

Tartu Ülikool  
Psühholoogia osakond

Roland Pihlakas

## **Klassikalise ja operantse mõtlemise modelleerimine**

Bakalaureusetöö

Juhendajad: Maarja Kruusmaa, Jüri Allik  
Läbiv pealkiri: naturaalse mõtlemise modelleerimine

Tartu 2007

## Kokkuvõte

Töö käsitleb klassikalise ja operantse tingimise modelleerimist kui naturaalse mõtlemise modelleerimise olulist alaülesannet.

Loodud mudel on aluseks edaspidistele teoreetilistele ja empiirilistele uurimustele. Lisas on kirjeldatud loodud mudeli võimalikke edasiarendusi ning ideid, mis võiksid aidata lahendada ülesandeid edasistes uuringutes. Tööga on kaasas mudeli lähtekood ning tabelitena mõned läbiviidud katseseeriad.

**Märksõnad:** tingimine, blokeerimine, kustumine, iseeneslik taastumine, diskriminantseid stiimulid, operantsed kinnituskavad, aheldamine, eeltingimine, taipamine.

Lisaks on kirjeldatud edasise arenduse võimalusi: tööriistade kasutamine, märgiline suhtlus ja mõtlemine, internaliseerimine.

## **Abstract**

The paper discusses the modelling of natural thinking, specifically the modeling of classical and operant conditioning as necessary part of the task.

The work gives a base for further theoretical and empirical research of these issues. In the addendum feasible extensions to the current realisation of the model are listed and general hypotheses for future research are posited. The source code and tables representing some performed test series are included.

**Keywords:** conditioning, blocking, extinction, spontaneous recovery, discriminant stimuli, operant reinforcement schedules, chaining, preconditioning, insight.

In the addendum are described possibilities of future development: tool use, symbolic communication and thought, internalization.

## Sisukord

Sissejuhatus .....	4
Eeldused .....	6
Mõtlemise tasandid .....	7
Piaget' sensomotoorse staadiumi alatasandid .....	8
Parker ja Gibsoni teooria primaatide füüsiliste võimete kohta.....	9
Thomas' õppimisvõimete hierarhia.....	10
Mudel .....	11
Stiimuleid ja tegevusi esindavate kanalite eeltöötlus.....	15
Klassifitseerija .....	15
Kajamälu .....	19
Assotsiatiivne sõlm .....	21
Klassikaline ja operantne tingimine .....	22
Klassikaline tingimine .....	23
Ennustusväärtuse lisatingimus.....	23
Valemid.....	24
Aeglase kustumise esimene valem.....	25
Blokeerimise valem .....	27
Aeglase kustumise teine valem, vastab blokeerimise nõudele .....	31
Iseeneslik taastumine .....	33
Operantne tingimine .....	36
Seadistuspunkti tegelike väärtuste dünaamika õppimine ning seadistuspunkti energia .....	36
Seadistuspunkti väärtuse ja energia ennustamine .....	39
Mitme tuleviku ajavahemiku operantne ennustamine .....	41
Operantse ennustuse energia mitme tuleviku ajavahemiku ennustamise korral .....	42
Mudeli õppimine ennustamiseks mitut tuleviku ajavahemikku .....	43
Diskriminantsed stiimulid.....	47
Kinnitavad stiimulid, aheldamine ja taipamine.....	51
Kokkuvõte .....	59
Viited.....	60

## Sissejuhatus

Kuigi peamiselt käsitlen klassikalist ja operantset tingitust ja tingimist, kasutasin töö pealkirjas sõna “mõtlemine”, kuna tingimine tähendab õppimist ja tingitus tegutsemist. Termin “mõtlemine” on lühike ning rõhutab mõlema aspekti olemasolu. Lisaks rõhutab sõna “mõtlemine” asjaolu, et mudel on paratamatult kognitiivsem, kui biheiviorism, mille toel see ehitatud on. Eriti kognitiivseks muutub mudel aheldamise ja taipamise peatükis.

Mõtlemise definitsioon Vögotski järgi:

- Mõtlemine on kogemuse seesmine organiseerimine
- Mõtlemine on tegevuse seesmine organiseerimine

Kirjutasin töö nõnda, et seda saaksid lugeda nii arvutiteaduse kui psühholoogia valdkonna inimesed. Seega olen mõningaid mõisteid selgitanud Lisas pikemalt lahti.

Psühholoogias üldteada ning seega pigem üldharivate mõistete tutvustused, samuti töö raames modelleeritud omaduste suhtes mittekeskseid teemasid (näiteks Toomela/Vögotski mõtlemise tasandid, mudeli potentsiaal edasiarenduse perspektiivis – mõned hüpoteesid) on tõstetud Lisasse.

Samuti on Lisas mudeli aspektid, mis tulenevad mudelist, kuid pole mudel ise (klassikalise tingimise diagrammid, operantsed kinnituskavad ja operantse tingimise diagrammid).

Artikli põhitekst sisaldab juba loodud mudeli põhjendamiseks ja iseloomustamiseks otseselt vajalikku probleemide analüüsi ning pakutud lahendusi.

Töö teine iseärasus on teemade hulk. Käsitlen põhjalikumalt ühte teemade põhigruppi ning lühidalt-ülevaatlilikult tervet hulka teisi teemasid, et positsioneerida mudelit ning näidata edasise arenduse võimalusi või ülesandeid. Ka põhjalikumalt käsitletud teemad võiksid igauks eraldi artiklit väärada, kuid minu eesmärk on hetkel:

- Leida originaalsed ideed omaduste teostamiseks, järgmisena võrrelda neid alternatiividega. See meetod võimaldab pikema viivitusega jõuda uute loovate lahendusteni.
- Veenduda, kas ja kuidas need elemendid koos töötada saavad. Leida elemendid, mis sobivad kokku, juhul kui selles osas on piiranguid. Lisaks, mõned elemendid ei ole üksi kasulikud, vaid ainult suurema komplekti sees.
- Leida üldine mudeli struktuur ja luua selleks sobiv programmi teostus. Käesoleval juhul tekkis tulemuseks muuhulgas ka paar metaprogrammi lakoonilisema, paindlikuma ja tähelepanu säästvama programmeerimise tarbeks.

Selle struktuuri loomise järel võib täpsustada puuduvaid detaile, võrrelda alternatiivsete intelligentsete süsteemidega.

- Luua tutvustav ülevaade seonduvatest küsimustest
- Vormistada omamoodi projekt edasisteks uuringuteks

Oma eelmises samateemalises kirjatöös (Pihlakas, 2005) loetlesin üles seniseid tehisintellekti uurimissuundi, tutvustasin arengulist lähenemist, kirjeldasin mõtlemise tasandeid Toomela ja Vögtotski järgi ning pakkusin välja mudelid mõtlemise tasanditele, mida nimetasin “1, 2a ja 2b” – mis võimaldasid mõelda sensorsetest üksitunnustest ning tunnuste kombinatsioonidest nii “alt üles” kui “ülalt alla” printsiibil. Töö lõpus loetlesin üles hulga tulevase edasiarendus- ning uurimisteemasid.

Ühest küljest pean oluliseks arengulist lähenemist, mis tähendab, et omaduste lisamisega tuleb tegeleda õiges järjekorras, teisest küljest pean oluliseks luua kõigepealt võimalikult üldine raam ja alles seejärel tulevikus pilti täiendada iteratiivselt detailsemaks minnes. Kohe üksikute detailidega põhjalikult tegelemine oleks nõ “evolutsiooniline” meetod ning ei võimaldaks *tõhusalt* tekitada süsteeme, kus üksikud elemendid ehk täiendused ei oma üksi alati märgatavat väärtust ehk ei anna vahetut tulemust uurimusartikli näol (eriti arvestades juba olemasolevate erinevate ideede rohkust tehisintellektis). Vastandatuna sellele, “intelligentne programmeerija” saab nõ ühe õhtuga luua programmi, kus on sadu omavahelisi sõltuvusi ning valida elemente lähtuvalt vajalikkusest – mitte populaarsusest, nende kasutatavuse sagedusest terves maailmas üle erinevate eesmärkide.

Arenguline lähenemine ongi siinkohal tasakaalustuseks, et ülalkirjeldatud konstrueerimisel oleks siiski toetuspinna all ja verstepostid märgitud. Rõhutaksin veel, et mudelit on siiski võimalik jagada osadeks ning arenduse käigus loomulikult testisin programmi osi eraldi ja järkjärguliselt. Arengulise lähenemise tutvustus on Lisas.

Seepärast jätkan käesolevat uurimust käsitledes esmajärjekorras kõige olulisemaid mõtlemise omadusi ja võimeid, mis on vajalikud ülejäänuteni jõudmiseks, järgides üldist arengulist suunda, kuid mitte koheselt kõikidesse võimalikesse teekonna kõrval asuvatesse teemadesse laskudes.

Käesolevas töös on kesksel kohal klassikaline ja operantne tingimine, primitiivse taipamise võime tekkimine (aheldamine), diskriminantsed stiimulid ning Lisas mitmest objektist mõtlemise võime tekkimine. Neile lisanduvad mõningad muud teemad ja võrdlused, mis on kasulikud käesoleva töö kesksete võimetega seoses silmas pidada.

Töö ei käsitle mudeli võimalikke seoseid närvisüsteemi arhitektuuri ja dünaamikaga.

## Eeldused

Mõtlemist modelleerides olen rakendanud mõned eeldused maailma kohta, kus mudel tegutseb. Need on:

- Tajude hierarhiline ülesehitus. See ei ole käsitletud käesoleva töö mudelis, kuid võib lugeda oluliseks eelduseks, mida naturaalse mõtlemise modelleerimises rakendada saab ja tasub. Seda eeldust on kasutatud näiteks Pihlakas, 2005 ja *Hierarchical Temporal Memory* (HTM) mudelis (George & Hawkins, 2005) <sup>1</sup>.
- Sündmuste ühesuunaline ajaline järgnevus ning asjaolu, et sama järgnevat sündmust võib põhjustada mitu alternatiivset eelnevat sündmust – põhjus ennustab tagajärge, kuid mitte vastupidi. See võib paista elementaarne, kuid mitte kõik võimalikud mudelid ei kasuta seda eeldust – näiteks Hebbi reegel – ning võivad enda kontekstis sellegipoolest head ja õiged olla.
- Mudel õpib, ennustab ja üldistab juba peale minimaalset kogemust. Täiendavad kogemused võimaldavad täpsustada seaduspärasid ning õppida erandeid. Taoline üldistamine toimub stiimulite tugevuste, kestuste ja ennustatavate tagajärgede määrade juures.
- Ennustused, mida mudel teeb, ennustavad esialgu sündmuse määra, mitte tõenäosust. Täiendava treenimise käigus on võimalik õpetada mudelile “tõenäosuste” ennustamist.
- Füüsilise maailma pidevus. Kui on teada kahe ühele skaalale paigutatava väärtuse vahel liikumise viis, siis saab seda rakendada ka ülejäänud selle skaala punktides asuvate väärtuste ehk *olekute* vahel liikumiseks. Analoogselt kolmanda eeldusega on võimalik õppida erandeid.
- On piisav ja isegi kasulik, et õpitakse ainult seoseid, mis vastavad blokeerimise ehk *nonredundancy* nõudele. Selle eelistest lähemalt Lisas.

## Mõtlemise tasandid

Eelmises kirjatöös (Pihlakas, 2005) pakkusin välja läheneda mõtlemise modelleerimise probleemile mõtlemise tasandite teooriate järgi.

Kuna käesolev töö keskendub Toomela kirjeldatud tasandite üldjaotuse järgi sensoorsete tunnuste kombinatsioonidest mõtlemise võimega mudeli loomisele, siis pole suurema hulga erinevate autorite poolsete kõrgemate mõtlemise tasandite kirjelduste võrdlemine siinkohal fookuses ning muutub aktuaalsemaks alles edaspidi.

Toomela kirjeldatud üldist tasandite jaotust saab vaadata detailsemalt Lisast. Töö põhitekstis toon Piaget sensomotoorse staadiumi alatasandite loetelu ning veel mõned analoogsed.

Et positsioneerida käesolevat mudelit mõtlemise arengu laiemas kontekstis ja näidata edasise arenduse suunda ja järjekorda, toon siinkohal lühidalt naturaalse ja kultuurilise mõtlemise tasandite loetelu Vögotski/Toomela järgi (Toomela, 2003b, 2005; Vygotsky & Luria, 1994b; Vygotsky, 1997):

Naturaalne mõtlemine:

1. Sensoorsed üksiktunnused
2. Sensorsete üksiktunnuste õpitavad kombinatsioonid
3. Objektid

Kultuuriline, märgiline mõtlemine:

4. Sünkreetsed märgid
5. Objektimärgid ja objekti omaduste märgid
6. Komplekssed mõisted
7. Teadusmõisteline mõtlemine
8. Süsteemne mõtlemine

Vajaduse korral vaata Lisast, mis erinevus on naturaalsel ja kultuurilisel mõtlemisel ning kultuurilisel mõtlemisel ja metamõtlemisel.

Käesolev mudel toetub suuresti Skinneri operantse tingimise kirjeldustele, ehk biheivioristlikele uurimustele. Vögotski ja Skinner sobivad minu meelest hästi kokku selles mõttes, et Skinner oskas teha lihtsaks suure tüki naturaalse mõtlemise kirjeldust – selleks piisas vaid stiimulite ja tegevuse aspekti heast kirjeldamisest – mida sai nüüd ka tõesti mudeli arenduseks hästi kasutada. Vögotski tuleb mängu peale naturaalse mõtlemise väljaarendamist ning näitab, mis moodi toimub üleminek naturaalsest mõtlemisest kultuurilisele ning edasi, ilma täiesti uut mudelit ehitamata. Mudel Vögotski teooriaid hetkel ei modelleeri, kuid see kontekst on kasulik silmas pidada.



### ***Piaget' sensomotoorse staadiumi alatasandid***

1. Refleksid
2. Primaarsed skeemid, suunatud oma kehale (või lihtsalt teatud liigutused, ilma suunda täpsustamata, näiteks kõndmine)
3. Sekundaarsed skeemid – suunatud välistele objektidele
4. Sekundaarsete skeemide koordineerimine – hierarhilised skeemid, eesmärgipärased vahend-eesmärk tegevused.
5. Kolmanda astme skeemid – tegevused suunatud objektide suhestamisele
6. Vaimsed representatsioonid

(Tomasello & Call, 1997; Miller, 1993)

Käesolevas mudelis 1. staadiumit esindab superviisor või mudelit treeniv inimene, seega mudeli õppiva osa väline süsteem. Mudelis on teostatud 2. ja 4. alatasand.

Mudeli edasiarendusena saab diskriminantsete stiimulite hulgast eristada selliste diskriminantsete stiimulite klassi, mis esindab objekte, millele parajasti mingi tegevus on suunatud, see eeldab mõningast visuaalset töötlust ning võimaldab 3. alatasandi.

Lisaks kirjeldan Lisas, kuidas mitme diskriminantse stiimuli sidumise võime ning 3. alatasandi sihtmärkstiimulite võime ühendamise tulemusena on tekib vähemalt mõningane tööriistakasutuse võime. Kui lisada objektide ruumiliste suhete taju, siis on saavutatud täielikult 5. alatasand. Selle alatasandi võimed on ühtlasi objekte omavahel suhestava tööriistakasutuse võimaldajaks (Tomasello & Call, 1997).

6. alatasandi võime kui objektide püsivuse võime peaks olema Tomasello & Call, 1997 järgi ka näiteks kodustatud koertel. Mudel seda alatasandit pigem ei käsitle, kuigi klassikalise tingimise viimases peatükis võib olla puudutatud osa seonduvast temaatikast.

Objektipüsivus on eraldi võime objektide manipuleerimise võimest: Mõnede uuringute järgi on inimestel võime manipuleerida objekte enne objektipüsivuse teket (Uzgiris & Hunt, 1974, viidatud: Tomasello & Call, 1997). Ning mõned uuringud on näidanud, et mõned objektipüsivuse aspektid esinevad inimestel enne kui neil on olnud mingit märkimisväärset kogemust objektide manipuleerimises (Baillargeon et al, 1985, viidatud: Tomasello & Call, 1997). Samale järeldusele vihjab imetajate, näiteks koerte, kes manipuleerivad objekte väga vähe, objektipüsivuse võime (Tomasello & Call, 1997).

Mõned uurijad on väitnud, et inimesed kasutavad probleemilahenduseks ja objektipüsivuseks representatsioone, samas tavalised ahvid kasutavad rohkem tajul põhinevaid strateegiaid.

Sellise eristuse muudab probleemsemaks, et ka koerad on võimelised järgima objektide

nähtamatuid ringipaigutusi – tekib küsimus, kas inimahvide 6-nda alatasandi kõik võimed on ikkagi representatsioonidel põhinevad. (Tomasello & Call, 1997). Igal juhul on see huvitav erinevus ja väärib edaspidist täpsustamist.

6. alatasandiga seondub võime mõista teatud tööriista ja objekti vahelisi suhteid ilma füüsilise manipuleerimiseta ning võime isegi tööriista muuta või valmistada, nõnda et see vastaks olukorra nõuetele (Tomasello & Call, 1997).

### ***Parker ja Gibsoni teooria primaatide füüsiliste võimete kohta***

1. ja 2. alatasand: lihtne mõistmine, käe-suu koordineerimine
3. Käe-silma koordineerimine, sekundaarsed tsirkulaarsed operatsioonid
4. Skeemide rakendamine üksikutele objektidele, näiteks toidu valmistamine
5. Objektide püsivus, kolmandad tsirkulaarsed operatsioonid, tööriistad
6. Vaadeldud uudsete tegevuskavade hilisem imiteerimine; tegevuste mentaalsed representatsioonid.

(Tomasello & Call, 1997)

Parker ja Gibsoni 3. alatasand on sisuliselt vastavuses Piaget 3. alatasandiga.

4. alatasand: erinevalt Piaget tasandite loetelust on siin nõue, et skeemid on rakendatud objektidele. Selle suhtes on Piaget tasandite loetelu eristavam, kuna skeemide rakendamise võime ja objektidele suunatud tegevuste võime on mõlemad mudelis sõltumatud.

5. alatasand: käesolevas töös ei ole uuritud objektide püsivuse nähtust. See võib ehk tekkida, kui mudel on piisavalt mahuka tajukogemusega ja seega saab ennustada objektide liikumist. Kuid see vajab eraldi kinnitust. Omadus võib eeldada ka tuge tajusüsteemilt, näiteks HTM-sarnase struktuuri olemasolu.

6. alatasand: tegevuskavade hilisem imiteerimine tähendab minu jaoks, et primaat on võimeline mõistma teise isendi liigutusi, ehk tal on olemas mirror-rakud<sup>2</sup>. Ta ei rakenda õpitud kogemust kohe (ei ahvi), vaid vajaduse tekkides ja eeldusel, et ta vähemalt ebamääraselt seostab, mis moodi õpitud põhjuse-tagajärje ahel talle momendi vajaduse suhtes kasulik on. Mudeli arendamise seisukohast tähendab see mirror-rakkude või mõne ligikaudse ekvivalendi lisamist (näiteks mootorikat kirjeldav raadiosignaal isendite vahel, kuigi see ei anna mudelile veel võimet õppida inimeste liigutustest).

## **Thomas' õppimisvõimete hierarhia**

1. Habituaatsioon
2. Signaali õppimine (klassikaline tingimine)
3. Stiimul-vastus õppimine (operantne tingimine)
4. Aheldamine (operantsed tegevusjärgnevused)
5. *Concurrent discriminations* / kaasnevad stiimulid (diskriminantstiimulid)
6. *Affirmative concepts* / Kinnitavad kontseptid: absoluutsed ja relatiivsed
7. Tingimuslikud kontseptid, konjunktiivsed kontseptid, disjunktiivsed kontseptid
8. Bikonditsionaalsed kontseptid: "siis ja ainult siis"

(Tomasello & Call, 1997)

Käesoleva mudeli puhul on olemas võimed kuni 5-nda tasemeni.

Thomas' järgi võimed kuni 5-nda tasemeni esinevad mingis vormis paljudel selgroogsetel. Pole ka erilist tõendusmaterjali, et need võimed oleksid primaatidel enam arenenud. 4-nda ja 5-nda taseme poolest on erinevus, kui palju palju aheldamisi või kaasnevaid stiimuleid liik suudab moodustada, kuid ka selles osas pole tõendeid, et primaadid oleksid eelisolukorras (Tomasello & Call, 1997).

6. taseme poolest on mudelis olemas absoluutsed kontseptid, juhul kui mudel ühendada tajusüsteemiga, mis on võimeline tajusid klassifitseerima (osa klassifitseerimist on ka järgnevas mudelis kirjeldatud). See võime peaks samuti imetajatel ühtmoodi esinema (Tomasello & Call, 1997).

6. taseme relatiivsed kontseptid – võimaldab kahe stiimuli võrdlemist ning on eksklusiivselt primaatide võimeks ja vastab Piaget tertsiarsetele operatsioonidele (Tomasello & Call, 1997). See ei ole hetkel mudelis. Oletatavasti on selle võime tekke eelduseks tähelepanu juhtiva protsessi lisamine: võime liigutada pilku vaheldumisi ühele ja teisele objektile ning sel toel õppida tajus toimuva muutuse kui teatud liiki stiimuli tähendust.

7. taseme kontseptid peaks olema esindatavad kui diskriminantsete stiimulite eriliigid, juhul kui mudel võimaldaks tegevusega siduda mitu diskriminantset stiimulit – omadus, mis on plaanitud mudeli järgmiseks edasiarenduseks. Need kontseptid ei nõua stiimulite võrdlemist, küll aga seda, et nii stiimuli olemasolu kui ka puudumine saaks omada tähendust (viimane aspekt on käesolevas mudelis olemas, puudu on mitme stiimuli ühendamine). Need kontseptid võivad samuti nõuda tähelepanu olemasolu, kuna tõenäoliselt ei asu vastavad objektid ühes tükis.

8. tase vajab edaspidist täpsustamist.

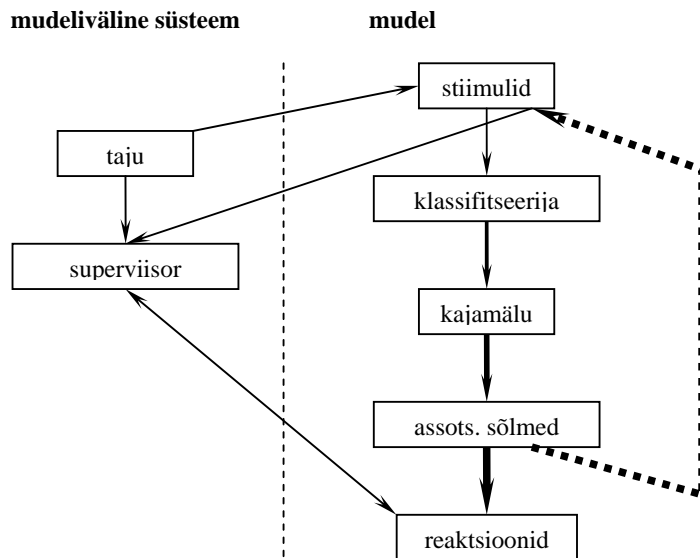
## Mudel

Järgnevalt toon kaks joonist mudeli üldise struktuuri kohta.

Mudel koosneb piirkondadest, mis talletavad sissetulevat infot ning suhestavad seda varasemaga. Piirkonnad on ühendatud kanalitega. Piirkondadest väljuv infovoog on enamasti suurema mahuga kui sissetulev – seda väljendab joonte jämedus. Erandiks on assotsiatiivsetest sõlmedest lähtuvad kanalid, mis võivad olla potentsiaalselt väga mahukad, kuid ei pea – seega viimaste joonte jämedus pole sama, mis tegelik kogemuste hulk. Praktikas võib oletatavasti kõige mahukamaks kujuneda  $S \rightarrow S$  seoste piirkond, ehk isendi osalusest sõltumatuid välismaailma põhjuslikke järgnevusi õppiv piirkond – neil isenditel, kellel see piirkond esineb ja on täielikult kasutuses (vaata peatükki: Kinnitavad stiimulid, aheldamine ja taipamine).

Jooniste järel selgitan lühidalt, mida piirkonnad teevad. Järgnevates artikli peatükkides kirjeldan detaile.

### Klassikaline tingimine:

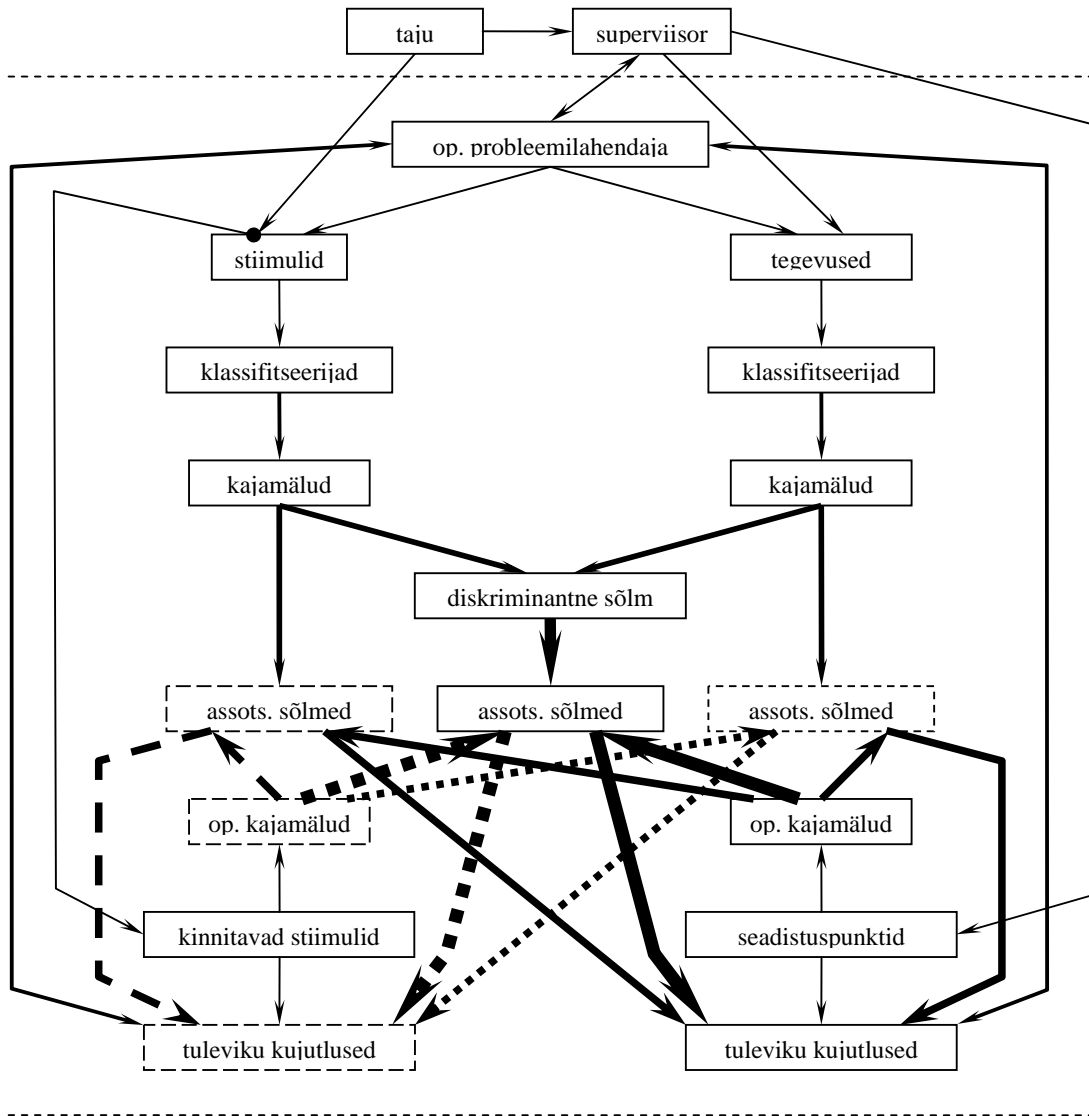


### Piirkondade selgitused:

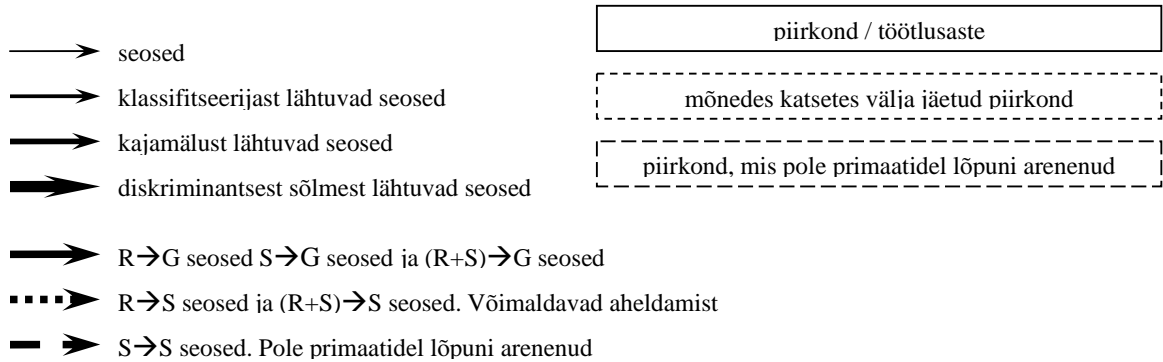
- Taju – stiimulid, mis saadetakse mudeli sisenditesse.
- Superviisor – süsteem, mis tekitab reaktsioone tajule (tingimata stiimulitele) ning treenib mudelit (mudel panus on, et see õpib tingitud stiimulite tähendusi).
- Stiimulid – mudeli sisendid.
- Klassifitseerija – võib osadel juhtudel mudelist välja jätta. Osadel juhtudel on tarvilik stiimulit omaduste järgi täiendavalt liigendada.

- Kajamälu – säilitab ajalise jälje varem esinenud stiimulitest või nende liigendustest. Võimaldab seostada hilisemaid reaktsioone ja stiimuleid varasemate sündmustega.
- Assots. sõlmed – assotsiatiivsed sõlmed, säilitavad infot stiimulite või nende liigenduste tähenduste ehk vastavate reaktsioonide kohta.
- Reaktsioonid – mudeli väljundid, mida superviisor treenib ja mudel ennustama õpib.
- Hetkel mudeli teostuses ei ole ühendatud seoseid assots. sõlmed → stiimulid, mis võimaldaksid teisest tingimist. Need seosed saab soovi korral lisada.

Operantne tingimine:



**Legend**



Operantse tingimise jooniselt on ära jäetud klassikalise tingimise seosed.

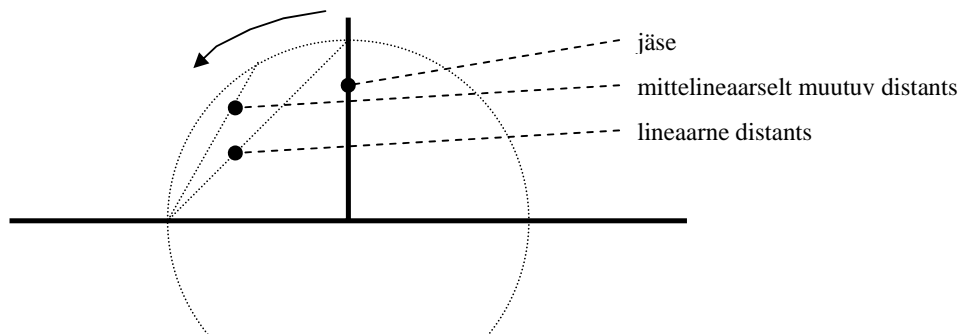
**Lisandunud piirkonnad:**

- Diskriminantne sõlm – seostab tegevuste ja diskriminantstiimulite kombinatsioone mudeli sisesteks stiimuliteks ehk representatsioonideks.
- Op. kajamälud – operantsed kajamälud. Säilitavad ajalise jälje seadistuspunkti või kinnitava stiimuli varasematest seisunditest, kuid töötavad natuke erineva printsiibiga kui stiimulite ja tegevuste kajamälud.
- Seadistuspunktid – esindavad näitajaid, mida mudel üritab hoida määratud tasemel. See tase määratakse väljaspoolt mudelit tulevas informatsioonis.
- Kinnitavad stiimulid – stiimulite täiendavad esindused, mis võimaldavad stiimulitel omandada kinnitava jõu.
- Tuleviku kujutlused – sisaldavad mudeli ennustusi seadistuspunktide ja kinnitavate stiimulite olekute kohta mitmel järgneval ajavahemikul.
- Op. probleemilahendaja – operantne probleemilahendaja. Tegeleb seadistuspunktide tegelike väärtuste eelistatud väärtustel hoidmisega ehk operantse eesmärgi suhtes optimaalsete tegevuste leidmisega ning kinnitavatele stiimulitele energiatega määramisega (selgitus vastavas peatükis).

## **Stiimuleid ja tegevusi esindavate kanalite eeltöötlus**

### **Klassifitseerija**

Stiimulite ja tegevuste tugevused on mudeli jaoks esitatud analoog-kodeeringus ehk reaalarvudena. Need väärtused klassifitseeritakse. Sooritatud operantse tingimise katsetes stiimuli tugevus osadel juhtudel tähendab stiimuli kestust (see arvutatakse superviisoris). Klassifitseerimine on kasulik, kuna võimaldab õppida tegevusi, mille tulemused pole lineaarsed. Näiteks kehtib see paljude ringliikumiste kohta, mille eesmärk on jäset või midagi muud viia soovitud punktini.

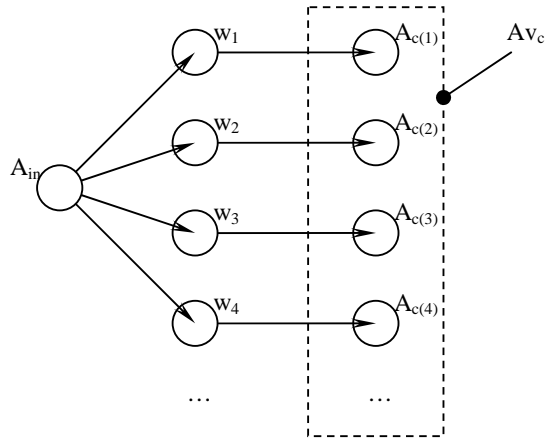


Hetkel tarvilik põhjus on ajaliste intervallide õppimine: mõned stiimulid või tegevused annavad tulemuse ainult siis, kui nad kestavad vähemalt teatud aja. Sel moel toimib ka näiteks intervall-kinnituskava, kus diskriminantstiimul (aeg viimasest kinnitusest) peab olema kestnud teatud aja, et talle järgnev tegevus tulemuslik oleks.

Klassifitseerijat sisaldavat mudelit on alati võimalik lihtsustada lineaarseks ehk mitteklassifitseerivaks mudeliks, seades klasside arvu võrdseks ühega – sisuliselt kaotades andmete töötlemise liinist klassifitseeriva komponendi. Samas poleks algusest peale lineaarsusega piiratud mudelit võimalik panna õppima mittelineaarseid seoseid.

Klassifitseerija ehitus:





$$\text{a) } A_{c(i)} = \begin{cases} 1, & \text{kui } i = \arg \min_j [\text{abs}(w_j - A_{in})] \\ 0, & \text{ülejäänud juhtudel} \end{cases}$$

$$\text{b) } A_{c(i)} = \begin{cases} 1, & \text{kui } w_j \leq A_{in} \\ 0, & \text{ülejäänud juhtudel} \end{cases}$$

$$A_v = (A_{c(1)}, A_{c(2)}, \dots, A_{c(n)})$$

$A_{in}$  – klassifitseeritav väärtus

$w_i$  – klassi  $i$  “kaal”

$A_{c(i)}$  – väljundkaalude vektor, kajastab hetke sisendi klassi või klasse

$A_{c(i)}$  arvutamise (a) variant on potentsiaalselt kasutusel juhul, kui  $A_{in}$  esindab tugevuse muutu; variant (b) on kasutusel juhul, kui  $A_{in}$  esindab stiimuli integraali või kestust või stiimuli tugevust, kuna:

- Integraali või kestuse puhul: tegevuse kestust saab ainult pikendada või lõpetada. Selleks, et õppida käimasoleva tegevuse (mitteoodatud) tulemusetust, tuleb õppida, et möödunud tegevuse kestused ei andnud tulemust. Kui sellist õppimist ei võimaldataks (a-variant), siis omistatakse tulemusetus jätkuvalt kasvavatele kestustele, mitte nendele, mis olid tegevuse motivaatoriks, ning mudel üritaks tegevust järjest edasi. Seega kuigi mudel algul üldistab pikemalt lühemale tegevuse kestusele, on võimalik õppida erandeid ning selles seisnebki näiteks intervall-kinnituskava.
- Stiimuli tugevust esindava kanali korral annab omaduse, et tugevamat stiimulit õpitakse rohkem. See tähendab siinkohal: tugeva stiimuli tähendust üldistatakse ka nõrgemale stiimulile, kuid nõrgema stiimuli esinemisel aktiveeritakse vastavalt vähem vektori komponente ja vastav ennustus tuleb nõrgem. Kui aga samaaegselt tingida nõrka ja tugevat stiimulit, siis omandab tugevam stiimul rohkem seoseid. Samas treenides nõrgemat ja seejärel testides tugevamat stiimulit, on mudeli ennustus, et nõrgema stiimuli treenimise järel tugevam stiimul annab sama tugeva reaktsiooni kui oli treenitud nõrgemale stiimulile. Selle vastavust tegelikkusele tuleb veel kontrollida. Kuigi mudel üldistab esialgu tugevamalt nõrgemale, on võimalik täiendava treenimise õppida nõrgematele stiimulitele eraldi tähendused (**kolmas eeldus**).

Võib oletada, et (b)-valem võiks olla kasutusel ka stiimuli tugevuse muutu esindavas kanalises. See vajab edaspidist täpsustamist, käesoleva töö mudeli katsetes olid kasutusel ainult kestust või tugevust esindavad kanalid. Samuti võimalik, et osades tajukanalites ka stiimuli tugevust esindav kanal kasutab valemit (a).

Kestust esindava tajukanali korral on hetkel eeldatud, et sisendi minimaalne väärtus on 0; vastasel korral tuleks täiendavalt jälgida mudeli toimimist ja teha vajadusel täiendusi valemisse.

Klassifitseerija andmepunktide  $w(i)$  väärtused määratakse mudeli töö alguses juhuslike väärtustena ning neid võib mudeli töö käigus muuta mõne printsiibi alusel, näiteks Kohoneni iseorganiseeruva kaardi (SOM) valemiga<sup>3</sup>, kuid see pole käesoleva mudeli töötamise seisukohast keskse tähtsusega. Need kaalud võivad ka kogu töö jooksul mitte muutuda ning hetkel mudeli ülejäänud kvalitatiivsed omadused selle tõttu ei kannata. Käesoleva töö katsetes olid kaalud fikseeritud.

Klassifitseerijaks võib olla SOM, Grossbergi ART<sup>4</sup>, generatiivne topograafiline kaart GTM<sup>5</sup>, HTM (George & Hawkins, 2005), analoogselt ka minu seminaritöös kirjeldatud hierarhiline struktuur (Pihlakas, 2005) või mõni muu superviseerimata õppimisele ning vektorit väljundväärtusena võimaldavaks kohandatav komponent.

On samuti mõeldav, et valemiga (a) analoogselt töötava klassifitseerija väljundkaalude vektoris aktiveeritakse mitu elementi.

Hetke mudelis on sisend skalaarne, kuid stiimulite mustreid töötlevas süsteemis oleks tegu mitmedimensioonilise sisendiga. Siis muutub ka olulisemaks klassifitseerija valik.

Juhul kui õppimise tööfaasis kasutada kaalude muutmiseks taolist valemit, mille korral kaalud koonduvad treenitava sisendväärtuse suunas (näiteks SOM), annaks see ilmselt ühtlasi habituatsiooni nähtuse, mis muuhulgas on seotud adaptatsiooni nähtustega visuaalses tajus<sup>6</sup>. SOM tekitab ajuga sarnaseid jaotusi kõige tõhusamalt (Honkela, 1997), seetõttu oleks esimene hüpotees kasutada edaspidi seda valemit.

Klassifitseerija asub mudelis enne kajamälu, kuna sel moel saab kajamälu väljundpunktis asuv vektor sisaldada mitu aktiveeritud elementi. Kui klassifitseerija paikneks peale kajamälu väljundeid, siis aktiveeriks klassifitseerija vaid ühe elemendi enda väljundvektoris ning see kaotaks liiga palju informatsiooni toimunud sündmuste kohta. Arenduse käigus selgus, et mudel töötab esimese variandiga, kuid kehvasti või üldse mitte teise variandiga.

<sup>3</sup> [http://www.scholarpedia.org/article/Kohonen\\_Network](http://www.scholarpedia.org/article/Kohonen_Network)

<sup>4</sup> [http://neuron-ai.tuke.sk/NCS/VOL1/P3\\_html/node27.html](http://neuron-ai.tuke.sk/NCS/VOL1/P3_html/node27.html)

<sup>5</sup> <http://research.microsoft.com/~cmbishop/downloads/Bishop-GTM-Ncomp-98.pdf>

<sup>6</sup> Adaptatsiooni nähtuse tutvustus: [http://en.wikipedia.org/wiki/Neural\\_adaptation](http://en.wikipedia.org/wiki/Neural_adaptation)

Ajalistel suhetel on kaks aspekti: stiimuli kestus ning stiimulite vaheline intervall. Stiimulite vahelise intervalli õppimiseks on mudelis kajamälu, mida kirjeldan järgnevalt.

## Kajamälu

On vajalik, et assotsiatiivne võrk saaks luua seoseid sündmuste vahel, mis järgnevad teineteisele erinevate ajaliste hilinevistega. On tarvilik põhjus-tagajärje seoste õppimiseks. “Kajamälu” tähendab siinkohal, et mälu sisaldab sündmuse kaja – vastava sisendkanali varasemaid seisundeid. Ka mõnedes teistes klassikalise tingimise mudelites (Balkenius & Morén, 1998) on kasutatud analoogset konstruktsiooni ning nimeks pandud *delay-lines*. “Kajamälu” võib aga ei ole silmas peetud omama ühisosa teiste mälu-uurijate samanimelise terminiga.

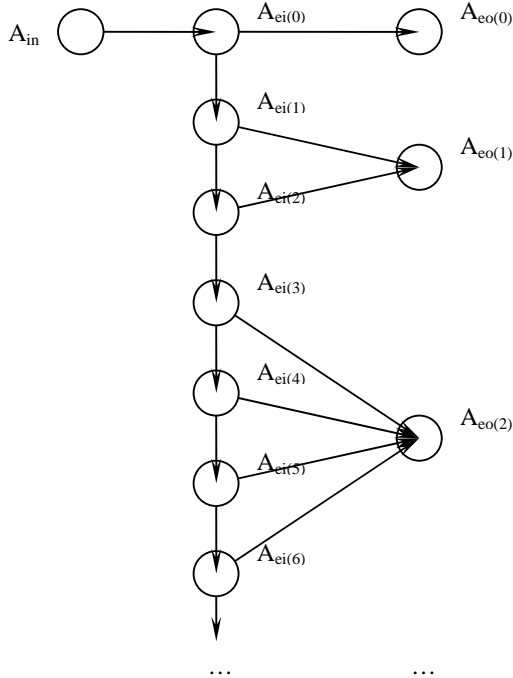
Kajamälu siinkohal ei tähenda ka, et tegu oleks auditoorse modaalsusega. Mudelis on igal sisendkanalil oma kajamälu, ning nende hulk on vastavalt kirjutatud mitmuses – “kajamälud”.

“Kajamälu”: iga raku kohta säilitatakse lühikeseks ajaks jälg viimastest olekutest. See võimaldab:

- Säilitada teatud kestusega ajaks rakkude laenglemisi varasematest ajavahemikest – võimaldab siduda hiljuti toimunud sündmusi olevikuga – õppida põhjus-tagajärje seoseid. Stiiimuli mõju võib kesta edasi peale stiiimuli lõppu (Skinner, 1965). Sisuliselt vajalik operantseks tingimiseks, kuna kinnitav stiiimul esineb hiljem kui sündmus, mille tähendust õpitakse.
- Klassikalises tingituses – tingitud reaktsioon tekib ligikaudu treenitud intervalli järel, mitte kohe (viide).
- Võimalik hilisem edasiarendus on õppida iga üksiku raku väljundväärtuste ajalist mustrit – on vajalik, et õppida stiiimuli lühikeste ajaliste järgnevuste tähendusi.

Kajamälu eristustihedus väheneb, mida kaugemale selles olevad sündmused käesolevast hetkest ajaliselt jõuavad. Omadus on motiveeritud sarnasest omadusest naturaalses õppimises (Skinner, 1965). Eristustiheduse vähenemise kõver aga ei ole käesoleval hetkel kuidagi faktiliselt põhjendatud vaid valitud mugavuse järgi.

Eristustiheduse vähenemine võimaldab logaritmiliselt väiksema andmemahu ning töömahuga töödelda ja säilitada ajaliselt kaugete sündmuste tähendusi. On mõeldav ja teostatav, et eristustihedus väheneb mingi muu skeemi järgi kui hetkel kasutatud.



$$A_{ei(0)} = A_{in}$$

$$A_{ei(i)}(t-1) \rightarrow A_{ei(i+1)}(t)$$

$$A_{eo(i)} = 1/2^i \cdot \sum_{j=2^i-1}^{2^{(i+1)}-1} A_{ei(j)}$$

kus

$A_{in}$  – stiimulit või tegevust esindav väärtus. Võib olla skalaarne (otse superviisorist või tajusüsteemist) või vektor (klassifitseerijat läbinud väärtus).

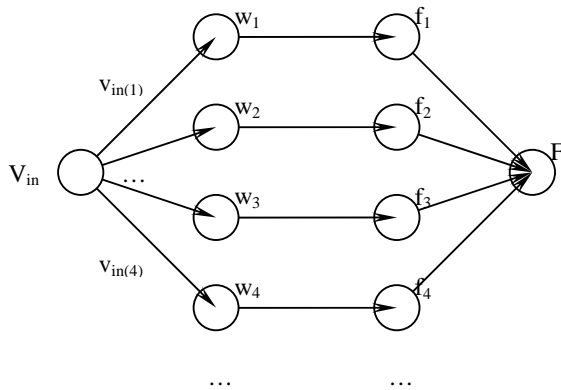
$A_{ei(i)}$  – kajamälu esimene aste (“siserakud”), ajaliselt kaugemad sündmused on eristatud sama tihedusega kui ajaliselt lähemad. Vastavalt skalaarne või vektor andmetüüp.

$A_{eo(i)}$  – kajamälu teine aste (“väljundrakud”), kajamälu väljundpunktid. Ajaliselt kaugemad sündmused on esindatud madalama eristustihedusega kui ajaliselt lähemad. Vastavalt kajamälu sisendile on andmetüüp skalaarne või vektor.

Iga kajamälu väljundrakk esindab kontsepti stiilis “ $n$  kuni  $m$  ajahetke tagasi olid selle kajamälu liini sisendis esindatud sündmused hulgast  $X$ , selle hulga elementide keskmine esinemissagedus oli ajavahemiku jooksul  $A_{eo(i)}$ ”. Iga kajamälu rakk esindab erinevat  $[n; m]$  vahemikku. Piirates kajamälu väljundrakkude arvu piirame mudeli poolt õpitavate erinevate kontseptide hulka.

Klassikalises tingimises reaktsiooni toimumise aeg järkjärgult läheneb tingitavale intervallile (Skinner, 1965). Selle reaktsiooni järkjärgulise lähenemise nähtuse põhjustab mudelis kestuse klassifikaator – algul üldistub kinnitus kõigile kestustele ja nendele kestustele järgnenud intervallidele, hiljem taandub seos ainult sobivatele tingitava intervalli ja kestuse kombinatsioonidele. Hetkelise stiimuli korral treenitakse mudelis kohe ainult minimaalse kestuse ja tingitava intervalli kombinatsioonile. Kas analoogne eristus / nähtus esineb ka naturaalses mõtlemises, vajab edaspidist täpsustamist.

## Assotsiatiivne sõlm



$$w_i = (s_i, n_i)$$

$$wv_i = \begin{cases} s_i / n_i, & \text{kui } n_i > 0 \\ 0, & \text{kui } n_i = 0 \end{cases}$$

$$f_i = v_{in(i)} \cdot wv_i$$

$$F = \sum_i f_i$$

$V_{in}$  – sisendvektor, klassifitseeritud väärtuse vektor, lähtub kajamälu ühest väljundpunktist. Juhul, kui klassifitseerija on vahepealt ära jäetud, siis sisaldab assotsiatiivne sõlm vastavalt ainult ühte andmepunkti.

$v_{in(i)}$  – sisendvektori skalaarne element  $i$

$w_i$  – andmepunktid, iga klassiga seonduma õpitud väärtus

$s_i$  – andmepunkti õpitud väärtuste summa komponent

$n_i$  – andmepunkti õpitud väärtuste samplite arvu komponent

$wv_i$  – andmepunktile vastava faktori skalaarne väärtus

$f_i$  – klassile  $i$  vastav skalaarse väljundväärtuse komponent ehk alafaktor

$F$  – skalaarne väljundväärtus, ennustatava tingimata stiimuli faktor

## Klassikaline ja operantne tingimine

Mudelis käsitletud klassikalise tingimise nähtused:

- blokeerimine
- kustumine
- kiirendatud taastumine
- iseeneslik taastumine

Operantse tingimise nähtused:

- kinnitus
- vältimine
- põgenemine
- diskriminantstiimulid
- aheldamine
- taipamine
- esmane ja teisene seos
- *interval*-kinnituskava
- *ratio*-kinnituskava
- *variable-interval* kinnituskava
- *variable-ratio* kinnituskava

Klassikalise ja operantse mõtlemise terminite tutvustust vaata Lisast.

Kinnituskavade tekke kohta on selgitus samuti Lisas.

## Klassikaline tingimine

Nõuded:

- kiire õppimine
- blokeerimine ehk *nonredundancy*, ennustusväärtuse lisatingimus
- kiire kustumine
- iga kord kiirenev taastumine
- iseeneslik taastumine ja selle nähtuse aeglane kustumine pikemas plaanis

Ülevaate klassikalise tingimise iseärasustest ning nende modelleerimiseks konstrueeritud mudelitest annab näiteks Balkenius & Morén, 1998 ning Johansson & Lansner 2002.

Järgnevalt pakun välja enda versiooni. Kommenteerin, et pakutav versioon ei väida, nagu oleks teised mudelid ebaõnnestunud või isegi võistlevad käesolevaga. Lähemal analüüsil võib nende vahel leida ühisosi kui erinevusi, kuid see võrdlus jääb peamiselt käesoleva töö raamidest välja. Erandiks on iseenesliku taastumise mudel, mis võiks olla kui täiendav samm edasi senistest mudelitest ning mille juures põhjendan, miks varasemad mudelid ei saaks seda omadust väljendada.

### ***Ennustusväärtuse lisatingimus***

Tingitava stiimuli poolt pakutav info peab olema *nonredundant* – lisama midagi infole, mida pakub mõni teine, varasemalt tingitud stiimul. Vastavat eksperimenti nimetatakse *blocking* (Balkenius & Morén, 1998; Johansson & Lansner 2002). Seda omadust omavad kõik klassikalise tingimise mudelid <sup>7</sup>.

Lisaks olen Lisas käsitlenud ja klassikalise tingimise diagrammidel demonstreerinud, mis kasu see nähtus annab õppivale süsteemile.

---

<sup>7</sup> Osaliselt sellest tingimusest tuleneb omakorda hulk teisi omadusi, millest võib pika ja tiheda ülevaate saada aadressidelt: [http://www.scholarpedia.org/article/Classical\\_Conditioning](http://www.scholarpedia.org/article/Classical_Conditioning) ja [http://scholarpedia.org/article/Computational\\_Models\\_of\\_Classical\\_Conditioning](http://scholarpedia.org/article/Computational_Models_of_Classical_Conditioning)



## **Valemid**

Tingitud stiimuli ennustamine toimub:

$$S_{ucs} = \sum_i^n F_{cs(i)} = \sum_j^m v_j \cdot w_j$$

kus

$S_{ucs}$  – tingimata stiimuli, väljundi väärtus

$F_{cs}$  – tingitud stiimuli väärtusele vastava assotsiatiivse sõlme ehk faktori väärtus

$w_j$  – andmepunkti  $j$  väärtus

$v_j$  – tingitud stiimuli klassi  $j$  aktivatsiooni määr

$n$  – tingitud stiimulite arv

$m$  – andmepunktide arv

Järgnevalt toon kaks valemit õppimise kahe aspekti – aeglane kustumine (kasutuses hiljem iseeneslikus taastumises) ning blokeerimise – teostamiseks. Kolmandaks toon valemi, mis neid aspekte ühendab.

Blokeerimise ja mitme muu klassikalise tingimise aspekti kohta on loodud mitmeid mudeleid (vaata näiteks: Balkenius & Morén, 1998; Johansson & Lansner, 2002). Käesoleva klassikalise tingimise valemite eesmärk on muuta õppimise ja kustumise kiirus sõltuvaks varasema kogemuse hulgast. Rohkem treenitud kogemust on pikas plaanis raskem kustutada (viide).

## Aeglase kustumise esimene valem

Tingitud stiimuli seose aeglasele kustumisele vastava õppimise võib saavutada:

üldistatult

$$F_{cs(i)}(t+1) \leftarrow \frac{s_{f(i)}(t) + S_{ucs} / n_p}{n_{f(i)}(t) + 1}, \text{ kui } n_p > 0$$

$$n_p = \sum_i v_{in(i)}$$

siit valem klassifitseerijat kasutava mudeli jaoks:

$$s_{f(i,j)}(t+1) \leftarrow s_{f(i,j)}(t) + S_{ucs} / n_p \cdot v_{in(i,j)}, \text{ kui } n_p > 0$$

$$n_{f(i,j)}(t+1) \leftarrow n_{f(i,j)}(t) + 1 \cdot v_{in(i,j)}$$

kui piirata õppimise kiirust, siis:

$$s_{f(i,j)}(t+1) \leftarrow s_{f(i,j)}(t) + \text{lslf}[s_{f(i,j)}(t), n_{f(i,j)}(t), S_{ucs} / n_p, v_{in(i,j)}]$$

$$n_{f(i,j)}(t+1) \leftarrow n_{f(i,j)}(t) + v_{in(i,j)} \cdot \text{lsl}$$

õppimiskiirust piirav funktsioon

$\text{lslf}(s, n, x, v) = o + (x \cdot v \cdot \text{lsl} - o) \cdot \text{lsl}$ , kus

$$o = \begin{cases} s \cdot \left( \frac{n + v \cdot \text{lsl}}{n} - 1 \right), & \text{kui } n > 0 \\ 0, & \text{vastasel juhul} \end{cases}$$

kus

**i** – faktori ehk assotsiatiivse sõlme indeks

**(t)** – diskreetne ajahetke indeks

**F<sub>cs(i)</sub>** – faktori *i* väärtus, üldistatud valemis

**S<sub>ucs</sub>** – tingimata stiimuli väärtus

**n<sub>p</sub>** – tingimata stiimulit ennustavate alafaktorite sisendaktivatsioonide summa

**s<sub>f(i)</sub>** – andmepunkti *i* treeningväärtuste summa, üldistatud valemis

**n<sub>f(i)</sub>** – andmepunkti *i* treeningväärtuste samplite arv, üldistatud valemis

**v<sub>in(i,j)</sub>** – assotsiatiivse sõlme *i* sisendvektori skalaarne element indeksiga *j*

**s<sub>f(i,j)</sub>** – assotsiatiivse sõlme *i* andmepunkti *j* treeningväärtuste summa

**n<sub>f(i,j)</sub>** – assotsiatiivse sõlme *i* andmepunkti *j* treeningväärtuste samplite arv

**lslf(s, n, x, v)** – andmepunkti *s*-komponendi paranduse *x*-i omandamise kiirust piirav funktsioon

**o** – neutraalne andmepunkti *s*-komponendi parandus, mille korral jääks andmepunkti

skalaarne koguväärtus pärast *n*-komponendi suurendamist samaks

**lsl** – õppimiskiiruse konstant, väärtusega vahemikus  $[0, 1]$ . 0-väärtuse korral õppimist ei

toimu, 1-väärtuse korral õppimise kiirus ei ole piiratud. Vahepealsete väärtuste korral

andmepunkti *s*-komponent muutub *ls* osa *o* ja *x*-i vahelisest distantsist.

Valem töötab põhimõttel: mida rohkem treeninguid on mingi faktor läbinud (suurem  $n_f$  väärtus), seda aeglasemini faktori väärtus muutub. Faktori väärtuse nullimiseks kulub samapalju treeninguid senisele vastupidise  $S_{ucs}$  väärtusega, kui faktori senise väärtuse saavutamiseks.

Sisuliselt õpib valem keskmist  $S_{ucs} / n_p$  väärtust.

Näide dünaamika kohta:

**TODO: tabel**

Binaarsete (väärtused kas 0 või 1) stiimulite korral ning eeldusel, et ühekorraga on aktiivne üks stiimul, ennustab valem ühtlasi sisuliselt Bayesi tõenäosust, et stiimuli korral esineb UCS. See valem ei vasta blokeerimise / *nonredundancy* nõudele. Samuti ei võimalda see kiiret taastumist ega kiiret kustumist.

## Blokeerimise valem

Tingitud stiimuli seose õppimise kooskõlas *blocking / nonredundancy* nõudega võib saavutada lahendades võrrandi:

üldistatult

$$\sum_i \frac{s_{f(i)} + x}{n_{f(i)} + 1} = S_{ucs}$$

seega

$$F_{cs(i)}(t+1) \leftarrow \frac{s_{f(i)}(t) + x}{n_{f(i)}(t) + 1}, \quad \text{kus}$$

$$F_{cs(i)} = s_{f(i)} / n_{f(i)}$$

ehk klassifitseerijat kasutavas mudelis:

$$\sum_i \sum_j f_{i,j}^* = S_{ucs}$$

$$f_{i,j}^* = \begin{cases} \frac{[s_{f(i,j)}(t) + x_s \cdot v_{in(i,j)}] \cdot v_{in(i,j)}}{n_{f(i,j)}(t) + 1 \cdot v_{in(i,j)}}, & \text{kui } n_{f(i,j)}(t) > 0 \text{ või } v_{in(i,j)} > 0 \\ 0, & \text{ülejäanud juhtudel} \end{cases}$$

võrrandi lahendus:

$$x_s = \frac{S_{ucs} - P^*}{N^*}, \quad \text{kus}$$

$$P^* = \sum_i \sum_j p_{i,j}^*$$

$$N^* = \sum_i \sum_j n_{i,j}^*$$

$$p_{i,j}^* = \begin{cases} \frac{s_{f(i,j)} \cdot v_{in(i,j)}}{n_{f(i,j)} + v_{in(i,j)}}, & \text{kui } v_{in(i,j)} > 0 \\ 0, & \text{vastasel juhul} \end{cases}$$

$$n_{i,j}^* = \begin{cases} \frac{v_{in(i,j)} \cdot v_{in(i,j)}}{n_{f(i,j)} + v_{in(i,j)}}, & \text{kui } v_{in(i,j)} > 0 \\ 0, & \text{vastasel juhul} \end{cases}$$

andmepunkti väärtuste uuendamine vastab valemile:

$$s_{f(i,j)}(t+1) \leftarrow s_{f(i,j)}(t) + x_s \cdot v_{in(i,j)}$$

$$n_{f(i,j)}(t+1) \leftarrow n_{f(i,j)}(t) + 1 \cdot v_{in(i,j)}$$

kui piirata õppimise kiirust, siis:

$$s_{f(i,j)}(t+1) \leftarrow s_{f(i,j)}(t) + x_s \cdot v_{in(i,j)} \cdot \text{Isl}$$

$$n_{f(i,j)}(t+1) \leftarrow n_{f(i,j)}(t) + v_{in(i,j)}$$

kus

$i$  – assotsiatiivse sõlme indeks

$j$  – alafaktori indeks

$f_{i,j}^*$  – uus faktori  $i$  alafaktori  $j$  väärtus

$x$  – võrrandi lahendamisel leitav väärtus

$x_s$  – alafaktoreid sisaldava võrrandi lahendamisel leitav väärtus

$F_{cs(i)}$  – faktori  $i$  väärtus, üldistatud valemis

$s_{f(i)}$  – andmepunkti  $i$  treeningväärtuste summa, üldistatud valemis

$n_{f(i)}$  – andmepunkti  $i$  treeningväärtuste samplite arv, üldistatud valemis

$s_{f(i,j)}$  – assotsiatiivse sõlme  $i$  andmepunkti  $j$  treeningväärtuste summa

$n_{f(i,j)}$  – assotsiatiivse sõlme  $i$  andmepunkti  $j$  treeningväärtuste samplite arv

$P^*, N^*, p^*, n^*$  – võrrandi lahendamiseks kasutatava valemi komponendid

$lsl$  – õppimiskiiruse konstant, väärtusega vahemikus  $[0, 1]$ .  $0$ -väärtuse korral õppimist ei toimu,  $1$ -väärtuse korral õppimise kiirus ei ole piiratud. Blokeerimise valemis on  $lsl$  konstanti kasutatud erinevalt, kui ülejäänud valemites, et valemil oleks kiirendatud taastumise omadus.

Nimetan ennustuse arvutamisel kasutatavaid liidetavaid faktoriteks, kuna see, mida blokeeriv õppimine teeb, on eesmärgilt sarnane faktoranalüüsile, selle erinevusega, et õppimine toimub *online*, ehk samal ajal tegevusega ja andmete laekumisega. Viimane asjaolu põhjustab paratamatult siiski erinevusi tavalisest faktoranalüüsist; siinkohal oli oluline rõhutada eesmärgi sarnasust.

Valem töötab põhimõttel: suurema samplite arvuga andmepunktide väärtusi muudetakse proportsionaalselt vähem kui väikese samplite arvuga andmepunktide väärtusi (samplite arv siinkohal tähendab andmepunkti järgmiseks tsüklikult omandatavat samplite arvu, seega senitreenimata andmepunkti samplite arv on  $1$ ), kuid siiski muudetakse kõigi andmepunktide väärtusi alati niipalju, et neist tulenevate faktorite summa vastaks treenimisel kasutatud  $S_{ucs}$  väärtusele. Seega pole säärast piirangut andmepunkti skalaarse väärtuse muutmisele, nagu esimeses, aeglase kustumisega valemis oli. Kõik andmepunktid võivad ühe ajahetke jooksul omandada senisest märgatavalt erinevad väärtused. Viimane avaldub kõige ilmekamalt juhul, kui esineb vaid üks CS ning kinnitav UCS on varasemast tingimisest erineva väärtusega: sel juhul edaspidi CS ei ennusta senini tingitud UCS väärtuste keskväärtust, nagu esimese valemi puhul, vaid viimast UCS väärtust.

Näide dünaamika kohta:

**TODO: tabel**

Sellel valemil pole vastavust iseenesliku taastumise aeglase kustumise nõudele, mille aga lisan järgmistes peatükkides.

See valem omab blokeerimise, kiire õppimise, kiire kustumise ja kiireneva taastumise omadusi.

Kuigi ülalolevat valemit on katsetatud ainult mõne klassikalise tingimise omaduse suhtes, peaks ta vastama ka mitmetele ülejäänutele, mis siinkohal aja puudusel välja jäid.

Seni olen teadlik ühest klassikalise tingimise omadusest, millele see valem ega ka järgnevates peatükkides olevad edasiarendused ei vasta: super-normaalne tingimine – kui tingida koos CS1+CS2, seejärel tingida üksi ning pidurdavalt CS1, siis see suurendab CS2 reaktsiooni, võrreldes juhtumiga, kui see oleks tingitud üksi ilma CS1-ta<sup>8</sup>.

Kui tingimine toimuks vastupidises järjekorras: CS1 → pidurdus, CS1+CS2 → UCS, siis mudelis tekiks samuti olukord, kus CS2 reaktsioon on suurem, kui ilma CS1-ta treenides oleks olnud.

Minu hetkehüpootees on, et ülalolev õppimise valem võib olla õige, kuid anomaalne nähtus annab märku täiendavatest seostest või protsessist, mis mudelis praegu puudub, ning mida tasuks edaspidi uurida. See omadus on väidetavalt olemas Rescorla-Wagneri originaalses mudelis (viide:

[http://www.scholarpedia.org/article/Computational\\_Models\\_of\\_Classical\\_Conditioning](http://www.scholarpedia.org/article/Computational_Models_of_Classical_Conditioning)), kuid selle mudeli puudus ja ühtlasi praeguse probleemi valguses huvipakkuv asjaolu on, et see mudel ei tööta reaajas (Johansson & Lansner, 2002).

Juhin veelkord tähelepanu, et UCS ennustus arvutatakse faktorite summast. Faktorite õppimine pole sama, mis tagajärgede tõenäosuste õppimine (vaata ka peatükki: Aeglase kustumise esimene valem). Mudel võib õppida tõenäosust, juhul kui vaadeldavale tagajärjele eelneb vaid üks põhjus, stiimul.

Enamasti on stiimuleid mitu, nende ennustused liituvad ning tõenäosuste summa asemel on ennustatud sündmuse määr, potentsiaalselt ka määraga  $> 1$ .

Siiski on blokeerimise valem täiendava treenimise käigus võimeline õppima, kui mõne sündmuse põhjused kokku liites ei põhjusta sündmuse määr tõusmist üle  $1$ . Seda on võimalik õppida diskriminantse sõlme lisamise abil, seostades koosinevad põhjused üheks diskriminantsete stiimulite komplektiks ning õppimisel omistades sellele komplektile negatiivse faktori väärtuse, mis liituks nende stiimulite poolt teineteisest sõltumatult tehtud (määraga  $> 1$ ) ennustuste summale:

$$\text{Ennustus} = \text{stiimul}_1 \times \text{kaal}_1 + \text{stiimul}_2 \times \text{kaal}_2 + \text{diskriminant}(\text{stiimul}_1, \text{stiimul}_2) \times \text{kaal}_3$$

Mudeli **neljas eeldus** on, et ennustavates seostes õpitakse esialgu määra. Erandite täiendava õppimise käigus on võimalik õppida tõenäosusi.

## Aeglase kustumise teine valem, vastab blokeerimise nõudele

Ühendamiseks ülalkäsitatud kaks omadust, ühendan ülalolevates õppimise valemite omadused.

Aeglase kustumise esimese valemi omaduseks oli, et iga andmepunkti väärtus saab muutuda vaid piiratud määral, sõltuvalt andmepunkti senisest samplite arvust ning  $S_{ucs}$  väärtusest.

Rakendan sarnase piirangu blokeerimise omadusega valemisse sel viisil, et uus valem õpib keskmist viga, mis jääb peale seda, kui  $S_{ucs}$  väärtusest lahutatakse ülejäänud faktoritele vastavad  $p^*$  võrrandi komponendid.

$$s_{f(i,j)}(t+1) \leftarrow s_{f(i,j)}(t) + lslf[s_{f(i,j)}(t), n_{f(i,j)}(t), x_{s(i,j)}, v_{in(i,j)}]$$

$$n_{f(i,j)}(t+1) \leftarrow n_{f(i,j)}(t) + v_{in(i,j)} \cdot lsl$$

kus

$$x_{s(i,j)} = \frac{S_{ucs} - P^* + p_{i,j}^*}{1 + N^* - n_{i,j}^*}$$

ja õppimiskiirust piirav funktsioon

$lslf(s, n, x, v) = o + (x \cdot v \cdot lsl - o) \cdot lsl$ , kus

$$o = \begin{cases} s \cdot \left( \frac{n + v \cdot lsl}{n} - 1 \right), & \text{kui } n > 0 \\ 0, & \text{vastasel juhul} \end{cases}$$

Ühe CS tingimise korral õpib valem ennustama  $S_{ucs}$  keskvärtust, erinevalt eelmises peatükis olnud blokeerimise valemist. Mitme CS korral esineb blokeerimise nähtus. Juhul, kui  $m$  arv stiimuleid tingitakse algusest peale korraga, käsitleb valem neid kui ühte, s.o iga neist saab  $1/m$  osa  $S_{ucs}$  keskvärtusest.

Analoogselt eelmistele õppimise valemitele käesolev valem rakendab kõigile assotsiatiivsetele sõlmedele ühe globaalse korrektsiooniparameetri:

$$(S_{ucs} - P^*, 1 + N^*)$$

Erinevus seisneb selles, et korrektsiooniparameetri mõju võib olla sõlme sisese täiendava info järgi lokaalselt piiratud.

Selle valemi eesmärk on saada lahti eelmises peatükis olnud blokeerimise valemi kustumise (õppimise) võrdlemisi konstantsest kiirusest – muuta õppimise ja kustumise kiirus sõltuvaks varasema kogemuse hulgast.

Näide dünaamika kohta:

**TODO: tabel**



See valem ei vasta kiire kustumise ja kiire taastumise nõuetele, sellest tulenevalt pole tal ka iseenesliku taastumise omadust.

## Iseeneslik taastumine

Pikas plaanis on rohkem kinnitatud seoseid raskem kustutada. See avaldub nende seoste iseeneslikus taastumises.

Ka operantse tingimise puhul kehtib, et rohkem kinnitatud seosed võtavad kauem aega kustutada (Skinner, 1965 ja Ferster & Skinner, 1957). See väide üksi veel ei täpsustaks, kas kustumise aeglusel võiks olla mingi piirang, ehk mudeli seisukohast, kas mudeli siseselt on omandavate seoste tugevusel lagi ning kustumine fikseeritud kiirusega või sõltub kustumine kõigi kogemuste statistilisest keskmisest s.o lisanud samplite arvu mõju.

### **TODO:**

+ **iseenesliku taasumise omadused, Rescorla, 2004**

Ükski Balkenius & Morén, 1998 kirjeldatud mudelitest ega Johansson & Lansner, 2002 mudel ei oma iseenesliku taastumise omadust.

Sissejuhatuseks pakun kolm hüpoteesi nähtuse põhjustena:

- kustutatud tegevus taastub, kuna teised tegevused on vahepeal suhteliselt nõrgenenud ning seega suhteliselt kustutatud tegevus saab taas ülekaalu
- kontekst on muutunud
- on toimunud “lühimälu” konsolideerimine või keskmistamine pikaajalise mälu, mille tulemusena püsimälu seose tugevus on küll nõrgenenud, kuid omaltpoolt on lühimälus olev mõjuv seose tugevus kasvanud, sest püsimälus olnud seos oli tugevam kui lühimälus.

See on efekti poolest analoogne ajalise konteksti muutumise ideele, mis vastandub füüsilise konteksti muutumisele (Bouton, 2004).

Järgnevalt tutvustan, mis moodi eelmistes peatükkides kirjeldatud valemeid saab iseenesliku taastumise omaduse tekitamiseks kasutada viisil, mis sarnaneb kolmanda toodud hüpoteesiga.

Iseeneslik taastumine eeldab vähemalt kahe mälusüsteemi olemasolu.

Üks mälusüsteem on tavaline blokeerimise nõudele vastav mälusüsteem, mida on modelleerinud ka mitmed varasemad mudelid (ülevaade: Balkenius & Morén, 1998) ja mulle teadaolevalt viimati Johansson & Lansner, 2002.

Devenport, 1998 järgi õpitud seose või kustutamise mõju (Rescorla, viide) on pöördvõrdeline tingimisest või kustutamisest möödunud ajaga. Siit tuleneb, et pikema aja tagant toimunud

tingimistel on kõigil enamvähem võrdne kaal ning käesoleva hetke CS ennustus vastab selle stiimuli varasemate tingimiste UCS keskmisele väärtusele.

Sellest keskmistamise omadusest on ajendatud esimene ja kolmas õppimise valem. Kolmas õppimise valem lisab hüpoteesi, et ka pikaajalises mälus olevad seosed vastavad blokeerimise nõudele – seega õpitakse mitte UCS väärtust, vaid  $x$ -paranduskomponenti kolmandas valemis. Taoline keskmistamine loomulikult ei vastaks seoste tugevusele, mis on pöördvõrdeline tingimisest möödunud ajaga, juhul kui möödunud aeg on veel lühike. Selleks on mudelis lühiajaline mälu. Kahe mälusüsteemi vahelise valiku dünaamikat kirjeldab hetkel valem:

$$f = f_{pl} \cdot (1 - ms) + f_{ps} \cdot ms$$

$$ms(t + 1) = ms(t) + mss \cdot [v_{in} - ms(t)]$$

kus

**f** – ennustatud alafaktori väärtus

**f<sub>pl</sub>** – “pikaajalise mälu” andmepunkti ennustus

**f<sub>ps</sub>** – “lühimälu” andmepunkti ennustus

**ms(t)** – ühe andmepunkti “lühimälu” ja “pikajalise mälu” vahelist valikut kontrolliv muutuja ajahetkel  $t$ .  $w=1$  korral toimub ennustus puhtalt “lühimälu” põhjal,  $w=0$  korral puhtalt “pikaajalise mälu” põhjal. Igal andmepunktil assotsiatiivses sõlmes ehk seosel on oma  $ms$  muutuja.

**v<sub>in</sub>** – andmepunkti sisendi ehk stiimuli tugevus

**mss** – ühelt mälusüsteemilt teisele ümberlülitumise kiirust kontrolliv konstant

Tõenäoliselt tuleb seda dünaamikat mudelis edaspidi viimistleda.

Valemi üldine dünaamika on hetkel selline, et stiimuli olemasolu korral liigub  $w$  väärtus  $[w; 1]$  vahemikus  $mss$  osakaalu jagu väärtuse  $1$  suunas, stiimuli puudumisel liigub  $w$  väärtus  $[0; w]$  vahemikus  $mss$  osakaalu jagu väärtuse  $0$  suunas. Nõrgema stiimuli korral on liikumise määr vastavalt väiksem.

“Lühimälu” esindab mudelis esimene blokeerimise valem (teine õppimise valem), “pikaajalist mälu” esindab mudelis aeglase kustumise ja blokeerimise nõudele vastav valem (kolmas õppimise valem).

Kummagi mälusüsteemi õppimine on praeguses mudelis jätkuvalt eraldi ning muutusteta,  $w$  kontrollib ainult tingitud stiimuli põhjal tingimata stiimuli ennustamist.

On tõenäoline, et ülal toodud mälusüsteemide vahel valimise dünaamika pole päris vastavuses Devenport, 1998 tulemustega. Võibolla aitab, kui edaspidi lisada mudelisse mitu lühiajalise mälu jälge.

Hetkel on oluline see: pikaajalise-lühiajalise mälu eristus mudelis on ajendatud mõttekäigust, et oleks üpriski raskesti usutav, kui aju säilitaks jälje kõigist tingimistest eraldi ja seda lõputuseni – tajumaailma maht võib olla suur, kuid siiski on piiratud, ning aju saab sellega toime tänu paralleeltöötlusele, kuid ajahetkede hulk sellele vastandatuna kasvab pidevalt ja paisutaks mälu nõuded palju suuremaks, kui seda oleks tajumaailma mahu probleem.

Võimalik, et ka pikaajalises mälus on mitmeid mälujälgi (näiteks episoodiline mälu) – käesoleva iseenesliku taastumise valemi ja mudeli eesmärk pole väita, et iseeneslik taastumine sõltub CS mitteesinemise ajast või ainult ajast, ega seda, et mälujälgi on kaks, vaid modelleerida paari iseenesliku taastumise omaduse saavutamiseks tarvilikku aspekti – mitme mälujälje olemasolu ning blokeerimise võimega keskmistava mälu õppimise valemit. Lisaks põhimõttelist mehhanismi (*ms*-muutuja) nende mälude vahel valimiseks, jättes eraldi küsimuseks, mis seda mehhanismi kontrollib. Käesolevas mudeli teostuses on kontrollijaks aeg CS esinemisest.

Teine mälusüsteemidest arvestab andmepunktide pikaajalist keskmist. Selleks on tarvilik, et see mälusüsteem sisaldab jälge kogemuste hulgast (andmepunkti  $n$ -komponent). Kuna varasemad mudelid seda komponenti ei sisalda, siis nad ei saa omada sellise mälusüsteemi omadusi.

## Operantne tingimine

Tänu operantsele tingimisele isend saab tegutseda aktiivselt vastavalt oma õpitud ootustele-motiividele, mitte ainult refleksidena välismaailma stiimulitele.

Kui klassikaline tingimine seostas stiimuleid reaktsioonidega või teiste stiimulitega, siis operantne tingimine seostab tegevusi, “reaktsioone” stiimulitega ehk tagajärgedega.

Kinnitustüüpide ja kinnituskavade tutvustust vaata Lisast.

## ***Seadistuspunkti tegelike väärtuste dünaamika õppimine ning seadistuspunkti energia***

Operantse tegevuse eesmärgi võib võrrelda homöostaasi<sup>9</sup> või PID-kontrolleri<sup>10</sup> ideega: eesmärk on saavutada või säilitada sisemine homöostaas, hoida väärtus tasakaalupunkti teatud eesmärgis-näitajas, mis asub seadistuspunktis. Seega seadistuspunkt sisaldab kahte komponenti: tegelik väärtus ja eelistatud väärtus.

Kui ma allpool kasutan terminit “seadistuspunkti väärtus”, siis pean silmas seadistuspunkti tegelikku väärtust, mitte eelistatud väärtust. Mudel õpib kontrollima tegeliku väärtuse dünaamikat.

Operantne tegevus ehk “operant” (Skinner, 1965) üritab deprivatsiooni korral saavutada kinnitavat stiimulit või suurendada selle määra, aversiooni korral eemaldada mingit stiimulit või ülemäärast stiimuli taset. Deprivatsioon ja aversioon on mudelis sama protsess, ainukeseks erinevuseks on, et ühel juhul püüeldakse seadistuspunkti positiivse diferentsiaali poole, teisel juhul negatiivse diferentsiaali poole.

Samuti on positiivne ja negatiivne kinnitus ja karistus sama protsess, mis opereerib seadistuspunkti erinevate väärtuste kontekstis ning omab erineva polaarsusega diferentsiaali:

- positiivne kinnitus: väärtus negatiivne, positiivne diferentsiaal
- negatiivne kinnitus: väärtus positiivne, negatiivne diferentsiaal
- positiivne karistus: väärtus neutraalne või positiivne, positiivne diferentsiaal
- negatiivne karistus: väärtus neutraalne või negatiivne, negatiivne diferentsiaal

Tingimisel mõjub ka väärtuse ootus: kui oodati positiivset väärtust (näiteks müra), kuid esines neutraalne, siis sellele eelnenud tegevus saab negatiivseks kinnitajaks (Skinner, 1965). See nähtus on kooskõlas ka klassikalise tingimise valemitega. Mudel kasutabki õppimiseks ainult klassikalise tingimise valemiteid ja tulenevalt klassikalise tingimise valemitest on analoogsed omadused ka mudeli ülejäänud ootustel-ennustustel.

Karistuse läbi õpitud kogemus võib mudelis mõnel juhul saada hiljem tegevust motiveerivaks kogemuseks ning kinnituse läbi õpitud kogemus vältimist põhjustavaks, juhul kui seadistuspunkti eelistatud väärtus peaks muutuma või tegelik väärtus liikuma teisele poole nulli tingimisel esinenud väärtusega võrreldes. Homöostaasi seadistuspunktide eelistatud väärtused on tavaliselt väga stabiilsed, kuid näiteks motoorsesse kontrolli rakendatuna oleks seadistuspunktid pidevas muutumises. See, kas mingi tingimine oli kinnitav või karistav, ilmneb alles konkreetsetes käitumisolukorras, mitte tingimise ajal – mis on kooskõlas ka Skinner, 1965 kirjeldatud lähenemisega.

Selguse mõttes lisan veel, et deprivatsioon või aversioon versus emotsioon on erinevad asjad: deprivatsioon ja aversioon on vajadused, emotsioon on hästi üldiselt võttes seisund, mis põhjustab teatud kaasasündinud käitumismustreid (Skinner, 1965).

Kui mõnele tegevusele järgneb korruga mitu erinevat kinnitajat (mitu seadistuspunkti liigub eelistatud väärtuse suunas), siis see tegevus saab üldistatud kinnitajaga sarnasel moel eriti tugevalt kinnitatud. Võib aga juhtuda, et sama tegevus saab ühekorraga kinnitatud ja karistatud, erinevate motiivide suhtes, sel juhul võib olla tulemuseks ostsilleeriv käitumine (kuna olles poolel teel ühe motiivi täitumise poole, see motiiv nõrgeneb võrreldes konkureerivaga) või osaline tegevus (Skinner, 1965). Eelnevast tuleneb, et seadistuspunktid on sõltumatud ning pole ühte globaalset kinnitust või karistust.

Seadistuspunktide väärtustes tasakaalu hoidmine on operantse mõtlemise üks olulisi erinevusi *reinforcement learningust (RL)*, mille eesmärk on maksimiseerida teatud näitajat (Sutton & Barto, 1998). *RL* põhimõttel töötavaid komponente annab küll vabalt ühendada nõnda, et nad kokkuvõttes midagi tasakaalus hoiaksid, kuid suur osa *RL* paradigmat ja sellest tulenevast mõtteviisist on mõjutatud siiski hinnangu maksimiseerimise vaatenurgast, mitte tasakaalu hoidmisest. Psühholoogia ajaloost vastab sellele pigem Thorndike teooria ning tema teooriat Sutton & Barto, 1998 raamatus ka tsiteeritakse, Skinnerit mitte.

Laiendades homöostaasi tasakaalupunkti jälgimise põhimõtet isendi motoorika juhtimise põhimõttele võib oletada, et mõnel juhul eelistatud väärtus (tasakaalupunkt) pole alati sama. Seega peab mudel õppima kõikidesse võimalikesse erinevatesse soovitud väärtustesse ehk tasakaalupunktidesse ehk olekutesse viivaid tegevusi.

Operantse tegevuse eesmärk on sarnane sellele, millega tegelevad kontrolliteooria mudelid. Analoogselt enamlevinud lihtsale kontrollisüsteemile PID-kontrollerile saab mudel kasutada sisendina stiimuli või tegevuse tuletist, hetkeväärtust, integraali. Käesolevas teostuses on

kasutuses ainult tegevuse kestus (pole täpselt sama, mis integraal) ning diskriminantsete stiimulite hetkeväärtus ja kestus.

Erinevalt PID-kontrollerist on mudeli õppimisfunktsioon keerukas, lisaks on olemas aheldamise ja taipamise võime.

Siit tuleb teine erinevus  $RL$ 'ust:  $RL$  õpib seisundite vahel liikumist nõnda, et igast seisundist teise seisundisse liikumist on tarvis eraldi õppida.  $RL$  õpib seisundite vahel liikumise tõenäosusi iga erineva tegevuse korral, kuid pole olemas seisundite väärtuste distantsi mõistet. Käesolev mudel õpib ühe seadistuspunkti väärtuste vahel liikumiseks väärtuse muutu iga erineva tegevuse korral.

Seega ühe seadistuspunkti või kinnitava stiimuli võimalike väärtuste ehk seisundite grupi elemendid saab paigutada intervallskaalale ning üldistada õpitud samme mistahes selle skaala punktide vahel liikumisele. Kuid kui mingite punktide vahel liikumises on erandeid, võrreldes selle üldistusega, siis saab seda täiendava õppimise käigus eraldi õppida, juhul kui võtta seadistuspunkti hetkeväärtus täiendavaks diskriminantseks stiimuliks. Selline seisundite intervallskaalale paigutamine on mudeli **viies eeldus** ning oluline on, et siinkohal erandite õppimiseks vajalikku diskriminantset stiimulit pole kohustuslik alati mudelisse ühendada, erinevalt  $RL$ -st.

Eeldus põhineb mõttekäigul, et füüsiline maailm on pidev ning selle eelduse rakendamine võimaldab õppida kiiremini: kui reinforcement learning vajab  $n$  seisundi vahel liikumiseks vähemalt  $n^2$  treenimist (tegevused igas seisundis, mis viivad igasse järgmisse seisundisse), siis käesolevas mudelis piisab 2 treenimisest (üks kummagi suuna jaoks).

Selline erinevus reinforcement learningust on analoogne Yudkowski, 2002 kriitikale semantiliste võrkude suunas, mille baasühikuks on mittepidevad mõisted, mille vahelised seosed tuleb eraldi õppida.

Seadistuspunkti olukorra hindamiseks arvutatakse selle energia. Seadistuspunkti energia ehk operantse energia valem:

$$Discr = A_o - P_o$$
$$E_o = abs(Discr)$$

$A_o$  – seadistuspunkti tegelik väärtus (*actual value*), antakse mudelivälise süsteemi ehk superviisori poolt

$P_o$  – eelistatud väärtus (*preferred value*), homöostaasi punkt. Antakse mudelivälise süsteemi poolt

**Discr** – ebakõla eelistatud väärtuse ja tegeliku väärtuse vahel (*discrepancy*).

$E_o$  – seadistuspunkti energia ehk operantne energia

Seadistuspunktiga (ja ka kinnitava stiimuliga) seonduvaid assotsiatiivseid sõlmi treenivaks tingimata stiimuliks on mudelis seadistuspunkti (või kinnitava stiimuli) tegeliku väärtuse muut, mis on arvatud valemiga:

$$Diff_o(t) = A(t) - A(t-1)$$

(**t**) – käesolev ajahetk

(**t-1**) – eelmine ajahetk (eelmine mudeli töötsüklil)

**Diff<sub>o</sub>** – tegeliku väärtuse muut ajavahemikul eelmise ajahetke ja käesoleva ajahetke vahel

Mudel õpib tegevuse tagajärgi ehk saab kinnitusi ka siis, kui hetkel deprivatsiooni ei esine ehk seadistuspunkt on eelistatud väärtusel või eelistatud vahemikus. Seega õppimine ei sõltu deprivatsioonist, kuid tegevuse käivitumine sõltub. Ka kustumiskõver ei sõltu sellest, kui suur oli deprivatsioon tingimise ajal (Skinner, 1965).

Teine huvitav tähelepanek on Skinneril see, et mõni kinnitav stiimul ei saagi deprivatsiooni kohe vähendada. Selle asemel tuleb mõelda, et stiimuli tüüp ise on neil puhkudel kinnitav (Skinner, 1965). Mudelis praegu sellist eristust ei ole, ehk täpsemini – seadistuspunktid esindavad neid väärtusi või stiimuleid, mida saab koheselt mõõta. Luua nende koheste väärtuste ning aeglaste väärtuste vahele täiendavaid seoseid on juba superviisori ülesanne ning sõltub mõõdetava nähtuse ja kinnitajate eripäradest.

### **Seadistuspunkti väärtuse ja energia ennustamine**

Käesolevas mudeli teostuses on igal seadistuspunktil üks väärtus, mille korral tema energia on null. Mudeli edasiarendusena on mõeldav luua seadistuspunkte, millel on punkti asemel eelistatud väärtuste vahemik, kus energia on null.

$$A_{op(i)} = A_{o(i)} + Diff_{op(i)}$$

$$E_{op(i)} = abs(A_{op(i)} - P_{o(i)})$$

**A<sub>op(i)</sub>** – seadistuspunktile *i* ennustatud väärtus

**A<sub>o(i)</sub>** – seadistuspunkti *i* tegelik väärtus

**Diff<sub>op(i)</sub>** – seadistuspunktile *i* ennustatud väärtuse muut

**P<sub>o(i)</sub>** – seadistuspunkti *i* eelistatud väärtus

**E<sub>op(i)</sub>** – seadistuspunktile *i* ennustatud energia



Lihtsa operantse probleemilahendaja tööülesanne on leida sellised tingitud stiimulite ehk tegevuste määrad, et kõigile seadistuspunktile klassikalise tingimise järgi ennustatud energia oleks järgmises ajahetkes minimaalne.

$$E_p = \sum_i E_{op(i)} \cdot w_{o(i)}$$

$E_p$  – kõigile seadistuspunktile ennustatud energiatega summa

$E_{op(i)}$  – ühele seadistuspunktile ennustatud energia

$w_{o(i)}$  – seadistuspunkti  $i$  kaal. Käesolevas teostuses  $w_{o(i)} = 1$

Mudel kasutab tegevuste määrade leidmiseks *simulated annealing*<sup>11</sup> algoritmi.

Seega

$$Av(t+1) = (A_1(t+1), A_2(t+1), \dots, A_n(t+1))$$

$$Av(t+1) \approx \arg \min \{E_p[Av(t+1)]\}$$

$\mathbf{Av}(t+1)$  – plaanitav ehk kujutletav tegevuste määrade vektor järgmiseks ajahetkeks

$A_i(t+1)$  – üks komponenttegevus järgmise ajahetke tegevuste vektoris

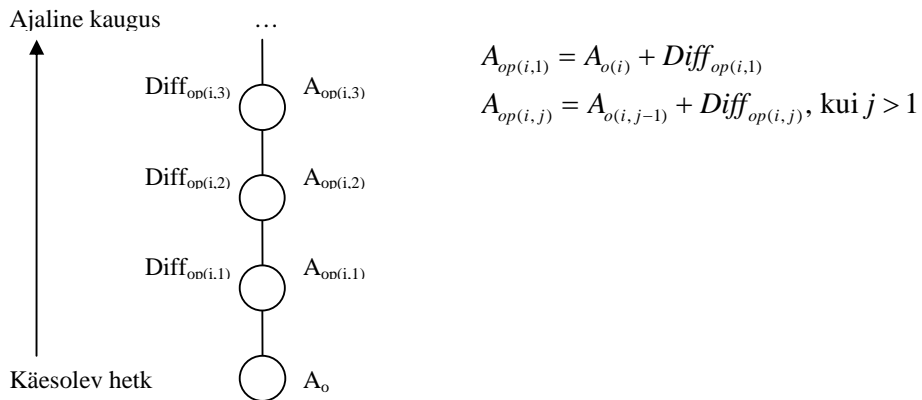
$\mathbf{n}$  – tegevuste hulk

$E_p[\mathbf{Av}(t+1)]$  – seadistuspunktile ennustatud summaarne energia tegevuste vektori  $Av(t+1)$  korral

### **Mitme tuleviku ajavahemiku operantne ennustamine**

Mõni tegevus võib anda vahetult ühe tulemuse ning pikemas plaanis erineva tulemuse. Juhul kui pikemas plaanis tekki tulemus kas nullib esialgse tulemuse, liigub hiljem vastupidises suunas või muudab tulemuse hoopis tugevamaks kui soovitud (seadistuspunkti väärtus liigub üle tasakaalupunkti), siis tuleb seda arvesse võtta. Samuti võib juhtuda, et tegevus annab soovitud tulemuse teatud ajalise hilinemisega. Seega oleks vahetult järgmise ajahetke ennustamine operantses mõttes tihti kasutu.

Seepärast sisaldab mudel kujutlust mitme tuleviku ajavahemiku kohta.



$A_{o(i)}$  – seadistuspunkti  $i$  käesoleva ajahetke väärtus

$Diff_{op(i,j)}$  – seadistuspunktile  $i$  tuleviku ajavahemiku  $j$  lõpuks ennustatud seadistuspunkti tegeliku väärtuse muut eelnevaga võrreldes

$A_{op(i,j)}$  – seadistuspunktile  $i$  tuleviku ajavahemiku  $j$  lõpuks ennustatud väärtus

Iga järgnev ennustatav tuleviku ajavahemik on mudeli praeguses teostuses kaks korda pikem kui varasem, ehk vähema eristusjõuga. On mõeldav ja teostatav, et mudel saab ka teistsuguseid eristustihedusi kasutada.

## Operantse ennustuse energia mitme tuleviku ajavahemiku ennustamise korral

ideaalis:

$$E_{p(i)} = \int_0^{fl} abs(A_{op(i,t)} - P_{o(i)}) dt$$

käesoleva teostuse puhul on kasutusel variant trapetsvalemist:

$$E_{p(i)} = \sum_{j=0}^{\log_2 fl} abs[(A_{op(i,j+1)} + A_{op(i,j)})/2 - P_{o(i)}] \cdot 2^j, \text{ kus}$$

$$A_{op(i,0)} = A_{p(i)}$$

kus

$A_{op(i,t)}$  – seadistuspunktile  $i$  tuleviku ajahetkeks  $t$  ennustatud väärtus

$P_{o(i)}$  – seadistuspunkti  $i$  eelistatud väärtus

$t$  – tuleviku ajahetke indeks

$fl$  – viimase ennustatava tuleviku ajahetke kaugus käesolevast hetkest (*futureline length*)

$E_{p(i)}$  – seadistuspunktile  $i$  ennustatud energia, arvestades kõiki ennustatud tuleviku ajavahemikke

$j$  – tuleviku ajavahemiku lõpu indeks

Tulevikku ennustatud ajavahemike seadistuspunktide energia korrutatakse kahe astmega, kuna mudeli eristustihedus kaugematele sündmustele väheneb ning iga järgmine tuleviku ajavahemik esindab mudeli hetketeostuses kaks korda pikemat ajavahemikku kui eelmine.

Käesolevast energia valemist peaks veelkord selgemini silma paistma operantse mõtlemise sarnasus kontrolliteooriaga ning erinevus  $RL'$  ust:

- Operantne mõtlemine üritab säilitada väärtust määratud tasakaalupunktis ka edaspidisteks ajavahemikeks, võimalikult väikese kõikumisega ümber selle. Kui ükski tegevus ei saavuta kohe soovitud seadistuspunkti väärtust, siis leitakse tegevus, mis selle pikemas plaanis saavutab, kuid ei ületa seda, või ei ületa teisele poole tasakaalupunkti kaugemale kui käesoleva hetke ebakõla juba on.
- On võimalik üldistada seadistuspunkti kahe väärtuse vahel liikumist mistahes väärtuste vahel liikumisele.

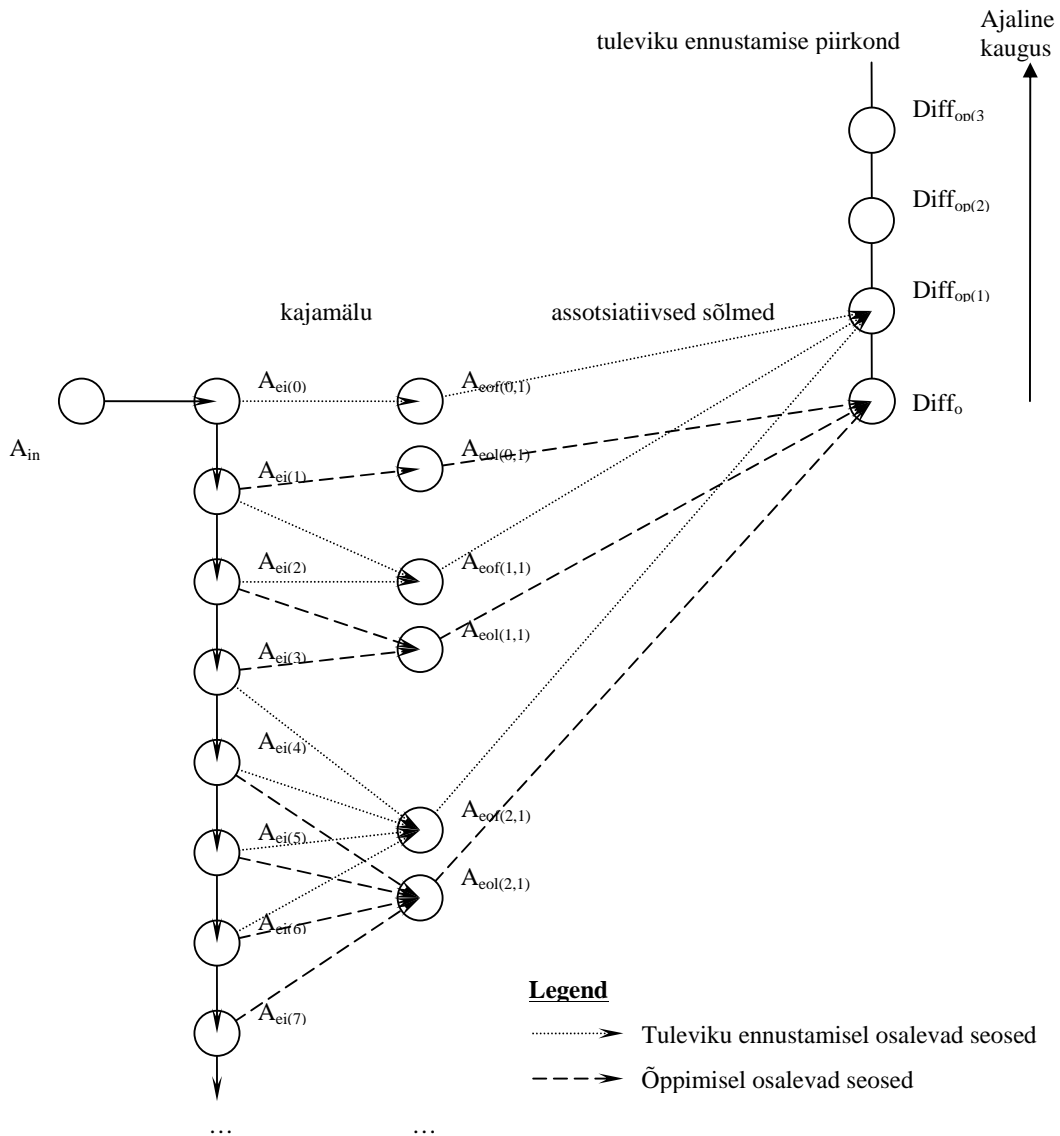
Seega eesmärk on saavutada soovitud väärtus võimalikult kiiresti ning võimalikult väikese tegeliku väärtuse tulevase kõikumisega.

Mudeli edasiarendusena on mõeldav, et arvestatakse ka tulevase muutusi eelistatud väärtuses, näiteks kiirelt muutuvate eesmärkidega motoorse kontrolli puhul.

## Mudeli õppimine ennustamaks mitut tuleviku ajavahemikku

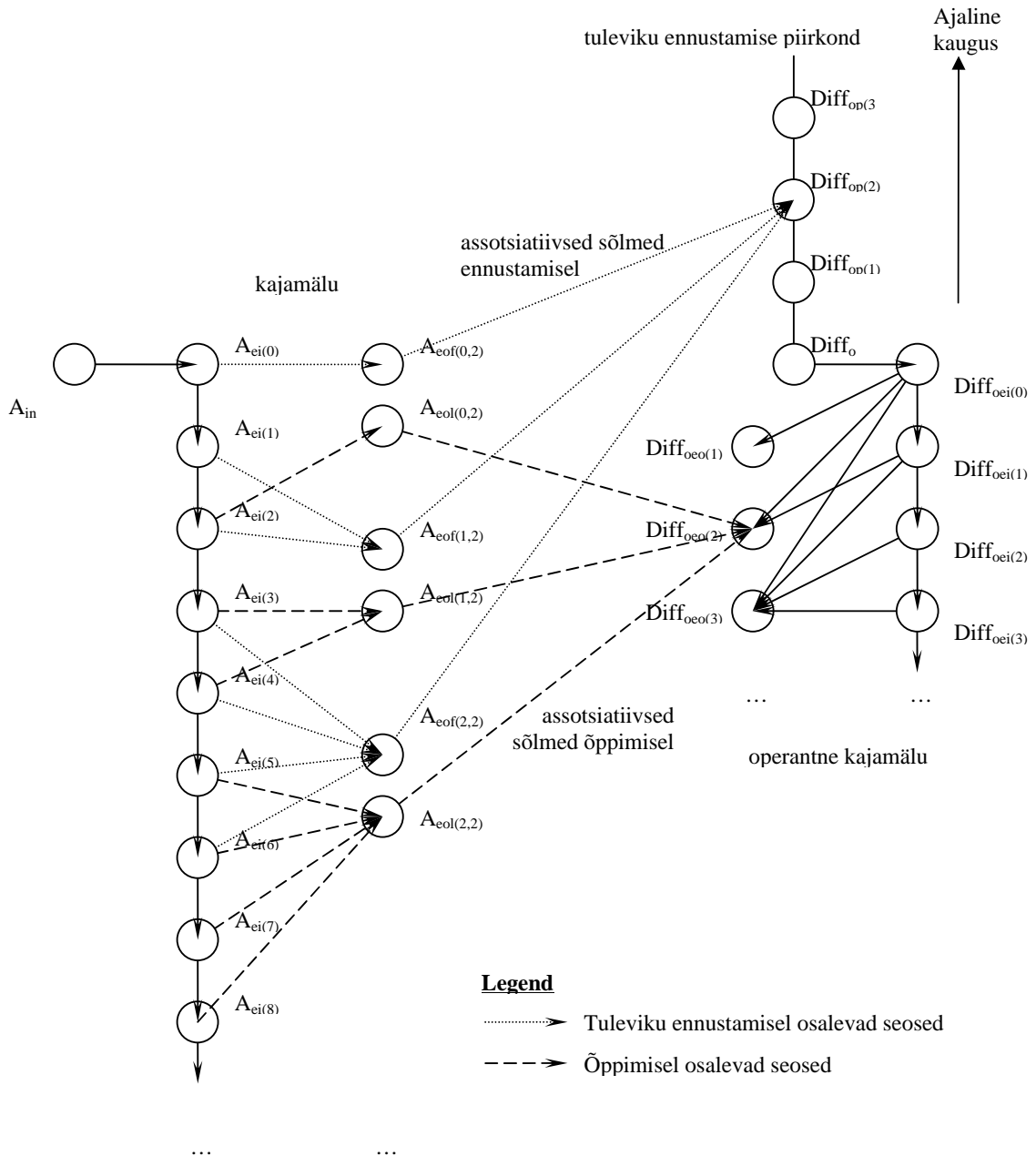
Järgnevalt toon detailsemate diagrammidega välja operantse kajamälu piirkonna ja tuleviku kujutlemise piirkonna suhted tavalise kajamälu ja assotsiatiivsete sõlmedega.

Alustuseks diagramm, mis kujutab esimese tuleviku ajahetke õppimiseks ja ennustamiseks tarvilikke seoseid:



Õppimisel ja tuleviku ennustamisel osalevad seosed kuuluvad paaridesse, igas paaris üks seos kummastki liigist. Iga õppimise ja ennustamise seosepaar läbib ühte ja sama antud paarile vastavat assotsiatiivset sõlme. Selle assotsiatiivse sõlme sisendid ja väljundid on õppimisel ja ennustamisel ühendatud erinevate piirkondadega.

Järgnevalt joonis, mis kirjeldab seoseid seadistuspunkti mitme tuleviku ajavahemiku muutude õppimiseks ja ennustamiseks:



Iga joonisel olev õppimisel ja tuleviku ennustamisel osalev seoste paar on seotud ühe sellele paarile vastava assotsiatiivse sõlmega. Õppimise tööfaasis liigub sellesse assotsiatiivsesse sõlme õppimiseks tarvilik informatsioon kajamälust (vaata joonist) ning teiselt poolt ennustatava seadistuspunkti kajamälust; tuleviku ennustamisel lähtub assotsiatiivse sõlme info kajamälu teistest piirkondadest (vaata joonist) ning liigub ennustatava seadistuspunkti tuleviku kujutlemise sõlmedesse.

$$A_{eof(i,j)} = 1/2^i \cdot \sum_{k=2^i-1}^{2^{(i+1)}-2} A_{ei(k)}$$

$$A_{eol(i,j)} = 1/2^i \cdot \sum_{k=2^i-1}^{2^{(i+1)}-2} A_{ei(k+2^i)}$$

$$Diff_{oei(0)}^f = Diff_o^f$$

$$Diff_{oei(j+1)}^f(t) \leftarrow Diff_{oei(j)}^f(t-1)$$

$$Diff_{oeo(j)}^f = \sum_{k=0}^{2^j-1} Diff_{oei(k)}^f$$

kus

$A_{eof(i,j)}$  – tingitud stiimuli kajamälu teine aste tuleviku ennustamise tarvis, ajaliselt kaugemad sündmused on esindatud madalama eristustihedusega kui lähemad.  $j$  on tuleviku ajavahemiku indeks, mida ennustatakse. Ülal toodud valem seda ei kasuta, küll aga on see tarvilik ennustuses osaleva assotsiatiivse sõlme määramiseks.

$A_{eol(i,j)}$  – tingitud stiimuli kajamälu teine aste õppimise tarvis, ajaliselt kaugemad sündmused on esindatud madalama eristustihedusega kui lähemad.  $j$  on tuleviku ajavahemiku indeks, mida ennustama õpitakse. Lisaks on  $j$  tarvilik õppimises osaleva assotsiatiivse sõlme määramiseks.

$Diff_{oei(j)}^f$  – seadistuspunkti või kinnitava stiimuli tegeliku väärtuse muudu kajamälu esimene aste, ajaliselt kaugemad sündmused on sama eristustihedusega kui lähemad.

$Diff_{oeo(j)}^f$  – seadistuspunkti või kinnitava stiimuli tegeliku väärtuse muudu kajamälu teine aste, ajaliselt kaugemad sündmused on vähema eristustihedusega kui lähemad. Õppimise tarbeks algab teise astme väärtuste arvutamine alati esimese astme esimesest rakust – oluline on tingitud stiimuli  $A_{in}$  kajade  $A_{eol(x)}$  ajalise suhte  $Diff_{oeo(j)}^f$ -ga ning  $A_{eof(x)}$  ajalise suhte  $Diff_{op(j)}^f$ -ga samasus.

Ajaliselt kaugemate sündmuste eristustiheduse vähenemine võiks olla kooskõlas Skinneri vaatlusega, et kinnitus peab olema vahetu ning vastasel korral kinnituse täpsus väheneb (Skinner, 1965) – mudeli puhul tähendab see, et see seostab pikema ajavahemikuga vastavalt rohkem toimunud sündmusi.

Õppimisel luuakse eraldi assotsiatiivne sõlm iga tegevust esindava raku indeksi  $i$ , kajamälu väljundi indeksi  $j$  ning ennustatava/õpitava tuleviku ajavahemiku indeksi  $k$  kombinatsiooni tarvis. Tulemuseks oli kolmemõõtmeline assotsiatiivsete sõlmede maatriks, mille indeksid on  $i, j, k$ .

Kui mudelis on mitu seadistuspunkti, siis määrab assotsiatiivse sõlme neljandana ennustatava seadistuspunkti indeks, mis aga ei mõjuta eelnevaid valemeid.

## Diskriminantsed stiimulid

Diskriminantsed stiimulid võivad olla sensoorsed stiimulid või teatud eelnevad tegevused, mis määravad käesoleva tegevuse tulemust. Diskriminantsetel tegevustel võivad olla omakorda omad diskriminantsed stiimulid.

Mõned tegevused annavad tulemusi vaid teatud lisatingimuste (diskriminantsete stiimulite) juures.

Kui neid stiimuleid parajasti ei esine, siis pole mõtet tegevust sooritada või ei teki teatud sündmuse või tegevuse toimumise järel tagajärje ootust (kuid taju puudumine võib tänu klassifitseerivale piirkonnale samuti olla tähendusega, olla stiimul).

Juhul kui esineb deprivatsioon, on sobiva diskriminantse sensoorse stiimuli ilmumine väliselt sarnane klassikalise tingituse stiimuli ilmumisele – stiimulile järgneb tegevus. Paljud kultuurilised ja märgilised “reaktsioonid”, nende õppimine (näiteks asjade nimetamine, korrutustabel, teksti valjusti lugemine, joonistuse kopeerimine) on selle järgi tegelikult operantne tegevus diskriminantse stiimuli juhtimisel (Skinner, 1965). Oluline on, et tegelikult on tegu tahtliku tegevusega, ilma et peaks aga motiivi pidevalt eksplitsiitselt stiimuli-reaktsiooni seosesse kaasama. Samas võib selline reaktsioon olla tugev või nõrk, kiire või aeglane sõltumata stiimuli tugevusest (Skinner, 1965).

*RL*'u kontsept *state* on sarnane diskriminantsele stiimulile. Erinevus on, et diskriminantseid stiimuleid saab olla (väga)palju ning igäüks võib neist seostuda tegevustega eraldi, *state* aga moodustatakse kõigi stiimulite kombinatsioonist.

Lisaks on tänu sellele võimalik õppimise käigus üldistatada diskriminantse stiimuli omadusi stiimuli väärtuse vahemikule, *state* puhul puudub aga seda moodustavate elementide väärtuste kui seisundite distantsi mõiste.

Iga diskriminantse stiimuli teatud olekut saab seada eesmärgiks eraldi ning need eesmärgid võivad olla pidevas muutumises – seega ka diskriminantsete stiimulite mõju tegevusele.

Diskriminantne stiimul on tingimus ühele tegevusele teatud eesmärgi korral; samas *RL state* määrab *policy* valiku (Sutton & Barto, 1998), ehk on globaalse ja väljaspool õppimist ka püsiva mõjuga tegevusele.

Kinnitavate stiimulite kui alaeasmärkide olekute saavutamisel saab kasutada sarnast stiimuli väärtuste skaala punktide vahelist liikumist, nagu seadistuspunkti puhul (vaata operantse energia peatükke).

Lisaks võivad muutuda diskriminantseteks stiimuliteks varem esinenud stiimulite kombinatsioone, tegevusi või stiimulite ja tegevuste kombinatsioon esindavad rakud.



Viimane juhtub näiteks *ratio*-kinnituskava puhul (Ferster & Skinner, 1957). Taolised diskriminantsete stiimulid ei ole väliselt vaadeldavad.

Diskriminantsete stiimulid jagan selguse mõttes diskriminantseteks tajudeks (s.h väliselt mittevaadeldavad kombinatsioonid esindavad rakud) ning diskriminantseteks tegevusteks. Esimesi on mõnel juhul võimalik tekitada mitmeastmelises operantses tegevuses, mida Skinner nimetas “aheldamiseks”. Viimaseid on võimalik tekitada otseselt (diskriminantse stiimulita tegevus on alati teostatav ehk sama tulemusega, v.a seoste iseenesliku taastumise aspekt).

Lisaks eristan sensoorse diskriminantstiimuli erijuhu – tegevuse sihtmärk-stiimuli. Tegevuse kõrgemal kavandamisel mitte ei kästa jäsemete konkreetsetel lihastel liikuda konkreetset viisil, vaid kästakse sooritada tegevust teatud punktide ruumis (analoogselt Georgopoulos jt., 1993 pakutule, mainitud: Kolb & Whishaw, 2003), mis võiks olla valitud ja pilgu keskmise hoitud kujutluspildiga sobivuse järgi. Seal ruumipunktis asuvat objekti või taju esindavat stiimulit nimetan sihtmärk-stiimuliks.

Käesolev mudel seda stiimuli liiki ei sisalda, küll aga on tarvilik edasiarenduseks, kuna võimaldab Piaget sensomotoorse staadiumi kolmanda alatasandi ning hiljem ka oletatavasti tööriistakasutuse võime (vaata Lisa).

Diskriminantsete stiimulite õppimiseks kombineeritakse mudelis tegevusi esindavate kajamälude väljundid tajusid esindavate kajamälude väljunditega.

Tegevust esindava kajamälu väljundeid võib kirjeldada vektorina:

$$A_{\text{eof}(i,k)} = (A_{\text{eof}(i,1,k)}, A_{\text{eof}(i,2,k)}, \dots, A_{\text{eof}(i,n,k)})$$

Analoogselt võib diskriminantstiimulit esindava kajamälu väljundeid kirjeldada vektorina:

$$S_{\text{eof}(i,k)} = (S_{\text{eof}(i,1,k)}, S_{\text{eof}(i,2,k)}, \dots, S_{\text{eof}(i,n,k)})$$

kus

$A_{\text{eof}(i,j,k)}$  – tegevust esindava raku  $i$  kajamälu teise astme väljund  $j$  ennustamiseks tuleviku ajavahemikku  $k$

$A_{\text{v eof}(i,k)}$  – tegevust esindava raku  $i$  kajamälu teise astme väljundite vektor ennustamiseks tuleviku ajavahemikku  $k$

Analoogselt saab konstrueerida definitsioonid sümbolitele:

$A_{\text{eol}(i,j,k)}$  ja  $A_{\text{v eol}(i,k)}$  – tegevust esindava raku jaoks, õppimisel

$S_{\text{eof}(i,j,k)}$  ja  $S_{\text{v eof}(i,k)}$  – diskriminantstiimulit esindava raku jaoks, tuleviku ennustamisel

$S_{\text{eol}(i,j,k)}$  ja  $S_{\text{v eol}(i,k)}$  – diskriminantstiimulit esindava raku jaoks, õppimisel

Diskriminantsete stiimulite õppimiseks luuakse assotsiatiivse sõlm iga tegevust esindava raku indeksi  $ai$ , selle kajamälu väljundi  $aj$ , diskriminantstiimulit esindava raku indeksi  $si$ , selle kajamälu väljundi  $sj$  ning ennustatava tuleviku ajahetke  $k$  kombinatsiooni tarvis; ühe piiranguga – diskriminantstiimulit seotakse ainult samal hetkel või peale vaadeldava diskriminantstiimuli kajamälu oleva sündmuse toimumist esinenud tegevuste kajaga, kuna tegevusega seotav diskriminantstiimul ei saa esineda peale tegevust – vastasel korral ei oleks võimalik seda tegevuse valikul arvesse võtta. Edaspidist uurimist vääriks, kas mingi stiimuli ootus järgnevasse ajahetkedesse võib olla naturaalses mõtlemises diskriminantstiimuliks käesoleva hetke tegevusele.

Seega hetkel piirang:

$$sj \geq aj$$

Tulemuseks on viiemõõtmeline assotsiatiivsete sõlmede maatriks, mille indeksid on  $ai$ ,  $aj$ ,  $si$ ,  $sj$ ,  $k$ .

Maatriksi osa, kus  $sj < aj$ , on hetkel kasutamata.

Kui mudelis on mitu seadistuspunkti, siis määrab assotsiatiivse sõlme kuuendana ennustatava seadistuspunkti indeks, mis aga ei mõjuta eelnevaid valemeid.

Eelnevalt kirjeldasin, kuidas leitakse tegevuse ja diskriminantstiimuli kombinatsioonile vastav assotsiatiivne sõlm. Järgnevalt kirjeldan, kuidas ühendatakse kahe kombinatsiooni valitud kajamälu väljundi väärtused, saataks edasi assotsiatiivsesse sõlme.

Nii tegevust esindava kajamälu kui ka diskriminantstiimulit esindava kajamälu rakkude väljunditeks on vektor, mis sisaldab kajamälu alguses klassifitseeritud väärtuse kaalude vektori kaja või kajade segu. Need vektorid korrutatakse, et iga diskriminantsetes sõlmes moodustatud kombinatsiooni element esindaks konjunktsiooni teatud tegevuse ja teatud stiimuli klassist:

$$M_{ai,aj,si,sj,k} = A_{eof(ai,aj,k)} \cdot S_{eof(si,sj,k)}$$

tekkiva kaalude maatriksi saab konvertida esituse mugavuse mõttes varemkirjeldatud assotsiatiivsele sõlmele vastavaks vektoriks:

$$V = (M_{0,0}, M_{0,1}, \dots, M_{0,q-1}, \\ M_{1,0}, M_{1,1}, \dots, M_{0,q-1}, \\ \dots \\ M_{p-1,0}, M_{p-1,1}, \dots, M_{p-1,q-1})$$

ehk lühemalt

$$V_i = M_{i \text{ div } p, i \text{ mod } q} \quad \forall i = 0, \dots, p \cdot q - 1$$

kus

**p** – tegevust esindava vektori suurus (maksimaalne klasside arv)

**q** – diskriminantstiimulit esindava vektori suurus

**div** – täisarvude jagamine

**mod** – jääk täisarvude jagamisest

Tegevust esindava kajamälu ning diskriminantstiimulit esindava kajamälu  $k$ -indeks on väärtuste kombineerimisel ühine.

Võib ette kujutada, et ka seadistuspunkti enda väärtus võib olla diskriminantstiimul. Näiteks, kui seadistuspunkt esindab mingi liigese tegelikku ja soovitud nurka propriotsepsiooni tajus, siis erinevate nurkade all sama tugev lihaspingutus võib anda erineva liigese nurga muudu. Praeguses töös sooritatud katsetes sellist seadistuspunkti diskriminantstiimuliks seostamist pole tehtud, kuid mudeli programm on paindlik ja võimaldab sellisele omadusele seadistada.

Samuti võib olla diskriminandiks varasem kinnitus või aeg selle toimumisest (Ferster & Skinner, 1957), sellist seadistust on mudeli juures kasutatud intervall-kinnituskava testimiseks.

## Kinnitavad stiimulid, aheldamine ja taipamine

Aheldamine on operantse tingimise termin, mis tähendab, et mingi kinnituse saamiseks tuleb sooritada tegevuste ahel. *Ratio*-kinnituskava on kõige lihtsam aheldamise näide.

Defineerin siinkohal ka sõna “taipamine”: Taipamine on isendi võime tekitada tegevusjärgnevusi, mida ta pole tervikuna kunagi varem sooritanud, küll aga on sooritanud või muul moel õppinud kõigi selle järgnevuse üksikute sammude tagajärgi eraldi. Järgnevuse alaelemendid tekitavad või aktiveerivad järgmise alaelemendi tegevuste diskriminantseid stiimuleid. Need järgmise alaelemendi diskriminantsed stiimulid on kinnitavateks stiimuliteks järgnevuse eelmise alaelemendi tegevusele või tegevus-stiimul seosele. Pean terminit “kinnitav stiimul” mõnevõrra segadusttekitavaks: isend õpib, et mingi sündmust tekitab mingi stiimuli ka siis, kui see stiimul ise veel pole operantse eesmärgiga seotud (eeltingimine). Kinnitav on see stiimul Skinner, 1965 vaatenurga järgi siiski, kuna kinnituse olemasolu selgubki alles selle poole püüdlamise olukorras. Ühes taipamise ahelas osalevaid alaelemente võib olla õpitud ka grupikaupa, kuid erinevas järjekorras ning võibolla pikemate ajaliste pausidega.

Eelnevates peatükkides kirjeldatud mudeli omadused lubasid vaid lihtsaid probleeme vajaduste täitmise juures lahendada. Kui mudel ei tulnuks ülesandega toime, kuna tegevus polnuks sobiva diskriminantstiimuli puudumise tõttu tulemuslik, siis mudel üritanuks järgmist, enda arvestuse järgi väiksema kasulikkusega varianti, seni kuni on midagigi tehtud või enam minimaalselt kasulikke variante ei ole. Oletatavasti sisaldunuks neis variantides või järgnes neile otsingukäitumine, kus isend liigub ringi lootuses, et mõnes asupaigas tingimused on sobivamad.

Otsingukäitumine on sellisele mudelile, mis mõtleb vaid üks samm ette, üks viimaseid väljapääse. Kuid suurema arukuse võimaluse korral sugugi mitte kõige efektiivsem. Näiteks, kui on olemas varasemaid kogemusi, võiks võimekam mudel neid kogemusi ka mitut tegevussammu nõudva tõkke kiireks lahendamiseks kasutada, ilma et oleks kogunud varasemalt tervet lahenduskäiku ühekorraga.

Toomela kirjeldab taipamist 3. tasandil (Toomela, 2003b), kuid ta peab silmas ahvide probleemilahendusvõimeid. Ahvide taipamisele on iseloomulik võime seostada mitut objekti. Samas on olemas tegevusahelaid, mis eeldavad igal sammul vaid ühe diskriminantstiimuli ning tegevuse seostamist, ja mida võiks oletada enamuse imetajate võimeks (vaata Tomasello, 1997 – Piaget 4. sensomotoorne tase). Selline võime mudeli ehitamise loogika järgi ka peab eelnema taipamises mitme objekti seostamise võimele ning ongi siinkohal teostatav.

Muu hulgas annab vastava mehhanismi loomine *ratio*-kinnituskava, kui aheldamise erijuhu, õppimise võime.

Tasub tähele panna, et taoline “probleemilahendus” esineb vaid siis, kui otsene tulemuslik tegevus puudub või katsed ebaõnnestuvad. See ei ole sama, mis inimeste tahtlik planeerimine ja tegevuse edasilükkamine.

Lisaks käesolev töö ei käsitle küsimust, mis keskkonnas isendid taolisi järgnevasi õpivad ning mis moodi ise õppimiseks sobivaid tingimusi loovad, näiteks mängu või katsetamise teel.

Probleemilahenduse käigus võivad aktiveeruda ka säärased ahelad ehk eesmärgipuu harud, milles sisaldub mittesoovitud samme – samme, mis annavad küll probleemilahenduse suhtes vajaliku tulemuse, kuid samas on muud moodi ebameeldivad või lausa kahjulikud. Need eesmärgipuu harud saavad väiksema kaalu, ja tekib suurem võimalus, et teised eesmärgipuu harud on prioriteetsemad. Aversiiivsus pidurdab vaid otse vahetus läheduses oleva eesmärgipuu haru, mitte kogu eelnevat puud.

Diskriminantseid stiimuleid esindavaid kanaleid saab tegevuse planeerimisel kasutada kahte moodi:

- a) Diskriminantsete stiimulite kajamälu sisu on selline, nagu senisest tajuinfost tulenevalt olema peaks; tegevuse planeerimisel leitakse käesoleval hetkel rakendatav tegevus, mis arvestades koosmõju diskriminantsete stiimulitega annaks minimaalse operantse energia ennustuse.
- b) Kujutleda ette erinevaid diskriminantsete stiimulite väärtusi (muutes kujutluse ajaks vastavate stiimulite kajamälu), kombineerida neid võimalike tegevuste väärtustega ning leida need diskriminantsete stiimulite väärtused, mis võimaldasid koosmõjus võimalike tegevustega saavutada minimaalse operantse energia ennustuse.  
Kui see viimati mainitud energia on madalam, kui viisil (a) saavutatav, siis püüelda vastava diskriminantse stiimuli väärtuse tekitamisele. Taoline alaeesmärkide seadmine võimaldab taipamise.

Viimane tähendab, et diskriminantseid stiimuleid esindavad rakud muutuvad ise samuti teatud mõttes seadistuspunktideks, ehk kinnitavateks stiimuliteks.

**Kajamälu funktsioneerimine operantse kujutlemise tööfaasis on järgnev:**

Operantse kujutlemise tööfaasis liiguvad kujutletavate tegevuste väärtused esimesse kajamälu väljundsõlme, samas kujutletavate diskriminantsete stiimulite väärtused levivad koheselt üle

terve kajamälu. Sel moel on võimalik kujutleda mistahes diskriminantse stiimuli ja tegevuse vahelisi ajalisi suhteid.

Kinnitav stiimul ei ole stiimul, mille osas tekib deprivatsioon, kui seda stiimulit pole hulk aega esinenud. Kinnitav stiimul saab kinnitavaks sellevõrra, kui võrd ta võimaldab hetkeolukorras mingile operantsele eesmärgile lähemale jõuda (Skinner, 1965).

Mudelis on need andmestruktuurilt erinevad tavalistest seadistuspunktidest: iga kinnitava stiimuliga seotakse grupp erinevaid väärtusi – iga väärtusega vastavuses erinev energia. See väärtustega seotav energia kajastab minimaalset ennustuste energiat, mida sellise diskriminantse stiimuli väärtuse korral parima tegevusega saavutada oleks võimalik. Nende energiatega kõver võib olla mittelineaarne.

Seega, kuigi kinnitava stiimuli erinevatel väärtustel on erinev energia, ei ole need energiad sama lihtsalt arvutatavad, kui seadistuspunktide energiad – pole ühte soovitud punkti, vaid selline kõver, millel võib asuda mitu miinimumi ning mudeli ülesanne on valida üks neist – milline on kõige saavutatavam või annab parima tulemuse ka kõiki teisi eesmärgi arvesse võttes.

Käesolev mudel arvestab olekute vahelisi distantse ning eeldab, et ühe stiimuli olekud on võimalik paigutada stiimuli väärtuse järgi ritta, liikumine nende olekute vahel on võimalik tehes lühemaid või pikemaid hüppeid üle vahepealsete olekute – soovitud olekuni ehk väärtuseni jõudmiseks tuleb teha tegevusi, mis nihutavad väärtust soovitud suunas, peatudes paratamatult teatud vahepealsetel väärtustel või võimaluse korral mitte.

Erinevates olukordades võib olla võimalik korraga teha erineva pikkusega hüppeid seadistuspunkti või kinnitava stiimuli väärtuses ning saab arvestada, kas ja kui võrd palju tuleb näiteks viibida ebasoovitavate väärtuste tsoonis, arvestades hetkel võimalikke tegevusvariante ning olemata kogunud varem täpselt sama olekute vahelist järgnevust.

Nagu ka seadistuspunktide puhul – kinnitavate stiimulite madalaima energia saavutamine on eesmärk, kuid energia ei saa olla kunagi alla nulli. Mida ihaldusväärsem on saavutamata stiimul, seda kõrgem tema energia. Juba saavutatud stiimulite ja saavutamata stiimulite olekute energia on mudelis 0.

Kinnitavate stiimulite olekute energiad arvutatakse, leides madalaima energiaga tegevuste komplekti, mida sellise diskriminantse stiimuli oleku puhul saavutada oleks võimalik:

$$e_{d(i)} = \min[E_p(Av, Dv_i)]$$

kus

$e_{d(i)}$  – oleku  $i$  energia

$\mathbf{Av}$  – tegevuste kujutletud määrade vektor

$\mathbf{Dv}_i$  – diskriminantsete stiimulite kujutletud määrade vektor, sisaldab väärtust  $v_{d(i)}$

$v_{d(i)}$  – olekule  $i$  vastav kinnitava stiimuli väärtus

$\mathbf{E}_p(\mathbf{Av}, \mathbf{Dv}_i)$  – ennustatud seadistuspunktide ja kinnitavate stiimulite summaarne energia tegevuste vektori  $\mathbf{Av}$  ja diskriminantsete stiimulite vektori  $\mathbf{Dv}_i$  korral

Hetke mudeli teostuses kõigi stiimulite tegelikele väärtustele vastavate olekute energiad liidetakse ning kõigi olekute energiad normaliseeritakse nõnda, et tegelikele väärtustele vastavate olekute energiatega summa oleks võrdne käesoleva hetke seadistuspunktide energiaga, et vältida olukorda, kus tsüklilised eesmärgipuud saavutavad suurema energia, kui seadistuspunktid. Tõenäoliselt tasub seda osa edaspidi veel arendada.

### TODO: valem

Valemid planeeritavate tegevuste poolt ennustatud kinnitavate stiimulite energia arvutamiseks:

$$E_{dp} = 1/n_{dp} \cdot \sum_{i=1}^{n_{dp}} e_{dp(i)}$$

$$e_{dp(i)} = \begin{cases} \max[0, 1 - \text{abs}(rs_i)] \cdot e_{d(i)}^* + e_c, & \text{kui } v_c \neq v_{d(i)} \\ e_c, & \text{kui } v_c = v_{d(i)} \end{cases}$$

$$e_{d(i)}^* = \min(0, e_{d(i)} - e_c)$$

$$rs_i = \frac{v_{d(i)} - v_p}{v_{d(i)} - v_c}$$

kus

$\mathbf{E}_{dp}$  – kinnitavale stiimulile ennustatud energia

$n_{dp}$  – kinnitava stiimuli olekute arv

$e_{dp(i)}$  – oleku  $i$  energia ennustuse korral

$rs_i$  – *remaining step's relative size*, ennustatud väärtuse kaugus vaadeldava oleku  $i$  väärtusest, suhestatuna käesoleva hetke väärtuse distantsiga vaadeldava oleku väärtusest

$e_{d(i)}^*$  – kui palju on oleku  $i$  energia väiksem stiimuli hetkeoleku energiast

$e_{d(i)}$  – oleku  $i$  energia, leitud mudeli eelmises tööfaasis

$e_c$  – kinnitava stiimuli hetkeoleku energia, leitud mudeli eelmises töösüklis

$v_{d(i)}$  – olekule  $i$  vastav kinnitava stiimuli väärtus

$v_c$  – kinnitava stiimuli väärtus käesoleval hetkel

**TODO: joonis interpoleeritud-ekstrapoleeritud energia graafikuga ja ennustatud väärtuse erinevate paiknemistega sellel.**

Ülalolevas valemis  $abs()$  osa tagab, et kui ennustatud väärtus liigub üle vaadeldava oleku väärtuse, siis ekstrapoleeritakse teisele poole väärtust sama energia kõvera kalle, peegelpildis.  $max()$  osa tagab, et kui ennustatud väärtus liigub eemale oleku väärtuse ja stiimuli käesoleva hetke väärtuse vahemikust, siis sellele liikumisele energia muutu ei omistata, samuti kui ennustatud väärtus liigub üle oleku väärtuse ning absoluutse distantsi poolest kaugemale kui on stiimuli käesoleva hetke väärtus oleku väärtusest, siis sealt distantsist edasi sellele liikumisele energia muutu ei omistata.

$e_{d(i)}$  \* ehk  $min()$  osa tagab, et käesolevast suurema energiaga olekute poole liikumist ei väldita, küll aga püütakse madalama energiaga olekute poole. Valemi see osa on halb juhul, kui on vaid üks madala energiaga olek ning juba asutaksegi selle juures. Sel juhul mudel võib hakata liikuma kõrgemate energiatega poole, ilma ise seda märkamata (ei esine vältimist kui tegevuse allasurumist). Juhul kui on mitu madala energiaga olekut kõrvuti, on võimalik nende vahel pendeldamine ning säärast probleemi ei teki (vältimine esineb).

Ülevaاتlikult on ülaloleva valemi mõte: kui tegevus nihutab kinnitava stiimuli väärtust senisest madalama energiaga oleku poole, siis vastavalt sellele, kui võrd palju absoluutse kauguse mõttes antud olekule senisest lähemale liigutakse, sellevõrra omistatakse tegevusele selle oleku madalamat energiat. Ülejäänud juhtudel (distant olekust suureneb või oleku energia on senisest stiimuli energiast suurem) oleku enda energiat ei arvestata ning kasutatakse stiimuli hetkeenergiat selle asemel. Lõpuks võetakse käesoleva planeeritud tegevusega seonduvatest olekute energiatest keskmine.

Planeeritud tegevuse koguenergia kinnitavate stiimulitega mudelis:

$$E_p = \left[ \sum_{i=1}^n E_{op(i)} \right] + \left[ \sum_{j=1}^m E_{dp(j)} \right]$$

kus

$E_p$  – ennustatud energia plaanitava tegevuse korral

$E_{op(i)}$  – seadistuspunkti  $i$  energia plaanitava tegevuse korral

$E_{dp(j)}$  – kinnitava stiimuli  $j$  energia plaanitava tegevuse korral

$n$  – seadistuspunktide arv

$m$  – kinnitavate stiimulite arv



Erilist liiki kinnitavateks stiimuliteks on üldistatud kinnitavad stiimulid, mis võimaldavad paljusid eesmärke saavutada (sellise stiimuli näiteks on raha). Need on sõltumatumad organismi seisundist ja on tegevust motiveerivad ka siis, kui mingi konkreetne deprivatsioon parajasti puudub, kuna vähemalt mõni muu sobiv vajadus võib olla aktiivne. Samuti võib selline üldistatud kinnitaja põhjustada eriti suure tõenäosusega tegevust, kui mitu deprivatsiooni esinevad korraga (Skinner, 1965). Mudelis see tähendab, et seadistuspunktide ja kinnitavate stiimulite energiad liidetakse (vastandatuna näiteks maksimumi võtmisele).

Olles leidnud kinnitavate stiimulite olekutele uuendatud energiad, leitakse lõpuks tegevus, mis ennustab minimaalset energiat arvestades diskriminantsete stiimulite tegelikke väärtusi:

$$e_{d(i)} = \min[E_p(Av, Dv_i)]$$

$$Av(t+1) \approx \arg \min\{E_p[Av(t+1), Dv]\}$$

kus

**Dv** – diskriminantsete stiimulite tegelikud väärtused

Selles viimases planeerimisfaasis diskriminantsete stiimulite väärtused ei levi üle kajamälu, vaid iga kajamälu väljundsõlm sisaldab seda väärtust, mis sinna mudeli põhitööfaasi (ehk väljaspool kujutlemist) jooksul tekkinud on.

Ülevaاتlikult: taipamine toimib sel teel, et mudeli probleemilahenduse esimese tööfaasi tulemusena “energiseeruvad” ehk saavad operantseks kinnitajaks need diskriminantsed stiimulid, mis võimaldavad omakorda eelmises töötüklis tekkinud eesmärgile lähemale jõuda ehk vähendada varem energiseerunud kinnitavate stiimulite ja seadistuspunktide energiat. Iga järgmise töötükliga energiseeruvad kinnitavad stiimulid, mis võimaldavad eelmises töötüklis leitud alaeesmärkidele (kinnitavate stiimulite väärtustele) lähemale jõuda. Mingi stiimuli tekitamine on sealhulgas lihtsalt stiimuli soovitud väärtusele lähemale jõudmise erijuht.

Operantses tegevuses analoogset nähtust nimetatakse aheldamine (Ferster & Skinner, 1957; Skinner, 1965).

Muu hulgas võimaldab see teekonna planeerimist.

Käesolev taipamise protsess on mehhaaniline ega nõua tähelepanu protsessi. Küll aga on see suuremate ülesannete korral ehk keerukamas keskkonnas arvutusmahukas. Tähelepanu aitaks seda optimeerida.

Eriti mahukaks muutub protsess siis, kui igas tegevuse ja tagajärje seoses osaleb mitu diskriminantset stiimulit (käesolevas mudelis moodustab iga kontsepti üks tegevus-stiimul paar) – ühest küljest võimaldaks see tööriistade kasutamise (lähemalt selle kohta Lissas), teisest küljest nõuab juba päris tungivalt tähelepanu või mõne muu optimeerimisprotsessi olemasolu – näiteks, et assotsiatiivses piirkonnas luuakse ainult neid seoseid, mille faktori väärtus ületab teatud määra.

Käesoleva mudeli planeerimisfaasis tuleb leida kõigi sisendite väärtuste ning tegevuste tugevuste kombinatsioonidele vastavad energiad, arvutusmahuks võib märkida  $O(n \times m)$ .

Mitme diskriminantse stiimuliga mudelis oleks arvutusmaht  $O(n \times m^q)$ , kus

**n** – iga tegevuse võimalike määrade arv

**m** – iga kinnitava ehk diskriminantse stiimuli võimalike väärtuste arv

**q** – iga tegevuse ja tagajärje vahelise seose õppimisel koosmõjuna arvestatavate diskriminantstiimulite maksimaalne arv.

Nüüd saan selgitada ka töö algusosas olnud operantse tingimise jaoks tarvis minevate piirkondade ja kanalite joonisel olnud väidet, et osad piirkonnad pole primaatidel lõpuni arenenud:

Tomasello & Call, 1997 ja Tomasello, 1999 järgi ahvid ei ole võimelised õppima väliste sündmuste järgnevusi, mille sammudes nad ise aktiivselt ei osale, ehk põhjus-tagajärg ahelate vahepealseid lülisid (ja / või neid teadmisi probleemilahenduses kasutama). Sama probleem esineb mõnevõrra inimlastel (Frye et al, 1995, viidatud: Tomasello & Call, 1997), kuid inimlapsed on selles siiski oluliselt võimekamad (Tomasello, 1999).

Tomasello, 1999 järgi see võiks lisaks jagatud tähelepanule olla üks väike kuid oluline lisandus evolutsioonis primaatidelt inimesele – võime mõista vahendavaid põhjuslikke seoseid ning seega ka sündmuste ahela algpõhjusti.

Vahendavate põhjuslike seoste õppimise võime tähendaks lisaks täiendavate algpõhjuste mõistmisele ka võimet luua uudseid viise, mis viivad varem vaadeldud mittetegevusliku sündmuste ahela mistahes vahepealse lülini, mitte ainult tingimata alglülini, mis hetkel ehk pole saavutatav (Tomasello, 1999).

Mudeli seisukohast eelnev tähendaks oletust, et ennustavad seosed  $S \rightarrow S$  vahel jäetakse vähemalt osadel juhtudel tekitamata, evolutsioonis arvatavasti (energia ja aja) optimeerimise eesmärgil. Lisaks see tähendab, et ei looda diskriminantseid komplekte, milles osaleva tegevuse määr oleks null (antud tegevust ei toimu), vastasel korral toimiksid sisuliselt  $S \rightarrow S$  ennustused sellegipoolest.

Samas käib see oletus vastu teisele tingimise ja eeltingimise nähtusele, mis sisaldab  $S \rightarrow S$  seoseid. Teine oletus seetõttu on, et seosed  $S \rightarrow S$  küll tekitatakse ja kasutatakse toimunud sündmuste tagajärgede ennustamisel, kuid operantse probleemilahenduse ehk kujutlemise tööfaasis neid seoseid ei aktiveerita – nende teadmistega ei opereerita.

Samas on tõenäoliselt kõigil imetajatel operantse probleemilahenduse käigus kasutusel seosed  $S \rightarrow G$  vahel, mida Skinner, uurides rotte ja tuvisid, nimetas tingitud kinnitajateks (nendega tingimata ei kaasne enam tegevust  $G$  saavutamiseks) (Skinner, 1965); mujal on kasutatud ka terminit teisene kinnitav stiimul (väljendamaks mõtet, et primaarne kinnitav stiimul on see, mis mehhaaniliselt kutsus esile  $G$ ).

Nende kahe löigu informatsiooni põhjal võiks tuletada, et enamasti puuduvad kolmanda järgu tingitud kinnitajad, ehk probleemilahenduses ei kasutata seoseid  $S \rightarrow S \rightarrow G$  (ehk juhul, kui kolmanda järgu kinnitavad stiimulid pole seostatavad isendi tegevustega – pole ühtlasi õpitavad kui diskriminantsed stiimulid tegevusahelas). Kuivõrd see oletus tõele vastab, vajab edaspidist kontrollimist.

$S \rightarrow S$  seostele võiks funktsionaalselt vastata ajus *angular gyrus* piirkond, mis inimestel on märgatavalt suurem, kui primaatidel (viide).

## Kokkuvõte

Mudeli väärtus:

- blokeerimine.
- seoste iseeneslik taastumine ehk ajalise konteksti aspekt.
- stiimulite ja tegevuste kestuste tähenduste õppimine.
- ajaliste suhete õppimine (kajamälu).
- operantne käitumine füüsilises maailmas, kiirendatud olekute vahelise liikumise õppimine (vaata: eeldused).
- operantne tegevus kui kontroll vastandatuna *reinforcement learning*'u väärtuse maksimeerimise paradigmale.
- diskriminantstiimulid v *reinforcement learning*'u *state*. Võimaldab konteksti kiiremini õppida, kuna iga diskriminantstiimulit saab seostada tegevusega eraldi, *state* on aga kõiki stiimulite konjunktsioon.
- Kinnitavad stiimulid, aheldamine ja taipamine, s.h operantsed kinnituskavad
- **TODO: veel eeldustest tulenevaid väärtusi.**
- mudel on edasiarendatav, vaata Lisa.

Empiiriliselts saab mudeli realiseeringu toimimist edaspidi võrrelda naturaalse süsteemide toimimisega erinevate ülesannete juures. Samuti saab võrrelda mudeli tõhusust teiste tehisintellekti süsteemidega.

Mudel on võrdlemisi paindlik, muu hulgas tänu oma sisemisele diferentseeritusele ning struktuurile, ja selgivate täpsustuste järgi on võimalik mudelit veelgi täiendada.

## Viited

- Balkenius, C. & Morén, J. (1998). Computational Models of Classical Conditioning: A Comparative Study. URL: <http://citeseer.ist.psu.edu/246220.html> [14.05.2007].  
Lund, Sweden: Lund University Cognitive Science.
- Bouton, M. E. (2004). Context and Behavioral Processes in Extinction. Learning & Memory 11: 485-494. URL: <http://www.learnmem.org/cgi/content/full/11/5/485> [14.05.2007].
- Devenport, L. D. (1998). Spontaneous recovery without interference: Why remembering is adaptive. Animal Learning & Behavior 26: 172-181.
- Ferster, C. B. & Skinner, B. F. (1957). Schedules of Reinforcement. Prentice-Hall, Inc.
- George, D. & Hawkins, J. (2005). A Hierarchical Bayesian Model of Invariant Pattern Recognition in the Visual Cortex. Menlo Park, CA: Department of Electrical Engineering, Stanford University & Redwood Neuroscience Institute. URL: <http://www.stanford.edu/~dil/invariance/Download/GeorgeHawkinsIJCNN05.pdf> [14.05.2007].
- Honkela, T. (1997). Self-Organizing Maps in Natural Language Processing. Helsinki, Finland: Neural Networks Research Centre, Helsinki University of Technology. URL: <http://www.cis.hut.fi/~tho/thesis/> [14.05.2007].
- Johansson, C. & Lansner, A. (2002). An Associative Neural Network Model of Classical Conditioning. URL: [http://www.nada.kth.se/~cjo/publications/An Associative Neural Network Model of Classical Conditioning.pdf](http://www.nada.kth.se/~cjo/publications/An_Associative_Neural_Network_Model_of_Classical_Conditioning.pdf) [14.05.2007]. Stockholm, Sweden: Numerical Analysis and Computer Science, Royal Institute of Technology.
- Kolb, B. & Whishaw, I. Q. (2003). Organization of the Motor Systems. Fundamentals of Human Neuropsychology: 120-145. New York: Freeman.
- Miller, P. H. (1993). Theories of developmental psychology. New York: Freeman.
- Pihlakas, R. (2005). Mõtlemise tasandid ja nende modelleerimine. Esimesed kaks tasandit. (seminaritöö). Tartu, Eesti: Tartu Ülikooli Psühholoogia osakond.
- Rescorla, R. A. (2004). Spontaneous Recovery. Learning & Memory 11: 501-509. URL: <http://www.learnmem.org/cgi/content/full/11/5/501> [14.05.2007].
- Skinner, B. F. (1965). Science and human behavior. New York: The Free Press.
- Sutton, R. R. & Barto, A. G. (1998). Reinforcement Learning: An Introduction. Cambridge: MIT Press. URL: <http://www.cs.ualberta.ca/~sutton/book/the-book.html> [14.05.2005].
- Tomasello, M. & Call, J. (1997). Primate cognition. New York: Oxford University Press.

- Tomasello, M. (1999). The cultural origins of human cognition. Cambridge: Harvard University Press.
- Toomela, A. (2003b). Development of Symbol Meaning and the Emergence of the Semiotically Mediated Mind. A. Toomela (Ed.), Cultural Guidance in the Development of the Human Mind: 163-209. Westport, CT: Ablex Publishing.
- Toomela, A. (2005). Decision-making with incomplete information: Systemic and non-systemic ways of thinking in medicine. R. Bibace, J. Laird, & J. Valsiner (Eds.), Science and medicine in dialogue: Thinking through particulars and universals. Westport, CT: Greenwood.
- Vygotsky, L. S. & Luria, A. (1994b). Tool and symbol in child development. (originaal: 1929). R. van der Veer, & J. Valsiner (Eds.), The Vygotsky reader. (99-174). Oxford: Blackwell.
- Vygotsky, L. S. (1997). Thought and Language. (originaal: 1934). Alex Kozulin (Ed.). Cambridge, Massachusetts: MIT Press.
- Yudkowski, E. S. (2002). Levels of Organization in General Intelligence. Singularity Institute for Artificial Intelligence. Goertzel, B. & Pennachin, C. (Eds), Artificial General Intelligence. URL: <http://www.singinst.org/LOGI/LOGI.pdf> [14.05.2007].