

Kooperáció és intelligencia

Tanulás többágenses szervezetekben/1

Erősen strukturált kooperatív szervezet – (együtt tervezünk)

Modell: a közösség - egyszerű ágensekben gazdag közösség,
ahol egy-egy ágens kicsi, specializált, és tudásalapú

Single Function Agents – SiFA

egy SiFA ágens architektúrája:

funkció	mire képes, pl. <i>Tanács, Elemzés, Kritika, Értékelés, Becslés, Tervkészítés, Szelekció, Javaslat</i>
cél	mire van közvetlen hatással (valamilyen objektum, vagy az előállítása)
szempont	<i>funkció alkalmazása a célra</i> , de milyen perspektívából: költség, gyárthatóság, teherbírás, pl. anyagok szelektálása teherbírás szempontjából

egy SiFA ágens tudása:

tervezői tudás (a tényleges csapatbeli feladata, l. architektúra)

konfliktusfeloldó tudás - konfliktusok detektálása, osztályozása, feloldása

kommunikációs tudás - beszédaktusok szerkesztése és interpretálása

lokális tár

Az ágensközösség koordinálása

Globális tár körül szerveződik:

Globális tár (BB) – „design board” =

- tervezés specifikációja,
- a megoldás aktuális állása,
- javaslatok és döntések.

Opportunista ütemezés – egy ágens működésbe lép, ha (látja, hogy):

- a feladatának előfeltételei teljesülnek,
- az eredményeire igény jelentkezett.

Kommunikáció – **KQML ACL** (beszédaktusok, alkusz nélkül)

felkérés egy alternatíva elfogadására

jelentés, hogy egy kényszer (nem) teljesül

kérdés a preferenciákra vonatkozólag

javaslat

javaslat (javasolt paraméterérték) **elutasítása**

....



KQML

Ágens típusok

Szelektor	alternatívák között választ, preferenciák és kényszerek függvényében
Becslő	hiányos információ miatt paraméterértékeket becsül
Értékelő	tervezési célok és paraméterek alapján paramétereket értékel egy adott szempont alapján, pl. „jó”
Kritikus	potenciális problémákat, gyenge döntéseket azonosít, <i>milyen kényszerek sérültek</i>
Dicsérő	pozitív megjegyzéseket tesz, <i>miért fontos?</i>
Javasló	kritikák és kontextusok alapján alternatív megoldásokat, ill. cselekvéseket ajánl

A közös feladat: **rúgótervezés**,

több anyagra vonatkozó tervezési kísérletek - hőmérséklet-tartomány, húzószilárdság, ellenállás/ ütközés, ellenállás/ sokk, megengedett munka stressz, rugalmassági modulus, élettartam, tűrési határok, „végtelen sokszor” alkalmazható stressz, keménység, elektromos vezetés, mágneses tulajdonságok, korrózióvédelem, méret, alak, átmérő, költség, ...

Konfliktusok: konfliktusok ágensek között (eltérő funkciókból adódóan), pl.

Becslő - Szelektor/Javasló

- nem becsülhető a választásra vonatkozó kevés információ miatt

Értékelő - Becslő

- az értékelés nem megbízható/ nem pontos a szegényes becslés miatt

Szelektor – Szelektor

- eltérő preferenciák eltérő választáshoz vezetnek

Szelektor – Becslő/Értékelő

- bizonyos kritériumok hiányában a választás nem végezhető el
(pl. becsült érték mellett nincs annak statisztikai megbízhatósága
– várható érték, szórás, torzítás)

Kritikus – Becslő/Értékelő

- alacsony kvalitású munkát nem fogad el, pl. a tervezés későbbi fázisai miatt

Kritikus – Szelektor

- a kritika alapját (sérült kényszer) szelektornak át kellene vennie

....

Konfliktus típusok

Kevesebb információ van egy ágens kimenetében,

mint amire egy másik ágensnek szüksége lenne,
minden olyan ágens esetében, amely mások kimenetére alapoz:
pl. Értékelő-Becslő, Szelektor-Becslő, Kritikus-Szelektor, ...

Információ kvalitása nem elegendő

az átvevő ágens szempontjai szerint:
pl. Értékelő-Becslő, Szelektor-Értékelő, Kritikus-Becslő, Kritikus-Értékelő

Szegényes processzási modell(je az ágensnek), főleg Kritikus, Javasló, Szelektor

Eltérő preferenciák,

ha több ágens verseng, hogy ua. a mennyiségnek az értékét rögzítse, főleg
Szelektorok

Sérült kényszerek, főleg Kritikusok és Szelektorok, Javaslók

....

Konfliktusok feloldása

Információ kibővítése (→ **kommunikáció**)

Megegyezés bizonyos értékek kérdésében (→ **kommunikáció**)

Információ kvalitásának emelése

Átállás jobb processzási módszerekre

....

Konfliktusok érzékelés

Globális megoldás tárban

Szelektor konfliktusok

Szelektor nem tudja a preferált értékét megválasztani más ágens választása miatt → **tárgyalás** inicializálása

Kritikus konfliktusok

Kritikus ellenőrzi a számon tartott kényszereket, sérülés esetén kinyomozza, melyik paraméter (ágens) a ludas → **tárgyalás** inicializálása

Dicsérő szerepe:

állásfoglalása nyomást jelenthet konfliktusfeloldás során,
hogy ki inkább vonja vissza a javaslatát,

...

SIFA tanulási lépések – mások megtanulása

Egy A ágens egy B ágens koncepcionális leírását tanulja:

- 1. fázis:** Egyedi kölcsönhatás B-vel $\Rightarrow$ esetképzés (tanulás tárgyalásból)
esetindex = B döntése mögött húzódó tervezői igények
esettartalom = B konkrét válasza minden eset = egy tanulópélda
- 2. fázis:** Esetek integrálása a B eddig megtanult modelljébe
(tanulás tárgyalás végeztével), leképzés:
különböző tervezési feltételek $\rightarrow$ B opciói/preferenciái

Nehézségi fokok (a kontextus függés és rejtett stratégiák miatt)

(nyílt) dicsérő $\rightarrow$ kritikus $\rightarrow$ szelektor (burkolt)

Tervezési feltételek ugyanazok, de más-más ágenstípusról mást tanul

Szelektor

csak **részleges** preferencialista: a preferenciákat tárgyalás közben tárják fel, javaslat elfogadás után, **a többi preferencia rejtett**, a feltárt preferenciák tárgyalásfüggők, ...

Kritikusok

nem elfogadható javaslatok által képviselt értéktartományok, rögzített modell, ...

Dicsérők

legkönnyebb, amilyen értékeket dicsérnek, rögzített modell

Tanulás hatása, célja: tárgyalás intenzitásának mérséklése későbbiekben

MAS hatás: tanulás során mérséklődik a kommunikáció (kevesebb konfliktus) egyre nehezebben tanul, mert kevesebb a példa

- Kritikusok jelenléte:

Szelektor tanulja a Kritikusokat, abból adódóan kevésbé tárja fel a preferenciáit a további kölcsönhatásokban, **Szelektor-ról való tanulás nem lesz pontos**

Jobb eredmény: Szelektor tanulás Kritikusok nélkül, majd finomítás kritikusok jelenlétében.

Mozgócélp probléma:

- A tanul B-ről, B tanul A-ról, meg másokról
- A tanul implicit módon C-ről, B által, abból, amit a B tanult a C-ről ha C javaslatokat tesz a B Szelektornak, akkor az A ágens, amely a B modelljét tanulja, B válaszaként feldolgozza a B és a C következtetéseit

MAS hatás: Tanulással a kölcsönhatások mértéke 2/3-ra csökkent

Koaktív tanulás

coacting (pszichológia) – egyedi szereplés, az ugyanilyen feladattal küszködő mások közepette

tanulók aktív kölcsönhatása, de minden tanuló önmagáért tanul, nem csoporteredményről van itt szó

tanulás **eltérő** példahalmazokon

kommunikáció → példák, közlés a hallgató érdekében
megítélés firtatása
→ példák megítélése

tanulás **azonos** példahalmazon

kommunikáció → példák megítélése

fontos: **kritikus példák cseréje**

kritikus példák minősítése – majdnem, mint a felügyelt tanulás
(egymás részére)

Kooperáció és intelligencia, BME-MIT

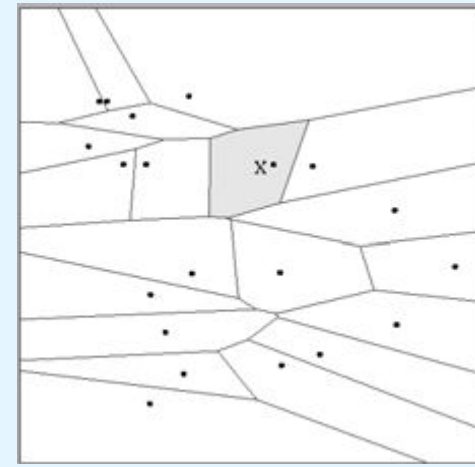
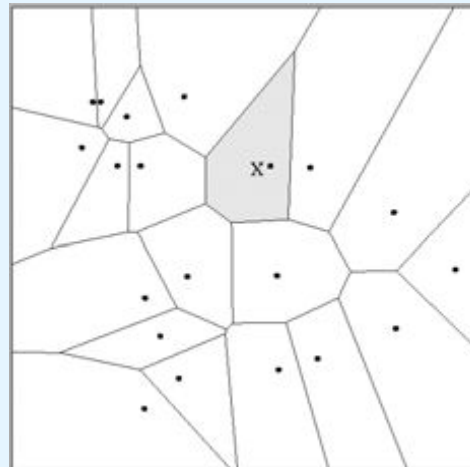
Kísérlet: osztályozástanulás - IBL – Instance-Based Learning

(memory based, 1-Nearest Neighbor, k-Nearest Neighbor, Kernel Regression, ...)
egyszerű fogalom reprezentáció, alacsony inkrementális tanulási költség,
kis tárigény, fogalompéldák felkérésre, folytonos fv. tanulása,
nemlineárisan szeparálható kategóriák tanulása, ...

egy új eset osztályozása (jóslása) =
tárolt, az új esetre leginkább hasonlító esetek
osztályozása alapján

- távolság definíciója?
- hány szomszéd?
- súlyozás (opcionális)?
- illesztés a lokális pontokra?

Távolság:



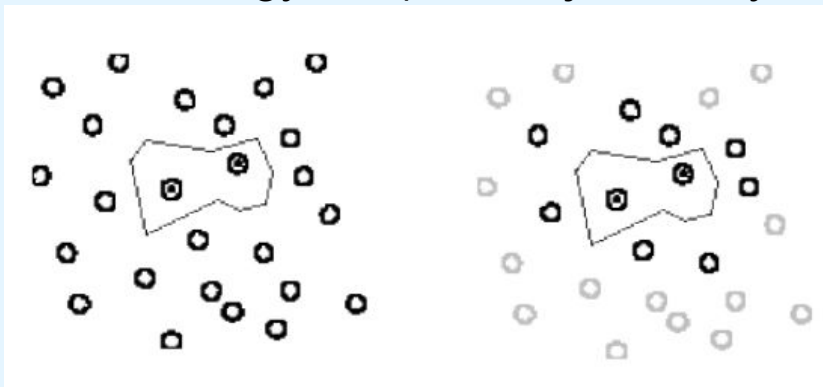
Kísérlet: osztályozástanulás prototípus tanulással

IBL – Instance-Based Learning

IB1 – minden megfigyelt eset tárolása (zajvédett, memória!)

IB2 – csak a rosszul osztályozott esetek tárolása (zajérzékeny, memória ok)

IB3 – minden rosszul osztályozott eset tárolása, korrekt/nem korrekt osztályozók adminisztrálása (szignifikancia teszt), rossz jóslók elhagyása (nem zajérzékeny, memória ok)



Ágens:

igazmondó

jóhiszemű

kooperatív

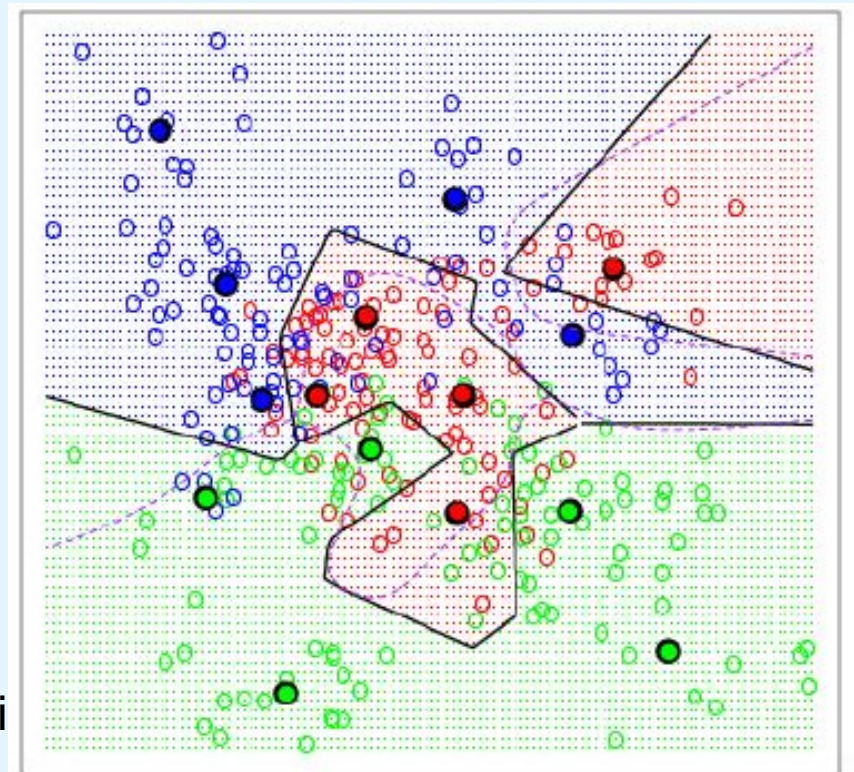
jóindulatú

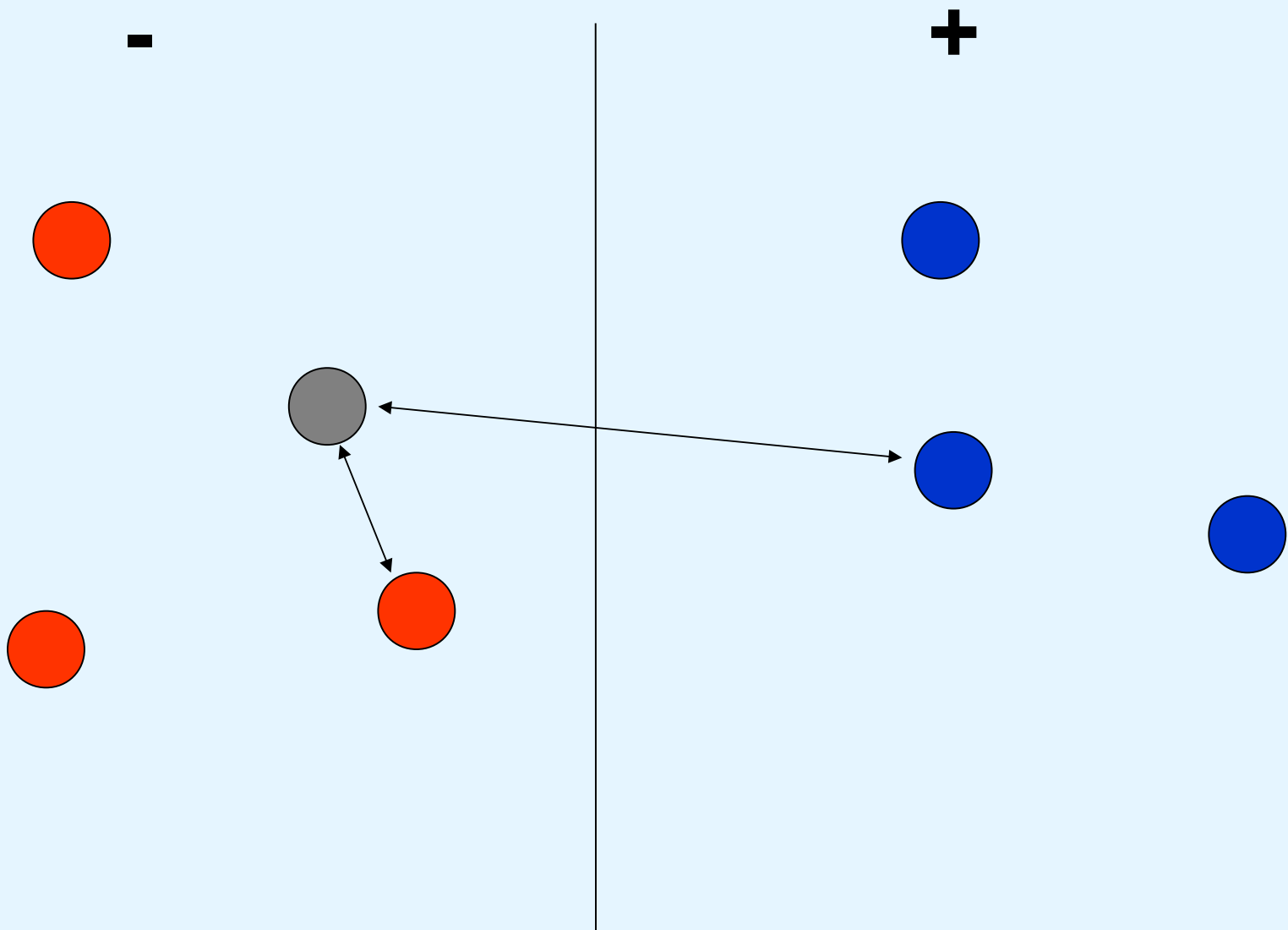
hasonló (azonos) tudáskezelés

hasonló (azonos) probléma

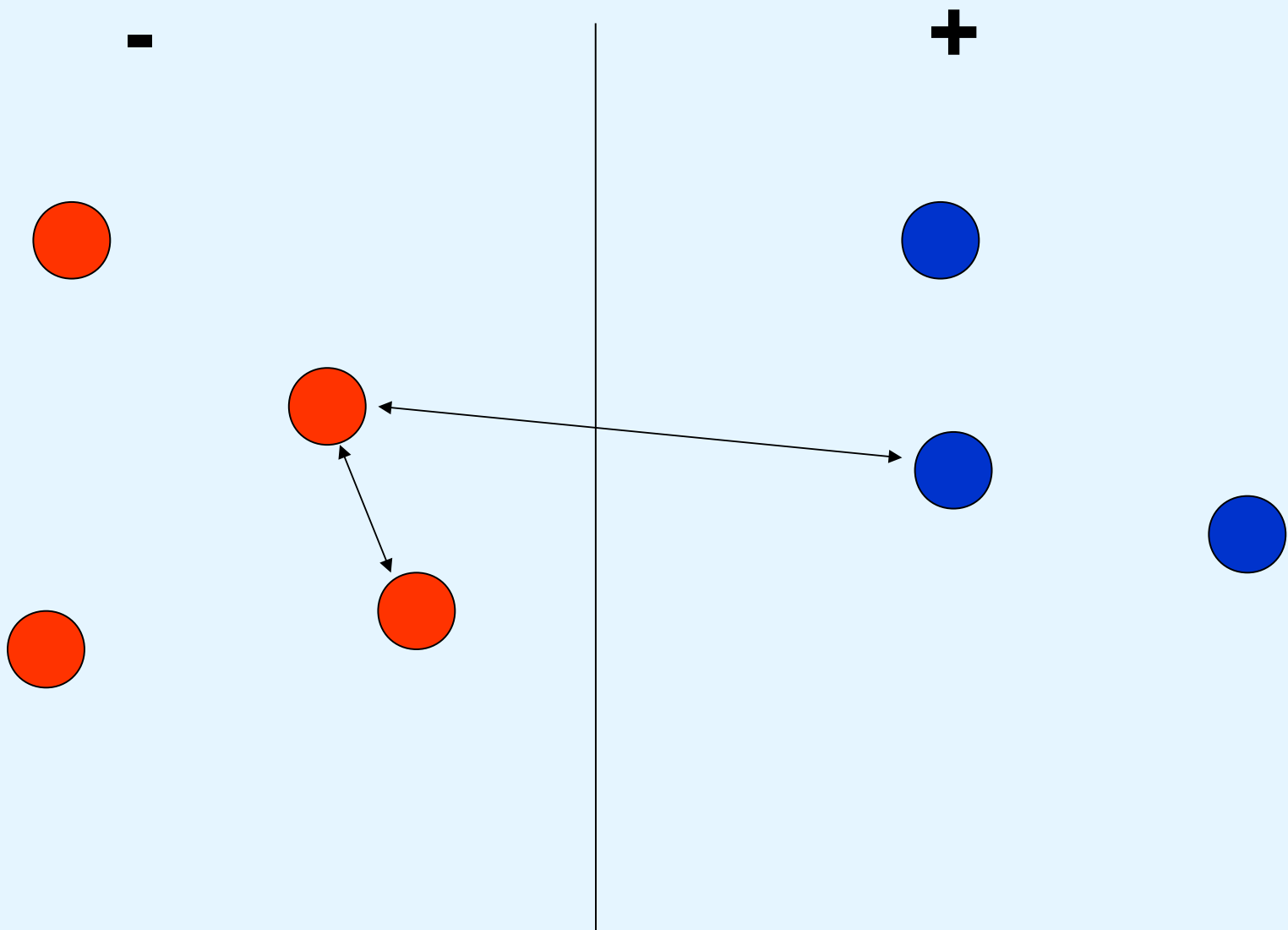
megoldó készség

Kooperáció és i

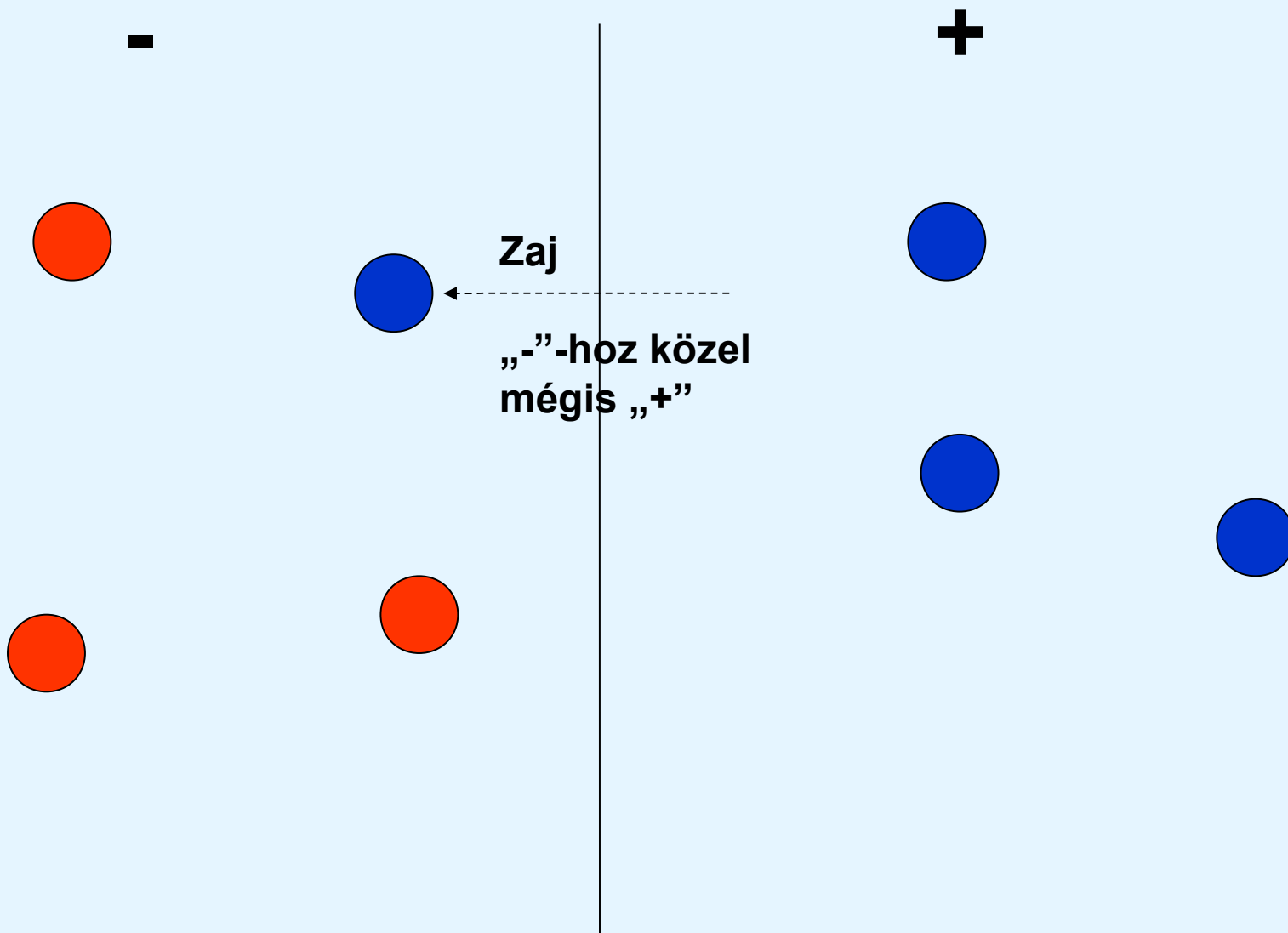


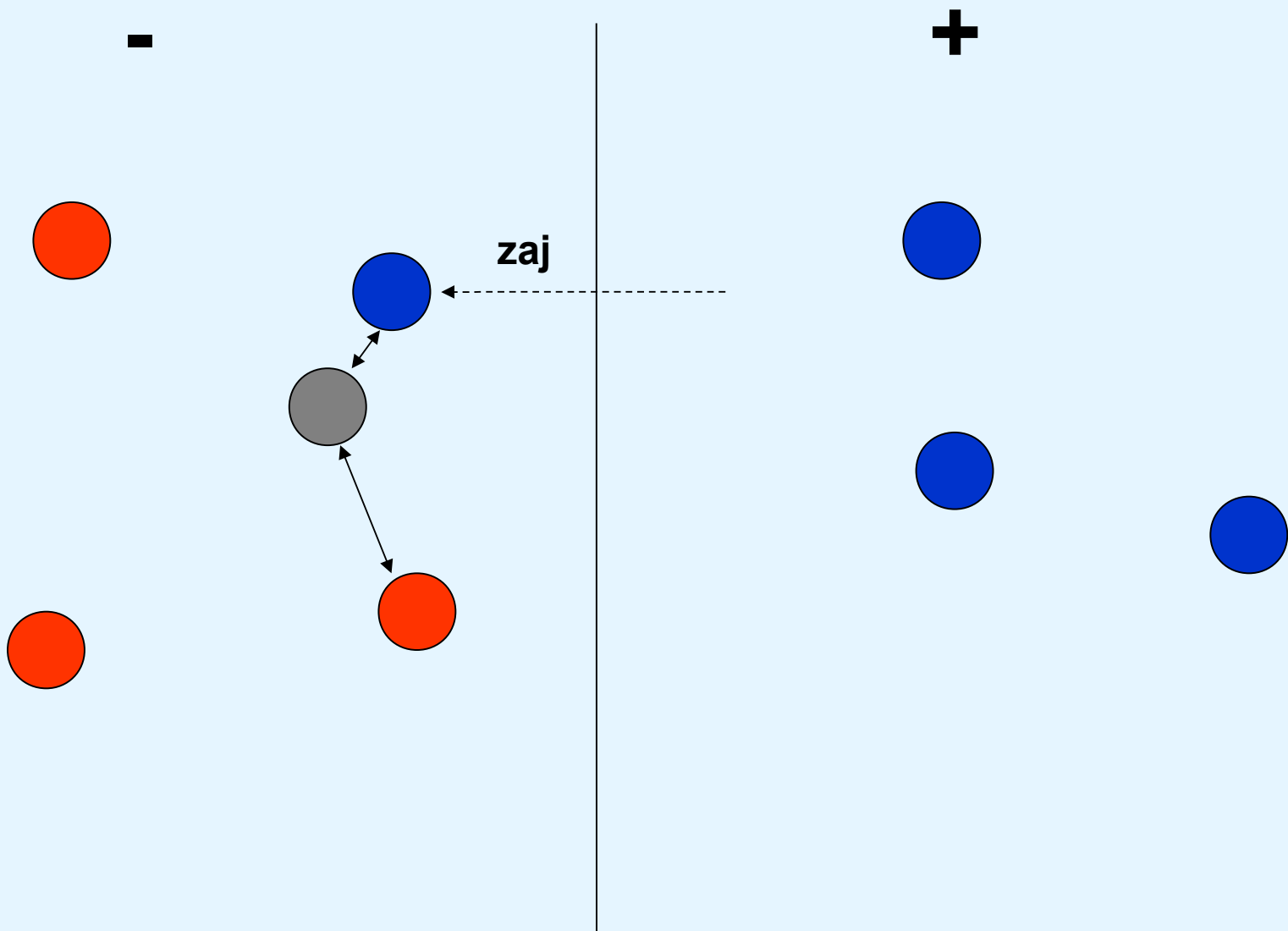


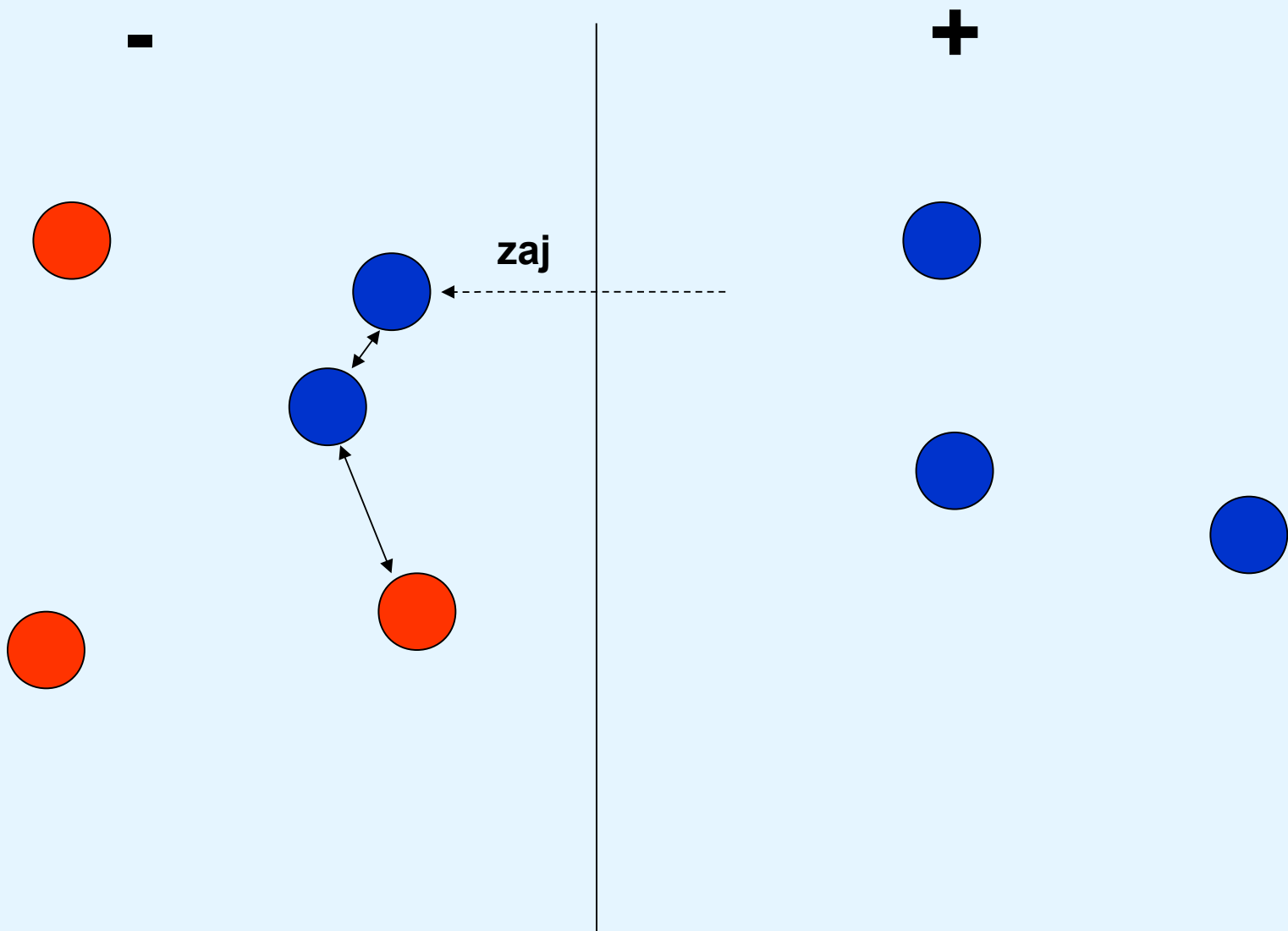
Kooperáció és intelligencia, BME-MIT



Kooperáció és intelligencia, BME-MIT







Kooperáció és intelligencia, BME-MIT

Ugyanaz a tanítóhalmaz, eltérő magántanulási stratégiák

IB1	-S	-S	+S	+S
IB2	-S	+NS	-S	+NS

vizsgált eset

osztályozó példa

S – store, NS – non store,

P – pass, G - get (ask for), D - drop

↓	IB1	-S	-S	+S	+S
↓	IB2	-NS	+NS	-S	+NS

IB2 nem tárolja az IB1 által sem jól osztályozott példát,
IB2 javul. **IB2 a példát nagy valószínűséggel zajnak hiszi.**

IB1	-S	-S	+SP	+S
IB2	-NS	+NS	-S	+NS

Ha IB2 rosszul osztályoz, amikor IB1 helyesen,
IB2 eltárolja az IB1 által átadott osztályozó példát.
További javulás, egyes problémákon jobb, mint IB3.

IB1	-S	-SD	+SP	+SP
IB2	-NS	+NS	-SD	+NS

IB1 eldobja a rossz eredményt (IB2 másképpen osztályoz) adó osztályozó példát. Ha mindketten jól működnek, IB1 átadja a leginkább hasonló tárolt példát, amit az osztályozásnál használt, és azt az IB2 eltárolja. IB2 eldobja az osztályozó példáját, ha rosszul működik, és átveszi az IB1-ét. Legjobb jóslási képesség zajos eredményeken. Megnőtt tár.

IB1	-S	-SD	+SP	+NSP
IB2	-NS	+NS	-SD	+NS

Helyes osztályozás esetén az új példát nem tárolják.
IB1 nem jobb, de tár kevesebb. IB2 jobb, mint egy független IB2, majdnem IB3 hatékonysági szinten.

Ugyanaz a stratégia (IB2), de eltérő tanítóhalmazok

A $-NS \uparrow -SG \uparrow +NS +NS$
B $-S \mid +NS -S +NS$

A csak akkor kérdez **B**-től az ő eredményét, ha neki nem sikerült. Ha **B** helyesen osztályozza az **A** tanító példáját, **A** megkéri (get) és eltárolja a **B** osztályozó példáját.

Ha **B** szerint sem korrekt a példa, akkor **A** eldobja azt (feltehetően zajos). Zajos adatokon **A** jobb, mint **B**, fele annyi példát tárolva.

Ez nem más, mint a 'véleménykikérés', 'segítségnyújtás' **B** részéről.

A $-NS -SG +NS +NSG$
B $-S +NS -S +NS$

A minden esetben kérdez rá a **B**-re, és eltárolja a **B** osztályozó példáját, ha az korrektül osztályoz. **A** további javulása (pozitív döntések megerősítése).

A $-SP -SP +NS +NS$
B $-SP +NS -SP +NS$

Mindkét együtt cselekvő minden rosszul osztályozott példáját másnak elküldi. Így mindenki azonos szinten integrálja magába a közösség tudását – **tanult tudás megosztása**.

MAS hatások

Egy adott **M** példaszám **N** db ágensre oszlik el, ami – független tanulás esetén – az egy-egy ágensnek a tanulási görbét ellapítaná, mivel egy-egy ágens a tanulóhoz nem **M**, hanem **M/N** példát használ fel.

Az együttműködés révén az egyedi (**M/N** példából tanuló) ágensnek tanulási görbéje az **M** példából önállóan tanuló ágens tanulási görbével lesz azonos (ugyanolyan meredek), **az együttműködés tehát a példahiányt képest pótolni, valahogy úgy, mintha egy nyers háttértudást jelentene az ágensnek.**

(gondolkodjunk itt a háttértudás példapótló hatására, ld. alap MI kurzus, tanulóval foglalkozó fókák, magyarázat alapú tanulás).

Katasztrófa-elhárítás



Kooperáció és intelligencia, BME-MIT

Természeti katasztrófák és az információs technológia tapasztalata
(1995-ös **Kobe-i földrengés**, több, mint 6 ezer áldozat, a város 1/5-e megsemmisült):

1. A feltételezett felskálázhatóság nem volt elegendő,
2. Szükségsegítség központok (Emergency Response center) és dolgozói sérülnek/ megsemmisülnek,
3. Kommunikációs kiesések, zavarok,
4. Lakosság és önkéntesek információs elszigeteltsége,
5. Nem elegendő támogatás a döntéshozatalhoz.

Tanulás valós-idejű stratégiai csapatmunkában

(CMUnited - Szoftver labdarugó bajnok 97-99, DARPA, majd **RoboCup-Rescue**)

Csapat (team)

- tárgyalás, megegyezés a közös tervekben a cselekvések indítása előtt
- **dinamikus, valós-idejű környezet, korlátozott kommunikáció**

PTS – Periodic Team Synchronization

**autonóm működés kommunikáció kiesése esetén,
mégis cselekvés a közös (csapat) cél felé**

(futball, kórház/ gyármenedzsment, több űrhajóból álló misszió,
keresés/mentés (SAR search and rescue), csatatéri műveletek, hálózati
csomag routing, stb.),

„periodikus” szinkronizálás, korlátlan kommunikációval (off-line)

dinamikus, valós-idejű (on-line) működés

csapat teljesítménye romlik, ha valamelyik ágens működése romlik
a cél esélye romlik, vagy az elérése időben kihúzódik,

nem megbízható kommunikáció

üzenet érkezése valamilyen valószínűséggel,
üzenetvétel csak bizonyos időközönként,
kommunikációfüggőség költsége

(ha üzenetvétel nélkül nem tud cselekedni)

Ágens felépítése és tudása (az alap architektúra egy változata)

a világalapota (felfrissül szenzorok, megjósolt cselekvések alapján)

hozzáférhető belső/külső viselkedés számára

belső állapot (ágens állapotváltozói)

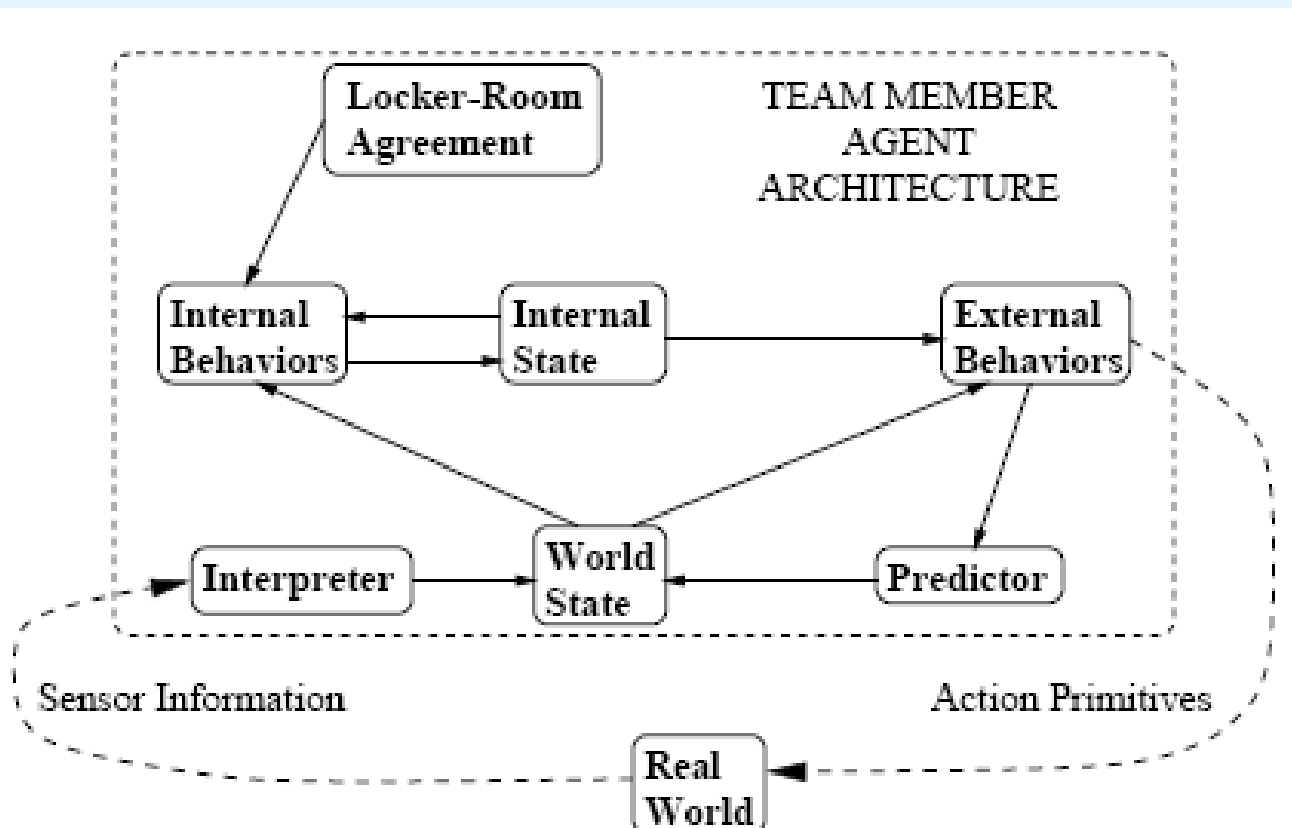
belső viselkedés (kognitív)

külső viselkedés (kommunikációs + fizikai)

öltözői taktikai megbeszélés (LRA - locker-room agreement)

amikor a csapat képes elszigetelten „szinkronizálni”

– rugalmas csapatstruktúra, inter-ágens protokollok,



Csapat (team) struktúra

rugalmas ágensszerepek + szerepcsere protokollok

csapatformációk szerephalmazokkal

többlépéses, többágenses tervek meghatározott
körülményekhez

– set-plays

Szerep (viselkedés) specifikálása

'merekv'

'rugalmas' (bizonyos autonómia)

Kihívások:

öltözői megbeszélések reprezentációja és végrehajtása;
hogyan állapítjuk a szerepváltás és formációváltás legalkalmasabb időpontját;
hogyan biztosítani, hogy minden ágens azonos formációt vesz fel;
hogyan biztosítjuk, hogy egy formáció minden szerepe aktív.

Formáció

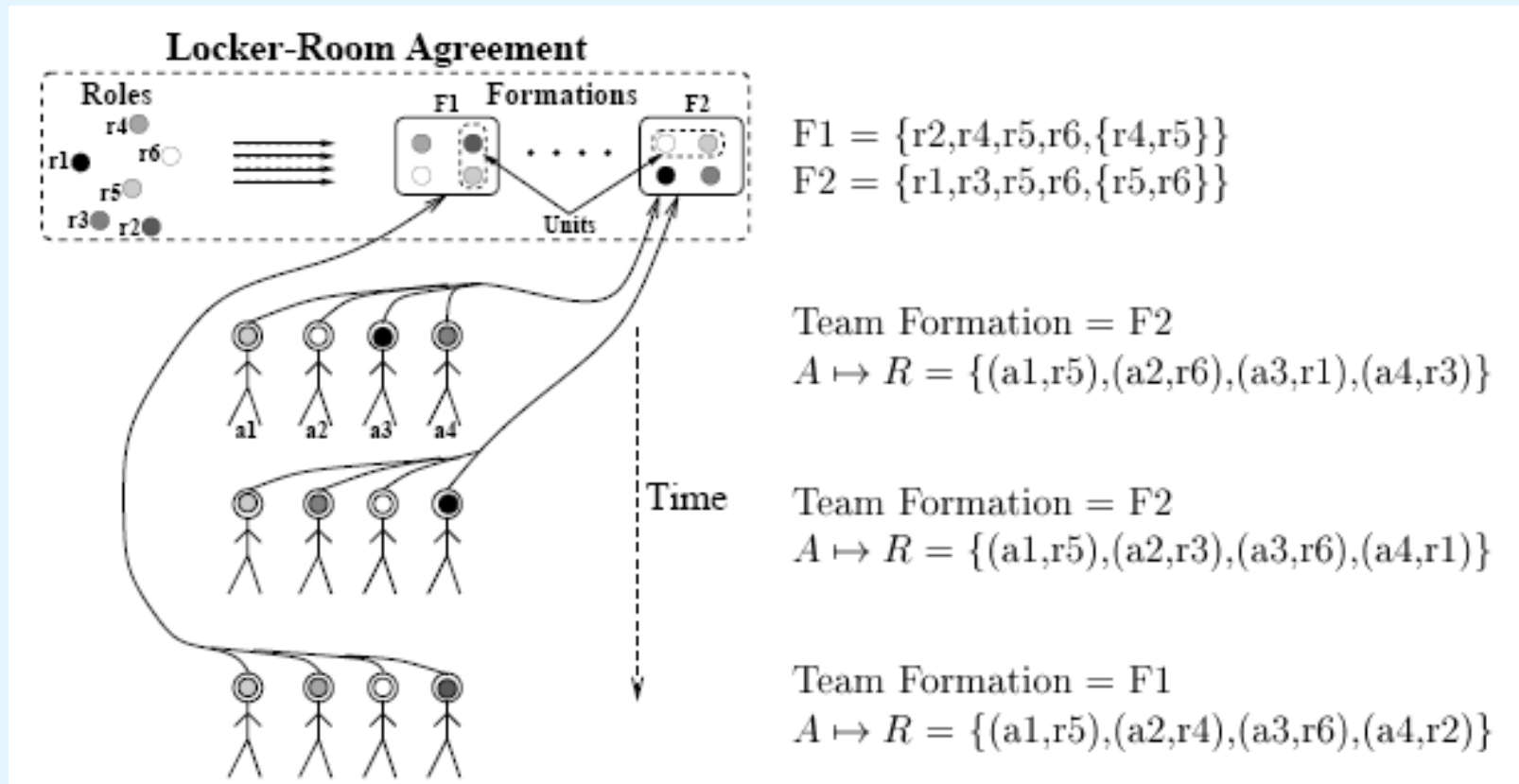
$$F = \{R, \{U_1, U_2, \dots, U_k\}\}$$

R – szerepek

U – egységek, egység = a szerepek részhalmaza, „kapitány” szerep

LRA = kezdő formáció, kezdő ágens-szerep leképzés, formációváltás futásidejű triggerek, korlátos kommunikáció, az aktuális kép (ki milyen egységben, szerepben) lehet torz, robusztus cselekvés, periodikus szerepismertetés,

Pl.



Kommunikáció (kihívások)

egyedi csatorna, alacsony sávszélesség, megbízhatatlan kommunikáció

azonosítás: melyik üzenet melyik ágensnek szól

aktív zavarás ellenséges ágensek részéről, 'álruhás' üzenetek

LRA – kódolási megállapodások

alacsony sávszélesség miatt kerülni kell, hogy többen egyszerre beszéljenek
maximalizálása annak a körülménynek, hogy a csapat azonos formációt használ, de a kommunikációs kiesések miatt az ágensek autonómok

Üzenettípusok:

Message Target	Response requested	
	no	yes
Single agent	a1	b1
Whole team	a2	b2
Part of team	a3	b3

kommunikáció – segíti felfrissíteni a világ- és belső állapotot, nem feltétele a cselekvésnek (robosztusság kommunikációs kiesésekre)

Csapatszinkronizálás valós-időben

LRA + ami átjön az üzenetekből

Labdarugó problématerület (eredeti kísérlet) – CMUnited-97

22 ágens

egyetlen megbízhatatlan kommunikációs csatorna

egy beszélő ágenszt mindenki hall, irányát nem szükségképpen azonosíthatóan, csak bizonyos távolságon belül, üzeni csak bizonyos időközönként tud

Objektummozgás

mozgás, gyorsulás, romlás,
csökkenés, ..., ütközés

Zaj

Játékos teherbíró képessége:

újratermelhető, nem újratermelhető.

Ágens érzékelése

aura információ

a látás precizitása

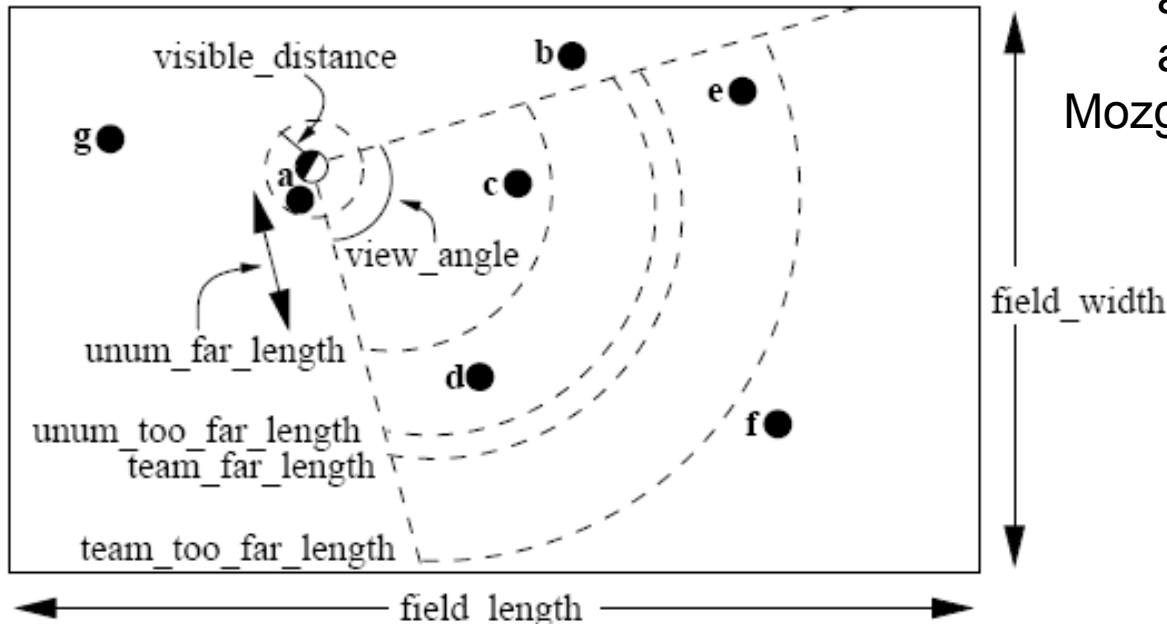
a hallás precizitása

Mozgás Fordulás

Előrerohanás

Rúgás

Fogás



Érzékelés és cselekvés aszinkron.

Látás korlátozott, környezet részben hozzáférhető

Kommunikáció csak közeli hallott üzenetekkel

Véges teherbíró képesség

Zaj

beavatkozásban

objektumok mozgásában

vizuális érzékelésben

Érzékelés, vagy üzenet kiesése

Utasítás végrehajtására nincs garancia

(végrehajtást szenzorikusan kell verifikálni)

Valós-idejű

(aura információ aszinkron módon érkezik)

Pozicionálás

strategic position by attraction and repulsion (SPAR)

taszítás (repulsion) az ellenségtől

taszítás (repulsion) csapattársaktól

vonzás (attraction) aktív csapattársakhoz és labdához

vonzás (attraction) az ellenség kapujához

Kooperáció és intelligencia, BME-MIT

Mérkőzések – RoboCup-XX, www.robocup.org

CMUnited-97

formáció: 4-4-2 8-2-0 (vége felé, ha győz)

3-3-4 (vége felé, ha veszít)

negyeddöntő (29 csapat, AT Humboldt), össz: 67-14

CMUnited-98

bajnok (34 csapat), össz: 66-0

CMUnited-99

bajnok (37 csapat):_



<http://www.robocup2014.org/>



RoboCup-Rescue típ. projektek

Tanulható csapatevékenység kiterjesztése heterogén szerepkörű, kritikus folyamatokkal teli környezetben célokat megvalósító ágensközösségekre.

Természeti katasztrófák és az információs technológia tapasztalata (1995-ös Kobe-i földrengés, több, mint 6 ezer áldozat, a város 1/5-e megsemmisítve):

1. A feltételezett felskálázhatóság nem elegendő,
2. Szükségsegítség központok (Emergency Response center) és dolgozói sérülnek/megsemmisülnek,
3. Kommunikációs kiesések, zavarok,
4. Lakosság és önkéntesek információs elszigeteltsége,
5. Nem elegendő támogatás a döntéshozatalhoz.

Szimulációk (**RobotCup-Rescue Simulation**) célja:

1. A szükséges információ begyűjtése, akkumulálása, továbbítása, szelektálása, elemzése, összegzése, szétküldése.
2. Megfelelő döntéstámogatás.
3. Rendszer elosztottság megnövelt megbízhatóság és robusztusság érdekében.
4. Folyamatos cselekvés biztosítása normális körülményektől a szükségállapotig.



19-20 JULY 2014
Setup Day

21-24 JULY 2014
Competitions

25 JULY 2014
Symposium



[HOME](#)

[LEAGUES](#)

[SYMPOSIUM](#)

[PARTICIPANTS](#)

[VISITORS](#)

[RoboCup Rescue](#)

ROBOCUP RESCUE

Disaster rescue is one of the most serious social issue which involves very large numbers of heterogeneous agents in the hostile environment. The intention of the RoboCupRescue project is to promote research and development in this socially significant domain at various levels involving multi-agent team work coordination, physical robotic agents for search and rescue, information infrastructures, personal digital assistants, standard simulator and decision support systems, evaluation benchmarks for rescue strategies and systems that are all integrated into a comprehensive system in future.

Kooperáció és intelligencia, BME-MIT

USAR (Urban Search And Rescue) Arenas_

Reference Test

Arenas for

Autonomous

Mobile Robots_



<http://www.nist.gov/el/isd/testarenas.cfm>



Specific robotic capabilities:

- Negotiate compromised and collapsed structures
- Locate victims and ascertain their conditions
- Produce practical sensor maps of the environment
- Establish communications with victims
- Deliver fluids, nourishment, medicines
- Emplace sensors to identify/monitor hazards
- Mark or identify best paths to victims

Kooperáció és intelligencia, BME-MIT

Simulated Victims: Simulated victims with several signs of life such as form, motion, head, sound and CO2 are distributed throughout the arenas requiring directional viewing through access holes at different elevations.

Yellow Arena: For robots capable of fully autonomous navigation and victim identification, this arena consists of random mazes of hallways and rooms with continuous 15° pitch and roll ramp flooring.

Orange Arena: For robots capable of autonomous or remote teleoperative navigation and victim identification, this arena consists of moderate terrains with crossing 15° pitch and roll ramps and structured obstacles such as stairs, inclined planes, and others.

Red Arena: For robots capable of autonomous or remote teleoperative navigation and victim identification, this arena consists of complex stepfield terrains requiring advanced robot mobility.

Blue Arena: For robots capable of mobile manipulation on complex terrains to place simple block or bottle payloads carried in from the start or picked up within the arenas.

Black/Yellow Arena (RADIO DROP-OUT ZONE): For robots capable of autonomous navigation with reasonable mobility to operate on complex terrains.

Black Arena (Vehicle Collapse Scenario): For robots capable of searching a simulated vehicle collapse scenario accessible on each side from the RED ARENA and the ORANGE ARENA.

Aerial Arena (< 2 KG, VTOL sUAS): For small unmanned aerial systems under 2 kg with vertical take-off and landing capabilities that can perform station-keeping, obstacle avoidance, and line following tasks with varying degrees of autonomy.

Rétegezett tanulás (Layered Learning)

közvetlen bemeneti adat – kimenet leképezés tanulása nem kivitelezhető

- dekompozíció, taszk feltörése rétegekre,
- más-más koncepció tanulása rétegenként, lokális tanulás – lokális koncepciók

alulról felfelé taszk dekompozíció adva van

alulról felfelé tanulás, amíg nem éri el a teljes probléma-komplexitást, altáaszok, granuláltság – a konkrét tartomány függvénye.

adatokból tanulás, adaptálás – tanulás minden szinten

gépi tanulás: ha kézi (szimbolikus) hangolás nem megy, adaptálás, ha a feladat előre alul-definiált, ha dinamikus.

tanulás eredménye egy szinten – a következő szint példái

minden szinten tanulás a következő szintet közvetlenül befolyásolja:

- a tanító példák szerkesztése,
- a tanításhoz szükséges tulajdonságok biztosítása,
- a kimeneti halmaz nyesése



Rétegezett tanulás (Layered Learning)

Formalizmus

tanulás: **H** a hipotézisek halmaza, belőle egy **h** hipotézis tanulása

H: állapottulajdonságok **S** $\rightarrow$ kimenetek **O**

rétegezett tanulás:

hierarchikus taszkrétegek, egy réteg: **Li** = (**Fi**, **Oi**, **Ti**, **Mi**, **hi**)

Fi – az állapottulajdonságok bemeneti vektora

Oi – a kimenetek halmaza

Ti – tanító példák az **Li** taszk számára (**(f,o): f** $\in$ **Fi** $\rightarrow$ **o** $\in$ **Oi**)

Mi – a gépi tanulás algoritmus **Li** rétegben, **Fi** – **Oi** leképezést legjobban leíró hipotézis megválasztása **Ti** alapján,

hi – tanulás eredménye

<u>Réteg</u>	<u>Stratégiai szint</u>	<u>Viselkedés</u>	<u>Példa</u>
1	Robot	egyéni	labdafogás
2	1-1 játékos	több-ágens	átadás értékelés
3	1-N játékos	team	átadás válsztás
4	team formation	team	stratégiai poz.
5	team-ellenség	adverz	stratégiai adapt.

Robot football

	<u>viselkedés</u>	<u>példa</u>
L1	egyedi	labda megfogása
L2	több ágens	labdaátadás értékelése
L3	csapat	labdaátadás kiválasztása

**L1 – Labda megfogása: a megtanulása (sebesség információ nélkül)
könnyebb, mint a 'beprogramozása'**

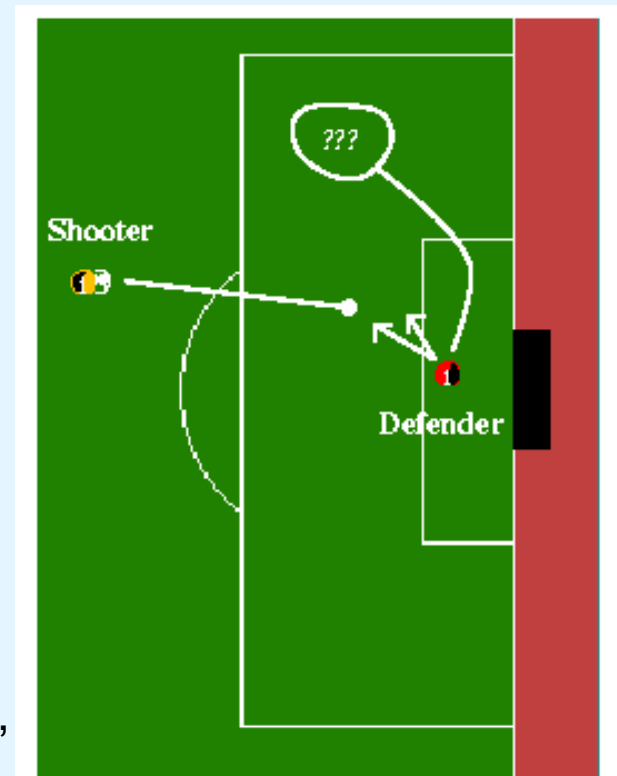
F1 – labda távolsága-1, -2, szög

O1 – megfordulás szöge

T1 – kapura lövések,
minősítés: megfogta, gól, mellément

M1 – neurális háló

h1- kb. 5000 példa után 84% megfog, 9% gól



Kooperáció és intelligencia,

L2 – Labdaátadás értékelése: több ágenses viselkedés tanulása, labdaátadás egy másik játékosnak – sikerül-e? átveszi? A tanító példák szerkesztésénél a passzoló és az ellenségek tudják a h1-t.

F2 – 174 tulajdonság, játékosok, ellenségek pozíciói, lényeges, lényegtelen, ... hadd szelektálja a tanuló algoritmus!

O2 – [-1, 1], szándékolt átadás vagy siker v. kudarc: 0 ... 1 megbízhatósági tényezővel

T2 – véletlenül helyezett játékosok, h1 labda kezelési képességgel, siker (társ megfogta), kudarc (ellenségek megfogták), félrement (senki), véletlen elhelyezésnél a passzok 51% sikeres,

M2 – döntési fa tanulása (irreleváns bemenetek kisselektálása),

h2- döntési fa, osztályozás + (-1 ... 1) megbízhatósági tényező, összes átadás 65%-ka, sikeresnek becsült átadások 79%-ka valójában sikeres,



Sample Decision Tree Output

→ Teammate 2 : Success with confidence 0.8
Teammate 3 : Failure with confidence 0.6
Teammate 4 : Success with confidence 0.3

Kooperáció és ir

L3 – Labdaátadás szelektálása: Együttműködő /ellenséges csapatviselkedés,

h2 felhasználása tanító minták szerkesztéséhez: ha rendelkezik labdával, kinek adja át (a legsikeresebbnek ítélt passz általában nem a stratégiailag legjobb, azért dönthet, átadja, vagy kapura lő).

Óriási a döntési tér (ágens kurrens pozíciója, társai, ellenségei kurrens pozíciói, társak képessége az átadás átvételére, ellenségek képessége az átadás megfogására, társak döntési képessége, ellenségek stratégiája), a döntés jósága csak a csapat hosszú idejű sikereiből, a tér redukálása **h2**-ből.

F3 – (játékos pozíciója, ...), durva helyzetbontás, 1 cselekvésfüggő tulajdonság minden átadáshoz (h2 eredménye minden lehetséges megfogó társhoz)

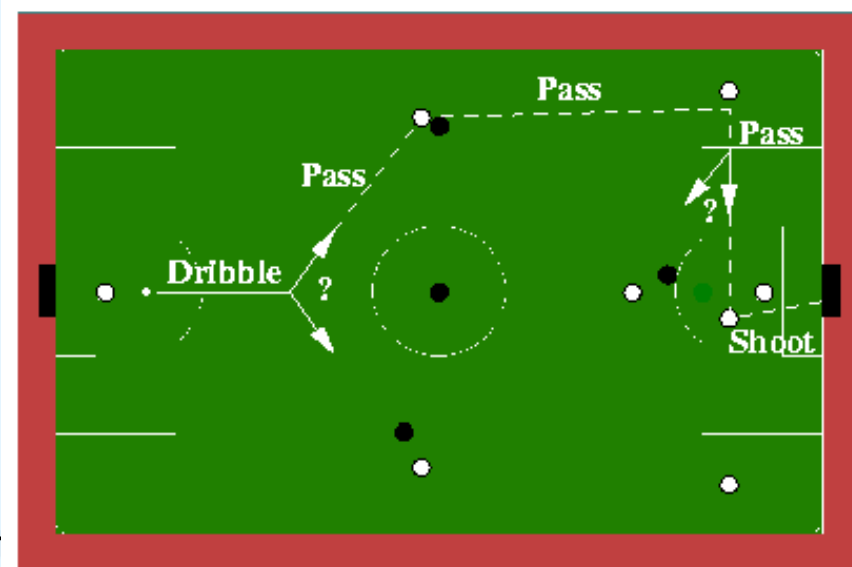
O3 – (lövés) $\cup$ (melyik társ),

T3 – valós mérkőzésekből,

M3 – TPOT-RL: on-line, többágenses, megerősítéses tanulás, Q-tanulás mintájára

(csapatbontott, átláthatatlan területek),

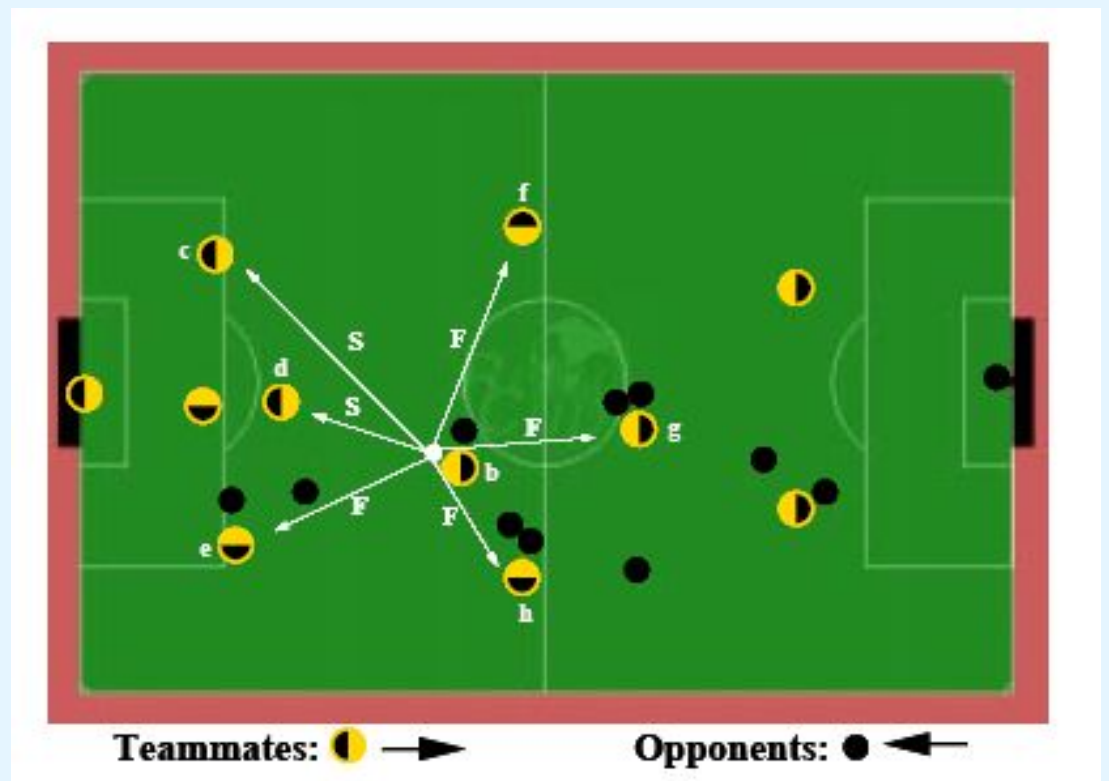
h3- elosztott átadás szelektálási stratégia,



- **Teammate**

- **Defender**

Kooperáció és in



RL algoritmus kihívásai:

1. on-line
2. nagy állapotteret kezel korlátozott tanulással
3. tanulni képes igen késleltetett megerősítésből
4. változó helyzeteket/koncepciókat kezel
5. team-bontott helyzetekben működik
6. „elmosódó átmeneteket” kezel

TPOT-RL = Team-Partitioned, Opaque-Transition Reinforcement Learning

Hogy nézne ki egy „rendes” megerősítéses tanulás?

**egyedi ágens - nincs kontrollja a helyzet egészén,
a cselekvése folytatásán sem!**

**„láncolt cselekvések” – a cselekvését más ágens cselekvése követi, stb.
a cselekvés utáni állapot ismeretlen (másoktól, ellenségtől függ)**

A tanult stratégia: állapot $\rightarrow$ cselekvés leképezés (Q tanulás)
 $S \rightarrow A$

Mivel az állapottér igen nagy lehet (és így a tanulás kevés példából nehéz v. lehetetlen), az állapotteret általánosítani kell kisebb komplexitású (dimenzióban redukált) V tulajdonságtérre (feature space) (F_2 , ld. előbb).

$$S \rightarrow V \rightarrow A$$

Tanulás lépései:

1. állapottér általánosítás: $s \in S \rightarrow v \in V$
- 2. értékfüggvény tanulás $Q: (v, a) \rightarrow R$**
3. cselekvés megválasztása, megerősítés érkezése, Q felfrissítése

1. Állapottér általánosítás

állapottér partíció (egy-egy játékos nem az egész pályán „érdekelt”), az eggyel lejjebbi szinten történő tanulás eredménye (mi lényeges az alacsonyabb szintű jellemzők között).

2. Értékfüggvény tanulás - „elvi” (idő-különbség Q -tanulás):

$$Q(s, a) = Q(s, a) + \alpha (r + \gamma \max_{a'} Q(s', a') - Q(s, a))$$

kihasználja a „modell” ismeretét, azaz hogy a cselekvés hatására milyen állapotba kerül az ágens környezete: $T: (S, A) \rightarrow S$,

de ez itt nem alkalmazható!

átláthatatlan környezetben való (modell nélküli) tanulás (tulajdonképpen „naiv” hasznossági tanulásnak felel meg):

$$Q(v, a) = Q(v, a) + \alpha (r - Q(v, a))$$

Mi a megerősítés?

$r \in [-Q_{\max}, Q_{\max}]$, $t_{\lim}$ -belül a jövőből, mindkettő problémafüggő

Probléma: a megerősítés túl ritka! (gól, mérkőzés-, bajnokság végeredménye)
ágens: kb. 10 tanuló példa/ 10 perces játék nagyon kevés!

hatékony tanulás kell!

Megerősítés kérdése még egyszer:

- gól, igazi megerősítés, de ritka,
- „**belső**” megerősítés, környezeti megfigyelések alapján
(jelen esetben a labda pályája alapján)
rúgás pillanatában megjegyzi a labda pozícióját: x_t

$$r = R(\text{labda pozíció} \mid t_{\lim})$$

$r =$ a labda out-ra megy (a gól is az): $t + t_0$ ($t_0 < t_{\lim}$), de minél
később, annál kevesebb visszkapja a labdát: $t + t_r$ ($t_r < t_{\lim}$)

a labda még mindig a pályán: $t + t_{\lim}$

az ágens a labda átlagos pozícióját számolja (megfigyelésből)

ha $x_{\text{avg}} > x_t$ $r > 0$, annál nagyobb, minél közelebb a (ellenfél) kapuhoz

ha $x_{\text{avg}} < x_t$ $r < 0$ annál nagyobb, minél közelebb a (saját) kapuhoz