



دانشگاه صنعتی شریف

دانشکده مهندسی کامپیوتر

## هوش مصنوعی

بهار ۱۴۰۰

استاد: محمدحسین رهبان

گردآورندگان: نیما فتحی - سروش وفایی تبار

### Markov Decision Process

مهلت ارسال: ۱۹ خرداد

تمرین هفتم بخش دوم

- مهلت ارسال پاسخ تا ساعت ۲۳:۵۹ روز مشخص شده است.
- همکاری و همفکری شما در انجام تمرین مانعی ندارد اما پاسخ ارسالی هر کس حتما باید توسط خود او نوشته شده باشد.
- در صورت همفکری و یا استفاده از هر منابع خارج درسی، نام همفکران و آدرس منابع مورد استفاده برای حل سوال مورد نظر را ذکر کنید.
- لطفا تصویری واضح از پاسخ سوالات نظری بارگذاری کنید. در غیر این صورت پاسخ شما تصحیح نخواهد شد.

### سوالات نظری (۱۰۰ نمره)

۱. (۵۰ نمره) کدام یک سریع تر است؟ value iteration یا policy iteration؟ با کمک نوشتن معادلات دلایل خود را مطرح کنید.
۲. (۵۰ نمره) در روش policy extraction استفاده از مقادیر  $Q$  بهتر است یا  $V$ ؟ اگر فرقی ندارند اثبات کنید در غیر این صورت دلایل خود را کامل توضیح دهید.