

Versengő ágensek konfliktusai - játékelmélet

Játékelméleti modellek általában túl egyszerűek a tudás-intenzív feladatokkal küszködő intelligens ágensek világához, de bizonyos alapvető problémákat jó illusztrálnak.

Egy ilyen alapvető “iskolapélda” az un. **börtönrab paradoxon**.

(M. Flood, M. Dresler, 1950, Rand Corp., A. Tucker, Stanford)

Két embert tartoztatnak le, gyanúsítottként, egy fegyveres rablás színhelyéhez közel. Mindkettőnél fegyver van, de fegyverviselési engedélyük nincs. Vizsgálati fogságba, külön cellákba kerülnek, összebeszélni nincs idejük/alkalmuk.

Tudják, hogy ha kitartanak és tagadnak mindent, kaphatnak egy-egy fél évet tiltott fegyverviselésért. Ha önként bevallják a rablást, akkor erre való tekintettel minimális két-két éves büntetésre számíthatnak.

Azonban az államügyész mindkét fogolynak a következő alkut ajánlja: ha ő a másik ellen vall és a bűnösségét tanúsítja, szabad, a másik pedig maximum büntetést, azaz húsz évet kap.

Mit tesznek erre fel a foglyok ?

A helyzet játékelméleti módszerekkel is elemezhető, megalkotva a lehetséges kimenetek mátrixát. Milyen büntetésre számíthatok?

	<i>Én bevallom</i>	<i>Én nem vallom be</i>
<i>Ő bevallja</i>	24 hó (24)	240 (0)
<i>Ő nem vallja be</i>	0 (240)	6 (6)
<i>Átlag</i>	12	123

Az én szemszögömből nézve (ha a másik lépései egyformán valószínűek)

Tehát, ha várható, átlagos nyereségre törekszem, akkor a baloldali oszlopot kell választanom és a hozzá tartozó cselekvést végrehajtanom.

Ezzel azonban a (6 hó, 6 hó) optimumot elkerülve a játék a (24 hó, 24 hó) állapotba kerül, holott elvileg az optimum elérhető volt.

	<i>Én bevallom</i>	<i>Én nem vallom be</i>
<i>Ő bevallja</i>	24 hó (24)	240 (0)
<i>Ő nem vallja be</i>	0 (240)	6 (6)
<i>Átlag</i>	12	123

De így is lehet: nekem a 0 a legjobb (ehhez neki vinnie kell 20 évet, hátha tényleg ilyen buta), tehát bevallom. 0-ra számítva azzal találom magam, hogy ő is lépett, és az elvárt 0 helyett 2 éves helyzetbe kerültem.

Játékelmélet elemei

Játék $\Gamma = (\mathbf{N}, \{\mathbf{S}_i\}_{i \in \mathbf{N}}, \{\mathbf{u}_i\}_{i \in \mathbf{N}})$ (normál alak, extenzív alak)

Stratégiák és kombinációk $\mathbf{S} = \{(s_1, s_2, \dots, s_n) \mid s_1 \in \mathbf{S}_1, s_2 \in \mathbf{S}_2, \dots, s_n \in \mathbf{S}_n\}$

Hasznosság $\forall i \in \mathbf{N}$ -re $u_i : \mathbf{S} \rightarrow \mathfrak{R}$

Szimmetrikus játék $\mathbf{S}_1 = \mathbf{S}_2 = \dots = \mathbf{S}_n$

$$u_i(s_1, s_2, \dots, s_n) = u_{\pi(i)}(s_{\pi(1)}, s_{\pi(2)}, \dots, s_{\pi(n)})$$

Zéró-összegű játék
(„legvéresebb” a verseny) $\forall s \in \mathbf{S}$ -re $\sum_{i \in \mathbf{N}} u_i(s) = 0$

2 ágens esetén

$$\mathbf{S}_1 = \mathbf{S}_2, \forall s \in \mathbf{S} \quad u_1(s_1, s_2) = u_2(s_2, s_1), \quad u_1(s_1, s_2) = -u_2(s_1, s_2)$$

Dominancia Az $i \in N$ játékos $s_i, s_i^* \in S_i$ stratégiái esetén $s_i^* \succ s_i$, ha

$\forall s_{-i} = (s_1, s_2, \dots, s_{i-1}, s_{i+1}, \dots, s_n) \in S_{-i}$ esetén

$u_i(s_i^*, s_{-i}) \geq u_i(s_i, s_{-i})$ teljesül, továbbá

$\exists s'_{-i} \in S_{-i}$ úgy, hogy $u_i(s_i^*, s'_{-i}) > u_i(s_i, s'_{-i})$

Nash-egyensúly (NE) - a legjobb válaszstratégiát játszani

$s^* = (s_1^*, s_2^*, \dots, s_n^*) \in S$ stratégia-kombináció NE, ha

$\forall i \in N$ és $s_i \in S_i$ esetén

$u_i(s_i^*, s_{-i}^*) \geq u_i(s_i, s_{-i}^*)$

Szigorúan dominált stratégiák
iterált eliminációja

	2		
	x	y	z
1			
u	(6, 3)	(1, 5)	(0, 6)
v	(1, 7)	(2, 8)	(2, 6)

Nemek harca (Battle of sexes)

1 \ 2	Vers.	Koop.
Vers.	(1, 1)	(3, 2)
Koop.	(2, 3)	(0, 0)

Gyáva nyúl (Chicken)

1 \ 2	Bátor	Gyáva
Bátor	(0, 0)	(3, 1)
Gyáva	(1, 3)	(2, 2)

Forintpárosítás (Matching pennies)

1 \ 2	Fej	Írás
Fej	(1, -1)	(-1, 1)
Írás	(-1, 1)	(1, -1)

Vezérürü (Leader)

1 \ 2	Megy	Vár
Megy	(0, 0)	(3, 2)
Vár	(2, 3)	(1, 1)

Kevert stratégia: egy valószínűségi eloszlás a tiszta stratégiák felett

$$\forall s_i \in S_i \text{-re } \sigma_i(s_i) \geq 0, \text{ és } \sum_{s_i \in S_i} \sigma_i(s_i) = 1 \text{ (tiszta: } \sigma_i(s_i) = 1)$$

Kevert stratégia
kombinációk

$$\sigma = (\sigma_1, \sigma_2, \dots, \sigma_n) \in Q = Q_1 \times Q_2 \times \dots \times Q_n$$

Hasznosság

$\forall i \in N$ –re $u_i : Q \rightarrow \mathfrak{R}$, ahol $\sigma \in Q$ esetén

$$u_i(q) = \sum_{s=(s_1, s_2, \dots, s_n) \in S} \sigma_1(s_1) \cdot \sigma_2(s_2) \cdot \dots \cdot \sigma_n(s_n) \cdot u_i(s)$$

Kevert Nash-egyensúly (KNE) $\sigma^* = (\sigma_1^*, \sigma_2^*, \dots, \sigma_n^*) \in Q$ kevert NE, ha

$\forall i \in N$ és $\sigma_i \in Q_i$ esetén $u_i(\sigma_i^*, \sigma_{-i}^*) \geq u_i(\sigma_i, \sigma_{-i}^*)$, ahol

$$\sigma_{-i}^* = (\sigma_1^*, \sigma_2^*, \dots, \sigma_{i-1}^*, \sigma_{i+1}^*, \dots, \sigma_n^*)$$

Azaz $\forall i$ -re $\sigma_i^* \in \arg \max_{\sigma_i \in Q_i} u_i(\sigma_1^*, \sigma_2^*, \dots, \sigma_{i-1}^*, \sigma_i, \sigma_{i+1}^*, \dots, \sigma_n^*)$

(Neumann, 1928): Ha egy 2-személyes zéró-összegű játékban teljesül a minimax-feltétel, akkor a játéknak van nyeregpontja (egyensúlya), és a játék értéke a minimax-feltétel két oldalán álló kifejezés közös értéke.

$$\max_{q_1 \in Q_1} \min_{q_2 \in Q_2} u(q_1, q_2) = \min_{q_2 \in Q_2} \max_{q_1 \in Q_1} u(q_1, q_2) = u(q_1^*, q_2^*)$$

(Nash, 1951): Ha egy n-személyes játék eredeti S_i stratégiahalmazai végesek, akkor a keveréssel létrejövő stratégiahalmazok szorzatán definiált játéknak van **legalább egy σ kevert Nash-egyensúlya** (tiszta stratégia is kevert stratégia!).

Kevert σ stratégiák tulajdonsága: A σ^* stratégia-együttes akkor és csak akkor alkot Nash-egyensúlyt, ha minden játékosnak maximális hasznot adnak a kevert stratégia pozitív valószínűséggel játszott tiszta stratégiái, azaz:

$$\begin{aligned} u_i(s_i, \sigma_{-i}^*) &= u_i(s'_i, \sigma_{-i}^*), & s_i, s'_i &\in S_i^+(\sigma^*) & S_i^+(\sigma) &\subseteq S_i \\ u_i(s_i, \sigma_{-i}^*) &\geq u_i(s'_i, \sigma_{-i}^*), & s_i &\in S_i^+(\sigma^*), s'_i &\notin S_i^+(\sigma^*) \end{aligned} \quad (*)$$

Börtönrab Paradoxon

$(D,D) = NE$

D – domináns stratégia

minden (x,y) , (D,D) -t kivéve Pareto optimális

(C,C) maximálja a szociális jólét függvényt

	2		
1		D	C
D	(2, 2)	(4, 1)	
C	(1, 4)	(3, 3)	

MAS – Tárgyalás versengés
közepette

Mechanizmus-tervezés: szociális jóléti függvény

(Szavazás)

(Aukció)

Megegyezés keresése/elérése: **Tárgyalás**

(Érvelés)

Megegyezés elérése speciálisan megtervezett tárgyalásos protokollokkal

Befagyásmentes legyen

Terminálódjon

Biztos sikerre vezessen

Szociális jólétet maximalizáljon (vagy valami mást?)

Pareto-hatékony legyen

(nincs olyan más megegyezés, ami legalább egy ágens hasznosságát megnövelne úgy, hogy más ágensek hasznosságát nem csökkenti)

Individuális racionális legyen

(a protokoll követése ágensek önérdelke (benne lenni jobb))

Stabil

(a protokoll meghatározott viselkedésre készíti az ágenseket)

Egyszerű

(optimális stratégia komplexitása legyen alacsony, kiszámítható)

Elosztott

(egy egyedi ágens bukása a protokollt ne buktathassa

(single point of failure))

...

Tárgyalás komponensei

Javaslat halmaz: ágensek által megtehető javaslatok

Protokoll: megadja az adott helyzetben legális javaslatok meghatározását

Ágensek stratégiái

Megegyezés elérésének szabályai (tárgyalás terminálási szabályai)

Tárgyalás fajtái

Egy tételű

Többtételű (javaslatok mennyisége exponenciális, javaslatok összevetése nehéz)

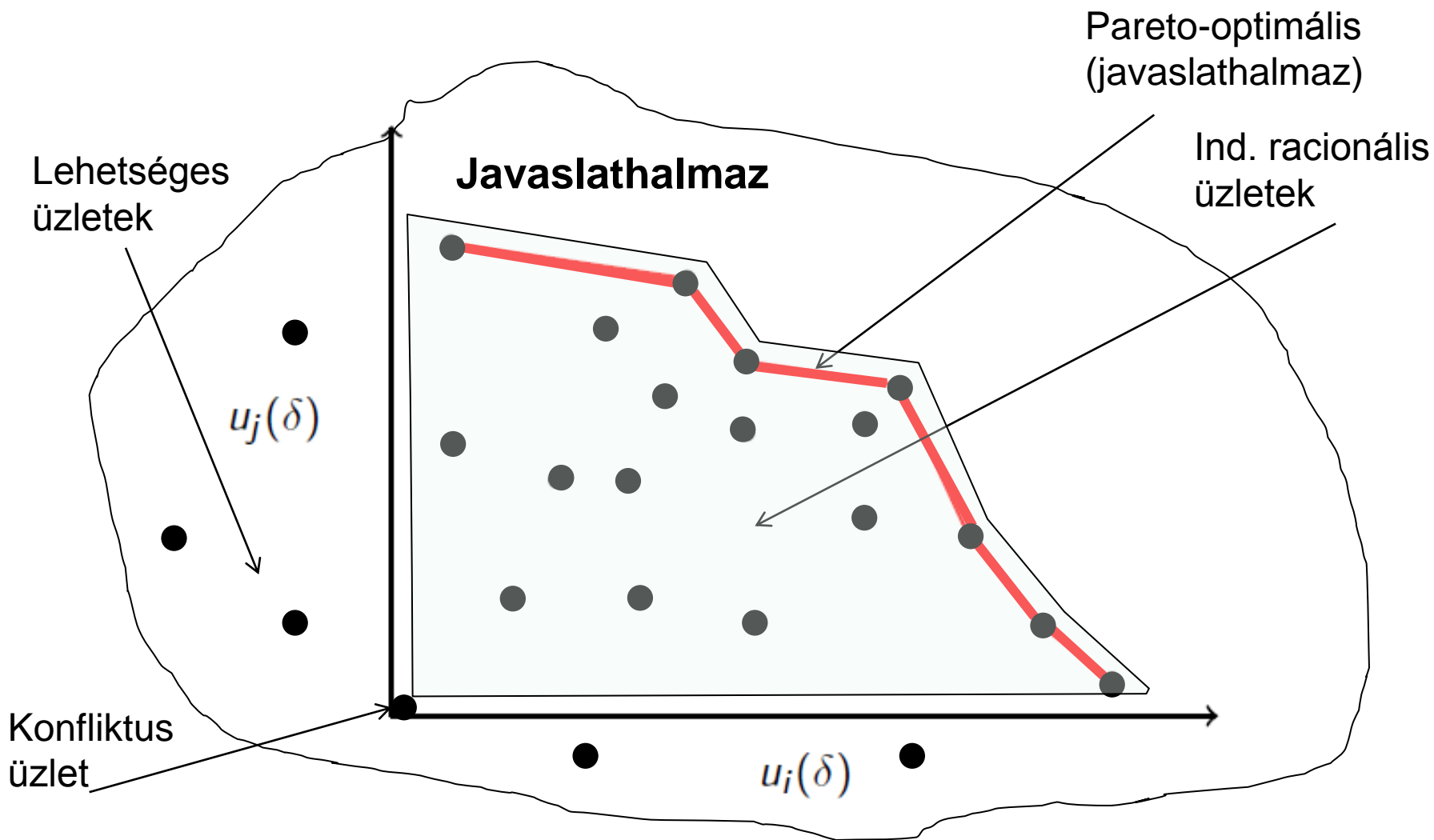
1-1 tárgyalás (tip. többágenses rendszerekben)

N-1

N-M

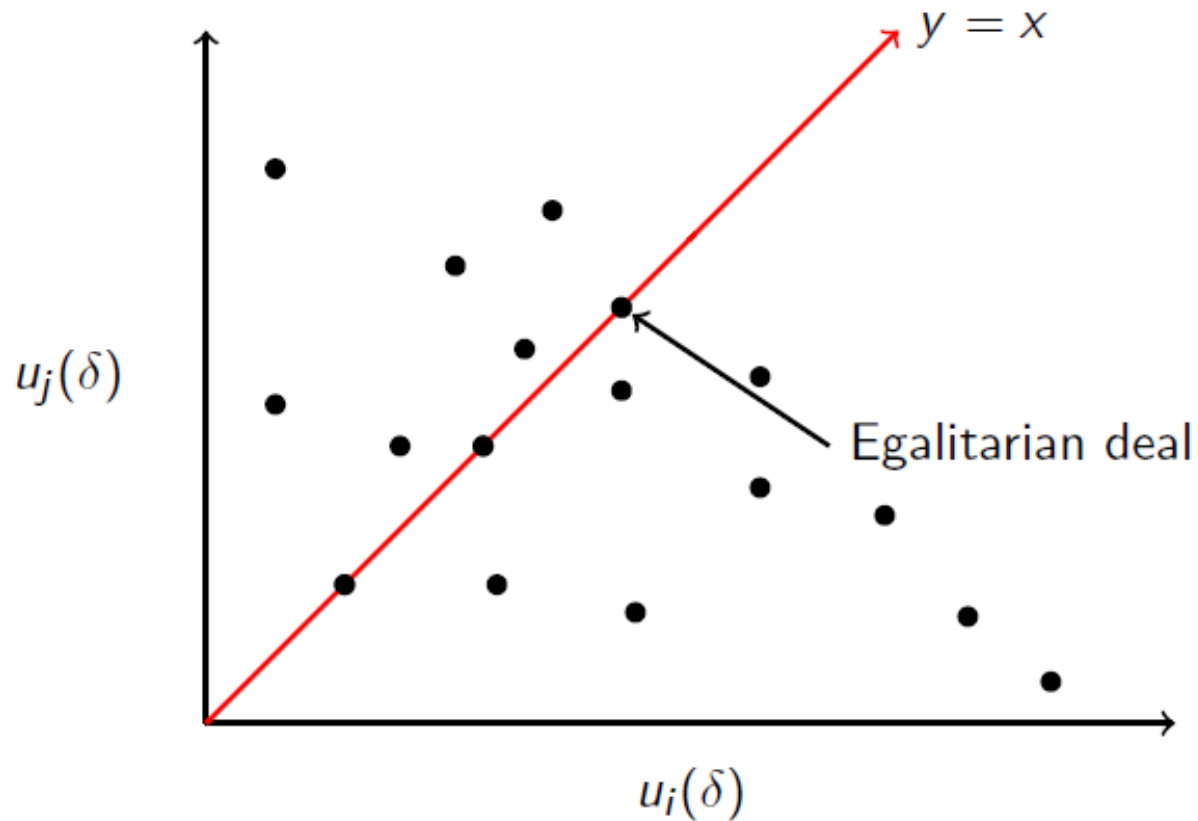
Stratégiai megoldás lehetőségei

Pareto-optimális, de melyik legyen?



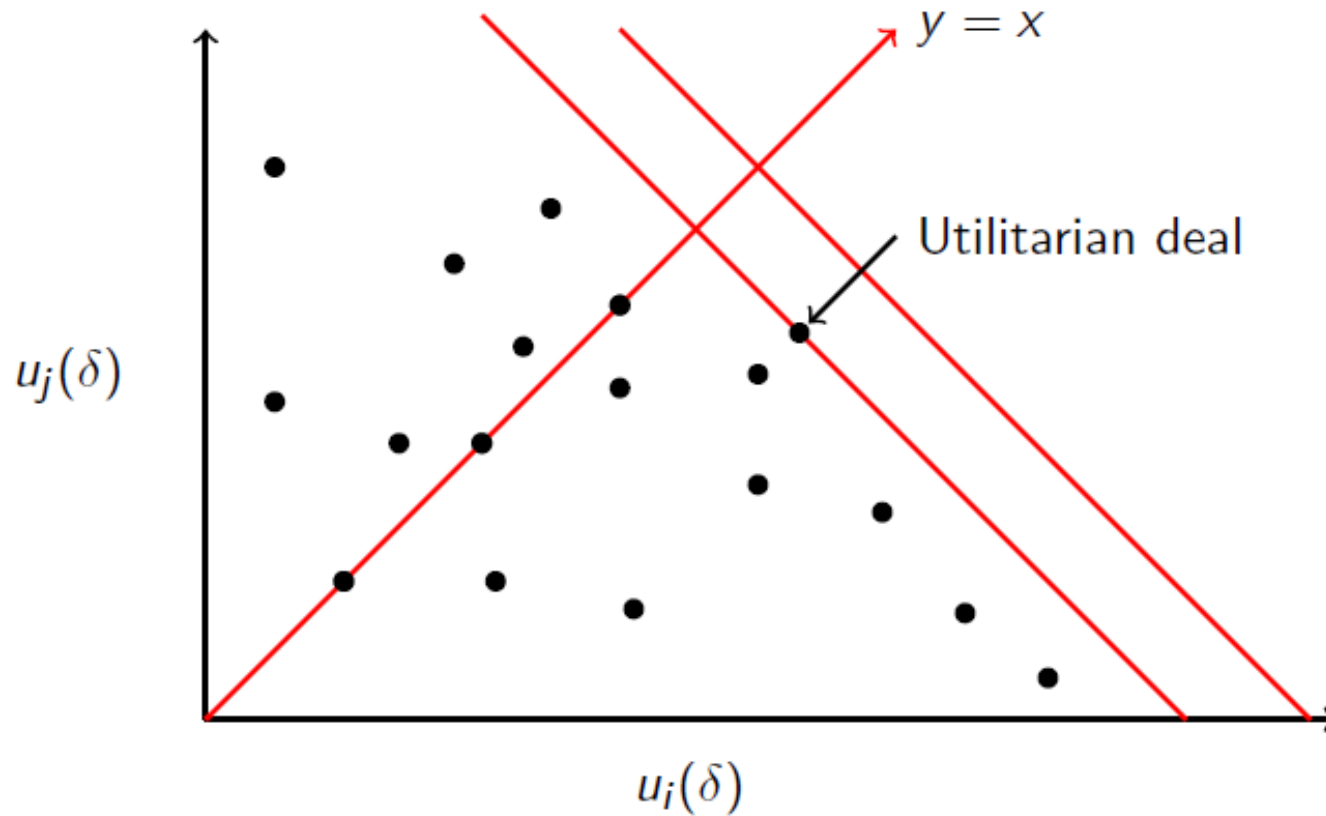
Egalitárianus

$$\delta = \operatorname{argmax}_{\delta' \in E} \sum_i u_i(\delta'), \quad E = \left\{ \delta \mid \forall_{i,j} u_i(\delta) = u_j(\delta) \right\}$$



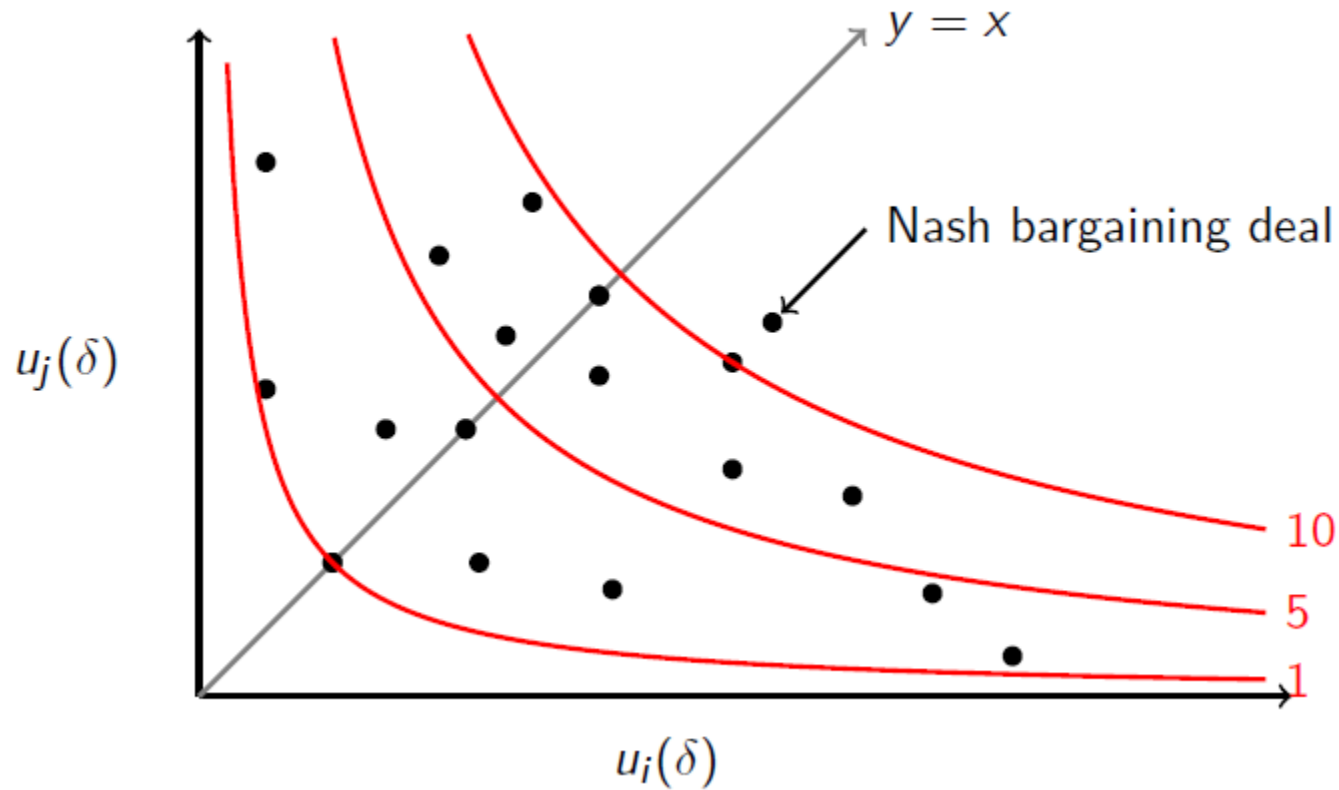
Utilitárianus

$$\delta = \arg \max \sum_i u_i(\delta)$$



Nash-alku

$$\delta = \arg \max_{\delta'} \prod_i u_i(\delta')$$



Monoton Engedmény Protokoll

Fordulók

Megegyezés, ha 1. ágens olyan δ_1 üzletet és 2. ágens olyan δ_2 üzletet javasol, hogy vagy $u_1(\delta_2) \geq u_1(\delta_1)$, vagy $u_2(\delta_1) \geq u_2(\delta_2)$, azaz „a másik javaslata legalább ilyen jó, mint az enyém”

Megválasztás: üzlet maximális hasznossággal

Ha legális üzletre nincs lehetőség: konfliktus-üzlet

Monoton Engedmény Protokoll – Zeuthen (1930) stratégia

Milyen legyen egy jó stratégia?

Mit kellene javasolni az első fordulóban?

Egy-egy fordulóban kinek kellene engedményeket tennie?

Ha egy ágens engedményt tesz, mennyit engedjen?

Monoton Engedmény Protokoll – Zeuthen (1930) stratégia

Milyen lenne egy jó stratégia?

Mit kellene javasolni az első fordulóban?

Mindenki az ő leginkább preferrált üzletét.

Egy-egy fordulóban kinek kellene engedményeket tennie?

Annak az ágensnek, aki legkevésbé hajlandó kockáztatni a konfliktust.

Ha egy ágens engedményt tesz, mennyit engedjen?

Épp annyit, hogy megszűnjen ilyennek lenni
(azaz a konfliktust legkevésbé kockáztatni hajlandó)

Ha kevesebb: újra kell engedményt tennie – nem hatékony

Ha több – hasznosságot pazarol.

Az i ágens hajlandósága konfliktust kockáztatni magas, ha:

az aktuális javaslata és a konfliktus hasznosság különbsége alacsony
(a konfliktussal nem veszít sokat)

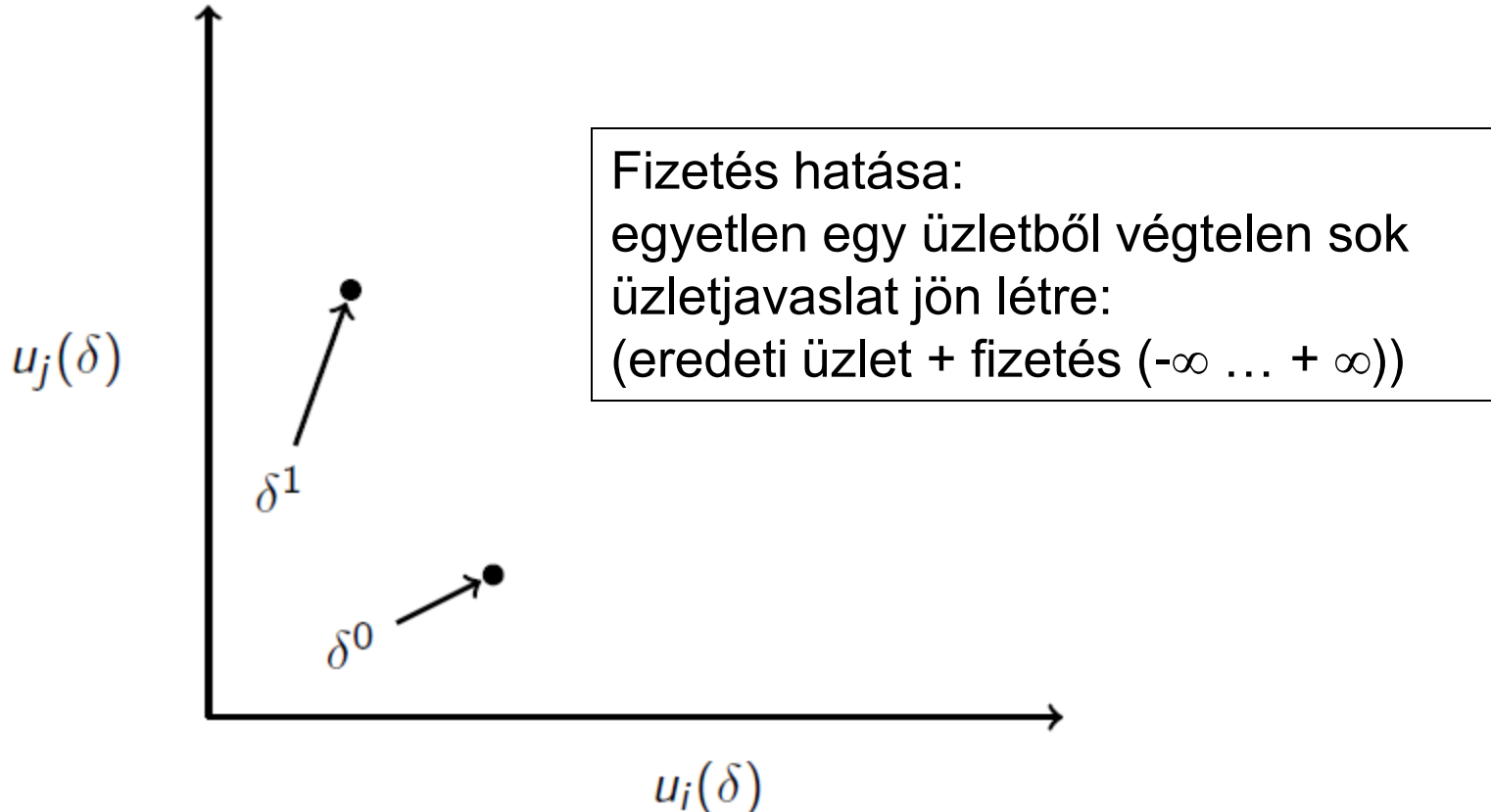
az aktuális javaslata és az ellenfél javaslata hasznosság különbsége magas
(az engedménnyel sokat veszít)

Feladathozzárendelési probléma - Task Allocation Problem

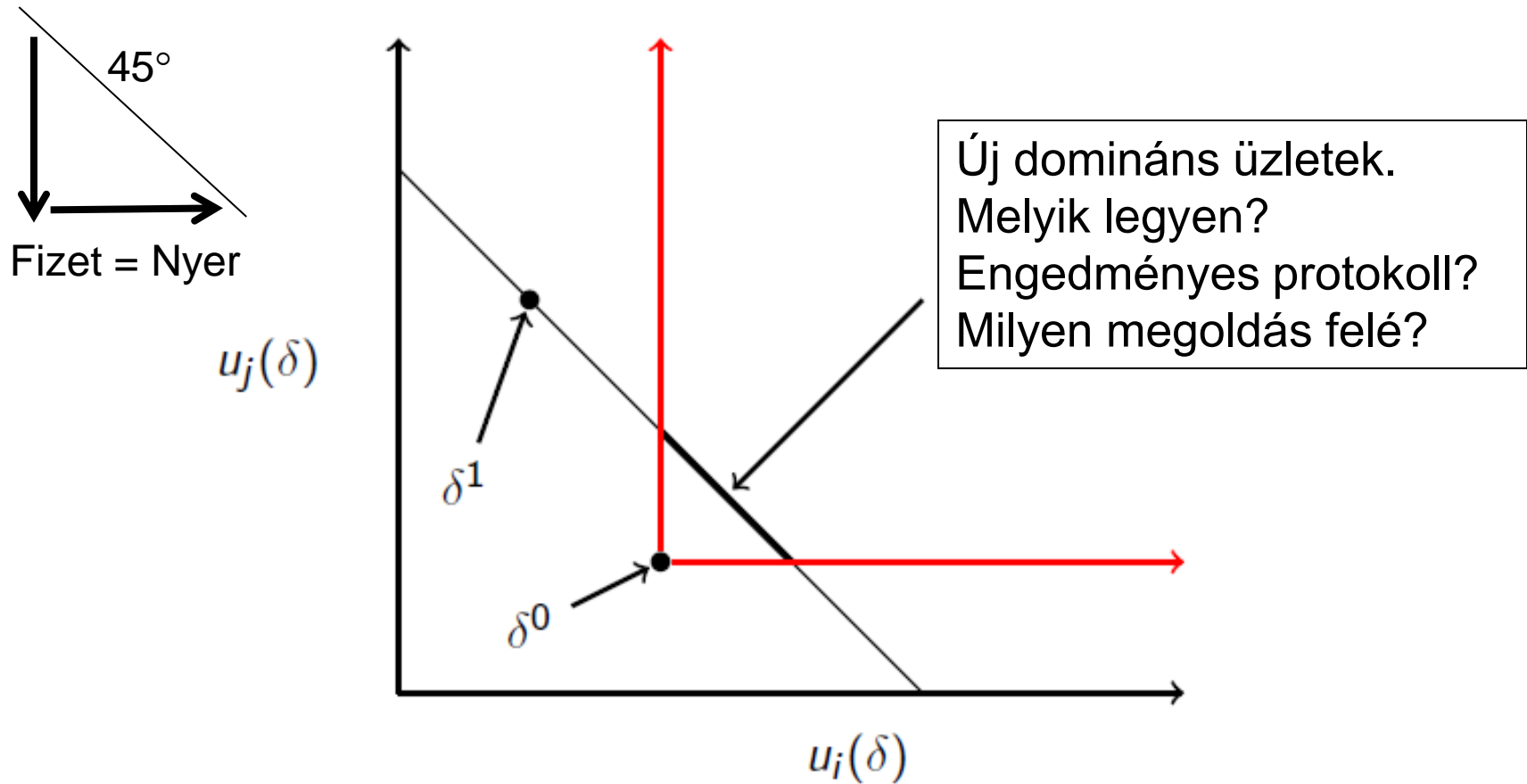
Mi van, ha nemigen van egy jó üzlet,
amit lehetne javasolni és elfogadásra találni?

Fizetések:

(pénzbeli) fizetések bevezetése – növeli a lehetséges üzletek választékát



Feladathozzárendelési probléma - Task Allocation Problem



MAS – Tanulás

(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), 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, nagyon sok konkrét algoritmus (tudásreprezentáció függvényében)

Tények/kihívások

Hipotézistér

Többágenses környezetben tanulás szintén egy keresés a hipotézisek terében, de most a tér (a többi ágens miatt) szokotlanul nagy.

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 mások által meghatározott kontextusban. Ilyen tanulás 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.

Tények/kihívások

Megerősítés

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

Stacionaritás

Több, parallel és konkurens módon tanuló ágensek 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.

Érdemhozzárendelés

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

Tények/kihívások

Nash-egyensúly

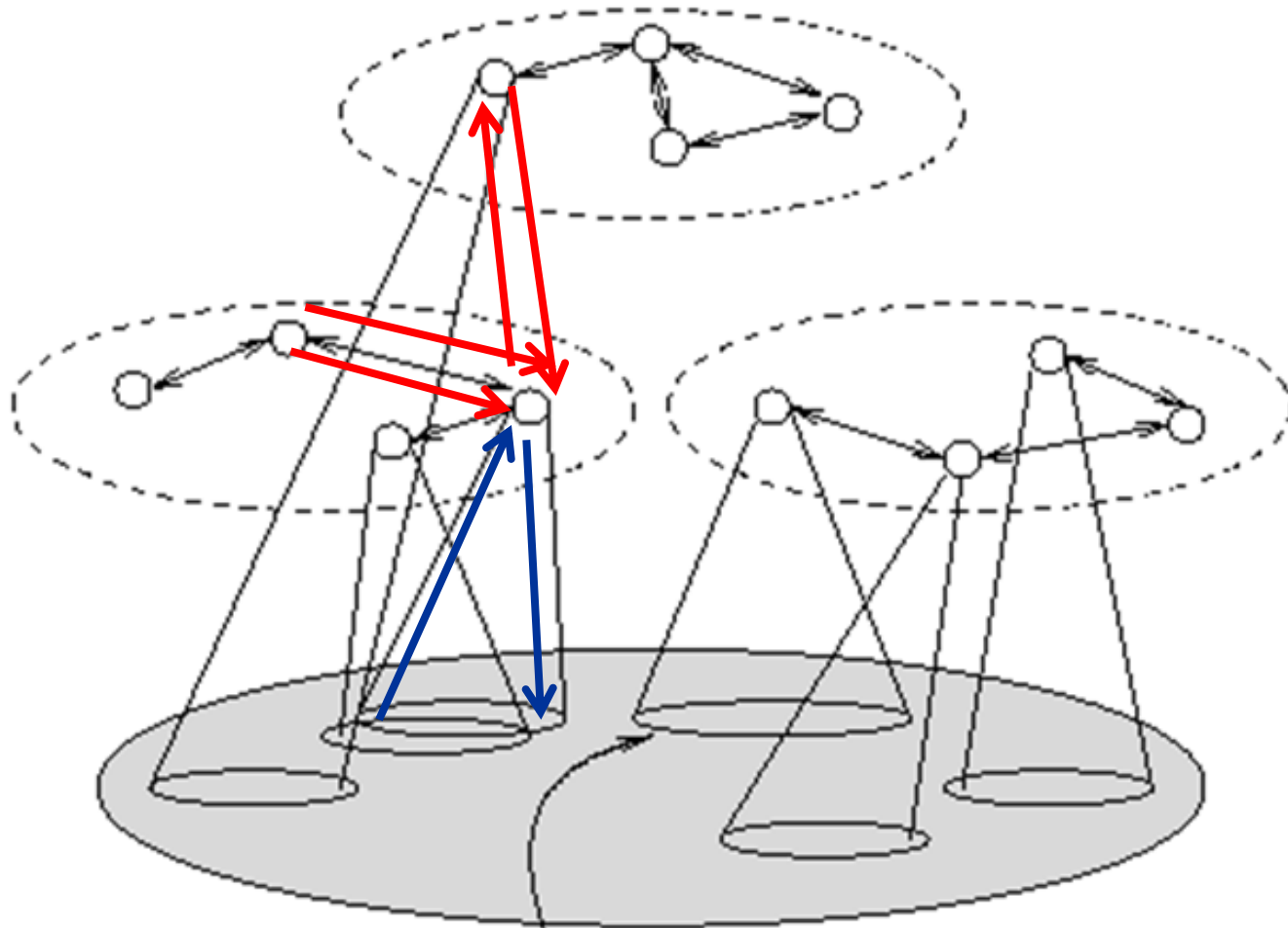
Stacionaritás nélkül az optimálisan megtanulthoz való konvergencia (példák növekedésével) nem biztosítható. Nash-egyensúly egy alternatíva, de láttuk, hogy ez nemigen az optimális szociális jólét alternatívája. És mi van, ha több Nash-egyensúly van?

Konvergencia

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

Kommunikáció

Kommunikáció egy információforrás. Információ 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?



KEY



organisational relationship



interaction



agent

sphere of influence

Environment

Együtműködés tanulása

Feladat tanulása

Kihívások

tanulás egymásról: mivel más ágensek a környezet részei, azok modelljét meg kell tanulni a környezeti hatások hatásos jóslása érdekében (hogyan fognak válaszolni, mit lépnek, mi a szándékuk, ...)

tanulás közösségről és környezetről: egy szervezet minőségileg több, mint az egyedek összege, az ágens egyedek kitanulását követően meg kell tanulni a szervezetük speciális métavonásait

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

tanulás torzított szociális struktúrákban/ struktúrák által:

ld. korábban: elfogultság, ..., torzítás – objektív, v. szubjektív

régi 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 pl. a (bizonytalan) kommunikáció

Kihívások

Tanulni

másokról

mert ez kell a konfliktusok, a fölösleges kommunikáció mérsékléséhez, a tárgyalásokhoz szükséges erőforrások mérsékléséhez, az ágens tárgyalási készség javulásához, az együttműködés optimalizálásához, ...

másoktól

mert érdemes, hiszen ők is intelligens, racionális egyedek, talán a környezetet másképpen, más szempontból jobban érzékelik, a tudásuk jól kiegészítheti a miénket, ...

mások ellenére

mert az együttműködési igény, kényszer nem mindig van jelen, a tanulásunk pedig jobb versenyhelyzetbe helyez minket, és így másokat rosszabban, ...

mások segítségével

mert érdekeltek lehetnek abban, hogy a csapat minden tagja hatékonyabb és ügyesebb legyen, így a közös megoldás is jobb lesz, ...

Kihívások

Tanulás és tárgyalás

Tanulás	tárgyalás	-ból
	tárgyalás	közben
	tárgyalás	végeztével

- Tanulás tárgyalásból, feltétlenül szükséges egy **stabil ágensközösség**
- ugyanazok az ágensek több tárgyalásban is részt vesznek
 - bizonyos ágenstípusok halmaza állandó (különben lehetetlen tanulni)

Ágens **mit tartson meg** a tárgyalás végeztével

a beérkező információ permanens megtartása az eredeti formájában

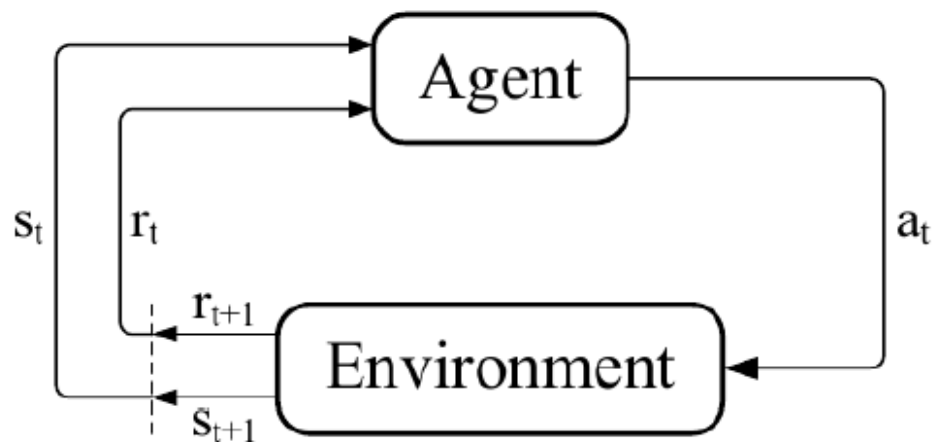
a beérkező információ ideiglenes tárolása, kompilálása és beintegrálása

a beérkező információ törlése, miután az ágens azt felhasználta

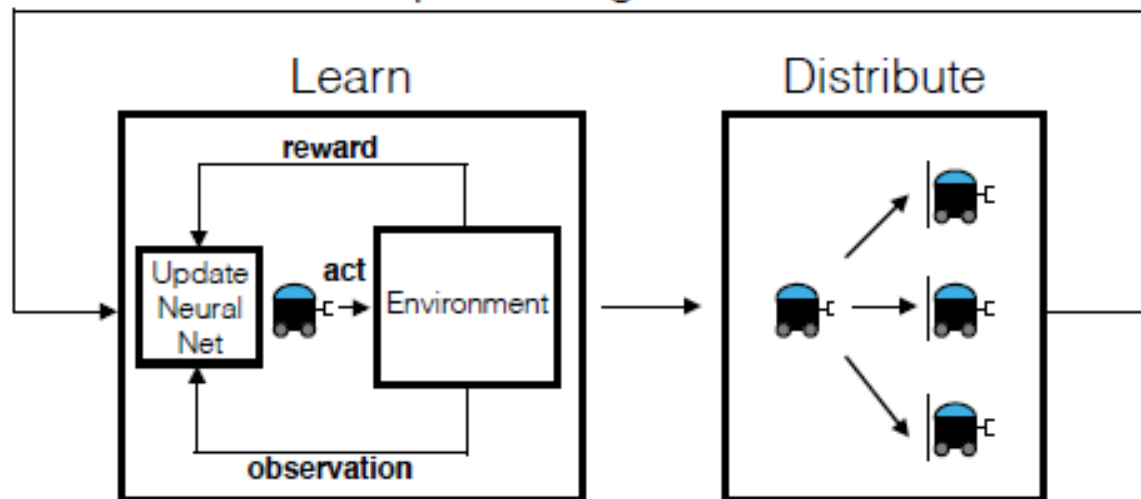
(processzállás és komplexitás terhe)

MARL – Multi Agent Reinforcement Learning

Többágenses megerősítéses tanulás

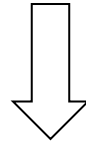
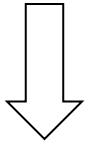


Improved Agent Policies



Egy ágenstől több ágensig

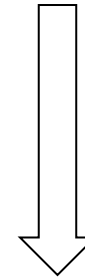
$$Q_{k+1}(s_k, a_k) = Q_k(s_k, a_k) + \alpha_k \left[r_{k+1} + \gamma \max_{a'} Q_k(s_{k+1}, a') - Q_k(s_k, a_k) \right]$$



Mások cselekvései is

$$Q_{k+1}(s_k, \mathbf{a}_k) = Q_k(s_k, \mathbf{a}_k) + \alpha_k \left[r_{k+1} + \gamma \max_{\mathbf{a}'} Q_k(s_{k+1}, \mathbf{a}') - Q_k(s_k, \mathbf{a}_k) \right]$$

Valami más, ami a cselekvések baráti,
vagy adverz jellegére utal és eszerint
számítja ki a várható jövőt.



$$Q_{k+1}(s_k, \mathbf{a}_k) = Q_k(s_k, \mathbf{a}_k) + \alpha_k \left[r_{k+1} + \gamma XYZ - Q_k(s_k, \mathbf{a}_k) \right]$$

1 ágens \implies 2 ágens \implies N ágens

Játék lehet **modell-alapú**: ágens először megtanulja az ellenfél stratégiáját, majd talál rá a legjobb választ.

Lehet **model-nélküli** is, amikor ágens az ellenfélre jó választ adó stratégiát tanulja meg anélkül, hogy az ellenfél stratégiáját explicite kitanulná.

Egy ágens megerősítését generáló függvényt ρ_i .

- **teljesen kooperatív** ágens rendszerek $\rho_1 = \dots = \rho_n$

- **teljesen versengő** ágens rendszerek (két ágens, zérus összegű)

és $\rho_1 + \rho_2 + \dots + \rho_n = 0$, több ágens esetén

$$\rho_1 = -\rho_2$$

- **vegyes** ágens rendszerek (általános összegű, ahol semmilyen feltétel nem adható).

$$\rho_1 + \rho_2 + \dots + \rho_n > 0$$

Minden zérus-összegű mátrix játéknak van NE-ja tiszta stratégiákban.

(Neumann)

Minden általános összegű mátrix játéknak van NE-ja. (Nash)

Minden teljesen versengő sztochasztikus játéknak van NE-ja (Shapley)

Minden általános (vegyes) sztochasztikus játéknak van NE-ja. (Fink)

- MAS - Szervezet tanulása ...
- Tervkészítés ...
 - Emociók ...
 - Ember-gép közös nyelv ...
 - Mixed initiative team ...