



- پرسش‌های خود را می‌توانید در کوثرای درس و یا با ایمیل دستیار آموزشی مطرح کنید.

## بخش اول

لطفاً بین سه سوال زیر به دو سوال پاسخ دهید.

۱. (۲۵ نمره) بنظر شما تفاوت مسائل *Markov Decision Process* یا *MDP* با دیگر مسائل برنامه ریزی پویای ذکر شده در درس تحقیق در عملیات چیست؟  
آیا می‌توانید مسائل *MDP* را به شکل یک مدل برنامه ریزی خطی بنویسید که در شرایط معادله بلمن صدق کند؟
۲. (۲۵ نمره) پاداش کل مورد انتظار با تخفیف ( *discounted expected total reward* ) را در افق محدود در نظر بگیرید. عدد  $N$  را تعداد افق‌ها فرض کنید. لطفاً بررسی کنید که در اینجا می‌توان شرط دوم مطرح شده در کلاس را از شرط یک بدست آورد یا نه. اگر  $N$  یک متغیر تصادفی مستقل از مدل *MDP* است، این حالت را دوباره بررسی کنید. (این مدل را *stochastic terminated MDP* گوئیم).
۳. (۲۵ نمره) توضیح دهید چطور می‌شود یک مساله درخت تصمیم‌گیری ( *decision tree* ) را به کمک معادله بهینگی به عنوان یک مساله *MDP* نوشت.  
توجه کنید که معادله بهینگی که در ادبیات به آن معادله بلمن هم گفته می‌شود از مهمترین قضایا در حل مسائل برنامه ریزی پویای تصادفی یا همان *MDP* است. برای مطالعه بیشتر بخش ۱۴.۲.۱ کتاب مرجع درس را مطالعه فرمایید.

## بخش دوم

از بین سوالات اول تا هفتم فصل چهاردهم کتاب مرجع (آپلود شده در کوثرای درس)، لطفاً به دو سوال پاسخ دهید. (هر سوال ۲۵ نمره دارد). نوشتن بخش سوم تمرین شماره پنج اجباری نیست.