

MESTERSÉGES INTELLIGENCIA

MESTERSÉGES ÉLET

ÁGENS RENDSZEREK

Tadeusz P. Dobrowiecki

Kooperatív rendszerek jegyzete alapján

[https://wiki.sch.bme.hu/bin/view/Infoszak/KoopRsz?CGISES
SID=7080697d09172c65fa6540daf02f3055](https://wiki.sch.bme.hu/bin/view/Infoszak/KoopRsz?CGISES
SID=7080697d09172c65fa6540daf02f3055)

Milyen rendszer egy ágens?

- *"Ágens akármi lehet, amit úgy lehet értelmezni, hogy szenzoraival a környezetét érzékeli és a beavatkozó szerveivel a környezetébe beavatkozik" (Russel, Norvig, 1995)*
- *"Autonóm ágensek olyan számítási rendszerek, amelyek valamilyen komplex dinamikus környezetben tartózkodnak, érzékelnek és ebben a környezetben autonóm módon cselekednek, és ily módon olyan taszkokat vagy célokat valósítanak meg, amire megtervezték őket" (Maes, 1995)*

Multi-agens rendszerek

- A több - akár egymástól független - ágensből (az egyszerűség kedvéért, pongyolán: cselekvőből) álló rendszereket multi-agens rendszereknek nevezzük.
- Multi-agens rendszerekkel igen gyakran találkozunk a bennünket körülvevő mind természetes, mind mesterséges környezetben:
 - az emberi társadalom maga, és különböző szerveződései, továbbá a gazdaság szereplői,
 - a rovartársadalmak és egyéb állatkolóniák, ...
 - Bonyolult hardver és szoftverrendszerek...
- A rendszerek esetében az egyik legfontosabb, legérdekesebb folyamat a **koordináció**.
- Ez az a mechanizmus, melynek segítségével a rendszer szereplőinek tevékenysége összehangolódik a külső korlátok betartása, illetve valamilyen közös cél elérése érdekében.
- (Pl. a hangyák élelemkeresésének szervezése,
- pl. a vastaps spontán kialakulása is (már emotív elemek függvényében).
- A hálózatok és az elosztottság terjedésével a multi-agens rendszerek egyre gyakoribbak a számítógépes világban is (hagyományosan a kliens-szerver architektúrát, illetve elosztott problémamegoldást (a szereplők egyazon probléma megoldásán fáradoznak) jelent.
- Az internet, s - vele együtt - a "nyílt hálózat" gondolata előtérbe helyezi az egymástól független, csak részben közös célokkal rendelkező rendszereket lényeges koordinálási problémákkal.

Példák:

Internet ágensei (rendszerek milliói, mindenféle ember/gép, általában töredékes, dinamikus információhalmaz

Narrow casting” - (az MI egyik új kihívása): az emberek és az ágensek ama szűk halmazának az identifikálása, akik a hirdetett információban vagy szolgáltatásban érdekeltek, vagy fordítva: egy ilyen szolgáltatást képesek nyújtani.

Cyberspace-ben

- információ gyűjtő brokerek,
- kereső rendszerek,
- Web robotok –
 - információ szűrők,
 - Usenet News szűrők,
 - elektronikus levelezés szűrők, ...
 - információ szolgáltatások: elektronikus könyvtárosok, adatbázis-kezelők
- **szoftver robotok**
- **operációs rendszer ágensek (naplózás figyelés)**

Örökölt rendszerek (legacy systems)

- korábban telepített rendszerek ágens-ített változatai: pl. távközlési hálózatok menedzsmentje, számítógéppel vezérelt gyártás, szállításszervezés, légi irányítás, orvosi rendszerek, betegfelügyelet kritikus feladatait végző rendszerek. A változó ipari/üzleti feltételek miatt ezeket a rendszereket frissíteni kellene, csak az átírásuk túlságosan költséges. „**Agent wrapper**” csomagolás, hogy az új rendszerekkel legyen képes együttműködni.
- FTP ágensek,
- **mérnöki ágensek**
- **digitális könyvtárak** - könyvtár, mint az információs ágensek közössége:
 - 'user-interface' ágens, 'user profile' ágens, adatbáziskérés megfogalmazása, adatbázis (információgyűjtő hely), interface ágens, információs szolgáltató ágens, thesaurus ágens, ...

Tipizált Üzenetváltó ágensek (*Typed-Message agents*):

- csoportként definiálva, csoportnak információt kell cserélnie a taszkok végrehajtásához és ehhez

megosztott protokollal kell rendelkeznie, ahol az **üzenetszemantikának** legalább egy részét tipizálni kell feladat független módon.

MI, ágensek a Web-en

Intelligens ágensek alapvető feladatai

- **információ keresés problémája**
- **biztosító ágens (*supply*)** ⇒ eredeti
⇒ közvetítő - indexelő, átalakító, lefordító, ...
- **igénylő ágens (*demand*)**
- **alkusz ágens (*broker*)**
- információ “eladása” - profit: “olcsón” venni
- “vásárlása” - profit: “drágán” eladni

Részletes példa:

hatékony információ alkusz (igényes MI feladat)

- segítség (ágensnek, embernek, ehhez kapcsolattartás kell) a kérés **megfogalmazásánál**
- információs források **azonosítása**,
- kérés megválaszolásának **megtervezése**, adott információs források mellett,
- a terv **végrehajtása**,
- a válaszok megadása kliensnek: (**interpretált**)
 - a válasz hogyan kapcsolódik a kéréshez,
 - fogalmak definíciója,
 - alternatív kérések ötlete.
 - **kliens ágens / ember** nem ismeri: az információs források nyelvét és hozzáférését,

ágens rendszer fogalma

- a **környezetének szerepe:**
 - a környezet a cél és a feladat forrása, sikeresség megmérettetője, az ágens intelligenciája “környezet-relatív”,
 - környezet felskálázása \Rightarrow ágens/ember társadalom (**e-society**),
- **deliberatív képességek:**
 - tudás (akár bizonytalan) kezelése,
 - cselekvés szintű “előrelátás” – tervkészítés, de itt most az információszerzést kell megtervezni,
 - tanulás - megerősítéses
 - tudási-ntenzívítás,
 - technológiai és társadalmi **tranziensek**.

Ami erre ráépül

Ágens (advanced) tulajdonságok - viselkedés, képességek

Ágensközösségek és a problémái – együttműködés, konfliktusok

Ágensmodellek – tudás és képesség modellezése

Nyelvek és protokollok

Architektúrák, eszközök, szabványok, ...

- Általában egy-egy fejlesztés kissé eltérő meghatározáshoz vezet, de: (közös vonások az 'ágens' szó hagyományos értelmezéséből):
- valami, ami **cselekszik**, vagy képes cselekedni,
- valami, ami **mások helyett**, de azok beleegyezésével cselekszik.

Ágensek egy lehetséges osztályozása (ágensek világa):

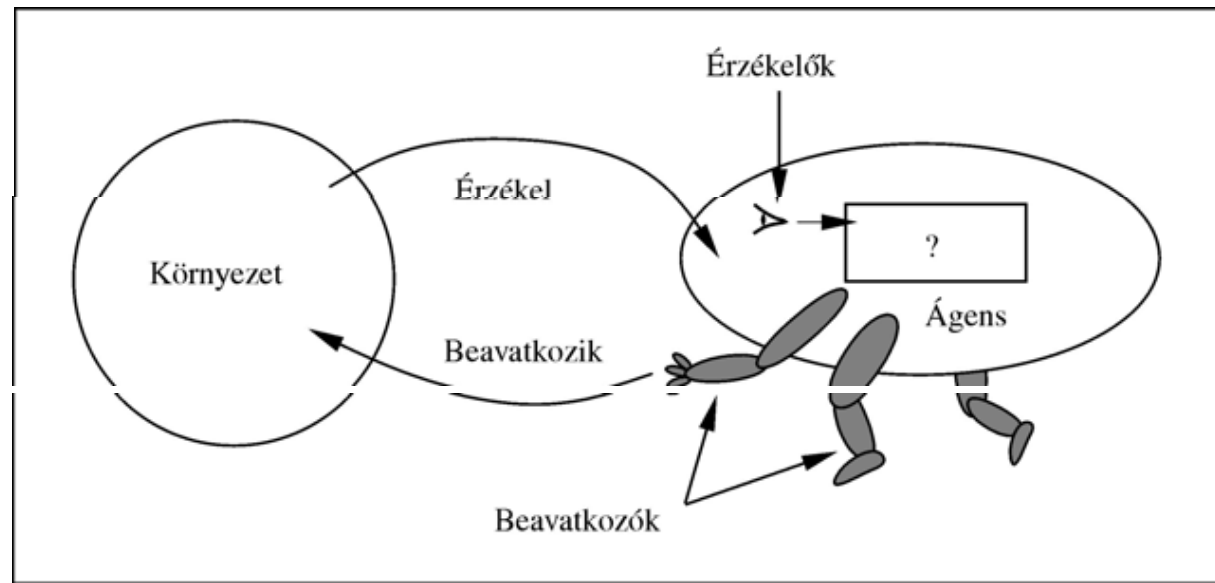
- **autonóm ágens:**
 - biológiai ágensek
 - **robotikus ágensek**
 - számítási ágensek
- 'artificial life' ágensek
- szoftver ágensek
- vírusok
- **szórakoztató ágensek**
- **taszk-specifikus ágensek**

továbbá:

ÁGENS RENDSZEREK

Autonóm ágens (egy lehetséges definíció):

- egy környezetbe beágyazott és a környezet részét képező rendszer, amely ezt a környezetet érzékeli, és ennek függvényében, időben folyamatosan, agendájának megfelelően cselekszik úgy, hogy ezzel a később érzékelt környezet állapotát befolyásolja.



Lényeges tulajdonságok:

- Az autonóm ágens egy **környezetbe van beágyazva**, abból **kiemelve**, más környezetben (amihez az érzékszervei 'nem jók') **nem ágens**;
- **környezetfüggő** az ágens **racionalitása**, és így '**intelligencia szintje**' is;
- minden szoftver ágens egy program, de nem minden program egy ágens: **lényegi különbség**, hogy az **ágens kimenetei (célorientáltan) befolyással vannak a későbbi bemeneteire**, és hogy a működése időben 'nyújtott'.

Intelligens ágensek elmélete

- 'Gyenge' ágens definíció

(az un. 'gyenge' MI nyomán: az intelligens viselkedés csupán külső reprodukálása)

- **Alapvető tulajdonságok:**

- **'kitartó' (*persistent*)** - folyamatosan konzisztens belső állapottal rendelkezik;
- ***autonóm*** - nagy fokú kontrollt gyakorol a saját belső állapota és akciói felett;
- ***önálló*** - direkt emberi beavatkozás nélkül működik;
- ***reaktív*** - érzékeli a környezetének változásait és reagál azokra;
- ***szociális*** - kapcsolatban áll emberekkel, ill. más ágensekkel;
- ***kommunikál*** - képes információt cserélni más rendszerekkel;

Az ágens aktív, nemet is mondhat, lazán csatolt.

- **Formálisan, tartalmilag, temporálisan független a válaszában.**
- **A kommunikáció egy cselekvésfajta,** következtethető, tervezhető, ...
- A kommunikációnak szemantikus értelmezése van: az üzenetet meg kell érteni (kontextusában), miért jött, miért ilyen, hogyan kell rá reagálni...

Kiegészítő tulajdonságok ('magasabb intelligencia' irányában):

- • **kezdeményezés**: - (*pro-active, data-directed execution*) cél-orientált, opportunista viselkedés (felhasználói feladat hiányában maga fogalmazza meg a feladatait);
- • **mobilitás** - képes helyről-helyre mozogni, megtartva saját belső állapotát;
- • **következtetés** - alapvetően logikai tudásreprezentációval dolgozik, tehát tudnia kell logikai módon következtetni;
- • **tervkészítési készség** - a fentiekből értelemszerűen következik;
- • **tanulás, adaptáció**;
- • **párbeszéd** - ahhoz, hogy lássuk, hogy ágens megosztja-e a céljainkat és képes azokat megvalósítani, párbeszédre van szükség, amely tisztázza az intenciókat és képességeket; a párbeszéd eredménye a megegyezés;
- • **igazmondás** - szántszándékkal nem hazudik a környezetének (segítőkéssz, ha többen vannak);
- • **jóindulat** - megkísérel teljesíteni mások kéréseit;
- ...

- **Erős' ágens definíció** ('erős' MI nyomán - a 'tisztá' intelligencián túlmenően, más 'emberibb' vonások is)
- ... a fentieken túlmenően, az ágens koncepciójának kidolgozása, megvalósítása olyan elvek alapján, melyeket többnyire emberekre alkalmaznak, mint: tudás, vélemény, szándék, meggyőződés, kötelességtudat, ... esetenként teljesen 'emberi érzelmekkel' modellezik az ágenst:
- **emocionális ágens.**

- **Intencionális megközelítés** - nagyon bonyolult rendszerekről úgy is beszél(het)ünk, hogy képesek vagyunk a viselkedésüket megmagyarázni vagy megjósolni anélkül, hogy a működésüket megértenénk.
- ...
- Informatikai rendszer intencionális modellje: ...
- Ember intencionális modellje: ...
- ***Egy modell legyen a lehető legegyszerűbb, de annál egyszerűbb azért mégsem (Einstein)***

Elméletek, architektúrák és nyelvek

Ágens, MAS – Multi Agent Systems

- **Koordinálás, együttműködés,
konfliktusfeloldás**
- **CMAS – Cooperative MAS
(együttműködő)**
- **SMAS – Self-interested MAS
(kompetitív)**

Ágenselméletek

ágens modell - alapvetően egy logikai modell

Modellezni kellene:

- **információs attitűdök**

- tudás
- hiedelem (*belief*)
- kölcsönös információs attitűd

- **pro-attitűdök**

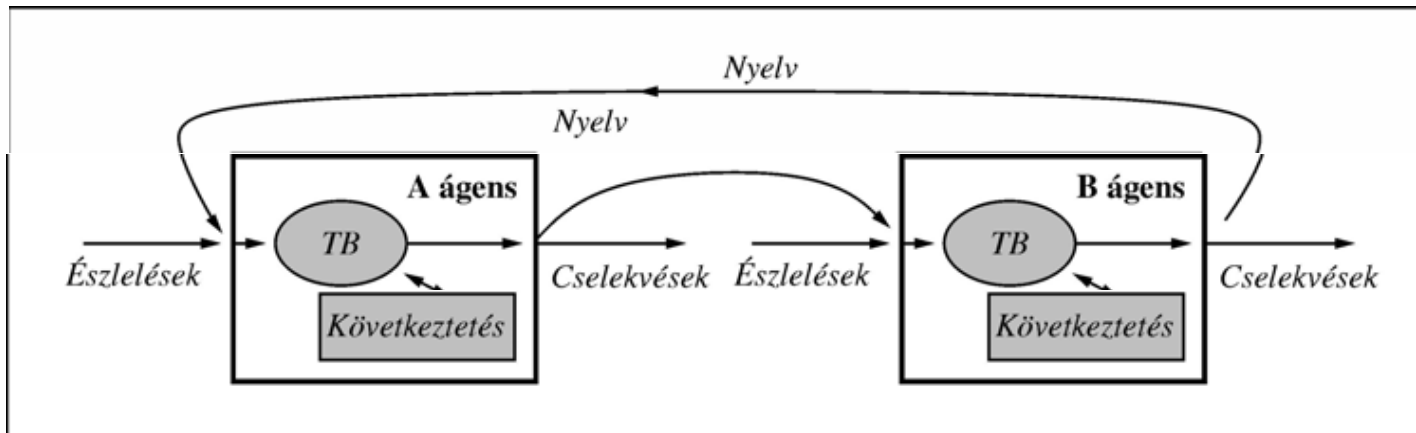
- kíváncsiság
- szándék
- kötelesség
- kötelezettség
- választás
- kollektív pro-attitűdök

- **információs attitűdök (hozzaállások) modalitásai**

modalitás (elsőrendű logika kiterjesztése), például:

- '*Júlia azt hiszi, hogy P* ' (*)
- egy naiv átírása logikai nyelvre: '*Hiszi(Júlia, P)*'
- de ez csak látszólag egy predikátum kalkulusbeli állítás, mert az egyik argumentum (P) már önmagában is egy logikai állítás, és ez nincs megengedve;
- Predikátum kalkulus állításai az un. extenzionális állítások, 'igazságfukcionálok'. Igazságértékük csak a bennük szereplő termek igazságértékétől függ.
- **Teljesen világos, hogy az (*) állítás értéke nem függ a P állítás logikai értékétől.**
- Ha nem tudunk dolgozni az elsőrendű logikával, akkor azt ki kell terjeszteni (modális logikák, modalitások), viszont akkor annak a bizonytalan eldönthetősége el fog veszni. Megoldás: modális operátorokkal kiterjesztett logika, amely más vonatkozásaiban viszont leszűkített, hogy a bizonyítás kezelhető legyen.

- **pro-attitűdök modálításai**
 - **cselekvések** reprezentációja
 - **idő** reprezentációja
- Nyelv:
- audio jelzés
- üzenetforgalom
- fizikai jezés (süketnémák nyelve, gesztusok, mimika, ...)



Ágens architektúrák

- **BDI – Belief, Desire, Intention** (gyakorlati humán következtetés, 1987) , az ágens egy célorientált entitás, racionálisan cselekszik.
- **Belief – Hiedelem – információs attitűd**, az ágens által tudné vélt információ reprezentálása.
- **Desire (Goal) - Kíváncsalom (cél) – motivációs attitűd**, célállapot explicit reprezentációja – a cselekvés végrehajtásának az oka.
- **Intention (Plan) – Szándék (terv) – deliberatív attitűd**, a cél elérésének eszközei.
- Hibrid, *Érzelmi*, **Modális logikai modellek..**

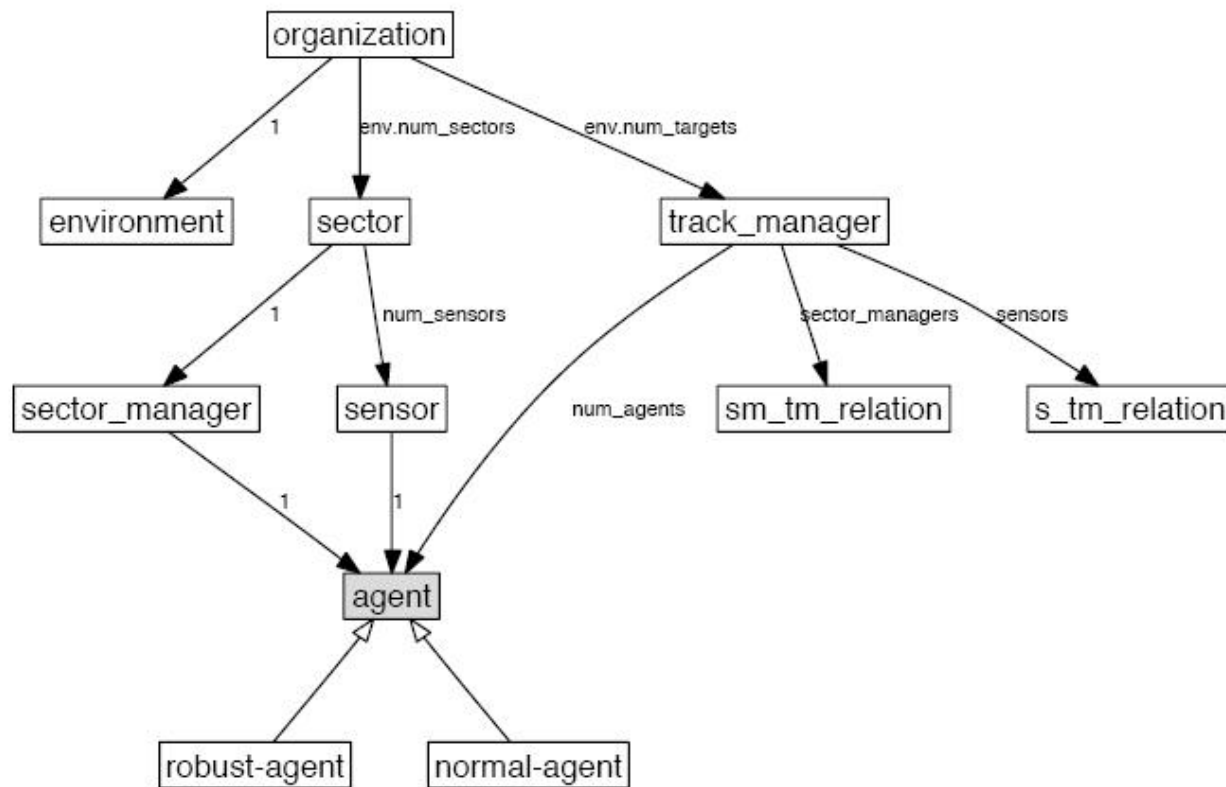
Ágens nyelvek

- Intelligens ágensek együttműködése
- kommunikáció - a tudás és az információ kicserélése (**KQML** – Knowledge Query and Manipulation Language)
- kooperáció; - az együttműködés lehetséges formái
- koordináció; - az együttműködés összehangolása, tervezés

Ágens implementációk

- Milyen „szervezethez” fog tartozni (milyen legyen a „szociális” kézségkészlete),
- kikkel lesz kénytelen kapcsolatban lenni.
- az un. **Platform Szolgáltatások**
 - Normatív
 - Opcionális ágensek.

Elosztott szenzor hálózat példa



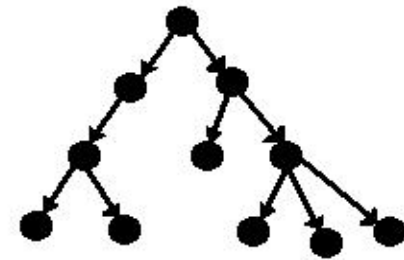
Ágensszervezetek – Multi Agent Systems

- **Szervezet:** szerepek, relációk, tekintélystruktúrák együttese - konkrét *célt* szolgál
- absztrakt: **valamilyen jellegű komplexitás csökkenése**
 - (munka, következtetés, kommunikáció, keresés, ...)
- kialakítása, jellemzői a rendszer viselkedését befolyásolja
- *mérsékelt kölcsönhatások*
- „*erő a sokaságban*”
- *bizonytalanság* csökkentése/lekezelése
- *redundancia* csökkentése, vagy explicit módon megnövelése
- olyan *magas szintű célok* formálizálása, amikről az egyedi ágensnek nincs tudomása
- növelt kommunikáció, számítási overhead
- *csökkentett* az eredő rugalmasság, *reaktivitás*
- egy új *komplexitási réteg*

- **Számítási szervezetelmélet**
(Nincs szervezet, ami alkalmas lenne minden helyzetben)
- **Szervezeti önkialakítás** – a rendszer dinamikus, adaptív módon szelektál és adaptál egy adott szervezeti sémát.
- A rendszerépítés módozatai:
 - Hierárchia
 - Holárchia
 - Koálició
 - Team
 - Kongregáció
 - Közösség
 - Federáció
 - Piac
 - Mátrix
 - Összetett ...

Hierárchia

– koncepcionális ágens-**fa**struktúra



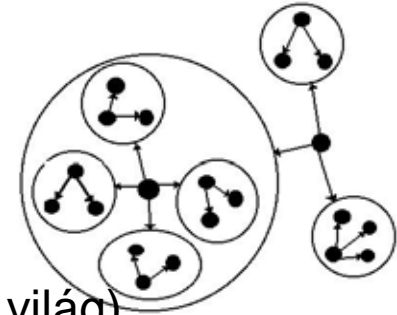
- magasabban lévő ágens – átfogóbb rálátás a problémára
- kölsönhatás – csak az összekötött entitások között
- fa alján előállított **adat** tipikusan halad **felfelé**, szélesebb rálátást biztosítva
- vezérlés lefelé** halad, magasabb szintű ágensek utasításokat adnak alsó szintű ágenseknek

Hierárchia típusai:

- **egyszerű** – egy db döntéshozó gyökér
- **egyenletes** – a tekintély delegálva van a rendszer különböző részeibe:

– a döntés ott történik, ahol kellő adat áll rendelkezésre a következtetéshez és kellő szervezeti hatalom van a döntéshez:

- minden szint = **szűrő**, csak a szükséges adatot (*explicit*) és parancsot (*implicit*) engedi át felfelé a hierarchiában
- **több-felosztású** – a szervezet bontása „termékek” mentén (pl. szolgáltatások), minden részlegnek teljes kontrollja van az ő terméke felett,
 - egyszerűbb döntéshozatal,
 - csökkentett kölcsönhatások.

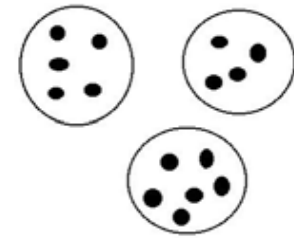


Holarchia

- Holon = *holos* (egész) + *on* (része)
- ágyazott, önismétlő szervezetek (fizikai, biológiai, társadalmi világ)
- többszintű csoportosított hierarchiák
- minden csoportosítás jellege a tagjaiból ered, de tőlük eltérő
- minden holon – egyedek halmaza, de egyben egy magasabb halmaz eleme.**
- A parancs-lánc felfelé halad

Jellemzők:

- heterogén** – egyedi holonok maguk döntenek el a taszkok végrehajtását
- A holonok **autonómok** abban, hogy kell legjobban teljesíteni a rájuk bízott taszkat
- rugalmasság** a viselkedés megválasztásában, konfliktusos taszkok menedzselésében
- A holon viselkedés adaptív, alakul a körülményekhez, pótlagos koordinálás nélkül mérsékli a megbízó tudásigényét
- Az ilyen tudás hiányában **nehéz az egész rendszer hatékonyságát megjósolni!**



Koalíció

- ágenshalmaz minden részhalmaza lehet koalíció
- **rövid életű, célorientált**
- szerveződés a cél érdekében, a cél megszűnésével a koalíció megszűnik
- ágens „dezertálása” – a koalíció a megjelölt célját már nem biztos, hogy teljesíti, a kritikus tömeg elfogy
- keletkezhet **együttműködő**, de **önérdekelt** ágensek között is
- **lapos** struktúra, ám lehetnek „vezető ágensek” is
- formálása után a koalíció tekinthető egy **egyedi, atomi** entitásnak
- Lehetségesek az átlapolódó koalíciók.

Jellemzői:

- **motiváció:** legalább néhány résztvevő értéke valamilyen dimenzió mentén *superadditív*
- részvételi költségek *szubadditív*ak lehetnek
- **egyedhasznosság** maximizálása **csoportos munkával**

Koalíció *alakítás* problémája

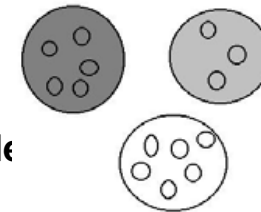
- olyan ágensrészhalmaz megválasztása, amely koalíció hasznosságát maximálja (érték mínusz költség)

Team

- **együtműködő** ágensek csoportja, akik megegyeztek, hogy együtt dolgoznak egy **közös cél** érdekében
- **teamhasznosság** maximizálása
- **elvárható koordináció** = ágensek cselekvései konzisztensek a céllal és azt támogatják
- egy ágens – egy vagy több szerep, eseményfüggő, a magassintű célok általában relatíve konzisztensek maradnak
- tipikusan reprezentálandó: *megosztott célok, kölcsönös hiedelmek, teamszintű tervek*
- **robosztusság, rugalmasság,**
- **Jellemzők:**
 - A team **nagyobb célokat** képes elérni, mint az egyedi ágens
 - **Rugalmassága biztosítja a** munka hatékonyságát bizonytalan környezetben, előre nem látott feltételek mellett
 - Következtetési képesség az ágensek közötti kölcsönhatások következményeivel kapcsolatban
 - **Erőteljes kommunikáció**
- **Alakítása nagy kihívás:**
 - magas célok kielégítéséhez hogyan kell hozzárendelni az ágenseket
 - végrehajtás alatt megtartani ezen **ágensek konzisztenciáját**
 - **teamstruktúra felülbírálása**, ahogy az ágensek és a körülmények változnak
 - Mi történik, ha ágens team-et alakít?
 - ismerni kell a potenciális résztvevőket, biztosítani az **ágensfeldedezést**.



Kongregáció (szín: feltehetően más érdekszférák)



- **lapos laza** szervezet, további hasznok reményében
- alapvetően (ez feltételezés) **hosszú életű** és **nem egy konkrét cél érdek**
- hasonló, vagy akár komplementer ágensek kapcsolata
- az **együttműködő partnerek könnyebb megkeresése**
- egyes ágenseknek **nincs** szükségszerűen **közös** vagy rögzített célja, de kell egy **stabil képesség-halmaz**, vagy követelmény-halmaz
- az ágensek **racionálisak**, lokális hosszú távú hasznosságukat maximalizálják (nincs szó csoportos, vagy globális díjazásról, megerősítésről)
- Az ágens **igénye lokális** (magán) **hasznosság növelése** – csatlakozás valamely kongregációhoz
- Az ágensek jönnek, mennek, de szükséges egy **minimális kritikus tömeg**, hogy a kongregáció teljesítse a célját
- elegendő haszonnövekedés a kongregációból, ami a keresés, csatlakozás, stb. költségeit túlszárnyalja
- A kongregáció előnye: a keresési **komplexitás csökkentése**, **kölcsönhatások mérséklése**
- más kongregációkhoz tartozó ágensek **nem kommunikálnak**, de nem szükségképpen kizáróak.
- *Jellemzők:*
 - könnyebb megtalálni partnereket, a keresendő **populáció nagysága csökkent**
 - nagyobb a **hasznosság, megbízhatóság a szorosabb kötelékek** révén kongregáción belül
 - nagyobb az **információ-megosztás** kongregációtagok között
 - **büntetés az ágenseknek** az elkötelezés csökkentéséért
 - A kongregáció **nem** szükségképpen tartalmazza a kölcsönhatás szempontjából **optimális** partnereket
 - A kongregáció: **kompromisszum** – idő és költség csökkentése, minőség, rugalmasság mérséklése a cél
- *Alakítás:*
 - **önérdekű cselekvés** – lokális hasznosság növelése, kudarc valószínűségének csökkentése

Az ágensek modellezéséről (általában)

- A feladat nehéz
- Melyik modell? Egyféle? Hibrid?
- Milyen tulajdonságok? Megfigyelhetők?
- Hogyan? Múltbeli? Játék közben? Időintervallumban?
- Rekurzív modell? (Ő is modellez engem)
- Hogyan építsem? (folyamatosan, következő epizódban)
- Félreismerés? (zaj, eltérő domain szemantika, dinamika helytelen értelmezése, ...)

Mit tartalmaz egy ágens modellje

- Megfigyelhető tulajdonságokat
- Szenzorikusan észlelhető tulajdonságokat
- Feltételezett hiedelmek halmazát
- Rekurzíók leírását
- Feltételezett cselekvéseket
- A jövőben végrehajtandó lehetséges cselekvéseket,
- Egyéb célokat, terveket, ...(erősen felépítésfüggő)

Osztályozás a modellezés folyamán

- **Megfigyeltek (A)**

- direkt módon megfigyelhetők valamilyen időintervallumban a megfigyelő (modellező) ágens által
- Értékek, időbeli változások

- **Kitaláltak (B)**

- nem figyelhetők meg, de következtethető, kiszámolható a szerepük, viselkedésük, (a priori tudás alapján).

Hogyan zajlik a modellezés (általában)

- Az **A** típusú információk frissítése a modellben viszonylag egyszerű: szenzorok adatait tároljuk.
- A **B** típusú információk előállítása nehéz ügy:
- **f**: (**A** információk, tudás az ágensről általában, korábbi **B** információk) \rightarrow **B**
- **f** vajon milyen függvény?

Fontos lehet:

- A terv felismerése (plan recognition)
- A terv azonosítása
- A viselkedés gráfok, predikciók, valószínűségi hálók építése
-
- és ezek kombinációi, ötvözött megoldásai.

Példák:

Online modellezés

- Pl. a játékelméletben minden forduló után újraértékel és új jóslást készít az ellenfél lépésére (pl. nyereségtáblázattal) Keveredik a modellezés és a játék (rekurzió).

Offline

- Csak a következő játékban használja a most szerzett tudást. Jobban értékelhető a modellezés, nincs rekurzió.

Hibaforrások félreismerése:

- A modellező ágens szenzor hibája
- A domain modellek eltérő volta, hibái
- A modellezett dinamizmusból fakadó hibák (pl. elmozdul, vagy változik)
- Rossz modell feltételezéssel él a modellező (pl. racionális – nem racionális)
- Szándékos félrevezetés
- Mások hibáinak átvétele

Főbb problémák:

- Rekurzió (A szerint B szerint A ...)
- Modellműködés komplexitása és időigény
- A megfelelő verifikáció és validáció hiánya
- ...