

Tartu Ülikool  
Psühholoogia osakond

Roland Pihlakas

## **Klassikalise ja operantse mõtlemise modelleerimine**

Bakalaureusetöö

Juhendaja: 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.

Töö 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, diskriminantsed 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.

Finally, 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.

Sissejuhatus .....	4
Miks modelleerida mõtlemist .....	6
Tasandite kirjeldus .....	7
Piaget' sensomotoorse staadiumi alatasandid .....	9
Parker ja Gibsoni teooria primaatide füüsiliste võimete kohta .....	10
Thomas õppimisvõimete hierarhia .....	11
Mudel .....	13
Stiimulite ja tegevuste kanalite eeltöötlus .....	16
Klassifitseerija .....	16
Kajamälu .....	20
Assotsiatiivne sõlm .....	23
Klassikaline ja operantne tingimine .....	24
Klassikaline tingimine .....	25
Ennustusväärtuse lisatingimus .....	25
Valemid .....	26
Aeglase kustumise esimene valem .....	27
<i>Blocking / nonredundancy</i> valem .....	28
Aeglase kustumise teine valem, vastab <i>blocking</i> nõudele .....	30
Iseeneslik taastumine .....	32
Operantne tingimine .....	34
Operandi väärtuste õppimine ning operandi energia .....	34
Operandi väärtuse ja energia ennustamine .....	37
Mitme tuleviku ajavahemiku operantne ennustamine .....	39
Operantse ennustuse energia mitme tuleviku ajavahemiku ennustamise korral .....	40
Mudeli õppimine ennustamiseks mitut tuleviku ajavahemikku .....	42
Vältimine ja põgenemine .....	45
Diskriminantsed stiimulid .....	47
Aheldamine ja taipamine .....	51
Operantsed kinnituskavad .....	59
Fixed-ratio .....	61
Fixed-interval .....	63
Variable-ratio ning variable-interval kinnituskavad .....	65
<i>Variable-ratio</i> .....	65
<i>Variable-interval</i> .....	68
Mudel ei käsitle .....	69
Kokkuvõte .....	69
Viited .....	70

## 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.

Mõtlemise definitsioon:

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

(L. S. Vögotski, 1926 tsiteeritud: Toomela, 2003c).

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 mittekesksed teemad (Toomela/Vögotski mõtlemise tasandid, mudeli potentsiaal edasiarenduse perspektiivis – mõned hüpoteesid) on enamasti tõstetud Lisasse. 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 igäüks eraldi artiklit väärida, 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 elementide 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 mudelitega.
- Luua tutvustav ülevaade seonduvatest küsimustest
- Vormistada omamoodi projekt edasisteks uuringuteks

Oma eelmises samateemalises kirjatöös (Pihlakas, 2005a) loetlesin üles seniseid tehisintellekti uurimissuundi, tutvustasin arengulist lähenemist, kirjeldasin mõtlemise tasandeid Toomela ja Vögtotski järgi ning pakkusin välja mudelid mõtlemise tasanditele “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 / täiendused ei oma üksi märgatavat väärtust. Vastandatuna sellele, “intelligentne programmeerija” saab nõ ühe õhtuga luua programmi, kus on sadu omavahelisi sõltuvusi. Arenguline lähenemine ongi siinkohal tasakaalustuseks, et ülalkirjeldatud konstrueerimisel oleks siiski toetuspinda all ja versta postid märgitud.

Seejärel 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 asuvasse teemadesse laskudes.

Käesolevas töös on seega kesksel kohal mudeli tasandid 2 ja 3, vastavalt 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 alguses toon uuesti tasandite loetelu ning Lisas lühikirjelduse, nii nagu neist modelleerimise suhtes relevantsete omaduste kokkuvõtte olen teinud, sest see info on baasiks kõigele edasisele. Seejärel järgnevad ülesandepüstitused mudeli tasanditele 2c ja 3 ning seejärel tutvustan lahendusi nõutud modelleerimiseks.

Töö ei käsitle mudeli võimalikke seoseid närvisüsteemi arhitektuuri ja dünaamikaga.

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

- Tajude hierarhiline ülesehitus. See ei ole käsitletud käesoleva töö mudelis, kuid

- Sündmuste ühesuunaline ajaline järgnevus (põhjus ennustab tagajärge, kuid mitte vastupidi. See võib paista elementaarne, kuid mitte kõik võimalikud mudelid ei kasuta seda eeldust).
- Mudel õpib, ennustab ja üldistab juba peale minimaalset kogemust. Täiendavad kogemused võimaldavad täpsustada seaduspärasid ning õppida erandeid. Taoline üldistamine toimub stiimulite tugevuste, kestuste, ennustatavate tagajärgede määrade, ... **TODO** juures.
- Ennustused, mida mudel teeb, ennustavad esialgu sündmuse määra (vastavalt *blocking/nonredundancy* nõudele), 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 punktidele. Analoogselt kolmanda eeldusega on võimalik õppida erandeid.
- On piisav või isegi kasulik, et õpitakse ainult seoseid, mis vastavad *nonredundancy* nõudele. Selle eelistest lähemalt diskriminantsete stiimulite peatüki lõpupoole.

Neurovõrkudele ei ole seni autorile teadaolevat probleemilahendajaid ehitatud. Kuid just neurovõrgud on need mudelid, mis suudaksid paremini olla kontaktis sensoorse/motoorse pidevalt muutuva ning mitmekesise maailmaga ning mis võimaldavad ka suuremat paindlikkust objektide käsitlemisel, näiteks erinevate objektide *sarnaste omaduste taipamist*.

### **Miks modelleerida mõtlemist**

- Ülesehitav lähenemine võimaldab küsida küsimusi, mida kirjeldavas lähenemises ei taibanud küsida. On võimalik luua tehislikud katsealused:
  - Kes on rangema kontrolli all.
  - Kelle protsessid/mõttekäigud jälgitavad aeglustatult. Keerukate protsesside mõistmiseks on see hädavajalik (Wagoner & Valsiner, 2003).
  - Võimaldab saada infot sisemiste vastasmõjude, andmete liikumise kohta.
  - Lisada või eemaldada elemente ja nende omadusi.
- Rakendused robotikas, autonoomsed süsteemid:
  - Naturaalse mõtlemise modelleerimine on tõenäoliselt lihtsaim ja harivat tulemust andev viis luua esimesi primaatidele või inimestele lähedase võimekusega tehisintellekte. Selle põhjal saab hiljem juba luua kohati veel puhtamaid, optimaalsemaid ja disainitumaid intellekte. Senise mudeli põhjal ma ei toeta väiteid, nagu oleks naturaalne mõtlemine suurelt jaolt evolutsiooniline segadus ja “spageti kood”.

- Tõhusam alternatiiv mõnede planeerimisülesannete lahendamiseks, mida hetkel tihti lahendatakse kasutades *reinforcement learning'ut* – tegevus pidevas füüsilises, motoorses maailmas vs diskreetse maailmas (vaata Lisa).
- Väidetavalt selgem ja turvalisem paradigma autonoomsete süsteemide eesmärkide, eesmärgipärase käitumise disainimiseks (vaata Lisa).

## Tasandite kirjeldus

Eelmises kirjatöös (Pihlakas, 2005a) 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 Lisast. Töö põhitekstis toon Piaget sensomotoorse staadiumi alatasandite loetelu ning veel mõned analoogsed.

Lisaks tahaksin kohe tuua välja olulise erinevuse Piaget ning Vögotski lähenemiste vahel, mis puudutab märgilise mõtlemise tasandeid – nimelt Piaget järgi mõtlemise areng eelneb kõne arengule ning varajane kõne on egotsentriline (viide), kuid Vögotstil eelneb mõtte arengule kõne areng, kus kõne on sotsiaalne – sotsialiseerumisel kõne internaliseerumise läbi areneb ka mõtlemine (viide). See on oluline perspektiiv silmas pidada käesoleva mudeli seisukohast – nimelt mudel peaks olema võimeline märgilise mõtlemise tasanditele arenema läbi sotsiaalse suhtluse – õppima ennustama, mis teised talle ütlevad, ning hakkama ise endale neid asju ütleva, ehk internaliseerima, ja hakkama läbi sisemise kõne oma tegevust ja mõtlemist suunama; samuti õppima internaliseerima muid väliseid protsesse ning õpitud skeemidena sisemiselt läbi töötama (Toomela, 1996; Vygotsky, 1994a; Vygotsky & Luria, 1994b).

See pole ainult nõue, vaid ka võimalus, kuna ennustamine on mudeli ülesanne niikuinii ning märgilise mõtlemise eeldused tekivad samuti muudel põhjustel.

Vögotski ja Piaget kirjelduse erinevus ei tähenda, et Piaget siinkohal täiesti eksib – täiesti võimalik, et kõne ja teatud mõtlemise võimete arenguks tööpoolest on tarvis siiski mõningaid täiendavad arenguid naturaalses mõtlemise mehhanismides, kuid see selgub edaspidi.

Samuti ei pruugi Piaget ja Vögotski seisukohad tegelikult olla nii vastandlikud, kui kirjeldasin. Ka Vögotski jaoks sai naturaalne mõtlemine edasi areneda, kuna ta rääkis kahe arengusuuna põimumisest, mitte ühe asendumisest teisega (viide).

Võimalik, et kõne ja teatud mõtlemise võimete arenguks on tarvilik või kasulik mõningate täiendavate omaduste areng naturaalses mõtlemise mehhanismides, kuid vastavad detailid selguvad edaspidi, märgilise mõtlemise modelleerimise uurimuses.

Pigem on see tekst mõeldud sellise tõlgenduse ühe poole vältimiseks (Piaget: “sisseehitatud mehhanismid on need, millel põhineb mõtlemise võimete areng, nende abil ja hiljem areneb väline kõne”) ja teise poole võimaluste esile toomiseks (Võgotski: “mõte areneb kui internaliseeritud kõne”).

Internaliseerimine on oluline arengusuund silmas pidada: praegune mudel tegeleb ennustamisega ning internaliseerimine on oma olemuselt samuti ennustamine. Kui internaliseerimine võimaldab selgeks õppida teatud tüüpi sündmusi, mis juhivad tähelepanu, tegevust ja mõtlemist algselt väliselt (märgiline suhtlus) (viide, Võgotski, Educational Psychology raamat), siis ei ole vaja tingimata eeldada keeruliste “mõtlemisoskuste” kaasasündinud olemasolu, ega neid mudelisse sisse programmeerida. Samas Piaget jäigema tõlgendamise järgi arendada mõtlemist enne suhtlust eeldaks rohkem mudelile oskuste sisseprogrammeerimist.

Et positsioneerida käesolevat mudelit mõtlemise arengu laiemas kontekstis ja näidata edasise arenduse suunda ja järjekorda, toon naturaalse ja kultuurilise mõtlemise tasandite loetelu:

Naturaalne mõtlemine:

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

Kultuuriline, märgiline mõtlemine:

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

Tasandite kirjelduse jaoks vaata Lisa või Pihlakas, 2005.

### **Lühike selgitus eelnevate tasandite kohta.**

Naturaalne mõtlemine on see, mis esineb loomariigis ning on kaasasündinud, evolutsioonis arenenud. Naturaalne mõtlemine on iseäralik selle poolest, et andmete liikumine toimib teatud seaduspärasuste järgi ka märgilise programmi ja kogemuse puudumisel (paljudel loomadel polegi märgilise mõtlemise võimet või nad ei kasuta seda – (viide, Tomasello)).

Kultuuriline mõtlemine on info liikumise erijuhtum, kus osa andmete vahelistest seostest vastab märgi definitsioonile (definitsioon + viide). Märgiline protsess, mis jookseb naturaalse mõtlemise mehhanismidel, kasutab selle võimalusi ning juhib andmete liikumist selles. Seega märgiline mõtlemine on kui tähelepanu liikumise programm, naturaalne mõtlemine kui riistvara (mis sisaldab andmeid).

Naturaalne vs märgiline: Naturaalne kujuneb evolutsioonis sammhaaval teatud “sisseehitatud” omadustest, nende võimest töödelda sensomotoorseid kogemusi. Hiljem toimub märgiliselt mõtlema hakanud indiviidi areng õppimise käigus teatud muude omaduste/protsesside lisandudes.

Metamõtlemine on märgilise mõtlemise erijuht, kus märgiline protsess juhib muuhulgas otseselt iseennast (viide). Kaudne analoog arvutimaailmas oleks koodi genereerimine, modifitseerimine. Metamõtlemine tekib alates teadusmõistelise mõtlemise tasandist. Tavaline märgiline mõtlemine ei ole otseselt võimeline ennast süstemaatiliselt jälgima ja muutma, kuigi igal tegevusel loomulikult on mõju potentsiaalselt mistahes andmetele mudelis.

### ***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, 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: 5. alatasandi võimed on ühtlasi objekte omavahel suhestava tööriistakasutuse võimaldajaks (Tomasello, 1997).

6. alatasandi võime kui objektide püsivuse võime peaks olema Tomasello, 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, 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, 1997). Samale järeldusele vihjab imetajate, näiteks koerte, kes manipuleerivad objekte väga vähe, objektipüsivuse võime (Tomasello, 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 inimesed 6-nda alatasandi kõik võimed on ikkagi representatsioonidel põhinevad. (Tomasello, 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 isegi tööriista muuta või valmistada, nõnda, et see vastaks olukorra nõuetele (Tomasello, 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, 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 oleks 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>1</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 momenti 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).

Tegevuste mentaalsete representatsioonide termini seonduvust mudeliga ei oska hetkel detailsemalt kommenteerida, kuid võimalik, et seotud märgilise mõtlemise protsessidega.

### **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. Kinnitavad kontseptid: absoluutsed ja relatiivsed
7. Tingimuslikud kontseptid, konjunktiivsed kontseptid, disjunktiivsed kontseptid
8. Bikonditsionaalsed kontseptid: "siis ja ainult siis"

(Tomasello, 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, 1997).

6. taseme poolest on mudelis olemas absoluutsed kontseptid, juhul kui mudel ühendada tajusüsteemiga, mis on võimeline tajusid klassifitseerima (osa klassifitseerimist on ka

---

<sup>1</sup> [http://scholarpedia.org/article/Mirror\\_Neurons](http://scholarpedia.org/article/Mirror_Neurons)

järgnevas mudelis kirjeldatud). See võime peaks samuti imetajatel ühtmoodi esinema (Tomasello, 1997).

6. taseme relatiivsed kontseptid – võimaldab kahe stiimuli võrdlemist ning on eksklusiivselt primaatide võimeks ja vastab Piaget tertsiarsetele operatsioonidele (Tomasello, 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 oleks 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).

8. tase vajab edaspidist täpsustamist.

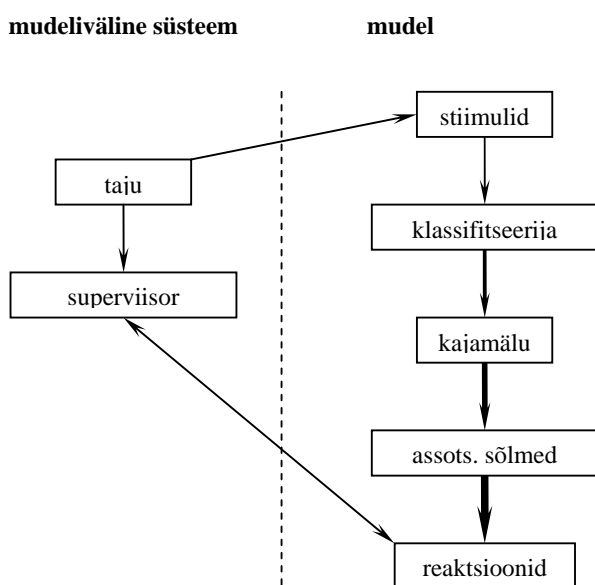
## Mudel

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.

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

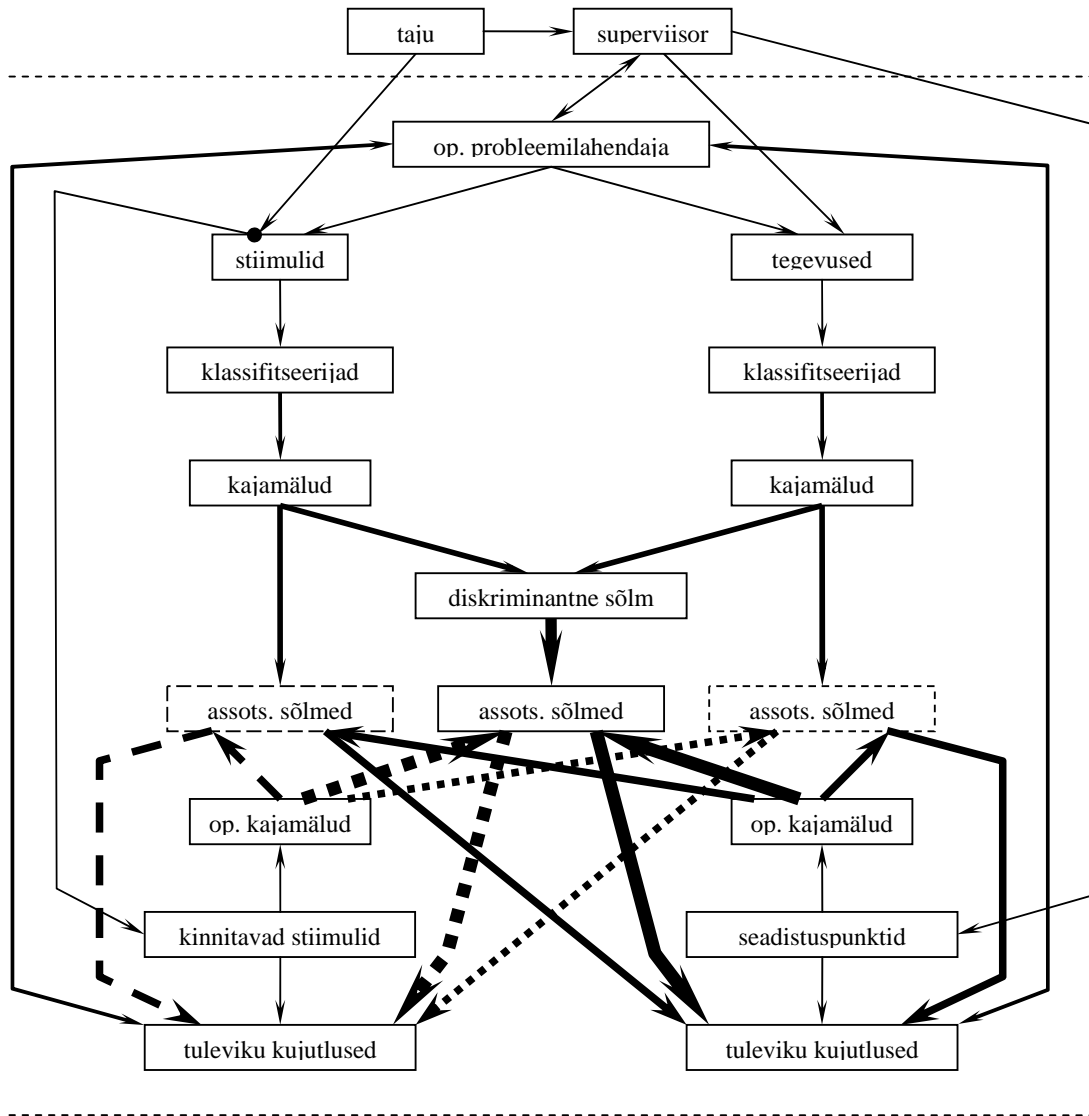
### Klassikaline tingimine:



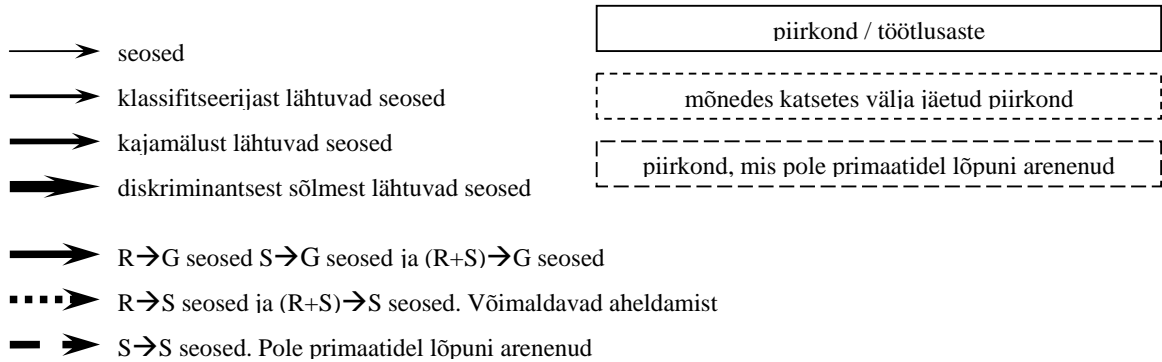
### Piirkondade selgitused:

- 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 / nende liigendustest. Võimaldab seostada hilisemaid reaktsioone varasemate sündmustega.
- Assots. sõlmed – assotsiatiivsed sõlmed, säilitavad infot stiimulite / nende liigenduste tähenduste ehk vastavate reaktsioonide kohta.
- Reaktsioonid – mudeli väljundid, mida superviisor treenib ja mudel ennustama õpib.

Operantne tingimine:



**Legend**



Operantse tingimise jooniselt on ära jäetud klassikalise tingimise seosed.

**Lisandunud piirkonnad:**

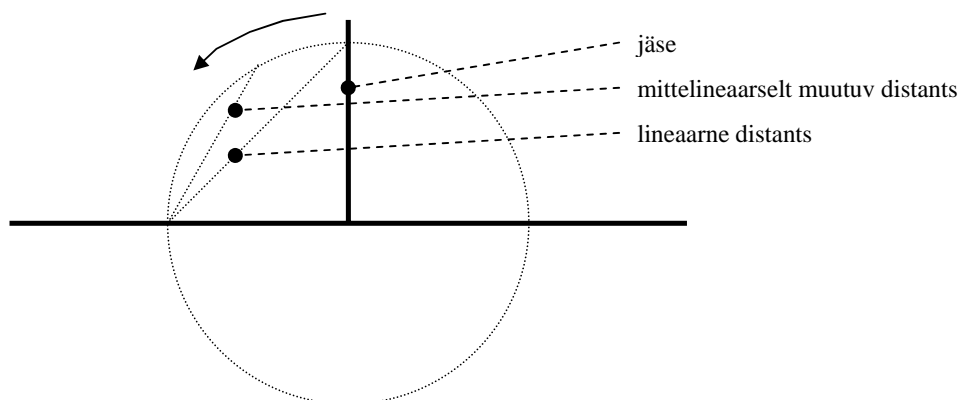
- Diskriminantne sõlm – seostab tegevuste ja diskriminantstiimulite kombinatsioone mudeli sisesteks stiimuliteks / representatsioonideks.
- Op. kajamälud – operantsed kajamälud. Säilitavad samuti ajalise jälje seadistuspunkti või kinnitava stiimuli varasematest seisunditest, kuid töötavad natuke erineva 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 seadistuspunkti tegeliku väärtuse eelistatud väärtusel säilitamise ehk operantse eesmärgi suhtes optimaalsete tegevuste leidmisega ning kinnitavatele stiimulitele energiatega määramisega (selgitus vastavas peatükis).

## ***Stiimulite ja tegevuste kanalite eeltöötlus***

### **Klassifitseerija**

Stiimulite ja tegevuste tugevused on mudeli jaoks esitatud analoog-kodeeringus ehk reaalarvudena. Need väärtused klassifitseeritakse. Hetkel stiimuli tugevus tähendab stiimuli kestust (see arvutatakse superviisoris).

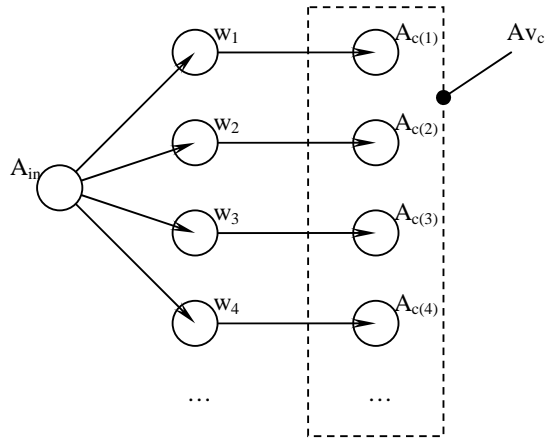
Klassifitseerimine on kasulik, kuna võimaldab õppida tegevusi, mille tulemused pole lineaarsed. Näiteks kehtib see paljude ringliikumiste kohta, mille eesmärk on jäset või midagi muud viia soovitud punkti.



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

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

Klassifitseerija ehitus:



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

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

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

$A_{in}$  – 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 muutust; variant (b) on kasutusel juhul, kui  $A_{in}$  esindab stiimuli integraali või kestust või stiimuli tugevust, kuna:

- Integraali või kestuse puhul: tugevuse kestust saab ainult pikendada või lõpetada. Selleks, et õppida käimasoleva tugevuse (mitteoodatud) tulemusetust, tuleb õppida, et möödunud tugevuse kestused ei andnud tulemust. Kui sellist õppimist ei võimaldata, siis omistatakse tulemusetus jätkuvalt kasvavatele kestustele, mitte nendele, mis olid tugevuse motivaatoriks, ning mudel üritaks tugevust järjest edasi. Seega kuigi mudel algul üldistab pikemalt lühemale tugevuse kestusele, on võimalik õppida erandeid ning selles seisnebki intervall-kinnituskava.
- Stiimuli tugevust esindava kanali korral annab sarnase omaduse, kui klassikalises tingimises – tugevamat stiimulit õpitakse rohkem (viide: Rescorla-Wagner). See tähendab käesolevas mudelis: tugeva stiimuli tähendust üldistatakse ka nõrgemale stiimulile, kuid nõrgema stiimuli esinemisel aktiveeritakse vastavalt vähem vektori komponente ja vastav ennustus tuleb nõrgem. 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 valem võiks olla kasutusel ka stiimuli tugevuse muutu esindava kanalis. See vajab edaspidist täpsustamist, käesoleva töö mudeli variandis on kasutusel ainult integraali esindav kanal. Samuti võimalik, et osades tajukanalites ka stiimuli tugevust esindav kanal kasutab valemit (a).*

- **TODO:** vaata ka

<http://digitalcommons.libraries.columbia.edu/dissertations/AAI3115343/>

<http://wwwlib.umi.com/dissertations/preview/3115343>

Integraali 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 printsibi alusel, näiteks Kohoneni iseorganiseeruva kaardi (SOM) valemiga (viide), 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.

On samuti mõeldav, et valemiga (a) analoogselt töötava klassifitseerija väljundkaalude vektoris aktiveeritakse mitu elementi. Klassifitseerijaks võib olla SOM, Grossbergi ART (viide), radiaalbaas-funktsioon-võrk (viide), Numenta HTM (viide), analoogselt ka autori seminaritöös kirjeldatud hierarhiline struktuur (Pihlakas, 2005) või mõni muu superviseerimata õppimisele ning vektorit väljundväärtusena võimaldavaks kohandatav komponent.

Hetke mudelis on sisend skalaarne, kuid stiimulite mustreid töötlevas võrgus 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. Samal ajal sensitiseerimise nähtus on pigem seotud neurokeemiliste protsessidega, mitte seoste struktuuriga <sup>2</sup>.

SOM tekitab ajuga sarnaseid jaotusi kõige tõhusamalt <sup>3</sup>, 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

<sup>2</sup> <http://en.wikipedia.org/wiki/Sensitization>

<sup>3</sup> <http://mlab.uiah.fi/~timo/som/thesis-som.html>

väljundeid, siis aktiveeriks klassifitseerija vaid ühe elemendi vektoris ning see kaotaks liiga palju informatsiooni toimunud sündmuste kohta. Arenduse käigus selgus, et mudel töötab esimese variandiga, kuid kehvasti või üldse mitte teise variandiga.

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

## Kajamälu

Vajalik, et assotsiatiivne võrk saaks luua seoseid sündmuste vahel, mis järgnevad teineteisele erinevate ajaliste hilinemistega. On tarvilik põhjus-tagajärje seoste õppimiseks.

“Kajamälu” tähendab siinkohal, et mälu sisaldab sündmuse kaja. Ka mõnedes teistes klassikalise tingimise mudelites (viide: Balkenius) on kasutatud analoogset konstruktsiooni ning nimeks pandud *delay-lines*.

Kajamälu siinkohal ei tähenda, et tegu oleks auditoorse modaalsusega. Mudelis on igal sisendkanalil oma kajamälu, ning nende hulk on kohati 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 sündmuse olevikuga – õppida põhjus-tagajärje seoseid. Pihlakas, 2005 mudel näiteks õppis seostama vaid täpselt samaaegseid sündmusi. Stiimuli mõju võib kesta edasi peale stiimuli lõppu (viide: Skinner). Sisuliselt vajalik operantseks tingimiseks, kuna kinnitav stiimul 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 stiimuli lühikeste ajaliste järgnevuste tähendusi.

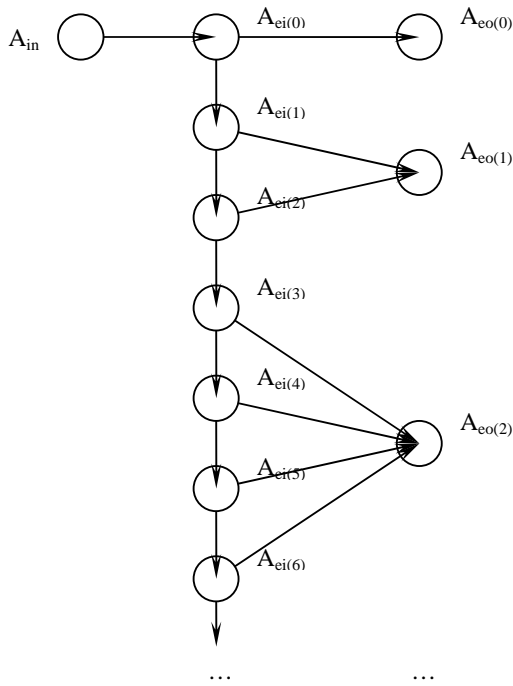
Igal stiimuleid esindaval ning tegevusi esindaval kanalil on “kajamälu”, rakud, mis esindavad vastava kanali varasemaid seisundeid. Visuaalsete rakkude puhul vastaks see ligikaudu seni kasutusel olnud terminoloogias ikoonilise mälu mõistele, auditoorsetel rakkudel kajamälule [vaata: viide, Tulving].

Kuna mudel esialgu ei täpsusta, kas töödeldav info on visuaalne või auditoorne ning võiks olla potentsiaalselt ühte kui teist, siis valisin sõna “kajamälu”, kuna see on iga üksiku kanali puhul adekvaatsem – üks kanal ei saa sisaldada ikoonilist infot. Tuleks tähele panna, et siin töös kasutatav “kajamälu” mõiste erineb siiski klassikalisest “kajamälu” mõistest, kuna puudutab konkreetselt kanali tasandit, mitte auditoorselt terviklikku kajamälu, kuigi viimane tõenäoliselt kerkib esile just taolisest rakutasemel töötavast kajamälust [viiteid?].

Kajamälu eristustihedus väheneb, mida kaugemale selles olevad sündmused käesolevast hetkest ajalisel jõuavad. Omadus on motiveeritud sarnasest omadusest naturaalses õppimises

(viide). Eristustiheduse vähenemise kõver aga ei ole käesoleval hetkel kuidagi faktiliselt põhjendatud vaid valitud mugavuse järgi.

Eristustiheduse vähenemine võimaldab logaritmiliselt väiksema andmemahtu 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 / 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.

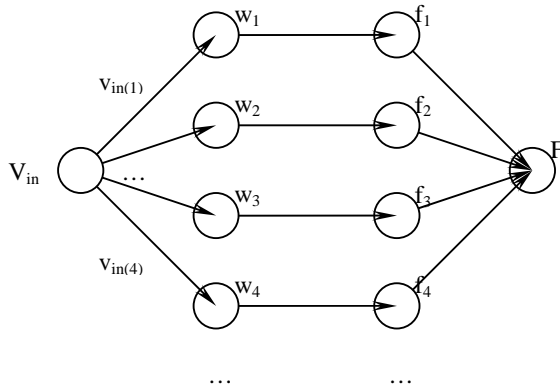
$A_{eo(i)}$  – kajamälu teine aste (“väljundrakud”), kajamälu väljundpunktid. Ajaliselt kaugemad sündmused on esindatud madalama eristustihedusega kui ajaliselt lähemad. Vastavalt kajamälu sisendile on andmetüüp skalaarne või vektor.

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  $A_{eo(i)}$ ”. 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 või koguni ainult minimaalse kestuse ja tingitava intervalli kombinatsioonile, juhul kui tingitud stiimulit esitatakse vaid hetkeks.

## 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/faktor.

$s_i$  – andmepunkti õpitud väärtuste summa

$n_i$  – andmepunkti õpitud väärtuste samplite arv

$wv_i$  – andmepunktile vastava faktori skalaarne väärtus

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

$F$  – skalaarne väljundväärtus, ennustatava tingimata stiimuli faktor

## Klassikaline ja operantne tingimine

Klassikalise tingimise terminid:

- *blocking*
- kustumine
- kiirendatud taastumine
- iseeneslik taastumine
- esmane ja teisene seos

Operantse tingimise terminid:

- *interval*-kinnituskava
- *ratio*-kinnituskava
- *variable-interval* kinnituskava
- *variable-ratio* kinnituskava
- kinnitus
- vältimine
- põgenemine
- aheldamine
- diskriminantstiimul

Klassikalise ja operantse mõtlemise terminite tutvustust vaata Lisast.

## Klassikaline tingimine

Nõuded:

- kiire õppimine,
- *blocking* ehk *nonredundancy*
- 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, 1998 ning Johansson & Lansner, (2002?) - viide.

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 saakski seda omadust väljendada.

Käesoleva mudeli ülejäänud omaduste katsetulemuste kvalitatiivsest seisukohast ei oma õppimise valemi valik tähtsust (viide katsetulemustele), mis on hea, kuna näitab, et ülejäänud omadusi oli võimalik katsetada võrdlemisi sõltumatult õppimise valemist. Küll aga võib oletada, et õppimise valemi valik mõjutab mudeli käitumist keerulisemates oludes kui katsetingimused, mis keskendusid spetsiaalselt võimalikult vähestele omadustele korraga.

### ***Ennustusväärtuse lisatingimus***

(contingency explanation, nonredundancy), näide

1960ndatel Robert Rescorla näitas, et CS – UCS seose tekkimiseks ei piisa nende paari panemisest. Ta seadis lisatingimuseks, et CS peab olema hea signaal UCS esinemisele. Seda lähenemist nimetatakse *contingency explanation of classical conditioning*. (TODO: ei saa aru, mis moodi see erineb varasemast teooriast?)

Teised uurijad seejärel leidsid (TODO: kes täpsemalt?), et

Tingimata stiimuli CS poolt pakutav info peab olema *nonredundant* – lisama midagi infole, mida pakub mõni teine, varasemalt tingitud stiimul. Vastavat eksperimenti nimetatakse *blocking* (viide).

## Valemid

Tingitud stiimuli ennustamine toimub:

$$S_{ucs} = \sum_i^n F_{cs(i)}$$

kus

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

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

$n$  – tingitud stiimulite arv

Järgnevalt toon kaks valemit õppimise kahe aspekti – aeglane kustumine (kasutuses hiljem iseeneslikus taastumises) ning *blocking* – teostamiseks ning viimaks valemi, mis neid aspekte ühendab. *Blocking* ja mitme muu (TODO: ehk vaja neid ka siin töös üles lugeda?) aspekti kohta on loodud mudeleid (viited: Balkenius, Johansson & Lansner, etc). Käesoleva mudeli 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

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

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

ehk

$$F_{cs(i)}(t+1) \leftarrow \frac{s_{f(i)}(t) + S_{ucs}}{n_{f(i)}(t) + 1}$$

siit valem klassifitseerijat kasutava mudeli jaoks:

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

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

kus

i – faktori / assotsiatiivse sõlme indeks

(t) – diskreetne ajahetke indeks

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

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

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

$v_{in(i,j)}$  – assotsiatiivse sõlme i sisendvektori skalaarne element indeksiga j

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

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

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.

Näide dünaamika kohta:

**TODO: tabel**

Binaarsete (väärtused kas 0 või 1)  $CS_i$  ja UCS korral ning eeldusel, et ühekorraga on aktiivne üks CS, ennustab valem sisuliselt Bayesi tõenäosust et stiimuli  $CS_i$  korral esineb UCS.

See valem ei vasta *blocking* / *nonredundancy* nõudele. Samuti ei võimalda see kiiret taastumist ja kiiret kustumist.

### **Blocking / nonredundancy valem**

Tingitud stiimuli seose õppimine kooskõlas *blocking / nonredundancy* nõudele 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:

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

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

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

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

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 / alafaktorite summa

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

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

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

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

Alafaktori väärtuse muutmisel kasutatakse samuti kaalu  $v_{in(j,j)}$ .

Valem töötab põhimõttel: suurema samplite arvuga faktorite väärtusi muudetakse proportsionaalselt vähem kui väikese samplite arvuga faktorite väärtusi (samplite arv siinkohal tähendab faktori järgmiseks tsükliks omandatavat samplite arvu, seega

senitreenimata faktori samplite arv on 1), kuid siiski muudetakse kõigi faktorite väärtusi alati niipalju, et nende summa vastaks treenimisel kasutatud  $S_{UCS}$  väärtusele. Seega pole sellist piirangut faktori väärtuse muutmisele, nagu aeglase kustumise valemis oli. Kõik faktorid võivad ühe ajahetke jooksul omandada senisest märgatavalt erinevad väärtused.

**TODO: õppimiskiiruse pidurdamise konstant valemisse**

Näide dünaamika kohta:

**TODO: tabel**

Sellel valemil pole vastavust iseenesliku taastumise aeglase kustumise nõudele. See valem omab *blocking*, kiire õppimise, kiire kustumise ja iga kord kiireneva taastumise omadusi.

Juhin veelkord tähelepanu, et UCS ennustus arvutatakse faktorite summast. Faktorite õppimine pole sama, mis tagajärgede tõenäosuste õppimine (vaata jaotust: 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 määraga  $> 1$ .

Siiski on *blocking* 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  $> 1$ ) ennustustele:

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

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

## Aeglase kustumise teine valem, vastab *blocking* nõudele

Ühendamiseks ülalkäsitletud kaks omadust, ühendan ülalolevates õppimise valemites olnud korrigeerimisparameetrid  $x$  ja  $S_{ucs}$ . Aeglase kustumise esimese valemi omaduseks oli, et iga faktori väärtus saab muutuda vaid piiratud määral, sõltuvalt faktori senisest samplite arvust ning  $S_{ucs}$  väärtusest. Rakendan sarnase piirangu *blocking* omadusega valemisse:

üldistatult

$$x = \arg \min \left\{ \text{abs} \left[ S_{ucs} - \sum_i \frac{s_{f(i)} + lf(x, S_{ucs})}{n_{f(i)} + 1} \right] \right\}, \text{ kus faktori muutu piirav funktsioon}$$

$$lf(x, l) = \begin{cases} x, & \text{kui } x \leq l \text{ ja } l \geq 0 \\ x, & \text{kui } x \geq l \text{ ja } l \leq 0 \\ l, & \text{ülejäanud juhtudel} \end{cases}$$

on eeldatud, et  $x$  ja  $l$  on alati sama märgiga.

seega võib kirjutada ka:

$$lf(x, l) = \min(\max(x, -l), l)$$

Klassifitseerija sisaldavas mudelis:

$$x_s = \arg \min(\text{abs}\{S_{ucs} - \sum_i \sum_j f_{i,j}^*\})$$

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

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

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

argmin() osa tuleneb sellest, et piiratud õppimise tõttu enam pole võimalik võrrandis faktorite summat alati  $S_{ucs}$  väärtusega võrduma panna. Kui kõigi faktorite  $n$  on piisavalt suur, siis liigub faktorite summa õppimise käigus  $S_{ucs}$  poole ( $S_{ucs}$  ning faktorite summa vahe absoluutväärtus väheneb), kuid siiski jääb vahe nullist suuremaks, erinevalt algsest *blocking* õppimise valemist. Sellisel juhul toimib valem, nagu ülal olnud esimene aeglase kustumise valem. Kui esineb uudseid faktoreid (nende  $n$  on väike), siis nendele toimib käesolev valem kui *blocking* õppimise valem.

Ülalolevat  $x$  või  $x_s$  väärtust on võimalik leida ilma rekursiivse arvutuseta (ilma erinevaid väärtusi läbi proovimata), käesolevas mudeli teostuses on  $x_s$  leidmine teostatud sorteerimise ning seejärel ühetasemelise for-tsükliga. Paralleelarvutuses on võimalik leida  $x$  või  $x_s$  väärtus ühe tööttsükliga.

Analoogselt eelmistele õppimise valemitele käesolev valem rakendab kõigile assotsiatiivsetele sõlmedele ühe globaalse korrektsiooniparameetri. Erinevus seisneb selles, et korrektsiooniparameetri mõju võib olla sõlme sisese info järgi lokaalselt piiratud ning parameetri enda arvutamine käib keerulisemalt kui eelmistes valemites.

Selle valemi eesmärk on saada lahti kustumise/õppimise kiiruse konstantsest parameetrist (erandiks esimese *blocking* valemi kiirenev taastumine), muuta õppimise ja kustumise kiirus sõltuvaks varasema kogemuse hulgast.

Näide dünaamika kohta:

**TODO: tabel**

**TODO: võrdlus Rescorla-Wagneri mudeliga (viide).**

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 Skinner, 1957). See väide kahjuks üksi veel ei selgita, 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 lisandub samplite arvu mõju.

Ükski Balkenius, 1998 (viide) kirjeldatud mudelitest ega Johansson & Lansner, 2002? (viide) mudel ei oma iseenesliku taastumise võimet.

TODO: iseenesliku taastumise täpsem kirjeldus + viide.

Sissejuhatuseks pakun kaks 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 / keskmistamine püsिमäluga, mille tulemusena püsिमälu seose tugevus on küll nõrgenenud, kuid omaltpoolt on lühimälus olev mõjuv seose tugevus kasvanud, sest püsिमälus olnud seos oli tugevam kui lühimälus. See on analoogne ajalise konteksti muutumise ideele, mis vastandub füüsilise konteksti muutumisele (viide:  
<http://www.learnmem.org/cgi/content/full/11/5/485>)

Järgnevalt tutvustan, mis moodi eelnevalt kirjeldatud valemeid saab iseenesliku taastumise omaduse tekitamiseks kasutada viisil, mis mõnel määral sarnaneb teise toodud hüpoteesiga.

### TODO:

+ iseenesliku taastumise omadused, Rescorla, 2004

+ paralleel: Devenport, 1998 + eelised

### Valem: TODO

Iseeneslik taastumine eeldab kahe mälusüsteemi olemasolu.

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

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

Käesoleva iseenesliku taastumise valemi eesmärk pole väita, et iseeneslik taastumine sõltub ajast või ainult ajast, vaid modelleerida paari iseenesliku taastumise omaduse saavutamiseks tarvilikku aspekti – mitme mälujälje olemasolu ning mehhanismi nende vahel valimiseks, jättes eraldi küsimuseks, mis seda mehhanismi kontrollib. Käesolevas mudeli teostuses on kontrollijaks aeg.

## Operantne tingimine

Tänu operantsele tingimisele isend saab tegutseda aktiivselt vastavalt oma õpitud ootustele-motiividele, mitte ainult refleksidena välismaailma stiimulitele. + Vaata viiteid @ Alg Mind lk 152

Nimetatakse ka *instrumental conditioning*. Õpib tegevuste ja tagajärgede vahelisi seoseid, võrrelduna klassikalise tingimisega, mis õpib sündmuste vahelisi seoseid ning muuhulgas kaudselt stiimuli ja tegevuse vahelisi seoseid.

**Klassikalise ja operantse tingimise tüübid.**

**TODO: loetelu.**

Kinnitustüüpide tutvustust vaata Lisast.

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

Operantse tegevuse eesmärk on saavutada või säilitada sisemine homöostaas, hoida väärtus tasakaalupunktis teatud eesmärgis-näitajas, mis asub seadistuspunktis (viide). 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 deprivatsiooni korral üritab 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 korral 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).

Tulenevalt klassikalise tingimise valemitest on analoogsed omadused ka mudelil.

Karistuse läbi õpitud kogemus võib mudelis mõnel juhul saada hiljem tegevust motiveerivaks kogemuseks ning kinnituse läbi õpitud kogemus vältimist põhjustavaks, 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 mootorika puhul on seadistuspunktid pidevas muutumises. See, kas mingi tingimine oli kinnitav või karistav, ilmneb alles konkreetse käitumisolukorras, mitte tingimise ajal.

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 korraga mitu erinevat kinnitajat (mitu seadistuspunkti saavutab eelistatud väärtuse), siis see tegevus saab 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*, mille eesmärk on maksimiseerida teatud näitajat (viide).

*Reinforcement learning* põhimõttel töötavaid komponente annab küll vabalt ühendada nõnda, et nad kokkuvõttes midagi tasakaalus hoiaksid, kuid suur osa *reinforcement learningu* paradigmat ja sellest tulenevast mõtteviisist on mõjutatud siiski hinnangu maksimiseerimise vaatenurgast. Psühholoogia ajaloost vastab sellele pigem Thorndike teooria (viide) ning tema teooriat raamatus (viide) ka tsiteeritakse, Skinnerit mitte.

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

Operantse tegevuse eesmärk on sarnane sellele, millega tegelevad kontrolliteooria mudelid. Analoogselt enamlevinud lihtsale kontrollisüsteemile PID-kontrollerile (viide) 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.

Teine erinevus *reinforcement learning*'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 distantsi mõistet.

Käesolev mudel õpib ühe seadistuspunkti väärtuste vahel liikumiseks väärtuse muutumise iga erineva tegevuse korral.

Seega ühe seadistuspunkti või kinnitava stiimuli võimalike väärtuste / seisundite grupi elemendid saab paigutada intervallskaalale ning üldistada õpitud samme mistahes selle skaala punktide vahel liikumisele. Kuid kui mingite punktide vahel liikumisel 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**.

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 järgmisesse 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 energia ehk operantse energia valem:

$$Discr = A_o - P_o$$

$$E_o = abs(Discr)$$

$A_o$  – seadistuspunkti tegelik väärtus (*actual value*), antakse mudelivälise süsteemi 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 arvutatud valemiga:

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

(t) – käesolev ajahetk

(t-1) – eelmine ajahetk (eelmine mudeli töötsükkel)

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). 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 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> – ennustatud seadistuspunkti i väärtus

A<sub>o(i)</sub> – seadistuspunkti 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 (tegevuste) määrad, et kõigile seadistuspunktile ennustatud energia oleks järgmises ajahetkes minimaalne.

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

E<sub>p</sub> – kõigile seadistuspunktile ennustatud energiategi summa

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

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

Mudel kasutab tegevuste määrade leidmiseks *simulated annealing* algoritmi (viide).

Seega

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

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

$$E_p[Av(t+1)] = Ef\{g_o[Av(t+1)]\}, \quad \text{kus}$$

$$A_{op}(t+1) = g_o[Av(t+1)] \quad \text{ning}$$

$$Ef(g_o) = E_p$$

$Av(t+1)$  – tegevuste määrade vektor järgmiseks ajahetkeks

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

$n$  – tegevuste hulk

$E_p[Av(t+1)]$  – seadistuspunktile ennustatud summaarne energia tegevuste vektori  $Av(t+1)$

korral

$A_{op(i)}$  – ennustatud seadistuspunkti väärtus

$Ef(g_o)$  – mudeli seadistuspunktide summaarne energia ennustuse  $g$  korral

$g_o(Av)$  – mudeli seadistuspunktide ennustamise funktsioon tegevuste vektorist  $Av$  operantse kujutlemise tööfaasis

$E_p$  – kõigile seadistuspunktile ennustatud energiategade summa

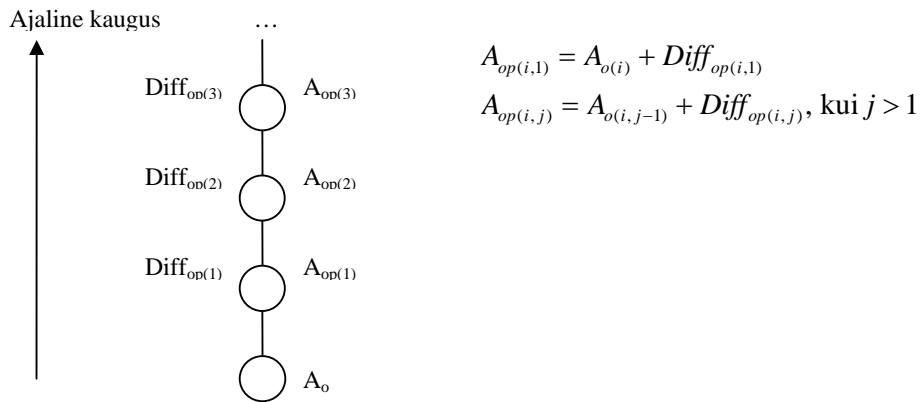
**TODO:  $Ef(g)$  ja  $g(Av)$ ,  $A_{op}$ ,  $E_p$  sümbolika ning indeksid**

**TODO: Simulated annealing kasutamise toetuseks uurimus ajus toimuvast**

**“juhuslikust” laenglemisest (viide) ja selle analoogia mudeliga.**

### **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.



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

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

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

Iga järgnev 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

$$A_{op(i,j)} = A_{o(i)} + \sum_{k=1}^j Diff_{op(i,k)}$$

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

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

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

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

$t$  – tuleviku ajahetke indeks

$j$  – tuleviku ajavahemiku lõpu indeks

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

$Diff_{op(i,j)}$  – operandi  $i$  ennustatud väärtuse muut ajavahemiku  $j$  jooksul

$Diff_{op(i)}$  – ennustatud operandi  $i$  väärtuse muut

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

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 ja *reinforcement learning*'u erinevus:

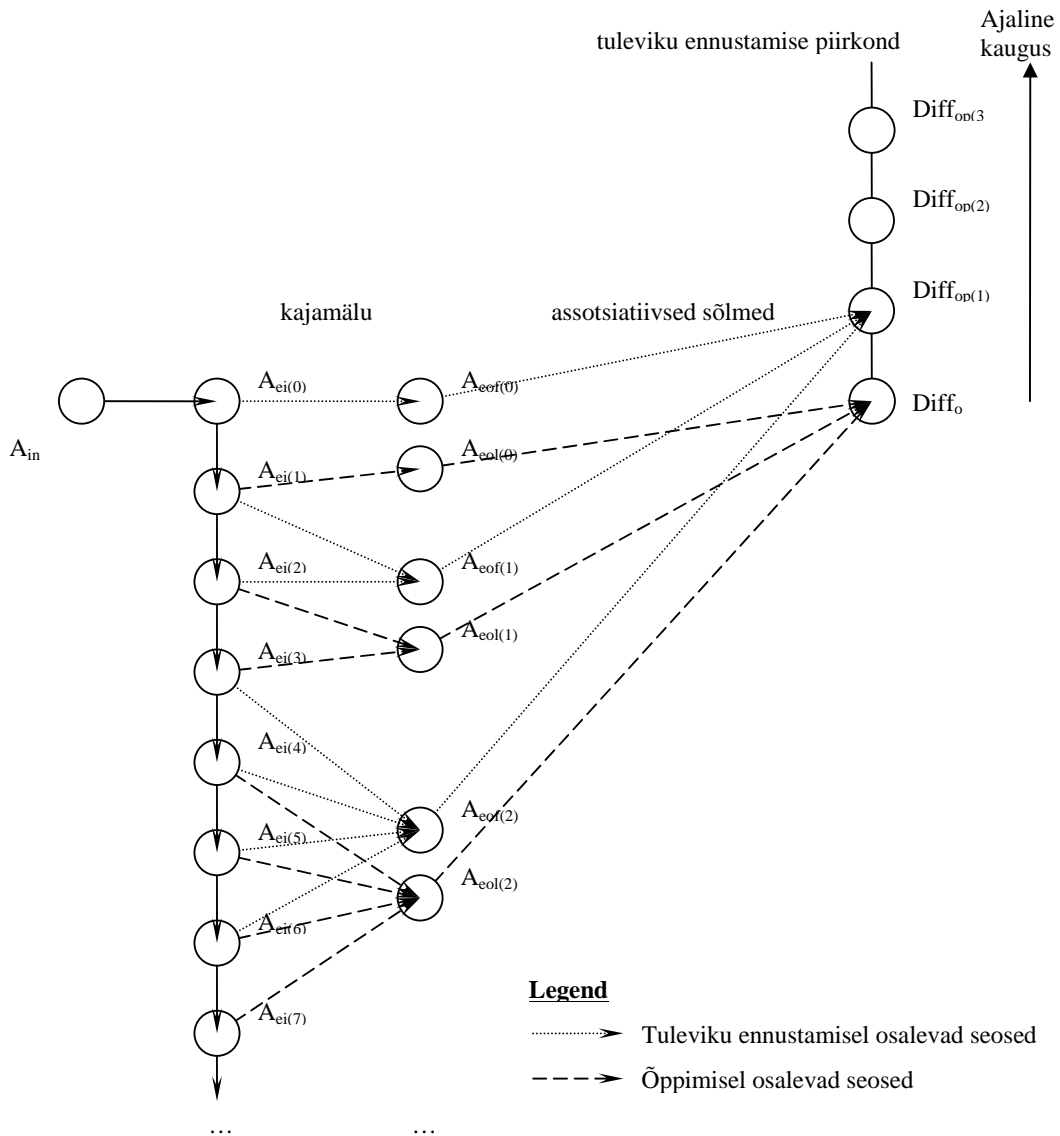
- a) Operantne mõtlemine üritab säilitada väärtust määratud tasakaalupunktis ka edaspidisteks ajavahemikeks, võimalikult väikese kõikumisega ümber selle. Kui ükski tegevus ei saavuta kohe soovitud seadistuspunkti väärtust, siis leitakse tegevus, mis selle pikemas plaanis saavutab, kuid ei ületa seda,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

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

## Mudeli õppimine ennustamiseks mitut tuleviku ajavahemikku

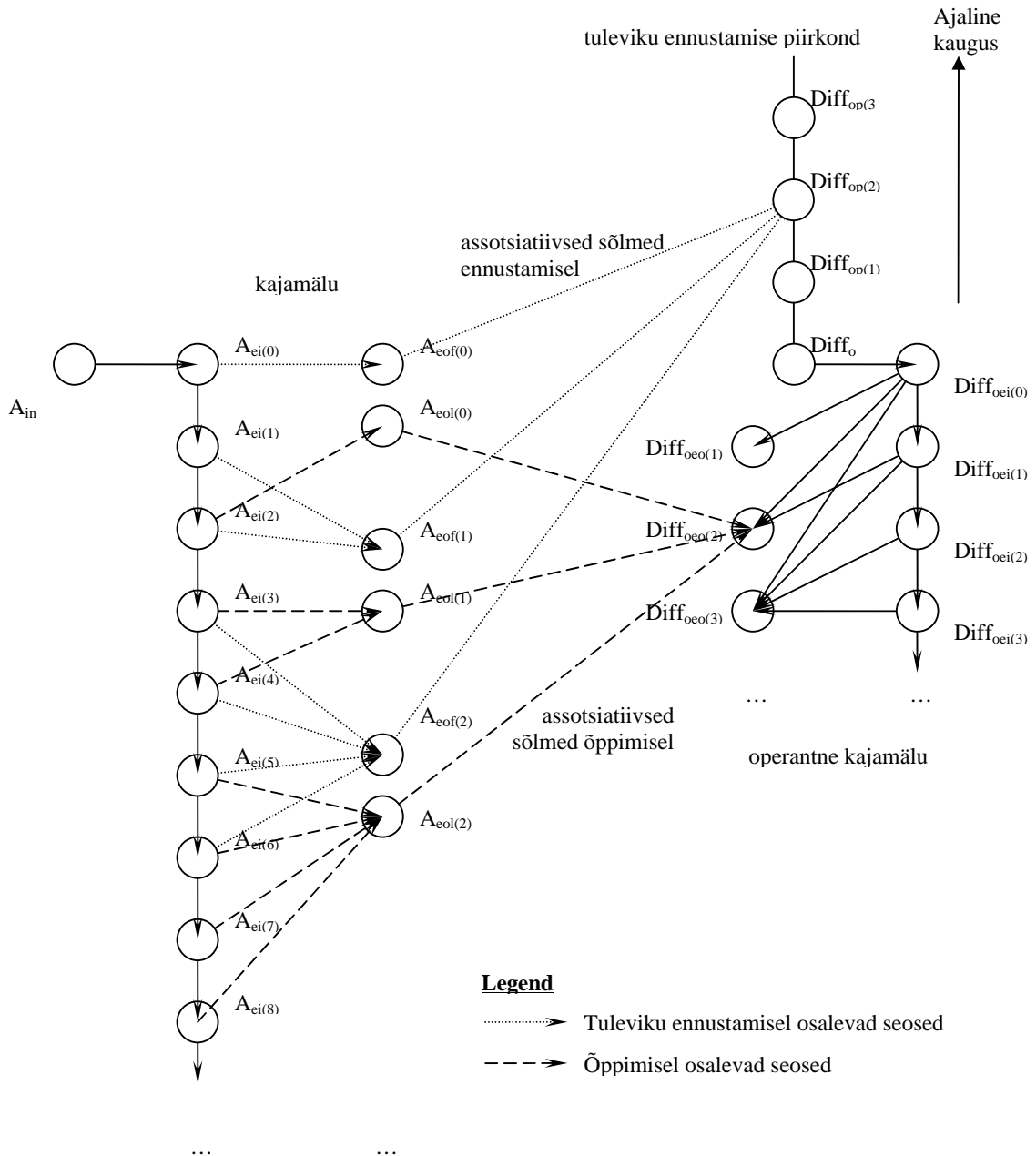
Järgnevalt toon detailsemate diagrammidega välja operantse kajamälu piirkonna ja tuleviku kujutlemise piirkonna suhted kajamälu ja assotsiatiivsete sõlmedega.

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



Iga õppimise ja ennustamise seosepaar läbib ühte ja sama sellele paarile vastavat assotsiatiivset sõlme. Kuid 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:



Joonis n.

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.

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

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

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

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

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

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.

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

$Diff_{oo(j)}$  – seadistuspunkti 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  $A$  kajade  $A_{eol(x)}$  ajalise suhte  $Diff_{oo(j)}$ -ga ning  $A_{eof(x)}$  ajalise suhte  $Diff_{op(j)}$ -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).

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  $k, i, j$ .

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

## **Vältimine ja põgenemine**

**TODO: peatüki paigutus artiklis.**

Võib ette kujutada kolme sorti vältimist:

- vältimine kui “tasakaalustamine”. Võib esineda enne ebameeldivat sündmust või pärast. Viimasel juhul nimetatakse põgenemiseks.
- vältimine kui “ära hoidmine”
- soovitud eesmärgi suunas tegutsemine ajapiirangu raames

Esimesed kaks tegevust võivad erineda: kui kõrvetan omal näpu vastu kuuma panni sanga, siis järgnevalt panen ta külma vette. See on sisult tasakaalustamine või traditsioonilise nime järgi “põgenemine”. Hoopis teistsuguse loomuga on tegevus, kui kasutan sanga haaramiseks pajalappi. Seda saab teha ainult enne sündmust ning peale sündmust ei muuda ta midagi olematuks.

Osa tasakaalustavaid tegevusi annavad tulemuse ka enne välditavat sündmust: sirutades käe välja enne kukkuma hakkamist või peale kukkuma hakkamist annavad mõlemad tasakaalustava tulemuse.

Praegune mudel suudab õppida ainult esimest sorti tegevusi. Teine vältimise liik nõuab täiendavat komplekti seoseid, et eristada tegevusi, mis aitavad nii enne kui pärast ebameeldivat sündmust, tegevustest, mis aitavad ainult enne. Neid seoseid ma praegu mudelisse ei lisanud ning see osa vajab edaspidist viimistlemist.

Kolmas vältimise liik: mingi soovitud tegevus on tulemuslik vaid siis, kui see viiakse läbi enne teatava ajaintervalli möödumist mingi stiimuli esinemisest. Naturaalses mõtlemises see esineb (Skinner, 1965). Sellise vältimise tarvis peaks olema mudelil võime mitte ainult mõista tegevuse mõju erinevatele tuleviku ajavahemikele, vaid ka planeerida tegevusi ja stiimulite väärtusi erinevatesse tulevikku hetkedesse ja seega võrrelda alternatiivsete tegevuskavade erinevaid tulemusi. See omadus ei eelda teist vältimise liiki, kuid eeldab mudelilt suuremat keerukust ja arvutusvõimsust planeerimise poolel ning jääb samuti edaspidiseks arenduseks. Täpsustuseks veel, et taolised tegevuskavad pole sugugi sama, mis aheldamist võimaldav protsess, mida kirjeldan allpool.

### ***Operantse tegevuse kustumine***

Operantne tegevus kustub aeglasemini kui omandatakse (Skinner, 1965).

Seda saab selgitada selle läbi, et piisab nõrgast seosest ehk tegevusele omistatud energiast, et tegevus käivitataks. Tegevustel on madal lävi, mis siiski on kõrgemal nullist, sest vastasel korral kustutatavad tegevused ei lõpeks kunagi muud moodi kui ainult karistamise või teiste tegevuste ülekaalu läbi.

Millest see lävi sõltub, jääb hetkel väljapoole mudelit. Selge on, et lävi on võrdlemisi madalal, ja oletatavasti stabiilne, kuna Skinneri järgi organism annab iga kinnituse kohta tagasi keskeltläbi kindla arvu operantseid tegevusi (Skinner, 1965), ning iga tegevus omakorda vähendab deprivatsiooni väikesel määral.

Mida suurem on deprivatsioon, seda sagedasem sagedasem on tegevus, seda pikem kustumiskõver ning rohkem tegevusi kustumise jooksul (Skinner, 1965). See on kooskõlas mudeliga, sest seadistuspunkti suurema energia korral on ka selle väikest edenemist ennustav tegevus suurema energiaga.

Kustumisega seondub ka küllastatuse teema: kui deprivatsioon kaob, siis tegevus seniks ajutiselt lõpeb. See aga ei tähenda, et kinnitamise järel tegevus alati kaob – kinnitus ei pruugi deprivatsiooni selles suhtes vähendada märkimisväärselt (Skinner, 1965).

Skinner kirjutab, et kustutatav tegevus hakkab esinema väiksema sagedusega (Skinner, 1957). Selle nähtuse esinemiseni saavad viia mitmed asjaolud ning hetkel ma ei ole neid kõiki analüüsinud ega mudeli operantset kustumiskõverat testinud.

## **Diskriminantsed stiimulid**

Diskriminantsed stiimulid võivad olla sensoorsed stiimulid või teatud eelnevad tegevused, mis on tarvis sooritada enne vaadeldavat tegevust. Viimastel võivad olla omakorda omad diskriminantsed stiimulid.

Juhul kui esineb deprivatsioon, on 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).

Lisaks võivad muutuda diskriminantseteks stiimuliteks varem esinenud stiimulite kombinatsioon või stiimulite ja tegevuste kombinatsioon esindavad rakud. See juhtub näiteks ratio-kinnituskava puhul (viide: Skinner). Taolised diskriminantsed stiimulid ei ole väliselt vaadeldavad.

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).

Diskriminantsed stiimulid jagan selguse mõttes diskriminantseteks tajudeks (s.h väliselt mittevaadeldavad kombinatsioon 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).

Lisaks eristan sensoorse diskriminantstiimuli erijuhtu – tegevuse sihtmärk-stiimuli.

Tegevuse kõrgemal kavandamisel mitte ei kästa jäsemete konkreetsetel lihastel liikuda konkreetset viisil, vaid kästakse sooritada tegevust teatud punktile ruumis (analoogselt Georgopoulos jt., 1993 pakutule, mainitud: Kolb & Whishaw, 2003. Veel viiteid:), mis võiks olla valitud kujutluspildiga sobivuse järgi. Seal ruumpunktis 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 tegevust esindavate kajamälude väljundid tajusid esindavate kajamälude väljunditega.

Tegevust esindava kajamälu väljundeid võib kirjeldada vektorina:

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

Analoogselt võib diskriminantstiimulit esindava kajamälu väljundeid kirjeldada vektorina:

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

kus

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

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

Analoogselt saab konstrueerida definitsioonid sümbolitele:

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

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

$S_{eol(i,j,k)}$  ja  $Sv_{eol(i,k)}$  – 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 teda tegevuse valikul arvesse võtta. Edaspidist uurimist vääriks, kas mingi stiimuli ootus järgnevasse ajahetkedesse võib olla naturaalses mõtlemises diskriminantstiimuliks käesoleva hetke tegevusele.

Seega hetkel piirang:

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

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, saatmaks 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.

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

tekkiva kaalude maatriksi saab konvertida esituse mugavuse mõttes varem kirjeldatud 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

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

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.

Käesolev mudel on robustne, luuakse kõik võimalikud assotsiatiivsed sõlmed vaadeldud maatriksisse. Edaspidise arenduse ja uuringute käsitleda oleks, millal on võimalik jätta võrdlemisi “ebahuvitavaid” faktoreid sisaldavad seosed loomata ja millise protsessi (naturaalses mõtlemises, kui see protsess esineb) läbi seoste maatriksit puhastatakse assotsiatiivsetest sõlmedest või nende alaosadest, mis ehk ei sisalda enam piisavalt olulisi faktoreid (seosed on kustunud). Sel juhul on andmestruktuurina maatriksi asemel parem kasutada *hashmap*’i või binaarotsingu puud kasutatavat assotsiatiivset massiivi.

Samuti tuleb uurida, kas ja kuidas võiks teatud tähelepanu protsess vähendada maatriksi läbiarvutamist tuleviku ajahetkede ennustamisel või ka õppimisel, jättes valikuliselt ajutiselt välja ülejäänud osa faktorite arvutamise või uuendamise.

Et mudel õpib faktoreid, mis vastavad *nonredundancy* nõudele, siis võiksid shansid selliste mehhanismide toimimiseks olla olemas: kui mingitele stiimulitele vastavad faktorid on head ja ennustavad tagajärgi piisavalt, siis see tähendab, et ülejäänud samal ajal esinenud stiimulid märkimisväärseid seoseid tagajärgedega ei omandagi. Samas näiteks Bayesi õppimise teel omandatud seostel selline selektiivsus puudub ja ebarealistlikult suured andmemahud kõigi asjade omavaheliseks seostamiseks on kerged tekkima, seega on ka tähelepanu protsessile suurem koormus (viide: Yudkowski, 2002).

Võib ette kujutada, et ka seadistuspunkti enda väärtus võib olla diskriminant. Näiteks kui seadistuspunkt esindab mingi liigese tegelikku ja soovitud nurka propriotseptiooni tajus, siis erinevate nurgade 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 annab sellisele omadusele seadistada.

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

## 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.

**TODO: selgitus, et aheldamise õppimisel toimuvad tegevused kohe õiges järjekorras.**

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 eelmisele alaelemendile. Pean terminit “kinnitav stiimul” mõnevõrra segadusttekitavaks: isend õpib, et mingi tegevus tekitab mingi stiimuli ka siis, kui see stiimul ise veel pole mingi operantse eesmärgiga seotud (eeltingimine, viidata?). Mingis taipamise ahelas osalevaid alaelemente võib olla õpitud ka grupikaupa, kuid erinevas järjekorras ning võibolla pikemate ajaliste pausidega.

On selge, et mõned autorid kasutavad sõna “aheldamine” ka sõna “taipamine” tähenduses. “Taipamise” terminit rõhutan, eristamaks mudeli võimet õppida tingitud tegevusjärgnevusi (aheldamist) nende järgnevuste või järgnevuse elementide ringiorganiseerimise ja erinevate järgnevuste ühendamise võimest. Taipamist võimaldav mehhanism võimaldab ka aheldamise, samas fikseeritud tegevusahelate tingimist oleks võimalik teostada mehhanismiga, mis ei võimaldaks kirjeldatud taipamist. Lihtne aheldamise mehhanism võimaldaks jääkade ahelate tingimist. Taipamist võimaldav mehhanism muudab kõigi teadaolevate ahelate lülid ringiühendatavateks. Uusi ahelad tekitades saab luua tegevusjärgnevusi, mida pole varem kinnitatud, luua uudseid seoseid õpitud lülide vahele.

Köhleri uuringute eesmärk oli uurida shimpanside võimet eesmärgi huvides vaimselt ringiorganiseerida maailma, mida ta nimetas “*insight*” (Tomasello, 1997). Shimpanside tajuvõimed on kahtlemata mitmekesisemad, võimaldavad rohkem seoseid ja manipulatsioone kui käesolevas mudelis, kuid kogemuste “kujutlemise” teel alternatiivseteks tegevusjärgnevusteks ringistruktureerimise põhiidee jääb minu jaoks hetkel samaks.

*Insight* ehk taipamine on Köhleri termin – vältimatult kognitiivne manipulatsioon; aheldamine Skinneri termin, rohkem biheivioristlik – lihtsal juhul, varem esinenud käitumisahela kordamine.

Piaget nimetas säärast protsessi “vigade ettenägemiseks” või “vaimseks katse-eksituse meetodiks”, kuna inividivid vaimselt proovib erinevaid stsenaariume läbi (Tomasello, 1997).

Praeguses mudelis “kujutlemine” pole midagi sellist, mida saaks visuaalselt näha ühe liikuva pildina või alternatiivsete teekondade ükshaaval läbimõtlemissena – selles osaleb palju kujutlusi korraga ning see pole mudeli jaoks probleem. Kui visuaalsed ja ainult ükshaaval läbimõeldavad kujutlused naturaalses mõtlemises esinevad, siis tulevad nad mudelisse hiljem ning võibolla näiteks ruumiliste rotatsioonidena jms, mille sooritamist ei pruugi olla alati võimalik paralleelselt arvutada, eriti kui kaasata protsessi tähelepanu ja märgiline kontroll.

Eelnevad 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 järgneb 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 ühekorruga.

Toomela kirjeldab taipamist 3. tasandil, kuid ta peab silmas ahvide probleemilahendusvõimeid. Ahvide taipamisele on iseloomulik võime seostada mitut objekti. Samas on olemas põhimõtteliselt analoogseid tegevusahelaid, mis eeldavad igal sammul vaid ühe diskriminantstiimuli ning tegevuse seostamist, ja mida võiks oletada enamuse imetajate võimeks (viide: vaata Tomasello, 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 selle võime loomine ratio-kinnituskava kui aheldamise erijuhi õppimise võime.

Erinevus III tasandist on, et jätkuvalt pole vaja lahendada probleeme, mis eeldavad mitme objekti või eraldiasuva tunnuste kombinatsiooni omavahel seostamist – on tarvis vaid stiimulite ja tegevuse (mis on erinevad modaalsused) õiget suhestamist. Samas on olemas sarnasus III tasandiga – lahendatakse probleeme, mille lahendus pole otse ühe tegevuse järel olemas ja teatud vajaliku tegevuse võimaldamiseks on enim, sõltuvalt olukorrast, vaja sooritada teatud teisi tegevusi.

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

Diskriminantseid stiimuleid esindavaid kanaleid saab tegevuse planeerimisel kasutada kahte moodi:

- a) Diskriminantsete stiimulite kajamälu sisu on see, mis 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.

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, igatühega vastavuses erinev energia. See väärtustega seotav energia on minimaalne ennustuste energia, 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 eesmärke arvesse võttes.

Kinnitavate stiimulite alaeesmärgina vajalike väärtuste poole püüdlemine on seega sarnasem *reinforcement learningule*, kus sisendi iga võimaliku väärtusega saab siduda mingi oleku ning anda olekule hinnangulise väärtuse. Siiski jääb erinevuseks distantide arvutamise võimalus.

Käesolev mudel arvestab olekute vahelisi distantse ning eeldab, et ühe sisendi olekud on võimalik paigutada sisendi väärtuse järgi ritta, liikumine nende olekute vahel on võimalik tehes lühemaid või pikemaid hüppeid üle vahepealsete olekute – soovitud olekuni/sisendi väärtuseni jõudmiseks tuleb teha tegevusi, mis nihutavad sisendi väärtust soovitud suunas, peatudes paratamatult teatud osadel 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 stsoonis, arvestades praegu võimalikke tegevusvariante ning olemata kogunud varem täpselt sama olekute vahelist järgnevust. Viimane oleks õppimisel probleem suurusega  $O(n^2)$ , – käesolevas mudelis piisab vaid olekute vaheliste distantide – erilist liiki suhete teadmisesest, mis on õppimisel probleem suurusega  $O(n)$ , kus  $n$  on olekute arv.

Vastavalt eelkirjeldatule, tegevused ennustavad muu hulgas kinnitavaid stiimuleid, seega tuleb leida viis mingi plaanitava tegevuse ja vastava ennustuse tulemusena kinnitavatele stiimulitele tekkivate väärtuste energia (soovitavuse) arvutamiseks.

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 energia on hetkel mudelis 0.

### **TODO sõnastus: seadistuspunkti tegeliku väärtuse ennustamine differentsiaali abil.**

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

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

$e_{d(i)}$  – oleku  $i$  energia

$A_v$  – tegevuste kujutletud määrade vektor

$Dv_i$  – diskriminantsete stiimulite kujutletud määrade vektor, sisaldab väärtust  $v_{d(i)}$

$E_p(A_v, Dv_i)$  – ennustatud seadistuspunktide ja kinnitavate stiimulite summaarne energia tegevuste vektori  $A_v$  ja diskriminantsete stiimulite vektori  $Dv_i$  korral

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

$E_{dp}$  – kinnitavale stiimulile ennustatud energia

$n_{dp}$  – kinnitava stiimuli olekute arv

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

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

$e_c$  – kinnitava stiimuli hetkeenergia, leitud mudeli eelmises töösükli

$v_c$  – kinnitava stiimuli väärtus käesoleval hetkel

$rs_i$  – *remaining step's relative size*, ennustatud väärtuse kaugus vaadeldavast olekust  $i$  suhestatuna käesoleva hetke väärtuse distantsiga vaadeldavast olekust

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

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

**TODO: näidisega joonis dünaamika kohta.**

Ülalolevas valemis  $\text{abs}()$  osa tagab, et kui ennustatud väärtus liigub üle 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).

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

**TODO: näite vältimise esinemise ja mitteesinemise kohta.**

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.

**TODO: valemi täiendus lokaalsete järeldustsaavutatavate miinimumide vältimiseks.**

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

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

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

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

kus

Dv – tegelike diskriminantsete stiimulite väärtused

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

Ülevaاتlikult: taipamine toimib sel teel, et mudeli probleemilahenduse esimese tööfaasi tulemusena “energiseeruvad” ehk saavad operantseks kinnitajaks need diskriminantsete stiimulid, mis võimaldavad omakorda eelmises töösüklis tekkinud eesmärgile lähemale jõuda ehk vähendada varem energiseerunud kinnitavate stiimulite ja seadistuspunktide energiat. Iga järgmise töösükliga energiseeruvad kinnitavad stiimulid, mis võimaldavad eelmises töösüklis 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 (viide: Skinner).

Muu hulgas võimaldab see teekonna planeerimist.

Tehisintellektist oleks sellise protsessi analoogiks “*breath-first search*”. Selline valik “*breath-first search*” vs “*depth-first search*” (Winston, 1984) vahelt on kõige lihtsam, kuna mudel suudab meelde jätta kõik ühele otsingupuule tasemele kuuluvad väärtused.

Otsingumeetodeid on veel, kuid ülejäänud on sisuliselt pigem eelnimetatud kahe edasiarendused ning puhta *breath-first search* rakendumise korral jäävad kõrvale.

Tegelikus närvisüsteemis võivad esineda mõlemad protsessid (ja edasiarendused) selles mõttes, et mõni “energiseerimise” haru levib kiiremini kui teine. Seda võiks uurida edaspidi lähemalt. Eriti oluliseks muutub otsingu meetodi valik protsessides, milles osaleb tähelepanu ning mis seega ei saa töötada paralleelarvutusena.

**TODO: HMM. Ei ole *breath-first search*. Ja seega ei pruugi selline taipamine olla ka optimaalne. Puu ülemis taseme parimad lahendused dikteerivad alloleva puu energiatega kujunemist. Kauguse arvutamine eesmärgist võib natuke aidata, aga mitte täielikult.**

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

Eriti mahukaks muutub protsess siis, kui igal tegevuse ja tagajärje seosel on mitu diskriminantset stiimulit (käesolevas mudelis on igal seosel ainult üks) – ühest küljest võimaldaks see tööriistade kasutamise (lähemalt selle kohta Lisas), 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 / diskriminantse stiimuli võimalike väärtuste arv

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

Nüüd saan selgitada ka töö algusosas olnud operantse tingimise tarvis minevate piirkondade ja kanalite joonisel olnud väidet, et osad piirkonnad pole primaatidel lõpuni arenenud: Tomasello, 1997 järgi ahvid ei ole võimelised õppima väliste sündmuste järgnevusi, mille sammudes nad ise aktiivselt ei osale. Sama probleem esineb mõnevõrra inimlastel (Frye et al, 1995, viidatud: Tomasello, 1997)

Mudeli seisukohast see tähendaks oletust, et ennustavad seosed  $S \rightarrow S$  vahel jäetakse vähemalt osadel juhtudel tekitamata, evolutsioonis arvatavasti (energia) 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 teisese 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 / 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 (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 lõigu 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 ühtlasi õpitavad kui diskriminantsed stiimulid tegevusahelas). Kuivõrd see oletus tõe vastab, vajab edaspidist kontrollimist.

Erilist liiki kinnitavateks stiimuliteks on üldistatud kinnitavad stiimulid, mis võimaldavad paljusid eesmärgi saavutada (näiteks 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 energiad liidetakse (vastandatuna näiteks maksimumi võtmisele).

**Kajamälu funktsioneerimine operantse kujutlemise tööfaasis:**

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.

**Mudeli töötsükkel**

Mudel töötab sünkroonselt. Töötsükkel koosneb järgnevatest faasidest:

**TODO**

## Operantsed kinnituskavad

Kinnituskavade tutvustust vaata lisast.

Kinnituskavad annavad üllatavalt sarnaseid tulemusi, eriti keerukate kinnituskavade korral, õige mitmesuguste liikide isendite juures: tuvid, hiired, rotid, koerad, kassid, ahvid (Skinner, 1957). Seega võiks oletada, et kinnituskavad kasutavad toimimiseks võrdlemisi lihtsaid ja primaarseid mehhanisme, nähtuste keerukus on nende protsesside kombineerumise ja dünaamilise vastasmõju tulemuseks.

Järgneva kohta lühike sissejuhatav kommentaar: *ratio* ning intervall-kinnituskavad toimivad täiesti erinevate protsesside toel.

## **Fixed-ratio**

Skinner kirjeldab fixed-ratio tegevuste ahelat märgistusega:

$$S_{reinf}^D \cdot R \rightarrow S_1^D \cdot R \rightarrow S_2^D \cdot \dots \cdot S_{N-1}^D \cdot R \rightarrow S_{Mag}^D$$

kus

$S_{reinf}^D$  – esimene kinnitusjärgne olukord, milles toimuv tegevus põhjustab uue

tegevusjärgnevuse esimese tegevuse järgse olukorra / diskriminantstiimuli ( $S_1^D$ )

R – tegevus

$S_x^D$  – igale olukorras  $S_{x-1}^D$  sooritatud tegevusele järgnevat olukorda esindav

diskriminantstiimul

$S_{Mag}^D$  – toidumagasini ilmumine, olukord, kus linnul on võimalik tulemuslikult teri nokkida

(viide)

Seega Skinneri idee järgne fixed-ratio kinnituskava on analoogne “teekonnale”. Iga teekonnal sooritatud sammu ja eelneva olukorra kombinatsiooni esindab erinev diskriminantstiimul, mis sellele sammule järgneb.

Käesolevas mudelis on stiimulite arv piiratud mudeli sisendite arvuga ning neid ei teki juurde. Skinneri “olukorrad” on samas diskriminantstiimulid, mis on pigem õppimise käigus (uue olukord-tegevus kombinatsiooni ilmumise järel) juurde tekitatud süsteemi siseselt ning ei vasta ühelegi tajutavale stiimulile.

See tähendab, et süsteemi siseselt tekib palju stiimuli esindusi, mis esindavad esinenud tegevuse-olukorra kombinatsioone ning ühtlasi on ise stiimuliteks, millega omakorda tegevusseoseid luua saab. Käesolevas töös on teostatud esimene neist omadustest, kuid mitte teist (esindused tekivad, kuid on kasutuses ainult seoses tajutavate tulemuste ennustamisega, kuid ise ei ole omakorda kasutatavad diskriminantstiimulitena). Sedasi nõuaks mudel rohkem arvutusvõimsust ning uusi “tegevus” – “tegevusele eelnenud olukord” kombinatsioone esindavate rakkude diskriminantstiimuliks saamise omaduse lisamist.

Küll aga on lisatud tööle katse “teekonna” läbimisega, kus igale sooritatud sammule vastab katseplaanis ette nähtud uue stiimuli ilmumine tajusse (viide katsetele). See on sarnane Skinneri katseplaanile, kus rakendati täiendavat välist stiimulit ratio või intervalli progresseerumise esindamiseks. See väline stiimul võib olla nii diskreetselt kui pidevalt kasvav (Skinner, 1957). Seda ka mudeli poolest: pidevalt muutuva välise stiimuli poolt kinnituskava progressi tajutav esindamine on veel üks omadus, mis nõuab stiimuli tugevuse klassifikaatori olemasolu. Kui väliseks stiimuliks on kell, siis võib klassifikaator asuda ka

nägemissüsteemis erineva kaldega asuvate joonte detektorite näol või toimida ka keerulisemate mehhanismide toel, nagu näiteks HTM, kuid see ei muuda siinkohal põhimõtet. Skinner, 1957 katsetest paistab veel välja, et välise kontrolliva stiimuli olemasolul omandatakse intervall-kinnituskava “puhtamalt” ja kiiremini.

Ratio-kinnituskava toimimine on ühtlasi näide, kuidas ülalpool kirjeldatud aheldamine toimiks tsüklilisena paistvates olukordades – tegelikult lahenduseni viiv olukordade-tegevuste graaf mudeli siseselt ei oleks tsükliline.

*Ratio*-kinnituskavadele on omane, et on teatud maksimaalne *ratio*, mida isend on võimeline õppima. Piiri ületamisel tegevus kaob märgatavalt ning tekivad pikad pausid. See piirang on võrdlemisi kõrgel. (Skinner, 1965). Hetkel pole selge, mis moodi mudel seda aspekti kajastada võiks. Mudeli aegluse tõttu saab katsetada ainult lühikesi tegevusahelaid. Mida pikem on *ratio*-kinnituskava ahela pikkus, seda pikemaks muutuvad pausid tegevuste vahel (Skinner, 1957). See nähtus saab tekkida mitmetel asjaoludel ning jääb samuti käesoleva mudeli katsetest välja ning edaspidiseks uurimiseks.

*Fixed-ratio* kinnituskavaga kinnitatud tegevus võib sisaldada pause, mida väiksem deprivatsioon, seda pikemad pausid. Kuid kui tegevus algab, siis toimub ta ühtlase tempoga kuni kinnituseni (Skinner, 1957). See ühildub oletusega, et esimese (ja iga järgneva) tegevuse käivitamiseks on tarvilik, et selle tegevuse energia ületaks teatud läve. Kui esimese tegevuse energia selle läve ületas, siis ületavad selle läve ka kõik järgnevad ahelas osalevad tegevused, sest nendelt järgnevatelt lüüditelt esimene tegevus oma energia saigi.

Huvitav oleks aga katsetada *variable-ratio* kinnituskava tingimustes, kus deprivatsioon napilt ületab ahela esimese tegevuse käivitamiseks tarvilikku läve. Kuna ahelas energiad summeeruvad, peaks mudeli praeguse kinnitavate stiimulite energiateg arvutamise valemi järgi olema võimalik tekitada olukord, kus ahela esimesed tegevused toimuvad, kuid seejärel ahela sooritamine peatatakse. Üks põhjus, miks küsimus on huvitav, on selles, et mul on mõttes ka alternatiivseid, mõnevõrra keerukamaid ja optimaalse kontrolli kriteeriumi järgi täiuslikumaid valemeid kinnitavate stiimulite energiateg arvutamiseks.

**TODO: selgitus-arvutus, miks fixed-ratio on raskemini kustutatav kui fixed-interval, kuid kergemini kui variable-ratio.**

## ***Fixed-interval***

Fikseeritud intervalliga kinnituskava on esindatav mudelis kombinatsioonina:

“stiimuli  $w$  kestuse määr  $x$ ” + “tegevuse  $y$  määr  $z$ ”  $\rightarrow$  “tagajärg”. Kus  $w$  oleks “kinnitava stiimuli” puudumist esindav tajukanal ning  $x$  väärtuseks selle tajukanali mistahes väärtus, mille ajal või millest suuremate väärtuste ajal tegevuse  $y$  sooritamine määraga  $z$  annab antud tagajärje. See eeldab, et mudeli stiimulite sisendites on tajusüsteemi poolt edastatud iga relevantse stiimuli olekut esindama erinev tajukanal ning tajukanalis oleva signaali tugevus kasvab vastavalt selle taju kestusele. Mudeli ülesanne on õppida tajukanali  $w$  sisendväärtuse  $x$  ning tegevuse  $y$  väärtuse  $z$  kombinatsioonile järgnevat tagajärge, selle järgi tegevust juhtida. *Fixed-interval* kinnituskava puhul kanal, mille kestust mõõdetakse, võiks olla näiteks “kinnitust ei ole”.

Intervall-kinnituskava juures tuleb arvesse veel üks asjaolu: kuigi isend õpib, et liiga varajane tegevus ei ole tulemuslik, võiks ta seda “igaks juhuks” ikkagi jätkata, sest minimaalsel määral jääb kinnituse ootus ikkagi alles, seni kuni varajast tegevust pole lausa karistatud.

Et varajased tegevused lõpuks “kustuksid”, peaks isendil olema lävi, millest madalama energiaga tegevusi ei sooritata. Tulemusena isend ei soorita tegevusi, mis ei ennusta enam piisavalt suurt kinnituse määra. Seda läve käsitlesin operantse tegevuse kustumise peatükis. Esialgu “kustuksid” kõige lühema intervalli järel toimuvad tegevused, seejärel pikema intervalli järel, ning lõpuks domineerima saanud mittetegevus pikeneks kuni umbes treenitava intervalli kestuseni (Skinner, 1957).

*Fixed-interval* kinnituskava treenimine tekitab protsessi kulgedes mitmesuguseid tegevusrütme, kuna organismi tegevus on kinnituse ajaks võrdlemisi sarnane *fixed-ratio* kinnituskavale, neil hetkedel saab osalist kinnitust hoopis see tegevusahel, mitte intervalliseos. Seetõttu peale kinnitamist tegevuse sagedus jälle kasvab, hiljem taas langeb (Skinner, 1965 ja Skinner, 1957). Siinkohal ei ole aga ruumipiirangu tõttu eesmärgiks kogu protsessi iseärasusi üles lugeda ega ka mudelil läbi proovida. Esialgu on oluline ainult see, kas mudel on võimeline omandama lõpuks stabiilset intervall-kinnituskavale omast tegevusrütmi. Stabiilset intervall-kinnituskava iseloomustab, et peale tegevust isend ootab ligikaudu kinnitatava intervallini ning natuke enne intervalli lõppu, kui ta ei suuda kestust enam eristada, järgneb *ratio*-kinnituskavale omane aheldatud tegevus kuni kinnituseni (Skinner, 1957).

**(viide katsetulemusele).**

Intervall-kinnituskava erijuht on kinnituskava, kus tegevus peab toimuma teatud aja jooksul eelmisest tegevusest (omamoodi segu *ratio*-kinnituskavast ja intervall-kinnituskavast).

Tulemuseks on väga kiire tegevus, kiirem kui *variable-ratio* korral (Skinner, 1965). Neid katseid pole praegu mudeliga läbi viidud, kuid vältimise protsess peaks seda nähtust selgitama ja stiimulite kestuste klassifikaator võimaldama.

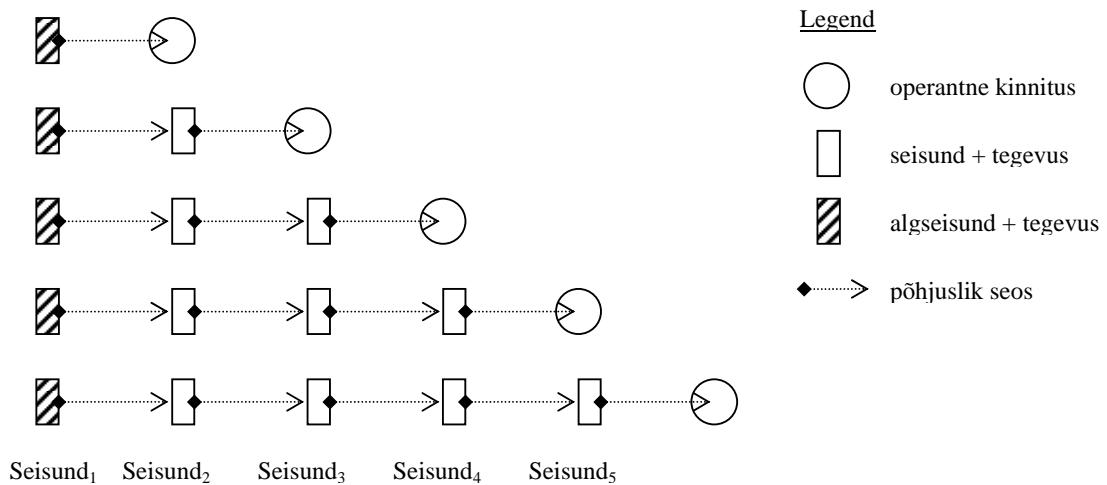
Suure intervalliga kinnituskava tavaliselt põhjustab mitmete algselt juhuslikult esinenud vahendavate tegevuste aheldamist (Skinner, 1957), mida võiks võtta märgina, et mudeli kestuse klassifitseerija poolt esindatavatel intervallidel tohib olla lagi ning pikema intervalliga kinnituskavad omandatakse juba aheldamise toel. See muidugi eeldab, et taoliseks aheldamiseks sobivad juhuslikud tegevused on igal vabamal momendil käivitumas.

### **Variable-ratio ning variable-interval kinnituskavad**

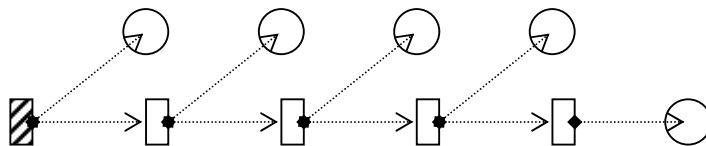
Kui tegevus annab tulemuse erinevates olukordades (erinev diskriminantne “olukord” ratio-kinnituskavas või erinev diskriminantne kestuse taju intervall-kinnituskavas), siis tekib vastavalt rohkem seoseid, mis ennustavad kinnitust, iga olukorra jaoks eraldi. Ükski neist ei pea ennustama kinnitust väga tugevalt, kuid seosed on olemas ning nende kõigi kustutamine on vastavalt aeganõudvam, arvestades, et kustumise kurvi alumises osas kustumise kiirus langeb.

### **Variable-ratio**

Põhjust, miks variable-ratio kinnituskavaga kinnitatud tegevused saavad suurema operantse energia, aitab kirjeldada järgnev joonis:



**Ehk:**



Ehkki *variable-ratio* kinnituskava võib pealevaadates jätta mulje kui ebakindlast kinnitusest, on tegelikult kinnituse saabumine seda kindlam, mida lähemale on tegevusahel jõudnud antud kinnituskava maksimaalsele *ratio* määrale. Seega esimesed ahela lülid saavad kinnitatud ligi  $1/n$  tõenäosusega, viimaste kinnituse tõenäosus läheneb  $1$ -le.

Ahela lülide energia summeerub esimese ahela lülini liikumisel. Ahela lülide endi vahelised seosed on alati täie tugevusega, kuna nad esindavad mudeli siseseid kontsepte, mis ei kao kuhugile (Seisund<sub>n+1</sub> tähendab: sooritasin just tegevuse olukorras Seisund<sub>n</sub>).

Kui kustumisel tegevuste ja kinnituste vahelised seosed nõrgenevad, siis energia vastavalt väheneb, kuid jääb siiski proportsionaalselt suuremaks, kui üksiku ja varemalt pidevalt kinnitatud seose energia peale kustumist oleks.

Sellise ahela energia enne kustutamist on esitatav valemiga <sup>4</sup>:

$$E = \sum_{k=1}^n \frac{1}{k}$$

seada nimetatakse  $n$ -s harmooniline number.

Kui  $n \rightarrow \infty$ , siis on tegu harmoonilise reaga ning

$$\lim_{n \rightarrow \infty} \sum_{k=1}^n \frac{1}{k} = \infty$$

Taolise tulemuse nägemine tekitab hüpoteesi, et mudelis peaks olema lisaks midagi, mis piirab energia kasvu väga suure maksimaalse *ratio* ehk  $n$  väärtuste korral, näiteks iga ahela tegevusele omistatakse ainult teatud osakaal järgmise kinnitava stiimuli ennustatud energia muudust. Sellise osakaalu olemasolu võiks olla ka selgituseks, miks piisavalt pika ahela korral *ratio*-kinnituskava tegevus lakkab. Hetkel seda täiendust mudelis teostatud pole ja ootab edasist uurimist.

Juhul, kui kustumise kõver on klassikalise tingimise valemite õige, lisanduks siia asjaolu, et ka iga  $1/n$  tõenäosusega kinnitatud sõlme seos on kaaluga rohkem kui  $1/n$  kinnitust, kuna kustumise kiirus kustumise kurvi alumises osas väheneb. Seega kõigi ahela lülide kinnitavate seoste mõju esimese tegevuse energiale on isegi suurem kui kõigi ahela lülide keskmiste kinnituste summa.

Vastavalt operantse planeerimise valemile, kui mingi seisund võimaldab tegevust, mis viib operantse kinnituseni, siis see seisund saab energia vastavalt oodatud tegevuse tulemuslikkusele. Seega on esimese tegevuse energia  $n \times w$ , kusjuures  $w > 1/n$ .

<sup>4</sup> [http://en.wikipedia.org/wiki/Harmonic\\_series\\_\(mathematics\)](http://en.wikipedia.org/wiki/Harmonic_series_(mathematics)) ja [http://en.wikipedia.org/wiki/Harmonic\\_number](http://en.wikipedia.org/wiki/Harmonic_number)

$$E(\text{Seisund}_5) = E_0 \cdot w_5$$

$$E(\text{Seisund}_4) = E_0 \cdot w_4 + E_0 \cdot w_5$$

$$E(\text{Seisund}_3) = E_0 \cdot w_3 + 2 \cdot (E_0 \cdot w_4 + E_0 \cdot w_5)$$

$$E(\text{Seisund}_2) = E_0 \cdot w_2 + 3 \cdot [E_0 \cdot w_3 + 2 \cdot (E_0 \cdot w_4 + E_0 \cdot w_5)]$$

jne.

üldistatult:

$$E(\text{Seisund}_i) = \sum_i^n E_0 \cdot w_i = E_0 \cdot \sum_i^n w_i$$

Variable ratio korral keskmiselt  $w = w_m = w_{m+1} = w_{m+2}$  jne, kus m on minimaalne kinnitatud ratio.

$$E(\text{Seisund}_i) = \sum_i^n E_0 \cdot w = E_0 \cdot \sum_i^n w = E_0 \cdot (n - i + 1) \cdot w$$

lihtsal juhul

$$w > 1/n$$

seega

$$E(\text{Seisund}_2) = E_0 \cdot (n - 1) / n$$

$$E(\text{Tegevus}_1) = E_0 \cdot w + E_0 \cdot (n - 1) / n = E_0$$

kus

$E(\text{Seisund}_i)$  – Seisundi i energia

$E_0$  – seadistuspunkti energia

$w_i$  – seisundis i sooritatava tegevuse ja operantse kinnituse seose tugevus (faktori väärtus).

NB. Seisundist seisundisse viivate seoste tugevus on 1, sest näiteks seisundis<sub>x</sub> sooritatud tegevus ongi seisundi<sub>x+1</sub> definitsioon.

n – seisundite koguarv (maksimaalne ratio)

Seisund<sub>1</sub> ei oma energiat, sest on juba saavutatud.

### ***Variable-interval***

Kui lugeda Skinneri (viide), siis on näha, et intervall-kinnituskava tulemusena tekib tegevus, mis natuke enne oodatava tasu hetke esineb korduvalt, kuni tasu ilmumiseni. Seega sisaldab intervall-kinnituskava enda komponendina ratio-kinnituskava. *Variable-interval* korral võib sinna omistada *variable-ratio* komponendi. Seega on *variable-interval* püsivam kui intervall-kinnituskava, kuid mitte nii püsiv kui seda on *variable-ratio*.

Võib küsida, miks *variable-interval* kinnituskava tegevuste intervall-komponent (ootamine) kiiremini ei kustu – vastus on analoogne *variable-ratio* kinnituskavale: iga üksik seos mingi intervalliga on nõrgem, kuid summaarselt on neid sellevõrra rohkem ja need seosed on oma kustumise kõvera alumises osas.

## Mudel ei käsitle

Kuigi mudelit saaks oletatavasti mugavalt edasi arendada primaatide tööriistakasutuse võimete vähemalt ühele osale (vaata Lisa), ei käsitle see järgnevaid primaatide mõtlemise aspekte:

- sotsiaalsed suhted
- momendil ka objektide vahelised ruumilised suhted (tertsiaarsed operatsioonid).

Cheney & Seyfarth, 1990, viidatud: Tomasello, 1997 hüpoteesi järgi objektide vaheliste suhete mõistmine arenes sotsiaalsete suhete mõistmisest. Evolutsiooniliselt võis järgnevus tõepoolest selline olla, kuid see veel ei täpsusta detaile, mis toimub mõtlemise mehhanismide sees.

Minu oletus on, et objektide vaheliste ruumiliste suhete mõistmine võib muu hulgas toetuda tähelepanule / pilgu sakaadidele – ühelt objektilt teisele liikuva pilgu sakaadi suund on ruumilise suhte kajastajaks.

Sotsiaalsed suhted on jäänud mudelist välja, kuna need on keerukamad, kui senise mudeli ülesehitamisel toeks olnud biheivioristlik lähenemine (mis ei väida, et sotsiaalseid suhteid ei saa siinsest komplekssema biheivioristliku lähenemise abil kirjeldada).

## Kokkuvõte

**TODO: Mida lisasin, eelised teiste mudelite ees.**

**Mida saab peale hakata sellega, rakendusi.**

Empiiriliselts saab mudeli realiseeringu toimimist edaspidi võrrelda naturaalsete süsteemide toimimisega erinevate ülesannete juures. Samuti saab võrrelda mudeli tõhusust teiste mudelitega.

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

## Viited

- Jüri Allik, Aavo Luuk. Nägemispsühholoogia.
- Talis Bachmann. Informatsioonitöötlus kognitiivses psühholoogias.
- Balkenius.
- Reinforcement Learning.
- Goal-Driven Learning:
- Dalbir Bindra. A Theory of Intelligent Behavior.
- Adam Eppendahl, Maarja Kruusmaa, Juri Gavshin. Avoiding irreversible actions
- Juri Gavshin, magistritöö
- Fagin, R., jt (2003). Reasoning About Knowledge. Cambridge, Massachusetts: MIT Press.**
- D’Inverno, M. & Luck, M. (2004). Understanding Agent Systems. Berlin, Heidelberg: Springer.
- Jeff Hawkins, Sandra Blakeslee. On Intelligence.
- Simon Haykin. Neural Networks – A Comprehensive Foundation.
- Wolfgang Köhler. Gestalt Psychology.
- Wolfgang Köhler. The Mentality of Apes.
- A. Leontjev. Inimese keel ja mõtlemine.
- Jaak Laineste. An extension of connectionist minds.
- Marcus, G. F. (2001). The algebraic mind: integrating connectionism and cognitive science. Cambridge, Massachusetts: Bradford/MIT Press.
- Miller, P. H. (1993). Theories of developmental psychology. New York: Freeman.**
- Russell, S. & Norvig, P. (2003). Artificial intelligence: a modern approach. Upper Saddle River: Prentice Hall.
- R. Rescorla, Wagner, ...
- Edasiarendused
- Skinner, B. F. (1965). Science and human behavior. New York: The Free Press.
- C. B. Ferster, B. F. Skinner. Schedules of Reinforcement.
- L. S. Võgotskij. Pedagogitseskaja Pshologia.
- Tomasello, M. & Call, J. (1997). Primate cognition. New York: Oxford University Press.
- Tomasello, M. (2000). The cultural origins of human cognition. Cambridge, London: Harvard University Press.
- Toomela doktoritöö.

- 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. (2003c). Otsustamine ebapiisava informatsiooni alusel: Süsteemne ja mittedüsteemne mõtlemine psühholoogias ja meditsiinis (HTEP.02.101). (2003 sügis, loengukursuse kiled). Tartu: Tartu Ülikooli Eripedagoogika osakond.
- 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.
- Weiss, G. (Ed.). (2000). Multiagent Systems. Cambridge, Massachusetts: MIT Press.
- Winston, P. H. (1984). Artificial intelligence. Reading, Menlo Park: Addison-Wesley Publishing Company.
- 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.
- Woodbridge, M. (2000). Intelligent Agents. G. Weiss (Ed.), Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence: 27-66. Cambridge, Massachusetts: MIT Press.
- Eliezer S. Yudkowsy. Levels of Organization in General Intelligence.
- ...
- ...