

Tanulás elosztott rendszerekben/2

Egy szervezet

- egy természetes, de „ellenséges” környezetben, ill.
- egy olyan környezetben, ahol más „ellenséges” szervezetek vannak

Kiindulás - katasztrófa-elhárítás

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.

**Dinamikus, „ellenséges” (nem kooperatív),
valós-idejű környezet,
korlátozott kommunikációval**

Tanulás valós-idejű stratégiai csapatmunkában

PTS – Periodic Team Synchronization

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

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

Időszakos koordinálás, korlátlan, zavartalan kommunikációval (off-line)

majd dinamikus, valós-idejű (on-line) működés

Nem megbízható kommunikáció

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

(ha üzenetvétel nélkül nem tud cselekedni)

(ha üzenetvétel miatt nem tud cselekedni)

Öltözői taktikai megbeszélés (LRA - Locker-Room Agreement)

Ágens felépítése és tudása (az alap architektúra egy változata)

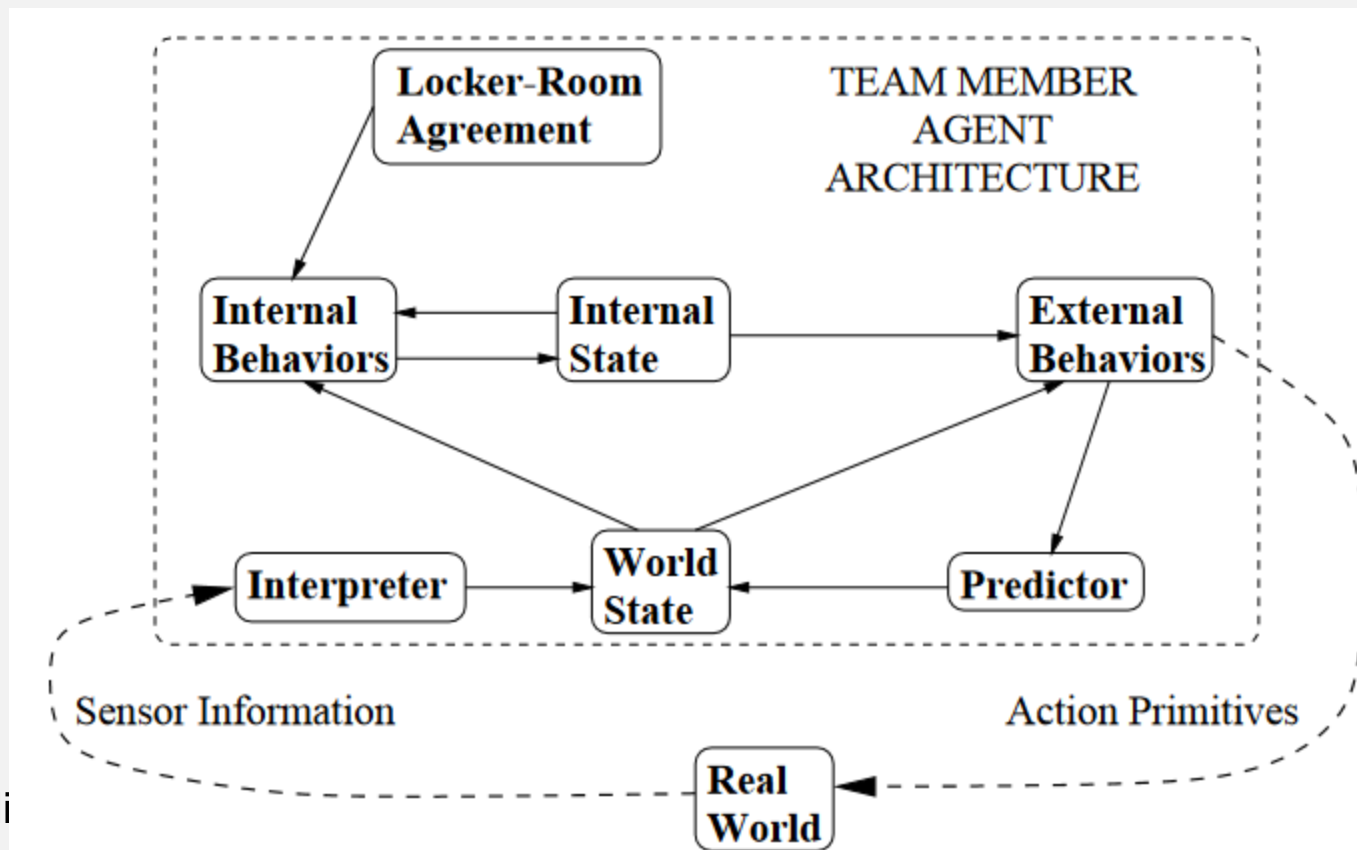
világállapota (felfrissül szenzorok, megjósolt cselekvések alapján)

hozzáférhető belső/külső viselkedés számára

belső állapot (ágens állapotváltozói)

belső viselkedés (kognitív)

külső viselkedés (kommunikációs + fizikai)



Csapat struktúra **rugalmas ágensszerepek + szerepcsere protokollok**
csapatformációk szerephalmazokkal
többlépéses, **többágenses tervek** meghatározott
körülményekhez – set-plays

Szerep (viselkedés) specifikálása: 'merev', 'rugalmas' (autonómia)

Kommunikáció (kihívások): egyedi csatorna, alacsony sáv szélesség,
megbízhatatlan kommunikáció, azonosítás: melyik üzenet melyik ágensnek
szól, aktív zavarás ellenséges ágensek részéről, 'álruhás' üzenetek,
LRA – kódolási megállapodások

Üzenettípusok:

üzenet címzettje: egyedi ágens, team része, egész team

válasz: nem szükséges, szükséges

Csapatkoordinálás: LRA + ami átjön az üzenetekből valós időben

Kihívások:

LRA megbeszélések reprezentációja és végrehajtása;

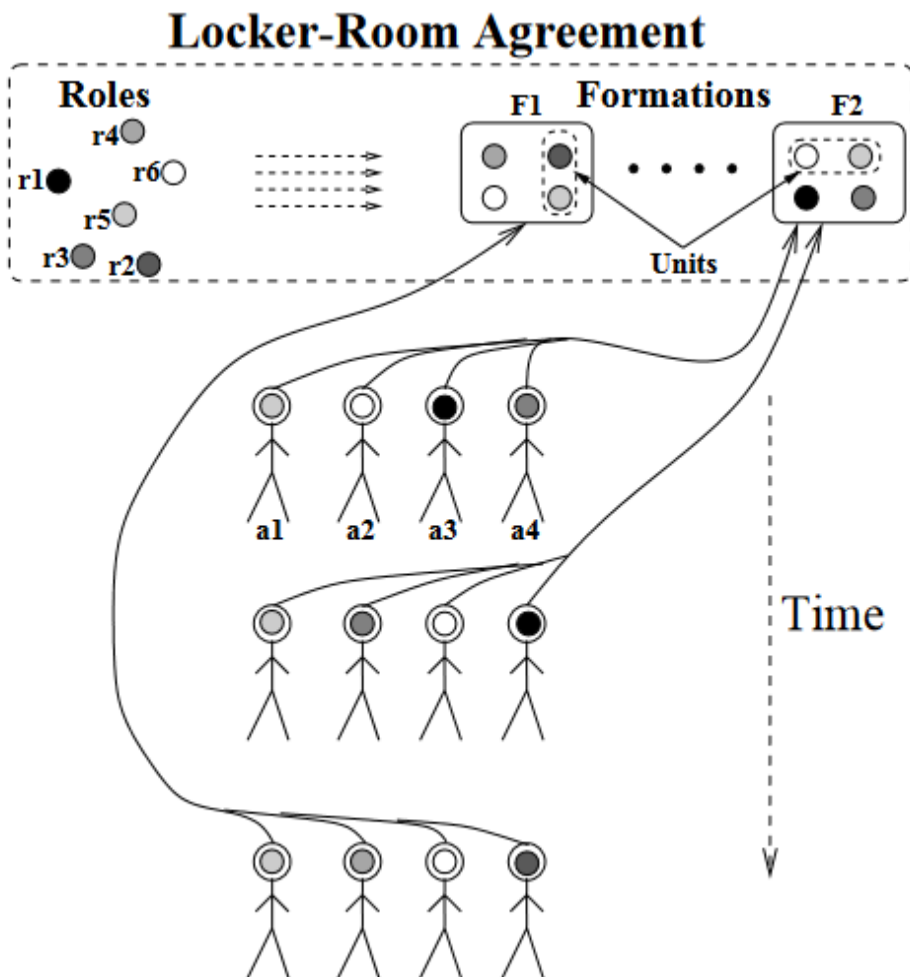
Hogyan állapítjuk a szerepváltás és formációváltás legjobb időpontját;

Hogyan biztosítjuk, hogy minden ágens azonos formációt vesz fel;

Hogyan biztosítjuk, hogy egy formáció minden szerepe aktív.

Formáció

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



$$F1 = \{r2, r4, r5, r6, \{r4, r5\}\}$$

$$F2 = \{r1, r3, r5, r6, \{r5, r6\}\}$$

Team Formation = F2

$$A \mapsto R = \{(a1, r5), (a2, r6), (a3, r1), (a4, r3)\}$$

Team Formation = F2

$$A \mapsto R = \{(a1, r5), (a2, r3), (a3, r6), (a4, r1)\}$$

Team Formation = F1

$$A \mapsto R = \{(a1, r5), (a2, r4), (a3, r6), (a4, r2)\}$$

Tanulás valós-idejű stratégiai csapatmunkában

(CMUnited – szg. labdarugó bajnok 97-99, DARPA (eredeti kísérlet))

MAS: 22 ágens

Kommunikáció: egyetlen megbízhatatlan kommunikációs csatorna, csak közelről hallott üzenetekkel.

Objektummozgás: mozgás, gyorsulás, lassulás, ..., ütközés.

Játékosok teherbírása: újratermelhető, nem újratermelhető, véges.

Ágens érzékelése: aura információ, látás precizitása, hallás precizitása, látás korlátos, környezet részben hozzáférhető.

Mozgások: fordulás, előrerohanás, rúgás, megfogás.

Érzékelés és cselekvés aszinkron, valós-idejű.

Zaj: beavatkozásban, objektumok mozgásában, vizuális érzékelésben.

Érzékelések, vagy üzenetek kiesése.

Utasítások végrehajtását szenzorikusan kell verifikálni.

Strategic position by attraction and repulsion (SPAR)

taszítás (repulsion) az ellenségtől, a csapattársaktól

vonzás (attraction) aktív csapattársakhoz, labdához, az ellenség kapujához.

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):

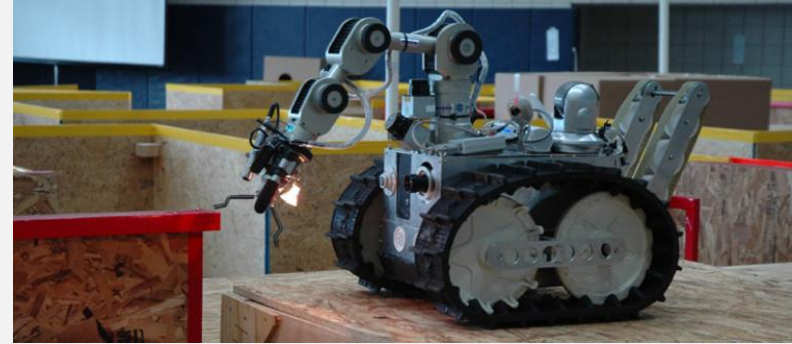
Games & Conferences
Robot World Cup

RoboCup-97 Nagoya



RoboCup-Rescue projektek

Tanulható csapattevékenység kiterjesztése kritikus környezetben célokat megvalósító ágensekre.



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

A szükséges információ begyűjtése, akkumulálása, továbbítása, szelektálása, elemzése, összegzése.

Megfelelő döntéstámogatás.

Rendszereosztottság megnövelt megbízhatóság/robusztusság érdekében.

Folyamatos cselekvés biztosítása normálistól szükségállapotig.



Intelligens Elosztott Rendszerek BME-MIT, 2019-2022

**DISASTER
ROBOTICS**

ROBIN R. MURPHY



USAR (Urban Search And Rescue) Arenas_

Test Arenas for Autonomous Mobile Robots_



<https://www.nist.gov/el/intelligent-systems-division-73500/performance-metrics-and-test-arenas-autonomous-mobile-robots>



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.



RoboCup@Home arena

Follow me, Clean up, ...

„Move to the LOCATION, find a person, and guide it to the exit. ...”

stb.

EGPSR

Endurance General Purpose
Service Robot Test



Egy szervezet tanulása (nézzünk bele)
(más szervezetek „ellenséges” környezetében)

Rétegezett tanulás (Layered Learning)

Közvetlen bemeneti adat – kimenet függvény tanulása nem megy

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

Alulról felfelé taszk dekompozíció adva van (feladathierarchia, ld. előbb)
alulról felfelé tanulás, amíg nem éri el a teljes probléma-komplexitást,
altáskok, granuláltság – a konkrét tartomány függvénye.

Adatokból tanulás, adaptálás – tanulás minden szinten

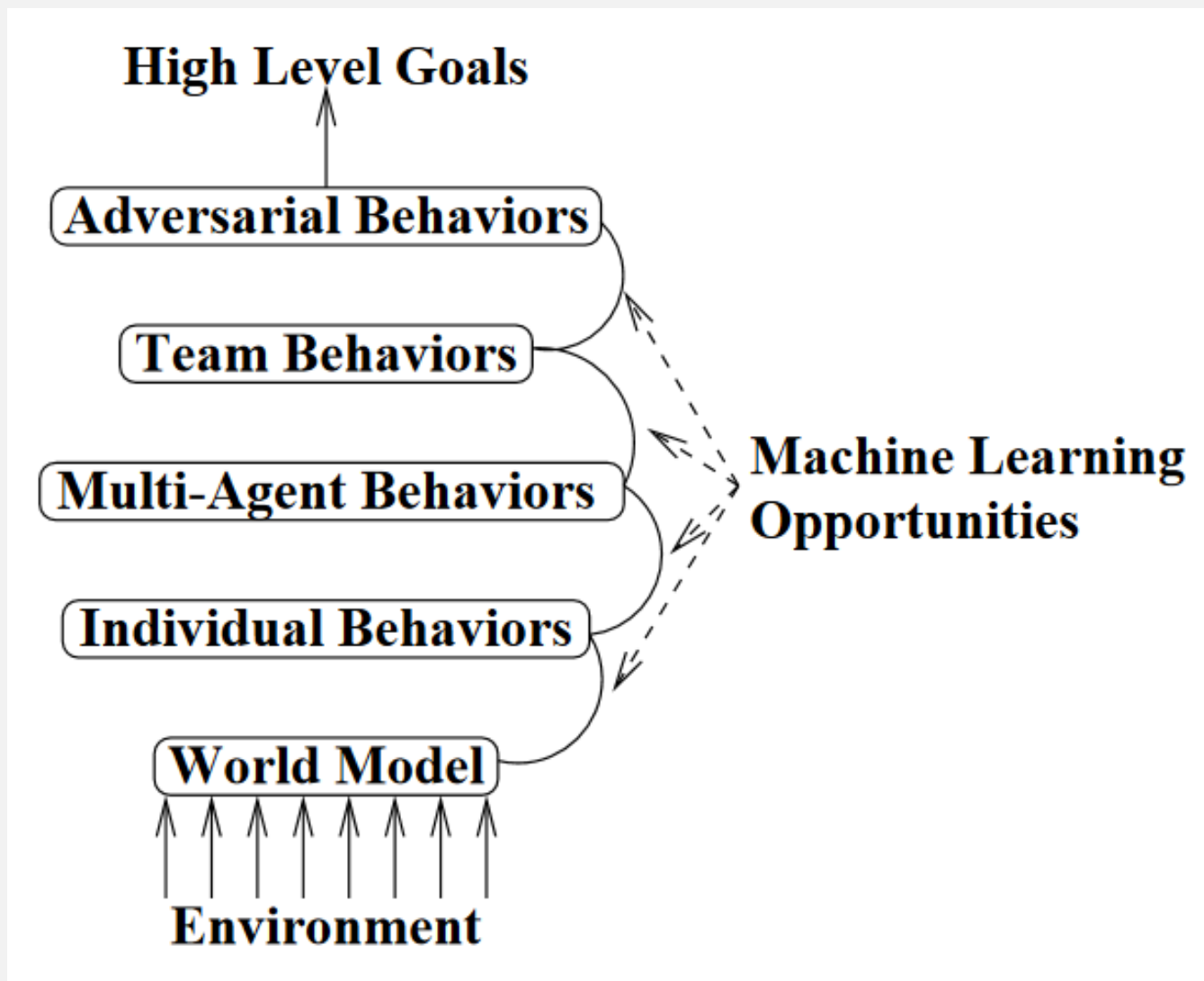
gépi tanulás: ha kézi (szimbolikus) hangolás nem megy,
adaptálódás, ha a feladat előre aluldefiniált, ha dinamikus.

Tanulás eredménye egy szinten – példák a következő szint számára

minden szinten tanulás a következő szintet közvetlenül befolyásolja:

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

Rétegzett tanulás (Layered Learning)



Rétegezett tanulás (Layered Learning)

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

H: állapottulajdonságok (attributúmok) **S** \rightarrow kimenetek **O**

Rétegezett tanulás:

hierarchikus taszkrétegek, egy réteg: $L_i = (F_i, O_i, T_i, M_i, h_i)$

F_i – az állapottulajdonságok bemeneti vektora

O_i – a kimenetek halmaza

T_i – tanító példák az **L_i** taszk számára (**(f,o)**: $f \in F_i \rightarrow o \in O_i$)

M_i – a gépi tanulás algoritmus **L_i** rétegben, **F_i – O_i** függvényt legjobban leíró hipotézis megválasztása **T_i** alapján,

h_i – tanulás eredménye

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

Robot futball

<u>viselkedés</u>	<u>példa</u>
L_1	egyedi labda megfogása
L_2	több ágens labdaátadás értékelése
L_3	labdaátadás kiválasztása

L_1 – Labda megfogása: a megtanulása (sebesség információ nélkül) könnyebb, mint a 'beprogramozása'

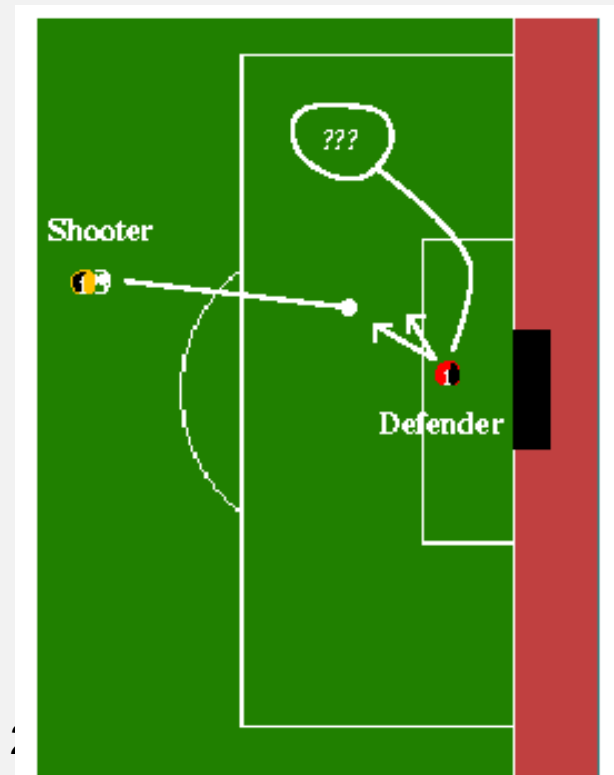
F_1 – labda távolsága-1,-2, szög

O_1 – megfordulás szöge

T_1 – kapura lövések,
minősítés: megfogta, gól, mellément

M_1 – neurális háló

h_1 - kb. 5000 példa után 84% megfog, 9% gól



L_2 – Labdaátadás értékelése: több ágenses viselkedés tanulása, labdaátadás egy másik játékosnak – sikerül-e? átveszi?

A tanítópéldák szerkesztésénél a passzoló és az ellenségek tudják a h_1 -t.

F_2 – 174 tulajdonság, játékosok, ellenségek pozíciói, lényeges, lényegtelen,
... **hadd szelektálja a tanuló algoritmus!**

O_2 – [-1, 1], szándékolt átadás vagy siker v. kudarc: 0 ... 1 megbízhatósági tényezővel

T_2 – véletlenül helyezett játékosok,
 h_1 labda kezelési képességgel,
siker (társ megfogta), kudarc
(ellenségek megfogták), félrement
(senki), véletlenül választott társnál
a passzok 51%-ka sikeres,

M_2 – **döntési fa tanulása (irreleváns bemenetek kiszzelektálása!),**

h_2 - döntési fa, osztályozás + (-1 ... 1)
megbízhatósági tényező, összes
átadás 65%-ka, sikeresnek (.8-.9)
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

L₃ – Labdaátadás szelektálása: Együttműködő/ellenséges csapatviselkedés

h₂ felhasználása tanító minták szerkesztéséhez: ha rendelkezik labdával, kinek adja át (a legsikeresebbnek ítélt passz általában nem a stratégiaileg 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 helyzetfelbontás, 1 cselekvésfüggő tulajdonság minden átadáshoz (h₂ eredménye minden lehetséges megfogó társhoz)

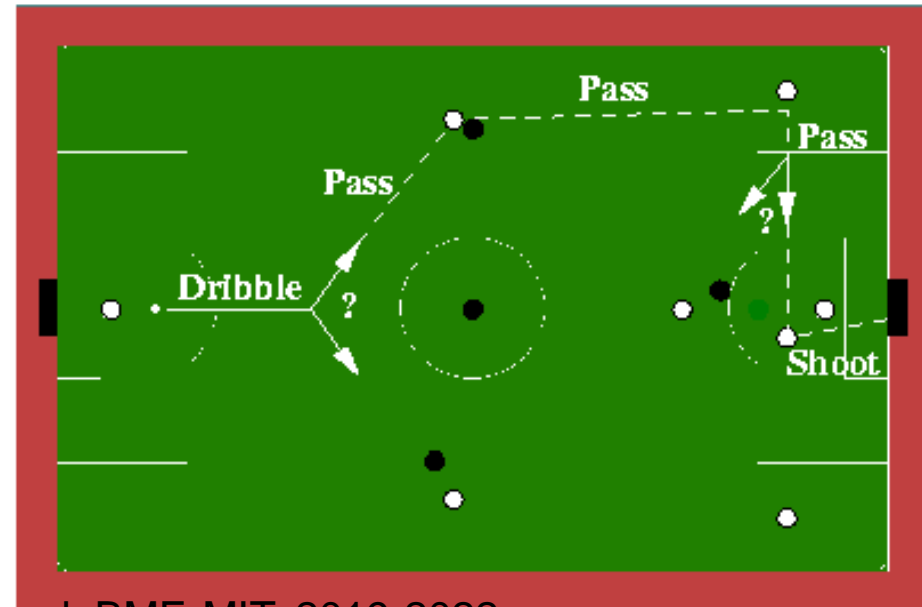
O₃ – (lövés) ∪ (melyik társ),

T₃ – valós mérkőzésekből,

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

(csapatbontott, átláthatatlan területek)

h₃- elosztott átadás szelektálási stratégia,



RL algoritmus kihívásai:

1. on-line
2. nagy állapottér kezelése korlátos tanulással
3. tanulni képes, de igen késleltetett megerősítésből
4. változó helyzeteket/koncepciókat kell kezelni
5. team-bontott helyzetekben működik
6. „elmosódó átmeneteket” kezel

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

Egyedi ágensnek nincs kontrollja a cselekvése folytatásán! A cselekvését más ágensek cselekvései követik. **A cselekvés utáni állapot ismeretlen** (másoktól függ). Átláthatatlan környezetben (modell nélküli) tanulás:

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

A megerősítés túl ritka! (gól, mérkőzés-, bajnokság végeredménye): kb. 10 tanuló példa/ 10 perces játék, nagyon kevés! **hatékony tanulás kell!**

- gól, igazi megerősítés, de ritka,

- „**belső**” megerősítés, környezeti megfigyelések alapján, rúgás pillanatában megjegyzi a labda pozícióját: x_t -t, $r = R(\text{labda pozíció} \mid t_{\text{lim}})$