\ clipped : \textsf{Moment} \times \textsf{Fluent} \times \textsf{Moment} \rightarrow \textsf{Formula} \\ initiates : \textsf{Event} \times \textsf{Fluent} \times \textsf{Moment} \rightarrow \textsf{Formula} \\ terminates : \textsf{Event} \times \textsf{Fluent} \times \textsf{Moment} \rightarrow \textsf{Formula} \\ prior : \textsf{Moment} \times \textsf{Moment} \rightarrow \textsf{Formula} \end{cases}$$

$$t ::= x : S \mid c : S \mid f(t_1, \ldots, t_n)$$

$$\phi ::= \begin{cases} q : \textsf{Formula} \mid \neg\phi \mid \phi \land \psi \mid \phi \lor \psi \mid \forall x : \phi(x) \mid \\ \mathbf{P}(a, t, \phi) \mid \mathbf{K}(a, t, \phi) \mid \\ \mathbf{C}(t, \phi) \mid \mathbf{S}(a, b, t, \phi) \mid \mathbf{S}(a, t, \phi) \mid \mathbf{B}(a, t, \phi) \\ \mathbf{D}(a, t, \phi) \mid \mathbf{I}(a, t, \phi) \\ \mathbf{O}(a, t, \phi, (\neg) happens(action(a^*, \alpha), t')) \end{cases}$$

Inference Schemata (Fragment)

$$\frac{\mathbf{K}(a, t_1, \Gamma), \;\; \Gamma \vdash \phi, \;\; t_1 \leq t_2}{\mathbf{K}(a, t_2, \phi)} \; [I_{\mathbf{K}}]$$

$$\frac{\mathbf{B}(a, t_1, \Gamma), \;\; \Gamma \vdash \phi, \;\; t_1 \leq t_2}{\mathbf{B}(a, t_2, \phi)} \; [I_{\mathbf{B}}]$$

$$\frac{\mathbf{K}(a, t, \phi)}{\phi} \; [I_4] \qquad \frac{t < t', \;\; \mathbf{I}(a, t, \psi)}{\mathbf{P}(a, t', \psi)} \; [I_{13}]$$

$$\frac{\mathbf{B}(a, t, \phi) \;\; \mathbf{B}(a, t, \mathbf{O}(a, t, \phi, \chi)) \;\; \mathbf{O}(a, t, \phi, \chi)}{\mathbf{K}(a, t, \mathbf{I}(a, t, \chi))} \; [I_{14}]$$

Differences from Prior Work

Differences from Prior Work

- Expressivity

  1. Can handle **states of minds** of other agents as goals.

  2. **Quantifiers** enable succinct representation of large (infinite) domains.

- Absence/scarcity of standardized domain knowledge.

  - Today's ML can't solve anomalous problems.

```
{:name        "The Purloined Letter"
 :description "Dupin's reasoning as he goes through the case"

 :assumptions {1 (Believes! g (hide m elaborate))
               2 (Believes! d (or (hide m elaborate) (hide m plain)))
               3 (Believes! m (Believes! g (hide m elaborate)))
               4 (if (Believes! m (Believes! g (hide m elaborate))) (hide m plain))
               5 (if (Believes! m (Believes! g (hide m plain))) (hide m elaborate))
               6 (Believes! m (Believes! g (hide m elaborate)))
               7 (Believes! d (if (Believes! m (Believes! g (hide m elaborate))) (hide m plain)))
               8 (Believes! d (if (Believes! m (Believes! g (hide m plain))) (hide m elaborate)))
               9 (Believes! d (Believes! m (Believes! g (hide m elaborate))))}


 :goal (Believes! d (hide m plain))}
```

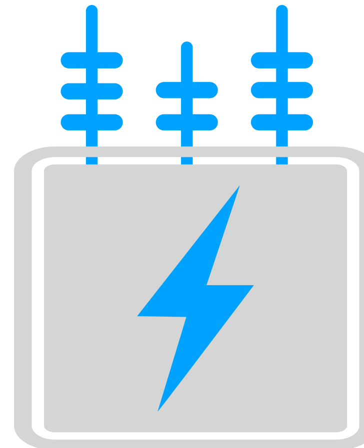
These model an infinite domain

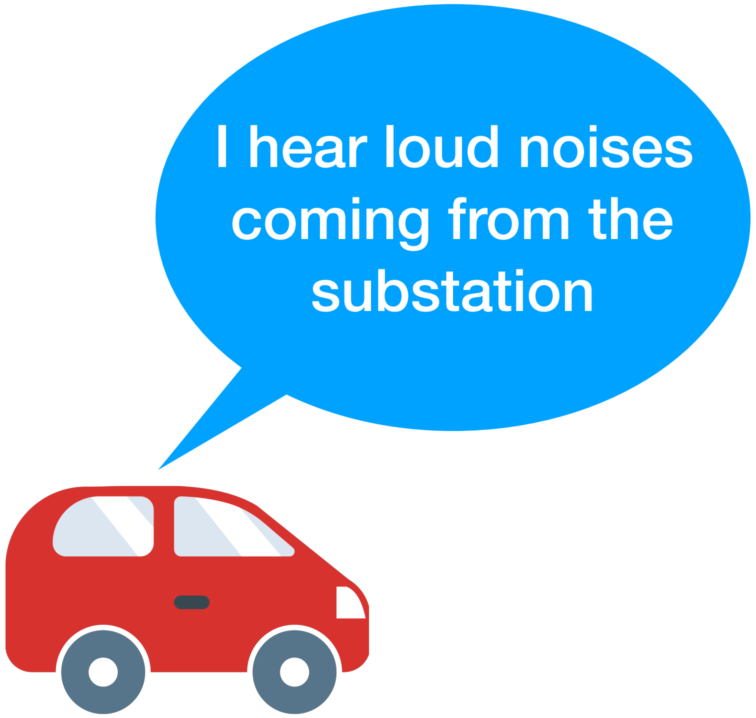
$$\forall x \exists y \mathbf{R}(x, y) \wedge$$
$$\forall x, y \neg (\mathbf{R}(x, y) \wedge \mathbf{R}(y, x)) \wedge$$
$$\forall x, y, z (\mathbf{R}(x, y) \wedge \mathbf{R}(y, z)) \rightarrow \mathbf{R}(x, z)$$

Applications to Smart Cities

Scenario 1: Power Outage

Scenario I: Power Outage

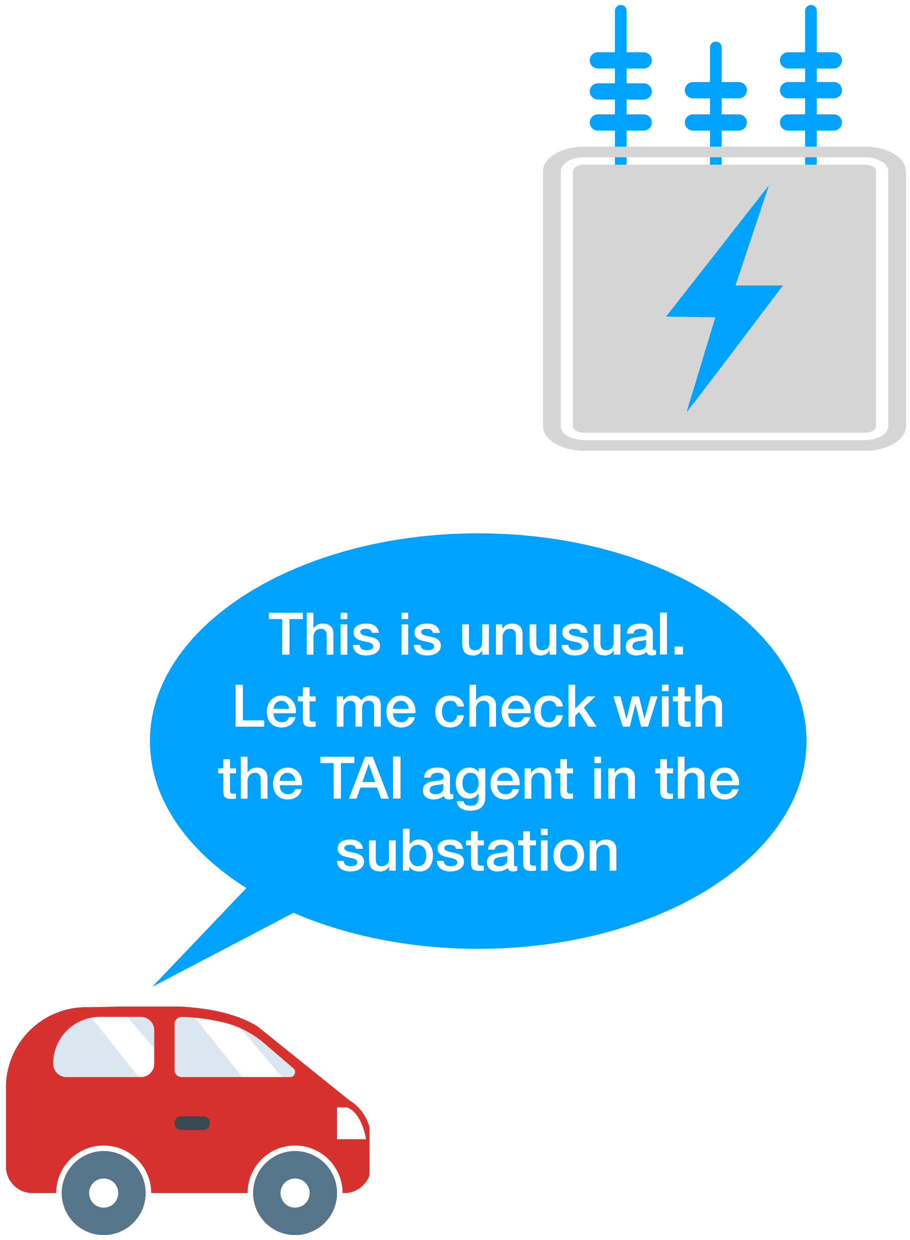




$$\mathbf{B}(a_c, t_0, noisy(substation) \rightarrow unusual)$$ f$_1$

$$\mathbf{P}(a_c, t_1, noisy(substation))$$ f$_2$

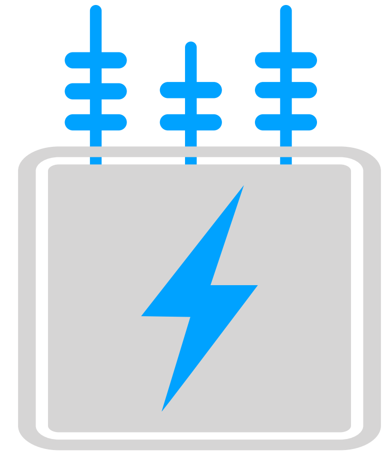
Scenario 1: Power Outage

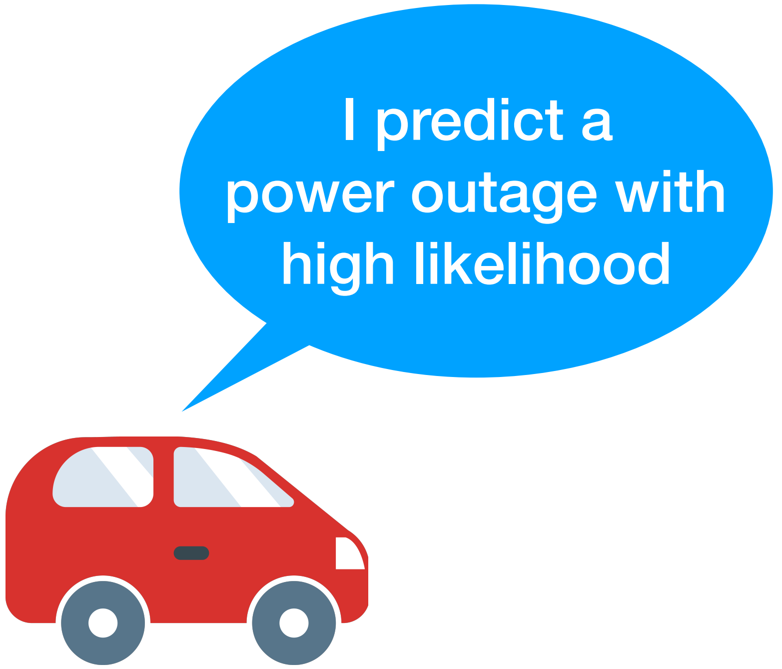


$$\forall t \ : \ \mathbf{O}\begin{pmatrix} a_c, t, unusual, \\ \qquad happens(action(a_c, consult(a_t)), t+1) \end{pmatrix} \quad \boxed{f_3}$$

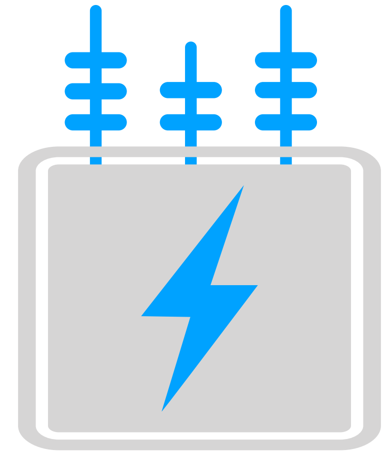
$$\forall t \ : \ \mathbf{B}(a_c, t, \boxed{f_3}\ ) \quad \boxed{f_4}$$

Scenario 1: Power Outage





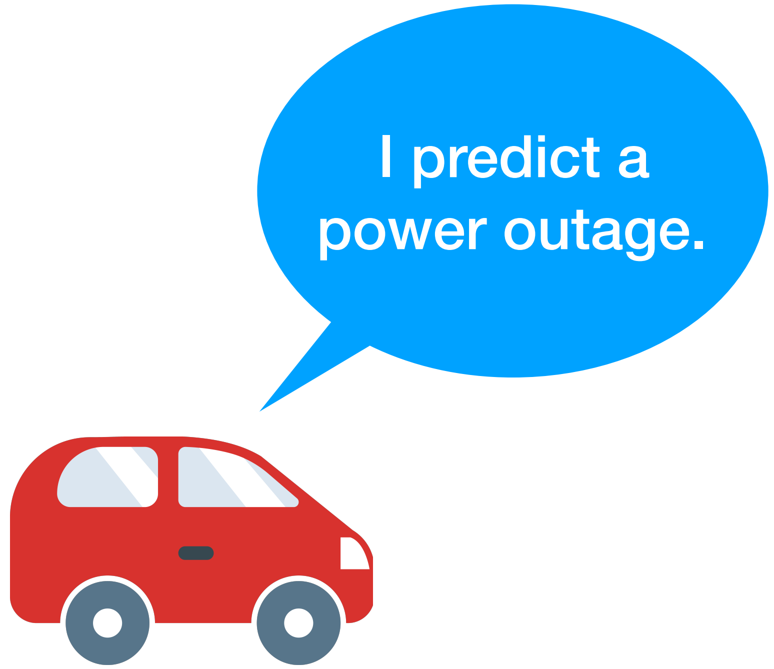
$$\forall a : \left(\begin{aligned} & happens(action(a, consult(a_t)), t_2) \\ & \quad \rightarrow \mathbf{B}(a, t_3, outage) \end{aligned}\right) \quad f_5$$



$$\forall t \;:\; \mathbf{O}\begin{pmatrix} a_c, t, outage, \\ \quad \mathbf{S}(a_c, a_h, outage, t+1) \end{pmatrix} \quad \boxed{f_6}$$

$$\forall t \;:\; \mathbf{B}(a_c, t, \boxed{f_6}\,) \quad \boxed{f_7}$$

Scenario 1: Power Outage



$$\mathbf{S}(a_c, a_h, outage, t_4)$$ f8

Scenario 1: Power Outage

$$\forall t \; : \; \mathbf{O}\begin{pmatrix} a_h, t, outage, \\ \quad \forall s \; : \; quantity(s) > 0 \end{pmatrix}$$

- Ask smart refrigerator which supplies are necessary
- Place an order online
  - If possible, request same-day drone delivery
  - If not, find a store along human’s route home, request store pickup
    - Tell car to reroute human to store before returning home

Scenario II:
Office Optimization

1. Scenario
   1. A new employee (*eric*) has been in office for orientation for 3 days
   2. Today, the boss (*bill*) has to assign *eric* a permanent desk
   3. *bill* employs TAI to help him determine the best desk for *eric*.

2. Therefore, TAI's goal is

$$\exists d \ : \ \begin{bmatrix} \mathbf{I}\left(bill, t^*, happens(action(eric, assigned(d)), t^* + 1)\right) \\ \wedge \ \forall e \ \begin{bmatrix} (happens(action(eric, assigned(d)), t^* + 1) \wedge pos\_emotion(e, t^* + 1)) \\ \rightarrow pos\_emotion(e, t^* + 2) \end{bmatrix} \\ \wedge \ \text{HighlyProductive(eric)} \end{bmatrix}$$

3. TAI determines, in a *tentacular* fashion, that desk 3 would be best for *eric*
   1. The building's smart thermostat was set to a temperature that most employees found comfortable. However, *eric*'s smart watch indicated to TAI that *eric* was cold. Desk 3 is close to the heater.
   2. The company's orientation software found that *eric* was most productive when at a desk away from loud-speaking employees. Desk 3 is in a quiet area (as determined by microphones in those areas).

A Comment on Ethics & Privacy

- Clearly could enable nefarious parties to
  - (Potentially unwittingly) influence humans
  - violate the privacy of humans, both consenting users and not
- Ethical/moral reasoning are at the forefront of DCEC/TAI
  - Include statements in TAI's contract to guarantee privacy of users' data
  - Utilize cutting-edge cryptographic methods for hiding sensitive data from TAI
    - homomorphic encryption
    - zero-knowledge proofs
    - differential privacy

Future Work

- Micro-domains and synthetic data in home automation, smart buildings and cities
- **Integration** of one more more common services with our reasoning systems and planners with a focus on:
  - Safety, Ethics, Efficiency
- Visual Question Answering (VQA) with Justifications
  - Human-understandable justifications from visual scenes

Thank You

Links

- TAI Project: http://kryten.mm.rpi.edu/TAI/tai.html
- Reasoner: https://github.com/naveensundarg/prover