

Tanulás elosztott rendszerekben/1

(Egyedi ágens) tanulásáról röviden

Célja: javulás (feladavégzésben), adaptalódás, robusztusság (környezet), kompenzálás, hibatűrés (ismerethiány, meghibasodás).

Miből: példák (forrásuk a környezet), háttértudás (forrása a fejlesztő), példák minősítése (kritikus, tanár).

Megközelítések: felügyelt, megerősítéses, felügyelet nélküli.

Tipikus **tanulási absztrakció:** induktív függvénytanulás (példák alapján), ami osztályozás (döntés), vagy regresszió (approximáció).

Absztrakt **tanulási algoritmus:** helyes függvény keresése hipotézisek terében.

Absztrakt tanulási algoritmus **problémái:** tanulási zaj, tanulási elfogultság, tanulás komplexitása (VKH tanulás), empirikus és háttér tananyag kölcsönhatása, és nagyon sok konkrét algoritmus (tudásreprezentáció függvényében).

MAS tanulás - Tények/kihívások

Hipotézistér

Itt a tanulás szintén keresés a hipotézisek terében, de a tér (a többi ágens miatt) most szokatlanul nagy.

Kialakuló (Emergent)

A kölcsönhatások bonyolultága miatt a tanult viselkedésben bekövetkező kis változások nem megjósolható (emergent, kialakuló) változásokhoz vezethetnek a szervezet viselkedésének egészében.

Játékelmélet

Többágenses szervezetben több független tanulási folyamat van jelen. Mindegyik ágens tanul és adaptálódik a mások által meghatározott kontextusban (tanulás = stratégia). Ez természetes módon játékelméleti modellekhez vezet, amelyek kapcsolata tanulással még nem tisztázott.

Tanítók

A kölcsönhatások említett komplexitása miatt nemigen esélyes felügyelt tanulásra törekedni. Többágenses rendszer együttes tanulásához nehéz megfelelő tudású kritikust találni.

Megerősítés

Amiatt ez egyik elterjedtebb tanulási forma a többágenses megerősítéses (kooperatív és versengő) tanulás és az evolúciós számítások (cooperative/competitive coevolution).

Érdemhozzárendelés

A környezeti változások az összes ágens együttes cselekvésének hatására következnek be. Ágensünk cselekvése mennyire volt ebben meghatározó? Milyen nagy az érdeme/bűne a közösség szempontjából kifejezetten előnyös/káros környezeti állapotok elérésében?

Globális megerősítés, lokális megerősítés, $\lambda u_{\text{saját}} + (1-\lambda) u_{\text{jólét}}$,

Wonderful Life Utility (team haszna nélkül), leszámoltatás helyett átlagolás feladatszekvenciák felett, ... *senkinek nem érdeke, de közérdek*

Stacionaritás

Több, parallel és konkurrens módon tanuló ágens jelenléte ágens környezetét nemstacionáriussá teszi, ami sérti a tradicionális (egyágenses) tanulási módszerek egyik alapvető feltételét.

Dinamika

A környezet mások tanulásának ütemében változik. Ha az ugyanilyen gyors, mint az ágensünk tanulása, akkor az soha nem éri el a „megtanult”-nak nevezhető állapotot.

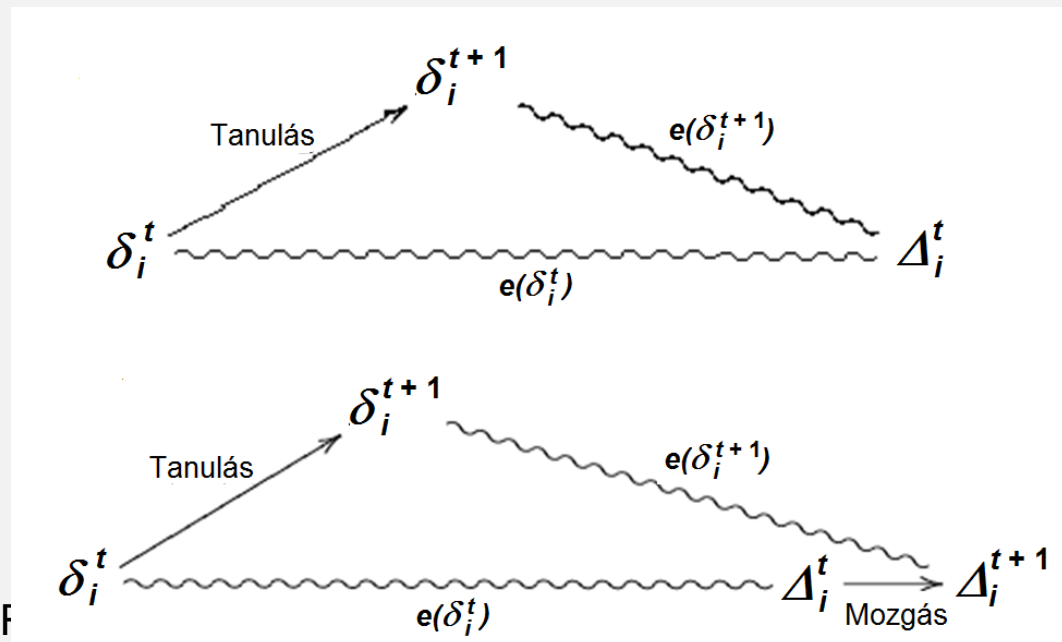
Konvergencia

Tanulás konvergenciája megtévesztő lehet és a mérnöki szempontból nem lehet egyedüli cél. Lehet, hogy a konvergens tanulás a szociális jólétre nézve (hasznosságban) veszteséges, a nem konvergens tanuláshoz képest.

A mozgó cél

Tanuló ágens környezete dinamikus, mert a benne lévő más ágensek is tanulnak és alakulnak át.

Mivel az ágensünknek éppen más ágensek viselkedését ki kell tanulnia, hogy az együttműködés/versengés miatt helyesen döntsön, azok tanulása a fő zavaró tényező.



Mi a **valószínűsége a sikeres megtanulásnak**, ha bizonyos ráhatások valószínűsége adott.

CLRI – elmélet (Change, Learning, Retention, Impact)

Változékonyság (changing rate – $c(i)$)

Tanulás sebessége (learning rate – $l(i)$)

Megtartás sebessége (retention – $r(i)$) helyes volt, és helyes marad

Illó sebesség, illékonyság (volatility – $v(i)$)

$$c(i) = \Pr[\text{változik} \mid \text{hibás volt}] \quad \forall_w c(i) = \Pr[\delta_i^{t+1}(w) \neq \delta_i^t(w) \mid \delta_i^t(w) \neq \Delta_i^t(w)]$$

$$l(i) = \Pr[\text{megtanulta a régit} \mid \text{hibás volt}] \quad \forall_w l(i) = \Pr[\delta_i^{t+1}(w) = \Delta_i^t(w) \mid \delta_i^t(w) \neq \Delta_i^t(w)]$$

$$r(i) = \Pr[\text{megtanulta a régit} \mid \text{helyes volt}] \quad \forall_w r(i) = \Pr[\delta_i^{t+1}(w) = \Delta_i^t(w) \mid \delta_i^t(w) = \Delta_i^t(w)]$$

$$v(i) = \Pr[\text{mozog}] \quad \forall_w v(i) = \Pr[\Delta_i^{t+1}(w) \neq \Delta_i^t(w)]$$

Nash-egyensúly

Stacionaritás nélkül az optimálisan megtanulthoz való konvergencia (a példaszám növekedésével) nem biztosítható. Azonban a tanulással valahová tartani kell.

Nash-egyensúly egy alternatíva, de láttuk, hogy ez nemigen egy optimális szociális jólét alternatívája.

És mi van, ha több Nash-egyensúly van?

Racionális NE \leftrightarrow optimális jólét, **NE koordinálás?**

Racionalitás igénye másodlagos a team optimális viselkedéséhez képest, kooperatív esetben hasznosságok korreláltak, mozgás globális NE felé lehetséges, más esetben koordinálás NE \rightarrow GE kérdéses, inkább \rightarrow NE

Kommunikáció

Kommunikáció egy információforrás, szükséges a tanuláshoz. Azonban kommunikáció erőforrásigényes, a tanulás is. Mi van, ha egy ágens korlátosan racionális? Milyen legyen akkor a kommunikáció?

Lehet tanulni kommunikáció érdekében is?

Direkt kommunikáció (dekódolható változás a környezetben)
(megosztott táruk, jeladás, üzenetváltás)

Indirekt kommunikáció (ferromonok, nyomok, pózok, ...)

Felskálázhatóság

Keresési tér mérete: nagy, heterogén, erősen kölcsönható MAS tanulása gyakorlatilag lehetetlen: egyedi tanulók tipizálása, heterogenitás mérséklése, viselkedési komplexitás mérséklése, viselkedések, **hasznosságok dekomponálása**

(pl. $Q(s, a1, a2, a3, a4) = Q(s, a1, a2) + Q(s, a3, a4)$)

Problémadekompozíció

viselkedések csoportosítása, **rétegezett tanulás**: először elemi viselkedés, majd egyre bonyolultabb, megerősítés-formálás: először egyszerűbb viselkedés előnyben részesítése, majd bonyolultabb, az előbbire alapozva, ...

koordinálási gráf: közös Q értékek részleges dekomponálása (ld. előbb)

Ellenség modellezése

ellenség-e? viselkedéseinek statisztikája?

Tanulás egymásról: mivel más ágensek a környezet részei, azok modelljét meg kell tanulni a környezeti hatások hatásos jóslása érdekében.

Tanulás közösségről: egy szervezet minőségileg több, ágensegyedek kitanulását követően meg kell tanulni a szervezetük speciális métavonásait.

Tanulás változó ágens egyedekkel való kölcsönhatásból: „tanulom a modelledet, de számodra én vagyok a környezeti tényező, így te tanulsz engem. Tanulás révén változom, a változó modelletemet tanulva, változol te is. A változó modelledet tanulva annál inkább én is változom, ...”

Csoporttagok felejtése csoport által: a megtanult információ érvényét veszti, ha az objektuma eltűnik, azonban ennek érzékelése egyáltalán nem triviális ott, ahol az „érezékelésnek” fő módja (bizonytalan) kommunikáció.

Tanulni

Másokról

Másoktól

Mások ellenére

Mások segítségével

Tanulás

tárgyalás**-ból**

tárgyalás **közben**

tárgyalás **végeztével**

Fetétlenül szükséges egy stabil ágensközösség

(„stacionárius szervezet”, különben lehetetlen tanulni).

MAS tanulás céljai

Erősen strukturált kooperatív szervezetben a „jóindulatú” konfliktusok forrásának és kerülésének a kitanulása. A tanulás célja = **a majdani fölösleges kommunikáció mérséklése**, azáltal a hatékonyságnövekedés.

Kooperatív, de laza, nem strukturált szervezetben induktíve tanuló ágensek együttműködése. A tanulás célja = **mások segítségével az egyéni tanulógörbéknek a javítása**.

Alapvetően kooperatív, laza, nem strukturált szervezetben tanuló ágensek. A tanulás célja = **koordinálást megtanulni az együttműködés fokozása érdekében**.

Erősen strukturált hierarchikus és kooperatív szervezetben, ami, mint egy metaágens, veszélyes környezetben ténykedik. A tanulás célja = egyéni, csoportos, **szervezeti hatékonyság növelése** szervezeti célok elérése érdekében.

Alapvetően versengő szervezetben. A tanulás célja = **mások minél jobb kitanulása** a minél hatékony „visszavágás” érdekében.