

# Representaatiot ja vahvistusoppimisalgoritmit

*Anna-Mari Rusanen, Otto Lappi,  
Jami Pekkanen & Jesse Kuokkanen*

”Mitä seuraavaksi kannattaisi tehdä?”, kysyy jokainen toimija jatkuvasti. Vastaus vaatii toimijalta monimutkaista tiedonkäsittelyä. Toiminnassa ei ole kyse vain teoista, liikkeistä tai liikeradoista vaan ennakoinnista, koordinoinnista ja valikoinnista. Tämän hienostuneen tiedonkäsittelyn ymmärtäminen on yhä tärkeämpää tekoäly- ja robotiikan tutkimuksessa, joissa toiminnanohjausta lähestytään algoritmisesti.

Vahvistusoppimisalgoritmit ovat nykyisin hyvin suosittu tutkimuksessa käytettävien algoritmien ryhmä. Vahvistusoppimisalgoritmit etsivät mahdollisimman hyviä lopputuloksia tuottavia toimintakäytäntöjä (engl. *action policy*). Niitä on sovellettu esimerkiksi robottikäsiin liikkeiden suunnitteluun, autonominen ajoneuvojen reittien valintaan ja pörssi-markkinoiden ennustamiseen. Kuuluisia vahvistusoppimisalgoritmeja ovat myös erilaiset pelisovellukset, kuten go- ja shakkipelisovellukset.

Vahvistusoppimiseen perustuva tekoäly- ja robotiikan tutkimus haastaa myös filosofien näkemyksiä toiminnanohjauksesta. Viime vuosina paljon huomiota ovat herättäneet erityisesti radikaalien enaktivistien, kuten Daniel Hutton ja Erik Myinin, väitteet, jotka näyttävät uuden tutkimuksen valossa vanhahtavina, epäuskottavina ja paikoin virheellisinä.

## Hutto, Myin ja toiminnanohjaus

Radikaalia enaktivismia edustavat filosofit Hutto ja Myin ovat viime vuosina herättäneet paljon huomiota poleemisella kirjoitustyyllillään. Uskollisine seuraajineen he ovat vaatineet räis-kyvällä tyyllillä kognitiontutkimukseen ”vallankumousta”, sen ”radikalisoimista” ja ”dogmaattisesta perustasta luopumista”. Heidän mukaansa kognitiontutkimuksessa ei ymmärretä toimin- nan keskeisyyttä tiedonkäsittelyssä. Sen sijaan tutkimuk- sessa sitoudutaan vanhentuneisiin uskonkappaleisiin. Hutton ja Myinin mukaan kognitiontutkijoiden pitäisi irrottautua näistä aikansa eläneistä tutkimuksen ”pilareista” ja uudistaa tutkimus- kenttensä teoreettinen perusta.<sup>1</sup>

Hutton ja Myinin mukaan pilareita on kaksi, representatio- nalismi ja komputationalismi. Representationalismi tarkoittaa oletusta, että tiedonkäsittelyä ei voida selittää ilman esittäviä mielen tiloja, representaatiota. Komputationalismi puolestaan tarkoittaa, että tiedonkäsittelyjärjestelmän toiminta selitetään laskennallisesti.

Hutton ja Myinin mukaan molemmista tulisi luopua. Tilalle he ehdottavat uutta viitekehystä, enaktivismia. Sen pääajatus on, että toiminta ja toimijuus voidaan selittää eräänlaisena kytkey- tymisenä: toimijan kyvyllä toimia ympäristössään tarkoituksen- mukaisesti, ympäristöön reagoimalla.

Hutto ja Myin vastustavat erityisellä vimmallalla ajatusta, että tällainen toiminnanohjaus tapahtuisi ulkoista todellisuutta esit- tävien tilojen, representaatioiden, avulla. Päinvastoin, heille toiminta on ympäristöön suuntautunutta reaktiivisuutta. Esi- merkkinä tästä he pitävät liikkeen ja motorisen toiminnan ohjausta. Myin ja Hutto kirjoittavat muun muassa, että ”[...] havaintokognitiivinen toiminta – tuulessa pyörähtelevän lehden jahtaaminen ja nappaaminen – kohdistuu objekteihin ja asiantio- loihin ilman, että se esittää niitä”<sup>2</sup>

Algoritmisen tutkimuksen näkökulmasta ”tuulessa pyörähte- levän lehden nappaaminen” on kuitenkin esimerkki pikemminkin

siitä, kuinka toiminnanohjaus tapahtuu esittävien mallien avulla. Liikkuvan lehden kiinni saaminen edellyttää toimijalta monimutkaista, saumatonta ja samanaikaista eri aistien, kehon koordinaation ja liikeratojen suunnittelun yhteistyötä. Näin hienostunut toiminnanohjaus ei tapahdu ”reaktiivisesti”, vaan tiedonkäsittelyjärjestelmän täytyy esimerkiksi ennakoida toiminnan päämääriä.

## Toiminnanohjaus ja vahvistusoppimisalgoritmit

Vahvistusoppimisalgoritmien tausta on behaviorismissa, joka pyrki selittämään mentaalaisia ilmiöitä ulkoisesti havaittavan käytöksen avulla. Viime vuosisadalla behavioristit huomasivat, että esimerkiksi eläimiä voidaan opettaa vahvistamalla tietyn ärsykkeen ja halutun reaktion välistä yhteyttä. 1900-luvun alun koirakokeistaan tunnettu Ivan Pavlov työtovereineen nimesi tämän oppimisen muodon klassiseksi ehdollistumiseksi, ja myöhemmin yhdysvaltalainen psykologi B. F. Skinner kehitti sen pohjalta mallin välineellisestä ehdollistumisesta.<sup>3</sup>

Varsinaisesti itse algoritmit kehitettiin 1970-luvulla, kun keksittiin tapa, jolla voitiin laskennallisesti kuvata mekanismi, jolla vahvistaminen tapahtuu. Se johti vahvistusoppimisalgoritmien kehittämiseen 1980-luvun taitteessa.

Vahvistusoppimisen ja ylipäänsä behaviorismin alkupe räinäinen motivaatio oli poistaa kaikki mielen sisältöihin viittaavat käsitteet toiminnan ja oppimisen kuvauksesta. Pitkään ehdollistumista tarkasteltiin näitä käsitteitä korostavan kognition tutkimuksen vastakohtana. Nykyisin vahvistusoppiminen nähdään yhtenä monista oppimisen muodoista, ja sille voidaan antaa myös kognitivistinen tulkinta.<sup>4</sup>

Vahvistusoppimisen tavoite on, että toimija oppii saavuttamaan toiminnallaan mahdollisimman suuren hyödyn eli maksimoimaan saamiensa palkkioiden määrän. Toimija saa jokaisesta toiminnosta tai teosta palkkion (engl. *reward*) tai rangaistuksen

(engl. *punishment*). Arvo (engl. *value*) tarkoittaa pitkän aikavälin arviota palkkion kokonaismäärästä. Oppiessaan toimija – algoritmi, kissa tai ihminen – päivittää toimintakäytäntöään joko heikentämällä tai vahvistamalla sitä. Behavioristien nimeämän teon seurauksen lain (engl. *the law of effect*) mukaan positiivinen palkkio vahvistaa palkkioon johtanutta toimintakäytäntöä. Jos toiminnasta rangaistaan, se puolestaan heikentää toimintakäytäntöä.

Käytännössä vahvistusoppiminen tapahtuu seuraavasti. Kuvitellaan, että kissaa opetetaan hyppäämään laatikkoon. Jos kissa hyppää laatikkoon, sitä palkitaan herkulla. Jos kissa kävelee pois, sitä ei palkita. Kissan näkökulmasta oppimisen tavoitteena on saada mahdollisimman paljon herkkuja, ei niinkään oppia hyppäämään laatikkoon. Kun kissa oppii, miten se saa herkkuja, se alkaa suosia käytäntöä, jonka avulla se saa mahdollisimman paljon palkkioita. Näin se oppii, että laatikkoon hyppääminen on hyvä tapa saada mahdollisimman paljon herkkupaloja eli maksimoida palkkio.

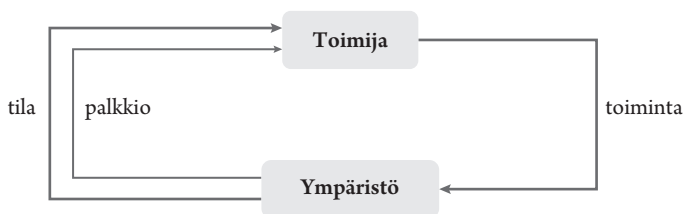
Eläimen tai ihmisen tapauksessa palkkio on koeolosuhteissa yleensä jokin herkku tai muu ulkoinen asia. Aivotutkijat olettavat, että nisäkkäiden – kuten ihmisen tai kissan – aivoissa palkkioita käsittelee neurofysiologinen, mielihyvään liittyvä mekanismi. Palkkio voi kuitenkin syntyä myös puhtaasti organismin sisäisten prosessien tuloksena. Se, mikä on palkitsevaa, ei siis palaudu johonkin ulkoiseen tekijään.

Palkkio ja arvo ovat organismiriippuvaisia eli niiden määrä riippuu organismin tilasta. Eri organismeilla on omat palkkiokriteerinsä. Se, mikä on palkitsevaa esimerkiksi kissalle, ei välttämättä ole palkitsevaa ihmiselle. Eikä se, mikä on yhdelle ihmiselle palkitsevaa, ole välttämättä kaikille ihmiselle yhtä palkitsevaa. Jos ihminen on nälkiintynyt, suklaa voi olla taivaallista. Jos ihminen on kylläinen, suklaa palkitsee vähemmän.

Kun siirrytään eläimistä tai ihmisistä tekoälyyn ja robotteihin, muuttuvat palkkiot neurofysiologisista mekanismeista numeroiksi ja ohjelmistokodeiksi. Algoritmin tasolla palkkio

on numeerinen signaali – matemaattinen suure. Robotille tai ohjelmistolle palkkio siis annetaan puhtaasti numerona, joka vahvistaa todennäköisyyttä toimia myöhemmin tavalla, jolla saa lisää palkkioita. Vastaavasti rangaistukset ovat negatiivisia palkkioita, eli ne heikentävät todennäköisyyttä.

Myös muut algoritmin kuvaukseen käytettävät käsitteet ovat matemaattisia. Niille voidaan antaa tulkinta teknisten käsitteiden avulla. Vahvistusoppimisalgoritmeja on useita, mutta perusmuotoinen algoritmi kuvataan usein seuraavan kaavion avulla:



*Algoritmi sisältää agentin eli toimijan, toiminnan, ympäristön, ympäristön tilasiirtymät sekä palkkion.*

Algoritmin käsitteet – toimija, ympäristö, palkkio, toiminta ja tila – eivät kuitenkaan vastaa arkikielen käsitteitä. Esimerkiksi ympäristö ei viittaa ulkoiseen, konkreettiseen ympäristöön. Ympäristö on määritelmä algoritmille annetusta ”ympäristöstä”. Se on osa algoritmia – ei sen ulkopuolella. Tietojenkäsittelytieteilijät puhuvatkin keinotekoisesta, ”synteettisestä” ympäristöstä. Synteettistä ympäristöä voidaan kuvata matemaattisena mallina, joka muodostuu joukosta ajallisesti toisiaan seuraavia maailman tiloja.

Vahvistusoppimisalgoritmeissa – monien muiden toiminnanohjauksen laskennallisten mallien tapaan – toiminta, kuten omenan poimiminen, puretaan algoritmin suorittamiksi laskennallisiksi askeliksi. Algoritmin avulla esimerkiksi robottikäsi voidaan ohjelmoida liikkumaan tietyllä tavalla. Tässä algoritmi voi

käyttää apuna mallia, jota se päivittää asteittain. Usein päivittäminen on helpointa havaintojen eli jonkinlaisen palautteen avulla.

Vahvistusoppimisen erikoisuus on, että malleja ei päivitetä vain havaintojen pohjalta, vaan toimija arvioi toimintoja myös tulevaisuuden mahdollisten palkkioiden näkökulmasta. Vahvistusoppimiseen perustuvat algoritmit pystyvät siten ennakoimaan tilanteita, ennen kuin ne varsinaisesti tapahtuvat.<sup>5</sup> Algoritmit arvioivat, mitä ne hyötyvät tietyistä toiminnasta. Ne eivät siis laske vain sitä, mitä ne seuraavaksi tekevät. Tästä syystä ne ovat kognitiivisesti kehittyneempiä kuin monet aiemmat laskekennalliset toiminnanohjausmallit, jotka laskevat vain seuraavia toimintoja eivätkä toimintojen seurauksia.

Vastaavasti vahvistusoppimisalgoritmit kykenevät tiettyyn pisteeseen asti reagoimaan myös tilanteeseen, jossa täytyy päivittää ja osittain vaihtaa toimintastrategiaa. Jos kissan pitääkin oppia hyppäämään laatikon sijasta tuolille, sen täytyy muuttaa strategiaa niin, että se soveltuu uuteen tilanteeseen. Tai, jos robottikäden ympäristö muuttuu, ja omenan paikka vaihtuu, robotin täytyy löytää uusi toimintastrategia. Eräissä vahvistusoppimisalgoritmien sovelluksissa on pystytty toteuttamaan tällaista toiminnansuunnittelua erityisellä suunnittelujärjestelmällä, jossa tulevaisuutta koskevan ennusteen avulla kyetään päivittämään toimintasuunnitelmaa askel askeleelta.<sup>6</sup> Tähän järjestelmä käyttää niin sanottua tulevaisuuden toimintaympäristöä koskevaa mallia. Sen avulla toimija voi arvioida toiminnan vaikutuksia. Malli koostuu muun muassa niin sanotuista tilasiirtymäsäännöistä sekä tulevia palkkioita koskevista ennusteista.<sup>7</sup>

## Vahvistusoppimisalgoritmit ja representaatiot

Vahvistusoppimisalgoritmien erikoisuus ja vahvuus on, että ne kykenevät laatimaan tulevaisuutta koskevia ennusteita arvioimalla saatavia palkkioita. Ne tarjoavat siten esimerkin kahdesta

erilaisesta representaatioiden ryhmästä, mallirepresentaatioista ja arvorepresentaatioista. Näistä kumpikaan ei varsinaisesti esitä ulkoista todellisuutta vaan sellaisia toiminnan päämääriä, joita ei ole konkreettisesti mielessä edes olemassa. Ne suuntautuvat siis tulevaisuuteen eivätkä ympäröivään todellisuuteen.

### *Mallirepresentaatiot*

Suunnittelujärjestelmien sisältämät mallit voidaan tulkita niin, että ne esittävät järjestelmän tulevaisuuden tiloja. Toiminnanohjausjärjestelmä muodostaa mallin, joka kertoo, kuinka algoritmin ympäristö – sille annettu kuvaus ympäristöstä – kehittyy tulevaisuudessa. Sen avulla järjestelmä arvioi, mistä toiminnosta se saa mahdollisimman suuren palkkion. Nämä mallit eivät viittaa konkreettiseen ympäristöön tai sen osiin. Ne esittävät algoritmille annettua ympäristöä ja sen mahdollista tulevaisuutta. Representaatioiden kohdetta ei siis ole olemassa, eikä se voi aiheuttaa representaatiota tai sen viriämistä.

Usein toiminnalle on kuitenkin keskeistä, että sovellus – vaikkapa robottikäsi – kykenee ohjaamaan toimintaa ulkomaailmassa. Jos robottikäden tehtävä on esimerkiksi poimia oikea omena oikeasta puusta, on sovelluksen löydettävä omena, osattava arvioida käden etäisyys omenaan sekä laskettava oikea tarttumisote. Ulkoinen toimintaympäristö asettaa siten toiminnan onnistumiselle reunaehdoja, jotka sovelluksen on otettava huomioon. Ulkoisen ympäristön ja toiminnanohjausjärjestelmän välillä täytyy siis olla jokin vastaavuus, jonka vuoksi sisäisen mallin avulla tapahtuva toiminta onnistuu myös ulkoisessa ympäristössä.

Filosofien intuitio on usein, että näissä tilanteissa toiminnanohjausjärjestelmän sisältämän ympäristön kuvauksen – tai algoritmista muodostuvan mallin – tulisi tavalla tai toisella kuitenkin olla yhteydessä maailmaan. Usein ajatus on, että ulkoisen todellisuuden täytyy aiheuttaa mallin todellisuus. Jos mallilla ja maailmalla ei ole yhteyttä, on täysin sattumanvaraista, mil-

loin algoritmi tulkitsee reunaehdot oikein ja milloin ei. Se tekee onnistumisesta siis ihmeen.

Tämän niin sanotun ihmeargumentin ytimessä on ajatus, että mallin onnistunut toiminta ei voi perustua ihmeille. Sen sijaan onnistunut toiminta osoittaa, kuinka mallien täytyy olla systemaattisessa yhteydessä maailmaan ja tavalla tai toisella perustua maailmasta saatavalle informaatiolle. Tästä seuraa, monet argumentoivat, että onnistunut toiminta oikeastaan osoittaa, miksi mallit ovat maailmaa esittäviä tiloja.

Robottiikan ja tekoälyn kohdalla ihmeargumentti ei kuitenkaan ota huomioon kahta asiaa. Ensiksi, nykyiset robotit – ja myös ihmisaivot – ovat monimutkaisia järjestelmiä, joissa on useita eri tehtäviä toteuttavia alijärjestelmiä. Osa niistä toimii havaintojärjestelminä. Ne tekevät ennusteita ympäröivästä todellisuudesta ja vastaanottavat myös sitä koskevaa informaatiota. Toiset taas ennustavat, miten toimintaa kannattaa ohjata. Kolmannet vastaavat muistitiedon tallentumisesta, neljännet kommunikaatiosta ja niin edelleen. Olennaista on, että myös muut järjestelmät kuin toiminnanohjausjärjestelmät käsittelevät tietoa. Siten toiminnanohjausjärjestelmät voivat saada joltakin toiselta järjestelmältä syötettä, joka puolestaan liittyy suoremmin ympäristöä koskeviin havaintoihin. Toiminnanohjausjärjestelmien itsensä ei siis tarvitse välttämättä havainnoida konkreettista ympäristöä.

Toiseksi, ihmeargumentti ei ota huomioon, että mallien ja ulkoisen todellisuuden välillä voi olla useita erilaisia syy- ja seuraussuhteita. Niistä kaikki eivät suinkaan ole esittäviä eli representaatiosuhteita.<sup>8</sup> Siten esimerkiksi palautemekanismien avulla tapahtuva muuttujien päivitys voi kyllä päivittää järjestelmää ympäristön informaation avulla, mutta päivitys itsessään ei ole representoinnin eli esittämisen muoto.

Käytännössä juuri näin ongelma usein ratkaistaan nykyrobotiikassa. Muuttujien päivittämiseen on tarjolla erilaisia teknisiä ratkaisuja. Tyypillisesti on jokin mekanismi, jonka avulla ulkomaailmasta saatava palaute päivittää toiminnanohjauksesta



vastaavan algoritmin muuttujia. Keskeistä on, että palaute ulkomaailmasta vaikuttaa lähinnä syy- ja seuraussuhteisiin, eikä sillä ole pääsyä algoritmin itsensä sisälle tai sen sisäisiin suhteisiin.

### *Arvorepresentaatiot*

Toinen representaatioiden ryhmä muodostuu mallien sisältämistä pitkän aikavälin palkkioiden arvioista eli arvoista. Nämä ”arvorepresentaatiot” esittävät matemaattisesti yksilöityä arviota palkkioiden kokonaismäärästä, eivät konkreettista palkkiota. Vaikka arvorepresentaatioiden kohteena oleva ”palkkio” operationalisoidaan eläinkokeissa usein herkkuna, representaatiot itsessään eivät silti viittaa niihin tai niiden fysikaalis-kemiallisiin piirteisiin (esim. herkun sisältämä glukoosi). Representaatiot eivät esitä herkkua vaan matemaattisesti yksilöityä arviota tulevaisuuden palkkiosta.

Operationalisoitu palkkio, esimerkiksi herkkuei *aiheuta* representaatiota tai sen sisältöä. Ei myöskään ole olemassa palkkiodetektoria, joka tunnistaisi palkkion ja muodostaisi siitä representaation. Palkkio ei yksinkertaisesti toimi signaalina, joka johtaa palkkiorepresentaation viriämiseen. Päinvastoin palkkiorepresentaation olemassaolo on edellytys sille, että ulkoisen todellisuuden objekti voi ylipäänsä toimia palkkiona.

Arvorepresentaatiot ovat organismiriippuvaisia. Odotettavissa olevan palkkion arvioitu määrä riippuu nimittäin osittain organismin tilasta. Esimerkiksi se, kuinka palkitseva makea herkkuei on, riippuu yksilön metaboliasta ja tilanteesta. Kuten algoritmin kuvauksen yhteydessä todettiin, sama herkkuei ei ole yhtä palkitseva kaikille tai kaikissa olosuhteissa.

Näiden organismi- tai tilanneriippuvaisten ominaisuuksien luonne on hyvin hankalasti käsitteellistettävissä. Ne muistuttavat joissain suhteissa Locken sekundaareja kvaliteetteja eli havaittaviin riippuvaisiin ominaisuuksiin, mutta niiden matemaattinen kuvaustapa vaikeuttaa niiden käsitteellistä jäsenystä filosofialle tuttuun viitekehysten sisällä.

Kognition- tai tekoälytutkijoilla ei myöskään ole tarjota näistä representaatioista sellaista kuvausta, jolla filosofit voisivat työskennellä. Tutkijat eivät yksinkertaisesti tiedä vielä riittävästi näistä järjestelmistä tai niiden sisältämistä representaatioista. Vaikka niiden matemaattiset ominaisuudet tunnetaan, teoreettista ymmärrystä niistä nimenomaan representaatioina ei vielä ole. Paremmin tiedetäänkin, mitä ne eivät ole: ne eivät ole havaintojen kaltaisia representaatioita.

Monissa filosofisissa teorioissa käsitteistö on algoritmisesta näkökulmasta usein arkikäsitteistöä, eikä se aina sovellukaan kovin hyvin algoritmisen tutkimuksen kuvaamiseen. Toiminnanohjaus on tästä yksi esimerkki. Se on tutkimusalue, josta filosofeilla on vahvoja intuitioita mutta ei sellaista käsitteistöä, jolla ilmiöitä voitaisiin hedelmällisesti tarkastella. Toiminnanohjaus on tässä suhteessa esimerkki myös siitä, kuinka algoritmit muuttavat tieteellistä tutkimusta ja kuinka se heijastuu myös tieteenaloja tarkasteleviin filosofian osa-alueisiin. Esimerkiksi toiminnanohjausrepresentaatioista ei ole vielä kokonaiskuva, ja jää nähtäväksi, millaisista osista se lopulta rakentuu.

## Hutto, Myin ja kärpäsdetektorit

Hutto ja Myin, samoin kuin monet muut filosofit, eivät ota huomioon, että toiminnanohjauksessa voitaisiin olettaa toiminnan päämääriä kuvaavia representaatioita. He yksinkertaisesti sivuuttavat kysymyksen, koska he ovat usein huomaamattaan sitoutuneet niin sanottuun kärpäsdetektorina tunnettuun representaatiokäsitykseen.

Kärpäsdetektorimallissa representaatiot määritellään mielen tai kognition tiloiksi, jotka esittävät ulkoista ympäristöä. Tyypiesimerkki tällaisesta representaatiosta olisi vaikkapa havaintojärjestelmän tuottama representaatio kissasta, kun katselet laatikkoon hyppäävää kissaa. Kun näet hyppäävän kissan,

kissa sekä osittain aiheuttaa havaintosi että toimii myös kohteena, jota aivoissasi syntyvä havaintotila esittää.

Kärpäsdetektorimallin juuret ovat 1950- ja 1960-lukujen havaintopsykologiassa, jossa tutkittiin erityisesti näköjärjestelmien reseptiivisiä kenttiä. Tutkijoiden tavoitteena oli selvittää, kuinka näköjärjestelmä valikoi informaatiota. Kokeissa tutkittiin esimerkiksi, kuinka sammakon silmien solut reagoivat pieniin, mustiin, liikkuviin ja kärpäsen kaltaisiin täpliin eli kuinka ne jäljittävät (engl. *detect*) kärpästen piirteitä. Sammakoiden havaintojärjestelmää avanneessa artikkelissa Jerome Lettvin kumppaneineen ehdotti, että piirteitä jäljittävät ”detektorit” eli niihin erikoistuneet näköaivokuoren järjestelmät sisältävät mekanismeja, joilla näköaivokuori muuttaa piirteitä koskevat yksinkertaiset havainnot monimutkaisiksi representaatioiksi.<sup>9</sup>

Kärpäsdetektorit levisivät nopeasti havaintopsykologiasta filosofiaan. Ne hallitsivat mielikuvia representaatioista useiden vuosikymmenien ajan. 1980-luvulla mielenfilosofit väittelivät antaumuksella siitä, kuinka (ganglio)solujen signaalien muunnosominaisuuksista seuraavat kärpäsdetektoriominaisuudet tai kuinka representaatioiden sisältö määräytyy suhteessa ulkoiseen todellisuuteen.

Hutton ja Myinin argumentit jatkavat tätä perinnettä. Myös heille representaatiot ovat havaintojärjestelmän tuottamia mielen tiloja, jotka viittaavat (vain) ulkoiseen todellisuuteen.<sup>10</sup> Heidän keskeisimmät argumenttinsa, kuten niin sanottu sisällön vaikea ongelma, käsittelevät lähinnä 1980-luvun kärpäsdetektoripohjaisten teorioiden vaikeuksia määritellä representaatioiden yksilöimisen kriteerejä ja niiden hyväksymis- tai totuusehtoja. Hutto ja Myin eivät edes varsinaisesti pohdi, voisivatko toiminnanohjaukseen liittyvät representaatiot olla muita kuin havaintotiloja tai suuntautua johonkin muuhun kuin ulkoiseen ympäristöön.

Havaintojärjestelmien lisäksi on kuitenkin myös muita tiedonkäsittelyn muotoja, jotka eivät ensisijaisesti esitä ulkoista ympäristöä. Esimerkiksi toiminnanohjauksen tehtävä ei ole varsinaisesti esittää ulkoista todellisuutta vaan turvata onnistunut

toiminta. Vahvistusoppimisen tapauksessa toiminnan onnistumista arvioidaan ennustamalla toiminnasta seuraavaa ennustetta maksimaalisesta palkkiosta tai ennustetta mallille annetun synteettisen todellisuuden kehityksestä. Nämä arviot puolestaan ovat esittäviä tiloja, representaatioita.

Olennaista on, että nämä representaatiot auttavat tiedonkäsittelyjärjestelmää valikoimaan toimintoja, eivät esittämään ulkoista todellisuutta. Ne eivät siis ole havaintojen kaltaisia tai synny havaintojen tapaan ulkoisten kohteiden aiheuttamien ärsykkeiden virittäminä. Esimerkiksi arvio saatavasta palkkiosta ei ole havaitun konkreettisen palkkion aiheuttama, eikä sitä vastaa ulkoisessa todellisuudessa mikään palkkio, joka voisi aiheuttaa palkkiota koskevan representaation viriämisen. Pikemminkin arvio palkkiosta tekee jostakin asiasta palkitsevan. Ei myöskään ole piirredetektoria, joka olisi erikoistunut jäljittämään palkkioiden tai ennustemallien piirteitä ympäristöstä.

Vahvistusoppimisalgoritmit ja niihin perustuva toiminnanohjaustutkimus tarjoavat siis esimerkin representaatioista, joiden avulla tiedonkäsittelyjärjestelmä voi suunnitella toimintaa. Nämä representaatiot eivät kuitenkaan ole havaintotilojen kaltaisia tai viittaa ulkoiseen todellisuuteen, vaan ne viittaavat arvioituihin toiminnan päämääriin. Niiden avulla ei viitata ulkoiseen ympäristöön vaan ohjataan toimintaa.

## Lopuksi

Hutto ja Myin seuraajineen ovat jo vuosien ajan hyökänneet toistuvasti kognitiontutkimuksen teoreettista ydintä vastaan. Heidän mukaansa toimintaa selitettäessä ei tarvita ulkoista todellisuutta esittäviä representaatioita vaan toiminta voidaan selittää jonkinlaisena reaktiivisena ympäristöön suuntautumisena. Robotiikan, tekoälyn – ja myös ihmisaivojen – tutkimus osoittaa kuitenkin, että tämä kritiikki ei perustu realistiselle käsitykselle kognitiontutkimuksen nykytilasta.

Monimutkaiseen toimintaan kykenevät robotit saattavat sisältää useita eri tehtäviä toteuttavia alijärjestelmiä. Kuten edellä kuvattiin, osa järjestelmistä saattaa toimia havaintojärjestelminä, jotka tekevät ennusteita ympäröivästä todellisuudesta ja vastaanottavat myös sitä koskevaa informaatiota. Toiset taas ennustavat, miten toimintaa kannattaa ohjata. Kolmannet vastaavat muistitiedon tallentumisesta, neljännet kommunikaatiosta ja niin edelleen. Tietoa käsittelevät siis toiminnanohjausjärjestelmän ohella muutkin järjestelmät. Toiminnanohjausjärjestelmät voivat saada syötettä näiltä muilta järjestelmältä.

Toiminnanohjausjärjestelmien tehtävä on ennakoida, suunnitella ja valikoida toimintaa, ja ne eivät välttämättä itse havainnoi konkreettista ympäristöä. Vahvistusoppimisalgoritmit tarjoavat yhden tavan tutkia tällaista toiminnan ennakointia laskennallisesti sekä simuloida tutkimuksen tuloksia konkreettisesti. Toiminnanohjausta koskevia oletuksia voi testata rakentamalla robotteja tai tekoälysovelluksia.

Nämä sovellukset ovat osoittautuneet poikkeuksellisen lupaaviksi, ja vahvistusoppimista pidetään laajalti yhtenä edistyskellisimpänä nykyisistä algoritmisista menetelmistä. Sille perustuvat sovellukset voittavat ihmisen monissa peleissä jo nyt, ja ne ennustavat usein osakemarkkinoita paremmin kuin ihmiset. Vahvistusoppimiseen perustuvat robottikädet pystyvät poimimaan esineitä, ja niillä ohjelmoidut ajoneuvot kykenevät navigoimaan liikenteessä.

Kyse ei siis ole yksittäisestä mallista vaan menestyksellisestä sovellusten sukupolvesta. Sen avulla tutkitaan täsmälleen samaa ilmiötä, jota enaktivistit markkinoivat osoituksena representatioiden tarpeettomuudesta. Silti toistaiseksi vahvimpien algoritmien valossa toimintaa ja toiminnanohjausta voidaan tarkastella myös representationaalisesti. Siten toiminnanohjauksen tutkimus ei osoita, että kognition tutkimuksen pitäisi päivittää uskonkappaleensa. Päinvastoin se osoittaa, että filosofien ja erityisesti enaktivistien tulisi päivittää representaatioita koskevat filosofiset intuitionsa.

## Representaatiot ja vahvistusoppimisalgoritmit

1. Hutto & Myin 2013.
2. Myin & Hutto 2015, 62.
3. Pavlov 1927.
4. Niv 2009.
5. Sutton & Barto 2018.
6. Doya 2008.
7. Doya 2008.
8. Ramsey 2007.
9. Lettvin ym. 1959.
10. Hutto 2006; 2013; Hutto & Myin 2020.

+ + +

- Doya, Kenji (2008). Modulators of Decision Making. *Nature Neuroscience* 11:4, 410–416.
- Hubel, David H. & Torsten N. Wiesel (1959). Receptive Fields of Single Neurons in the Cat's Striate Cortex. *The Journal of Physiology* 124:3, 574–591.
- Hutto, Daniel (2006). Knowing What? Radical Versus Conservative Enactivism. *Phenomenology and the Cognitive Science* 4, 389–405. <doi.org/10.1007/s11097-005-9001-z>
- Hutto, Daniel & Erik Myin (2013). *Radicalizing Enactivism: Basic Minds without Content*. Cambridge, MA: MIT Press.
- (2020). Deflating Deflationism about Mental Representation. Teoksessa Joulia Smortchkova, Krzysztof Dołęga & Tobias Schlicht (toim.): *What Are Mental Representations?* New York: Oxford University Press.

- Lettvin, Jerome Ysroael, Humberto Romersín Maturana, Warren Sturgis McCulloch & Walter Harry Pitts (1959). What the Frog's Eye Tells the Frog's Brain. *Proceedings of the IRE* 47, 1940–1951.
- Myin, Erik & Daniel Hutto (2015). REC: Just Radical Enough. *Studies in Logic, Grammar and Rhetoric* 41:54, 61–71.
- Niv, Yael (2009). Reinforcement Learning in the Brain. *Journal of Mathematical Psychology* 53:3, 139–154.
- Pavlov, Ivan Petrovitš (1927). *Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex*. Käänt. & toim. Gleb Vasilévich Anrep. London: Oxford University Press (ven. alkuteos 1927).
- Ramsey, William (2007). *Representation Reconsidered*. Cambridge: Cambridge University Press.
- Sutton, Richard S. & Andrew G. Barto (2018). *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press (ilm. alun perin 2014).