



یادگیری تقویتی

نیم سال بهار ۱۴۰۱-۴۰۲

اساتید: دکتر رهبان، آقای حسنی

تاریخ: ۵ اردیبهشت

کوییز چهارم

مدت: ۲۰ دقیقه

شماره دانشجویی:

نام و نام خانوادگی:

سوال ۱: سوال صحیح و غلط (۳۰ نمره)

پاسخ به این سوال نیاز به فرمول نویسی ندارد و صرفاً یک توضیح کوتاه کافی است.

(آ) روش های یادگیری تقویتی مبتنی بر مدل، علاوه بر value یا تابع policy، مدلی از محیط، از جمله احتمالات و پاداش های انتقال را یاد می گیرند.

(ب) معادله بلمن برای تابع state-value در یادگیری تقویتی مبتنی بر مدل مانند معادله بلمن برای تابع action-value است.

(ج) در یادگیری تقویتی مبتنی بر مدل، احتمالات و پاداش های انتقال در محیط ثابت و شناخته شده فرض می شوند.

(د) شبکه های عصبی بیزی با عدم قطعیت وزن را می توان به عنوان تقریبگر تابع در یادگیری تقویتی بیزی بدون مدل استفاده کرد.

(ه) در یادگیری تقویتی بیزی مبتنی بر مدل، استنتاج بیزی دقیق در مدل های پیچیده قابل حمل است و به روش های تقریبی مانند زنجیره مارکوف مونت کارلو یا استنتاج متغیر نیاز ندارد.

سوال ۲: Model-Based (۳۵ نمره)

(آ) در یادگیری تقویتی مبتنی بر مدل، تفاوت بین value iteration و policy iteration چیست و پیچیدگی محاسباتی هر روش چیست؟

(ب) در یادگیری تقویتی مبتنی بر مدل، معادله بلمن برای تابع state-value چیست و چه تفاوتی با معادله بلمن برای تابع action-value دارد؟

سوال ۳: Bayesian reinforcement learning (۳۵ نمره)

تفاوت بین توزیع های prior و posterior چیست و چگونه بر فرآیند یادگیری در یادگیری تقویتی بیزی تأثیر می گذارد؟