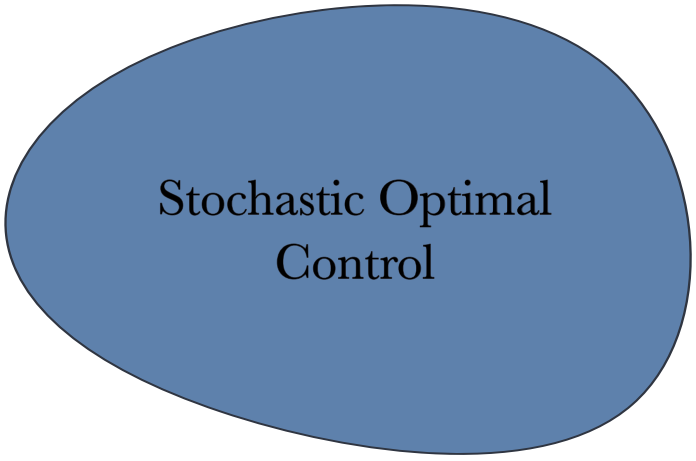
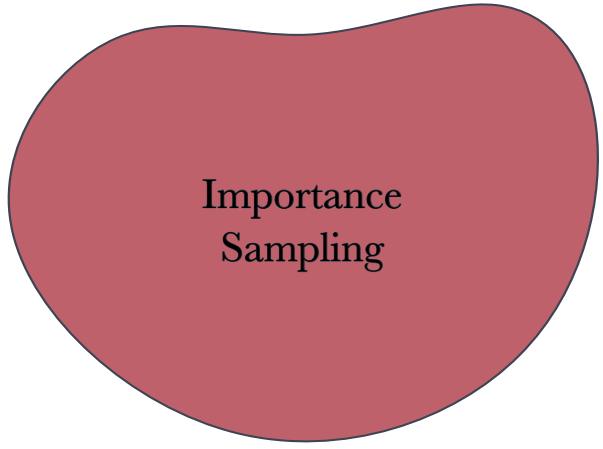
Connecting SOC with RL – Importance sampling

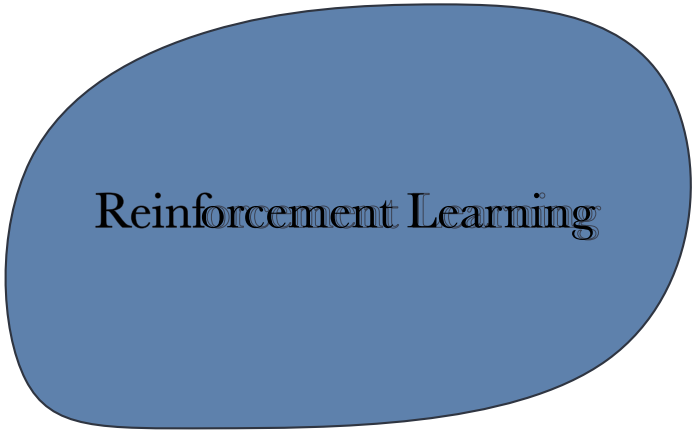
AI as a tool in Mathematics

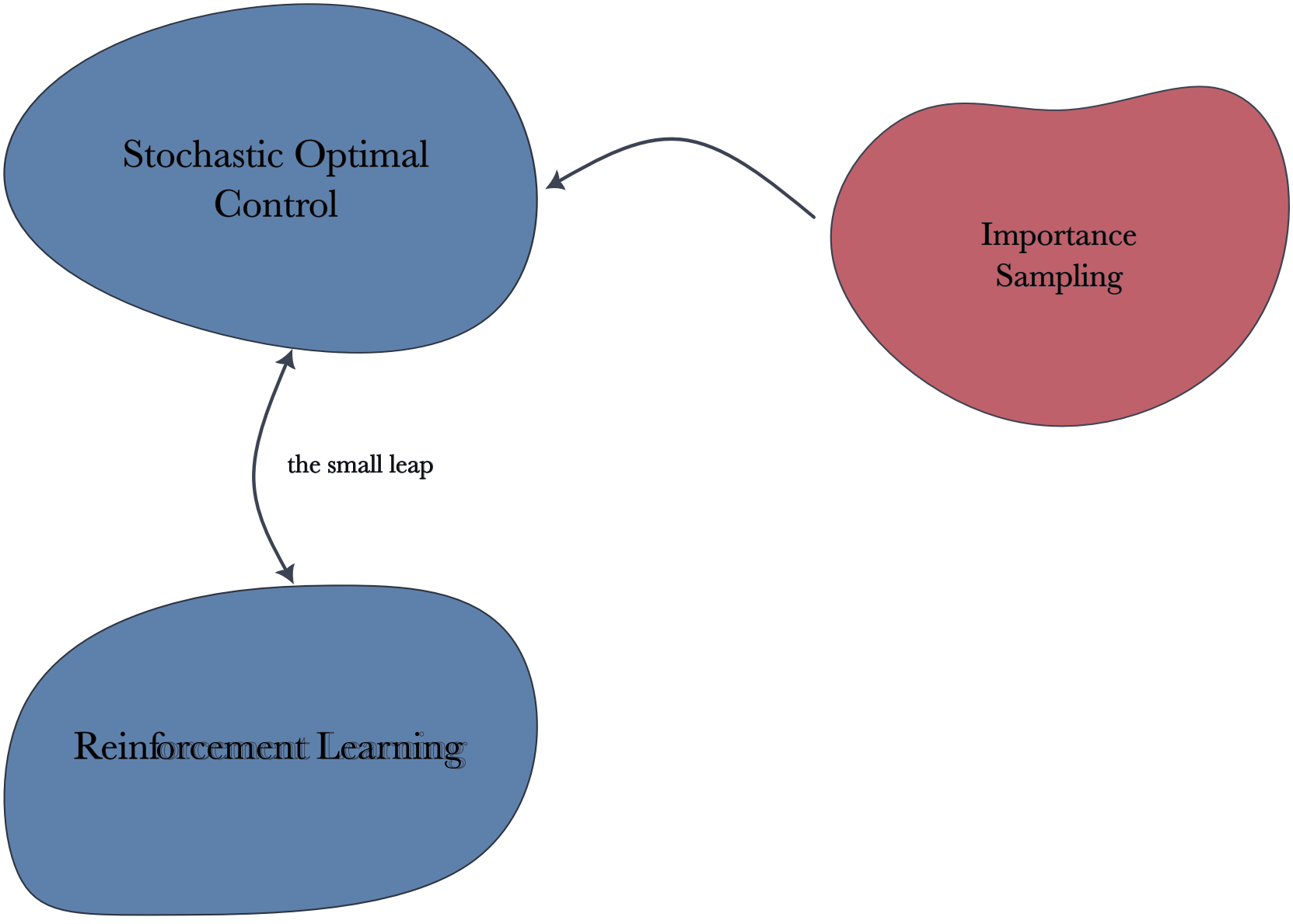
Alonso Cisneros

Freie Universität Berlin


Stochastic Optimal
Control


Importance
Sampling


Reinforcement Learning


Stochastic Optimal Control
Importance Sampling
the small leap
Reinforcement Learning

Outline

1. *Crash* course on RL
2. What is importance sampling
   - The connection to optimization
   - Optimal biasing
4. Optimal biasing as an RL problem

Crash course on Reinforcement Learning

Crash course on Reinforcement Learning

Figure 1: A miniopoly board

- The game has a state at turn t denoted s_t
- At a turn t players roll the dice
- The change in money after buying/paying rent/charging rent is recorded as a reward r_t

- The game has a state at turn t denoted s_t
- At a turn t players roll the dice
- The change in money after buying/paying rent/charging rent is recorded as a reward r_t

! Important

We train our robot to maximize the rewards as it takes actions exploring the space of states

Can you describe the state space?

- The game has a state at turn t denoted s_t
- At a turn t players roll the dice
- The change in money after buying/paying rent/charging rent is recorded as a reward r_t

> **Important**
>
> We train our robot to maximize the rewards as it takes actions exploring the space of states

Can you describe the state space?

How does it *look* like?

It's hard to describe the state space, but we can study the dynamics

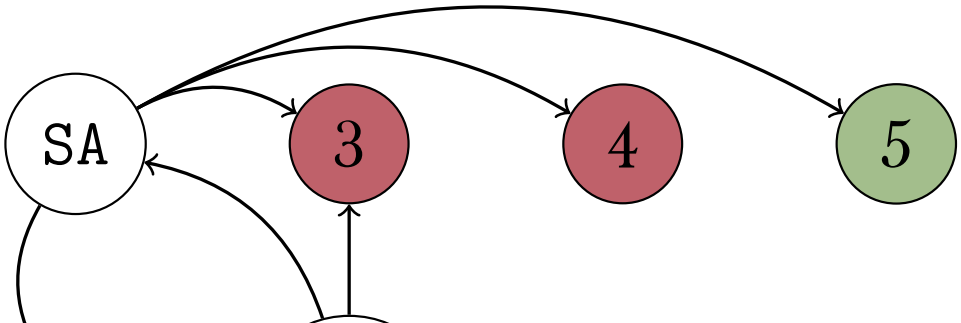
- The game has a state at turn t denoted s_t
- At a turn t players roll the dice
- The change in money after buying/paying rent/charging rent is recorded as a reward r_t

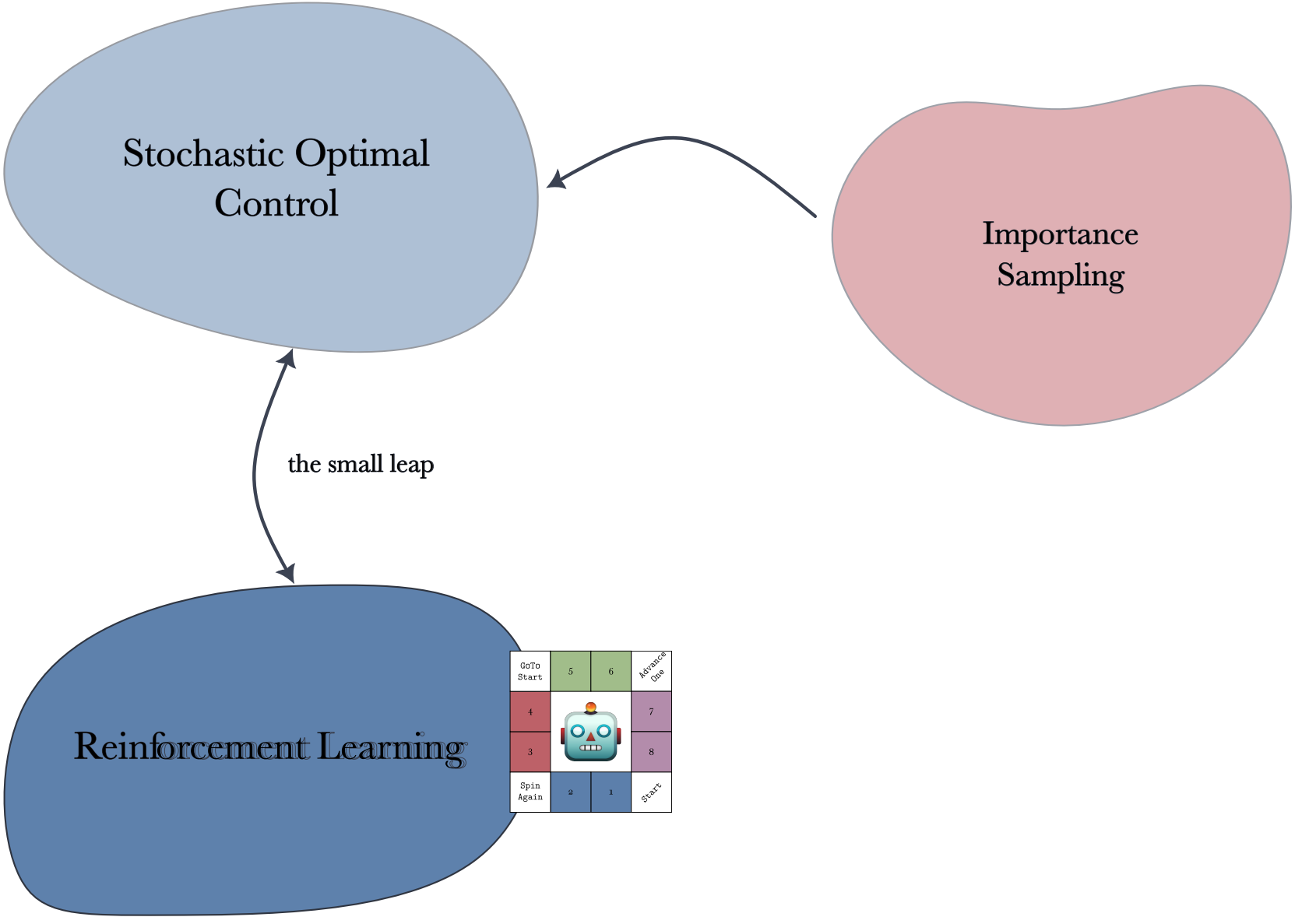
> **!** Important
>
> We train our robot to maximize the rewards as it takes actions exploring the space of states

Can you describe the state space?

How does it *look* like?

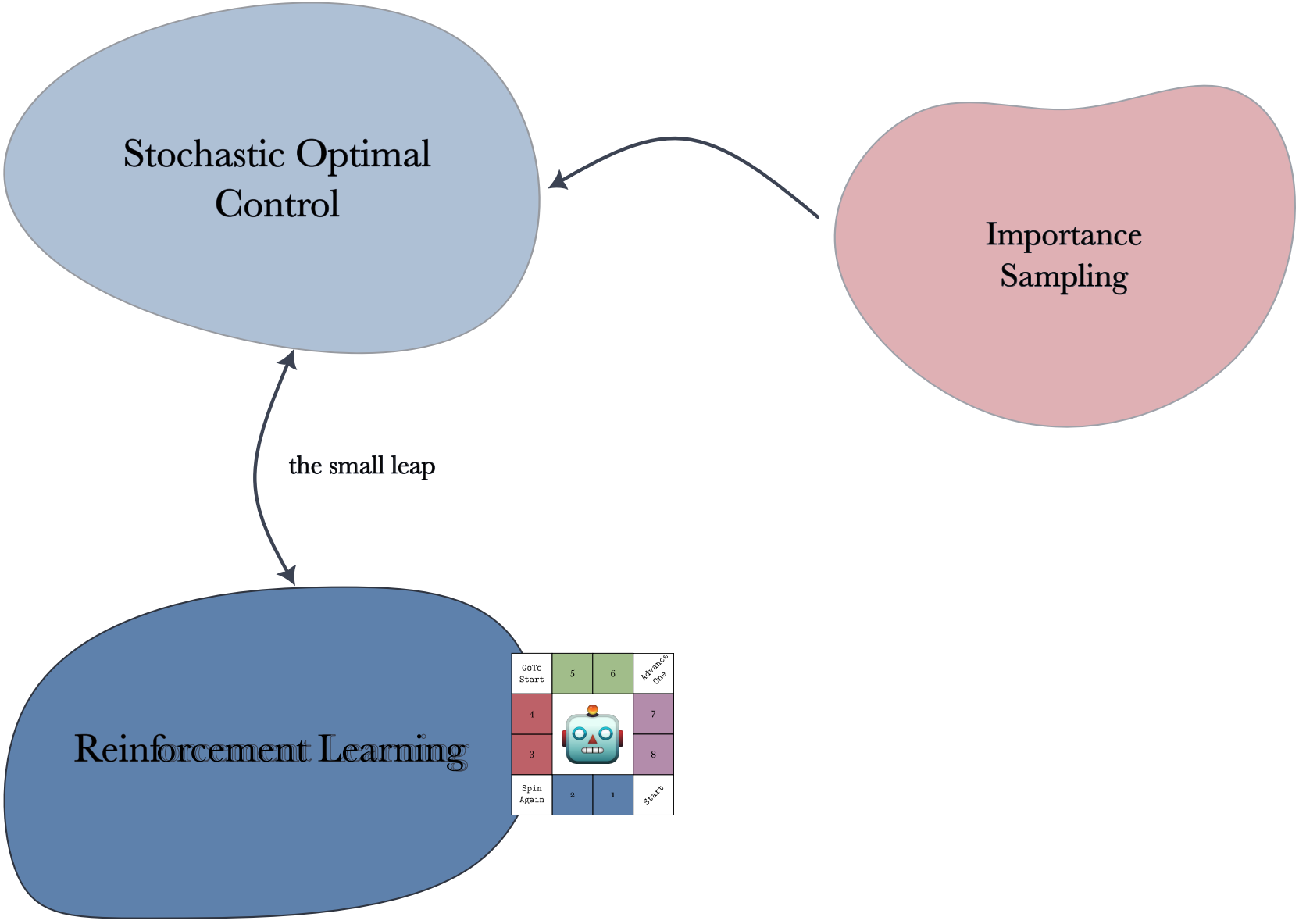
It's hard to describe the state space, but we can study the dynamics





What if we don't know how square transitions work?

- We calculated transition probability *with* the knowledge of the dice



What if we don't know how square transitions work?

- We calculated transition probability *with* the knowledge of the dice

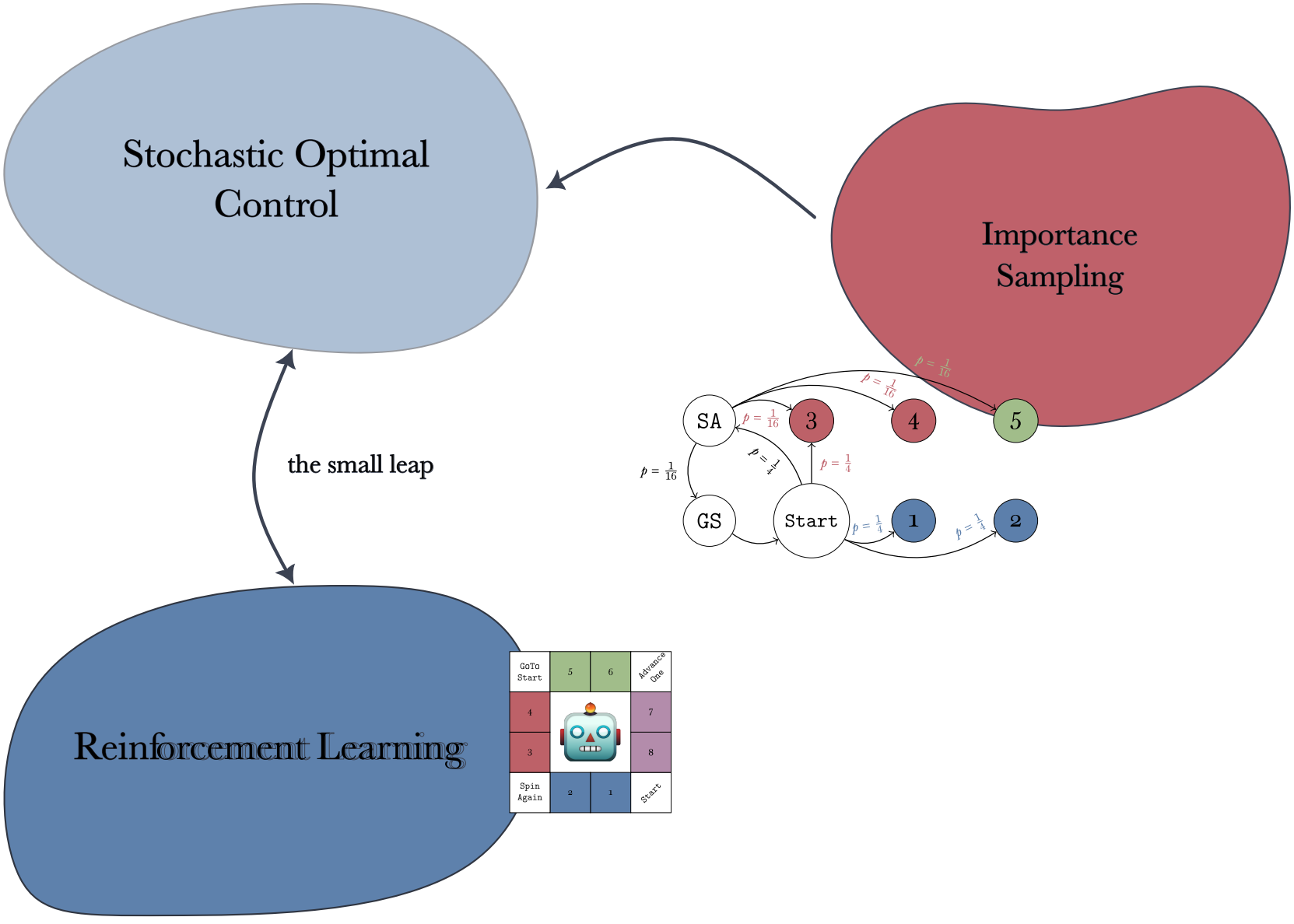


Figure 2: What we've covered so far

Importance Sampling

- We wanted to compute the expected reward of the robot after the entire game
- Not every problem is this well behaved
- This property is called *metastability*
- Importance sampling aims to remedy this

! Important

The general idea of importance sampling is to draw random variables from another probability measure and subsequently weight them back in order to still have an unbiased estimator of the desired quantity of interest