

# MIKSI KOGNITIOTIETEELLÄ EI OLE TEORIAA KESKUSPROSESSEISTA?

Pauli Brattico

Kognitiotiede, Helsingin yliopisto

*Aivojen ei-modulaaristen keskusprosessien tietämys on edistynyt hitaasti, sillä kaikki tunnetut keskusprosessien mallit ovat johtaneet laskennan kompleksisuuden ongelmaan. Yhteistä näille malleille on olettaa ihmisen keskuskognition olevan vain näennäisesti monimutkainen systeemi. Artikkelissa tarkastellaan hypoteesia, jonka mukaan aivot olisivat osittain aidosti monimutkainen järjestelmä. "Aito monimutkaisuus" määritellään algoritmisen informaatioteorian avulla.*

**Avainsanat:** Modulaarisuus, kompleksisuus, keskusprosessit, emergenssi, kompleksisuus, oppiminen, synnyntäisyys

## 1 JOHDANTO

Mentaalinen arkkitehtuuri jakautuu karkeasti modulaarisiin ja ei-modulaarisiin prosesseihin. Modulaaristen prosessien ymmärtäminen on lisääntynyt lupaavasti viimeisten vuosikymmenien aikana. Ei-modulaaristen keskusprosessien tapauksessa tilanne on toinen (Fodor, 1983; 2000). Ongelmana ei ole yrittämisen puute, sillä pelkästään tekoälytutkimuksen kenttä on sitoutunut tutkimaan keskusprosesseja kuten arkipäätelyä, ongelmanratkaisua, oppimista tai kielen ilmaisujen ymmärtämistä/käyttöä. Kielitieteessä funktionaalisen-typologisen sekä kognitiivisen kielentutkimus ovat ainakin osittain sitoneet itsensä keskusprosessien teorioihin,

sillä ne pyrkivät selittämään kielen piirteitä nojaamalla yleisiin kognitiivisiin mekanismeihin. Tällaisten pragmaattisten aspektien huomioiminen on välttämätöntä, jos kielen ominaisuudet halutaan selittää kielellisen kommunikaation tai kielellisen vuorovaikutuksen näkökulmasta (esim. Tomasello, 2003). Seuraavassa ensin lyhyt katsaus erilaisiin kognitiivisten keskusprosessien mallinnusyrityksiin sekä niiden ongelmiin. Pohdin samalla havaintojeni merkitystä kielitieteelle.

"Kognitiivisilla keskusprosesseilla" tarkoitetaan tässä artikkelissa kykyä integroida ja käsitellä tietoa useista lähteistä: aistimuksista, uskomuksista, aikaisemmista muistoista, tunteista, päätelyprosesseista ja niin edelleen. Esimerkiksi lauseen ymmärtäminen ja sen yhdistäminen kulloiseenkin tilanteeseen on prosessi, joka vaatii tiedollista integroitumista. Myös kielen käyttö on laajasti ajateltuna ilmiö, joka sisältää tällaisia integroivia prosesseja.

Kognitiotieteellinen keskusprosessien tutkimus alkoi varhaisista tekoälymalleista, jotka pyrkivät mallintamaan ongelmanratkai-

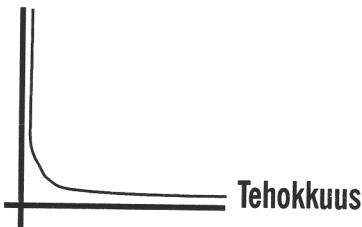
Artikkelin kirjoittamista on tukenut Suomen Kulttuurirahasto. Kiitän Otto Lappia, Lauri Alankoa, Ville Tuulosta ja Pertti Saariluomaa aihetta käsittelevistä keskusteluista. Materiaalin sisältöön on vaikuttanut myös vuonna 2002 pidetty tekoälyn filosofian kurssi tietojenkäsittelytieteen laitoksella; kiitän kurssin osallistujia aktiivisesta keskustelusta ja kritiikistä. Artikkelia koskeva kirjeenvaihto-osoite: Pauli Brattico, Kognitiotiede, Siltavuorenpenger 20, 00014 Helsingin yliopisto.

sua käyttämällä etsintää. Ideana on tutkia ja arvioida erilaisten vaihtoehtojen seurauksia ja valita niistä paras. Käytettäessä täydellistä hakua voidaan monissa tapauksissa löytää aina optimaalinen vaihtoehto. Täydellinen haku johtaa kuitenkin ihmisen kannalta triviaaleissa ongelmatapauksissa liian pitkäkestoiseen etsintäprosessiin hakuavaruuden kasvaessa räjähdysmäisesti.

Ongelma pyrittiin ratkaisemaan käyttämällä heuristisia malleja. Heuristiikan ideana on vähentää ongelmanratkaisuun kuluva aikaa rationaalisuuden kustannuksella. Heuristisesti toimiva ongelmanratkaisija ei pyri aina parhaaseen ratkaisuun, jolloin osa vaihtoehdoista jää tutkimatta ja päättelyyn jää aukkoja. Täysin rationaalisen systeemin vaatimus on selvästi epärealistinen myös ihmisen tapauksessa.

Heuristiikan ongelmana on, että kun järjestelmästä tulee riittävän tehokas, se on jo menettänyt liian suuren määrän rationaalisuutta. Vastaavasti jos rationaalisuutta eliminoidaan vain vähän, on hyöty tehokkuuden kannalta minimaalinen. Rationaalisuuden ja tehokkuuden välinen funktio on liian jyrkkä. Tilannetta voidaan havainnollistaa seuraavalla kuvalla:

## Rationaalisuus



Kuva 1. Monissa malleissa rationaalisuuden ja tehokkuuden suhde on liian jyrkkä ollakseen kognitiivisesti realistinen. Ihmisen aivot kuitenkin selviävät arkielämän ongelmatilanteista nopeasti ja luotettavasti.

Funktio kertoo, että riittävä rationaalisuus voidaan säilyttää vain hyvin tehottomalla järjestelmällä. Käytännössä tehokkaat heuristiikat ovat ihmiseen verrattuna irrationaalisia eivätkä siten realistisia kuvauksia ihmisen keskusprosesseista. Ne eivät selviä yksinkertaisistakaan tilanteista, joita ihmiset toisaalta ratkaisevat lähes intuitiivisesti (ks. Dreyfus, 1972; Pylyshyn, 1986).

Rationaalisuuden käsitettä on vaikea määrittellä, eikä sitä ole tarkoituksenmukaista tässä formalisoida. Kun puhumme ”kognition mallin rationaalisuudesta”, vertaamme mallin päättelykykyä ihmisen päättelykykyyn. Jotta malli olisi psykologisesti uskottava, täytyy sen päättelykyvyn muistuttaa ihmisen päättelykykyä. Karkeasti voidaankin sanoa, että nykyiset komputationaaliset mallit ovat rationaalisia ja tehokkaita juuri niissä tehtävissä, joissa ihmiset eivät ole, ja päinvastoin. Ne näyttävät kuvaavan kaikkea muuta kuin ihmisen luontaista kognitiota.

Rationaalisuus sinänsä on abstrakti käsite, joka soveltuu yhtä hyvin sekä biologisiin että keinotekoisin representatiivisiin järjestelmiin. Sillä tarkoitetaan tässä kykyä tuottaa tosia uskomuksia, joiden pohjalta on mahdollista toimia järkevästi. Irrationaalinen järjestelmä tuottaa epätosia uskomuksia. Ns. kehysongelma (frame-problem) kuvaa hyvin komputationaalisten mallien epärationaalisuutta (McCarthy & Hayes, 1969; Pylyshyn, 1986). Johonkin tavoitteeseen pääseminen edellyttää käytännössä aina jonkin toimintasarjan suorittamista tietyssä järjestyksessä. Jos haluan soittaa puhelimella, täytyy ensin nostaa luuri, ja vasta sitten näppäillä numero. Numeron näppäilyllyn ehtona on usein se, että numero etsitään ensin puhelinluettelosta sekä se, että käytössä on puhelin. Mutta entä jos puhelinluettelon selaaminen aiheuttaa puhelimen häviämisen? Silloinhan tehtävää ei kannata lähteä ratkaisemaan selaamalla puhelinluetteloa. Täytyy siis tutkia,

mitä seurauksia puhelinluettelon selaamisella voi olla, ja onko näiden joukossa puhelimen katoaminen. Tämän tutkimiseen menee valtavasti aikaa, sillä sama ongelma toistuu jokaisessa päättelyaskeleessa. Eri vaihtoehtojen ja niiden seurauksien punnitseminen johtaa kombinatoriseen räjähdykseen. Ihmisen kognitiolle puhelimella soittaminen ei tuota lainkaan tällaisia ongelmia. Päinvastoin, ihminen kykenee sulkemaan kaikki irrelevantit vaihtoehdot pois pohtimatta niitä. Esimerkiksi de Groot (1965) arveli tämä kyvyn shakinpeluussa perustuvan relevanttien vaihtoehtojen ”näkemiseen” eikä kognitiiviseen päättelyyn<sup>1</sup>.

Kuvassa 1 esitetty funktio esittää empiiristä löydöstä: se kertoo, miten tähän asti kehitetyt mallit ovat käyttäytyneet. Malleilla on aina liian vähän joko rationaalisuutta tai tehokkuutta. Yhteys rationaalisuuden ja tehokkuuden välillä ei siten ole välttämättä kuvan funktion kaltainen, joten tilanne voi olennaisesti parantua oikean heuristikan löydyttyä. Kun mitään ratkaisua ei kuitenkaan löydetty, siirryttiin luomaan ns. asiantuntijajärjestelmiä. Näiden ideana on saavuttaa riittävä rationaalisuus ja tehokkuus pienentämällä ongelmaympäristöä. Tällaisten systeemien ongelmanratkaisu soveltuu siten vain tietyn kapean erikoisalan piiriin. Tuon alan ulkopuolella ne ovat joko täysin toimintakyvyttömiä tai liian hitaita. Käytännössä malli toteutetaan antamalla sille riittävä määrä ennakkotietoa tehtävyympäristöstä. Malli kuvaa siten modulaaristunutta tietämystä, jolloin sen soveltuvuus keskuskognition malliksi on riittämätön<sup>2</sup>.

Yksi vaihtoehto on olettaa, että rationaalisuuteen ja tehokkuuteen tarvittava informaatio saadaan ympäristöstä. Oppiminen ei kuitenkaan ole ratkaissut ongelmia. Päinvastoin. Oppijan on havaintojensa perusteella tehtävä hypoteeseja ympäristönsä lainalaisuuksista. Hypoteesien kokeileminen on toi-

saalta aina etsintää: annettuna havaintoaines D, tehtävänä on löytää ja testata jotakin hypoteesia H, aivan kuten esimerkiksi shakinpeluussa on nykyinen peliasema (D) yhdistettävä parhaaseen siirtoon (H). Oppiminen on itse asiassa rationaalisuuden tavoittelua etsinnän keinoin, ja siten se on vain ongelmanratkaisun alatapaus. Käytännössä onkin niin, että oppiviin järjestelmiin pätevät kaikki edellä tarkastellut ongelmat. Jos ne ovat rationaalisia, ne ovat liian hitaita. Jos ne ovat riittävän tehokkaita, ne eivät opi mitään<sup>3</sup>. Kuten päättelyssä, irrelevantit vaihtoehdot täytyy oppimisessäkin sulkea etukäteen pois ilman että niitä täytyy eksplisiittisesti kokeilla.

Ainoa vaihtoehto on rajoittaa oppimista ja luoda asiantuntijajärjestelmiä, toisin sanoen luoda järjestelmiä jotka pystyvät oppimaan vain tietyn erikoisalan tietämystä rajoitetun määrän. Tämä johtopäätös on huomioitu kielitieteessä mm. generatiivisessa traditiossa, missä tavoitteena on muodostaa teoria pitkälle erikoistuneesta ja myötäsyttyisesti kehittyvästä kielimoduulista joka on erikoistunut kielen omaksumiseen ja valikoivaan kielellisen syötteen prosessointiin.

Erityisesti suomalaisessa kielitieteellisessä kirjallisuudessa tähän ehdotukseen on suhtauduttu varsin negatiivisesti, aivan kuin kognitiiviteollä olisi jo ratkaissut nämä oppimiseen liittyvät ongelmat. Näin ei kuitenkaan ole. Esimerkiksi Tomasellon (2003) mukaan lapsi voi oppia kielen, koska hänellä on kyky etsiä invariansseja aistimuksesta (”pattern-finding skill”). Tämä kyky perustuu analogioiden ja distributionaalisen analyysin tekemiselle (s. 162). Itkonen (1986) on samoilla linjoilla. Näistä teeseistä kaikki lienevät samaa mieltä. Oppimisen ongelma on siinä, että mikäli oppijan annetaan muodostaa vapaasti analogioita ja invariansseja aistimustensa perusteella, tämä johtaa loputtomaan etsintään. Vaihtoehto-

ja ja niiden seurauksia on tietysti liikaa. Miten tehdä ”järkeviä” hypoteeseja kokeilematta kaikkia muita? Toisin sanoen, miten rajoittaa etsintää menettämättä kaikkea rationaalisuutta?

Kuten generatiivisessa traditiossa ja kognitiotieteessä, myös Tomasellon mukaan tähän tarvitaan kielispesifejä rajoituksia (”constraints”, ks. luku 5.3), joista hän mainitseeikin jo alustavasti muutamia. Tämän lähestymistapa johtaa käytännössä asiantuntijajärjestelmään, missä etsintää rajoitetaan etukäteen. Itkonen (1986, s. 272) havainnollistaa analogioihin perustuvaa oppimista konkreettisemmin Chomskyltä lainatulla esimerkillä. Tarkastellaan lausetta (1).

(1) Mary bought a dog to play with

Jossain vaiheessa olemme omaksuneet tietoa, joka kertoo että infiniittisen lauseen tekijä on Mary. Voidaanko tällaista tietoa omaksua oppimalla? Itkosen mukaan tämä tilanne päätellään toisen, analogisen lauseen pohjalta. Esimerkiksi lauseessa (2) ”on mahdollinen vain yksi tulkinta: Maryn täytyy olla subjekti” (s. 272). Tätä tietoa sovelletaan analogisesti lauseeseen (1).

(2) Mary bought a ball to play with

Hypoteesien määrää on näin ollen rajoitettu etukäteen niin paljon kuin on käytännössä mahdollista: (A) lähtökohtana on lause (2) ja (B) vain yksi tulkinta on siinä mahdollinen. Mistä nämä rajoitukset ovat peräisin? Hieinan erilaisessa lauseessa (3) tulkinta muuttuu: nyt koirakin voi olla infiniittisen lauseen subjekti:

(3) Mary bought a dog to play with her

Miksi oppija ei olettaisi, että lause (1) on muodostettu analogisesti lauseesta (3) hil-

jentämällä lauseen lopussa oleva pronomini? Entä miksi lause (2) ei ole analoginen lauseen (3) kanssa? Miksi tulkinta muuttuu kun lauseen loppuun ilmestyy pronomini? Toisin sanoen, kun käytämme lausetta (2) analogian lähtökohtana, on meidän jo pitänyt tietää etukäteen että tässä tapauksessa – pronominin puuttuessa – pallo ei voi olla infiniittisen verbin subjekti / tekijä. Pragmaattinen tieto ei tässä tapauksessa auta, sillä lauseen (3) tulkinta ei mitenkään muutu vaikka ostamisen kohteena olisi pallo: tuloksena on vain lause, joka kuvaa ehkäpä jotakin tulevaisuudessa mahdollista tilannetta, missä pallotkin osaavat leikkiä. Generatiivisessa kieliteoriassa tähän ongelmaan suhtaudutaan pyrkimällä muotoilemaan teoria tyhjistä elementeistä joka rajoittaa analogioiden ja invarianssien muotoilemista. Pelkästään oppimiseen perustuva mekanismi joutuisi kokeilemaan äärettömän monia muita tilanteeseen soveltuvia hypoteeseja, sillä ellei kielestä tiedetä mitään etukäteen, voisi mikä tahansa niistä olla aivan hyvin totuudenmukainen. Oppiminen tukehtuu kombinatoriseen räjähdykseensä. Analogioiden tai invarianssien tehokas löytäminen siis vain olettaa, että oppimisrajoitusten ongelma on jo ratkaistu, se ei itse ratkaise sitä.

Tässä yhteydessä voitaisiin ehdottaa, että hypoteeseja kokeillaan niiden yksinkertaisuuden osoittamassa järjestyksessä. Kognitiotieteessä tämäntapaista ehdotusta on toisinaan kutsuttu evaluaatiometriikaksi. Sen tarkoituksena on järjestää hypoteesit jonkin periaatteen, esimerkiksi yksinkertaisuuden, perusteella. Tällainen evaluaatiometriikka edustaa siten etukäteen asetettuja rajoituksia, ja sitä pidettiin pitkään keskeisenä osana lapsen synnynnäistä tietoa eli universaalia kielioppia (Chomsky, 1965, s. 37-47).

Erityisen hyvin keskuskognition mallintamisen ongelma tulee esille mekaanisen konekäännöksen yhteydessä. Kone pystyy luo-

maan melko hyvän syntaktisen analyysin sekä kohde- että lähdekielestä, mutta ei useinkaan luomaan järkevältä kuulostavaa eli rationaalista käännoästä. Tämä johtuu integroivien prosessien puutteesta. Esimerkki (4) havainnollistaa tätä ilmiötä. Jälkimmäisen lauseen hän- pronominin merkitys lainataan edellisestä lauseesta. Tätä riippuvuussuhdetta olen merkinnyt alaindeksillä:

- (4) a. Lääkäri<sub>(i)</sub> sanoi potilaalle<sub>(i)</sub>, että hän<sub>(i)</sub> on sairas  
 b. Potilas<sub>(i)</sub> sanoi lääkärille<sub>(i)</sub>, että hän<sub>(i)</sub> on sairas  
 c. Potilas sanoi lääkärille<sub>(i)</sub>, että hän<sub>(i)</sub> on sairas pervo  
 d. Potilas<sub>(i)</sub> myönsi lääkärille<sub>(i)</sub>, että hän<sub>(i)</sub> on sairas pervo

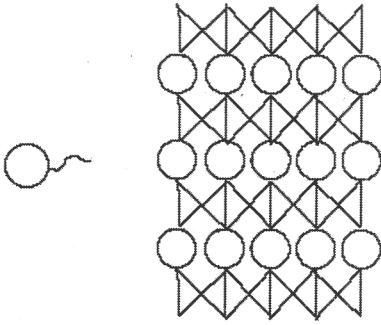
Lauseet (a–b) osoittavat, että pronominin viittauskohde ei määräydy lauseen lokaalisen ominaisuuksien perusteella. Pronomini yhdistetään potilaaseen kummassakin lauseessa, koska lauseen kuvaamassa tilanteessa potilas on todennäköisesti se, joka on sairas. Esimerkit (c–d) osoittavat, että tulkinta on sensitiivinen lauseen sanojen merkityksille. Ongelmana tässä on se, että sanan merkitys tai sen kuvaaman tapahtuman ominaisuudet eivät ole lauseen ominaisuuksia: niihin vaaditaan ensyklopedista tietoa maailmasta, toisin sanoen on kyettävä yhdistämään informaatiota monesta lähteestä kuten lauseen rakenteesta, ensyklopedisesta tiedosta, aikaisemmasta kontekstista, ja niin edelleen. Pronominin viittauskohdetta ei voi siis laskea pelkästään lauseen rakenteeseen nojautuen.

Tämän kaikkia kausaalisia/komputaationaalisia selitysyrittäviä luonnehtivan laskennan kompleksisuuden ongelman olemassaolo on tiedostettu jo kauan. Se on pysäyttänyt edistyksen keskuskognition tutkimuksessa. Myös sellainen kielitiede, joka tutkii ihmisten vuorovaikutusta, kommunikaatiota ja kielen kognitiivisia aspekteja, on sitonut itsensä keskuskognition teorioihin, joita ei toistaiseksi voida todentaa.

## 2 AIVOT JA KOMPLEKSISUUS

Kaikilla edellä mainituilla malleilla on ongelmana eräänlainen laskennallinen kompleksisuus: rationaalisuuden ja tehokkuuden suhde on liian huono. Optimistisen näkemyksen mukaan kompleksisuus on peräisin käytetystä algoritmista. Ratkaisu olisi entistä parempien algoritmien kehittäminen. Tätä näkemystä voidaan kritisoida siitä, että ratkaisua on yritetty keksiä jo 50 vuotta – turhaan. Ja kun jokin selitys ei onnistu, on jossain vaiheessa kosmeettisten korjausten sijaan yritettävä etsiä syytä perustavammalta tasolta. Tämä antaa aiheen etsiä ratkaisua toisesta suunnasta.

Kaikkia edellä tarkasteltuja malleja luonnehtii tietynlainen yksinkertaisuus: ne pyrkivät kuvaamaan keskusprosesseja käyttämällä mahdollisimman vähän informaatiota ("ohjelmakoodia"). Täydellistä hakua tekevät ohjelmat nojaavat rajoittamattomaan rekursioon, missä sama funktio "kutsuu itseään" tai matemaattisemmin ilmaistuna funktion määrittelyssä esiintyy funktio itsessään. Näin voidaan minimoida tarvittavan informaation määrä, kun monimutkainen toiminta typistyy elegantisti yhteen koodinpätkään. Heuristiikkojen tarkoituksena on rajoittaa funktion alaa, mutta informaation määrä pidetään edelleen alhaisena. Konnektionistisissa malleissa pyritään puolestaan luomaan monimutkaisia verkkoja koodaamalla yhden formaalin solun toimintaa ja monistamalla tätä solua. Näin syntyy monimutkaiselta näyttävä verkko, joka on edellä kuvatussa mielessä syvemmältä olemukseltaan yksinkertainen. Monet alan tutkijat pitävätkin juuri tätä äärimmilleen vietyä yksinkertaisuutta konnektionismin parhaana puolena (Rumelhart ym., 1996, s. 533). Tilannetta on havainnollistettu kuvassa 2.



Kuva 2. Monimutkaiselta näyttävä konnektionistinen verkko pitää sisällään vain vähän informaatiota: alkuperäinen ohjelma sisältää vain yhden formaalin solun ohjelmakoodin sekä solujen yhteyksien topografian. Verkon painokerroimet alustetaan usein pseudosatunnaisesti. Tällaiset homogeeniset konnektionistiset mallit nojaavatkin käytännössä aina oppimiseen, koska sellaisenaan ne ovat informaatioisällöltään lähes täysin tyhjiä. Konnektionistinen malli voi tietysti olla myös aidosti monimutkainen, olettaen että verkon painokertoimiin sisällytetään aitoa informaatiota.

Oppivassa järjestelmässä luotetaan siihen, että tarvittava informaatio saadaan automaattisesti ympäristöstä: tällöin malli itsessään voi olla alussa hyvin yksinkertainen. Asiantuntijajärjestelmät puolestaan koodaavat tietoa ympäristöstä, mutta koska informaation määrää pyritään pitämään vähäisenä, niistä tulee hyvin kapean alan osaajia. Tarvittavan alkuinformaation olematon määrä on siis yksi malleja yhdistävä tekijä. Voisiko se olla syynä siihen, että kaikkia malleja vaivaa laskennan kompleksisuuden ongelma?

Kun malliin on upotettu vain vähän informaatiota, mutta se näyttää kompleksiselta, puhutaan pseudokompleksisuudesta. Esimerkiksi piin desimaalikehitelmä on pseudokompleksinen: se vaikuttaa satunnaiselta, mutta pelkistyy matemaattiseksi kaavaksi joka on yksinkertainen. Käänteistä toimenpidettä kutsutaan emergenssiksi tai itseorganisaatioksi: monimutkaisuus ”emergeoittuu” yksinkertaisista periaatteista (Auang, 1998, Dennett, 1991, Holland, 1998, Kauffman, 1996, Simon, 1996, ks. myös Shalizi, 2001)<sup>4</sup>.

’Self-organization’ suggests a spontaneous increase of complexity occurring in a system with simple, generic (e.g. spatially homogeneous) initial conditions. (Bennett, 1988, s. 252).

Tällainen pseudokompleksisuus on kaikkia keskuskognition malleja luonnehtiva ominaisuus. Olisiko olemassa lukuja/olioita, joita ei voitaisi tuottaa aloittamalla pienestä määrästä alkuinformaatiota? Esimerkiksi suurin osa luonnollisista luvuista on sellaisia, ettei niitä voida lainkaan pelkistää. Niihin ei kätkeydy minkäänlaista säännönmukaisuutta. Kognitiotieteessä on siis tehty oletus, jonka mukaan ihmisen aivot ovat vain näennäisesti monimutkainen systeemi. Tällainen malli epäonnistuu kognitiivisten prosessien kuvaamisessa, mikäli ihmisen mieli/aivot ovatkin aidosti kompleksinen systeemi. Koska vaikuttaa siltä, että kaikissa malleissa on todella jokin pielessä, ja lisäksi että tämä ongelma liittyy jollain tavalla kompleksisuuteen, on mielekästä tarkastella tätä hypoteesia mahdollisena selityksenä: olisivatko aivot ja erityisesti niiden ei-modulaariset prosessit ”aidosti kompleksinen systeemi” (Sallo, 1998; 2000a; b)?

### 3 AITO JA NÄENNÄINEN KOMPLEKSISUUS BIOLOGISISSA SYSTEMEISSÄ

Kompleksisuutta voidaan mitata monella tavalla. Yhteistä näille kaikille tavoille on se, että monimutkaisuus voidaan yhdistää intuitiivisesti tarvittavan työn määrään. Laskennan kompleksisuus tarkoittaa sitä, että jonkin ongelman ratkaisemiseksi on tehtävä valtavan paljon laskentaa tai siihen vaaditaan suuri määrä muistia. Algoritmien kompleksisuus viittaa tarvittavan informaation määrään ja siten koodin spesifioimiseen tai riippumattomien valintojen tekemiseen (tästä enemmän jäljempänä). Myös intuitiivisesti kompleksisuus liittyy johonkin, mitä

on ”vaikea ymmärtää” tai minkä ymmärtäminen vaatii paljon työtä.

Äärellisen olion sisältämän informaation määrä voidaan määrittellä ns. Kolmogorov-kompleksisuuden avulla (ks. Li & Vitayi, 1997)<sup>5</sup>. K-kompleksisuuden ideana on määrittää lyhimmän tietokoneohjelman pituus, joka kuvaa/tuottaa kyseisen olion. Ohjelmointikieleksi valitaan ns. universaali Turingin kone  $f$ , koska se pystyy simuloimaan mitä tahansa laskennallista järjestelmää. Tarkasteltaessa olion  $x$  K-kompleksisuutta valitaan se ohjelmointikieli  $r$  (PASCAL, C, BASIC...) ja se ohjelma  $(p)$ , jotka yhdessä tuottavat mahdollisimman lyhyen kuvauksen oliolle  $x$ .

$$(1) K(x) : \min\{|rp| : f(rp) = x\}$$

Olio, jonka K-kompleksisuus on samaa luokkaa kuin olion itsensä koko, on satunnainen: siinä ei esiinny lainkaan säännönmukaisuuksia tai invariansseja joiden keksiminen pelkistäisi kuvausta. Korkea K-kompleksisuus ja satunnaisuus rinnastuu entropiaan: mitä enemmän satunnaisuutta, sitä enemmän entropiaa. Yksinkertainen olio on sellainen, jonka kuvausta voidaan pelkistää suhteellisen paljon. Esimerkiksi '010101...01' on yksinkertainen olio. Monet oliot ovat monimutkaisuudeltaan jotakin tältä väliltä. Pseudokompleksiset oliot ovat sellaisia, jotka vain näyttävät (ihmisistä tai ”kevyistä” satunnaistesteistä) monimutkaisilta, mutta joilla on itse asiassa hyvin lyhyt laskennallinen kuvaus.

Matemaattinen formalismi – ohjelma  $p$ , universaali kone  $f$  ja olio  $x$  – on tässä pelkkä abstraktio: sen tarkoitus on kuvata toisistaan riippumattomien valintojen määrää. Yksinkertaisimmillaan näitä valintoja kuvataan käyttämällä bittejä 0–1, mutta mikä tahansa muukin notaatio on mahdollinen. K-kompleksisuuteen ei vaikuta valittu notaatio. ”Tietokoneohjelma”  $p$  yhtälössä on siis

vain joukko valintoja, jotka yhdessä universaalien koneiden kanssa riittävät jonkin olion  $x$  tuottamiseen. K-kompleksisuus kertoo lisäksi, mikä on välttämätön määrä valintoja, kun tarkoituksena on tuottaa annettu  $x$ . Välttämätön ja riittävä määrä valintoja tarkoittaa itse asiassa sitä, että valinnat ovat keskenään täysin riippumattomia. Säännönmukaisuus taas tarkoittaa, että valintojen välillä on riippuvuusuhteita eli säännönmukaisuuksia. ”Valinnat” voivat viitata vaikka ihmisenkaltaisen robotin rakennusprosessiin tai aivoja kuvaavan mallin rakentamiseen<sup>6</sup>.

Kuten edellä tuli todettua, monet kognitiiviset keskuskognition mallit kuuluvat pseudokompleksisten olioiden joukkoon, samoin esimerkiksi sellaiset luvut kuin  $\pi$ . Myös dynaamiset mallit tuottavat pseudokomplekseja olioita (esim. fraktaalit). Olenaista on kuitenkin se, että suurin osa mahdollisista olioista (esimerkiksi luonnollisista luvuista) on kuitenkin hyvin kompleksisia, ja vain harvat yksinkertaisia. Suurin osa kognitiotieteilijöistä olettaa, että aivot ovat edellä määritellyssä mielessä perimmäiseltä olemukseltaan yksinkertaiset; osa tutkijoista on käyttänyt K-kompleksisuuden käsitettä tämän hypoteesin täsmentämiseen (ks. Chaitin, 1979: 7–8; Chater, 1996; 1999; Sambrook & Whiten, 1997). Kuten totesin, tämä on vain empiirinen oletus, ei matemaattinen välttämättömyys. Entä jos biologiset oliot ovat aidosti kompleksisia? On empiirinen kysymys, kumpaan joukkoon biologiset organismit mukaan lukien ihmisen aivot kuuluvat.

Näennäisen kompleksisuuden oletus voisi selittää, miksi pseudokompleksiset mallit epäonnistuvat keskuskognition kuvaamisessa. Lisäksi hypoteesi selittää, miksi epäonnistuminen heijastuu aina laskennan kompleksisuutena. Laskennallinen kompleksisuus, joka liittyy rationaalisuuteen ja tehokkuuteen, on yhteydessä Kolmogorov-komplek-

sisuuteen: Informaation lisääminen malleihin sekä ainakin teoriassa ratkaisee nykyisten mallien ongelmat että selittää, miksi nykyiset mallit ovat komputationaalisesti liian kompleksisia.

K-kompleksisuus on informaatiota, jota ei voi luoda tyhjästä. Tällainen kompleksisuus ei voi emergoitua. Tämä merkitsee sitä, että aivoihin koodautunut informaatio ei voi purkautua ulos pelkästään geeneistä. Ihmisen perimässä näyttää olevan hyvin vähän informaatiota K-kompleksisuuden avulla mitattuna (Li & Vitányi, 1997: luku 7.7). Mistä aivojen entropia siis on peräisin? Ympäristöön vetoaminen ei silläkään voi ratkaista ongelmaa, sillä käyttökelpoisen informaation poimiminen ympäristöstä johtaa laskennan kompleksisuuteen ja satunnaisen informaation poimiminen taas tuottaa epärationaalisen järjestelmän. Ainoa järkevä informaation lähde on se osa ympäristöstä, joka (1) pysyy vakiona organismista toiseen ja joka (2) tuottaa organismin käyttäytymisen ja selviytymisen kannalta järkevää informaatiota. Tällainen informaatio voisi olla peräisin esimerkiksi niistä monimutkaisista reunaehdoista, joissa alkionkehitys tai elämä tuntemassamme muodossa on ylipäänsä mahdollinen. Toisin sanoen kompleksisuus on systeemin koko alkutilassa sekä geenien lukuprosessissa, ei pelkästään geeneissä itsessään (Gottlieb, 1998; Oyama, 2000). Hypoteesi johtaa interaktionistiseen näkemykseen informaation alkuperästä:

Organismic form ... is not transmitted in genes any more than it is contained in the environment, and it cannot be partitioned by degrees of coding or by amounts of information. It is constructed in the developmental processes ... Chromosomal form is an interactant in the choreography of ontogeny; the "information" it imparts or the form it influences in the emerging organism depends on what dance is being performed when, where, and with whom. (Oyama, 2000: 26).

Tämän biologisen alkutilan kompleksisuus on viime kädessä peräisin evoluutiosta ja sitä ohjaavista mekanismeista. Biologisen informaation kompleksisuus on joko sen yksinkertaisuuden loogista syvyyttä (Bennett, 1988; Li & Vitányi, 1997: 7.7) tai se on peräisin fysikaalisesta entropian lisääntymisestä. "Itseorganisoituminen" voidaan tässä viitekehityksessä nähdä täysin satunnaisen, korkeaentropisen järjestelmän "kristallisoitumisena" sen entropiatason laskiessa (Wolfram, 1983). Kompleksisuuden alkuperä ja syntyminen on fysiikan ja biologian kannalta avoin ongelma.

Aivojen aito kompleksisuus on kognitiotieteen kannalta ongelmallista, sillä sen mukaan ihmisen kognitiota – ainakaan keskusprosesseja – ei voida pelkistää lainkaan lainomaisiin kuvauksiin. Kun tieteen tehtävänä on ympäröivän satunnaisuuden pelkistäminen, saattavat biologiset järjestelmät olla sellaisia, että niissä esiintyy runsaasti pelkistymätöntä entropiaa. Biologinen järjestelmä yhdessä ympäristönsä kanssa muodostaa funktionaalisen, mutta aidosti kompleksisen, korkeaentropisen systeemin.

Vaikka keskuskognition tutkiminen saataakin kohdata tällaisia vaikeuksia, on aivojen modulaaristen prosessien kohdalla vielä jonkin verran toivoa. Tämä on Fodorin (1983) modulaarisuushypoteesin keskeisin osa. Kielen tutkimuksen kohdalla se merkitsee rajoittumista modulaaristen järjestelmien tutkimiseen. Ihmiselle ominainen tilanetaju ja informaation tehokas integroitumiskyky ei ehkä perustu yhteen tai kahteen mentaaliseen prosessiin (kuten "oppiminen", "etsintä", "produktiosäännön soveltaminen"), vaan valtava tietomassa on elimellinen osa itse prosessia eikä sitä voida abstrahoida pois.



## VIITTEET

- Auyang, S. Y. (1998). *Foundations of complex-system theories: In economics, evolutionary biology, and statistical physics*. Cambridge, MA: Cambridge University Press.
- Bennett, C. H. (1988). Logical depth and physical complexity. Teoksessa R. Herken (toim.), *The universal Turing machine: A half-century survey*. Oxford: Oxford University Press.
- Chaitin, G. J. (1969). On the length of programs for computing finite binary sequences. *Journal of Association of Computer Machinery*, 13, 547–569.
- Chaitin, G. J. (1979). Toward a mathematical definition of "life". Teoksessa Levine, R. D. & Tribus, M. (toim.) *The Maximum entropy formalism*. Cambridge, MA: MIT Press.
- Chater, N. (1996). Reconciling Simplicity and Likelihood Principles in Perceptual Organization. *Psychological Review*, 103, 566–581.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press
- Crutchfield, James P. (1994a). The calculi of emergence: Computation, dynamics, and induction. *Physica D*, 75, 11–54.
- Crutchfield, James P. (1994b). Is anything ever new? Considering emergence. Teoksessa Cowan, G., Pines, D. & Melzner, D. (toim.), *Complexity, metaphors, models, and reality*. Massachusetts: Addison-Wesley, 479–497.
- Dennett, D. (1991). Real patterns. *Journal of Philosophy* 88, 27–51.
- De Groot, *Thought and choice in chess*. Mouton: The Hague.
- Dreyfus, H. (1972). *What computers can't do*. New York: Harper Row.
- Fodor, J. A. (1983). *The modularity of mind: An essay on faculty psychology*. Cambridge, MA: MIT Press.
- Fodor, J. A. (2000). *The mind doesn't work that way. The scope and limits of computational psychology*. Cambridge, MA: MIT Press.
- Gottlieb, G. (1998). Normally occurring environmental and behavioral influences on gene activity: From central dogma to probabilistic epigenesis. *Psychological Review*, 105, 792–802.
- Holland, John H. (1998). *Emergence: from chaos to order*. Oxford University Press: Oxford.
- Itkonen, E. (1986). Ajatuksia Chomskyn nykyvaiheesta. *Virittäjä*, 263–277.
- Judd, J. S. (1990). *Neural network design and the complexity of learning*. Cambridge, MA: MIT Press.
- Judd, S. J. (1996). Complexity of learning. Teoksessa P. Smolensky M. C. Mozer D. E. Rumelhart (toim.), *Mathematical perspectives on neural networks*. New Jersey: Erlbaum.
- Kauffman, Stuart (1995). *At Home in the Universe: The search for the laws of self-organization and complexity*. Oxford University Press: Oxford.
- Kolmogorov, A. N. (1965). Three approaches to the quantitative definition of information. *Problems of Information Transmission* 1, 1–7.
- Li, M., & Vitányi, P. (1997). *Introduction to Kolmogorov complexity and its applications*. New York: Springer-Verlag.
- Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review* 95, 492–527.
- McCarthy, J. & Hayes, P. J. (1969). Some Philosophical Problems from the Standpoint of Artificial Intelligence. Teoksessa B. Meltzer and D. Michie (Toim.) *Machine Intelligence* 4, Edinburgh: Edinburgh University Press, pp. 463–502.
- Minsky, M. (1986). *The society of mind*. New York: Simon and Schuster.
- Minsky, M. L., Papert, S. A. (1990). *Perceptrons*. Cambridge, MA: MIT Press. (2. painos.)
- Oyama, S. (2000). *The ontogeny of information: Developmental systems and evolution*. Durham, NC: Duke University Press.
- Pylyshyn, Z. (toim.) (1986). *The robot's dilemma: the frame problem in artificial intelligence*. Norwood: Ablex.
- Rumelhart, D. E., Drubin, R., Golden, R. & Chauvin, Y. (1996). Backpropagation: The basic theory. Teoksessa P. Smolensky, M. C. Mozer & D. E. Rumelhart (toim.) *Mathematical perspectives on neural networks*, s. 533–566.
- Salo, P. (1998). *Complexity and cognition: Kolmogorov complexity and its application to cognitive science*. Julkaisematon lisensiaatintutkielma, Helsingin yliopisto.
- Salo, P. (2000a). Psykofyysiset lait ja intentionaalisten tilojen redusoitumattomuus. *Ajatus* 57, 57–75.
- Salo, P. (2000b). Computation in the human mind? Teoksessa Hyötyniemi, H. (toim.) *Step2000 – Millennium of Artificial Intelligence*, 111–121.

- Salo, P. (2003). Kielen omaksuminen – kypsymistä vai oppimista? *Puhe ja kieli* 23, 65–76.
- Sambrook, T., & Whiten, A. (1997). On the Nature of Complexity in Cognitive and Behavioural Science. *Theory and Psychology*, 7, 191–213.
- Simon, Herbert A. (1996). *The Sciences of the Artificial*. Cambridge, MA: MIT Press. (3. painos).
- Shalizi, C. R. (2001). *Causal Architecture, Complexity and Self-Organization in Time Series and Cellular Automata*. Academic dissertation, University of Wisconsin.
- Solomonoff, R. J. (1964). A formal theory of inductive inference, part 1, part 2. *Information and Control*, 7, 1–22, 224–254.
- Tesauro, G. (1987). Scaling relationships in back-propagation learning: Dependence on training set size. *Complex Systems*, 1, 367–372.
- Tesauro, G., Janssens, R. (1988). Scaling relationships in back-propagation learning: Dependence on predicate order. *Complex Systems*, 2, 39–44
- Tomasello, Michael (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.
- Watanabe, O. (1992)(Toim.) *Kolmogorov complexity and computational complexity*. Berlin: Springer-Verlag.
- Wolfram, Stephen (1983). Statistical Mechanics and Cellular Automata. *Reviews of Modern Physics* 55, 601–644.
- [2] De Groot tarkastelee teoriaa, jonka mukaan tämä näkökyky palautuisi kaikkien/useiden vaihtoehtojen läpikäymiseen, ks. De Groot, 1965, s. 405 ja erityisesti alaviite 15. Automaation teoriassa kognitiivisten kykyjen nopeutumista on selitetty algoritmisen tiedon korvautumisella instanssipohjaisella kovakoodauksella; ks. Logan 1988, erityisesti s. 493.
- [3] Minsky (1986) ehdotti ratkaisuksi teoriaa, jossa koko mentaalinen arkkitehtuuri koostuu useista asiantuntijoista. Samanlaisen ratkaisuun ovat päätyneet evoluutiopsykologit, jotka kannattavat massiivisen modulaarisuuden teoriaa. Näissä teorioissa keskuskognition ongelma on yritetty ratkaista kieltämällä sen olemassaolo.
- [4] Oppimisteoreettinen kirjallisuus on valtava. Konnektionistisia oppimisvaikeuksia käsittelevät esimerkiksi Judd (1990, 1996), Minsky & Papert (1990), Tesauro (1987) ja Tesauro Janssens (1988). Stephen Judd tiivistää tutkimustuloksia seuraavasti: “It is widely acknowledged that as networks get larger and larger and deeper, their learning time grows prohibitively [...] whenever things get scaled up, the news regarding training time is so bad that we can hardly measure it”(Judd, 1996, s. 123). “Symbolisten mallien” kohdalla tilanne on sama.
- [5] Emergenssin (“putkahtamisen”) käsitteelle ei ole olemassa yksikäsitteistä määritelmää. Edellä mainitun lisäksi oletetaan usein, että emergoitunut prosessi antaa systeemistä jollain tavalla ”paremman kuvauksen” (Crutchfield, 1994a, b, Shalizi, 2001, luku 11) jolloin voidaan puolustaa emergenssiä reduktionismin vastateesinä. Jos kokonaisuus on kirjaimellisesti enemmän kuin osiensa summa, on systeemiin täytyntä ilmestyä informaatiota sen ulkopuolelta. Tämä ei ole mielekäs lähtökohhta emergenssin määrittelemiseksi, koska putkahtaneet ominaisuudet eivät näin ollen ole peräisin systeemistä itsestään ja väite, että kokonaisuus on enemmän kuin osiensa summa, on triviaalisti tosi.
- [6] Kolmogorov-kompleksisuuden idea syntyi 60-luvulla, kun kolme tutkijaa G. Chaitin (1969), F. Solomonoff (1964) ja A. Kolmogorov (1965) todistivat sen perusteoreemat toisistaan riippumatta.
- [7] Kolmogorov-kompleksisuuden soveltaminen biologiseen luontoon vaatii lisäoletuksia siitä, miten formalismi tulkitaan fyysikaalisesti ja/tai biologisesti. Näitä oletuksia voisi kuvata ”fyysikaaliseksi Churchin teesiksi”. K-kompleksisuuden sisältyä lisäksi universaalien Turingin koneen valinnasta johtuva ”virhetermi”, jonka olemassaolo vaatii kyseisen tulkinnan luomisen. Ongelmana ei ole niinkään jonkin tietyn tulkinnan antaminen, vaan valitun tulkinnan perusteleminen itse kompleksisuudesta riippumattomalla tavalla. K-kompleksisuus on lisäksi ominaisuus, joka ei itse ole efektiivinen. Nähdäkseni mikäli nämä seikat ovat ongelmallisia sen hypoteesin kannalta, että aivot olisivat aidosti kompleksinen systeemi, ovat ne yhtä ongelmallisia myös sen kannalta, joka väittää niiden olevan yksinkertainen.

## LOPPUVIITTEET

## **WHY COGNITIVE SCIENCE DOES NOT HAVE A THEORY OF CENTRAL PROCESSES?**

*Pauli Brattico, Cognitive Science, Department of Psychology, University of Helsinki, Finland*

Attempts to model the so-called non-modular central processes of the human mind/brain have resulted in the problem of combinatorical explosion. One common feature in all these models has been their low information content. This article explores the contrary hypothesis that the human brain/mind is a genuinely complex system. The notion of genuine complexity is defined with the help of the algorithmic information theory. The linguistic relevance of the hypothesis is discussed.

Keywords: Modularity, complexity, central processes, emergence, learning, innateness