Logical Particle Filtering

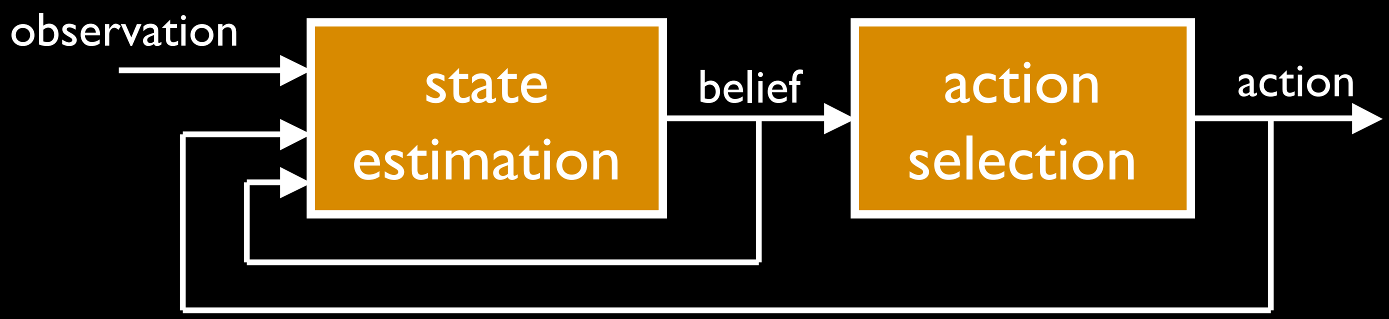
Luke Zettlemoyer

with Hanna Pasula and Leslie Pack Kaelbling

Computer Science and Artificial Intelligence Laboratory
Massachusetts Institute of Technology

CSAIL

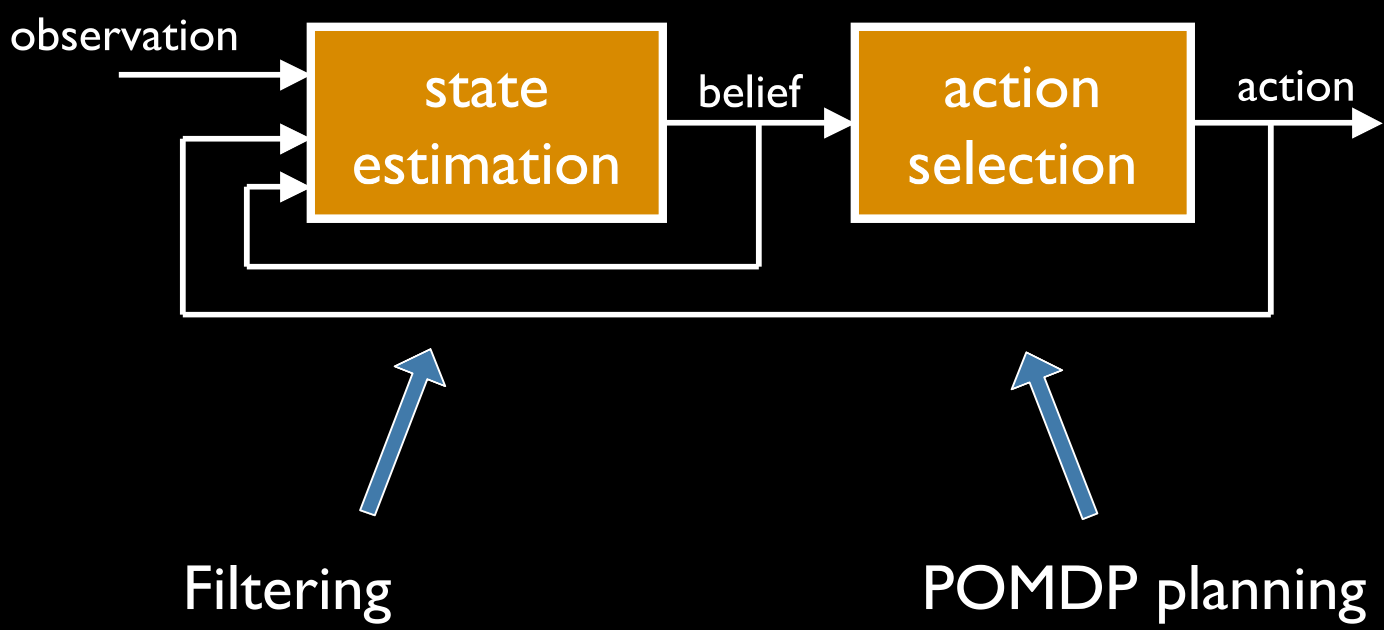
World-focused Inference Problems



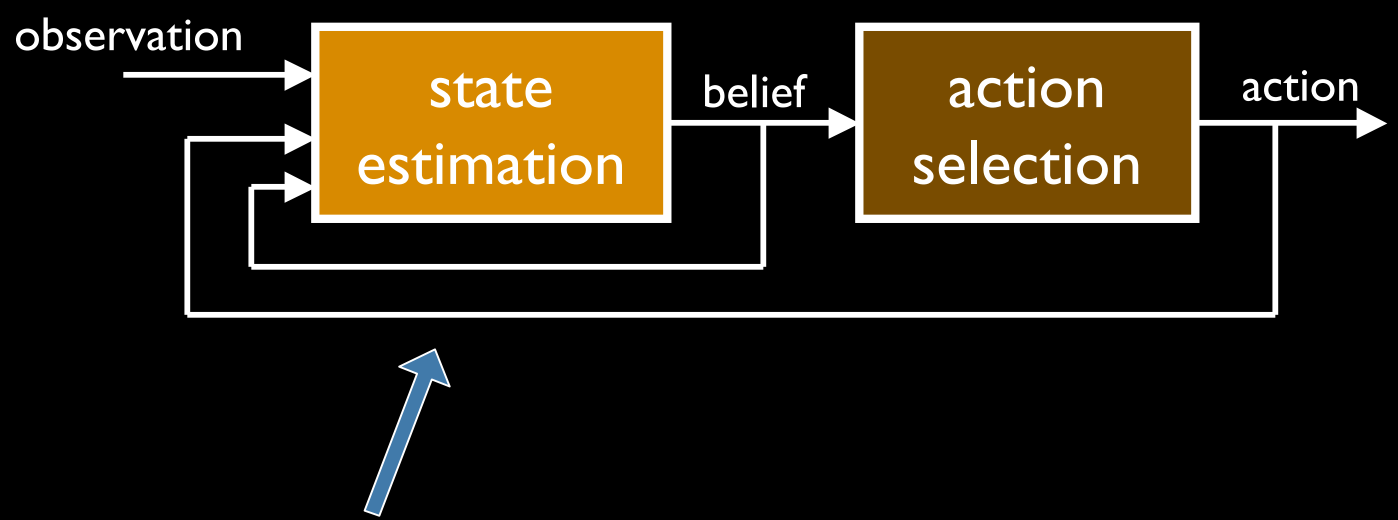
State estimation: What can I conclude about the current state s_t of the world given my history of actions $a_1 \ldots a_{t-1}$ and observations $o_1 \ldots o_{t-1}$?

Action selection: What action a_t should I take given my current beliefs about the state of the world?

World-focused Inference Problems



World-focused Inference Problems



Filtering

Representation

We need:

- model of world dynamics: $p(s_t|s_{t-1},a_t)$
  "What will happen if I do this action?"

- model of observations: $p(o_t|s_t)$

  "What is the world like when I get this observation?"

Eventually we want to learn these distributions, for now we assume they are given.

State estimation

Problem: given history of past observations and actions, what is the current state of the world?

Lazy: store everything, do inference when necessary

Eager: maintain an explicit representation of the current distribution over the state of the world ("filtering")

Exact logical filter

Current belief is a logical formula representing all possible worlds

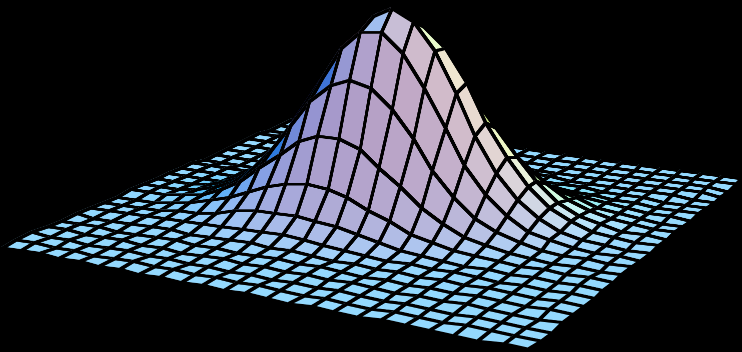
- On each step:
  - "progress" the old belief through logical description of world dynamics
  - conjoin the observation
  - simplify formula
- Issues
  - formulas become very complicated
  - every possible state represented with equal weight

Amir & Russell 03,
Shirazi & Amir 05

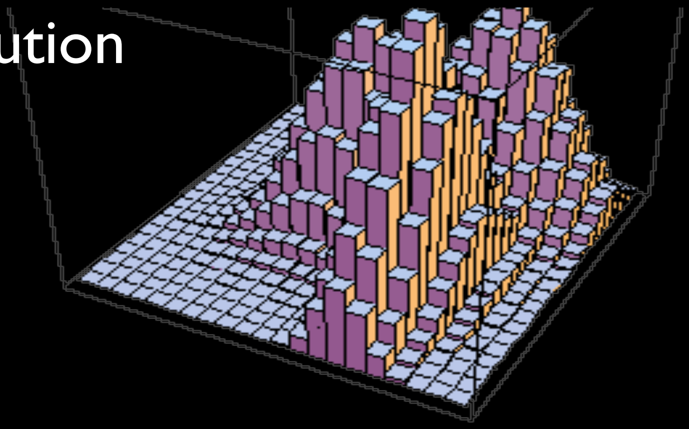
Representing the belief state

In the probabilistic case, we need: $p(s_t|o_1 \ldots o_t)$

Kalman filter: Gaussian

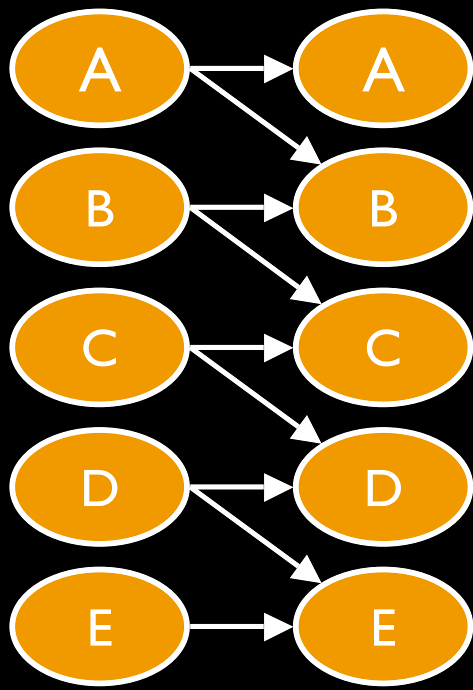


Histogram filter: discrete distribution



Bayesian network filter

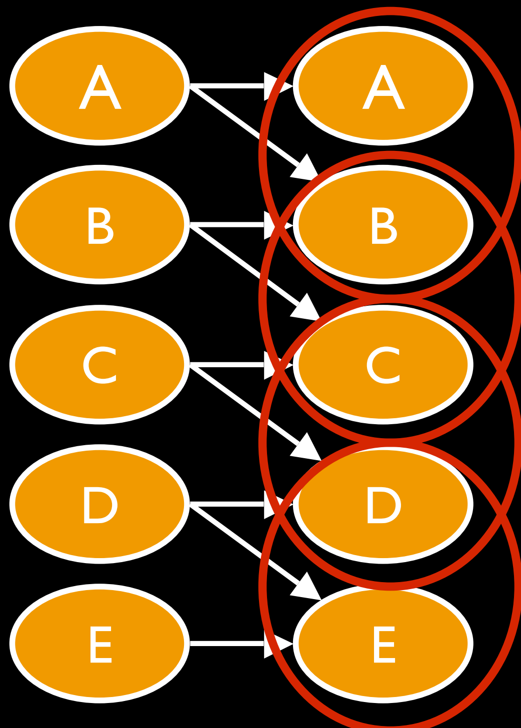
Even if dynamics are factored, joint distribution usually isn't



Boyen & Koller, 98

Bayesian network filter

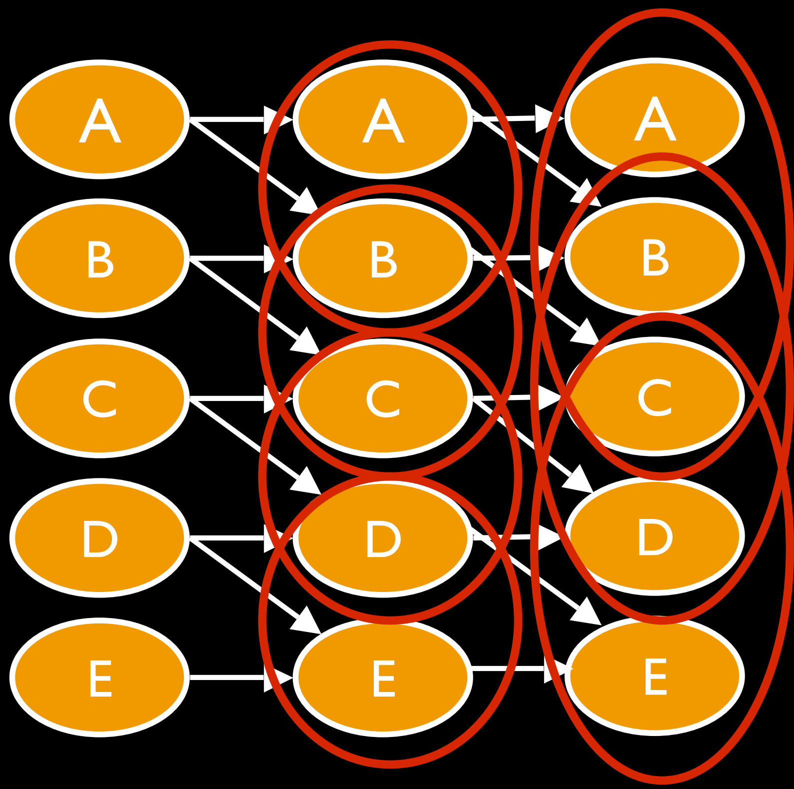
Even if dynamics are factored, joint distribution usually isn't



Boyen & Koller, 98

Bayesian network filter

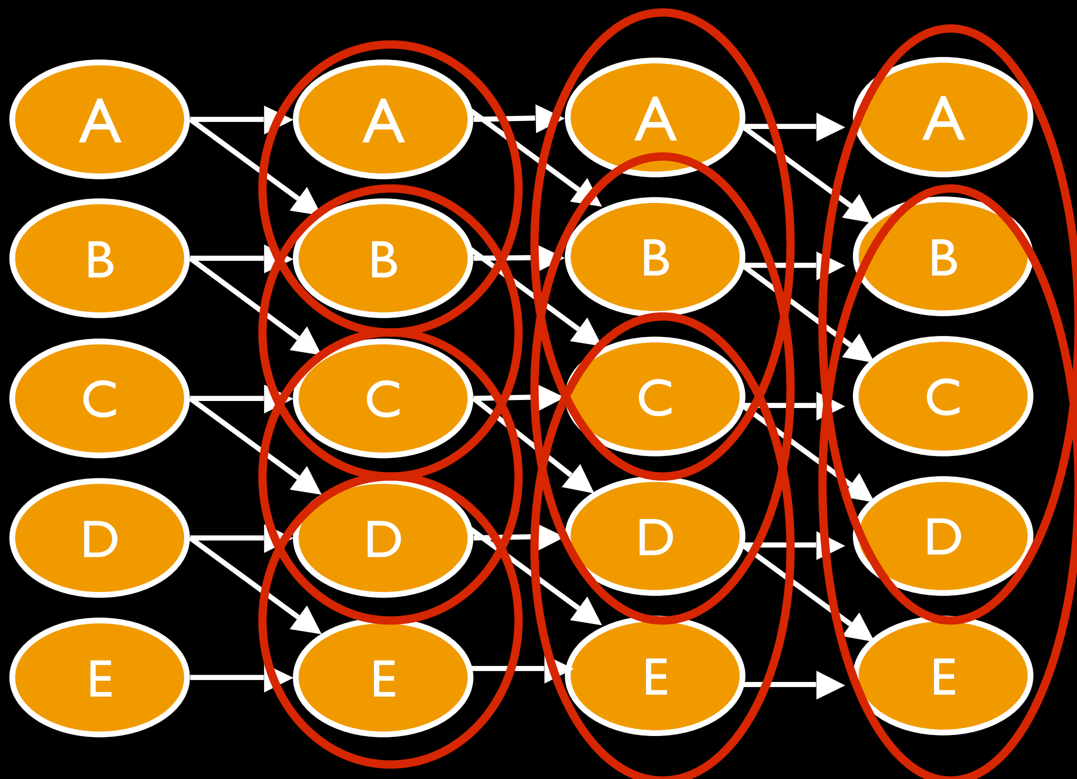
Even if dynamics are factored, joint distribution usually isn't



Boyen & Koller, 98

Bayesian network filter

Even if dynamics are factored, joint distribution usually isn't

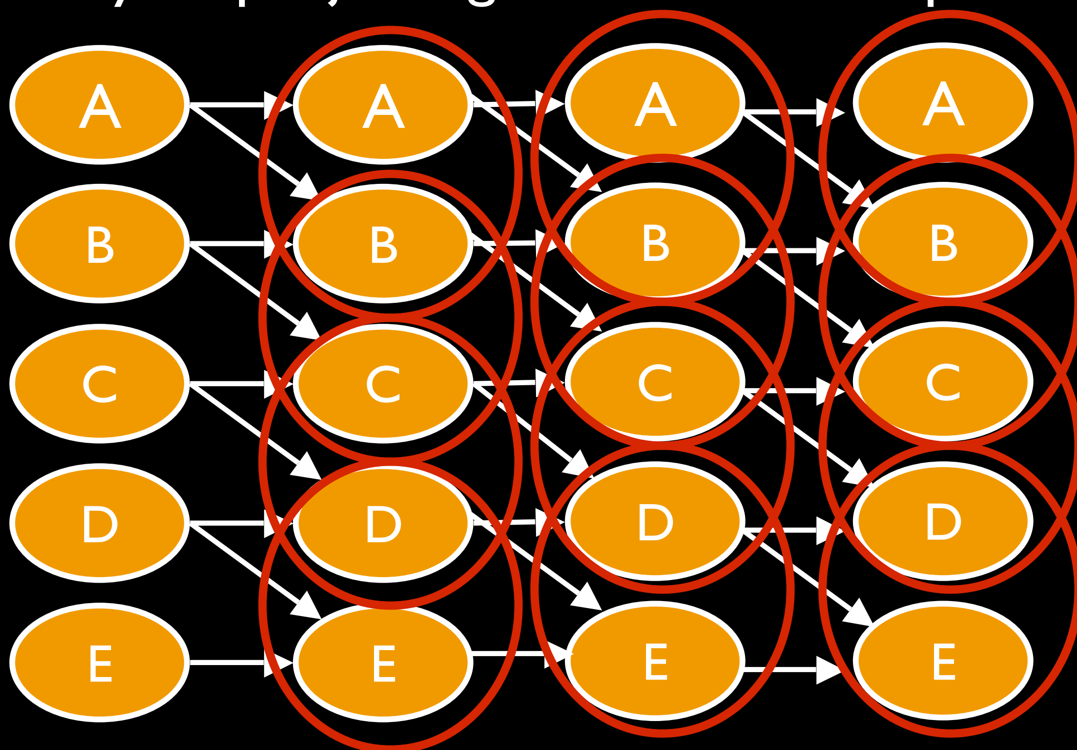


Boyen & Koller, 98

Bayesian network filter

Even if dynamics are factored, joint distribution usually isn't

Approximate by re-projecting into factored representation



Boyen & Koller, 98

Sampling simplifies

Whole probability distributions are

- hard to represent
- usually unnecessary

Spend your representational effort where it matters most!

Particle filter

- explicitly represent a set of states, drawn from the current belief distribution
- generates new set of samples efficiently from previous iteration

Problems with particles

Particles must represent every state with non-negligible probability

- Need *lots* of particles!

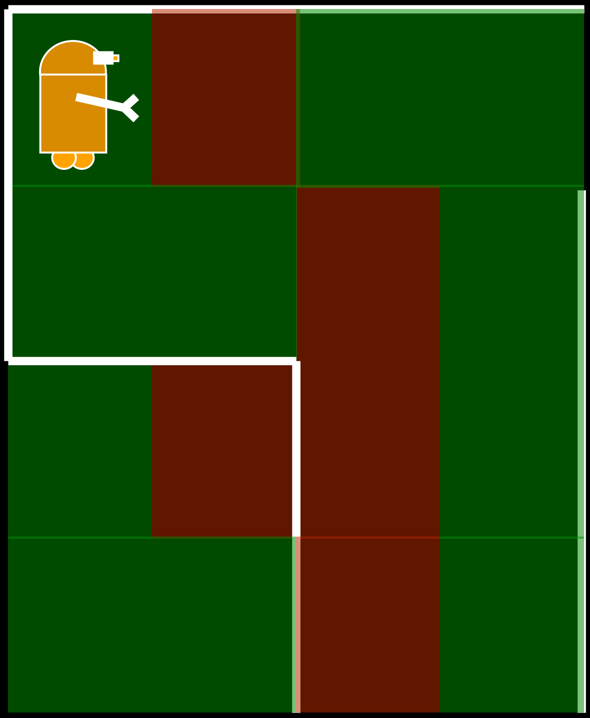
Each particle needs to specify a fully-ramified world state

- Need *big* (infinite) particles!

We want to be able to work in *infinite* worlds, with very little information.

An illustrative example

- R finds himself in an unknown world
- dimensions are unknown (possibly infinite)
- there are some walls between locations
- locations have appearance
- R moves (with error) through the world
- R observes (with error) the color at his location

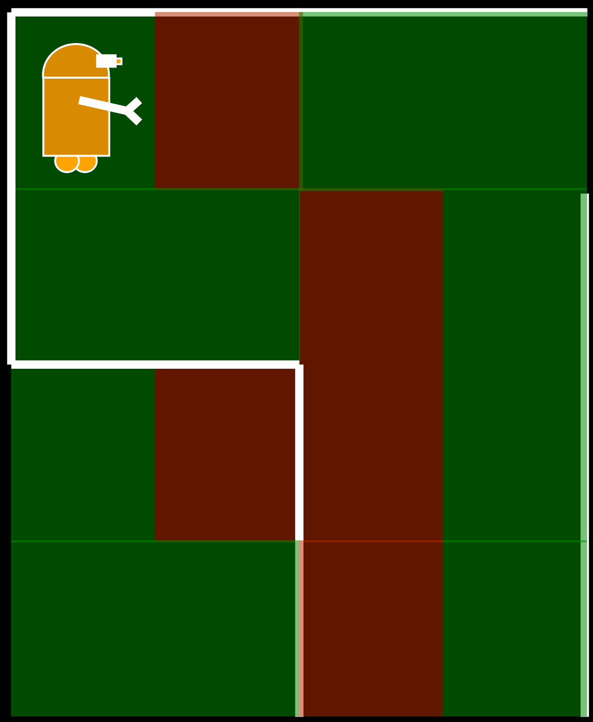


A day in the life

R is booted up,

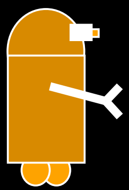
- sees a green square
- tries to move right
- sees a red square

What does R know about the world?



Exact logical filtering

R knows that either:



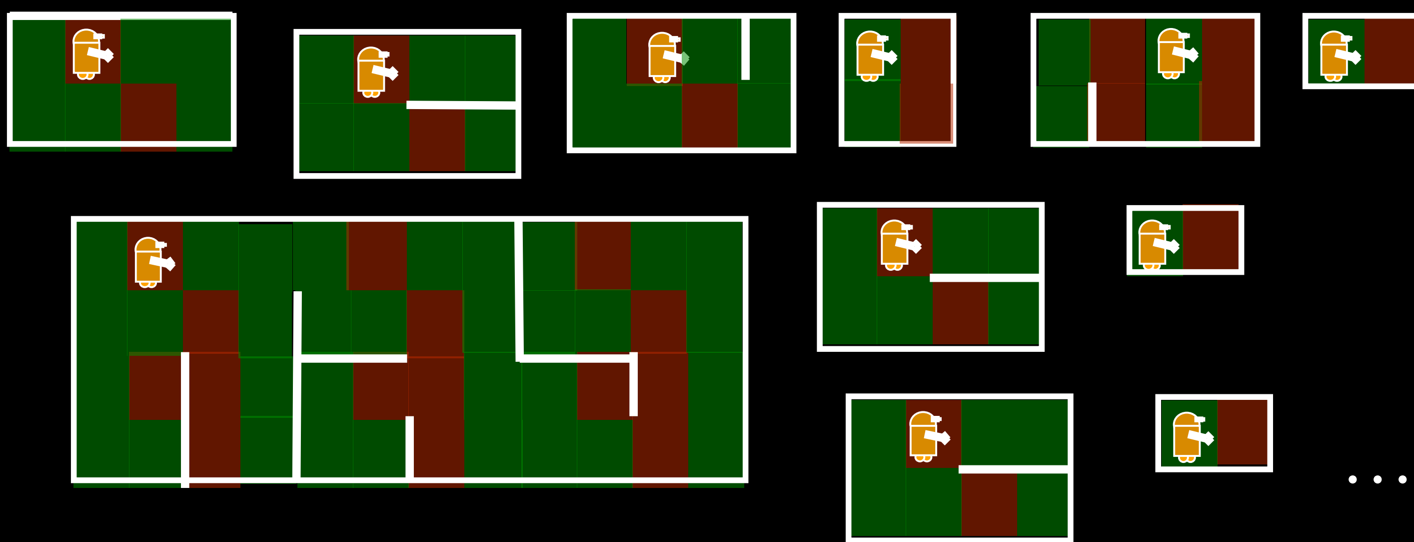
- there is one location that can look either red or green
- there is one location that can look red and another that can look green, and each has some number of walls around it

compact set representation

considers every possible state

Particle filtering

High probability worlds:



states are individually specified

only most likely represented

First-Order Particle Filtering

Particle filter in which

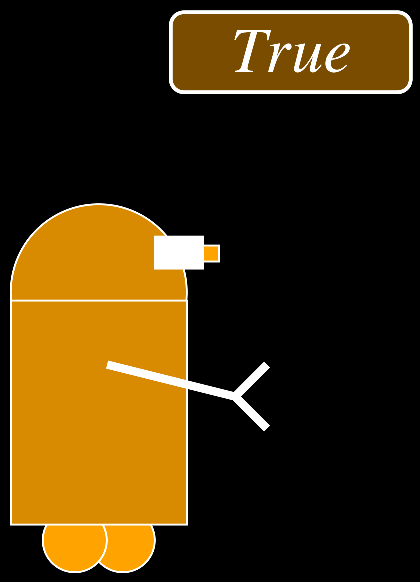
- each “particle” is a set of states compactly described by a logical formula
- there is a implicit distribution over the particles
- all states in the same particle depend on observation and transition history in the same way
  - they have the same $p(s_{1:t}, o_{0:t} | s_0)$

compact set representation

only most likely represented

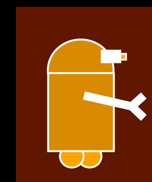
R wakes up

- One set of all possible worlds

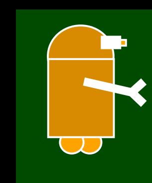


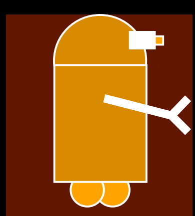
R sees a red square

$\exists x.at(R,x) \wedge red(x)$ 0.8



$\exists x.at(R,x) \wedge green(x)$ 0.2

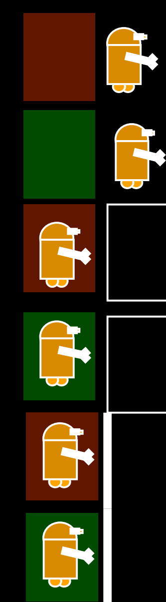


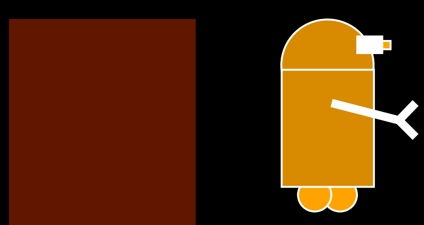


Prob of red given see red: 0.8

R tries to move right

$\exists x, y. at(R,y) \wedge red(x) \wedge rightOf(y,x)$	.504
$\exists x, y. at(R,y) \wedge green(x) \wedge rightOf(y,x)$	.126
$\exists x, y. at(R,x) \wedge red(x) \wedge rightOf(y,x)$	.056
$\exists x, y. at(R,x) \wedge green(x) \wedge rightOf(y,x)$	.014
$\exists x. at(R,x) \wedge red(x) \wedge \neg\exists y. rightOf(y,x)$	0.24
$\exists x. at(R,x) \wedge green(x) \wedge \neg\exists y. rightOf(y,x)$	0.06

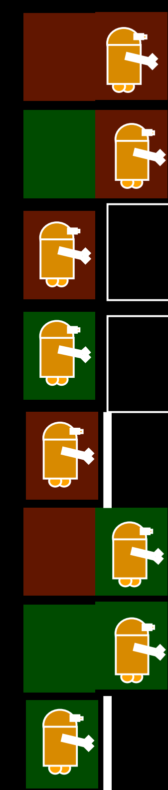


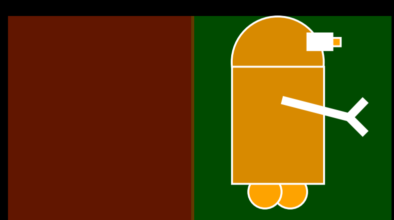


Prob of wall to right: 0.3
Prob of failing to move if there's a wall: 0.1

R sees a green square

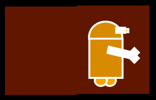
$\exists x, y. at(R,y) \wedge red(x) \wedge rightOf(y,x) \wedge red(y)$	.135
$\exists x, y. at(R,y) \wedge green(x) \wedge rightOf(y,x) \wedge red(y)$	.033
$\exists x, y. at(R,x) \wedge red(x) \wedge rightOf(y,x)$	.015
$\exists x, y. at(R,x) \wedge green(x) \wedge rightOf(y,x)$	.015
$\exists x. at(R,x) \wedge red(x) \wedge \neg\exists y. rightOf(y,x)$	.064
$\exists x, y. at(R,y) \wedge red(x) \wedge rightOf(y,x) \wedge green(y)$	.539
$\exists x, y. at(R,y) \wedge green(x) \wedge rightOf(y,x) \wedge green(y)$	.135
$\exists x. at(R,x) \wedge green(x) \wedge \neg\exists y. rightOf(y,x)$	.064





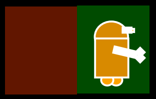
Samples instead of full partition

$\exists x,y.at(R,y) \wedge red(x) \wedge rightOf(y,x) \wedge red(y)$

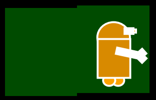


$\exists x.at(R,x) \wedge red(x) \wedge \neg\exists y.rightOf(y,x)$

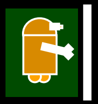
$\exists x,y.at(R,y) \wedge red(x) \wedge rightOf(y,x) \wedge green(y)$

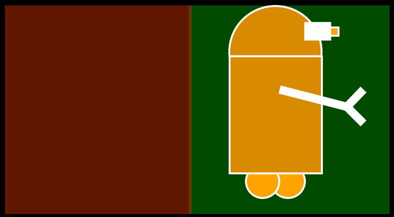


$\exists x,y.at(R,y) \wedge green(x) \wedge rightOf(y,x) \wedge green(y)$



$\exists x.at(R,x) \wedge green(x) \wedge \neg\exists y.rightOf(y,x)$





Logical Particle Filter

Initialization: For $i = 1 \ldots n$, set $h_o^{(i)} = true$

Filtering:

1. For $i = 1 \ldots n$, draw $h_t^{(i)} \sim q(h_t \mid h_{t\text{-}1}, o_t)$

2. Calculate important weights:
$$w(h_t^{(i)}) = p(o_t \mid h_t^{(i)})\, p(h_t^{(i)} \mid h_{t\text{-}1}^{(i)}) \,/\, q(h_t^{(i)} \mid h_{t\text{-}1}^{(i)}, o_t)$$

3. Normalize importance weights:
$$w'(h_t^{(i)}) = w(h_t^{(i)}) \,/ \sum_j w(h_t^{(j)})$$

4. Resample according to the $w'(h_t^{(i)})$

A particle filter where the $h_t^{(i)}$ are defined by the *logical* world dynamics.

Logical World Dynamics

Transition:

go-left:

$\exists!\, x \,.\, at(x)$

$\neg\exists\, y \,.\, left(x,y)$

$\rightarrow \{\ 1.0 : \text{no change}$

go-left:

$\exists!\, x \,.\, at(x)$

$\exists!\, y \,.\, left(x,y)$

$$\rightarrow \begin{cases} 0.9 : at(y), \neg at(x) \\ 0.1 : \text{no change} \end{cases}$$

...

Observation:

observe:

$\exists!\, x \,.\, at(x) \wedge red(x)$

$$\rightarrow \begin{cases} 0.9 : red \\ 0.1 : \text{green} \end{cases}$$

...

Technical Story

Rao-Blackwellization:

$$E_{P(x_1,x_2)} f(x_1,x_2) = E_{P(x_2)} E_{P(x_1|x_2)} f(x_1,x_2)$$
$$\approx \frac{1}{n} \sum_{\text{samples from} P(x_2)} E_{P(x_1|x_2)} f(x_1,x_2)$$

For us:

$$x_2 : \text{logical partition}$$
$$x_1 : \text{state within the partition}$$
$$f(x_1,x_2) : \text{Am I in room 6?}$$

created dynamically depending on observations

depends only on prior

Many other possible f

Discussion

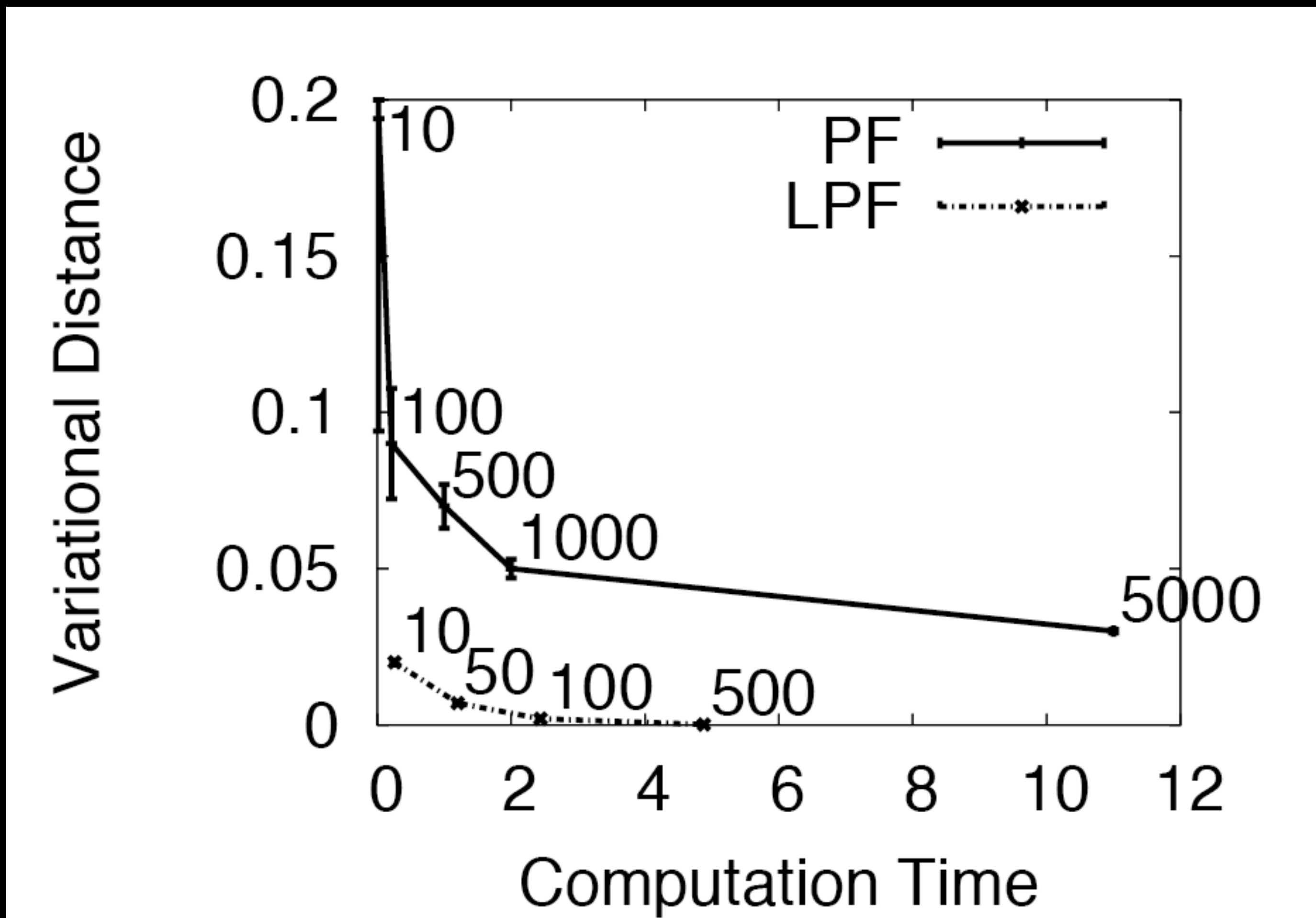
Guarantees:

- Computes correct estimate as number of particles grows to infinity
- Lower variance estimates
- Higher computational cost

Challenges:

- Logical formulae grow in size
- How many particles?

Preliminary experiments: approx. 5000 possible initial states



Related Work

- MCMC sampling over sets of states

  Milch & Russell, 06

- Particle filters within factored distributions

  Ng, Peshkin, & Pfeffer, 02

- Rao-Blackwellization in relational models

  Sanghai, Domingos & Weld, 05

- Relational hidden Markov models

  Kersting, De Raedt & Raiko, 06

Future Work

- More / better application domains
- More general logical formula
- Merging particles
- Relational POMDPS
- …

The End

Thanks for your time!