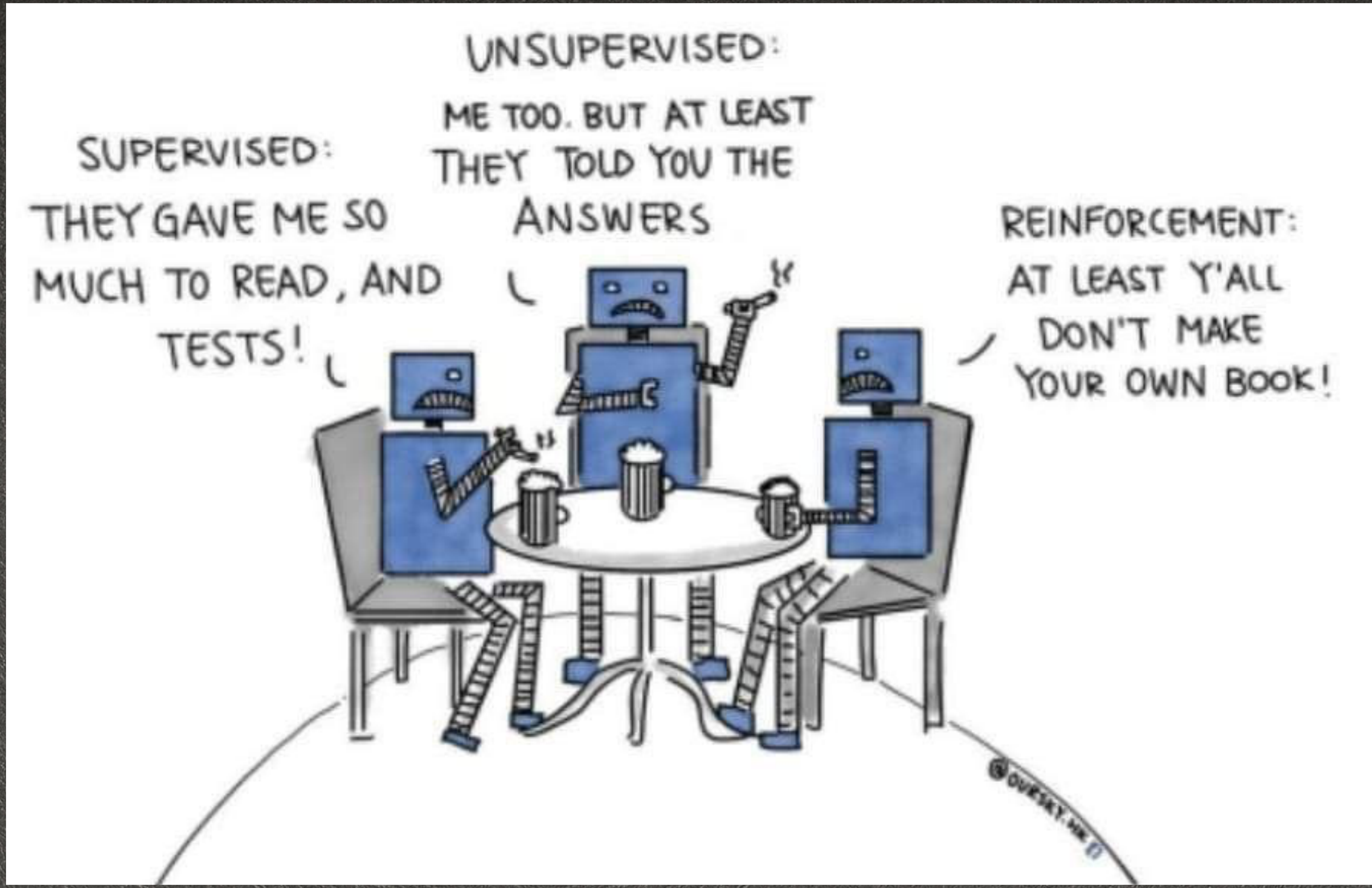
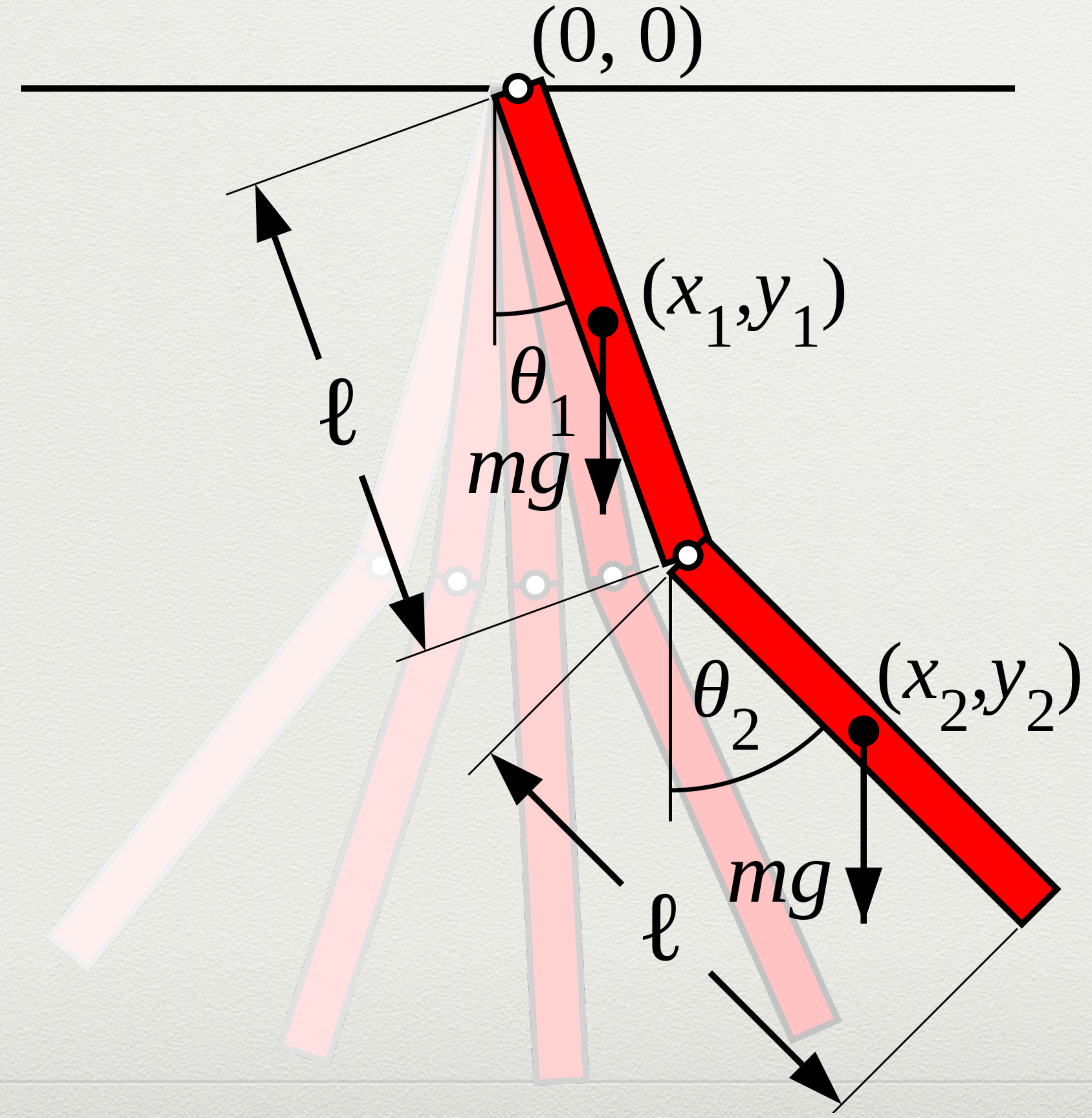


APRENDIZAJE REFORZADO

Física Computacional II

Ph.D. Santiago Eheverri Arteaga



$$L = \frac{ml^2}{6}\left(4\dot{\theta}_1^2 + \dot{\theta}_2^2 + 3\dot{\theta}_1\dot{\theta}_2\cos\left(\theta_1 - \theta_2\right)\right) + \frac{mgl}{2}\left(3\cos\theta_1 + \cos\theta_2\right)$$

RESEARCH ARTICLE

Top-down design of protein architectures with reinforcement learning

BY ISAAC D. LUTZ, SHUNZHI WANG, CHRISTOFFER NORN, ALEXIS COURBET, ANDREW J. BORST, YAN TING ZHAO, ANNIE DOSEY, LONGXING CAO, JINWEI XU, ELIZABETH M. LEAF, [...] DAVID BAKER +9 authors

• SCIENCE • VOL. 380, NO. 6642 • 14 APR 2023 : 266-273

As a result of evolutionary selection, the subunits of naturally occurring protein assemblies often fit together with substantial shape complementarity to generate architectures optimal for function in a manner not achievable by current design approaches. ...

RESEARCH ARTICLE

Controlling chaotic itinerancy in laser dynamics for reinforcement learning

BY RYUGO IWAMI, TAKATOMO MIHANA, KAZUTAKA KANNO, SATOSHI SUNADA, MAKOTO NARUSE, ATSUSHI UCHIDA • SCIENCE ADVANCES • VOL. 8, NO. 49 • 07 DEC 2022

Photonic artificial intelligence has attracted considerable interest in accelerating machine learning; however, the unique optical properties have not been fully used for achieving higher-order functionalities. Chaotic itinerancy, with its spontaneous ...

RESEARCH ARTICLE

Mastering the game of Stratego with model-free multiagent reinforcement learning

BY JULIEN PEROLAT, BART DE VYLDER, DANIEL HENNES, EUGENE TARASSOV, FLORIAN STRUB, VINCENT DE BOER, PAUL MULLER, JEROME T. CONNOR, NEIL BURCH, THOMAS ANTHONY, [...] KARL TUYLS

+23 authors • SCIENCE • VOL. 378, NO. 6623 • 01 DEC 2022 : 990-996

We introduce DeepNash, an autonomous agent that plays the imperfect information game Stratego at a human expert level. Stratego is one of the few iconic board games that artificial intelligence (AI) has not yet mastered. It is a game characterized by a ...

RESEARCH ARTICLE

Reinforcement learning with artificial microswimmers

BY S. MUIÑOS-LANDIN, A. FISCHER, V. HOLUBEC, F. CICHOS • SCIENCE ROBOTICS • VOL. 6, NO. 52 • 24 MAR 2021

Artificial microswimmers that can replicate the complex behavior of active matter are often designed to mimic the self-propulsion of microscopic living organisms. However, compared with their living counterparts, artificial microswimmers have a limited ...

Results / **1-20 of 658**

You searched for **reinforcement+learning** ✖

Sort

Most Relevant

Results Per Page

10

PhySH Concept

- ALL (658)
- Quantum computation (321)
- Numerical techniques (276)
- Particle data analysis (251)
- Machine learning (250)
- Quantum correlations, foundations & formalism (123)
- Quantum control (85)
- Quantum algorithms (82)

Show More

PhySH Discipline

PRE Editors' Suggestion — **1 citation** — PDF HTML

Gradient dynamics in reinforcement learning

Riccardo Fabbricatore and Vladimir V. Palyulin

Phys. Rev. E **106**, 025315 (2022) - Published 26 August 2022

Show Abstract +

PRE Editors' Suggestion — **19 citations** — PDF HTML

Deterministic limit of temporal difference reinforcement learning for stochastic games

Wolfram Barfuss, Jonathan F. Donges, and Jürgen Kurths

Phys. Rev. E **99**, 043305 (2019) - Published 10 April 2019

Show Abstract +

PRRESEARCH — **25 citations** — PDF HTML

Reinforcement-learning-assisted quantum optimization

Matteo M. Wauters, Emanuele Panizon, Glen B. Mbeng, and Giuseppe E. Santoro

Phys. Rev. Research **2**, 033446 (2020) - Published 18 September 2020

Show Abstract +

PRFLUIDS Editors' Suggestion — **38 citations** — PDF HTML

Controlled gliding and perching through deep-reinforcement-learning

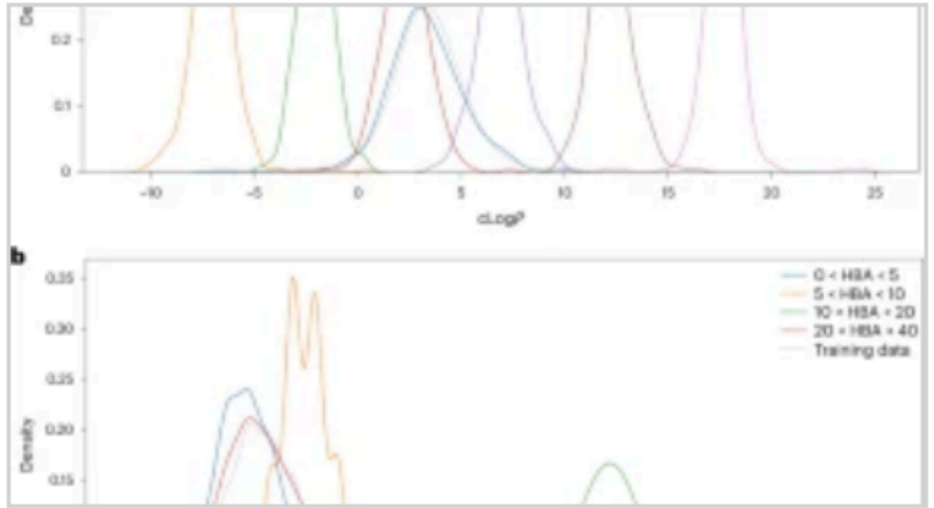
Guido Novati, L. Mahadevan, and Petros Koumoutsakos

Phys. Rev. Fluids **4**, 093902 (2019) - Published 6 September 2019

Research
03 Apr 2023
Nature Machine Intelligence
Volume: 5, P: 386-394

[Testing the limits of SMILES-based de novo molecular generation with curriculum and deep reinforcement learning]

Generative models in cheminformatics depend on molecules being representable as structured data, such as the simplified molecular-input line-entry system (SMILES). Mokaya and colleagues investigated how the choice of representation influences the quality of generated compounds, and found that string-based representations can hinder performance in a curriculum learning setting.

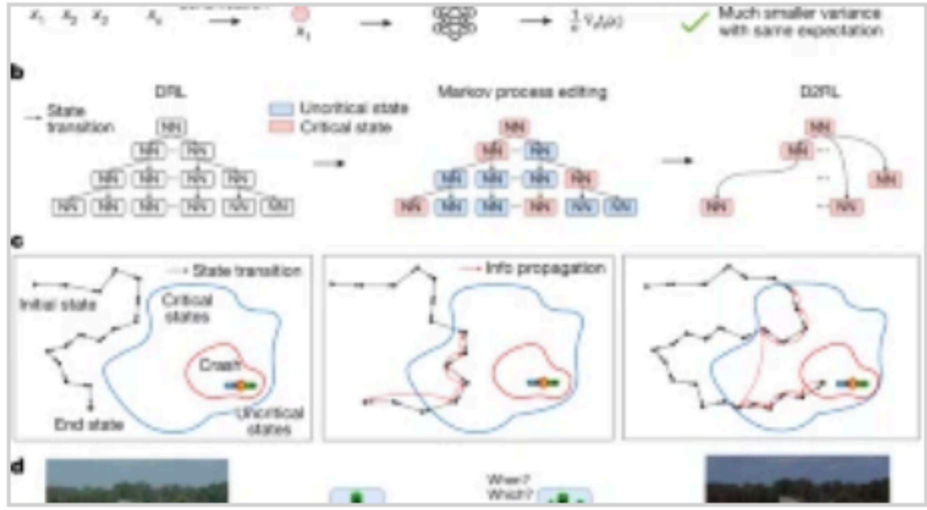


Maranga Mokaya, Fergus Imrie ... Charlotte M. Deane

Research
22 Mar 2023
Nature
Volume: 615, P: 620-627

Dense reinforcement learning for safety validation of autonomous vehicles

An intelligent environment has been developed for testing the safety performance of autonomous vehicles and its effectiveness has been demonstrated for highway and urban test tracks in an augmented-reality environment.

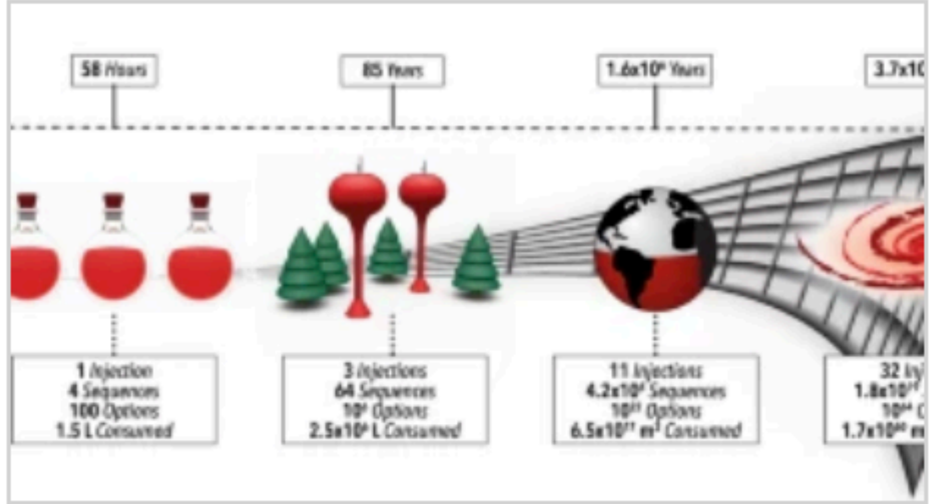


Shuo Feng, Haowei Sun ... Henry X. Liu

Research
Open Access
14 Mar 2023
Nature Communications
Volume: 14, P: 1-16

AlphaFlow: autonomous discovery and optimization of multi-step chemistry using a self-driven fluidic lab guided by reinforcement learning

Autonomous exploration of materials design space is hindered by its high dimensionality and the scarcity of data. In this work, we present AlphaFlow, a self-driven lab guided by reinforcement learning that enables accelerated discovery and optimization of multi-step chemistries.

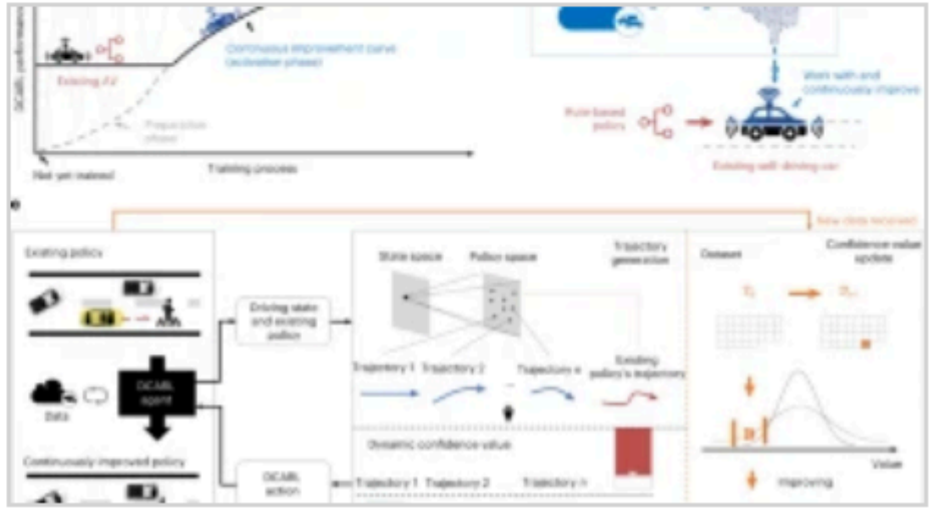


Amanda A. Volk, Robert W. Epps ... Milad Abolhasani

Research
23 Feb 2023
Nature Machine Intelligence
Volume: 5, P: 145-158

Continuous improvement of self-driving cars using dynamic confidence-aware reinforcement learning

Reinforcement learning is a powerful technique to learn complex behaviours, but in the context of self-driving vehicles it might result in unsafe behaviour in previously unseen situations. Cao et al. create a confidence-aware method that improves through reinforcement learning but reverts to safe behaviour when a situation is new.



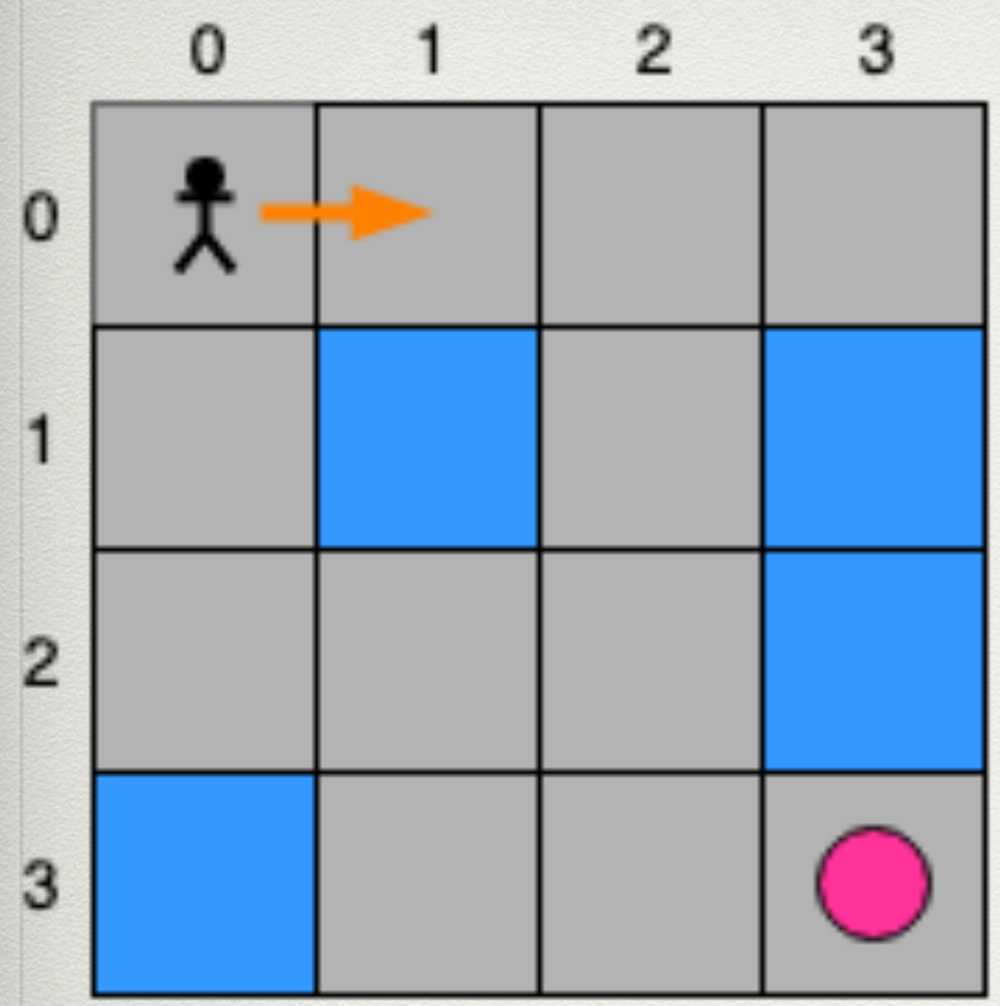
APRENDER AL INTERACTUAR CON EL ENTORNO

Aprendizaje reforzado

- Aprender a mapear situaciones en acciones para maximizar la recompensa.
- Un AGENTE realiza una ACCIÓN sobre un ENTORNO, la cuál puede MODIFICAR SU ESTADO y obtener o no una RECOMPENSA
- No se le dice cuales acciones tomar, pero cada acción puede afectar **no solo la recompensa siguiente sino las futuras.**
- Machine Learning supervisado aprende de ejemplos (los cuales tienen un "supervisor" externo), no de la interacción.
- Dilema Exploración vs Explotación: "Mejor malo conocido que bueno por conocer" pero si no conozco a nadie… 😅
- Ejemplos del día a día:
  - Hacer desayuno en casa ajena
  - 🐄 aprendiendo a caminar al nacer
  - Ajedrez jugado por un maestro (O juegos de mesa por un Friki)
- Necesario un constante monitoreo del entorno

Elementos básicos del aprendizaje reforzado

- **Estado** $s \in S$, **Acción** $a \in A$, **Recompensa** $r = R(s) \in \mathscr{R} \subset \mathbb{R}$
- **Política** $\pi = p(a \mid s) : A \times S \to [0,1]$: Define la forma como se comporta el agente en un tiempo dado. Reglas estímulo-respuesta. Función que transforma un estado del entorno en acciones a ser tomadas.
- **Función de recompensa** $f(a, s, r)$**:** Define el objetivo en el problema de aprendizaje reforazado. Función que toma el estado del entorno y la acción tomada y retorna un número que indica la deseabilidad de ese estado.
- **Función de valor** $V_\pi(s)$**:** Sumatoria de las funciones de recompensa posibles para el agente dado el estado actual. El objetivo es maximizar la función de valor.
    - Los métodos evolutivos como los algoritmos genéticos buscan directamente en el espacio de políticas sin considerar funciones de valor. (No interactúan de la misma forma con el entorno)
- **Modelo del entorno**: Función que dado un estado y una acción, predice el siguiente estado y siguiente recompensa



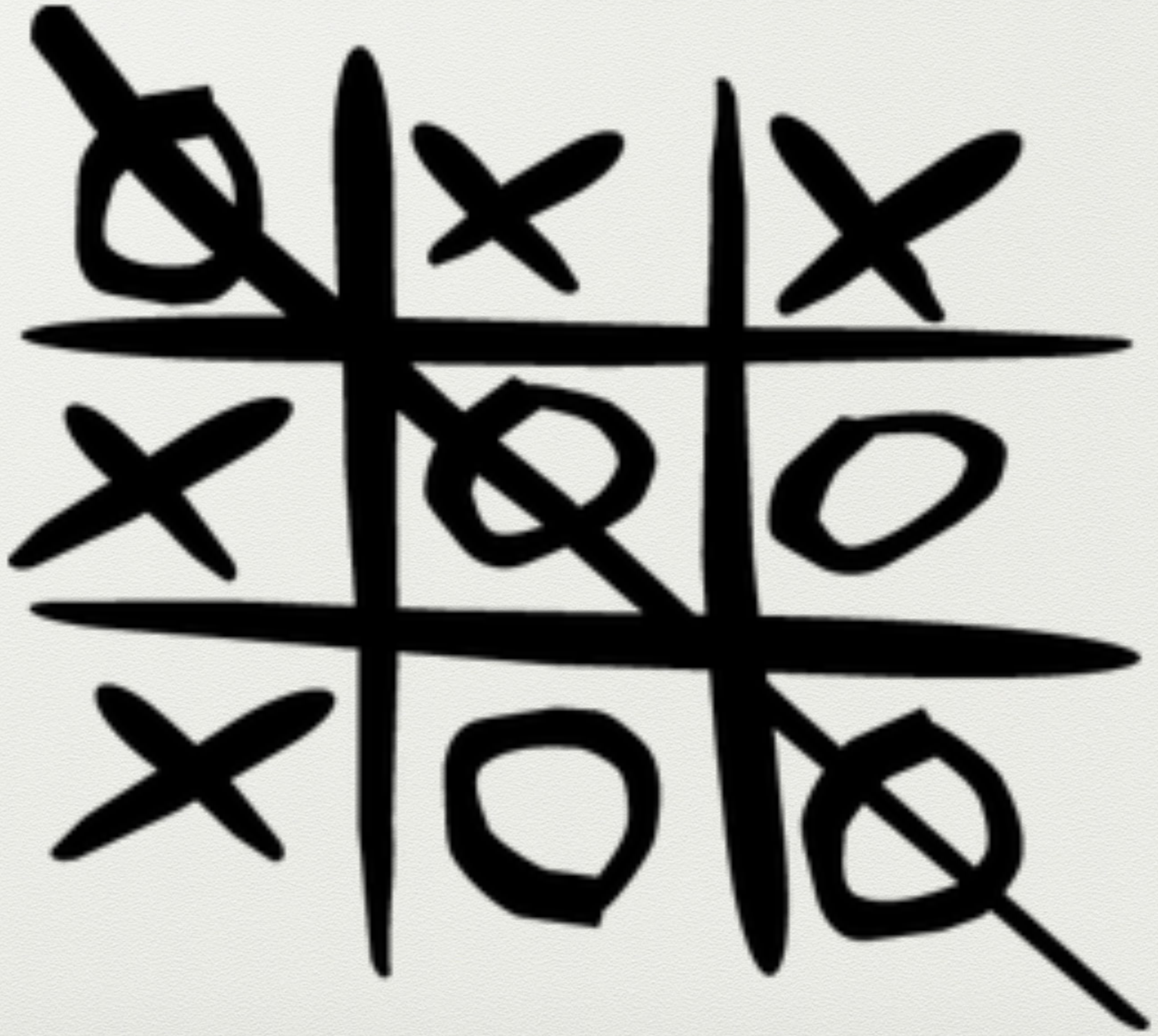
Legends:

A thick and safe layer of ice which you can walk over

A hole in the ice. Fall here and you're dead

The frisbee. Be the hero and get it back

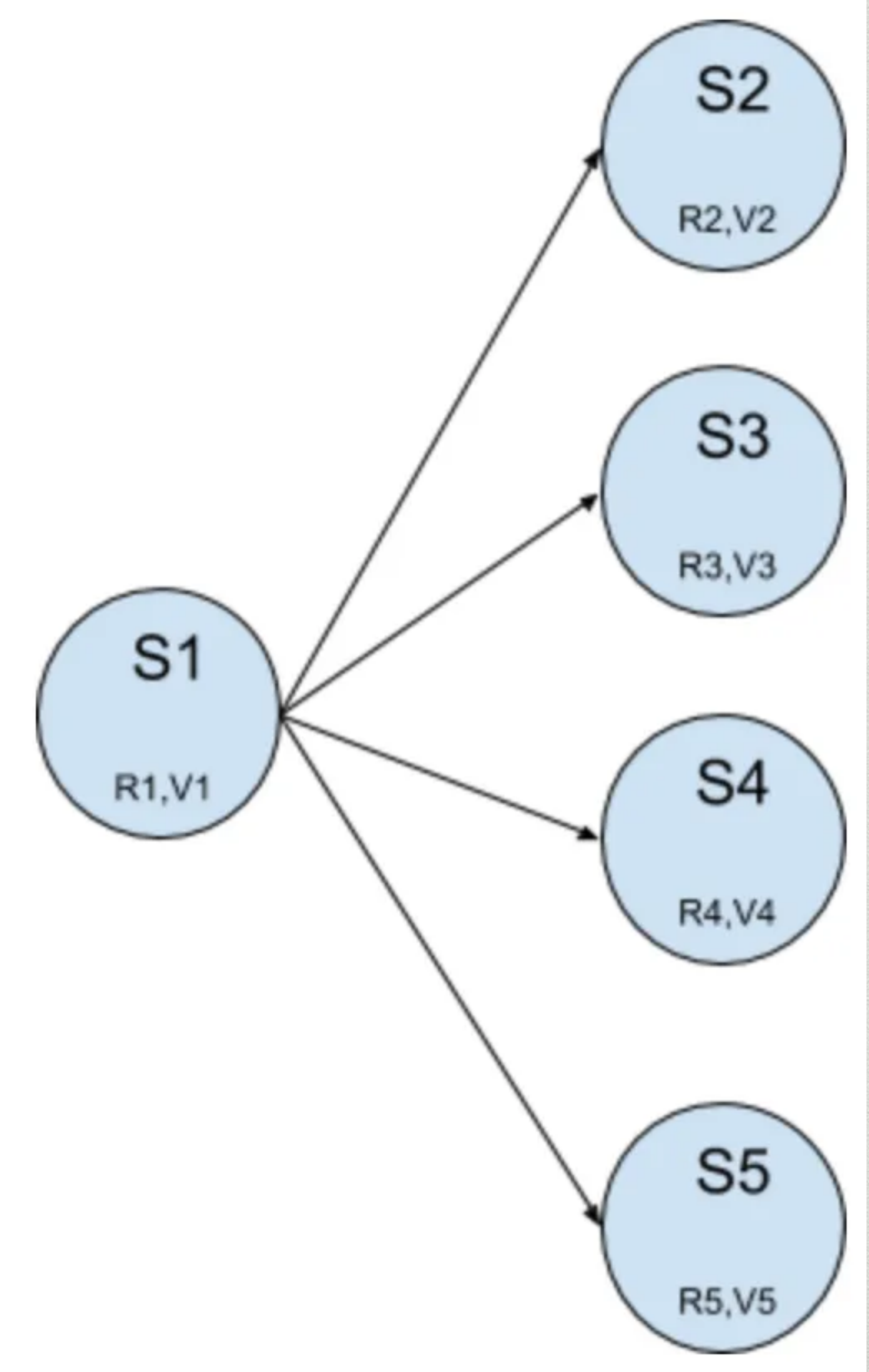
You



Otro tipo de redes…

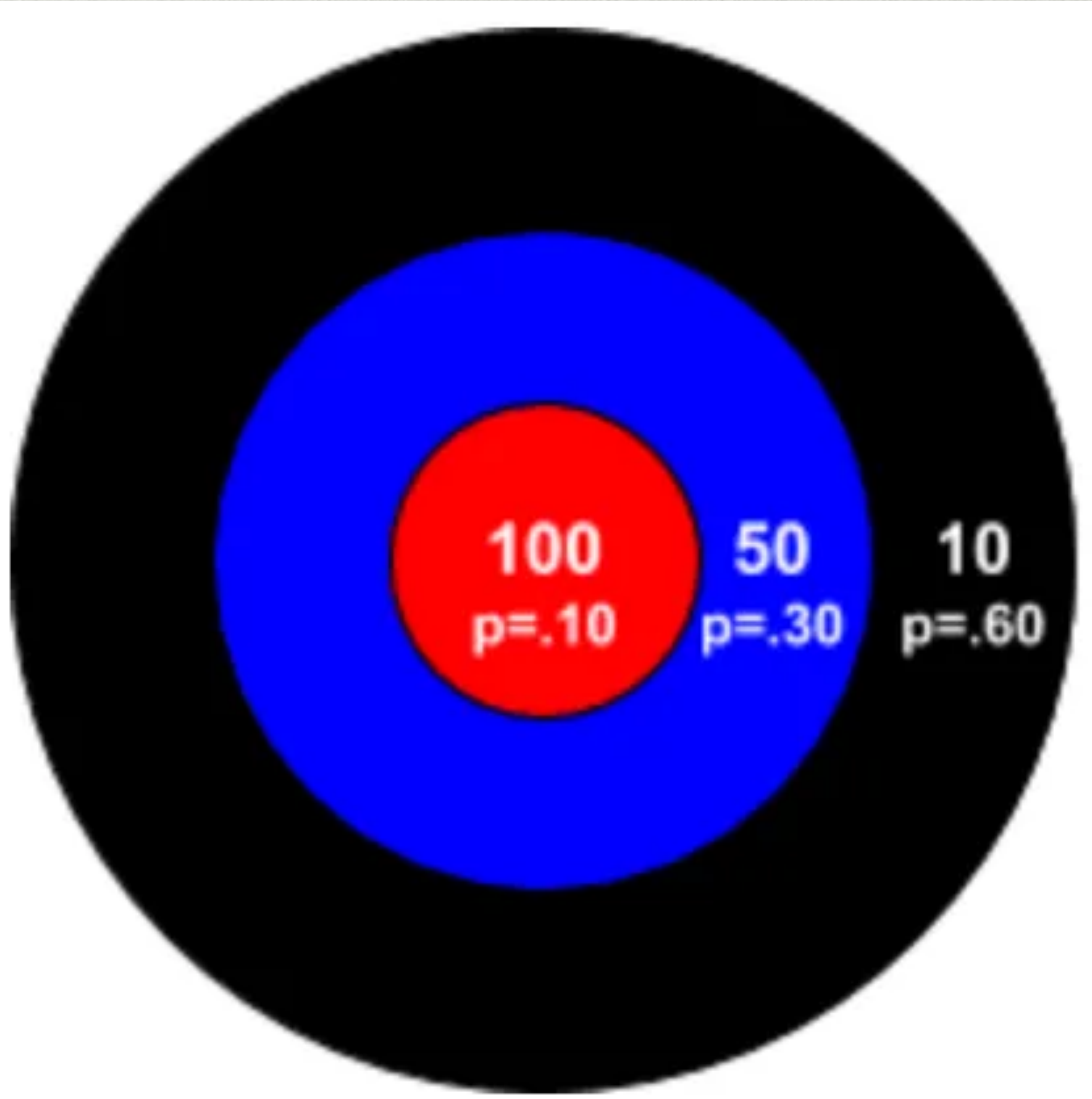
- La función de valor se suele definir como $V(s) = \sum_{a} \pi(a \mid s) \sum_{t}^{\infty} \gamma^t R(s_t)$ en donde $0 < \gamma < 1$ es el factor de descuento y determina la influencia de las recompensas futuras

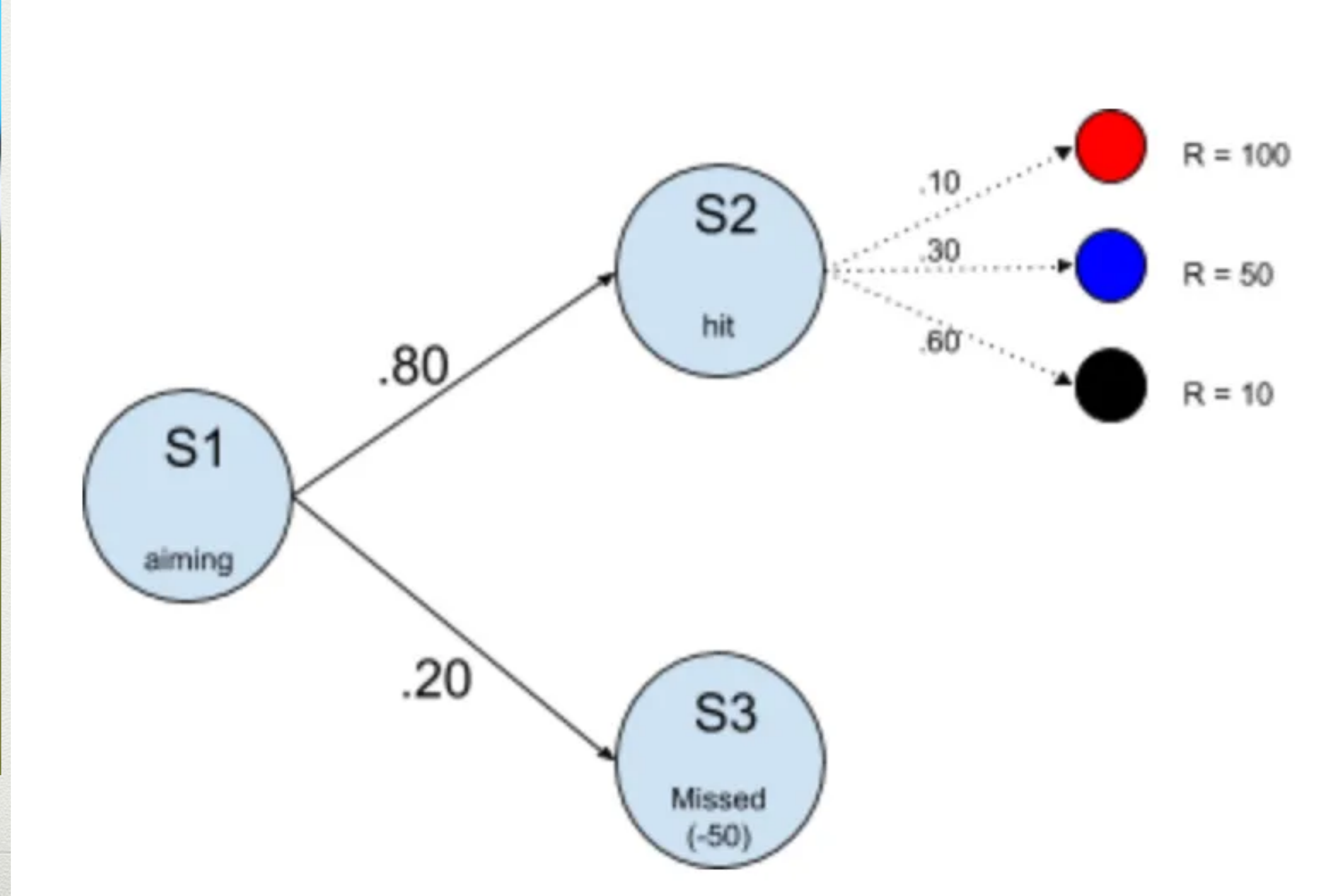
- En términos de los siguientes estados se tiene $V_\pi(s) = \sum_{a} \pi(a \mid s) \sum_{s' \in S} \sum_{r \in \mathscr{R}} p(s', r \mid s, a)(r + \gamma V_\pi(s'))$



Ejemplo de distribución de estados: Arquería

¿Y si se le para una mosca en la mano?





¿Y cómo encontramos π y V_π?

- Ver libros compartidos
- Ver https://gymnasium.farama.org/