Kooperáció és intelligencia

Tanulás többágenses szervezetekben/2

Tanulás több ágensből álló környezetben
-a mozgó cél tanulás problémája (alapvetően megerősítéses tanulás)

Legyen az ágens közösség formalizált leírása az alábbi:

N – az ágensek halmaza, $i \in N$ egy konkrét ágenst jelent,

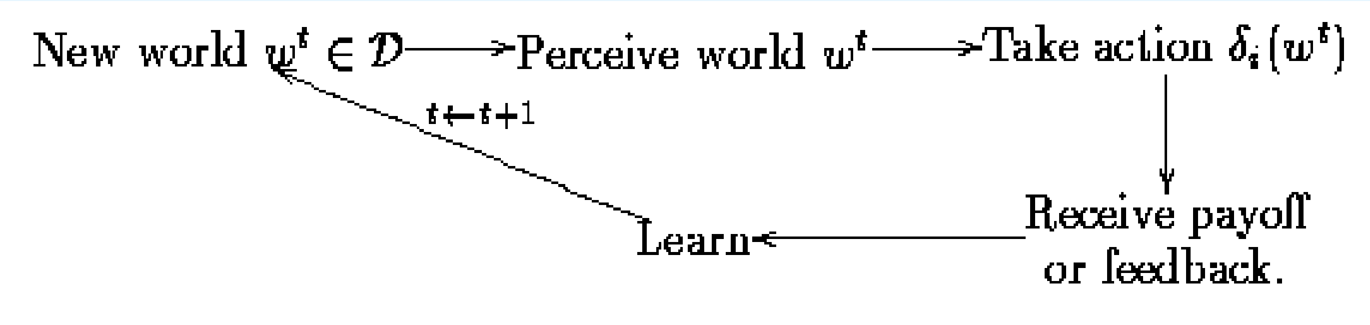
W – a világ állapotainak halmaza, $w \in W$ egy konkrét világot jelent,

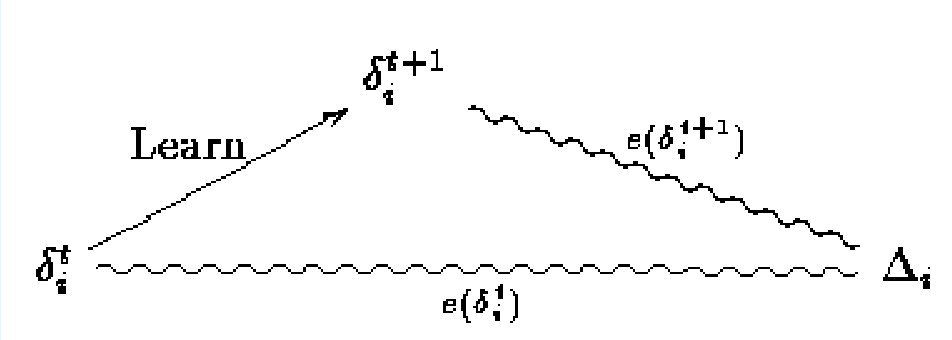
Ai – az i-edik ágens cselekvéseinek halmaza,

δ^t_i: W $\rightarrow$ *Ai* az i-edik ágens döntési függvénye,
hogy egy konkrét világállapotban milyen cselekvéshez folyamodjon,

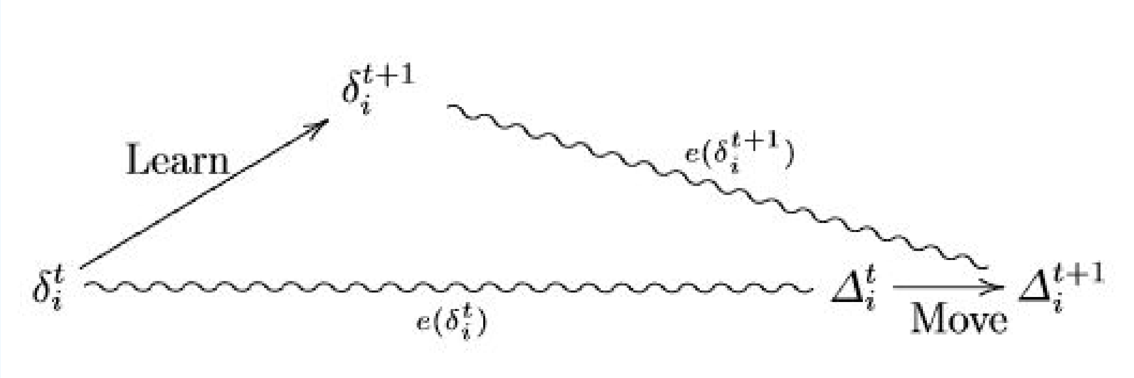
Δ^t_i : W $\rightarrow$ *Ai* az i-edik ágens cél (tanulandó ideális döntési) függvénye,

$e(\delta^t_i) = Pr\,(\delta^t_i(w) \neq \Delta^t_i(w) \mid w \in D)$ – az i-edik ágens hibája t-időben,
annak a valószínűsége, hogy az ágens rossz döntést fog hozni,
feltéve, hogy a környezeti állapotok egy *D* eloszlásból sorsolódnak,





A 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.



A döntési függvény δ^t_i lehet helyes, vagy hibás:

helyes volt	$\delta^t_i(w) = \Delta^t_i(w)$
hibás volt	$\delta^t_i(w) \neq \Delta^t_i(w)$
helyes lesz	$\delta^{t+1}_i(w) = \Delta^{t+1}_i(w)$
hibás lesz	$\delta^{t+1}_i(w) \neq \Delta^{t+1}_i(w)$
nem mozog	$\Delta^{t+1}_i(w) = \Delta^t_i(w)$
mozog	$\Delta^{t+1}_i(w) \neq \Delta^t_i(w)$
változik	$\delta^{t+1}_i(w) \neq \delta^t_i(w)$
nem változik	$\delta^{t+1}_i(w) = \delta^t_i(w)$
megtanulta a régit	$\delta^{t+1}_i(w) = \Delta^t_i(w)$
nem tanulta meg a régit	$\delta^{t+1}_i(w) \neq \Delta^t_i(w)$
új rácsúszott	$\delta^t_i(w) = \Delta^{t+1}_i(w)$
új nem csúszott rá	$\delta^t_i(w) \neq \Delta^{t+1}_i(w)$

A függvény megváltozása - a „tanulás” - javíthat, de ronthat is a dolgon.

A $\Delta^t_i(w)$ idő közbeni változása a „célpont” elmozdulása – a mozgó célpont.

Legyen akkor az ágens tanulási (helyes célfüggvény követési) hibája:

$$e(\delta^t_i) = \Sigma_w D(w) Pr(hibás\ volt)$$

Konkrét tanuló algoritmus helyett próbáljuk a **tanulás várható kivitelezhetőségét** valószínűségi alapon modellezni.

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

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

Definiáljuk a tanuló viselkedésére jellemző alábbi „mérőszámokat”:

- **változékonyság** (changing rate – c(i)):
  a helytelen leképzés „elmozdulása”, de mi felé?
  A helytelen leképzés javítása (helyes irányba történő változás a tanulás),

  c(i) = Pr(változik | hibás volt)

-**tanulás sebessége** (learning rate – l(i)):

  l(i) = Pr(megtanulta a régit | hibás volt)

$$\forall_w\ c_i = \mathbf{Pr}[\delta_i^{t+1}(w) \neq \delta_i^t(w) \,|\, \delta_i^t(w) \neq \Delta_i^t(w)]$$

$$\forall_w\ l_i = \mathbf{Pr}[\delta_i^{t+1}(w) = \Delta_i^t(w) \,|\, \delta_i^t(w) \neq \Delta_i^t(w)]$$

l(i) < c(i) **mert a tanulás csak változással érhető el,**
ha az ágens cselekvéseinek választéka bináris, akkor $l_i = c_i$, mert
a helytelen döntés alternatívája a helyes cselekvés,
(1 – l(i)) az a valószínűség, hogy a helytelen leképezés nem javul.

- **megtartás sebessége** (retention – r(i)): helyes volt, és helyes marad

r(i) = Pr(nem változik | helyes volt)

- **illó sebesség**, **illékonyság** (volatility – v(i)):

v(i) = Pr(mozog)

$$\forall_w \; r_i = \mathbf{Pr}[\delta_i^{t+1}(w) = \Delta_i^t(w) \mid \delta_i^t(w) = \Delta_i^t(w)]$$

$$\forall_w \; v_i = \mathbf{Pr}[\Delta_i^{t+1}(w) \neq \Delta_i^t(w)]$$

Foglalkozunk most az **ágens eredő hibájával**: $e(\delta_i^t) = \Sigma_{w \in D} D(w) \Pr[\delta_i^t(w) \neq \Delta_i^t(w)]$

A várható hiba t+1 időpontban függ attól,
hogy a célfüggvény változik-e (**a**), vagy sem, illetve,
hogy a döntési függvény t időpontban eredetileg helyes volt-e, vagy sem (**b**):

$$
\begin{aligned}
\Pr(E) = \; & \Pr(E \wedge a \wedge b) + \\
& \Pr(E \wedge a \wedge \neg b) + \\
& \Pr(E \wedge \neg a \wedge b) + \\
& \Pr(E \wedge \neg a \wedge \neg b) = \\
& \Pr(E \mid a \wedge b) \Pr(a \wedge b) + \\
& \Pr(E \mid a \wedge \neg b) \Pr(a \wedge \neg b) + \\
& \Pr(E \mid \neg a \wedge b) \Pr(\neg a \wedge b) + \\
& \Pr(E \mid \neg a \wedge \neg b) \Pr(\neg a \wedge \neg b)
\end{aligned}
$$

E = hibás lesz
a = nem mozog
b = helyes volt

$$
\begin{aligned}
&\mathbf{Pr}[\delta_i^{t+1}(w) \neq \Delta_i^{t+1}(w)] = \\
&\quad \mathbf{Pr}[\delta_i^{t+1}(w) \neq \Delta_i^{t+1}(w) \wedge a \wedge b] + \mathbf{Pr}[\delta_i^{t+1}(w) \neq \Delta_i^{t+1}(w) \wedge a \wedge \neg b] + \\
&\quad \mathbf{Pr}[\delta_i^{t+1}(w) \neq \Delta_i^{t+1}(w) \wedge \neg a \wedge b] + \mathbf{Pr}[\delta_i^{t+1}(w) \neq \Delta_i^{t+1}(w) \wedge \neg a \wedge \neg b],
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{Pr}[\delta_i^{t+1}(w) \neq \Delta_i^{t+1}(w)] = \; &\mathbf{Pr}[a \wedge b] \cdot \mathbf{Pr}[\delta_i^{t+1}(w) \neq \Delta_i^{t+1}(w) \mid a \wedge b] + \\
&\mathbf{Pr}[a \wedge \neg b] \cdot \mathbf{Pr}[\delta_i^{t+1}(w) \neq \Delta_i^{t+1}(w) \mid a \wedge \neg b] + \\
&\mathbf{Pr}[\neg a \wedge b] \cdot \mathbf{Pr}[\delta_i^{t+1}(w) \neq \Delta_i^{t+1}(w) \mid \neg a \wedge b] + \\
&\mathbf{Pr}[\neg a \wedge \neg b] \cdot \mathbf{Pr}[\delta_i^{t+1}(w) \neq \Delta_i^{t+1}(w) \mid \neg a \wedge \neg b].
\end{aligned}
$$

Pr(hibás lesz) = ahol az ilyen körülmények minden változatára meg kell
vizsgálni a hiba alakulását (közben Bayes tételhez is folyamodtunk):

A négy db valószínűség most:

$$\mathbf{Pr}[\delta_i^{t+1}(w) \neq \Delta_i^{t+1}(w) \,|\, \Delta_i^{t+1}(w) = \Delta_i^t(w) \wedge \delta_i^t(w) = \Delta_i^t(w)] = 1 - r_i.$$

Pr(hibás lesz | nem mozog ∧ helyes volt) = **1 – r(i)**

$$\mathbf{Pr}[\delta_i^{t+1}(w) \neq \Delta_i^{t+1}(w) \,|\, \Delta_i^{t+1}(w) = \Delta_i^t(w) \wedge \delta_i^t(w) \neq \Delta_i^t(w)] = 1 - l_i.$$

Pr(hibás lesz | nem mozog ∧ hibás volt) = **1 – l(i)**

$$\mathbf{Pr}[\delta_i^{t+1}(w) \neq \Delta_i^{t+1}(w) \,|\, \Delta_i^{t+1}(w) \neq \Delta_i^t(w) \wedge \delta_i^t(w) = \Delta_i^t(w)] = (r_i + (1 - r_i) \cdot B)$$

$$B = \mathbf{Pr}[\delta_i^{t+1}(w) \neq \Delta_i^{t+1}(w) | \delta_i^t(w) = \Delta_i^t(w) \wedge \Delta_i^{t+1}(w) \neq \Delta_i^t(w) \wedge \delta_i^{t+1}(w) \neq \Delta_i^t(w)].$$

Pr(hibás lesz | mozog ∧ helyes volt) = **r(i) + (1 - r(i)) B**
B = Pr(hibás lesz | mozog ∧ helyes volt ∧ nem tanulta meg a régit)

$$\mathbf{Pr}[\delta_i^{t+1}(w) \neq \Delta_i^{t+1}(w) \mid \Delta_i^{t+1}(w) \neq \Delta_i^t(w) \wedge \delta_i^t(w) \neq \Delta_i^t(w)]$$
$$= (1 - c_i)D + l_i + (c_i - l_i)F,$$

$$D = \mathbf{Pr}[\delta_i^t(w) \neq \Delta_i^{t+1}(w) | \delta_i^t(w) \neq \Delta_i^t(w) \wedge \Delta_i^{t+1}(w) \neq \Delta_i^t(w)]$$
$$F = \mathbf{Pr}[\delta_i^{t+1}(w) \neq \Delta_i^{t+1}(w) | \delta_i^t(w) \neq \Delta_i^t(w) \wedge \Delta_i^{t+1}(w) \neq \Delta_i^t(w) \wedge \delta_i^{t+1}(w) \neq \Delta_i^t(w) \wedge \delta_i^{t+1}(w) \neq \delta_i^t(w)].$$

Pr(hibás lesz | mozog $\wedge$ hibás volt) = **(1 - c(i)) D + l(i) + (c(i) - l(i)) F**

D = Pr(új nem csúszott rá | mozog $\wedge$ hibás volt)
F = Pr(hibás lesz | mozog $\wedge$ hibás volt $\wedge$ nem tanulta meg a régit $\wedge$ változik)

Pr(a ∧ b) = Pr(a | b) Pr(b)

Az eredő várható hiba akkor:

$$
\begin{aligned}
E[e(\delta_i^{t+1})] = E[\sum_{w \in W} \mathcal{D}(w) \mathbf{Pr}[\delta_i^{t+1}(w) \neq \Delta_i^{t+1}(w)]] = \sum_{w \in W} \mathcal{D}(w)(\\
\mathbf{Pr}[\Delta_i^{t+1}(w) = \Delta_i^t(w) | \delta_i^t(w) = \Delta_i^t(w)] \cdot \mathbf{Pr}[\delta_i^t(w) = \Delta_i^t(w)] \cdot (1 - r_i)\\
+ \mathbf{Pr}[\Delta_i^{t+1}(w) = \Delta_i^t(w) | \delta_i^t(w) \neq \Delta_i^t(w)] \cdot \mathbf{Pr}[\delta_i^t(w) \neq \Delta_i^t(w)] \cdot (1 - l_i)\\
+ \mathbf{Pr}[\Delta_i^{t+1}(w) \neq \Delta_i^t(w) | \delta_i^t(w) = \Delta_i^t(w)]\\
\cdot \mathbf{Pr}[\delta_i^t(w) = \Delta_i^t(w)] \cdot (r_i + (1 - r_i) \cdot B)\\
+ \mathbf{Pr}[\Delta_i^{t+1}(w) \neq \Delta_i^t(w) | \delta_i^t(w) \neq \Delta_i^t(w)] \cdot \mathbf{Pr}[\delta_i^t(w) \neq \Delta_i^t(w)]\\
\cdot (1 - c_i) D + l_i + (c_i - l_i) F.
\end{aligned}
$$

$E(\varepsilon(\delta^{t+1}_{i})) =$

$\Pr$ (nem mozog | helyes volt) $\Pr$ (helyes volt) $(1 - r(i))$

$+ \Pr$ (nem mozog | hibás volt) $\Pr$ (hibás volt) $(1 - l(i))$

$+ \Pr$ (mozog | helyes volt) $\Pr$ (helyes volt) $[r(i) + (1 - r(i))\, B]$

$+ \Pr$ (mozog | hibás volt) $\Pr$ (hibás volt) $[(1 - c(i))\, D + l(i) + (c(i) - l(i))\, F]$

ahol a kifejezések az ágens alábbi „hiba” problémáit jelentik:

- a helyes leképzést **megtartani nem tudja**,
- a helyes leképzést **nem tanulja meg**,
- a hibás leképzést **megtartja**, ill. **nem tartja meg, de rossz irányba változtatja**,
- **rossz irányba tanul**, **nem változik**, ill. **változik, de nem a tanulás irányába.**

Valószínűségek és a tényleges ágensek

A valószínűségek ugyanúgy takarják a konkrét tanuló algoritmust, mint a konkrét architektúrát és ágensprogramot, ill. a környezetnek hatását is.

További vizsgálathoz azokat a megfelelő valószínűségekbe kellene átönteni, melyek numerikusan az ágensek adottságainak felelnének meg.

Így a valószínűségi (hiba) egyenletek akár nagyon bonyolult, akár igen egyszerű összefüggésekhez vezethetnek.

Egyszerű szimuláció

ha a döntési, ill. a célfüggvény változik, akkor az új cselekvés megválasztása az A_i felett értelmezett egyenletes eloszlásból történjen, azontúl az ágensek tanulása legyen független, ilyenkor:

$$B = D = \frac{|A_i| - 2}{|A_i| - 1} \qquad F = \frac{|A_i| - 3}{|A_i| - 2}$$

$$\mathbf{Pr}[\delta_i^t(w) = \Delta_i^t(w) \wedge \delta_j^t(w) = \Delta_j^t(w)] = \mathbf{Pr}[\delta_i^t(w) = \Delta_i^t(w)] \cdot \mathbf{Pr}[\delta_j^t(w) = \Delta_j^t(w)]$$

$$E[e(\delta_i^{t+1})] = 1 - r_i + v_i \left(\frac{|A_i| r_i - 1}{|A_i| - 1} \right) + e(\delta_i^t) \left(r_i - l_i + v_i \left(\frac{|A_i|(l_i - r_i) + l_i - c_i}{|A_i| - 1} \right) \right)$$

egy lineáris összefüggést takar (**y = a x + b**, ahol az x a hiba és az y a jövőbeli várható hiba).

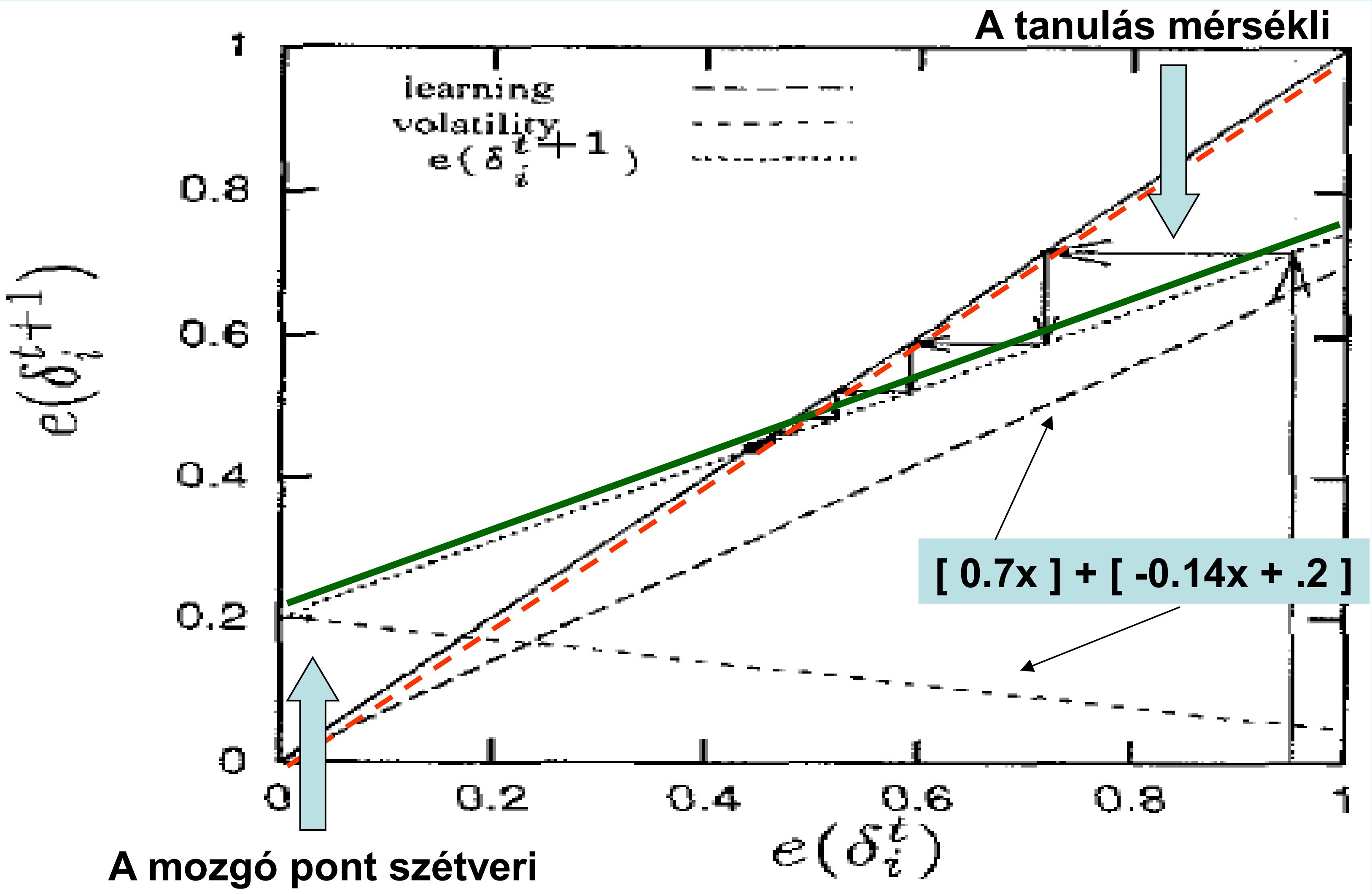
Az ágens hibája beáll a körülményektől függő szintre, amely az alábbi diagramból kiolvasható (v_i = 0.2, c_i = 1, l_i = 0.3, r_i = 1, minden ágens 20 cselekvéssel rendelkezik, a beállt hiba színt jelen esetben kb. 0.4):

$$E[e(\delta_i^{t+1})] = 1 - r_i + v_i \left(\frac{|A_i| r_i - 1}{|A_i| - 1} \right) + e(\delta_i^t) \left(r_i - l_i + v_i \left(\frac{|A_i|(l_i - r_i) + l_i - c_i}{|A_i| - 1} \right) \right)$$

Az összefüggésből és a diagramból meglátszik, hogy az ágens hibáját két „erő” formálja:

- **csökkenő** irányba a **tanulás** (l), és $y \cong v + x\,(0.7 - 0.7\,v)$
- **növekvő** irányba az **illékonyság** (v). $\cong [\,0.7x\,] + [\,-0.14x + .2\,]$

Mindkettő szempontjából a hiba kifejezés lineáris (tanulási egyenes pozitív, de $m < 1$ meredekségű, az illékonysági egyenes negatív meredekségű).
Az is látszik, hogy ha az ágens hibája nagyon kicsi, akkor a jósolt hibája majdnem egészében az illékonysággal magyarázható.


A tanulás mérsékli
learning
volatility
$e(\delta_i^t+1)$
[0.7x] + [-0.14x + .2]
A mozgó pont szétveri
0
0.2
0.4
0.6
0.8
1
$e(\delta_i^{t+1})$
$e(\delta_i^t)$

További vizsgálathoz vezessük be az **impakt** fogalmát:

$$\forall_{w \in W}\ I_{ji} = \mathbf{Pr}[\Delta_i^{t+1}(w) \neq \Delta_i^t(w) \,|\, \delta_j^{t+1}(w) \neq \delta_j^t(w)]$$

azaz az j-edik ágens döntésváltozása (mozgása) milyen hatással van az i-edik ágens tanulására.

Ennek segítségével meghatározható az eredő várható illékonyság, amivel viszont becsülhető a hiba várható alakulása az adott paraméterekkel (c_i, r_i, l_i) jellemzett ágensnél.

$$E[v_i^t] = 1 - \prod_{j \in N_{-i}} 1 - I_{ji}(c_j e(\delta_j^t) + (1 - r_j) \cdot (1 - e(\delta_j^t)))$$