

BME VIK Számítástudományi és Információelméleti Tanszék

Diplomamunka, önálló labor és TDK témajavaslat

Témakiíró:	Dr. Csáji Balázs Csanád
Beosztás:	tudományos főmunkatárs
Munkahely:	MTA SZTAKI
Email:	balazs.csaji@sztaki.mta.hu
Cím:	Megerősítéses tanulás
Leírás:	<p>A megerősítéses tanulás alapmodellje, hogy egy ágens (döntéshozó) egy bizonytalan (sztochasztikus) környezetben tevékenykedik, és a rendszertől kapott (állapot és költség) visszacsatolások alapján szeretne egy olyan döntési stratégiát találni, amely minimalizálja a hosszú távú (tipikusan várható diszkontált vagy átlagos) költségeit. A probléma matematikai leírására egy klasszikus modellt jelentenek a Markov döntési folyamatok (MDP), amelyek általános keretrendszerébe számtalan valós feladat átfogalmazható.</p> <p>A megerősítéses tanulás egyik célja a klasszikus dinamikus programozáson alapuló megoldások hátrányainak („curse(s) of dimensionality”) kiküszöbölése. A gyakorlatban előforduló feladatoknál tipikusan (1) a környezet dinamikája nem ismert (meg kell tanulni a megfigyelésekből), (2) az állapottér folytonos, vagy ha esetleg véges is, mérete óriási, így az értékelő függvény nem tárolható el közvetlenül a memóriában (csak egy approximációja), valamint (3) a (közelített) értékelő függvény frissítése is csak részlegesen történhet meg. További nehézségeket jelent, ha (4) a rendszer esetleg csak parciálisan megfigyelhető, vagy ha (5) a rendszer dinamikája időben megváltozhat.</p> <p>Feladat lehet például gyakorlati problémák megfogalmazása a megerősítéses tanulás keretrendszerében (MDP); különböző tanulóalgoritmusok implementálása, szimulációs kísérletek elvégzése és algoritmusok összehasonlítása; továbbfejlesztése (pl., strukturális tulajdonságok kihasználásával); modellek tanulása megfigyelési adatokból; változások felismerése a rendszer dinamikájában; adaptációs képességek vizsgálata; a felfedezés vs kihasználás probléma elemzése (például egyszerű, kis állapotterű, speciális szerkezetű feladatosztályok esetén); algoritmusok vizsgálata több tanuló ágens esetén, stb.</p> <p>Irodalom:</p> <ul style="list-style-type: none">- Sutton, Richard; Barto, Andrew: Reinforcement Learning: An Introduction. First edition: MIT Press. 1998. Second edition: 2017 (online draft).- Szepesvári, Csaba. Algorithms for Reinforcement Learning. Morgan & Claypool. 2010.- Dimitri P. Bertsekas, John Tsitsiklis. Neuro-Dynamic Programming. Athena Sci. 1996.- Csáji, Balázs Csanád; Monostori, László: Value Function Based Reinforcement Learning in Changing Markovian Environments, Journal of Machine Learning Research (JMLR), MIT Press and Microtome Publishing, Vol. 9, 2008, pp. 1679–1709.