Beliefs representation and management

Autonomous Software Agents

A.A. 2024-2025

Prof. Paolo Giorgini

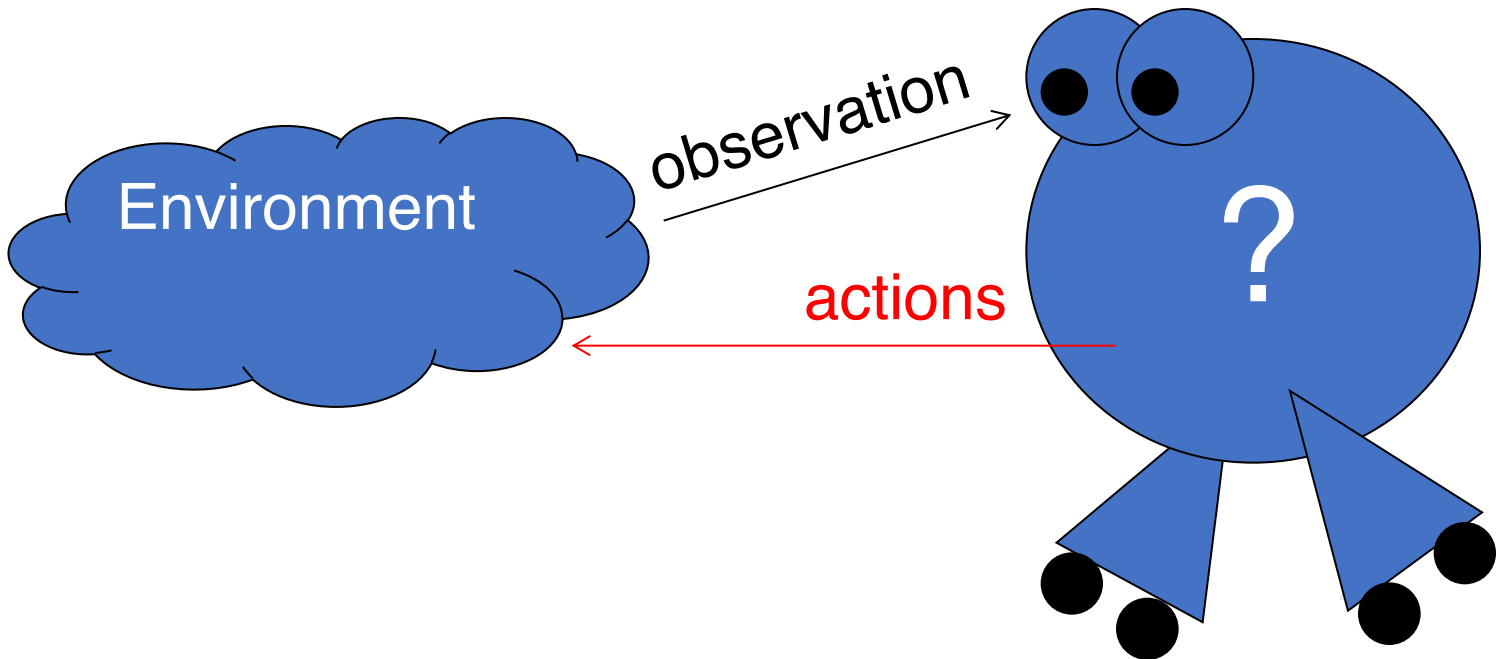
Dr. Marco Robol

UNIVERSITY OF TRENTO - Italy

Department of Information and Communication Technology

Belief Revision

- An agent revise beliefs in response to new information
  - But which ways of revising beliefs are “OK” and which are not?
- A belief revision theory is meant to provide a general answer, with a sense of “OK” that it specifies



An example

α: *All European swans are white*

β: *The bird caught in the trap is a swan*

γ: *The bird caught in the trap comes from Sweden*

δ: *Sweden is part of Europe*

If the agent's database is coupled with a program that can compute logical inferences:

ε: *The bird caught in the trap is white*

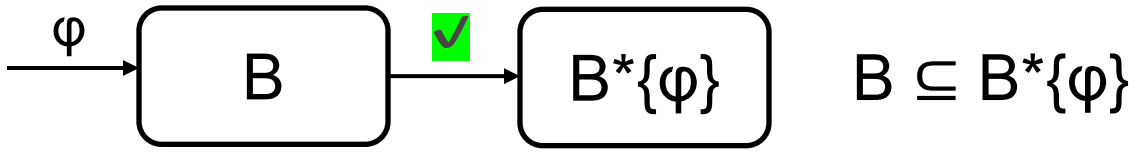
Now: the bird caught in the trap turns out to be black. This means that you want to add the fact ¬ε. But then the database becomes *inconsistent*.

Belief Revision

- To keep the database consistent, you need to *revise* it
  - Some of the beliefs in the original database must be retracted
  - You don't want to give up all of the beliefs since this would be an unnecessary loss of valuable information
  - choose between retracting α, β, γ or δ
- Problem: beliefs in a database have logical consequences, so which of the consequences to retain and which to retract
- If retract α, α has as logical consequences, among others, the following two:
- α’: *All European swans except the one caught in the trap are white*
  α": *All European swans except some of the Swedish are white*
- Do you want to keep any of these sentences in the revised database?

Belief Revision Theory

- A belief revision theory is, roughly, a theory saying which ways of belief revision are OK and which are not
- **Preservation**. If the information that agent A receives at time t is *compatible* with the set of the beliefs that A has right before t, then, right after t, agent A retains all of her beliefs in response to that information
- If ϕ is compatible with B, then B is a subset of $B * \phi$, where:
  - B is the set of one's beliefs right before the receipt of new information,
  - ϕ is the new information one receives,
  - $B * \phi$ is the set of one's new beliefs in response to new information ϕ



Belief Revision Theory

- **Preservation Thesis (the "Perfect Rationality" Version).** One is perfectly rational only if one has never violated, and would never violate, Preservation.
- *Example (Three Composers).* The agent initially believes the following about the three composers Verdi, Bizet, and Satie.
  - (a) Verdi is Italian
  - (b) Bizet is French
  - (c) Satie is French

  Then the agent receives this information
  - (e) Verdi and Bizet are compatriots

Example (Three Composers)

- The agent drops her beliefs (a) and (b), and retains the belief in (c) that Satie is French (after all, information (e) has nothing to do with Satie).
- The agent comes to believe the new information (e) that Verdi and Bizet are compatriots, while suspecting that Verdi and Bizet might both be Italian, and that they might both be French.
- At this stage, the agent does not rule out the possibility that Verdi is French (and, hence, a compatriot of Satie).
- So, what she believes at this stage is compatible with the following proposition.
- (f) Verdi and Satie are compatriots

Example (Three Composers)

- But then she receives a second piece of information, which turns out to be (f). Considering that she started with initial beliefs (a), (b), and (c) and received information (e) and (f), which jointly say that the three composers are compatriots, now she drops her belief (c).

- Information (f) is compatible with what she believes right before receiving this information, and she drops her belief in (c) nonetheless. So, this agent's second revision of beliefs violates Preservation. But there seems nothing in the specification of the scenario that prevents the agent from being perfectly rational. So, this seems to be a counterexample to the Preservation Thesis.

(a) Verdi is Italian

(b) Bizet is French

(c) Satie is French

B={a,b,c}

1 Belief Revision (+e) ----------

- (e) not compatible with B

(e) Verdi and Bizet are compatriots

(c) Satie is French

B*{e} = {c,e}

2 Belief Revision (+ f) ----------

- (f) compatible with B{e}*

(f) Verdi and Satie are compatriots

(e) Verdi and Bizet are compatriots

B*{e}*{f} = {e,f}

{c,e} $\not\subseteq$ {e,f} *violates Preservation, but still the agent seems to be perfectly rational*

Methodological questions

- How are the beliefs in the database ***represented***?
  - Most databases work with elements like *facts* and *rules* as primitive forms of representing information. The code used to represent the beliefs may be more or less closely related to standard logical formalism. A mechanism for belief revision is sensitive to the formalism chosen to represent the beliefs
- What is the relation between the elements explicitly represented in the database and the beliefs that may be ***derived*** from these elements?
  - It depends on the *application area*. The nature of the relation between explicit and implicit beliefs is of crucial importance for how the belief revision process is attacked
- How are the choices concerning what to ***retract*** *made*?
  - Eg: An agent can decide adopting the minimal change principle or on the base of the importance of the beliefs. In CS, Integrity constraints is a common way of handling the problem

Three kinds of Belief Changes

In AGM theory - Alchourrón, Gärdenfors, and Makinson (1985)

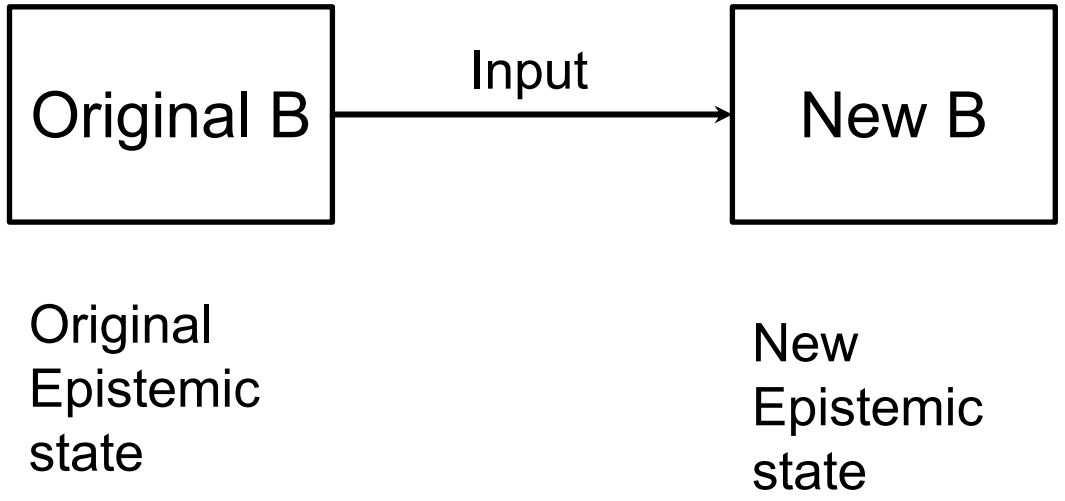
- *Expansion*: A new belief α is added to B together with the logical consequences of the addition (regardless of whether the larger set so formed is consistent). The new belief set is denoted by $B+\alpha$
- *Revision*: A new belief that is inconsistent with B is added, but, in order to maintain consistency of B, some of the old belief are deleted ($B\dot{+}\alpha$)
- *Contraction*: Some belief in B is retracted without adding any new belief. In order for the resulting B to be closed under logical consequences some other sentences from B must be given up $(B\dot{-}\alpha)$

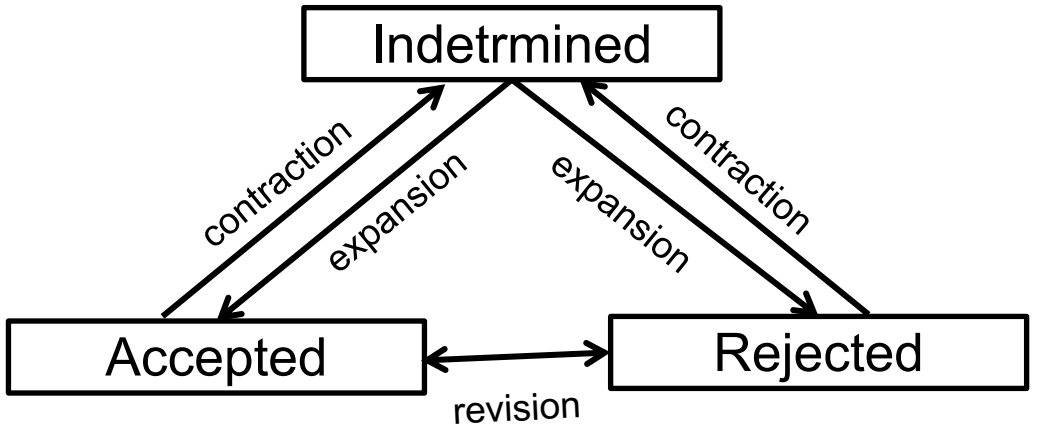
Three possibile attitudes for an agent:

Acceptance: $\alpha \in B$ (α is true)

Rejection: $\neg\alpha \in B$ (α is false, i.e. $\neg\alpha$ is true)

Indetermination: $\alpha \notin B \; nor \; \neg\alpha \notin B$





Belief Revision functions

- $B+\alpha$ can simply be defined as the logical closure of B together with α :

  $B+\alpha = \{\delta : B \cup \{\alpha\} \vdash \delta\}$

- Not so easy for revisions (and contractions)
  - There is no purely logical reason for making one choice rather than the other among the beliefs to be retracted, but we have to rely on additional information about these beliefs.
  - There are several ways of specifying the revision $B \dot{+} \alpha$.
  - General properties of a revision function has to be elaborated (i.e., algorithms can be found for computing revision functions – contraction functions)

Minimal change principle + others

- When we change our beliefs, we want to retain as much as possible from our old beliefs – we want to make a minimal change
  - Information is in general not gratuitous, and unnecessary losses of information are therefore to be avoided
  - This heuristic criterion may be called the criterion of informational economy
- + other criteria: consistency, preference, closure, primacy of the new information - not all of them are desirable at all times
- AGM postulates for contraction and revision
- Levi identity: $B\dot{+}\alpha = (B\dot{-} \neg\alpha) +\alpha$
  - (Revision = contraction + expansion)
- Harper identity: $B\dot{-}\alpha = (B\dot{+} \neg \alpha) \cap B$

Updates vs. Revisions

- Updating vs Revision - Katsuno and Mendelzon (1992)
  (1) new belief about a static world – **revision**
    - The information you receive corrects your knowledge about a world that has not changed
    - You need to modify your beliefs to align them with a reality that was already the case, but which you previously misunderstood
  (2) new belief about changes brought about by some agent – **updating**
    - Your beliefs were correct, but now the world has changed
    - The information you receive describes a change in the world
    - You need to modify your beliefs to reflect a new state of the world

Example Updating vs Revision 1/2

- In a room there is a table, a book and a magazine, and that either the book is on the table (β) , or (μ) the magazine is on the table, but not both, i.e., (β∧ ¬μ) ∨(μ∧¬β))
- A robot is then ordered to put the book on the table, and as a consequence, we learn that β. If we change our beliefs by revision, we should end up in a belief state that contains β ∧¬μ since β is consistent with B
- But, why should we conclude that the magazine is not on the table?
  - Only if you consider the **world to be static (Revision)**, then you are forced to discard μ, in order not to contradict your initial knowledge
    - → Your initial knowledge stated that both items cannot be on the table at the same time, so if you now know that β is true, you are compelled to reject μ

Example Updating vs Revision 2/2

- Let’s imagine that the **world has changed – updating**: the robot placed the book on the table
  - Previously, the magazine might have been on the table
  - Now, after the robot’s action, the book is on the table
- In this scenario:
  - There is no need to discard μ from the previous beliefs
  - The world has changed, so the initial beliefs no longer have to fully apply
  - Result: you can simply believe that β is true now, without having to infer ¬μ

Aspect	Revision	Update
Assumption	The world is **static**	The world has **changed**
Belief adjustment	You must **correct** your mistaken beliefs	You must **reflect** the new state of the world
New information	Reveals what was **already true**	Reveals what has **just changed**
Impact on beliefs	Must preserve consistency with prior knowledge	Prior knowledge may no longer apply
Example result	$\beta \wedge \neg\mu$	β (no need to reject μ)

Belief Revision

Let’s make it concrete

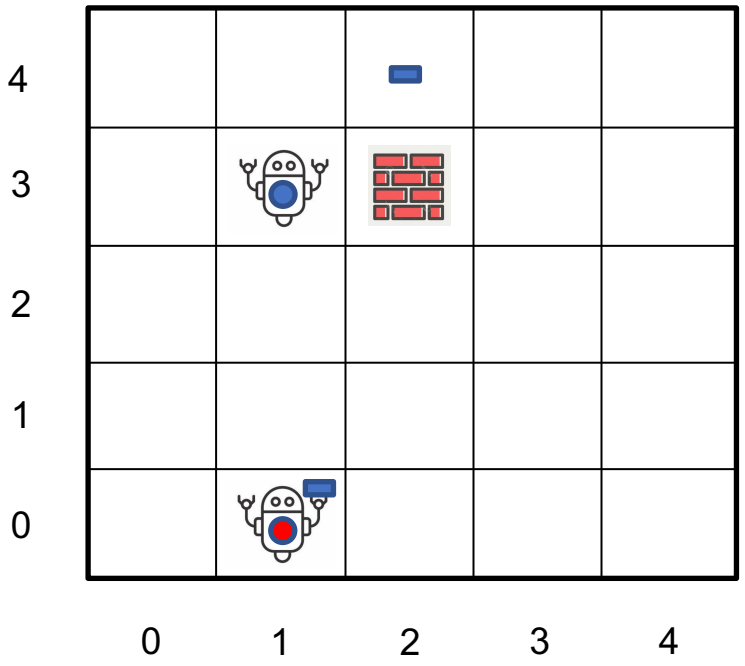
From sensors' data to Belief

- An agent acquires data from its sensors
  - E.g., temperature from the thermostat or its position from GPS

Time	Temperature	X	Y
09:00:00	8	3	3
10:00:00	9	3	4
11:00:00	11	3	5
12:00:00	13	4	6

  - Data can be stored as they are acquired
    - Acquisition time tells us actual values and it draws data evolution

Data completeness and correctness



Agent “Ag_1” (or ”Ag_2”)

Positions

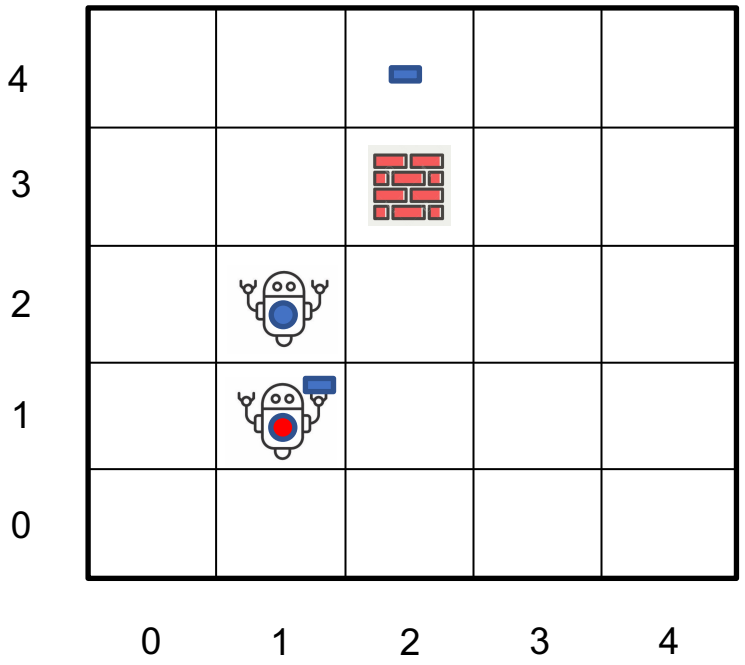
Time	Object	X	Y
1	Ag_1	1	0
1	Ag_2	1	3
1	Obst_1	2	3
1	Pack_1	2	4

Carry

Time	Agent	Pack
1	Ag_1	Pack_2

```
B = { In(Ag_1,1,0), In(Ag_2,1,3), In(Obst_1,2,3),
      In(Pack_1,2,4), carry(Ag_1,Pack_2) }
```

Data updating



Positions

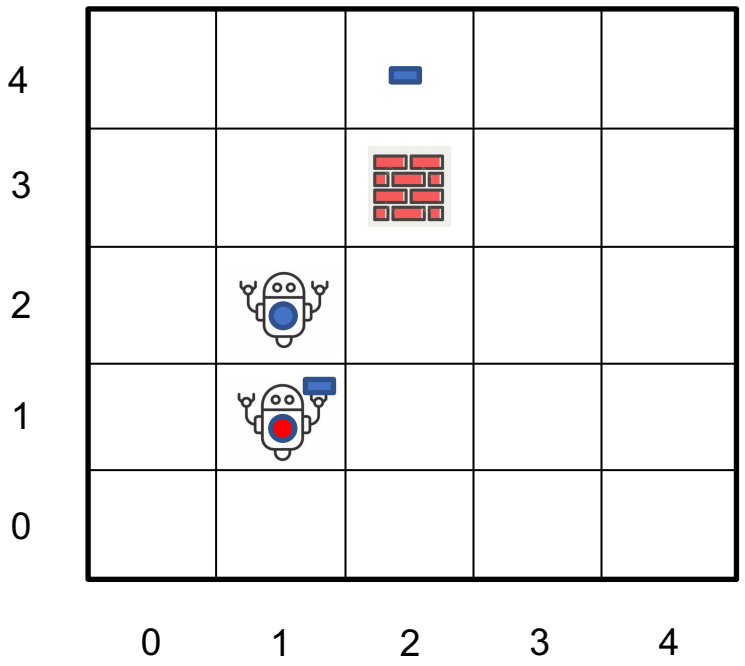
Time	Object	X	Y
~~1~~	~~Ag_1~~	~~1~~	~~0~~
~~1~~	~~Ag_2~~	~~1~~	~~3~~
~~1~~	~~Obst_1~~	~~2~~	~~3~~
~~1~~	~~Pack_1~~	~~2~~	~~4~~
2	Ag_1	1	1
2	Ag_2	1	2
2	Obst_1	2	3
2	Pack_1	2	4

Carry

Time	Agent	Pack
~~1~~	~~Ag_1~~	~~Pack_2~~
2	Ag_1	Pack_2

```
B = { In(Ag_1,1,1), In(Ag_2,1,2), In(Obst_1,2,3),
      In(Pack_1,2,4), carry(Ag_1,Pack_2) }
```

Derived Beliefs



Positions

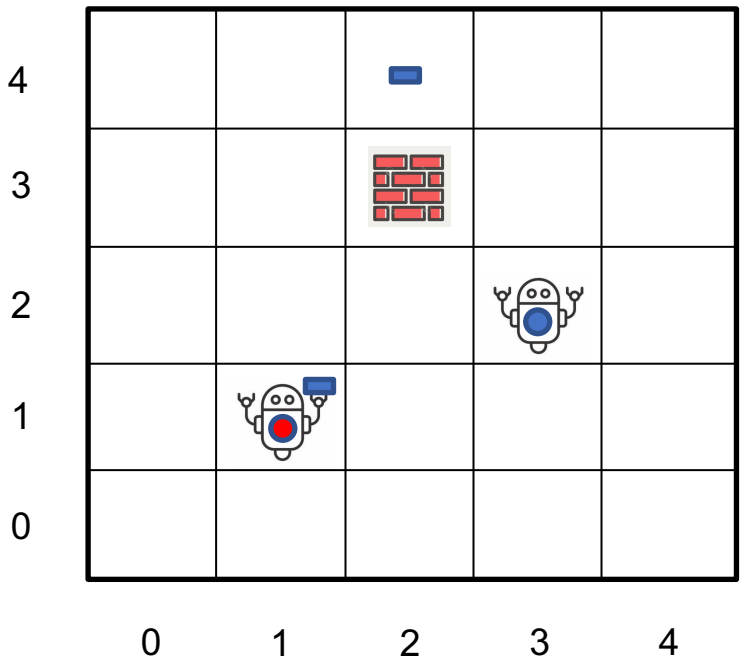
Time	Object	X	Y	Move
1	Ag_1	1	0	No
1	Ag_2	1	3	No
1	Obst_1	2	3	No
1	Pack_1	2	4	No
2	Ag_1	1	1	UP
2	Ag_2	1	2	DOWN
2	Obst_1	2	3	No
2	Pack_1	2	4	No

Carry

Time	Agent	Pack
1	Ag_1	Pack_2
2	Ag_1	Pack_2

```
B = { In(Ag_1,1,1), In(Ag_2,1,2), In(Obst_1,2,3),
      In(Pack_1,2,4), carry(Ag_1,Pack_2),
      move(Ag_1,UP),move(Ag_2,DOWN) }
```

Managing Inconsistencies



Positions

Time	Object	X	Y	Move
1	Ag_1	1	0	No
1	Ag_2	1	3	No
1	Obst_1	2	3	No
1	Pack_1	2	4	No
2	Ag_1	1	1	UP
2	Ag_2	1	2	DOWN
2	Obst_1	2	3	No
2	Pack_1	2	4	No
3	Ag_1	1	1	No
3	Ag_2	3	2	RIGHT
3	Obst_1	2	3	No
3	Pack_1	2	4	No

Carry

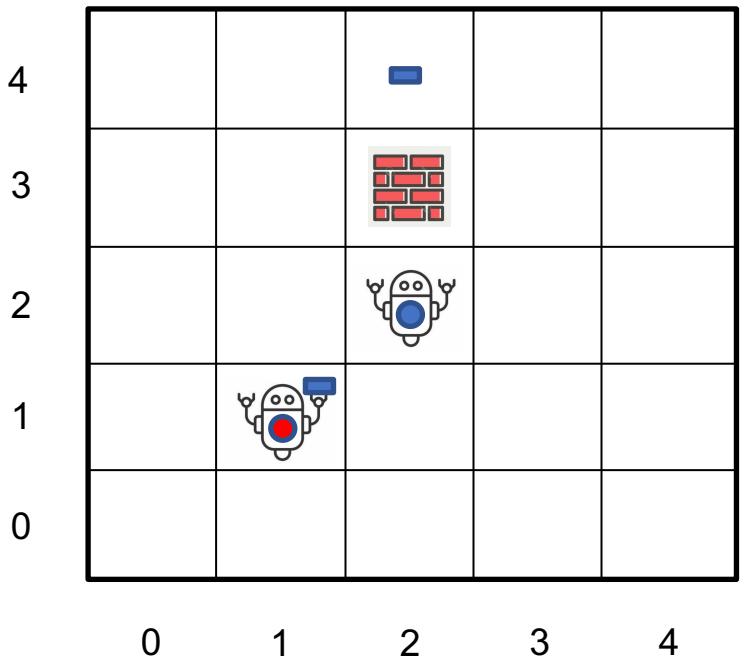
Time	Agent	Pack
1	Ag_1	Pack_2
2	Ag_1	Pack_2

Robot cannot move of two tiles in one step

Inconsistencies

- Inconsistencies can arise for several reasons
  - Sensors can send wrong data
  - Predefined rules are not anymore valid (evolution of the environment)
    - Now the robot can move of two tiles
  - Data are provided by other agents
    - They might lie or they could have wrong beliefs
- How much is it critical to solve the inconsistency?
  - Can we wait a little bit?
- How to solve them ?
  - Many different ways can be applied
    - `t=2 :In(Ag_2,1,2), move(Ag_2,DOWN)`
    - `t=3 :In(Ag_2,3,2), move(Ag_2,RIGHT)` ---→ `t=3 :In(Ag_2,2,2), move(Ag_2,RIGHT)`
- Policies/strategies to solve and manage inconsistencies should be part of the design

Managing Inconsistencies



Positions

Time	Object	X	Y	Move
1	Ag_1	1	0	No
1	Ag_2	1	3	No
1	Obst_1	2	3	No
1	Pack_1	2	4	No
2	Ag_1	1	1	UP
2	Ag_2	1	2	DOWN
2	Obst_1	2	3	No
2	Pack_1	2	4	No
3	Ag_1	1	1	No
3	Ag_2	2	2	RIGHT
3	Obst_1	2	3	No
3	Pack_1	2	4	No

Carry

Time	Agent	Pack
1	Ag_1	Pack_2
2	Ag_1	Pack_2

Solving the inconsistency
(most likely position for Ag_2?)

Managing inconsistencies

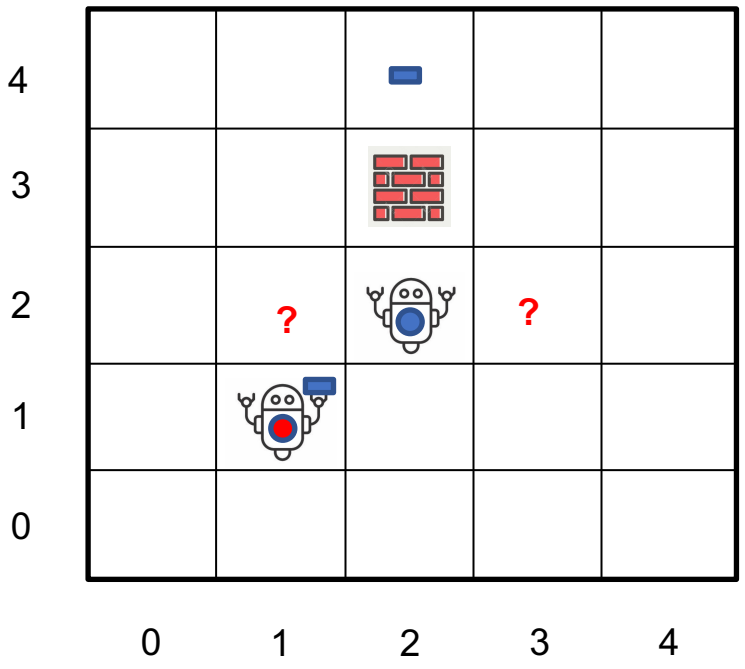
```
t=1: In(Ag_2,1,3)
t=2: In(Ag_2,1,2), move(Ag_2,DOWN)
t=3: In(Ag_2,3,2), move(Ag_2,RIGHT)
```

Possibile consistent sets:

```
t=1: In(Ag_2,1,3)
t=2: In(Ag_2,1,2), move(Ag_2,DOWN)
```

```
t=3: In(Ag_2,3,2), move(Ag_2,RIGHT)
```

Another example



t=1: In(Ag_2,**2**,**2**)
t=2: In(Ag_2,1,2), move(Ag_2,LEFT)
t=3: In(Ag_2,3,2), move(Ag_2,RIGHT)

Possibile consistent sets:

S_1

t=1: In(Ag_2,2,2)
t=2: In(Ag_2,1,2), move(Ag_2,LEFT)

S_2

t=1: In(Ag_2,2,2)
t=3: In(Ag_2,3,2), move(Ag_2,RIGHT)

More on the example

S_1

```
t=1: In(Ag_2,2,2)
t=2: In(Ag_2,1,2), move(Ag_2,LEFT)
```

what about t=3 ?

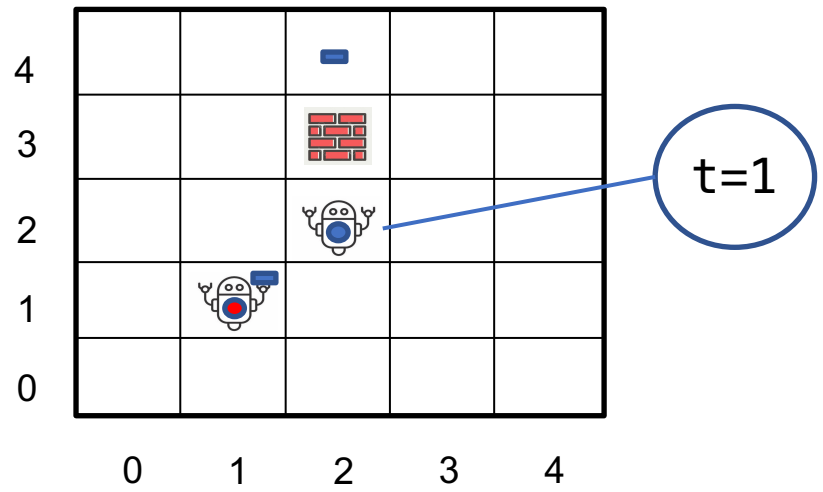
t=3: In(Ag_2,1,2) v In(Ag_2,0,2) v
In(Ag_2,1,3) v In(Ag_2,2,2)

S_2

```
t=1: In(Ag_2,2,2)
t=3: In(Ag_2,3,2), move(Ag_2,RIGHT)
```

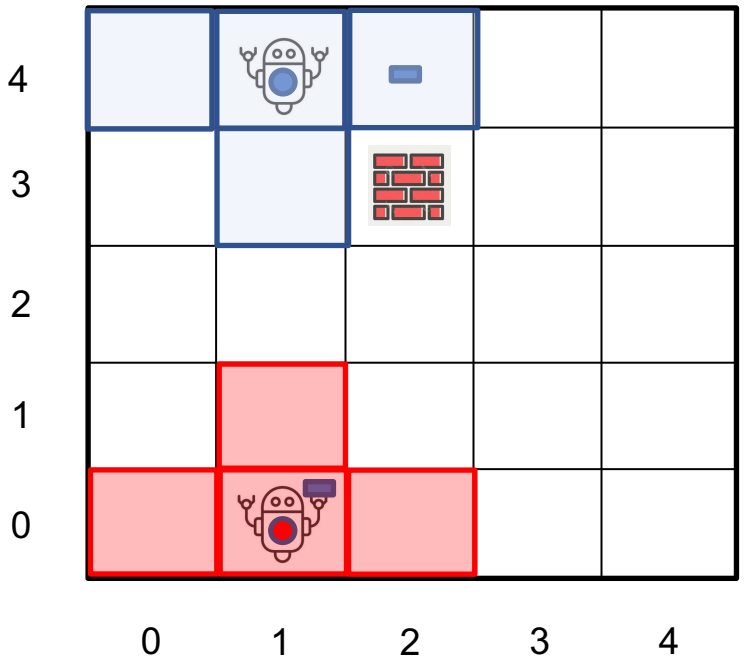
what about t=2 ?

t=2: In(Ag_2,2,2) v In(Ag_2,3,2)



After we had choose beween S1 and S2, should we update beliefs for t=3 and t=2, respectivly?

Partial view of the environment



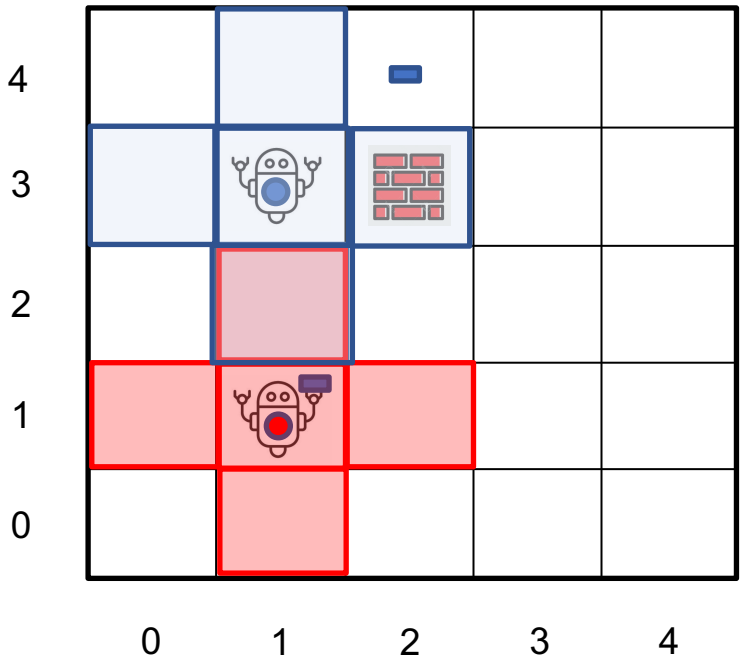
Agent 2

Time	Object	X	Y
1	Ag_2	1	4
1	Pack_1	2	4

Agent 1

Time	Object	X	Y
1	Ag_1	1	0

Partial view of the environment



Agent 2

Time	Object	X	Y
1	Ag_2	1	4
1	Pack_1	2	4
2	Ag_2	1	3
2	Obst_1	2	3

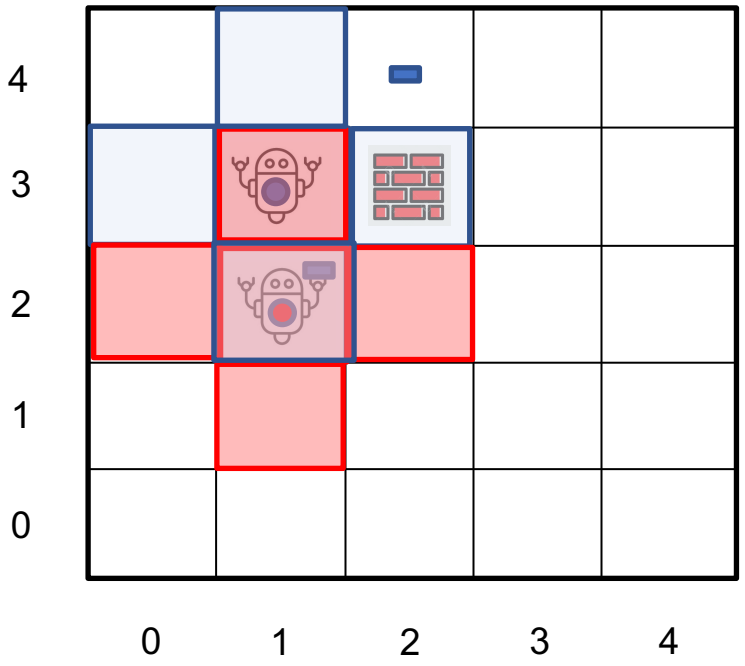
Agent 1

Time	Object	X	Y
1	Ag_1	1	0
2	Ag_1	1	1

At time t=2, is Pack_1 in (2,4)?

- YES ? NO? MAYBE? LIKELY?

Partial view of the environment



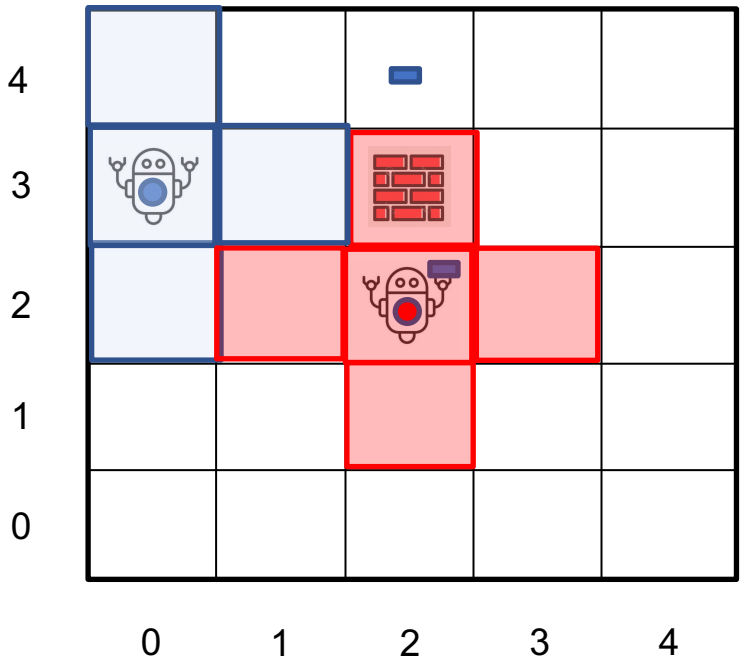
Agent 2

Time	Object	X	Y
1	Ag_2	1	4
1	Pack_1	2	4
2	Obst_1	2	3
2	Ag_2	1	3
3	Obst_1	2	3
3	Ag_2	1	3
3	Ag_1	1	2

Agent 1

Time	Object	X	Y
1	Ag_1	1	0
2	Ag_1	1	1
3	Ag_1	1	2
3	Ag_2	1	3

Partial view of the environment



Agent 2

Time	Object	X	Y
1	Ag_2	1	4
1	Pack_1	2	4
2	Obst_1	2	3
1	Ag_2	1	3
3	Obst_1	2	3
3	Ag_2	1	3
3	Ag_1	1	2
4	Ag_2	0	3

Agent 1

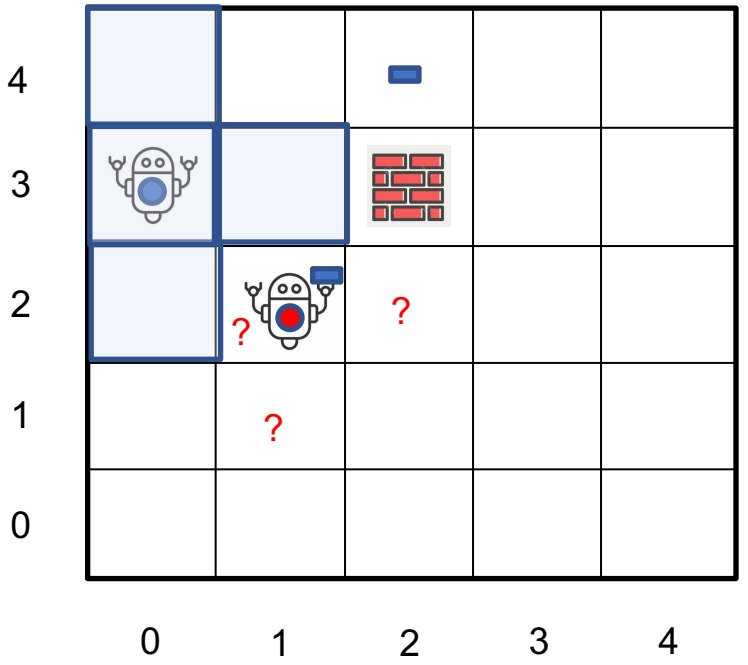
Time	Object	X	Y
1	Ag_1	1	0
2	Ag_1	1	1
3	Ag_1	1	2
3	Ag_2	1	3
4	Ag_1	2	2
4	Obst_1	2	3

They are not anymore there

As in the previous case

Agent 2

Time	Object	X	Y
1	Ag_2	1	4
1	Pack_1	2	4
2	Obst_1	2	3
2	Ag_2	1	3
3	Obst_1	2	3
3	Ag_2	1	3
3	Ag_1	1	2
4	Ag_2	0	3

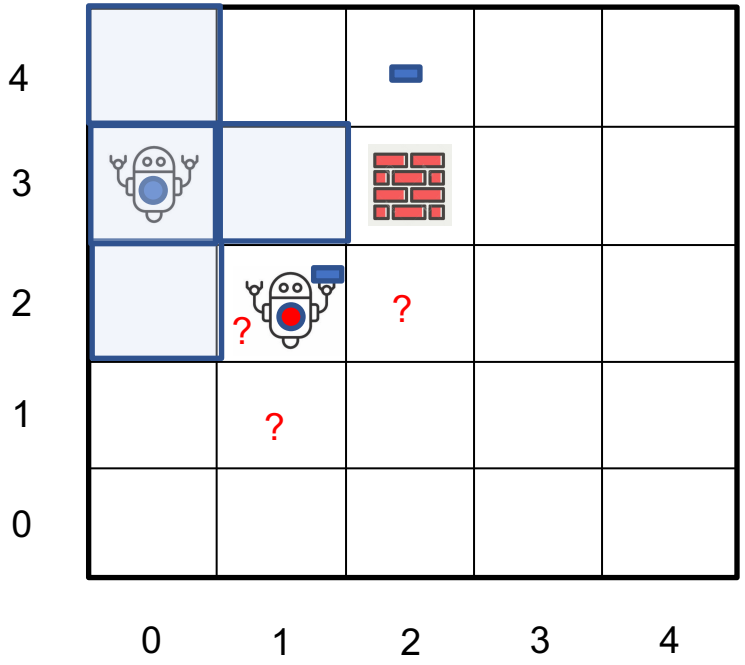


→ `t=4: In(Ag_1,1,2) v In(Ag_1,1,1) v In(Ag_1,2,2)`

Beliefs models

- Several beliefs models can be used
  - No memory: only beliefs based on current data
  - With memory: beliefs based on current data and keeping true not updated beliefs
  - With uncertainty
    - “the probability pack_1 that I saw long time ago is still in position (x,y) is very low”
    - ”the probability obst_1 is in position (x,y) is 1” (obst_1 is a wall a nobody can move it)
    - “I saw Ag_1 going in the direction of Pack_1, the probability Pack_1 is in position (x,y) is very low”
    - “Ag_1 was in position (x,y) and it was moving, the probability it is still there is very low”

In our example



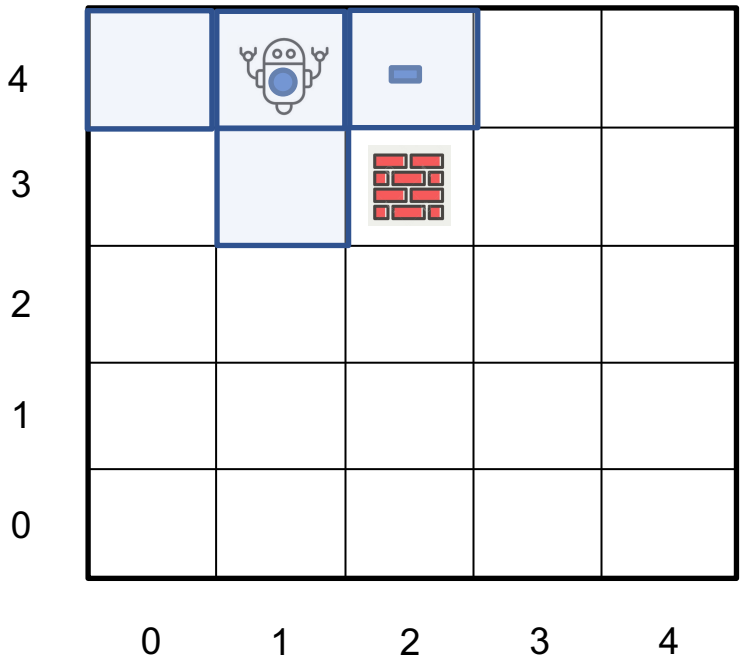
Agent 2

Time	Object	X	Y
1	Ag_2	1	4
1	Pack_1	2	4
2	Obst_1	2	3
2	Ag_2	1	3
3	Obst_1	2	3
3	Ag_2	1	3
3	Ag_1	1	2
4	Ag_2	0	3

→ `t=4: In(Ag_1,1,2) v In(Ag_1,1,1) v In(Ag_1,2,2)`

Time	Object	X	Y	Probability
4	Ag_2	0	3	1
4	Ag_1	1	2	0.33
4	Ag_1	1	1	0.33
4	Ag_1	2	2	0.33

Partial view of the environment



Agent 2

Time	Object	X	Y
1	Ag_2	1	4
1	Pack_1	2	4

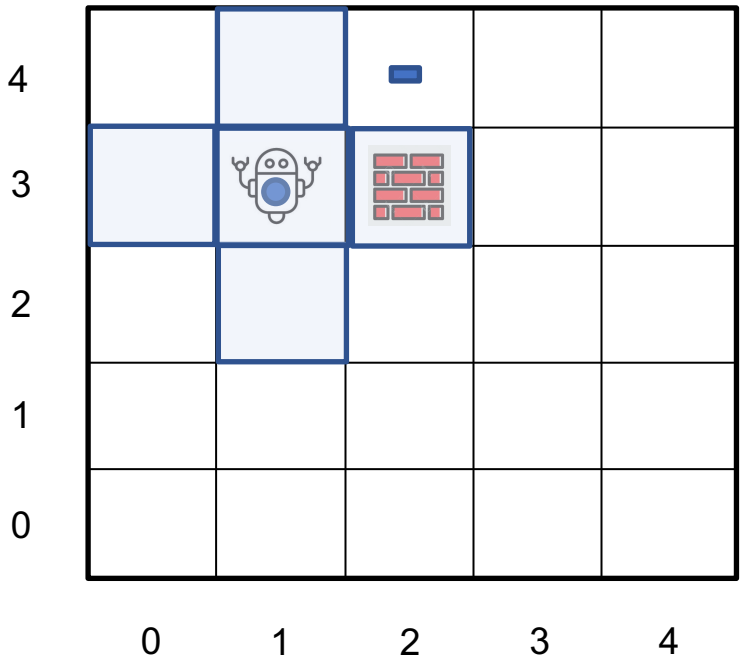
P(Pack_1@2,4) = 0.9

Pack_1 is at (2,4) with confidence **P(Pack_1@2,4) = 0.9`**

- Why 0.9?
- Agent 2 saw it (t=1)
- No info suggests it was picked up yet

it's **very likely** it is there

Partial view of the environment



Agent 2

Time	Object	X	Y
1	Ag_2	1	4
1	Pack_1	2	4
2	Ag_2	1	3
2	Obst_1	2	3

P(Pack_1@2,4) = 0.9

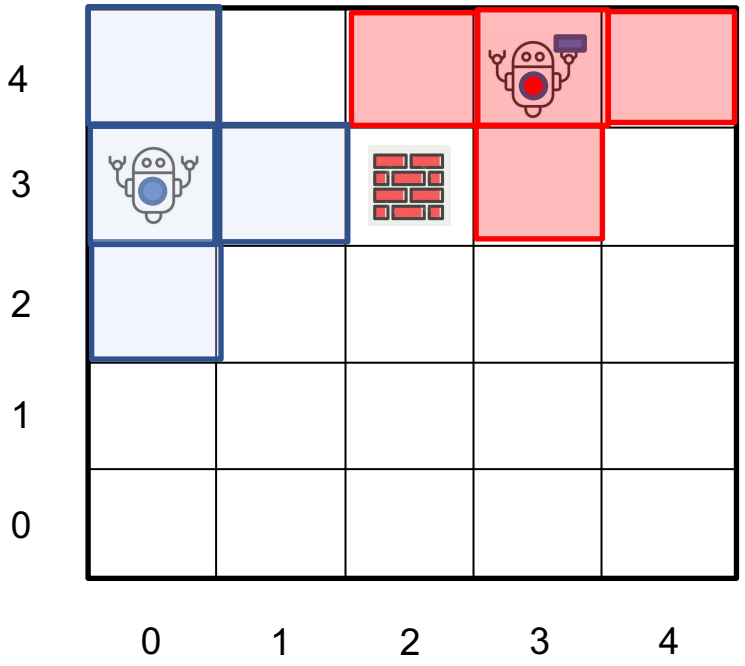
P(Pack_1@2,3 | not seen at t=2) = 0.5

Pack_1 is **not in Agent_2 view range** and it knows that **other agents may have picked it up**

P(P1@2,4 | not seen at t=2) = 0.5

- Partial observability
- Time elapsed since last observation
- Presence of other agents

Partial view of the environment



Agent 2

Time	Object	X	Y
1	Ag_2	1	4
1	Pack_1	2	4
2	Obst_1	2	3
2	Ag_2	1	3
3	Obst_1	2	3

Agent 1 enters view of (2,3)

Agent broadcasts:

¬In(P1,2,3) (Package not found at 2,3)

Agent A2 revises its belief:

- Previous: P(P1@2,4) = 0.5
- New info: A2 did not see Pack_1
- A2 trusts A1's sensor 80% of the time
- Use **Bayesian-style belief revision**

Bayesian-style belief revision

$$P(H|E) = \frac{P(E|H) \cdot P(H)}{P(E)}$$

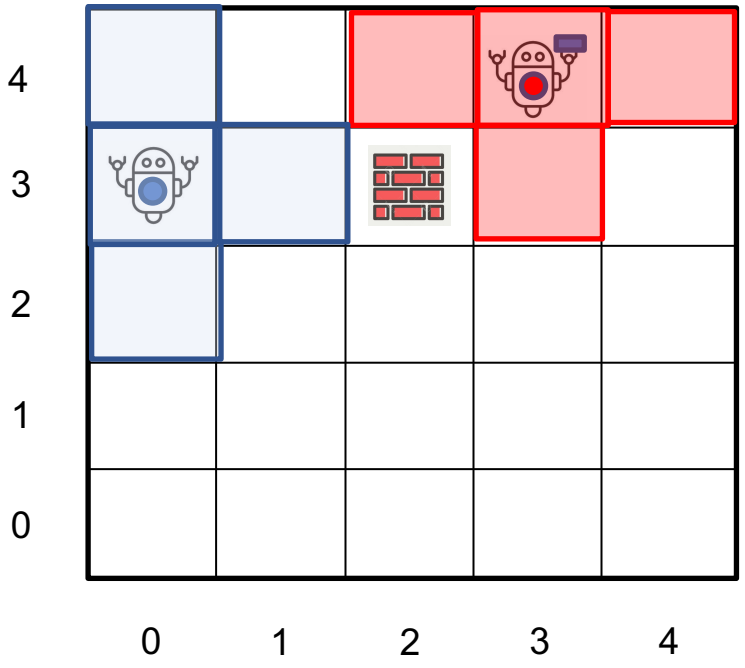
$$P(P1@2,4|\ \neg Seen) = P(\neg Seen\ |\ P1@2,4)\ x\ P(P1@2,4)/\ P(\neg Seen)$$

Assume:

- If P1 is at 2,4 → A1 has 80% chance of seeing it → P(¬Seen | P1@2,4) = 0.2
- If P1 is not at 2,4 → P(¬Seen | ¬P1@2,4) = 1

$$P(P1@2,4|\ \neg Seen) = (0.5x0.2)/[(0.5x0.2)+(0.5x1)] = 0.167$$

Partial view of the environment



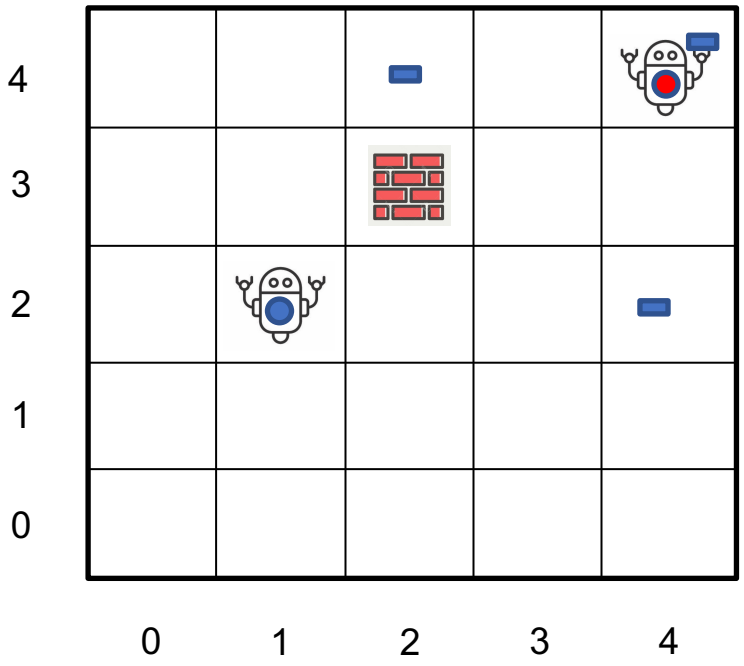
Agent 2

Time	Object	X	Y	P
1	Ag_2	1	4	
1	Pack_2	2	4	**0.9**
2	Obst_1	2	3	
2	Ag_2	1	3	
2	Pack_2	2	4	**0.5**
3	Obst_1	2	3	**1**
3	Ag_2	1	3	
2	Pack_2	2	4	**0.167**

Agent 2 knows that walls cannot be moved away

Since **P(P1@2,4) < threshold (e.g., 0.3)**, A2 **does not generate the option** to pick P1 up and **searches for another package instead**

Partial view of the environment



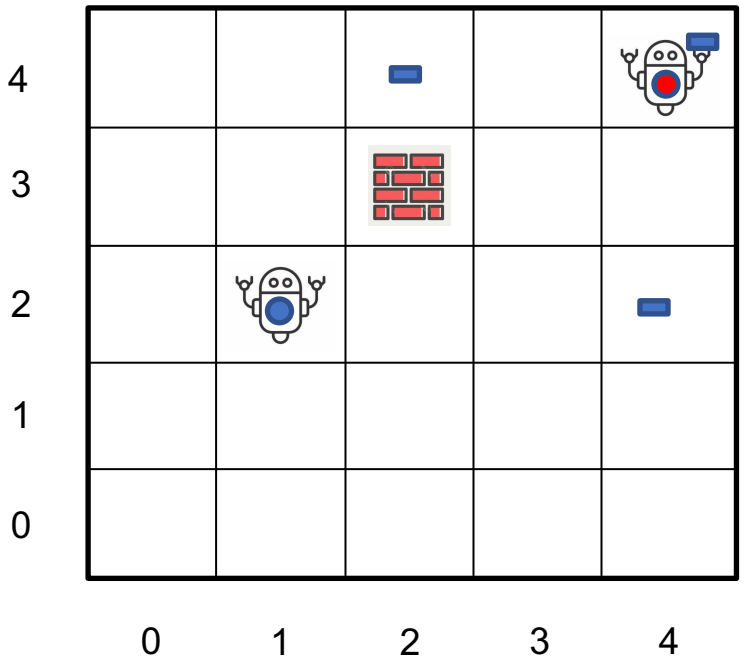
Agent 2

Time	Object	X	Y
1	Ag_2	1	2
1	Pack_1	2	4
1	Pack_2	4	2
1	Ag_1	4	4

Distance Ag_1, Pack_1, Pack_2 = 2

Distance Ag_2, Pack_1, Pack_2 = 3

More on Belief estimation



Agent 2

Time	Object	X	Y	P
1	Ag_2	1	2	
1	Pack_1	2	4	?
1	Pack_2	4	2	?
1	Ag_1	4	4	

How does Ag_2 estimate the probability that **Pack_1 is still at (2,4)** and **Pack_2 is still at (4,2)** at a future time, given that Ag_1 might have moved toward them in the meantime?

Belief Estimation

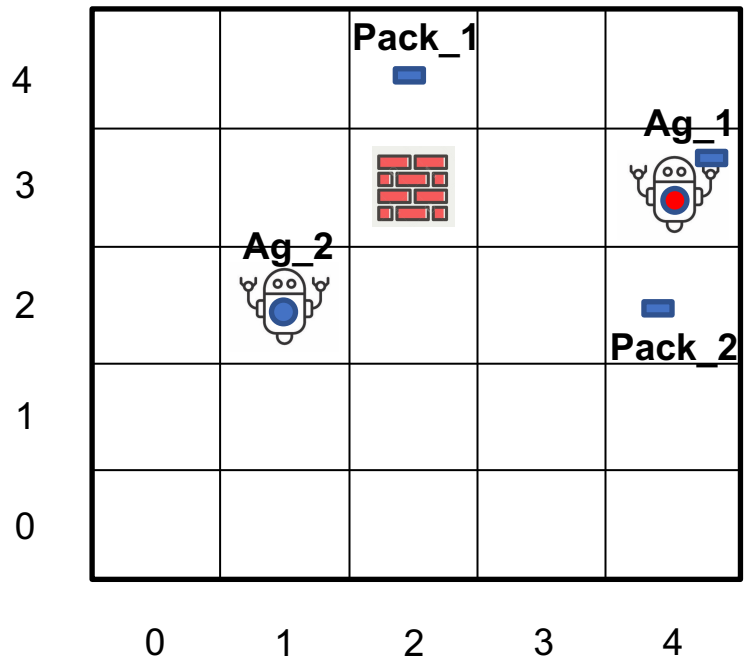
- **Belief estimation problem under uncertainty**, based on:
  - Agent **positions**
  - **Time** since last observation (t=1)
  - **Distance** to targets
  - Risk that another agent has already taken the package

Manhattan Distances

Package	Pos	Ag_1	Ag_2
Pack_1	(2,4)	2	3
Pack_2	(4,2)	2	3

- **Decay function for uncertainty**:
  - $D(d) = e^{-(\lambda \cdot d)}$
  - where *d* is distance from A2, and λ is a decay constant
- **Risk function for A1 stealing the package:**
  - $R = e^{-(\lambda \cdot dA1)}$
  - the closer A1 is, the **higher the risk** of pickup
- Let's pick λ = 0.3
  - A2's distance: 3 → $D = e^{-(0.3 \cdot 3)} \approx 0.406$
  - A1's distance: 2 → $P = 1 - e^{-(0.3 \cdot 2)} \approx 1 - 0.549 = 0.451$
  - Final belief (probability P1 is still there):
    - **P(P1 still there)** = D x P = 0.406 x 0.451 ≈ 0.183
    - P(P2 still there) = P(P1 still there) = 0.183
- **A2 has no strong preference** between the two

Now, consider …



Manhattan Distances

Package	Pos	Ag_1	Ag_2
Pack_1	(2,4)	3	3
Pack_2	(4,2)	1	3

- A1 is **closer** to **P2**
- Both agents are equally distant from **P1**

A2's final belief (probability P1 is still there):

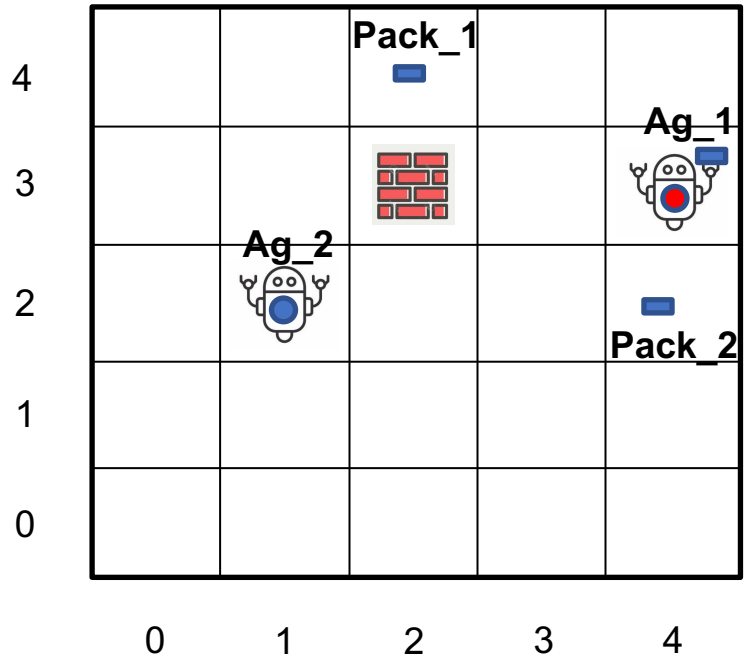
- **P(P2 still there)** = D x (1-R) = 0.406 x (1-0.741) ≈ **0.105**
- **P(P1 still there)** = 0.406 x (1-0.549)= **0.183**

A2 should prioritize going for P1, not P2

… and now

Package	$	Pos	Ag_1	Ag_2
Pack_1	5	(2,4)	3	3
Pack_2	2	(4,2)	1	3

P(P2 still there) = 0.105
P(P1 still there) = 0,183



Expected Value (EV)= P(available) x Reward

EV(P1)=0.183x5=0.915
EV(P2)=0.105x2=0.21

A2 should go for P1, because:
- Even though **it's equally distant**, and **A1 is closer to P2**
- **P1 offers a much better expected return**

But what about A1? Does it not should go for P1 as well?

- Let's define the probability that **A1 goes for a given package** as:

$$P(A1 \rightarrow P_i) = \frac{EV(P_i)}{\sum_j EV(P_j)}$$

EV(P1) = 5 ; EV(P2)=2 ; Total =7

P(A1->P1) = 0.714
P(A1->P2) = 0.286

A2's beliefs $$P(\text{available}) = D(d_{A2}) \cdot (1 - P(A1 \rightarrow P_i) \cdot e^{-\cdot d_{A1}})$$

P(P1) = 0.406 x(1-0.714x 0.406) = 0.288
P(P2) = … = 0.32

→

EV(P1)=0.241x5=1.445
EV(P2)=0.105x2=0.64

Even *after* considering that **A1 is likely to prioritize P1**, the **expected value of going for P1 is still higher** for A2. Now A2's **more informed, realistic belief update** that considers **A1's rationality**.

Beliefs and Introspective abilities

- What about beliefs concerning my intentions, desires, plans, actions? (Introspection)
  - If “I intend G, do I believe I intend to achieve G?”
  - Necessary to reasoning about intentions
    - If there is the opportunity to `pick_up(Pack_2)`
    - What about my current intensions? `B:{Intend(pick_up(Pack_1))}`
    - Are `Intend(pick_up(Pack_1))` and `Intend(pick_up(Pack_2))` consistent?
  - Not easy to implement
  - Synchronization between Beliefs – Intentions – Plans
  - Easy to get into self-contradictory reasoning

Beliefs about other agents' mental states

- Beliefs about other agents' beliefs
  - I believe you believe - `B:{believe(Ag_2,In(Pack_1,2,4))}`
  - Important for coordination, negotiation, and competition
  - We will se more on agents' communication and the speech act theory
- Beliefs about other agents' Intentions
  - I believe you Intend - `B:{intend(Ag_2,pick_up(Pack_1))}`
  - Reasoning about the others' behaviours (coordination, negotiation, and competition)
  - Prediction of others' actions – need to explore possible plans
- Beliefs about other agents' plans
  - I believe you have the plan of `B:{plan(Ag_2,[move(UP),move(RIGHT),pick_up(Pack_1)]}`
  - Usually related to intentions but not always intensions are known
    - I know it will follow that path, but I don't know why
  - Prediction of others' actions

For the project

- We do not use logical deduction

  B=$\{\alpha, \alpha \vdash \beta\}$ means also that $\beta \in B$

- We only use facts (true/false) and environmental contraints (integrety constraint like a DB)

```
fact:       t:   In(Ag_1,1,2)
constraint: t:   In(Ag_1,1,2)
            t+1: In(Ag_1,1,2) v In(Ag_1,1,1) v In(Ag_1,2,2)
```