

Üzleti Intelligencia

Megerősítéssel tanulás - beadandó feladatok

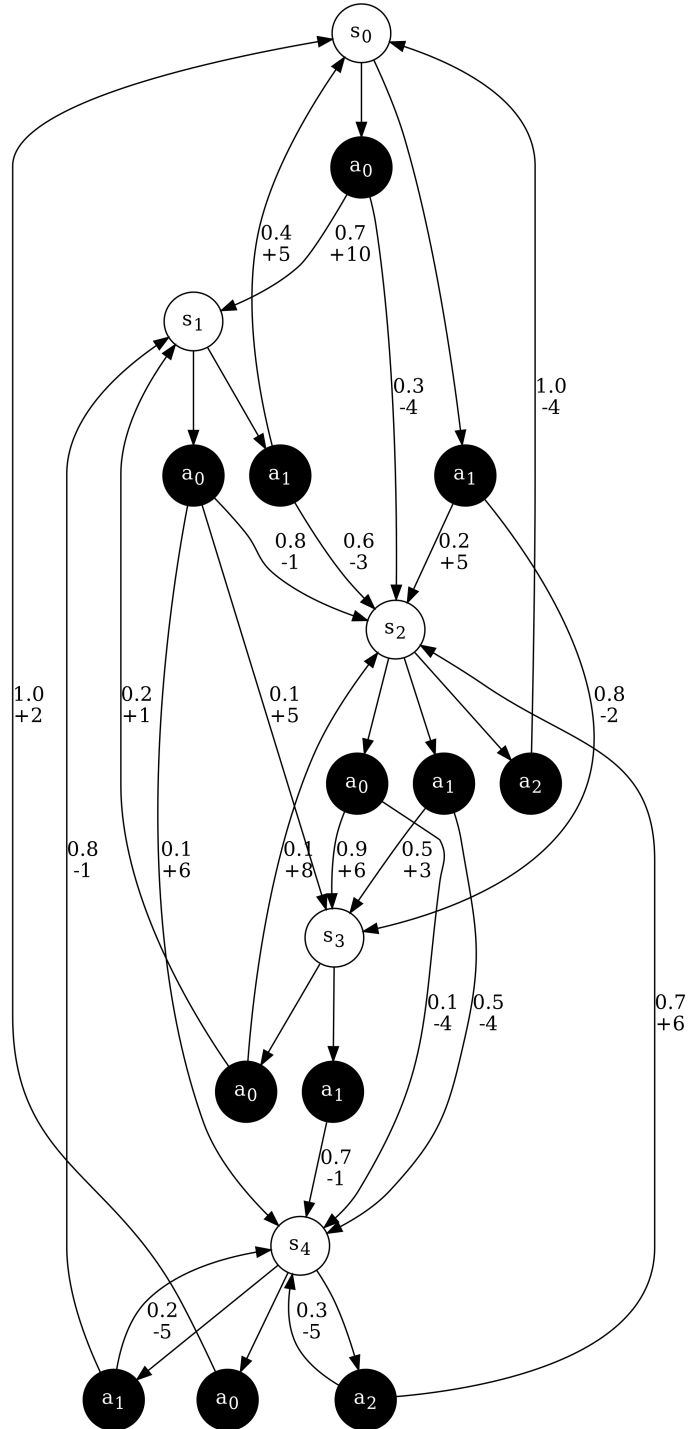
2023

A feladatok 1-3 skálán vannak osztályozva nehézség szerint, ahol 1 a legkönnyebb és 3 a legnehezebb. A feladatok véletlenszerűen sorsolódtak ki. A feladatok beadása Coospace felületen történik, ahol csak egy linket kell beküldeni, ami a feladatot megvalósító Git tárhelyre mutat. Késő beadás nem lehetséges, a dátum beadásakor feltöltött anyagok fognak leosztályozásra kerülni. Minden további információ a tantárgyi útmutatóban és az órán elhangzottakban találhatóak.

Az elégséges jegy feltételen, hogy minden részfeladat legalább 60%-os sikerességgel megoldásra kerüljön.

1 MDP

Implementálja az alábbi $MDP(S, A, P, R, s_0, \gamma)$ Markov döntési folyamatot. Az állapotok valószínűsége tizedes törtként szerepel, a jutalmak pedig + és - előjellel a kapcsolatokon. Használja az alábbi paramétereket: $s_0 = 0$, $\gamma = 0.98$



1.1 (A) feladat

- Adja meg minden állapot értékét a megfelelő Bellman egyenlettel. Ábrázolja diagramon az értékek konvergenciáját.
- Adja meg az optimális politikát Monte Carlo algoritmussal. Ehhez tartozóan implementálja a MC politika keresés és kiértékelés algoritmusát. Ábrázolja diagramon az értékek konvergenciáját.
- Oldja meg a környezetet Q -tanulással. Adja meg az optimális politikát a Q -táblából származtatva. Ábrázolja a Q -értékek konvergenciáját diagramon.
- Oldja meg a környezetet párbajozó mély Q -tanulással. Ábrázolja a jutalmak növekedését diagramon.
- Minden feladatrészhöz tartozóan írja le az eljárás elméleti alapjait és a saját tapasztalatait röviden egy Jupyter Markdown cellában.

1.2 (B) feladat

- Oldja meg a környezetet dinamikus programozással. Ehhez tartozóan implementálja a politika iteráció algoritmusát. Adja meg az optimális politikát. Ábrázolja az értékek konvergenciáját diagramon.
- Adja meg az optimális politikát temporális különbségek algoritmusával. Ehhez tartozóan implementálja a TD politika keresés és kiértékelés algoritmusát. Ábrázolja diagramon az értékek konvergenciáját.
- Oldja meg a környezetet dupla Q -tanulással. Adja meg az optimális politikát a dupla Q -táblából származtatottan. Ábrázolja az értékek konvergenciáját diagramon.
- Oldja meg a környezetet dupla mély Q -tanulással. Ábrázolja a jutalmak növekedését diagramon.
- Minden feladatrészhöz tartozóan írja le az eljárás elméleti alapjait és a saját tapasztalatait röviden egy Jupyter Markdown cellában.