iFM23

CAN-verify: Verification Tool for BDI Agents

Mengwei Xu (Newcastle), Thibault Rivoalen (ENAC),

Blair Archibald, (Glasgow), Michele Sevegnani (Glasgow)

Research Background

Engineered systems are becoming more complex and increasingly autonomous

Research Background

Engineered systems are becoming more complex and increasingly autonomous

roomba vacuum cleaner

care-o-bot robotic home assistant

parrot bebop 2 drone

Research Background

Autonomous Agent Systems

Benefits

improving efficiency **optimizing resource use** **…** **reducing human exposure**

Research Motivations

Engineered systems are becoming more complex and increasingly autonomous

the high complexity increases the probability of human design errors

Research Motivations

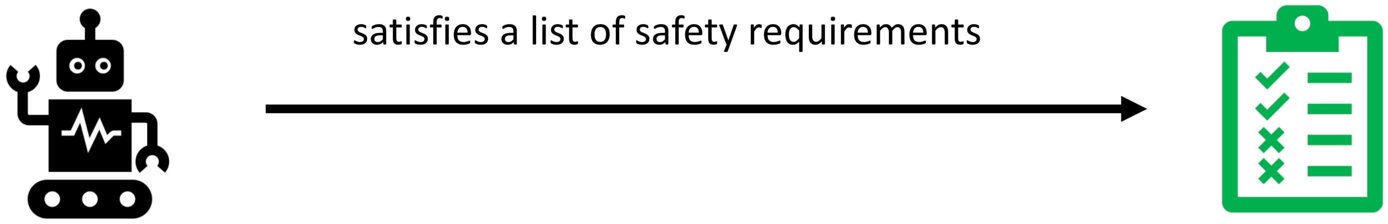
Engineered systems are becoming more complex and increasingly autonomous

the high complexity increases the probability of human design errors

the growing autonomy raises trustworthiness issues

Research Questions

Are these autonomous safe to deploy?



Research Questions

Are these autonomous safe to deploy?

- if an agent can successfully complete a mission
- if so, what chance is it under environmental uncertainty
- if there exists an optimal strategies for agent decision making while satisfying the properties

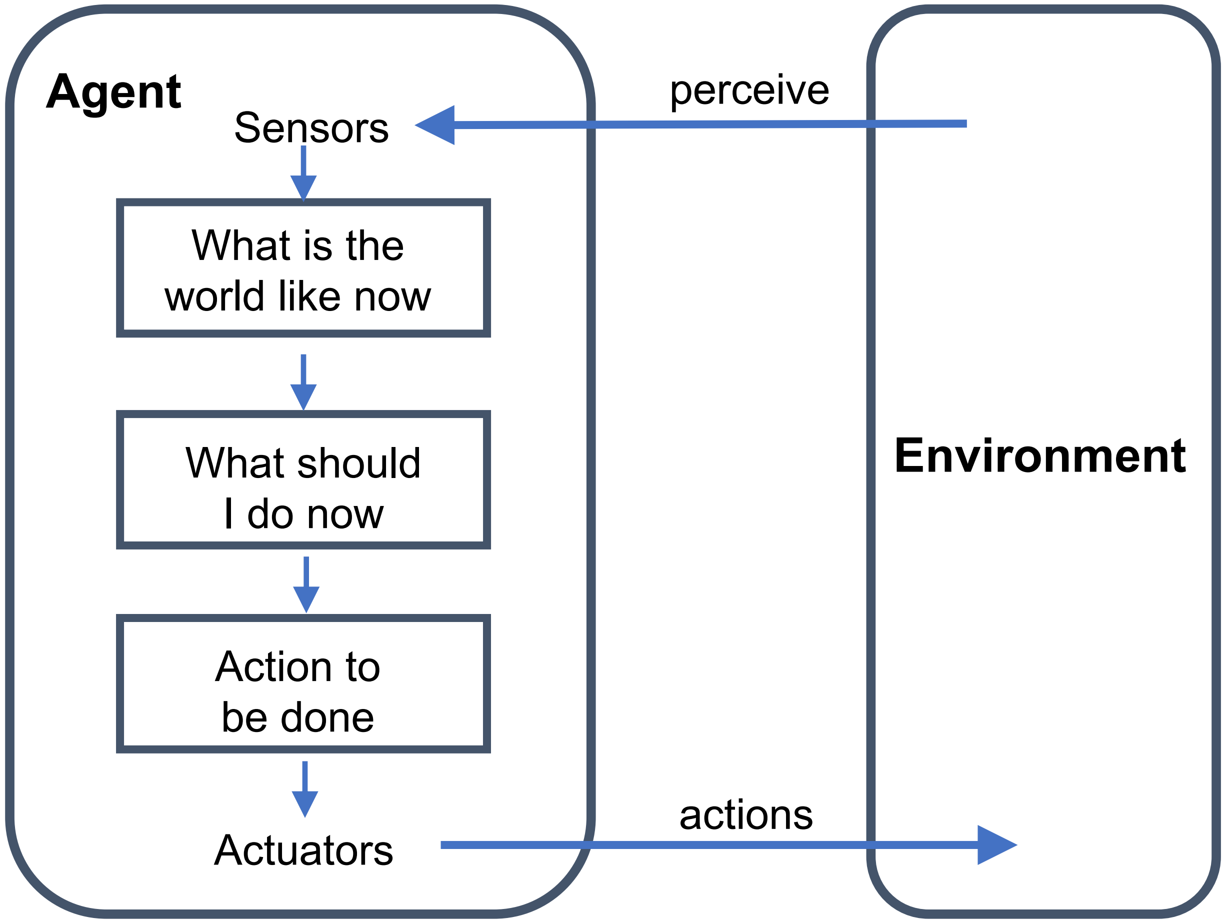
Autonomous Agent

An entity

which **perceives** its environment,
which **deliberates** accordingly,
which **takes actions** autonomously,

in order to achieve some objectives

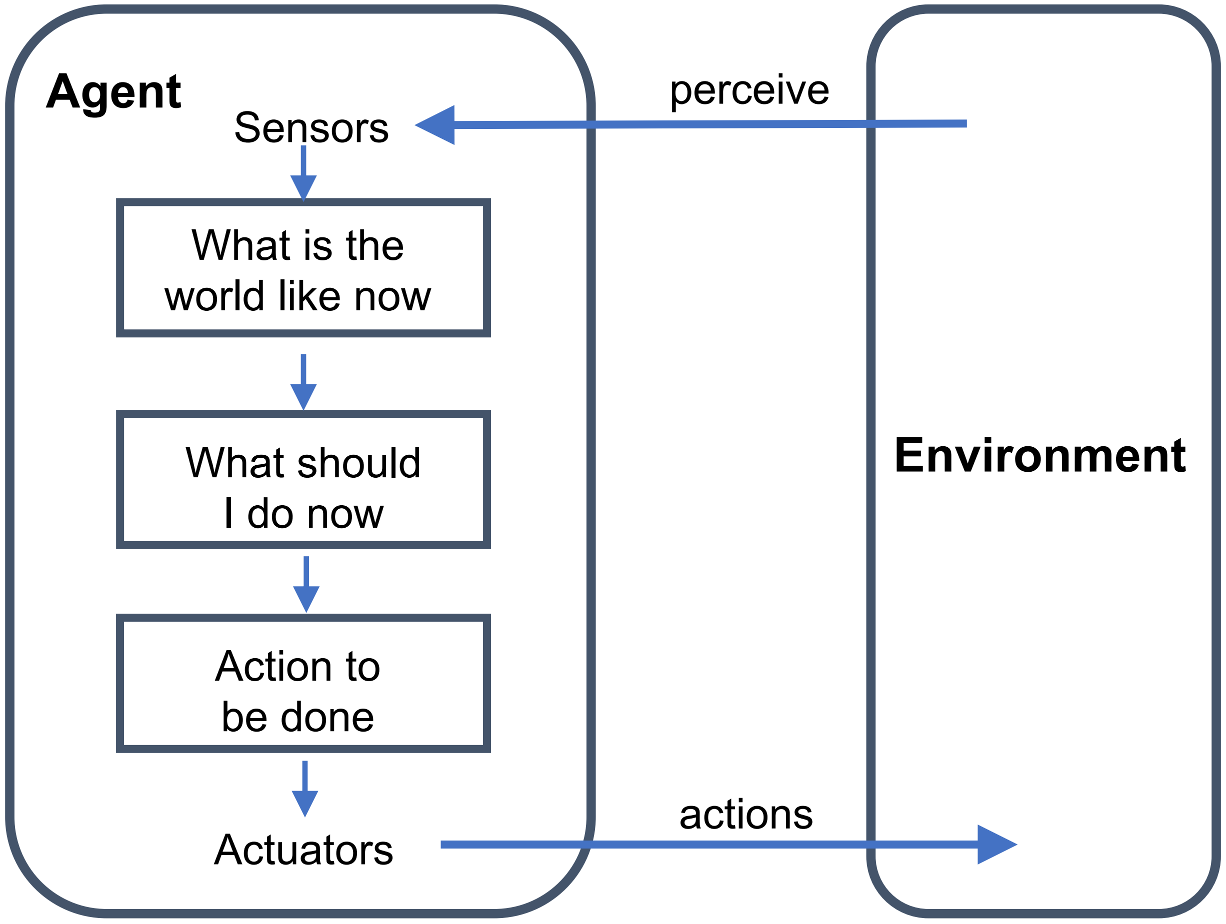
(Russel and Norvig, 2003)



Cognitive Agents

An autonomous agent which

must have explicit reasons for making the choices it does
should be able explain them if necessary



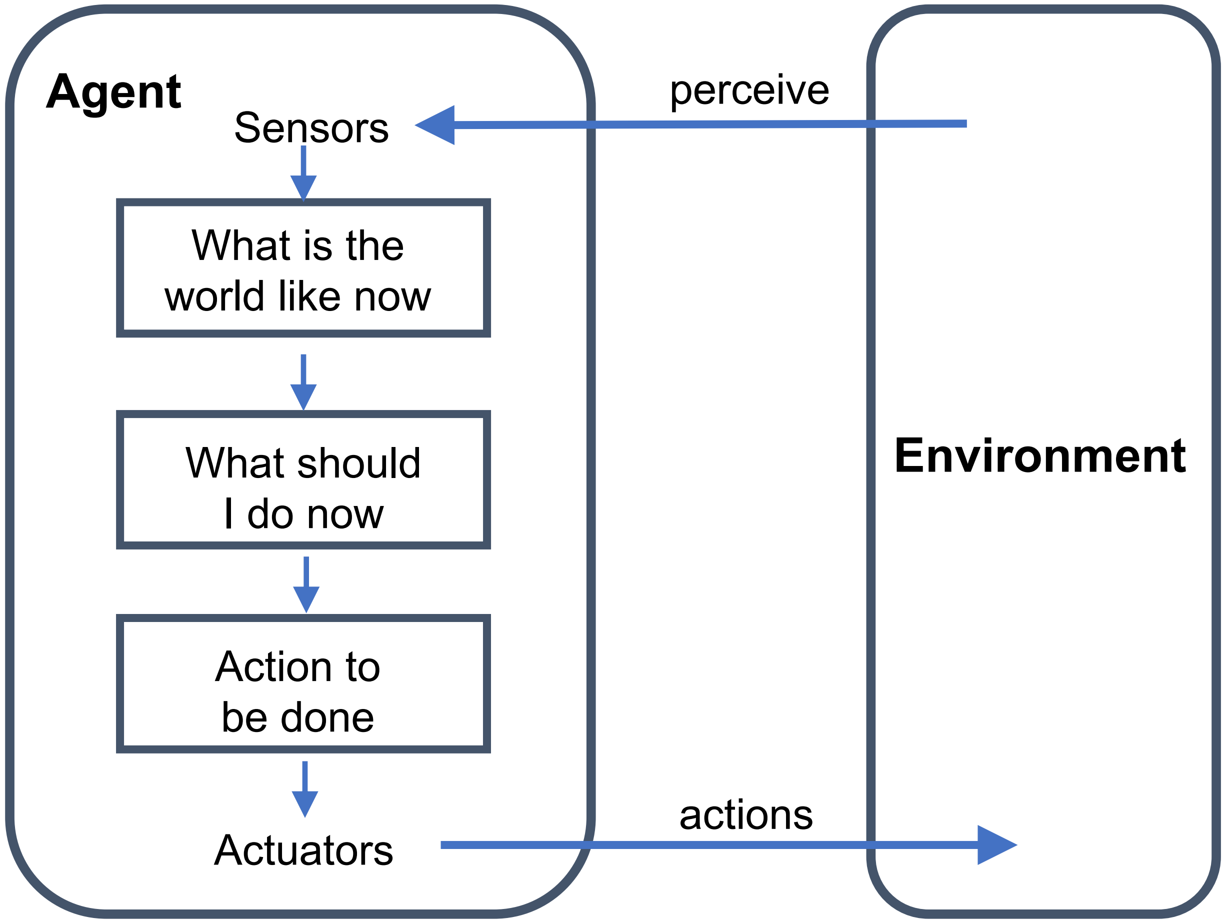
Cognitive Agents

An autonomous agent which

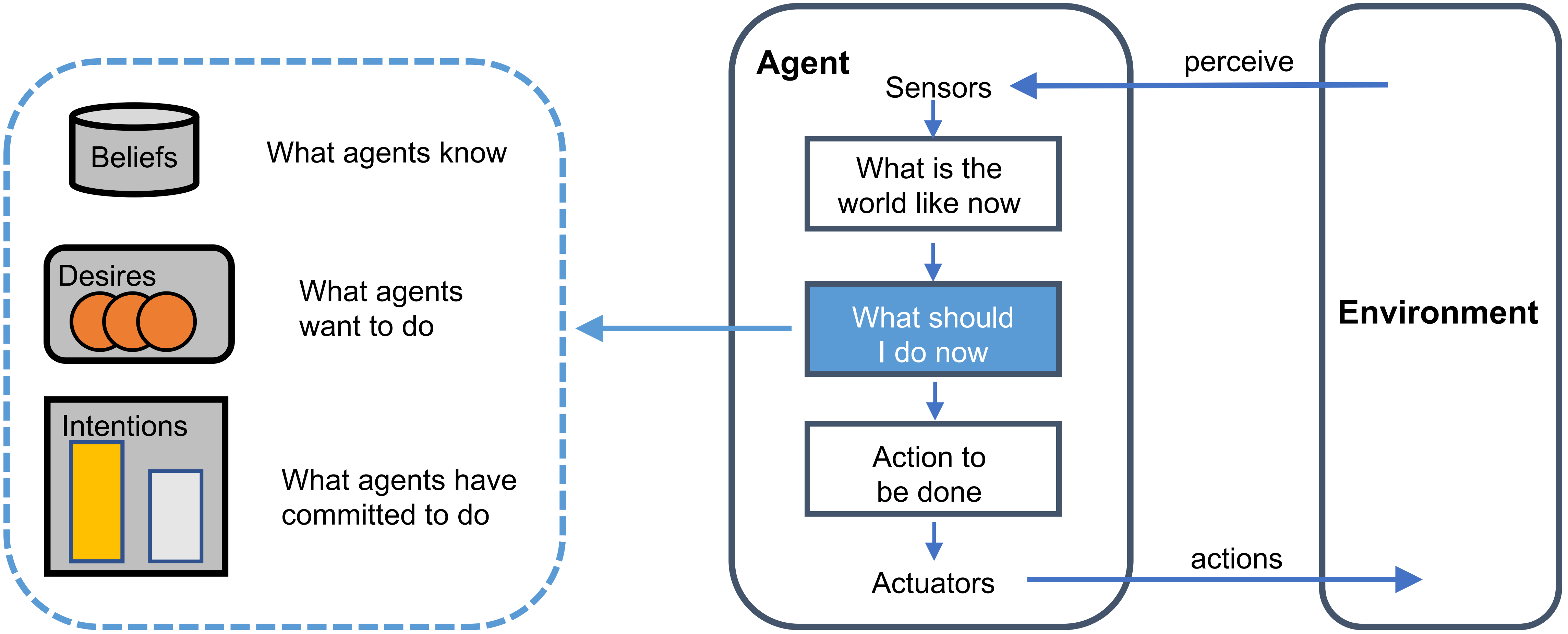
must have explicit reasons for making the choices it does
should be able explain them if necessary

Such agent are programmed by describing

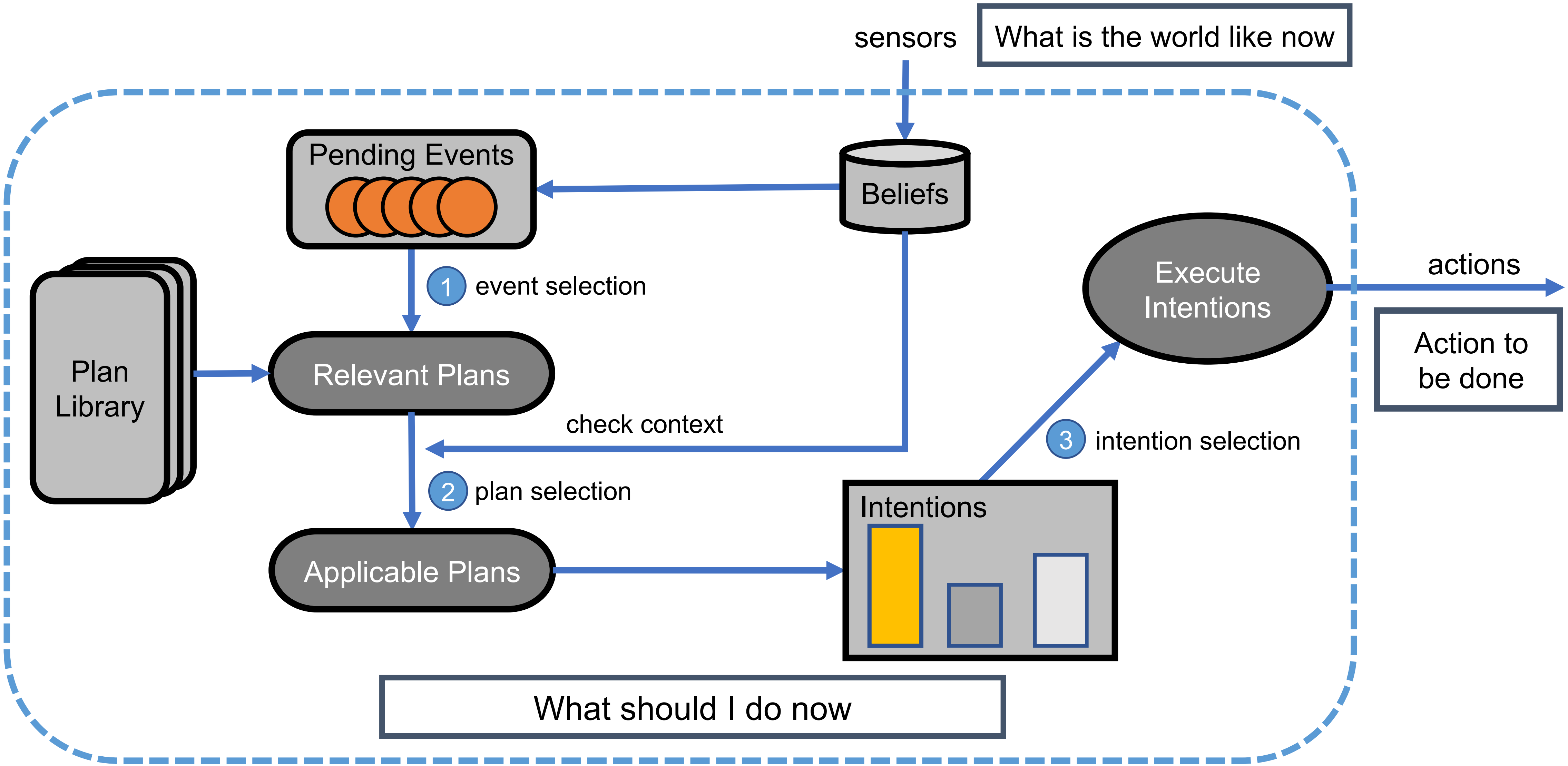
1. their **motivations** (goals, desires, intentions)
2. information (knowledge, beliefs)
3. and how these change over time



Beliefs-Desires-Intentions (BDI) Agents



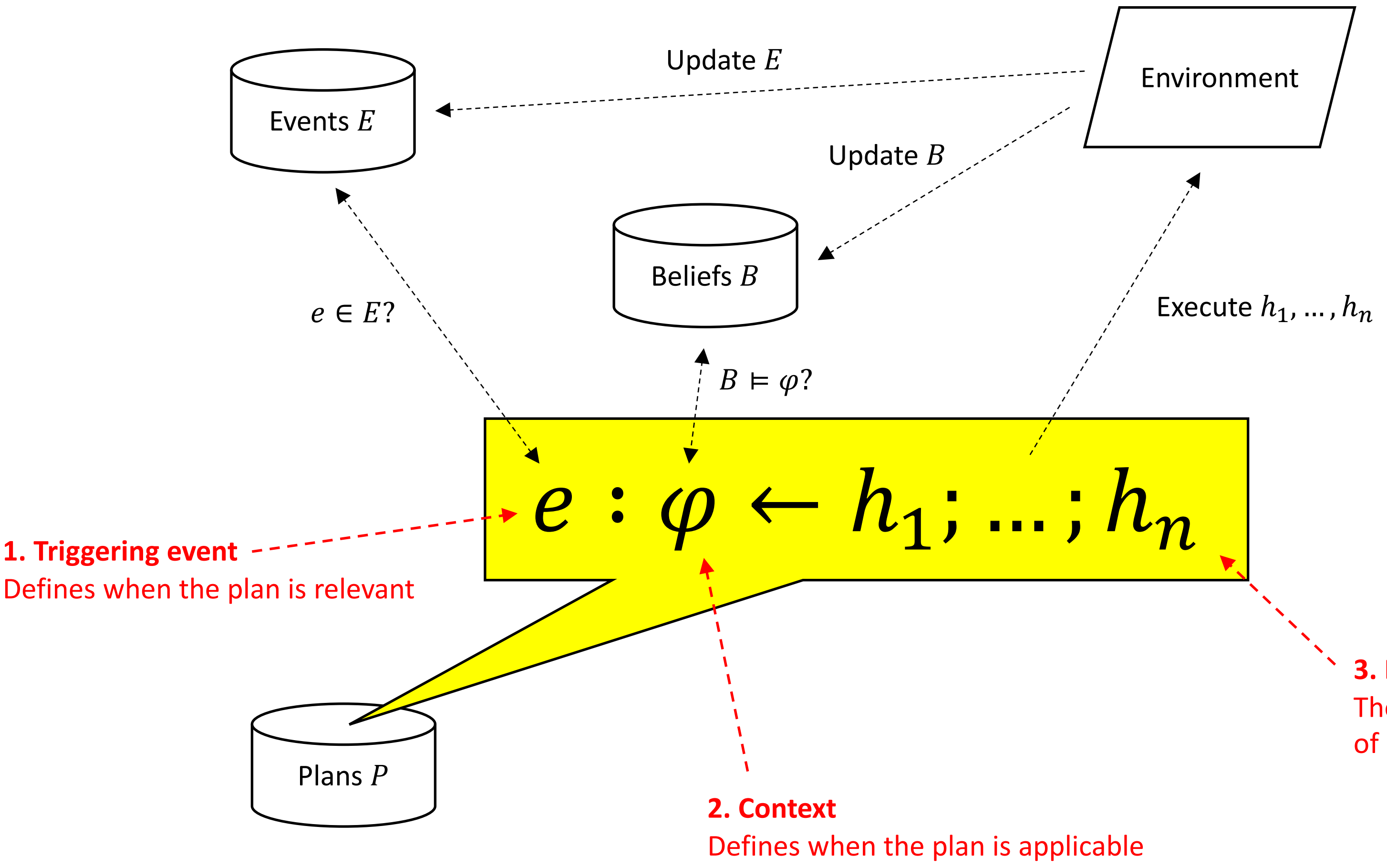
BDI Agents



sensors
What is the world like now
Pending Events
Beliefs
Plan Library
1 event selection
Relevant Plans
Execute Intentions
actions
Action to be done
check context
2 plan selection
3 intention selection
Intentions
Applicable Plans
What should I do now

BDI Agents

High level, Declarative, and Procedural – Plan library



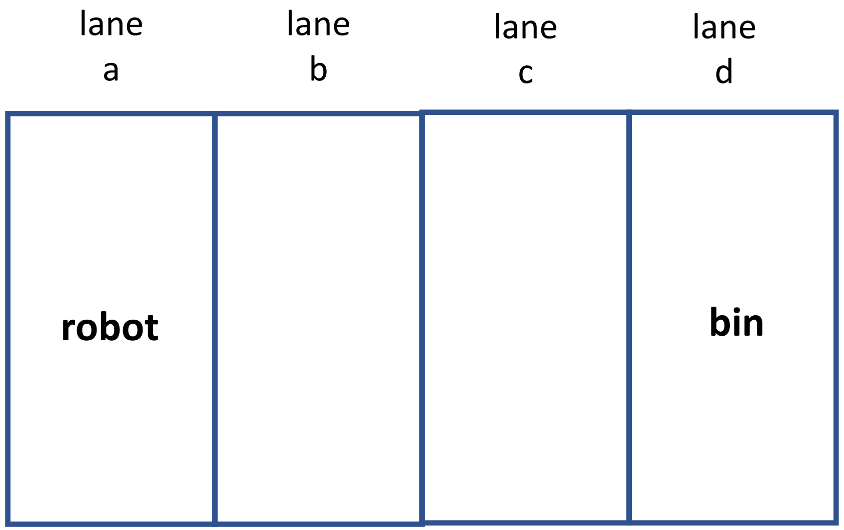
Events, beliefs, plans, actions, and goals all rely on logic programming…

BDI Agents

High level, Declarative, and Procedural

```
/* Initial base beliefs */

adjacent(a, b).
adjacent(b, c).
adjacent(c, d).
location(robot, a).
location(bin, d).
```

```
/* Plan library */

+location(waste, X)
	: location(bin, Y)
		<- !collect(waste); !deposit(waste, bin).

+!collect(X)
	: has(robot, X)
		<- stop.

+!collect(X)
	: not has(robot, X) & location(X, Y)
		<- !go_to(Y); pick_up(X).

+!deposit(X, Y)
	: has(robot, X) & location(Y, Z)
		<- !go_to(Z); drop(X, Y).

+!go_to(X)
	: location(robot, X)
		<- stop.

+!go_to(X)
	: location(robot, Y) & adjacent(Y, Z)
		<- move(Y, Z); !go_to(X).
```

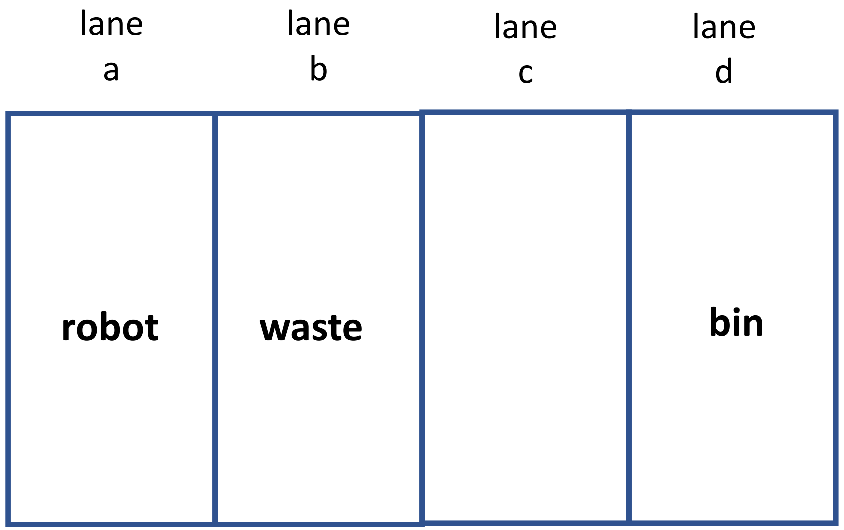


BDI Agents

High level, Declarative, and Procedural

```
/* Initial base beliefs */

adjacent(a, b).
adjacent(b, c).
adjacent(c, d).
location(robot, a).
location(bin, d).

location(waste,b).
```

```
/* Plan library */

+location(waste, X)
    : location(bin, Y)
        <- !collect(waste); !deposit(waste, bin).

+!collect(X)
    : has(robot, X)
        <- stop.

+!collect(X)
    : not has(robot, X) & location(X, Y)
        <- !go_to(Y); pick_up(X).

+!deposit(X, Y)
    : has(robot, X) & location(Y, Z)
        <- !go_to(Z); drop(X, Y).

+!go_to(X)
    : location(robot, X)
        <- stop.

+!go_to(X)
    : location(robot, Y) & adjacent(Y, Z)
        <- move(Y, Z); !go_to(X).
```



Research Questions (revisited)

Are autonomous systems safe to deploy?

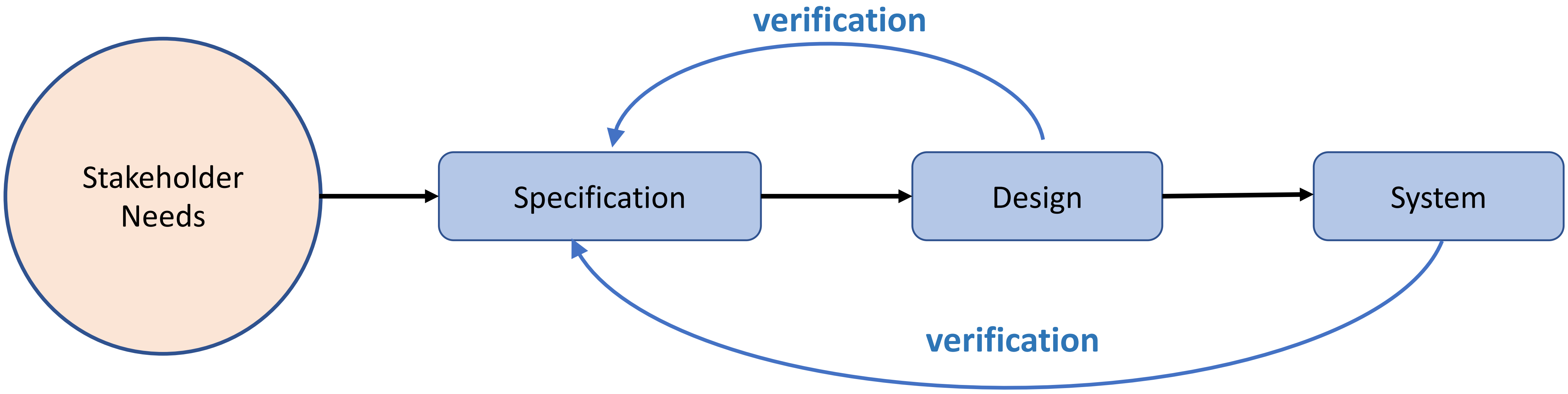
Writing correct BDI programs is not always easy

Plans can include complex constructs including

1. declarative goals,
2. failure recovery,
3. interleaved concurrency

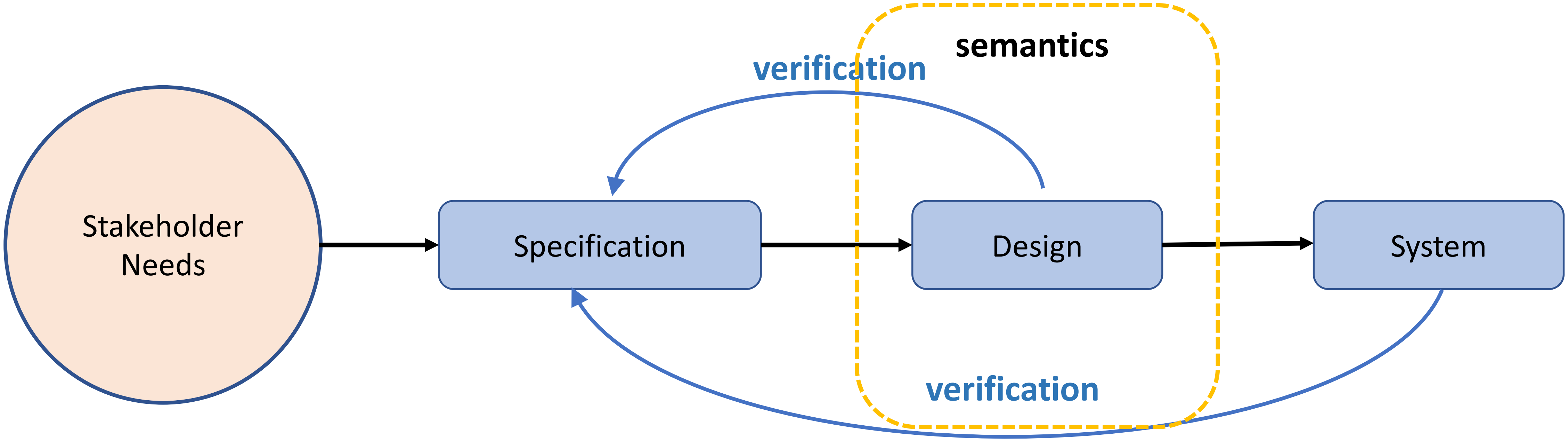
Research Methodology

Are autonomous systems safe to deploy?



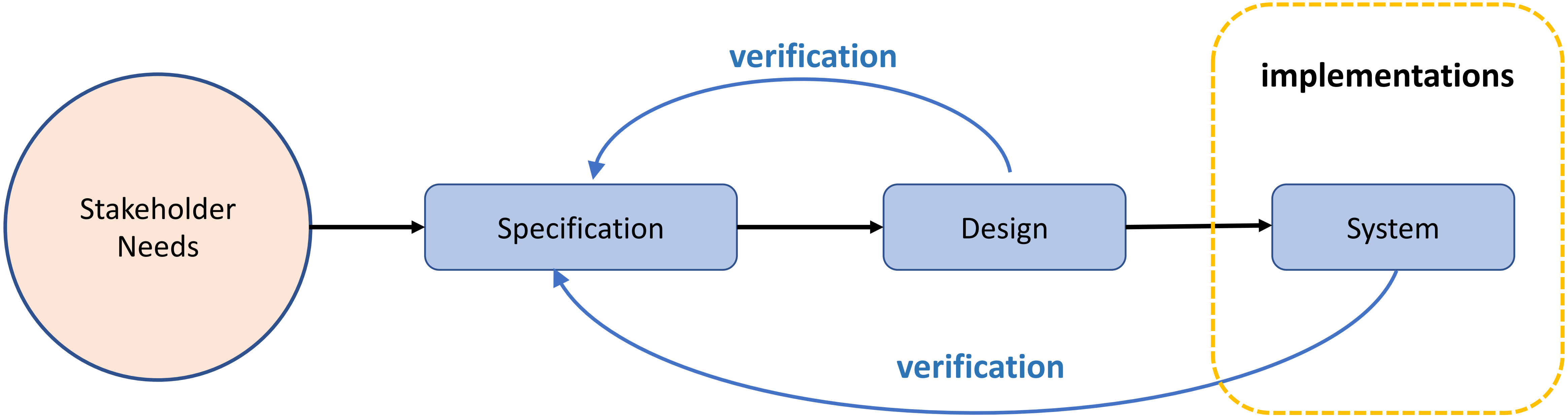
Research Methodology

Are autonomous systems safe to deploy?



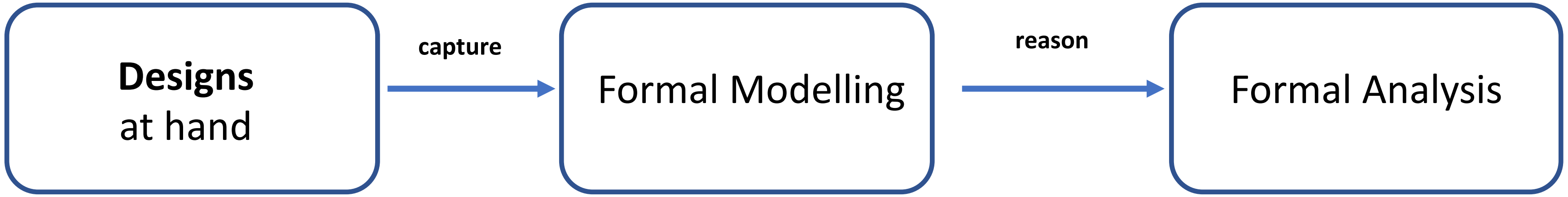
Research Methodology

Are autonomous systems safe to deploy?



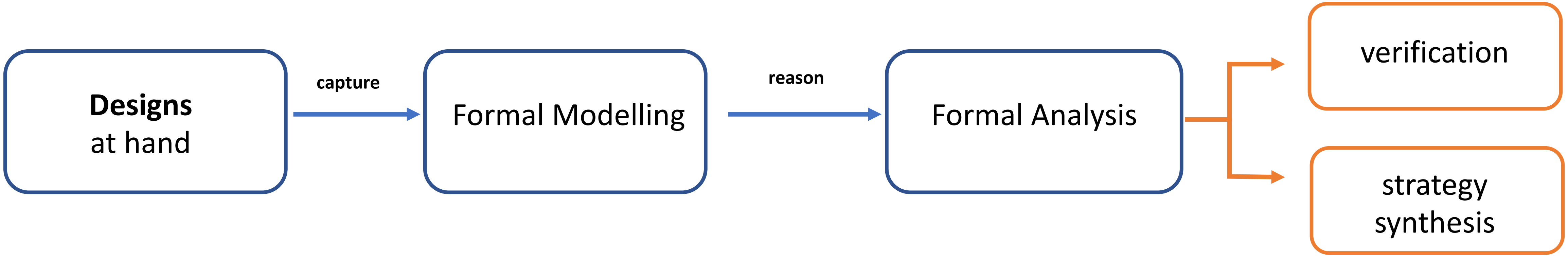
Research Methodology

Are autonomous systems safe to deploy?



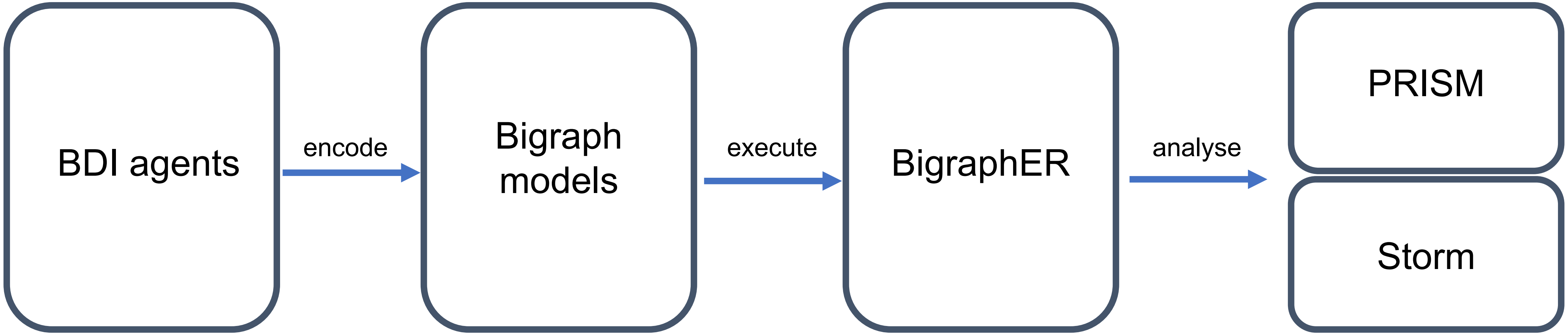
Research Methodology

Are autonomous systems safe to deploy?



Research Contribution

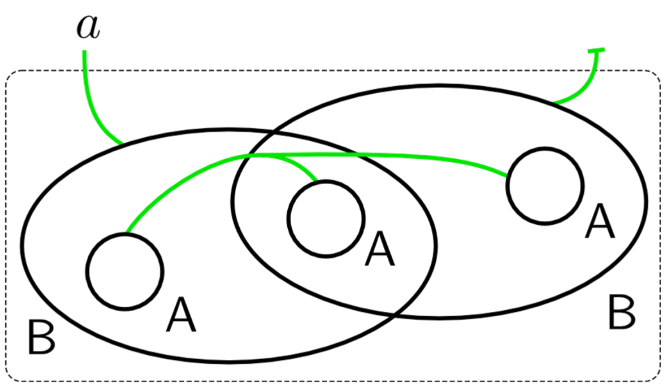
Are autonomous systems safe to deploy?



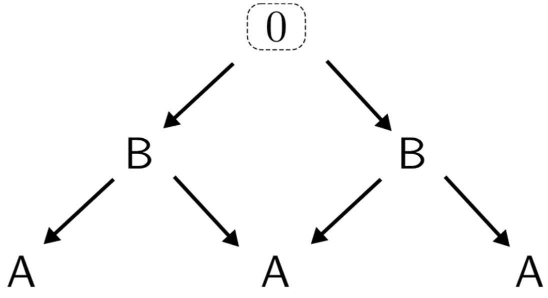
Bigraphs

1. Bigraph: superimposition of a place graph and a link graph
2. Place graph: directed acyclic graph - topological space - no distances - containment relation
3. Link graph: hypergraph - relationships between sets of entities

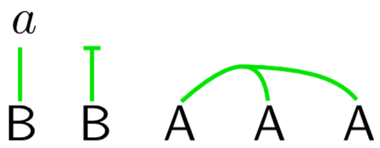
Robin Milner



Bigraph

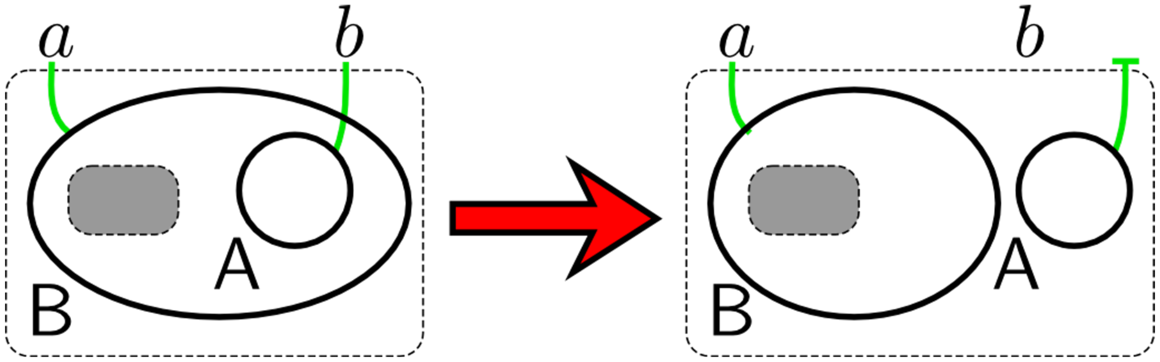


Place graph



Link graph

- A set of reaction rules specify the dynamics of the system
- How to apply a rule to a bigraph (rewriting):
  1. Identify occurrences of the lhs in the bigraph
  2. Substitute each of them with the rhs



BDI Encoding in Bigraphs

Science of Computer Programming 215 (2022) 102760

Contents lists available at ScienceDirect

Science of Computer Programming

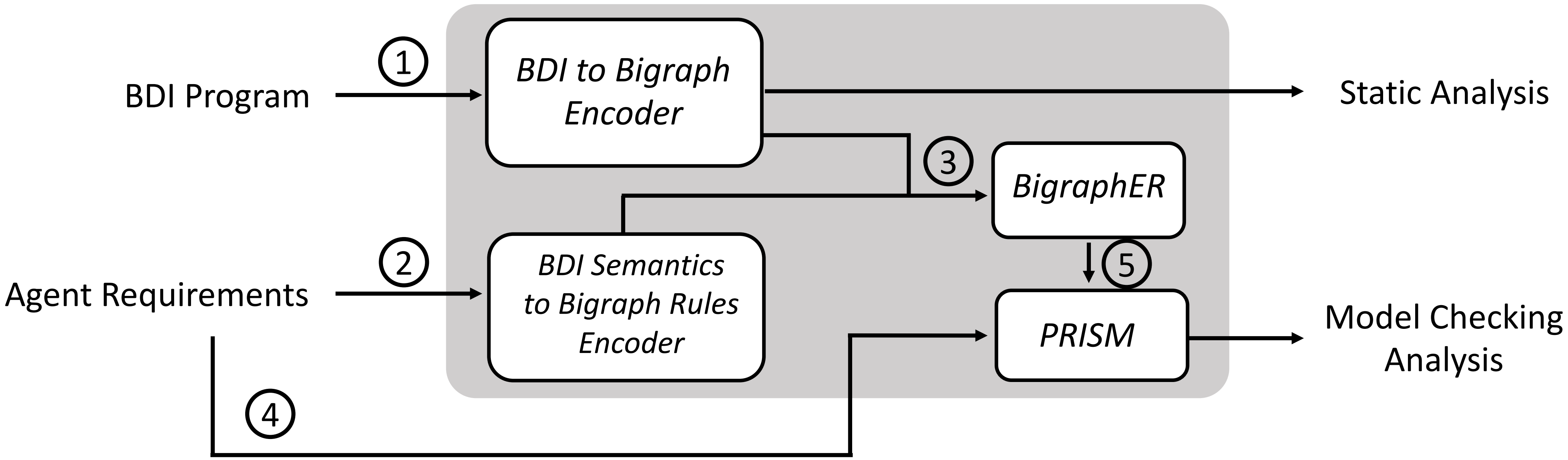
www.elsevier.com/locate/scico

Modelling and verifying BDI agents with bigraphs

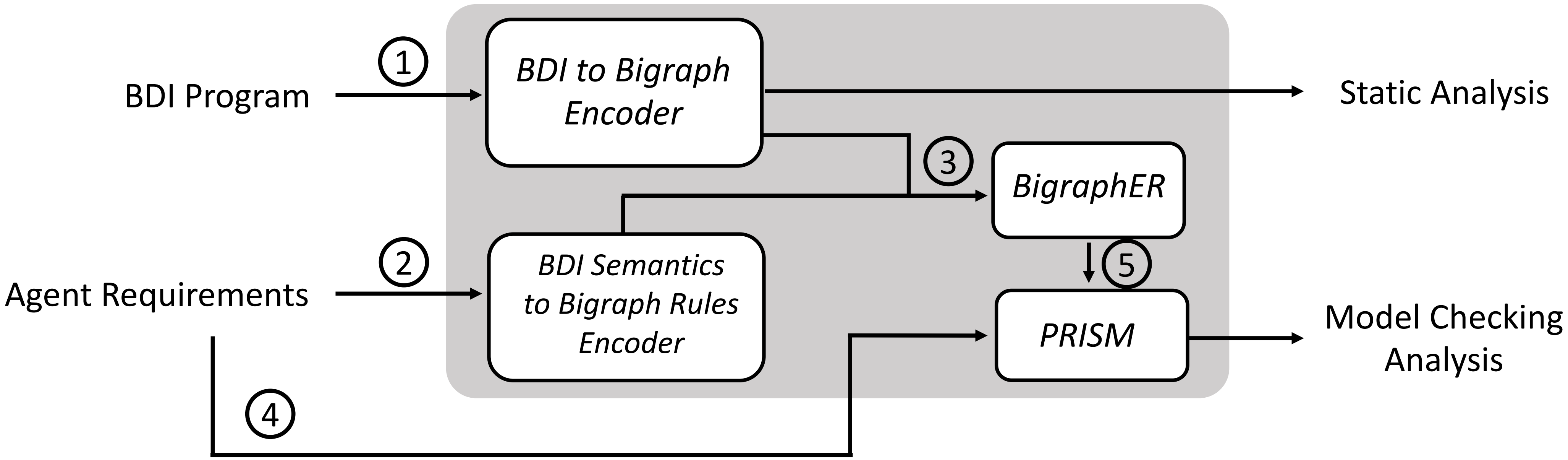
Blair Archibald, Muffy Calder, Michele Sevegnani, Mengwei Xu *

School of Computing Science, University of Glasgow, UK

Research Tool

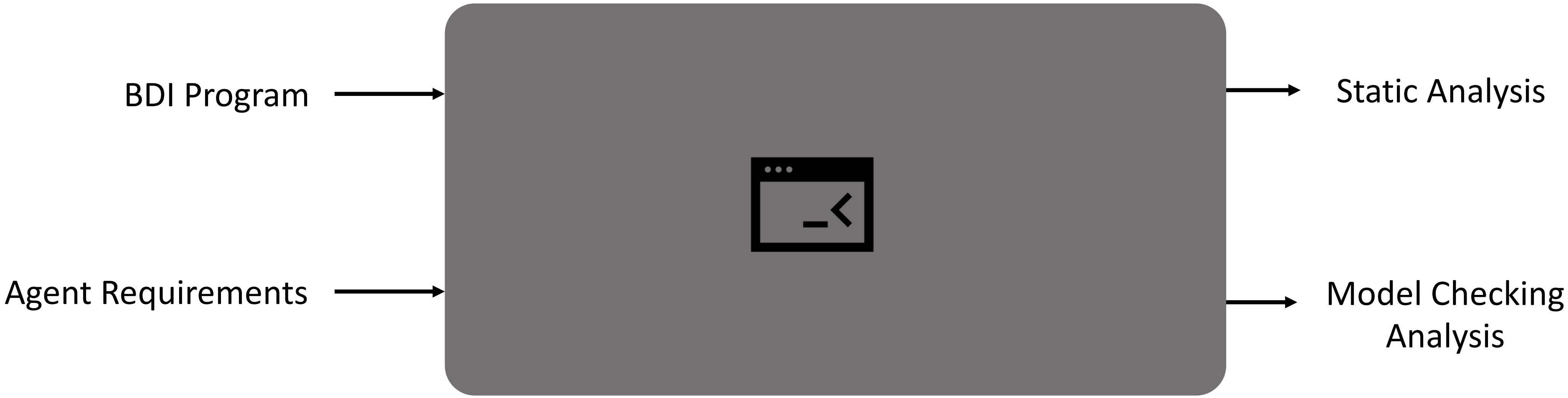


Research Tool



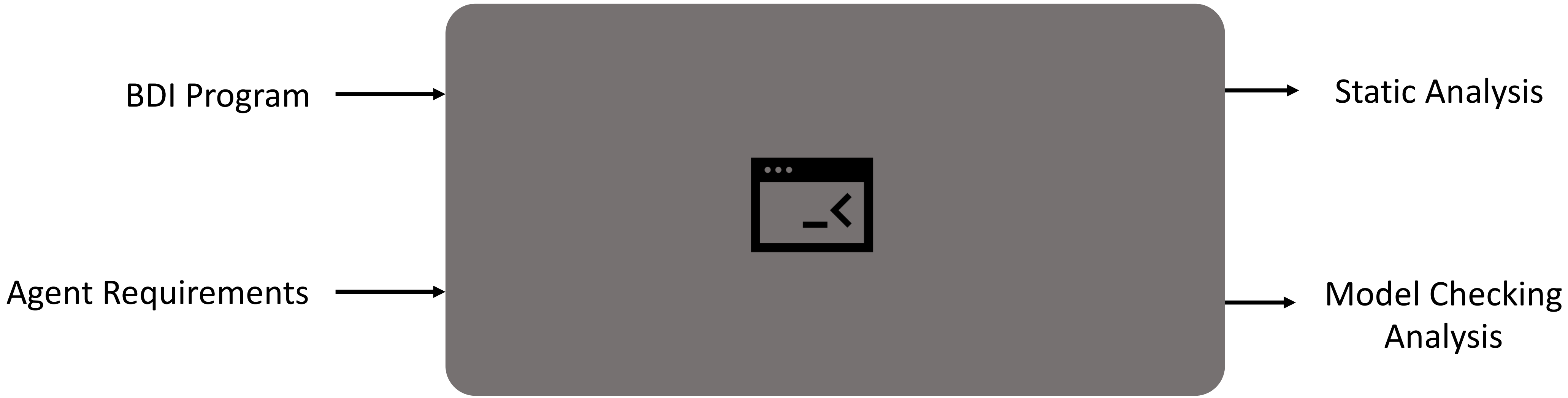
1. agent program compilation to bigraphs
2. predicate labelling in bigraph model
3. exhaustive execution of programs
4. built-in and user-defined belief-based specification formalisation in CTL
5. formal verification

Research Tool



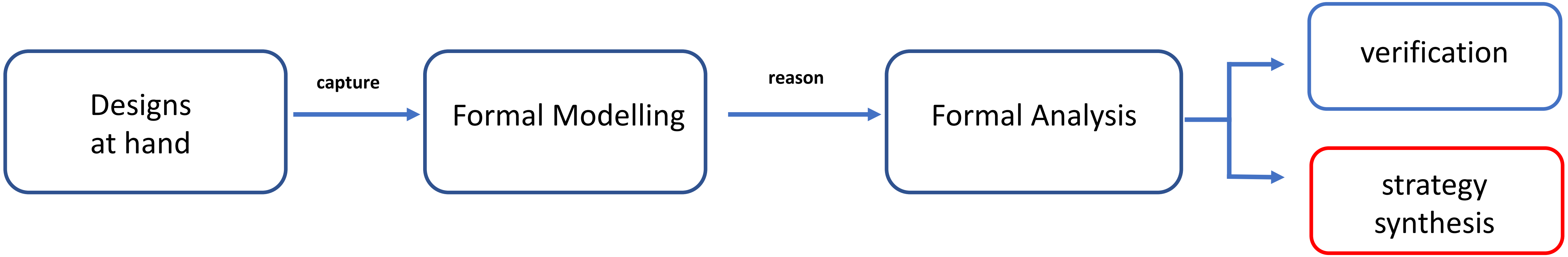
Research Tool

Dem



Future Work

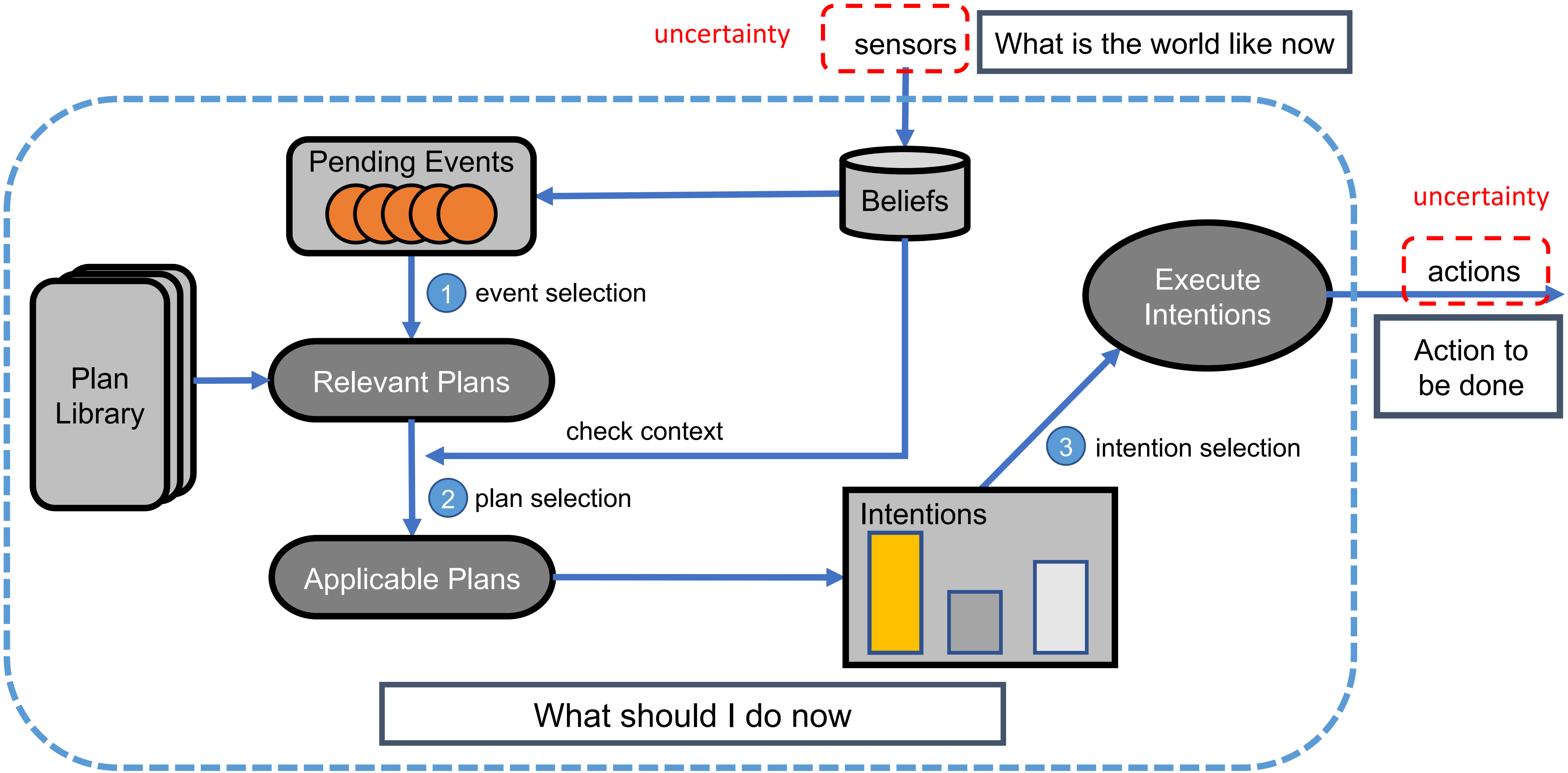
1. Quantitative Verification and Strategy Synthesis for BDI Agents. *NASA Formal Methods 2023*



Future Work

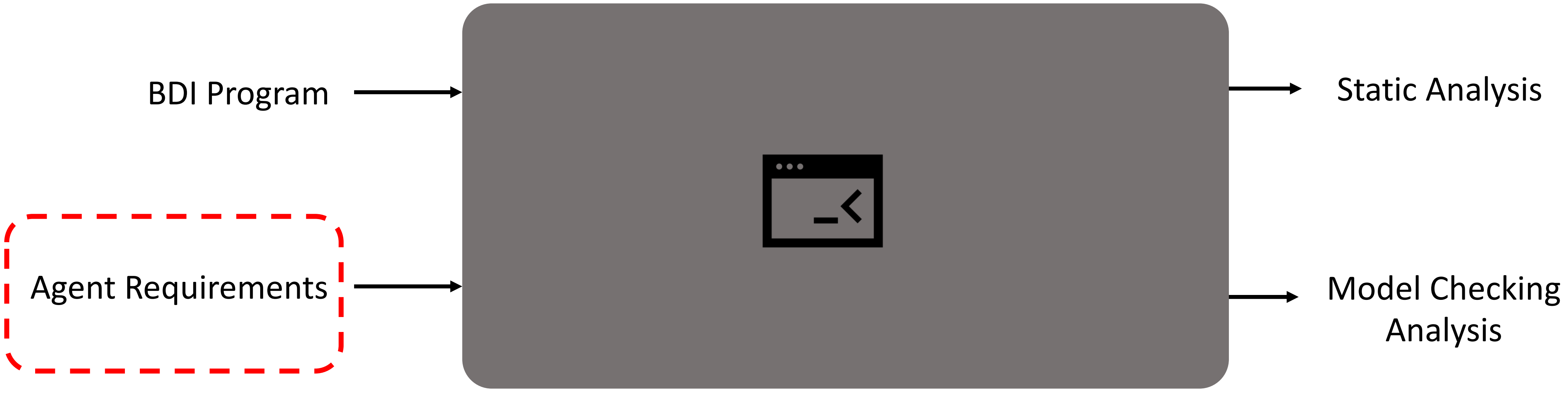
1. Probabilistic BDI Agents: Actions, Plans, and Intentions. *SEFM 2021*
2. Quantitative Modelling and Analysis of BDI Agents. *SoSyM 2023*

Future Work

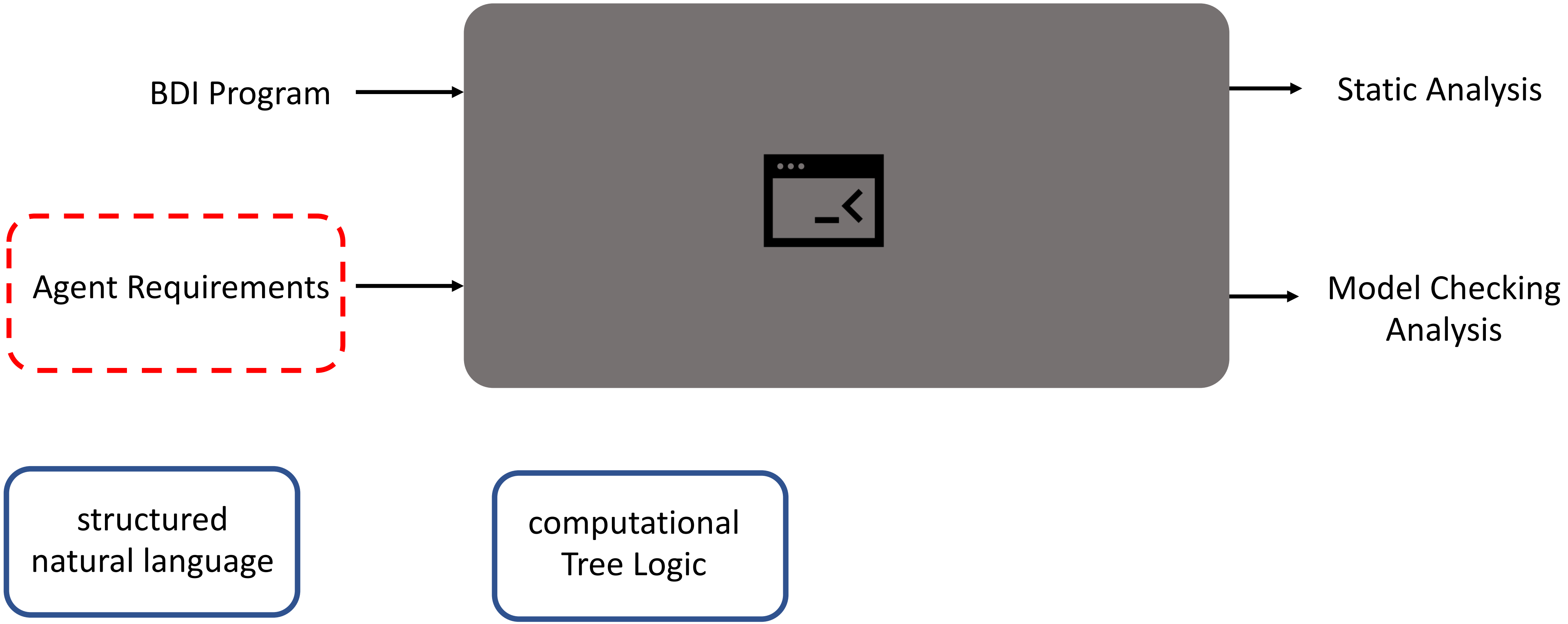


uncertainty
sensors
What is the world like now
Pending Events
Beliefs
uncertainty
Execute Intentions
actions
Plan Library
1 event selection
Relevant Plans
Action to be done
check context
3 intention selection
2 plan selection
Intentions
Applicable Plans
What should I do now

Future Work


BDI Program
Agent Requirements
Static Analysis
Model Checking Analysis

Future Work


BDI Program
Agent Requirements
Static Analysis
Model Checking Analysis
structured natural language
computational Tree Logic

Thank you
&
Questions