

به نام خدا



# یادگیری تقویتی در روباتیک

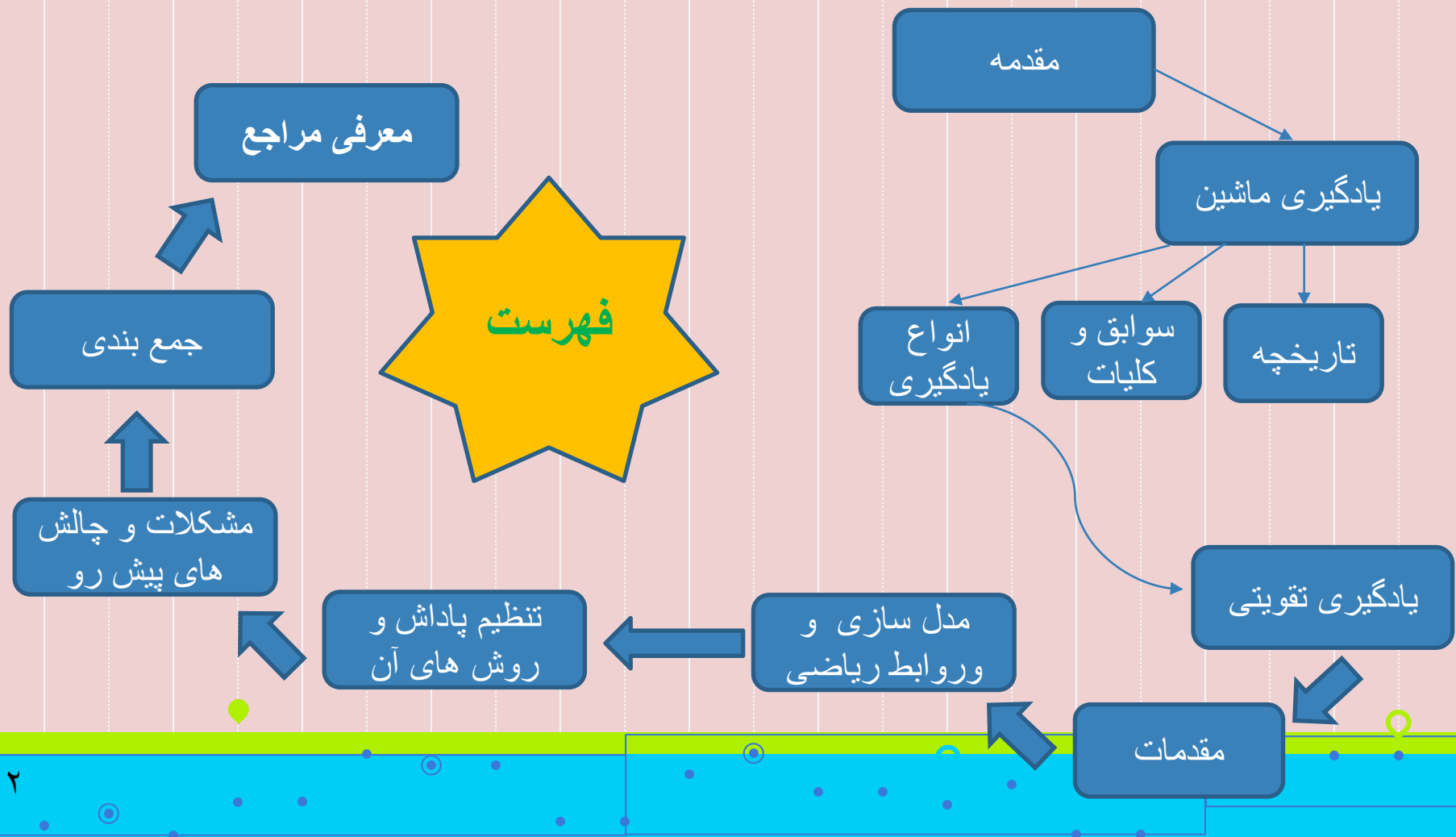
ارائه دهنده :  
یاسمن میرمحمد

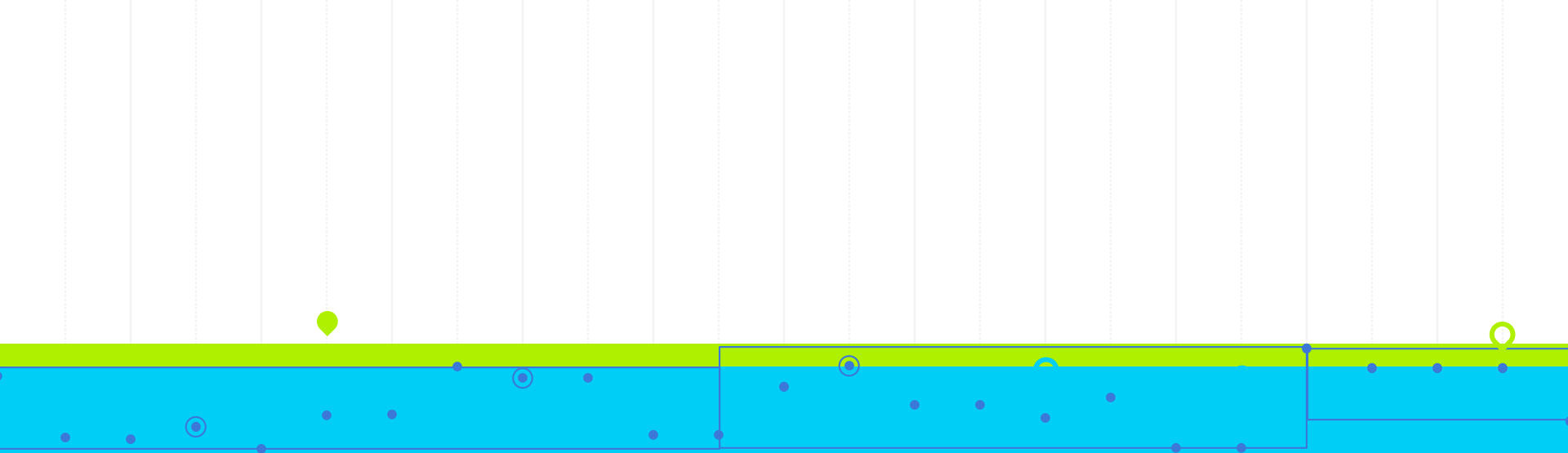
استاد راهنما:  
دکتر رضا صفابخش

پاییز ۹۶



دانشگاه صنعتی امیرکبیر  
( پلی تکنیک تهران )

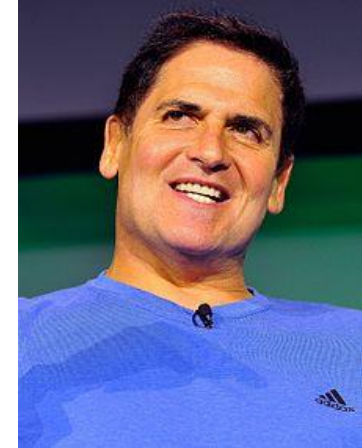




## مقدمه



“



هوش مصنوعی، یادگیری عمیق، یادگیری ماشین و هر مبحثی از این دست را که از آن “  
سر در نمی آورید یاد بگیرید؛ در غیر این صورت، در عرض سه سال آینده به یک دایناسور  
تبدیل خواهید شد!”  
مارک کوبان\_\_



مقدمه

یادگیری تقویتی

مدل مارکف

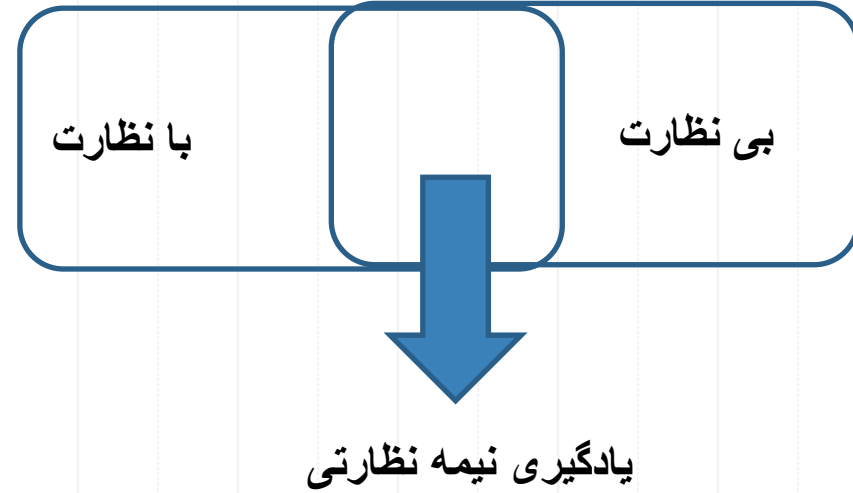
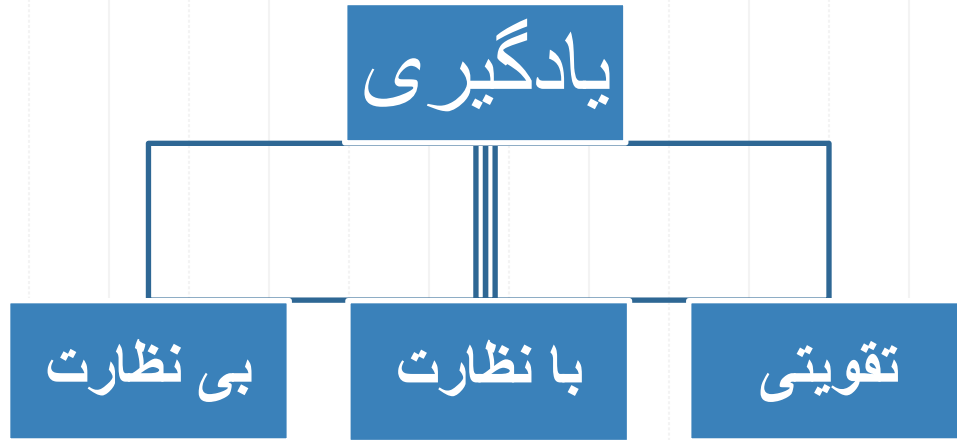
تنظیم پاداش

چالش ها

جمع بندی

منابع

# دسته بندی مسائل یادگیری ماشین



# مقدمه ای بر یادگیری تقویتی

## روانشناسی رفتارگرایی

بیشترین شباهت به یادگیری انسان  
بیشینه کردن پاداش

تجربه

تابع پاداش

مقدمه

یادگیری تقویتی

مدل مارکف

تنظیم پاداش

چالش ها

جمع بندی

منابع

# سوابق و کلیات



R. E. Bellman



D. P. Bertsekas

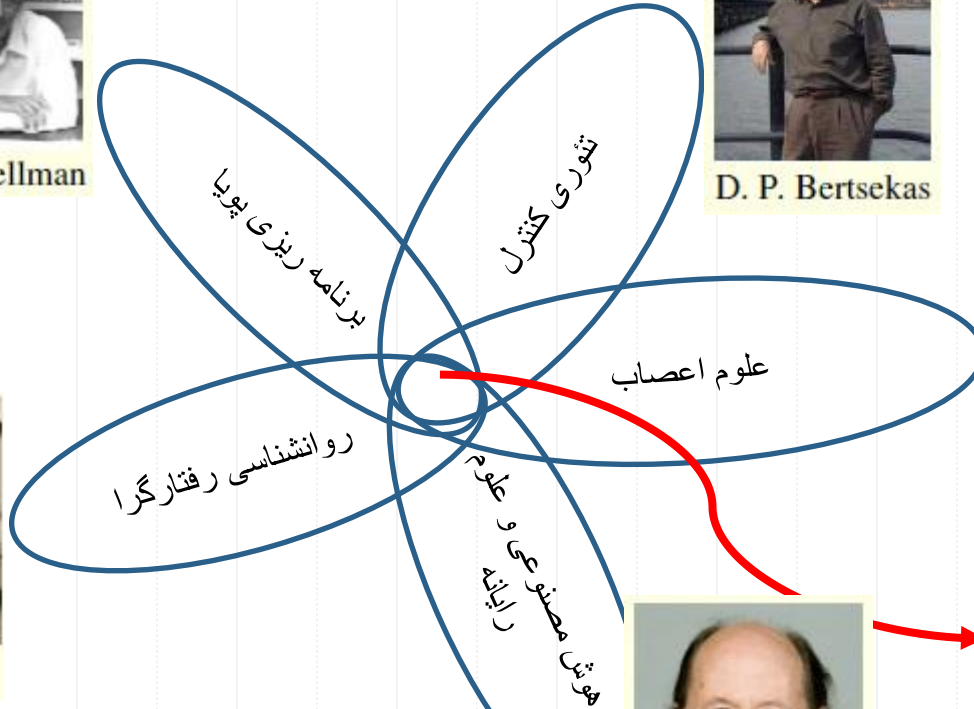


W. Schultz

یادگیری تقویتی



R. S. Sutton



B. F. Skinner

مقدمه

یادگیری تقویتی

مدل مارکف

تنظیم پاداش

چالش ها

جمع بندی

## محدودیت ها



برای یک سری مسائل با  
ویژگی های خاص

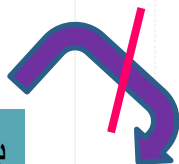


بیشترین سود در کوتاه مدت

## برتری



❖ استفاده از نمونه ها  
❖ تخمین توابع



بیشترین سود در دراز مدت

## انگیزه طرح



❖ خروجی مناسب  
❖ برچسب

مقدمه

یادگیری تقویتی

مدل مارکف

تنظیم پاداش

چالش ها

جمع بندی

منابع



# کاربردها

داده های بی برچسب  
رفتارهای ممتد و نامتناهی

تشبیه:

رابطه ی ریاضیات و فیزیک



یادگیری تقویتی

نظریه بازی ها

الگوریتم  
ژنتیک

نظریه کنترل

روباتیک

هوش  
ازدحامی

سامانه های  
چند عامله

مقدمه

یادگیری تقویتی

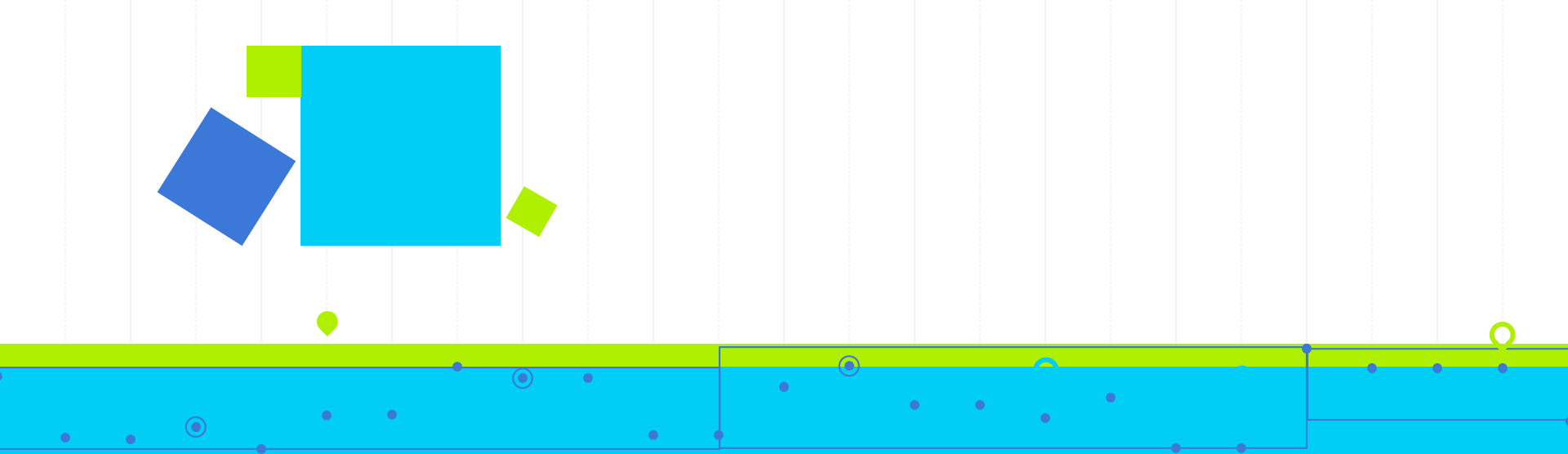
مدل مارکف

تنظیم پاداش

چالش ها

جمع بندی

منابع



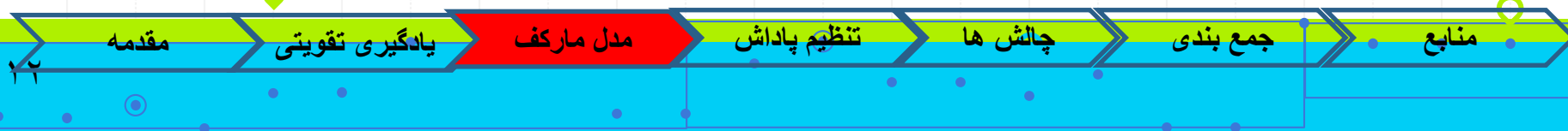
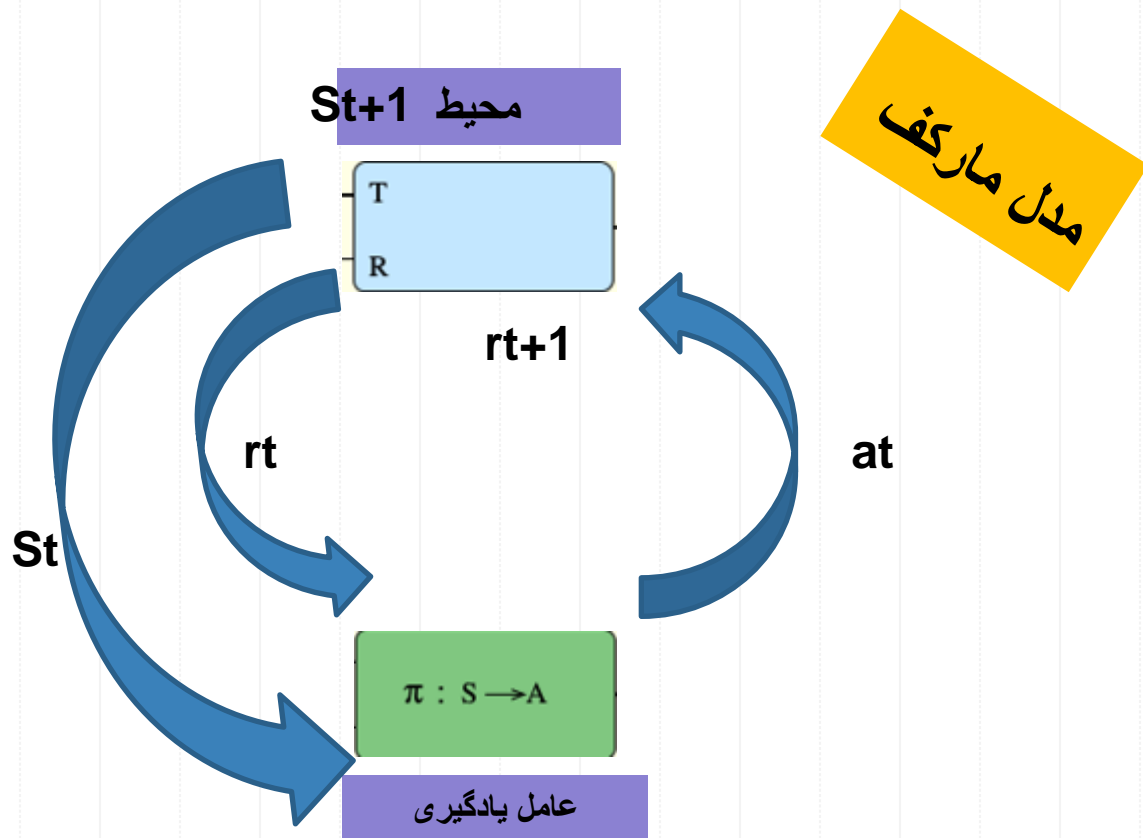
# ۱ مدل مورد استفاده و روابط ریاضی

S: مجموعه‌ای از حالات مختلف مساله

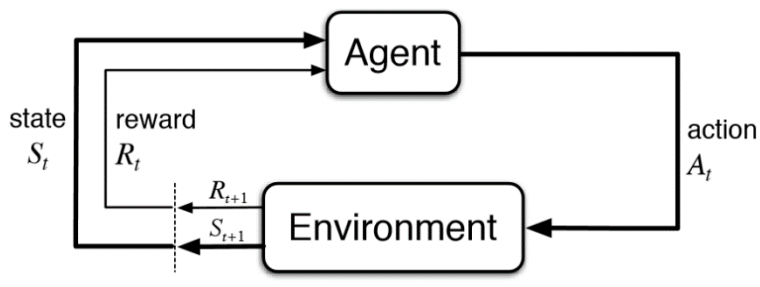
A: مجموعه‌ی تصمیماتی که میتواند اتخاذ شود.

T: یک تابع گذار  
 $T(s', a, s)$

R: یک تابع پاداش  
 $R(s, a)$



# مدل مارکف



$$R = \sum_{t=0}^{N-1} r_{t+1}$$

□

یافتن یک قاعده

$$J = \lim_{H \rightarrow \infty} E \left\{ \frac{1}{H} \sum_{h=0}^H R_h \right\}$$

$$J = E \left\{ \sum_{h=0}^H R_h \right\}$$

امید ریاضی

حالت بعدی و تابع پاداش ➡ حالت قبلی + تصمیم اتخاذ شده

مقدمه

یادگیری تقویتی

مدل مارکف

تنظیم پاداش

چالش ها

جمع بندی

منابع

# ۲ تنظیم پاداش

- محدودیت زمانی
- مدل سازی در دنیای واقعی

## معادله ی بلمن

تابع ارزش

$$V^{\pi}(s) = \sum_{s' \in \mathcal{S}} T(s, \pi(s), s') [R(s, \pi(s), s') + \gamma V^{\pi}(s')].$$

$$J(\pi) = \sum_{s,a} \mu^{\pi}(s) \pi(s,a) R(s,a) \quad .$$

مقدمه

یادگیری تقویتی

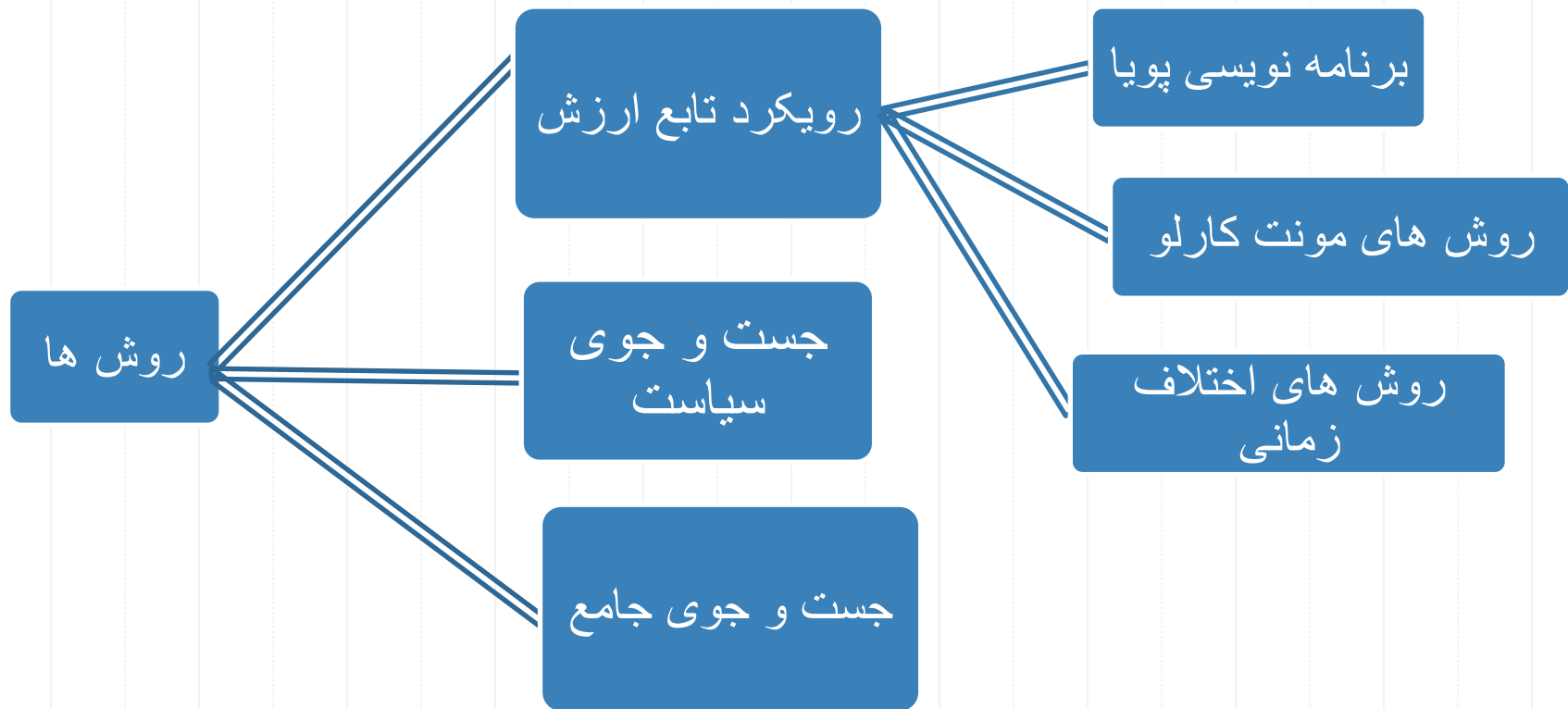
مدل مارکف

تنظیم پاداش

چالش ها

جمع بندی

منابع



مقدمه

یادگیری تقویتی

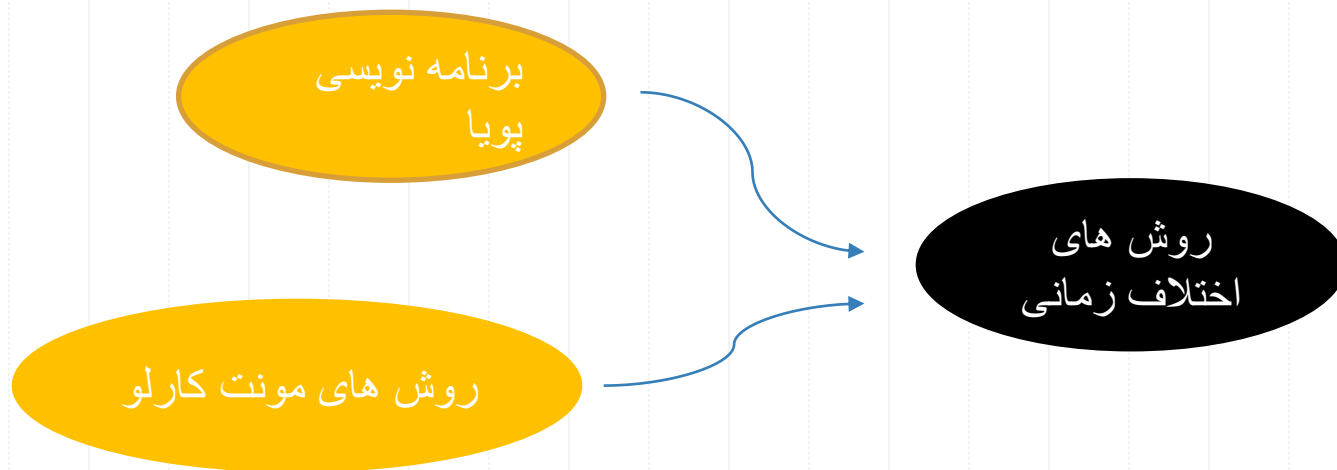
مدل مارکف

تنظیم پاداش

چالش ها

جمع بندی

منابع



مقدمه

یادگیری تقویتی

مدل مارکف

تنظیم پاداش

چالش ها

جمع بندی

منابع



## جست و جوی جامع

مشکلات



- تعداد حالات
- تنوع خروجی ها

- به ازای همه رویکردهای ممکن، در حین دنبال کردن آنها از پاداش‌ها نمونه برداری کن.
- رویکردی را که بیشترین مجموع پاداش را دارد انتخاب کن.

مقدمه

یادگیری تقویتی

مدل مارکف

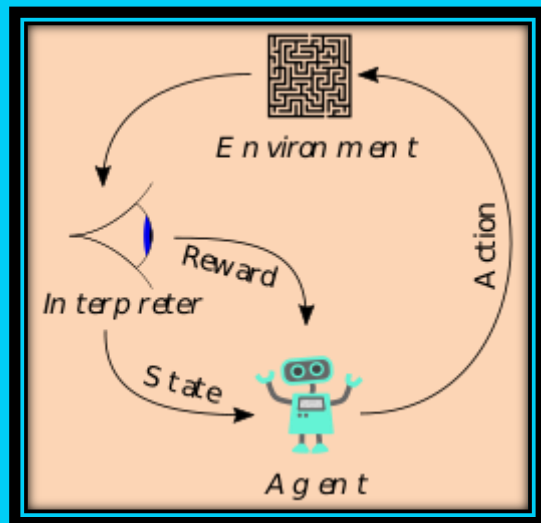
تنظیم پاداش

چالش‌ها

جمع بندی

منابع

# ۳ چالش ها و مشکلات



- مشکل محیط
- مشکل ابعاد
- مشکل شبیه سازی در دنیای واقعی
- مشکل مدل و الگوریتم
- مشکل تعیین هدف
- مشکل اکتشاف

مقدمه

یادگیری تقویتی

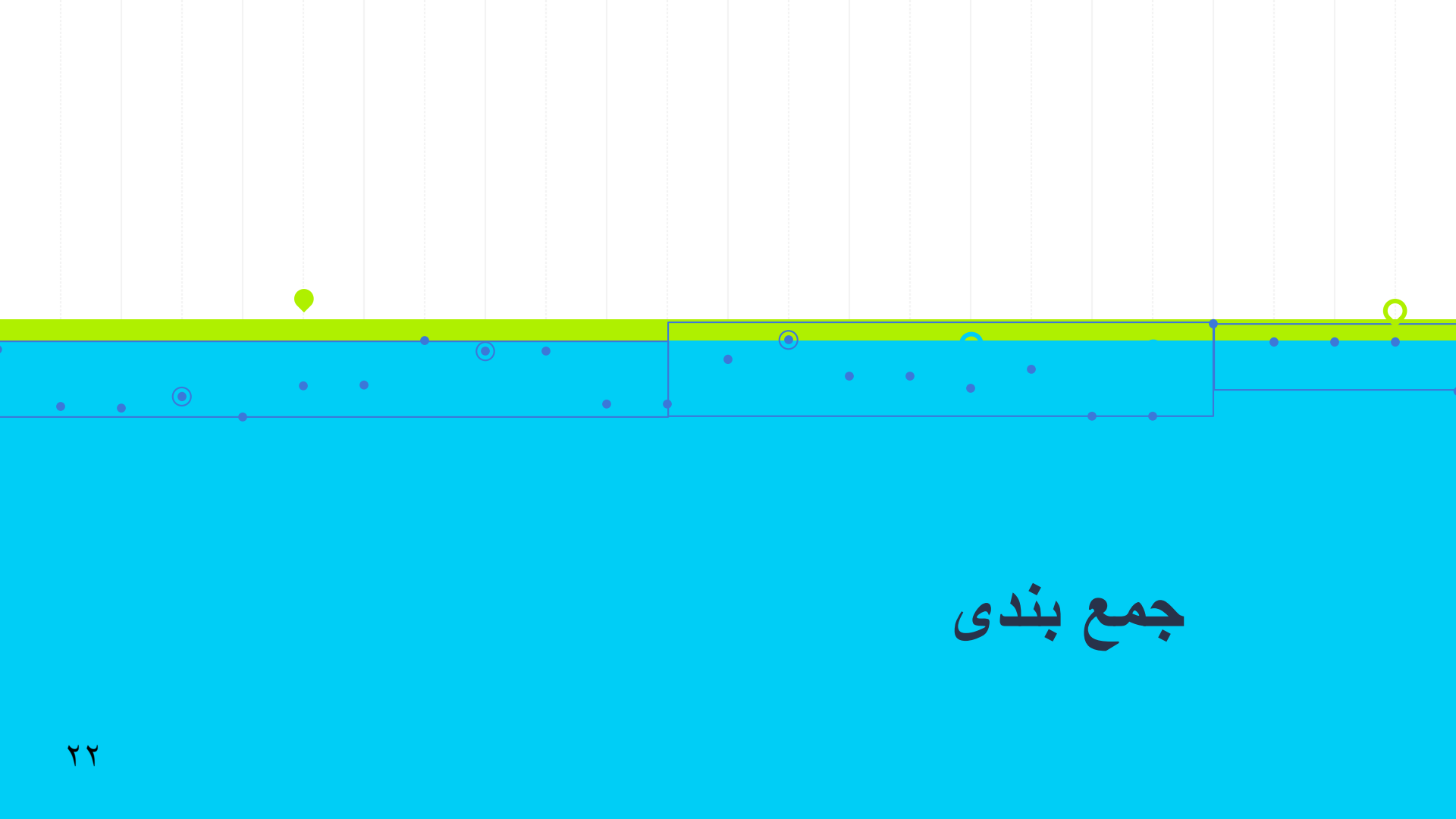
مدل مارکف

تنظیم پاداش

چالش ها

جمع بندی

منابع



جمع بندی

## جمع بندی

یادگیری  
ماشین

یادگیری  
تقویتی

تعامل با  
محیط+ماکزیم کردن  
تابع پاداش

مدلسازی  
رفتاری مساله

شرایط  
مدلسازی  
مارکف

روش های  
تنظیم پاداش

1. مشکل ابعاد  
2. مشکل اکتشاف

پایان!

مقدمه

یادگیری تقویتی

مدل مارکف

تنظیم پاداش

چالش ها

جمع بندی

منابع



“به محض اینکه ماشین ها توانایی تفکر را پیدا کنند، کاری انجام خواهند داد که هیچ انسانی به پایان نرسانده، و آن مفهوم حقیقی یادگیری است!”

\_\_آنتونی هینکس

مقدمه

یادگیری تقویتی

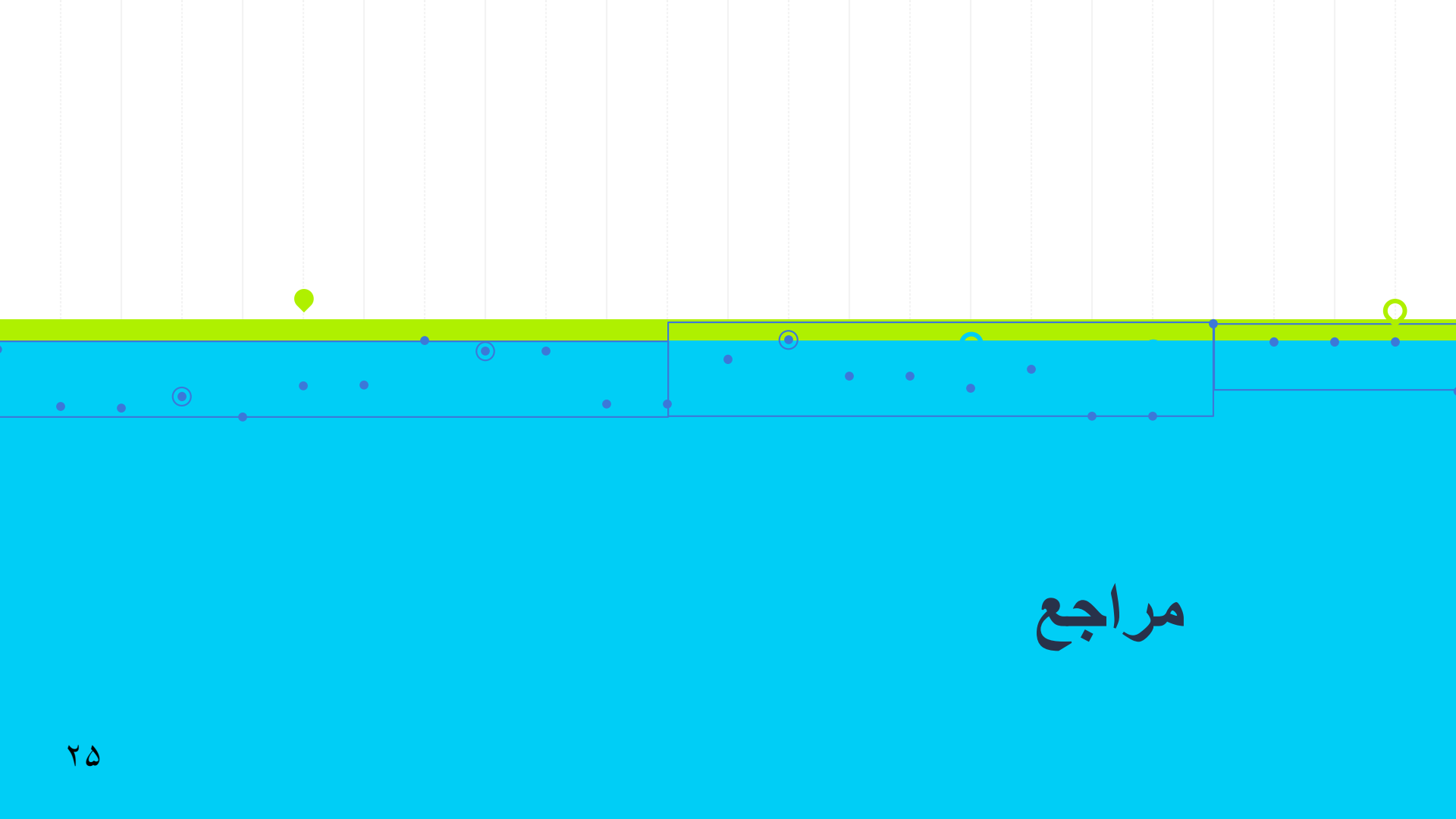
مدل مارکف

تنظیم پاداش

چالش ها

جمع بندی

منابع



مراجع

**[1] J Kober, JA Bagnell, J Peters ,” Reinforcement learning in robotics: A survey”, The International Journal of Robotics Research, vol. 32, pp. 1238-1274, September 2013**

**[ 2]Christopher. M Bishop, “ Pattern Recognition and Machine Learning” ( Information Science and Statistics),1 st ed, New York, Springer , August 2006**

**[3]Larry Wasserman, “All of Statistics”(A Concise Course in Statistical Inference), Springer Science & Business Media, December 2013**

**[4] Richard S. Sutton, J.Andrew Barto , “Reinforcement learning: An introduction”,US, MIT press , March 1998**

مقدمه

یادگیری تقویتی

مدل مارکف

تنظیم پاداش

چالش ها

جمع بندی

منابع



[5]Wikipedia.com,2017, Reinforcement Learning

Available at:

[https://en.wikipedia.org/wiki/Reinforcement\\_learning/](https://en.wikipedia.org/wiki/Reinforcement_learning/)

[6]Simple Beginner's guide to Reinforcement Learning & its implementation, Available at:

<https://www.analyticsvidhya.com/blog/2017/01/introduction-to-reinforcement-learning-implementation/>

[7] Wikipedia.com,2017 , Temporal difference learning

Available at:

[https://en.wikipedia.org/wiki/Temporal\\_difference\\_learning](https://en.wikipedia.org/wiki/Temporal_difference_learning)

مقدمه

یادگیری تقویتی

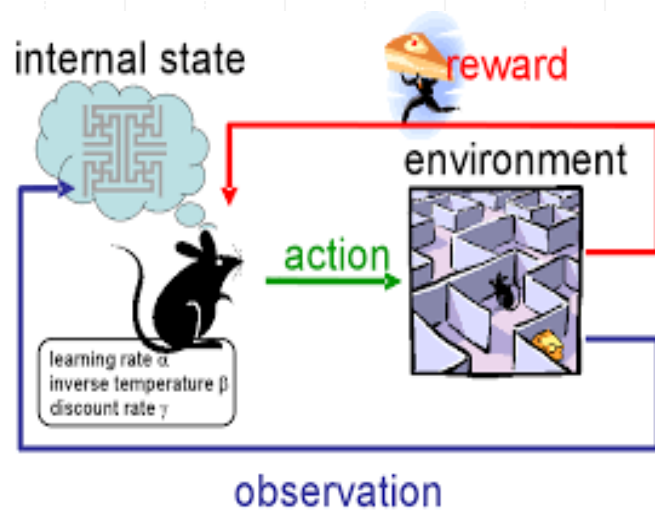
مدل مارکف

تنظیم پاداش

چالش ها

جمع بندی

منابع



سیاس از توجهتان!