An Illustrated Situation Calculus Abstraction for Iterative Explanatory Diagnosis

Christine Task[1],
Mark Wilson[2], Matthew Molineaux[1], David W. Aha[2]

[1]Knexus Research Corporation
[2]Navy Center for Applied Research in AI; Naval Research Laboratory

Research sponsored by OSD ASD (R&E)

Outline

“What’s Happening?”

Diagnosis Needs for Goal Reasoning in Partially Observable Environments

“Where Are We?”

The State-set Abstraction for Sets of Possible States

“How Did We Get Here?”

The Situation Calculus Abstraction for Possible Sequences of State Transitions

Putting It All Together

The Situation/State-set Space for Iterative Diagnosis in Goal Reasoning Agents

Using It

Insights, Potential Directions for Optimizations, Future Work

What's Happening?

Diagnosis Needs for Goal Reasoning in Partially Observable Environments

The Motivation:

An autonomous agent performing Goal Reasoning **needs** a reasonably accurate knowledge of its environment.

- In order to select appropriate **goals**
- In order to make effective **plans** to reach those goals.

What's Happening?

Diagnosis Needs for Goal Reasoning in Partially Observable Environments

The Motivation:

An autonomous agent performing Goal Reasoning **needs** a reasonably accurate knowledge of its environment.

- In order to select appropriate **goals**
- In order to make effective **plans** to reach those goals.

But in realistic settings, the agent **can't directly sense** everything in the environment.

- If the sniper doesn't appear on camera, how do you choose a goal that responds to it?
- If the ditch in the path doesn't appear on LIDAR scans, how do you create a plan to navigate around it?

What's Happening?

Diagnosis Needs for Goal Reasoning in Partially Observable Environments

The Motivation:

If the agent has a model of what can possibly happen in the environment, then it can use what it *can* sense to make inferences about what must be happening (or what must have already happened) in the parts of the environment that it cannot see.

What's Happening?

Diagnosis Needs for Goal Reasoning in Partially Observable Environments

The Motivation:

If the agent has a model of what can possibly happen in the environment, then it can use what it *can* sense to make inferences about what must be happening (or what must have already happened) in the parts of the environment that it cannot see.

To support this inference (using the **DiscoverHistory** algorithm [Molineaux and Aha 2015]) the agent creates and maintains an ***Explanation***, a hypothetical history of what the environment looks like **(states)**, and everything that's happened in the environment so far **(execution history)**.

Diagnosis Needs for Goal Reasoning in Partially Observable Environments

The Motivation:

If the agent has a model of what can possibly happen in the environment, then it can use what it *can* sense to make inferences about what must be happening (or what must have already happened) in the parts of the environment that it cannot see.

To support this inference (using the **DiscoverHistory** algorithm [Molineaux and Aha 2015]) the agent creates and maintains an ***Explanation***, a hypothetical history of what the environment looks like **(states)**, and everything that's happened in the environment so far **(execution history)**.

This history must be consistent with sensor readings. When a sensor reading conflicts with the current explanation, **a better explanation must be found**.

What's Happening?

Diagnosis Needs for Goal Reasoning in Partially Observable Environments

The Motivation:

If the agent has a model of what can possibly happen in the environment, then it can use what it *can* sense to make inferences about what must be happening (or what must have already happened) in the parts of the environment that it cannot see.

To support this inference (using the **DiscoverHistory** algorithm [Molineaux and Aha 2015]) the agent creates and maintains an ***Explanation***, a hypothetical history of what the environment looks like **(states)**, and everything that's happened in the environment so far **(execution history)**.

This history must be consistent with sensor readings. When a sensor reading conflicts with the current explanation, **a better explanation must be found**. ***This search may be broad, complex, slow, difficult, unpleasant…***

What's Happening?

Diagnosis Needs for Goal Reasoning in Partially Observable Environments

The Motivation:

To support this inference, the agent creates and maintains an ***Explanation***, a hypothetical history of what the environment looks like **(states)**, and everything that's happened in the environment so far **(execution history)**.

This history must be consistent with sensor readings. When a sensor reading conflicts with the current explanation, **a better explanation must be found**.

This search may be broad, complex, slow, difficult, unpleasant…

This work looks at the

Solution Space of Possible Explanations

How we can formalize it, understand it, depict it, and in the future, ***navigate it more efficiently***.

What's Happening?

Diagnosis Needs for Goal Reasoning in Partially Observable Environments

Explanation Solution Space,
Interdependent Sources of Uncertainty:

We don't know what **State** we're in

We don't know what **Events** led us to this state

What's Happening?

Diagnosis Needs for Goal Reasoning in Partially Observable Environments

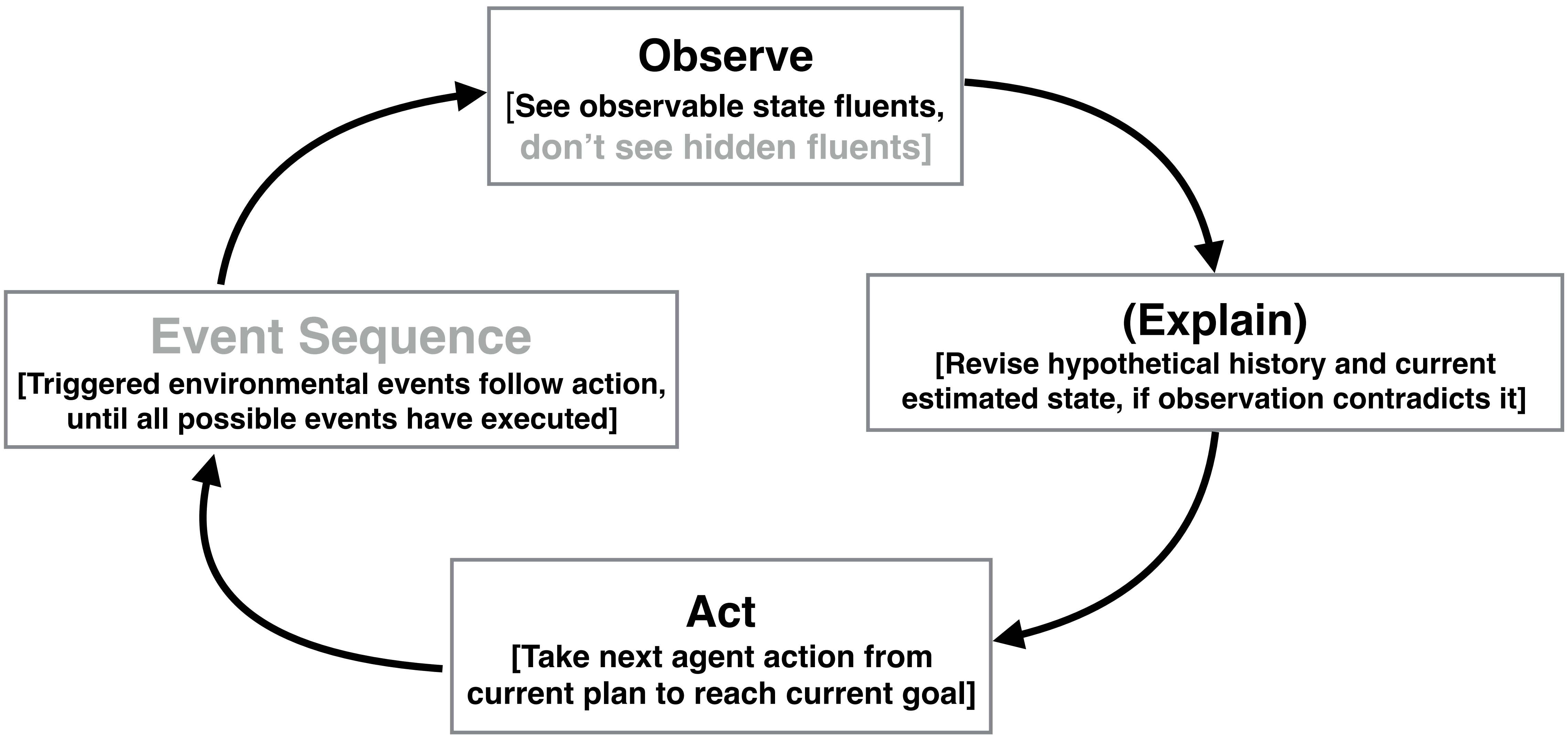
Our Problem Definition:

- '**Fluent**' refers to a fact about the environment (a predicate or proposition). Fluents are defined as either **observable** **or hidden**. The agent can read the current values of observable fluents by making an '**Observation**'.
- A **State** is a value assignment to all fluents.
- An agent **Action** changes the state (and is **observable**). Preconditions and effects for actions are known.
- An environmental **Event** changes the state (but is **not observable**). Preconditions and effects for events are known. Events happen immediately, deterministically, when their (possibly hidden) preconditions are satisfied.

What’s Happening?

Diagnosis Needs for Goal Reasoning in Partially Observable Environments

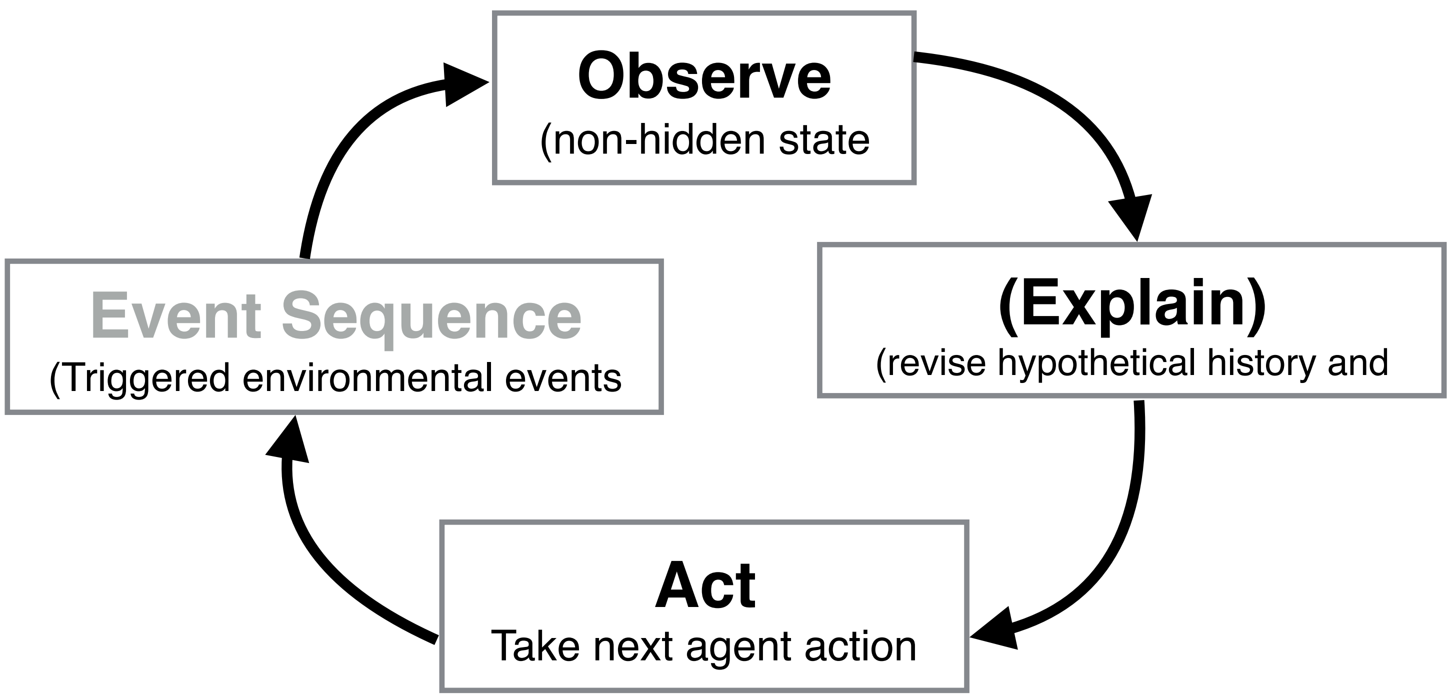
Our Problem is Iterative:



What's Happening?

Diagnosis Needs for Goal Reasoning in Partially Observable Environments

Our Problem is Iterative:



With our domain model, we use our Observable Execution History:

- observed initial state
- agent action 1
- (environmental events)
- observation 2

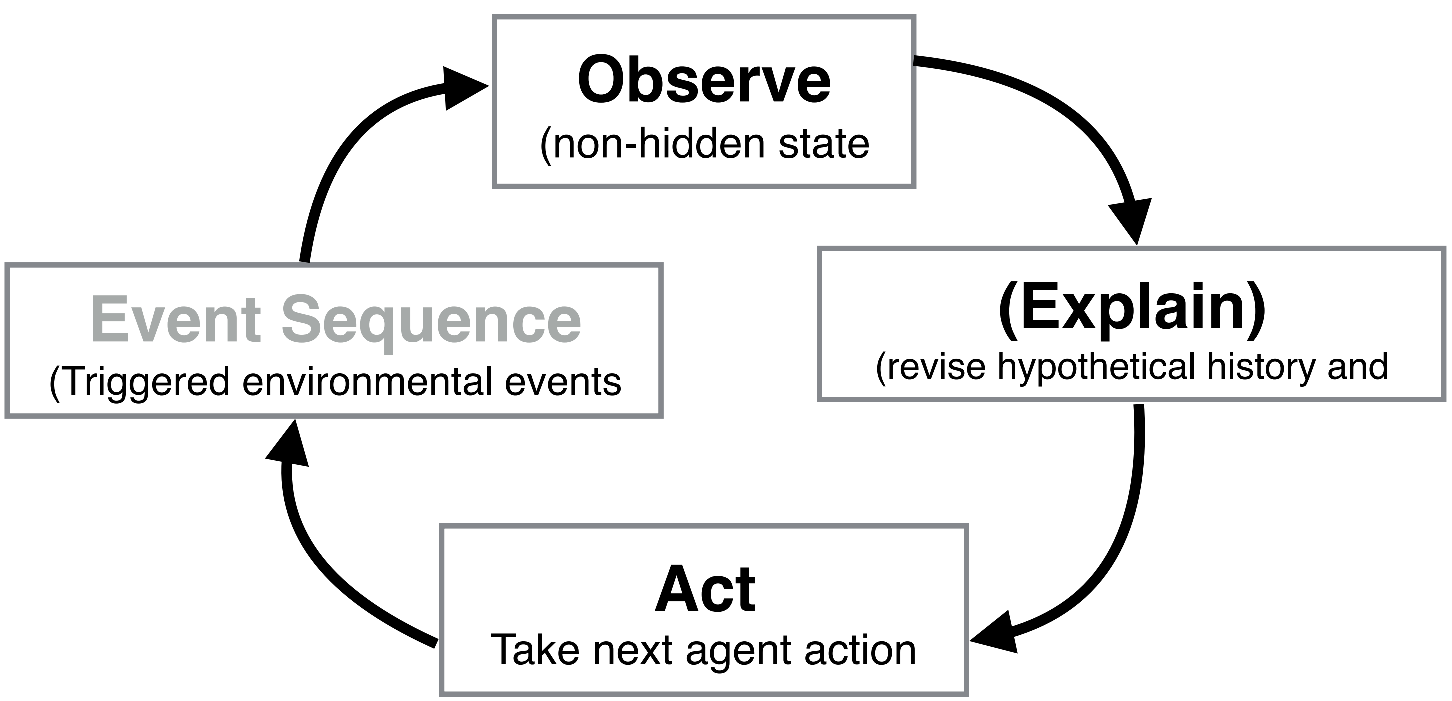
To Hypothesize a Full History (an 'Explanation'):

- true initial state
- agent action 1
- environmental events
- state 1

What’s Happening?

Diagnosis Needs for Goal Reasoning in Partially Observable Environments

Our Problem is Iterative:



With our domain model, we use our Observable Execution History:

- observed initial state
- agent action 1
- (environmental events)
- observation 2
- agent action 2
- (environmental events)
- observation 3

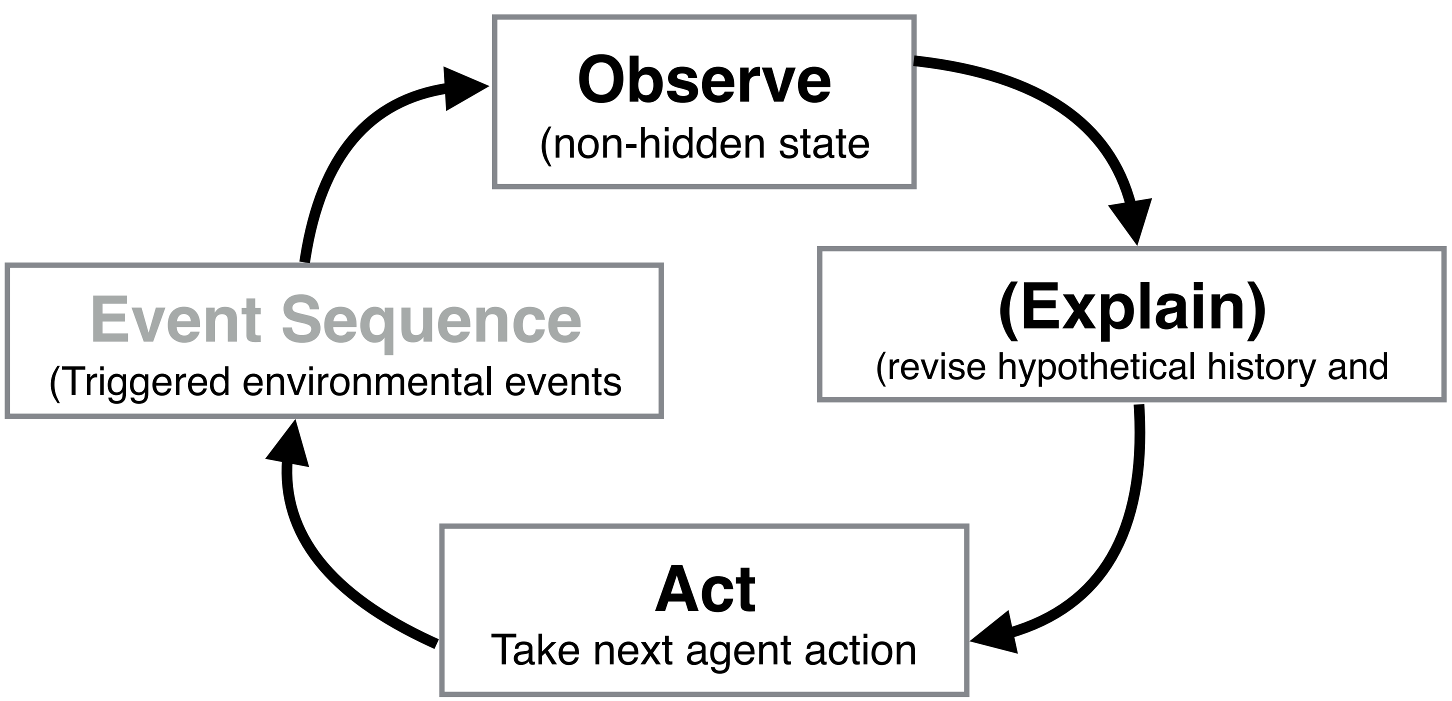
To Hypothesize a Full History (an ‘Explanation’):

- true initial state
- agent action 1
- environmental events
- state 1
- agent action 2
- environmental events
- state 3

What’s Happening?

Diagnosis Needs for Goal Reasoning in Partially Observable Environments

Our Problem is Iterative:



With our domain model, we use our Observable Execution History:

observed initial state
agent action 1
(environmental events)
observation 2
agent action 2
(environmental events)
observation 3
agent action 3
(environmental events)
observation 4
agent action 4

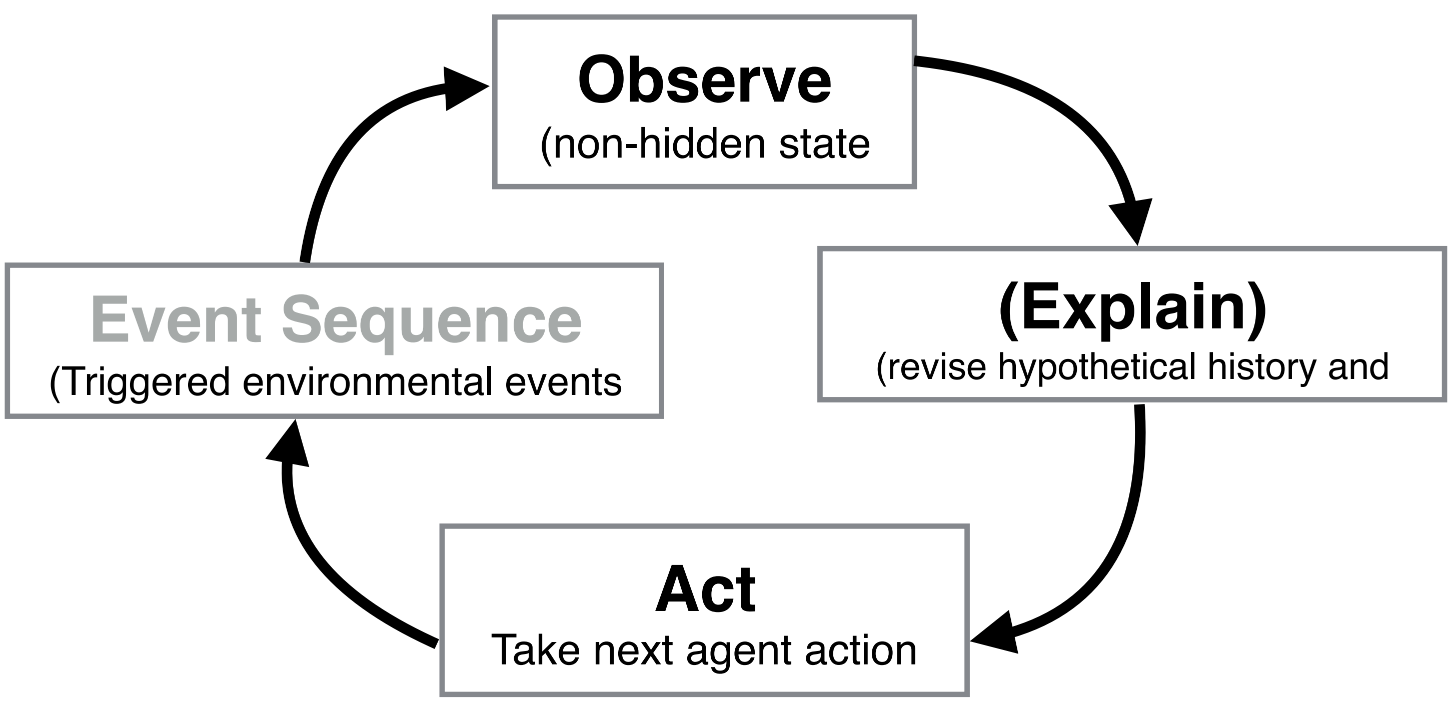
To Hypothesize a Full History (an ‘Explanation’):

true initial state
agent action 1
environmental events
state 1
agent action 2
environmental events
state 3
agent action 3
environmental events
state 4
agent action 4

What's Happening?

Diagnosis Needs for Goal Reasoning in Partially Observable Environments

Our Problem is Iterative:



With our domain model, we use our Observable Execution History:

observed initial state
agent action 1
(environmental events)
observation 2
agent action 2
(environmental events)
observation 3
agent action 3
(environmental events)
observation 4
agent action 4
(environmental events)
observation 5
agent action 5
(environmental events)
observation 6
agent action 6
(environmental…)

To Hypothesize a Full History (an 'Explanation'):

true initial state
agent action 1
environmental events
state 1
agent action 2
environmental events
state 3
agent action 3
environmental events
state 4
agent action 4
environmental events
state 5
agent action 5
environmental events
state 6
agent action 6
environmental…

What's Happening?

Diagnosis Needs for Goal Reasoning in Partially Observable Environments

Iterative Diagnosis Solution Space,
Interdependent Sources of Uncertainty:

We don't know what **State** we're in

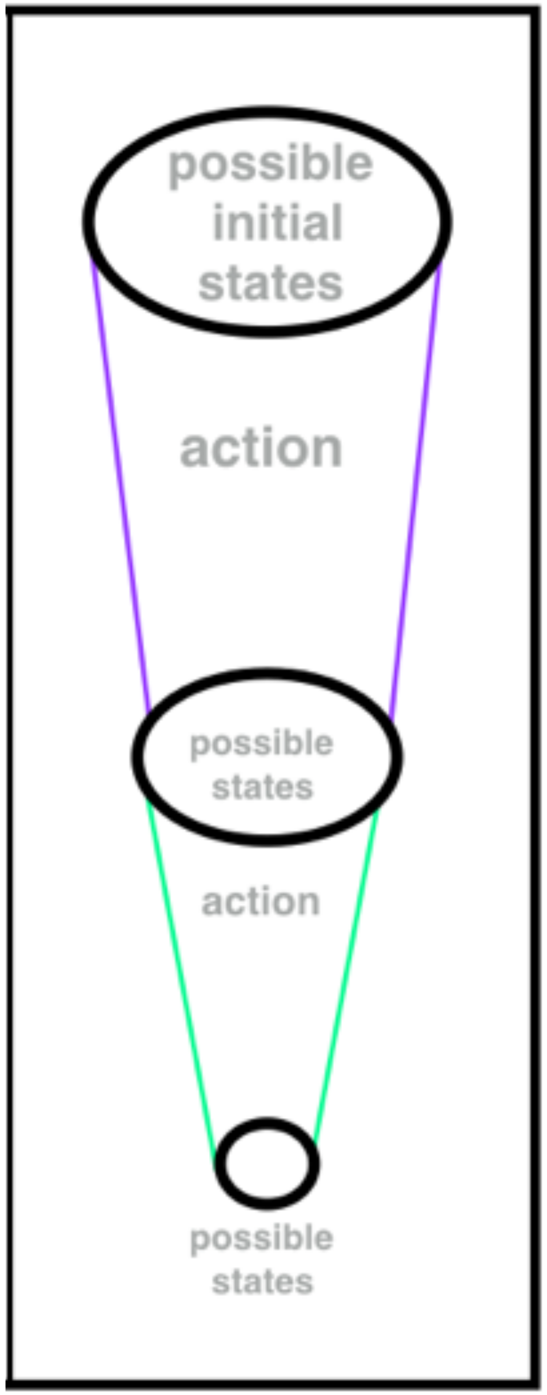
We don't know what **Events** led us to this state

The State-set Abstraction for Sets of Possible States

[Pang and Holte 2011]

Definition: A *state-set* is a set of possible states.

The *state-set* framework provides a method for depicting and reasoning over sets of possible states (and transitions between them)



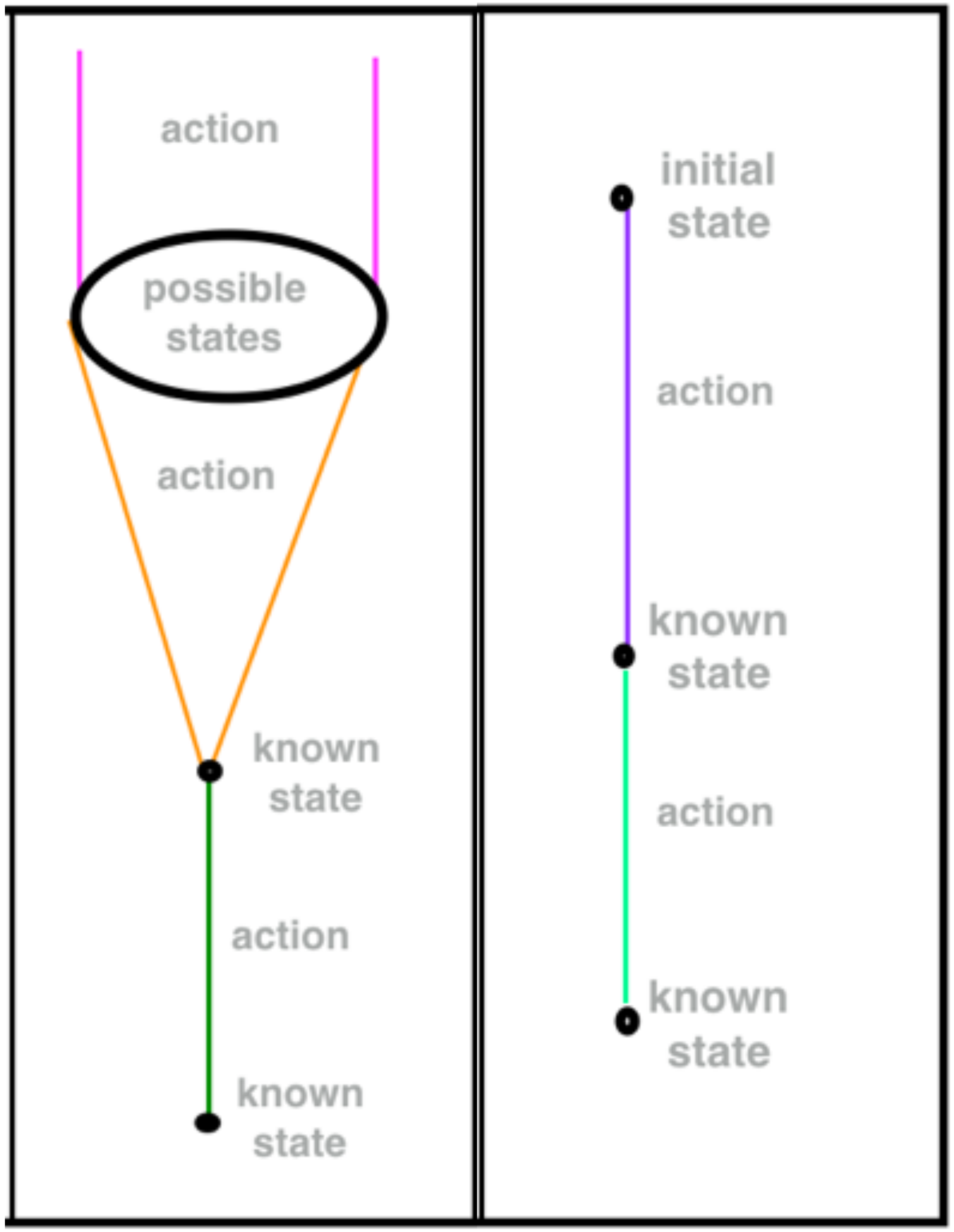
Where Are We?

The State-set Abstraction for Sets of Possible States

Theorem 1: In a deterministic context, a strongly connected sequence of state-sets** is always monotonically decreasing (i.e. non-increasing) in size.

Corollary 1: *Once we know what state we're in, we'll always know what state we're in.*

**Where a sequence [Set1, Action1, Set2, Action2] implies Set2 equals the intersection of image(Action1(Set1)) and domain(Action2()).



What's Happening?

Explanation Solution Space,
Interdependent Sources of Uncertainty:

We don't know what **State** we're in
But we know how to describe that uncertainty

We don't know what **Events** led us to this state

How Did We Get Here?

The Situation Calculus Abstraction for Possible Transition Sequences

[Lin and Reiter 1994]

Definition: A *situation* is a sequence of state transition functions (actions or events)—but not the states themselves!

The *situation calculus* is a formal logic for reasoning over these sequences.

[McIlraith 1997]

In the diagnosis community, useful concepts such as Observations, Hypothesized Initial States and Diagnoses (similar to Explanations) have been defined for the situation calculus.

Situation Calculus Symbols

- **Situations**: The complete sequence of actions (often indicated with the symbol s) that have occurred in the system up to a given point. The null, initial situation is denoted as S_0, and the distinguished function 'do' describes situation transitions: $s_2 = do(a, s_1)$ denotes the situation s_2 resulting from performing action a in situation s_1.
- **Objects** O: A finite set of typed objects that exist in the environment (examples: squad members, trees, the ASM robot itself)
- **Fluents** F: A set of predicates[2] over objects, with values that vary across situations. For this reason, the situation is always the last parameter in a fluent expression. For example, the truth value of the predicate fluent $Wounded(soldier1, s)$ indicates whether $soldier1$ is wounded after the action sequence denoted by situation s. Note that, by itself, s is generally not sufficient to determine the value of a fluent $F(x, s)$; this value is also dependent on the initial system state, T_{S0}, which we discuss below.
- **Actions** A: There is a finite set of action symbols. The behavior of these actions (i.e., their preconditions and effects) are encoded in the precondition and successor state axioms described below. The atomic expression $Poss(a, s)$ indicates whether action a is possible in situation s (and, as with fluents, the value of $Poss(a, s)$ is partially dependent on the initial system state T_{S0}.)

Situation Calculus Axioms

- **Foundational Axioms** Σ_{found} : The foundational axioms specify the domain-independent framework of the situation calculus, including the definition of situations (described informally above) and the framing axiom or domain closure axiom (described informally below). They also define the predecessor relation $s \sqsubset s'$, which holds if and only if s is a strict prefix of s' (recall that each situation encapsulates an entire action sequence, starting from the initial null situation, S_0)
- **Initial Constraints** T_{SC}^{S0} : A set of constraints on fluents, which all valid initial states must satisfy (for example, $\forall x \in Soldiers: \neg Wounded(x, S_0)$).
- **Successor State Axioms** T_{SS}: This set contains one pair of Successor State Axioms (SSA) for each fluent; it encodes the effects each (possible) action can have on the fluent's value[3]. These are of the form:
  $$F(x_1, \dots, x_n, do(a, s)) \equiv \Phi_F(a, x_1, \dots, x_n, s)$$
  $$\neg F(x_1, \dots, x_n, do(a, s)) \equiv \Phi_{\neg F}(a, x_1, \dots, x_n, s)$$
  where Φ_F is a formula uniform in s (ie, not referring to any predecessors of s), and $a, x_1, \dots, x_n$ are free variables spanning all applicable actions and parameter values for F. For example:
  $Wounded(x, do(a, s)) \equiv [Wounded(x, s) \vee (a = IsShot(x))]$
  $\neg Wounded(x, do(a, s)) \equiv [\neg Wounded(x, s) \vee (a = Treated(x))]$
- **Action Precondition Axioms** T_{AP}: This set contains one precondition axiom for each action symbol in the domain. These are of the form:
  $$Poss(a(x_1, \dots, x_n), s) \equiv \Pi_a(x_1, \dots, x_n, s)$$
  where Π_a is a formula uniform in s which defines all conditions under which a can be performed in s, and $a, x_1, \dots, x_n$ are free variables. For example: $Poss(IsShot(x), s) \equiv [UnderAttack(x) \wedge Exposed(x)]$
- **Unique Action Name Axioms** T_{UNA}: These axioms enforce unique names for actions.
- **Initial State** T_{S_0}: These axioms specify the complete set of initial fluent values for a given instance of the problem. Because situations specify action sequences rather than environmental states, T_{S_0} is necessary (in general) to compute which fluent values hold and which actions are possible in a given situation.

What's Happening?

Explanation Solution Space, Interdependent Sources of Uncertainty:

We don't know what **State** we're in

But we know how to describe that uncertainty

We don't know what **Events** led us to this state

But we can describe what histories are possible

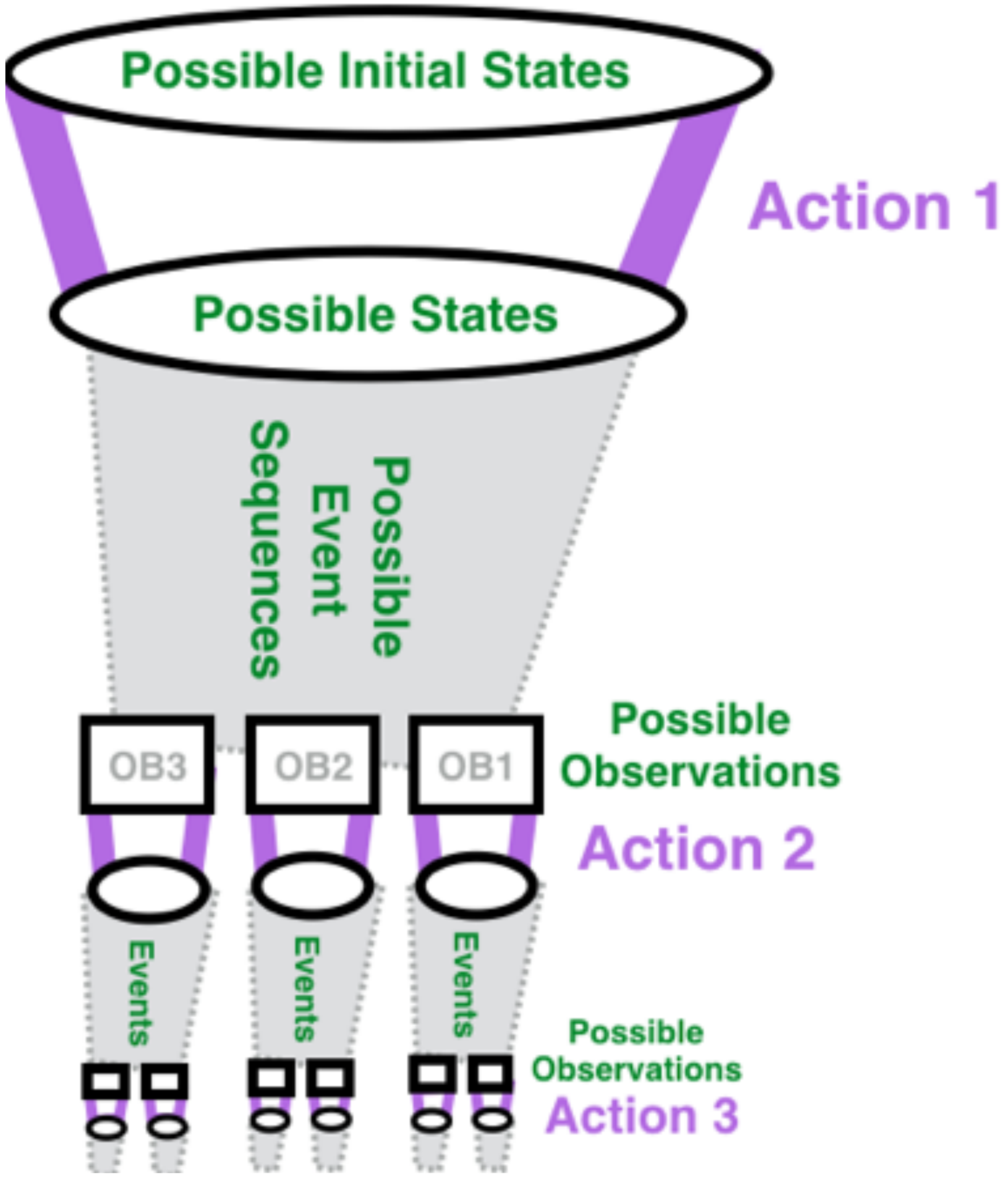
Putting It All Together

The Situation/State-set Space for Iterative Diagnosis in GR Agents

Now we can formally describe both sources of uncertainty: states (state-sets), and event/action sequences (situations).

This enables us to diagram (and define formally, omitted here) the solution-space consisting of all possible explanations for an agent at a given point in a given plan of actions.

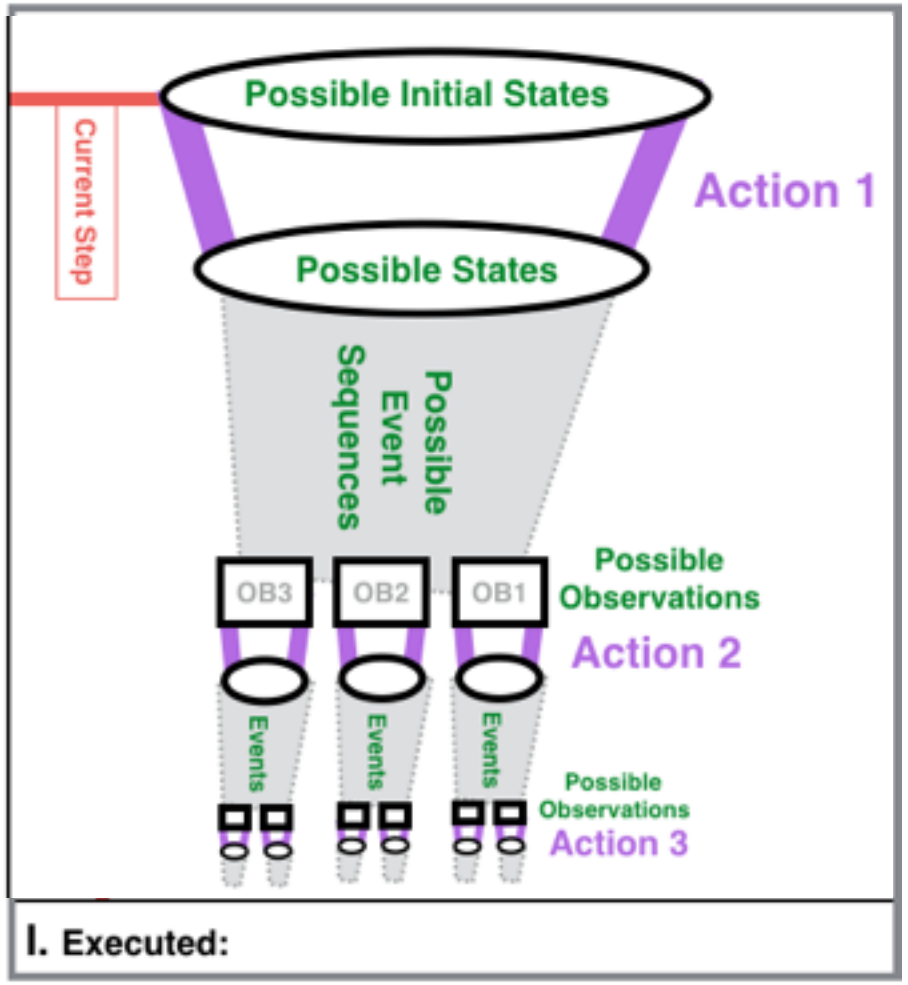
This is the space our explanation search algorithm needs to navigate efficiently.



Putting It All Together

The Situation/State-set Space for Iterative Diagnosis in GR Agents

The Space of Possible Execution Paths for: *Plan = [Action 1, Action 2, Action 3]*

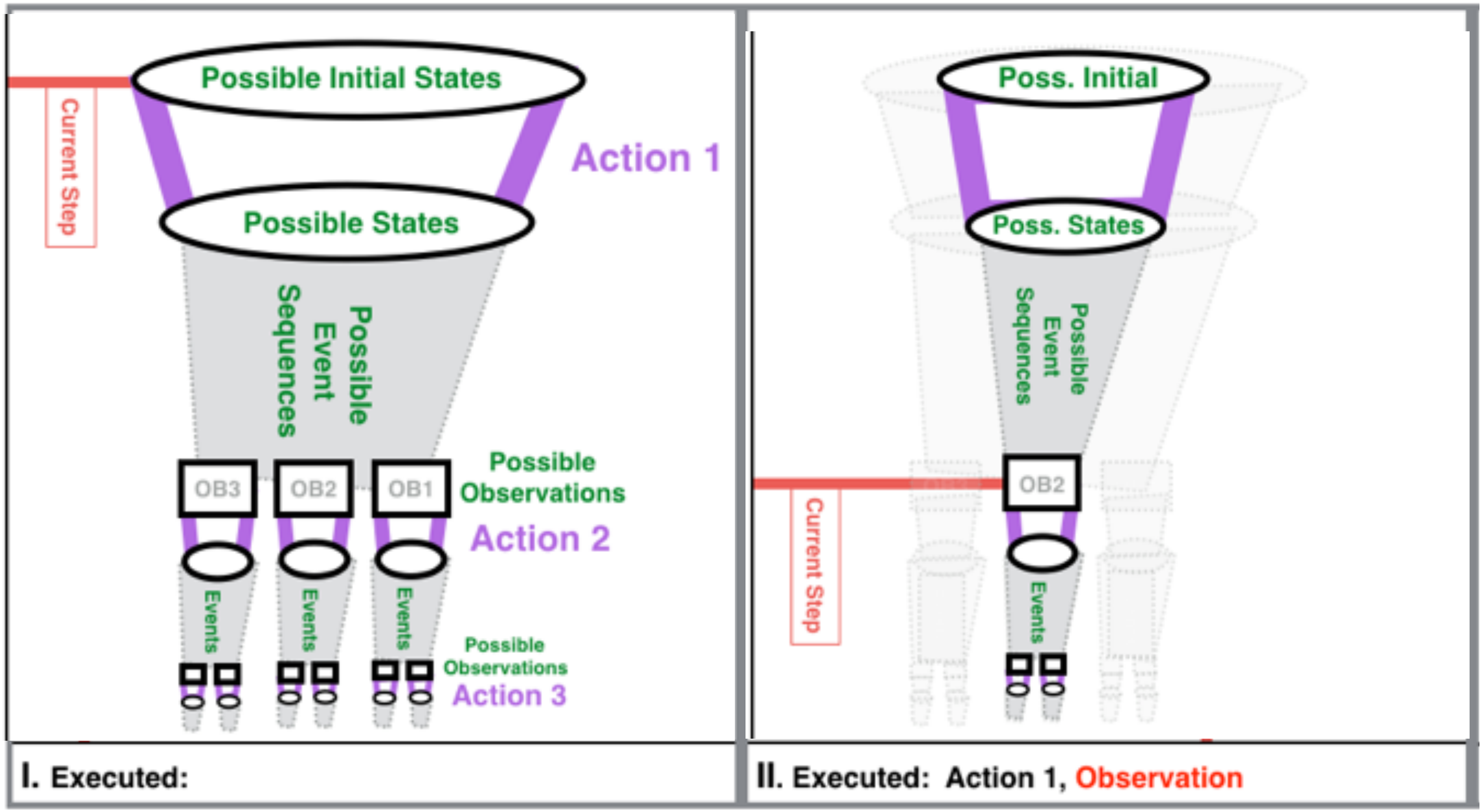


The situation/state-set space of possible explanations ***changes*** as our iterative execution progresses.

Putting It All Together

The Situation/State-set Space for Iterative Diagnosis in GR Agents

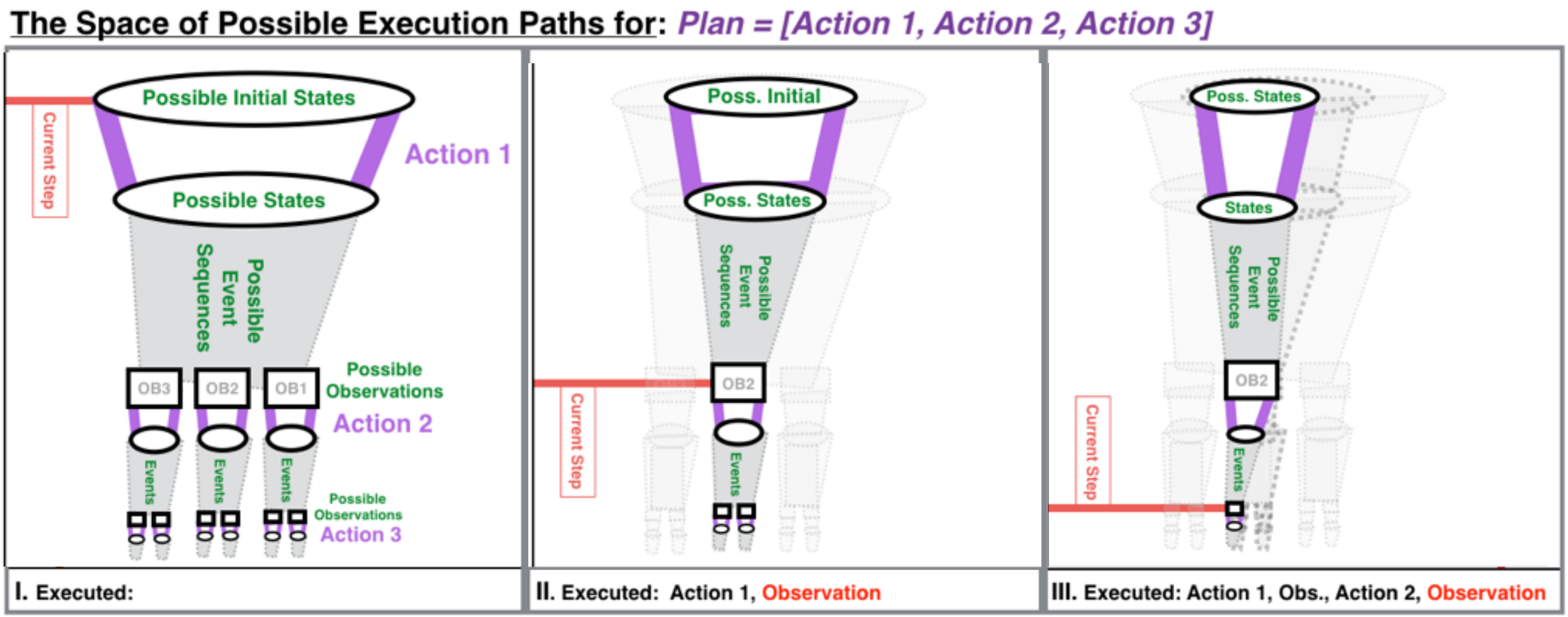
The Space of Possible Execution Paths for: *Plan = [Action 1, Action 2, Action 3]*



The situation/state-set space of possible explanations ***changes*** as our iterative execution progresses.

Putting It All Together

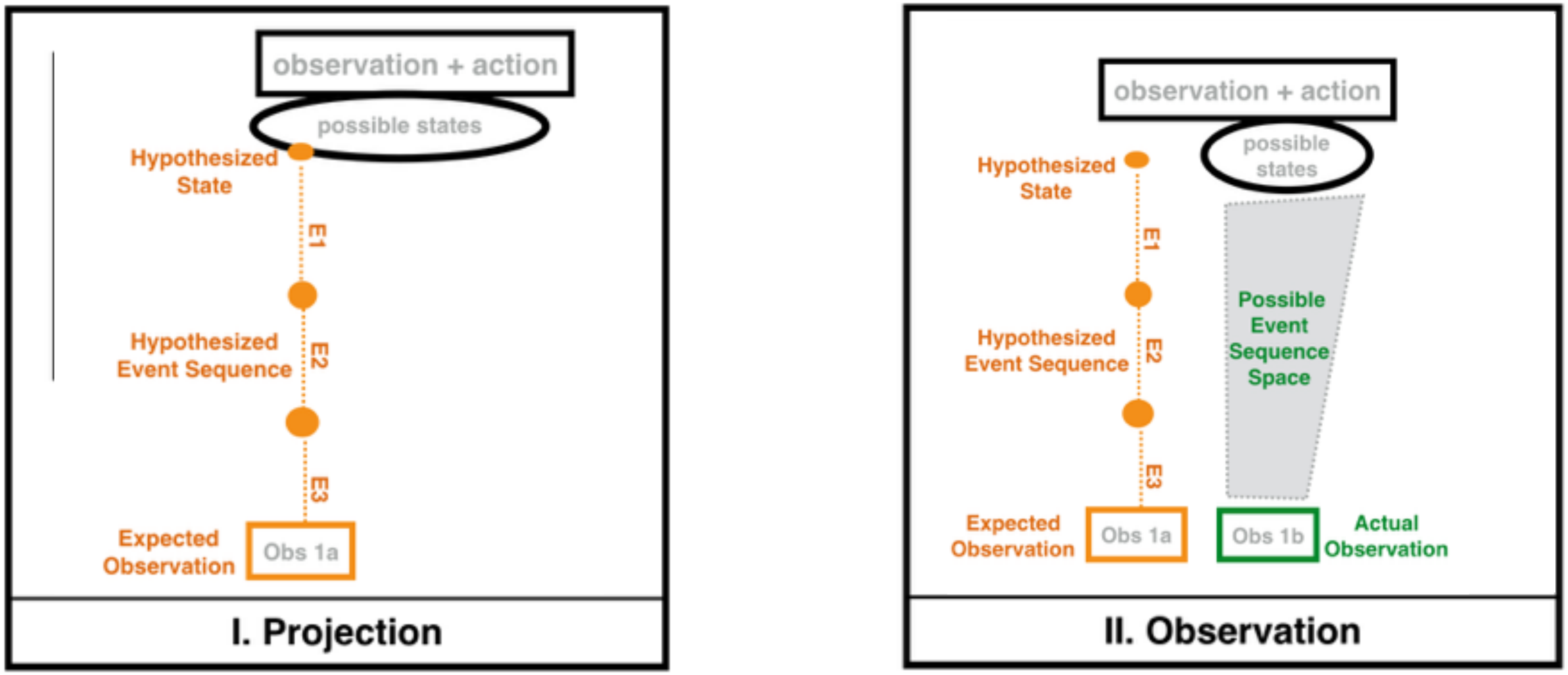
The Situation/State-set Space for Iterative Diagnosis in GR Agents



The situation/state-set space of possible explanations ***changes*** as our iterative execution progresses.

Putting It All Together

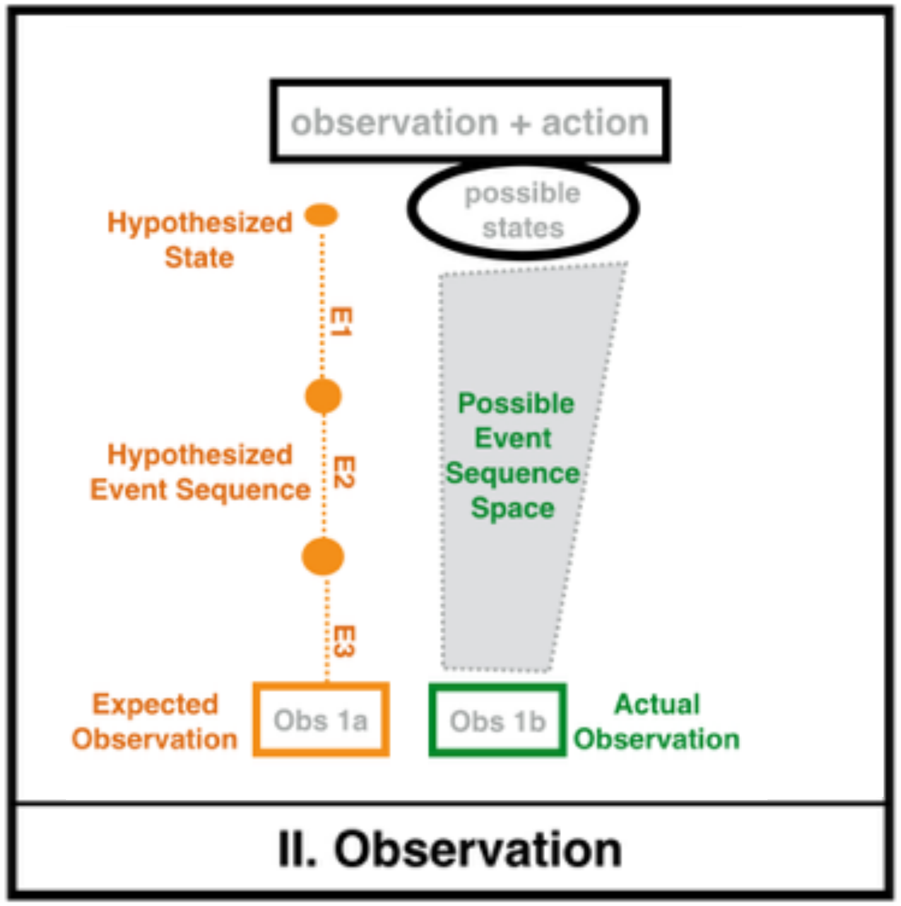
The Situation/State-set Space for Iterative Diagnosis in GR Agents

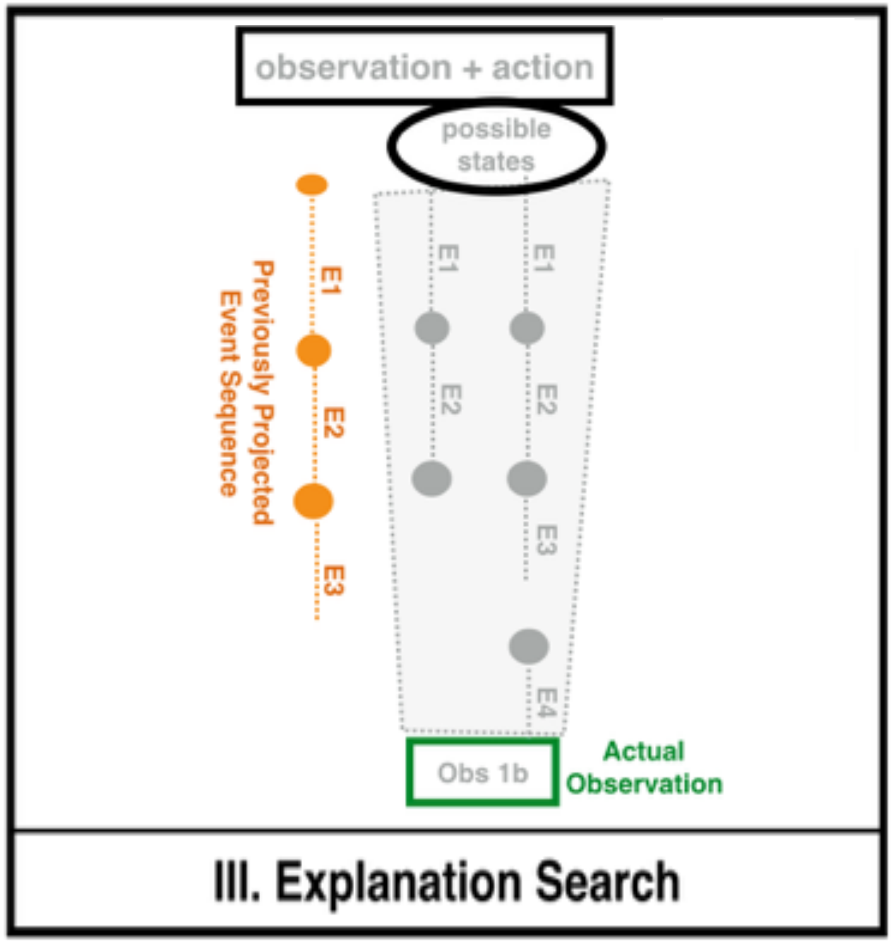


This formalism allows us to depict, and better understand, how the DiscoverHistory explanation search algorithm interacts with its solution space. [Molineaux and Aha 2015]

Putting It All Together

The Situation/State-set Space for Iterative Diagnosis in GR Agents

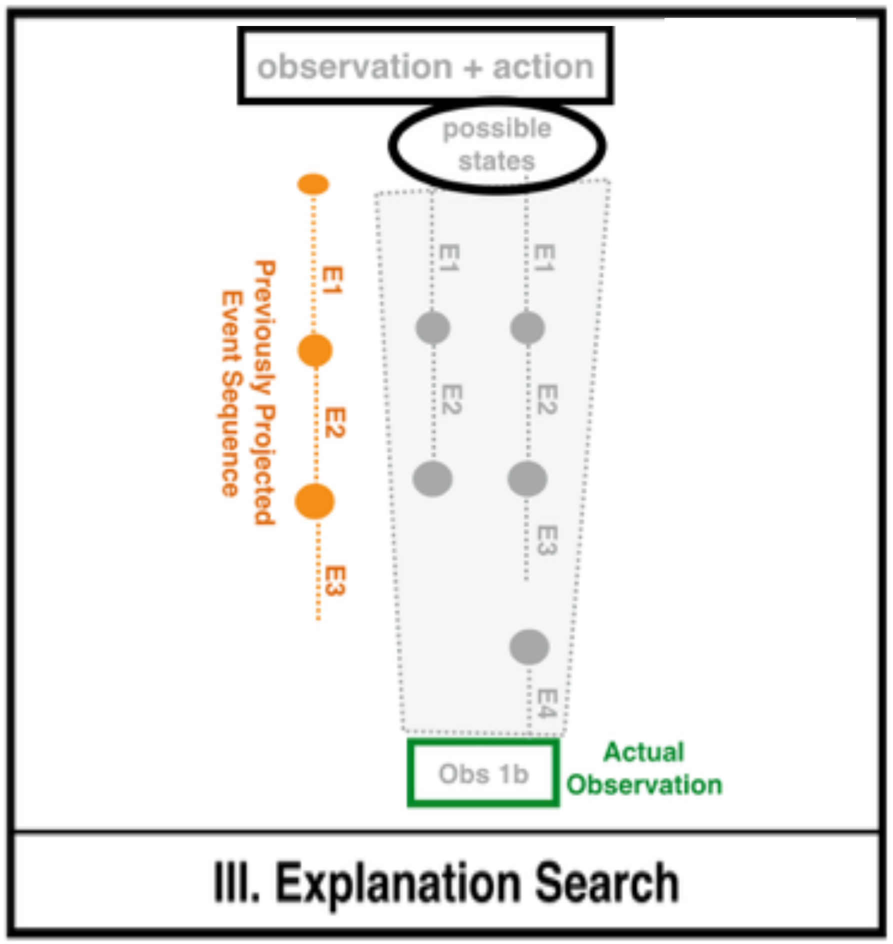




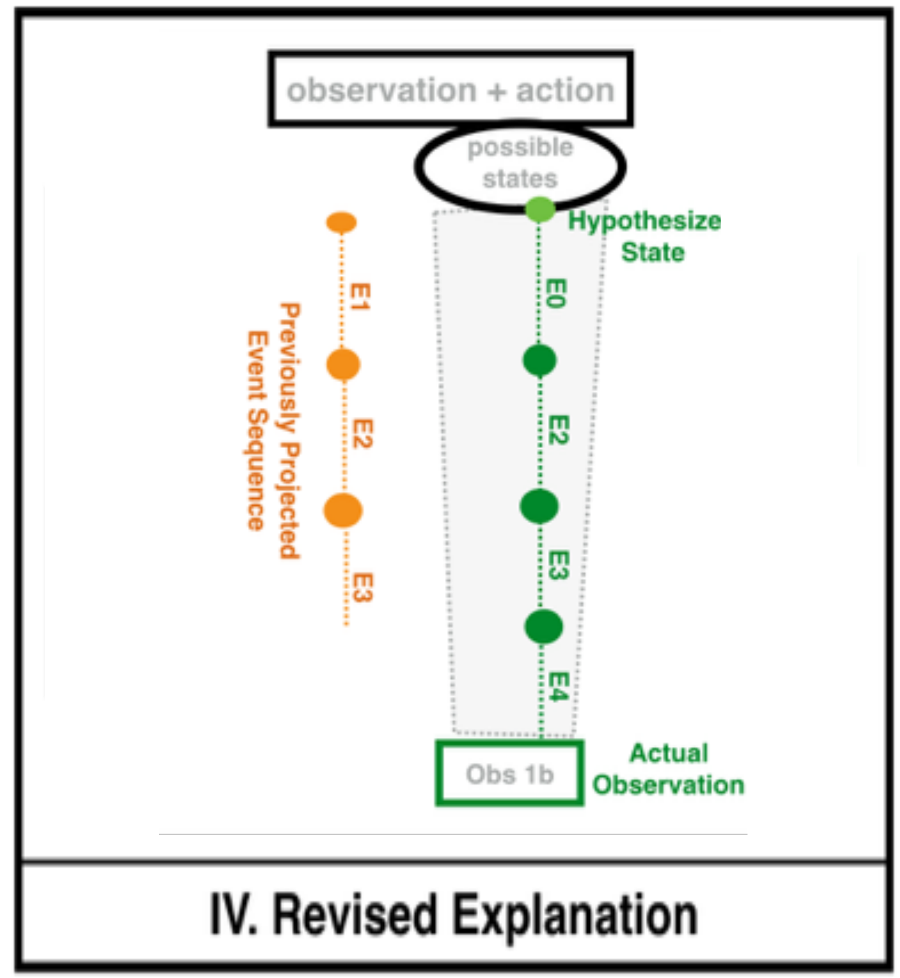
This formalism allows us to depict, and better understand, how the DiscoverHistory explanation search algorithm interacts with its solution space. [Molineaux and Aha 2015]

Putting It All Together

The Situation/State-set Space for Iterative Diagnosis in GR Agents



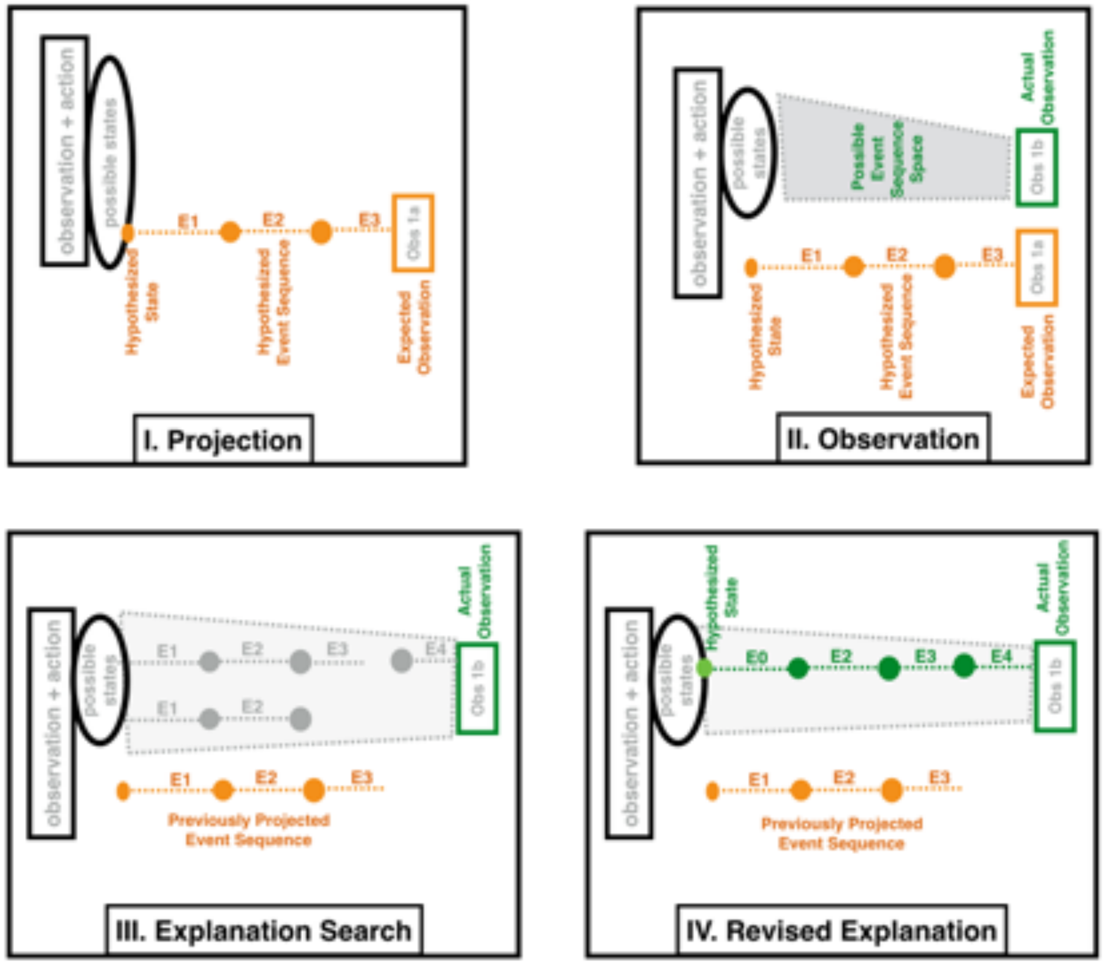
III. Explanation Search



IV. Revised Explanation

This formalism allows us to depict, and better understand, how the DiscoverHistory explanation search algorithm interacts with its solution space. [Molineaux and Aha 2015]

Insights, Potential Directions for Optimizations, Future Work

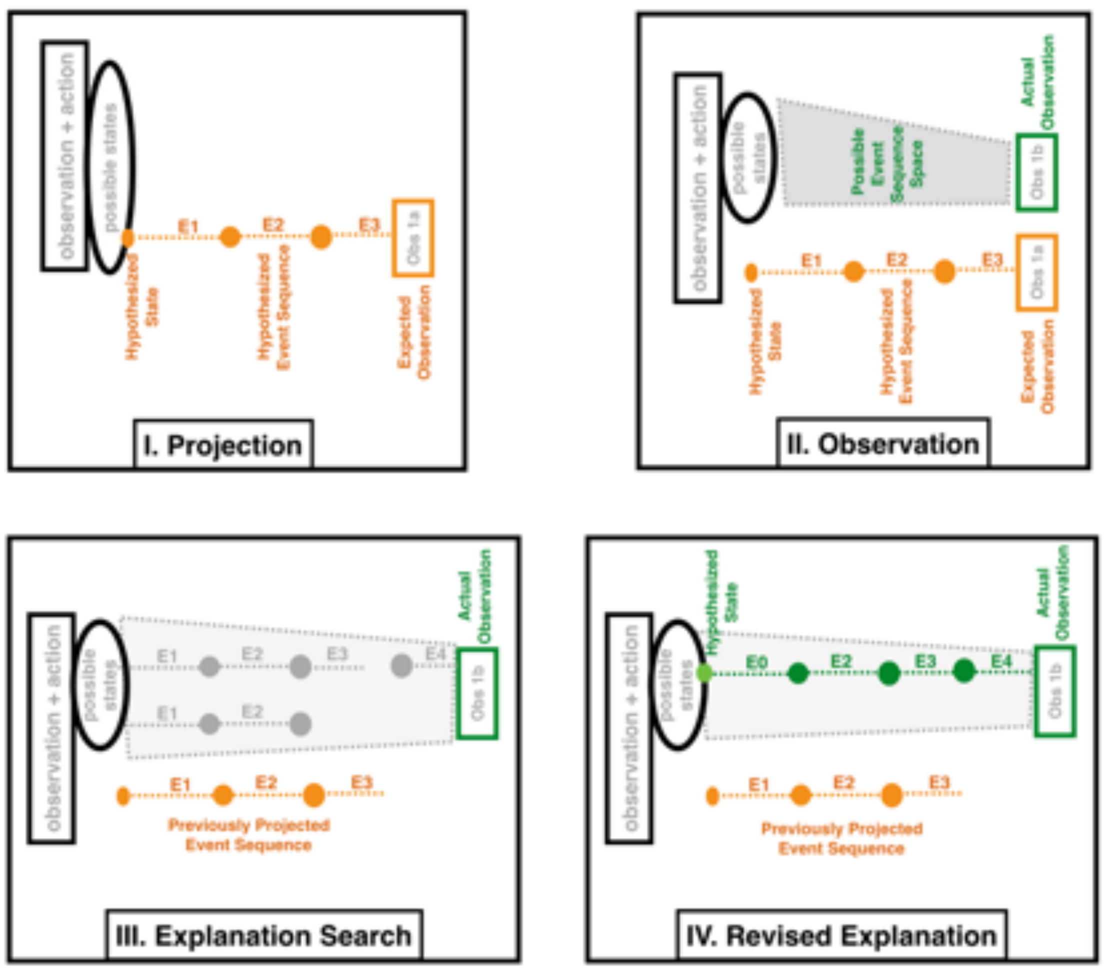


What insights might we take from this problem formalization?

- Guide explanation search to remain within space of possible event sequences (using case-based learned, or directly computed explanation sub-sequences?)

Using It

Insights, Potential Directions for Optimizations, Future Work

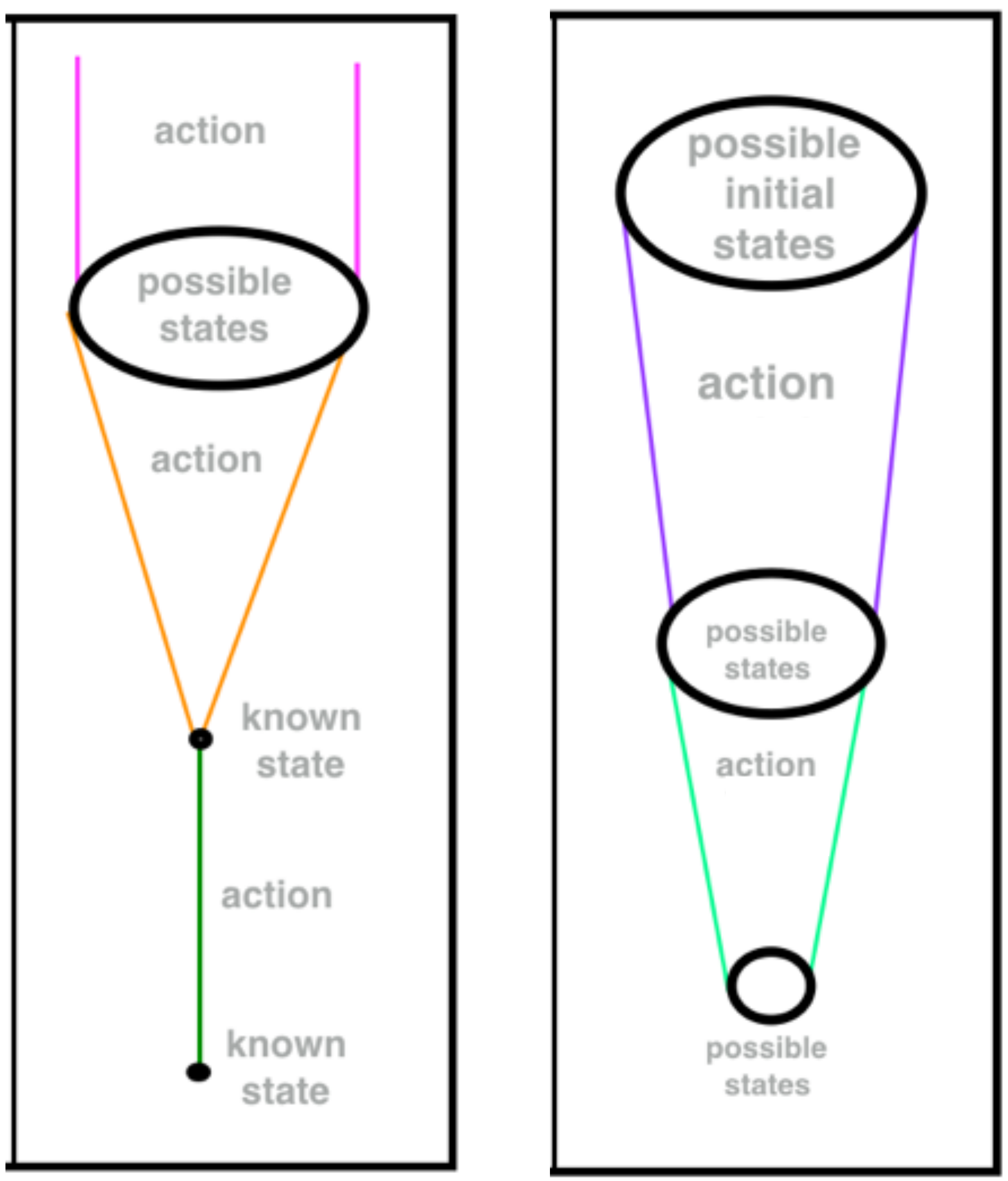


What insights might we take from this problem formalization?

- Guide explanation search to remain within space of possible event sequences (using case-based learned, or directly computed explanation sub-sequences?)
- Examine how decisions about domain modeling affect the size and complexity of the solution space that must be navigated?

Using It

Insights, Potential Directions for Optimizations, Future Work

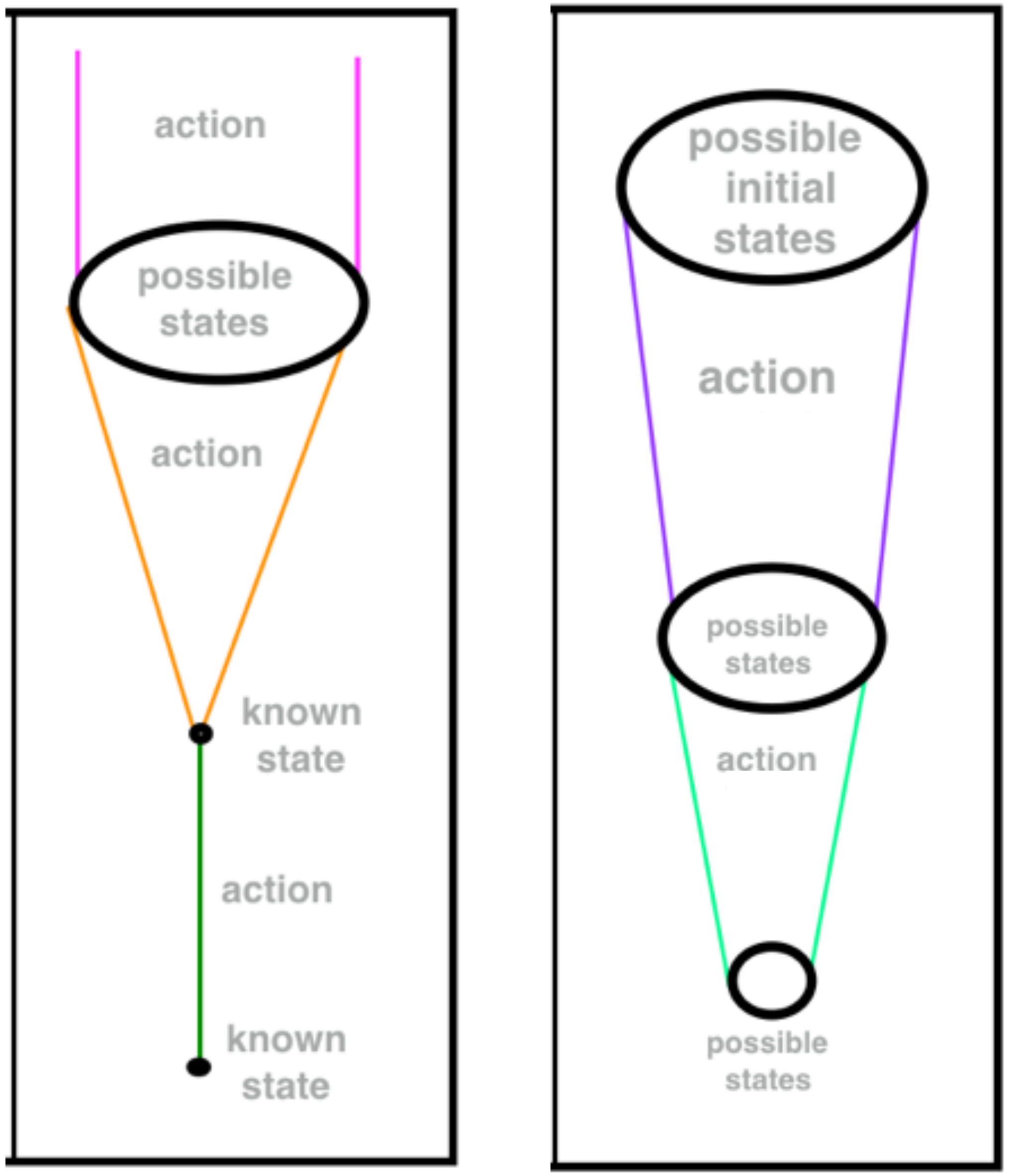


What insights might we take from this problem formalization?

- Quick sanity check: If we reach a fully known state in our execution, will we ever be able to correctly infer (with certainty) the event sequences and states that occurred before that state?

Using It

Insights, Potential Directions for Optimizations, Future Work

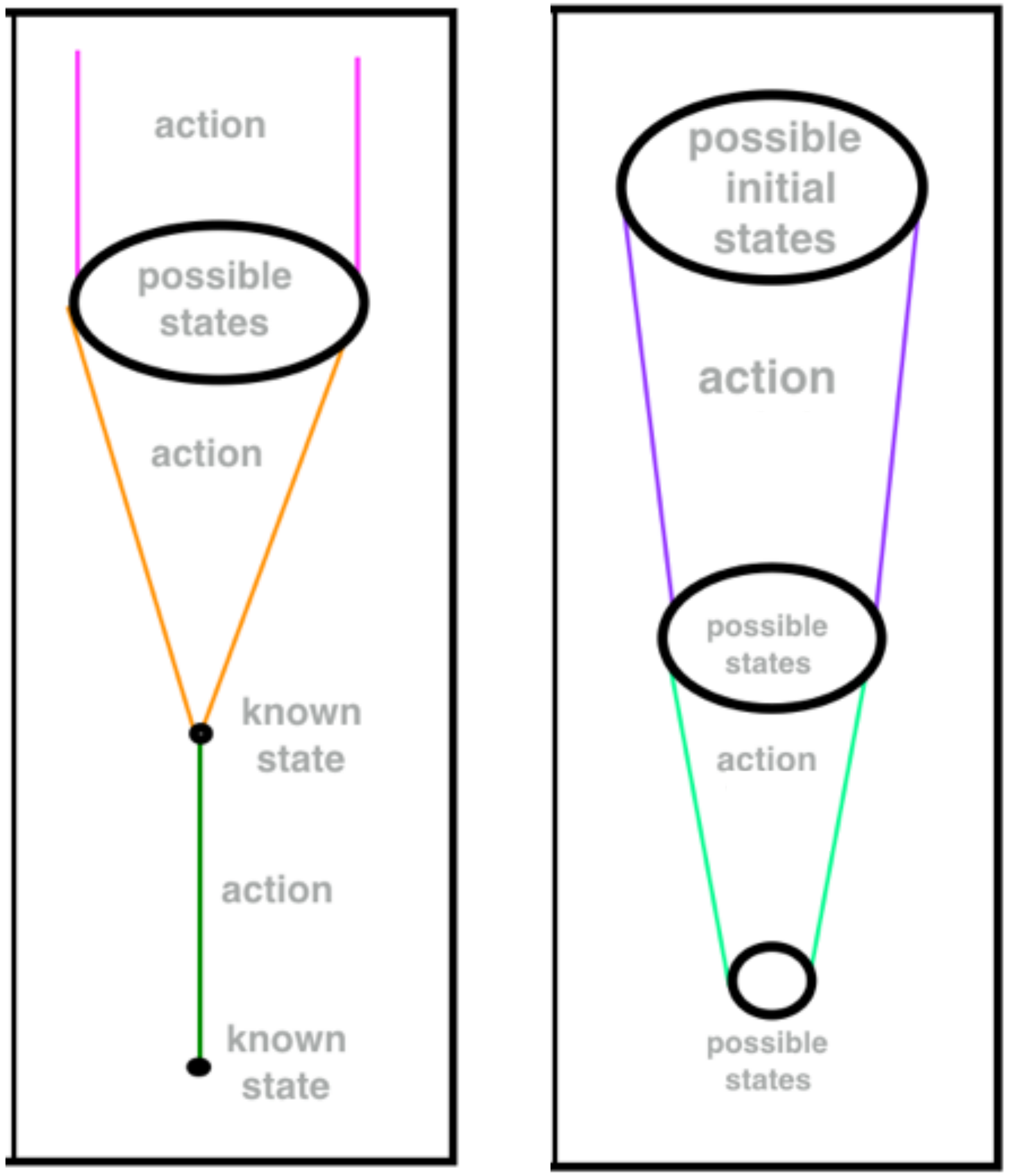


What insights might we take from this problem formalization?

- In general, the monotonically decreasing state-set sizes is an indicator of information about our history being destroyed. We cannot distinguish any more histories than we have current possible states.

Using It

Insights, Potential Directions for Optimizations, Future Work

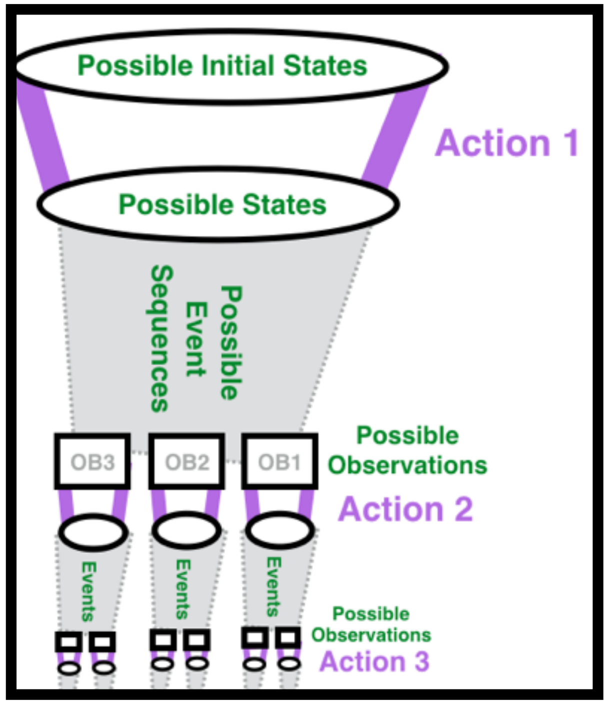


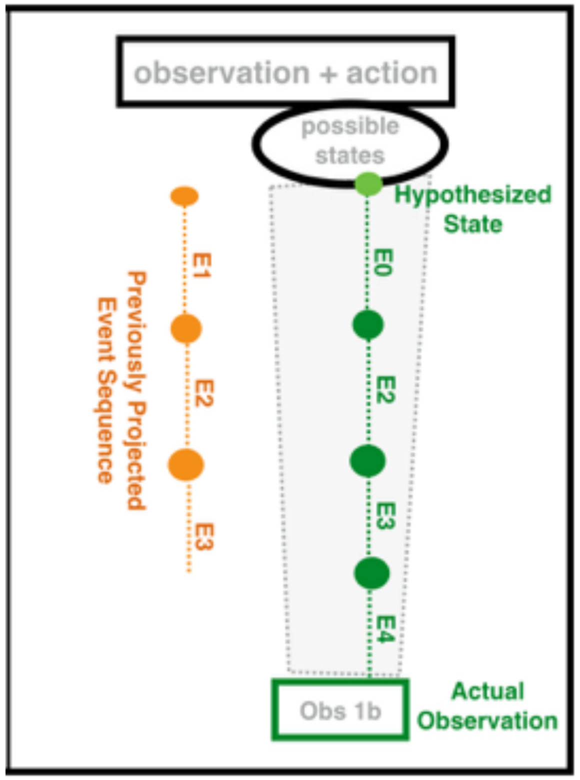
What insights might we take from this problem formalization?

- In general, the monotonically decreasing state-set sizes is an indicator of information about our history being destroyed. We cannot distinguish any more histories than we have possible current states.

- *Can we pick a good point to cut-off attempting to reasoning about our past?*

Questions?





Any questions?

…suggestions, ideas, insights, criticisms,
monologues, short poems, chili recipes…

Christine Task
Knexus Research Corporation
ctask@knexusresearch.com