

Karo Salminen

DANIEL C. DENNETTIN INTENTIONAALISUUDEN KÄSITYS

Tietojärjestelmätieteen
pro gradu -tutkielma
30.8.2004

Jyväskylän yliopisto
Tietojenkäsittelytieteiden laitos
Jyväskylä

TIIVISTELMÄ

Salminen, Karo

Daniel C. Dennettin intentionaalisuuden käsitys / Karo Salminen

Jyväskylä: Jyväskylän yliopisto, 2004.

97 s.

Tutkielma

Tässä tutkielmassa tarkastellaan Daniel C. Dennettin käsitystä intentionaalisuudesta. Intentionaalisuudella tarkoitetaan mielen kykyä ilmentää jotakin tai suuntautua johonkin.

Mielenfilosofiassa on puhuttu intentionaalisen puhettavan palautumattomuudesta relaatiologiikkaan eli fysikaaliseen kieleen. Tähän liittyy läheisesti intentionaalista puhetapaa käyttävän arkisen mielestä puhumisen eli arkipsykologian ongelma. Tässä työssä kuvataan Dennettin ehdottama näiden ongelmien ratkaisu. Lähteinä tarkastelulle on käytetty Dennettin tuotantoa ja aihe-alueeseen liittyvää kirjallisuutta.

Dennett on esittänyt ratkaisuksi intentionaalisen asenteen, jonka avulla tarkasteltavia olioita kohdellaan rationaalisina ja intentionaalisina systeemeinä. Dennettin mukaan intentionaalisuus liitetään olioon eli systeemiin attribuution avulla. Dennettin näkemyksen mukaan olio on intentionaalinen, jos sen käyttäytymistä voidaan ennustaa intentionaalisena systeeminä. Näkemys laajentaa intentionaalisuuden myös muihin olentoihin kuin ihmisiin, muun muassa eläimiin ja tietokoneisiin. Dennettin näkemyksen epäintuitiivisuus ja näennäinen epäjohdonmukaisuus on johtanut epäilyyn Dennettin intentionaalisuuskäsityksen mielekkyydestä.

AVAINSANAT: intentionaalisuus, intentionaalinen asenne, arkipsykologia, mielenfilosofia, kognitiotiede

SISÄLLYSLUETTELO

1 JOHDANTO	4
2 LÄHTÖKOHTIA DENNETTIN INTENTIONAALISUUDEN KÄSITYKSELLE	9
2.1 Intentionaalisuus käsitteenä.....	9
2.2 Intentionaalisuuden ongelma	12
2.3 Universaali darwinismi Dennettin presuppositioina.....	14
3 DENNETTIN INTENTIONAALINEN ASENNE.....	17
3.1 Kolme strategiaa olioiden toiminnan tarkastelulle.....	17
3.1.1 Strategioiden väliset suhteet.....	20
3.1.2 Ihmisten arkinen ilmiöiden hahmottaminen	23
3.2 Intentionaalisen systeemin käsite ja intentionaalinen strategia.....	24
3.2.1 Intentionaalisuuden tulkinnallisuus	29
3.2.2 Intentionaalisen strategian epätäydellisyys ja voima	33
3.3 Intentionaalisten systeemien tutkiminen.....	37
3.3.1 Intentionaalisten systeemien teoria.....	38
3.3.2 Alipersoonallinen kognitiivinen psykologia ja dekompositio- menetelmä	41
3.3.3 Semantiikan luominen syntaksista.....	42
3.3.4 Reduktion mahdollisuus	45
4 DENNETTIN NÄKEMYKSEN ERITYISKYSYMYKSET.....	50
4.1 Uskomuksen ja mielipiteen erottelu	50
4.2 Rationaalisuusoletus	53
4.3 Kritiikki kolmannen persoonan mukaisesta lähestymistavasta	60
4.4 Kiista alkuperäisestä ja johdetusta intentionaalisuudesta.....	61
4.4.1 Virvoitusjuoma-automaatin intentionaalisuus eri ympäristöissä.....	63
4.4.2 Intentionaalisuus Rinnakkaismaassa	65
4.4.3 Selviytymiskoneen intentionaalisuus.....	66
4.5 Intentionaalisten tilojen ja funktioiden määrittämättömyys.....	68
4.6 Intentionaalisten tilojen ontologinen asema.....	71
4.6.1 Kuvion olemassaolo.....	74
4.6.2 Epifenomenalismmin vaara	77
4.6.3 Suhde pragmatismiin.....	81
5 YHTEENVETO	83
LÄHDELUETTELO.....	88

1 JOHDANTO

Tämä tutkielma tarkastelee Daniel C. Dennettin intentionaalisuuden käsitystä. Niin kuin intentionaalisuus (engl. intentionality) on mielenfilosofian kiistellyimpiä alueita, on myös Dennett mielenfilosofian kiistellyimpiä ja arvostetuimpia keskushahmoja.

Intentionaalisuuden viitekehyksenä toimii traditionaalinen mieli-keho -ongelma (engl. mind-body problem) tai ongelma mentaalisen ja fyysisen (materiaalisen) suhteesta. Kaikki ovat yhtä mieltä siitä, että esimerkiksi kivenlohkareella on sellaisia fysikaalisia ominaisuuksia kuin paino, paikka, väri ja ajanhetki; toisin sanoen, että se on spatiotemporaalinen. Näin ei ole kuitenkaan laita esimerkiksi ajatuksen, että ”kivenlohkare on raskas”. On mieletöntä väittää, että tällaisella ajatuksella olisi tietty paino, paikka ja muita fysikaalisia ominaisuuksia. Juuri siksi ajatuksia nimitetäänkin mentaaliseksi erotuksena fysikaalisiin entiteetteihin. Ongelma syntyy siitä, miten aivot materiaalisena kappaleena voivat tuottaa jotakin mentaalista, kuten ajatuksia. Jos mentaalinen ja fysikaalinen ovat erilaisia substansseja, miten materiaalinen voi vaikuttaa mentaaliseen?¹

Intentionaalisuudessa on kysymys mentaalisen mielen ja materiaalisen maailman välisestä suhteesta eli relaation ongelmasta. Intentionaalisuus on mielen suuntautumista maailmaan, suuntautumista johonkin muuhun kuin itseensä; ajattelulla on olemassa aina jokin kohde, ja ajattelun ja kohteen välillä on jonkinlainen relatio. Ajatuksella on siis jokin sisältö (engl. content), ja tuo sisältö on ajatuksen kohteen (tieto)edustus. Kun mieli on suuntautunut johonkin olioon, kutsutaan kognitiotieteessä mielen edustusta kohteestaan

¹ Ks. yleisesti mieli-keho-ongelmasta esim. Bunge 1980. Ks. yleisesti (Dennettin näkökulmasta) kognitiotieteellistyneestä mielenfilosofiasta Dennett 1978b.

representaatioksi.² Erityisesti kielestä ja tietokoneista puhuttaessa intentionaalisuudelle yhtäläinen käsite on semantiikka (engl. semantics).³ Samoin kuin intentionaalisuus, semantiikka viittaa aina itsensä ulkopuolelle – maailmaan –, koska semantiikka vaatii tulkitsemisprosessin, jossa merkit liitetään merkkien ulkopuolisiin merkityksiin. Intentionaalisuus on 1800-luvulla eläneen Franz Brentanon (1973) irredusibiliteettiteesin mukaan ainoastaan ominaislaatuista mielelle, mentaalille. Intentionaalinen relaatio (tai semantiikka) ei ole sen luonteen vuoksi palautettavissa normaaliin relaatiologiikkaan (tai syntaksiin [engl. syntax]). Palautumattomuudesta on syntynyt ongelma siitä, miten ihmisten toimintaa voidaan selittää, jos pitäydytään tieteen naturalistisessa sitoumuksessa, joka pohjautuu relaatiologiikan ja fysiikan kaikkivoipaisuuteen. Tekoälyn puolella puhutaan usein vahvan tekoälyn (engl. strong artificial intelligence; Searle 1980) mahdollisuudesta – voidaanko sanoa tietokoneen todella ymmärtävän sen käsittelemää tietoaan – eli loppujen lopuksi siitä, voidaanko syntaksista luoda semantiikkaa.

Ihmisten toiminnan selittämisessä viitataan ihmisten ajatuksiin, uskomuksiin, haluihin ja muihin mentaalisiin tiloihin. Sanomme esimerkiksi henkilön ottavan sateenvarjon mukaansa ulosmenessään, koska hän uskoo, että ulkona sataa. Tällaista mielentiloihin viittaavaa selitystapaa käytämme arkisissa toiminnan kuvauksissamme, ja näin tekevät myös sekä sosiaaliset että psykologiset tieteet.

² Representaation käsite ja luonne ovat kriittisiä kognitiotieteen kannalta, minkä takia representaatiosta ollaan montaa mieltä eri yhteyksissä. Myös tämä työ tulee sisältämään Dennettin intentionaalisuuden käsityksen johdosta erilaisen representaatio-käsityksen, jota moni ei hyväksy.

³ Koska intentionaalisuus on tämän tutkielman fokus – tarkemmin sanottuna Dennettin intentionaalisuuden käsitys – en semantiikkaa tule laajemmin käsittelemään. John Haugeland on kirjoittanut hyvän johdannon semantiikkaan kognitiotieteen näkökulmasta artikkelissaan Haugeland 1981.

Yleisesti ottaen ratkaisuna palautumattomuudelle on pidetty kognitiotieteessä funktionalismia⁴ (engl. functionalism), joka sisällyttää intentionaalisuuden ja representaatioiden sisällöt toiminnallisuuteen ja sen erityistä muotoa, konefunktionalismia (engl. machine functionalism; esim. Putnam 1967). Konefunktionalismia vastaan on puhunut erityisesti Dennett, jonka näkemys on ehkä huomattavin, vallankumouksellisin ja omaperäisin (yksittäinen) kontribuutio intentionaalisuus-keskusteluun.⁵

Daniel C. Dennett⁶ on saanut paljon vaikutteita Ludwig Wittgensteinilta, Williard van Orman Quinelta ja Gilbert Rylelta. Samoin kuin nämä suurmiehet, Dennett on tuonut filosofian kenttään uutta: sen lisäksi, että hän on pyrkinyt syrjäyttämään konefunktionalismin ja mielen kielen (engl. language of thought; Fodor 1975) mukaisen representaatioiden tulkitsemistavan, hän on tuonut mielenfilosofiaan mukaan evolutiivisen näkökulman.⁷ Hän on tutkinut pääosin intentionaalisuutta (tai mielen ja representaatioiden sisältöä) ja tietoisuutta. Jo väitöskirjassaan *Content and Consciousness* (1969) Dennett jakoi mielenfilosofian tutkimuksensa kahteen osaan, intentionaalisuuteen ja tietoisuuteen (engl. consciousness). Näin siksi, että hän pitää intentionaalisuutta edellytyksenä tietoisuudelle. Tämänkaltaista jaottelua intentionaalisuuden ja tietoisuuden

⁴ Ks. funktionalismista yleensä ja sen eri muodoista esim. Block 1980. Ks. yleisesti funktionalismin ongelmista esim. Fodor & Block 1972 ja Block 1978.

⁵ Toisinaan Dennettin näkemystä on nimitetty homonkulaariseksi funktionalismiksi (engl. homoncular functionalism; Lycan 1981)

⁶ Daniel Clement Dennett syntyi vuonna 1942. Hän toimii tällä hetkellä muun muassa filosofian professorina Tufts'n yliopistossa Yhdysvalloissa. Hän on kirjoittanut lukuisia kirjoja mielenfilosofian saralta. Hänen kirjojaan ovat (alkuperäisnimiltään ja ilmestymisvuosiltaan): *Content and Consciousness* (1969), *Brainstorms: Philosophical Essays on Mind and Psychology* (1978), *The Mind's I: Fantasies and Reflections on Self and Soul* (1981, toimitettu yhdessä Douglas Hofstadterin kanssa), *Elbow Room: The Varieties of Free Will Worth Wanting* (1984), *The Intentional Stance* (1987), *Consciousness Explained* (1991), *Darwin's Dangerous Idea: Evolution and the Meanings of Life* (1995), *Kinds of Minds* (1996), *Brainchildren: Essays on Designing Minds* (1998) ja *Freedom Evolves* (2003). Hän on vaikuttanut myös tekoälyn ja kognitiivisen etologian tutkimukseen. Yksityiskohtaisempi ja laaja-alaisempi kuvaus Dennettistä ja hänen työstään löytyy artikkelista Brook & Ross 2002b.

⁷ Myöhemmin hän onkin eksplisiittisesti (esim. Dennett 1995a) siirtynyt linjalle, jota voisi kutsua post-darwinistilaiseksi filosofiaksi (Kilpinen 2003, 234 alaviite).

välillä moni muu mielenfilosofi ei tee.⁸ Väitöskirjansa jälkeen hän keskittyikin ensin intentionaalisuuteen ja kirjoitti myöhemmin laajasti huomiota herättäneen populaarin kirjan tietoisuudesta (*Tietoisuuden selitys* [1999]).

Yleisesti mieleen voidaan ajatella olevan kaksi mahdollista näkökulmaa: relationaalinen eli intentionaalinen ja fenomenaalinen (engl. *phenomenal*). Näistä fenomenaalinen näkökulma – joka usein rinnastetaan tietoisuuden kanssa – käsittelee sitä, millaista on olla tietyssä mielentilassa (esim. kivun tilassa), eli fenomenaalinen näkökulma tutkii kvalioita (engl. *qualia*). Tässä tutkielmassa tarkastellaan vain ensimmäistä mielen näkökulmaa, ja fenomenaalinen näkökulma jätetään pois tarkastelun ulkopuolelle.⁹ Näin siksi, että keskittyminen intentionaalisuus muodostaa pohjan myös esimerkiksi Dennettin käsitykselle tietoisuudesta.¹⁰

Tämän tutkielman pääkysymys on: ”mikä ja minkälainen on Dennettin käsitys intentionaalisuudesta?” Tätä kysymystä tarkennetaan kysymällä edelleen, miten Dennett sitoo yhteen fysiikan ja intentionaalisen kielen. Dennettin näkemystä kuvataan mahdollisimman tarkkaan ja annetaan perusteluja hänen usein intuition vastaiselle intentionaalisuuden käsitykselleen. Kuvauksen tarkkuudesta ja aiheen laajuudesta johtuen osa Dennettiin kohdistuneesta kritiikistä jätetään pois tai ainoastaan lyhyen käsittelyn varaan, mutta joitain hyvin olennaisia ja keskeisiä erityiskysymyksiä käsitellään kyllä. Olisi hedelmällistä verrata Dennettin intentionaalisuuden käsitystä muihin

⁸ Ks. esim. Siewert 2003, luku 9 ”Four Views of the Consciousness-Intentionality Relationship”.

⁹ Dennett on esittänyt näkemyksensä kvalioista esimerkiksi kirjassaan Dennett 1999 ja artikkelissaan Dennett 1988b.

¹⁰ Sanottakoon, että Dennettin mielestä intentionaalisuus on riittävä mentaalisuuden ehto. Tätä ei moni muu hyväksy (kuten John Searle), vaan mentaalisuudelta vaaditaan myös tietoisuutta. Ks. esim. Pierre 2003. Ks. myös 4.3 ”Kritiikki kolmannen persoonan mukaisesta lähestymistavasta”.

keskustelua herättäneisiin käsityksiin, mutta sekin joudutaan jättämään tarkastelun ulkopuolelle.¹¹

Vastauksien saamiseksi kysymyksiin ja Dennettin näkemyksen kuvaamiseksi tutkielma on jaettu kolmeen osaan. Ensimmäisessä osassa tarkastellaan intentionaalisuuden ja intentionaalisen kielen erityispiirteitä ja sen palautumattomuutta fysikaaliseen kieleen. Tarkoituksena on muodostaa käsitys intentionaalisuuden ja arkisen toiminnan selittämisen intentionaalisen puheen viitekehyksestä, jotta lähtökohdat olisivat otollisia Dennettin näkemyksen tarkastelemiseen. Tämän jälkeen siirrytään lyhyen Dennettin evoluutioteoreettisen taustan esittelyn lävitse itse Dennettin intentionaalisuuden näkemyksen pariin tutkielman toiseen osaan. Lopuksi kolmannessa osassa käsitellään tarkemmin Dennettin näkemykseen kuuluvia erityiskysymyksiä, joiden käsittely vaatii lähempää tarkastelua.

Aiheesta on aiemmin kirjoitettu yksi suomenkielinen pro gradu -tutkielma (Huttunen 1996), mutta kyseinen kirjoitus ei mene mielestäni tarpeeksi syvälle Dennettin näkemyksen käsittelyyn. Huttusen rajaus on myös erilainen kuin omani.¹² Annan tässä tutkielmassa painoa Dennettin intentionaalisuuden käsityksen laaja-alaisuudelle, ja käsittelen sitä mahdollisimman seikkaperäisesti. Päälähteinä käytän Dennettin teoksia *The Intentional Stance* (1987) ja *Brainstorms: Philosophical Essays on Mind and Psychology* (1978a).¹³

¹¹ Muilla käsityksillä tarkoitan erityisesti Jerry Fodorin, Paul ja Patricia Churchlandin, Richard Rortyn, Donald Davidsonin, John Searlen, Ruth Millikanin ja Fred Dretsken intentionaalisuuden ja intentionaalisten tilojen käsityksiä. Donald Davidsonin ja Ruth Millikanin käsitykset ovat hyvin lähellä Dennettin käsitystä, joten niiden vertailu Dennettin käsityksen kanssa olisi hyvin avartavaa. Jossain määrin tulen kuitenkin sivuamaan lyhyesti muun muassa Fodorin ja Churchlandien näkemyksiä tutkielman edetessä.

¹² Suurin puute Huttusen tutkielmassa mielestäni on sen rajauksen häilyvyys. Huttunen eksyy helposti laajemmin Dennettin ajatusten ja niiden kritiikkien syövereihin eikä pysy rajauksensa sisällä. Lisäksi ontologia-keskustelu on hänellä pintapuolinen. Kuitenkin yleisesti ottaen hänen työnsä on mielestäni erinomainen suomenkielinen katsaus Dennettin ajatteluun yleensä ja sen kritiikkiin.

¹³ Kumpikin näistä kirjoista on kokoelma artikkeleita, mutta joitakin artikkeleja on paranneltu. *The Intentional Stancen* artikkelien perään on lisäksi liitetty ”Reflections”-osuudet, jossa

2 LÄHTÖKOHTIA DENNETTIN INTENTIONAALISUUDEN KÄSITYKSELLE

Johdannossa tuotiin esille tutkielman aihealueen sijainti mielenfilosofisessa keskustelussa ja esiteltiin lyhyesti, mitä intentionaalisuudella tarkoitetaan. Tässä luvussa jatketaan aiheesta tarkentaen intentionaalisuus-käsitteen taustaa (luku 2.1 ”Intentionaalisuus käsitteenä”), sen ongelmallisuutta ja merkityksellisyyttä (luku 2.2 ”Intentionaalisuuden ongelma”). Lopuksi, ennen siirtymistä Dennettin käsitykseen intentionaalisuudesta, ja sen ratkaisusta luodaan silmäys hänen argumentointinsa taustalla vaikuttavaan evolutiiviseen ajatteluunsa (luku 2.3 ”Universaali darwinismi Dennettin presuppositiona”).

2.1 Intentionaalisuus käsitteenä

Käsitteellä intentionaalisuus on pitkä perinne filosofiassa, joka juontuu aina skolastikkoihin saakka. 1800-luvulla elänyt Franz Brentanon elvytti käsitteen uudelleen käyttöön pyrkien osoittamaan irredusibiliteettiesillään mentaalisen ja fyysisen yhtäläistämisen mahdottomuuden. Irredusibiliteettiesiä ei sinällään enää nykyään hyväksytä sen ongelmien takia, mutta käsite on jäänyt elämään sekä analyyttisessä että mannermaisessa traditiossa.¹⁴ Seuraavassa mukailen lähteitä Dennett & Haugeland 1987, Pierre 2003 ja Siewert 2003 intentionaalisuus-käsitteen valaisemiseksi.

”Intentionaalisuus” ymmärretään filosofisessa keskustelussa niin, että *intentionaalisuus* on joidenkin mentaalisten tilojen eli mielentilojen – ideoiden, ajatusten, uskomusten¹⁵, halujen ja niin edelleen – erityistä kykyä ilmentää

Dennett tarkastelee artikkeleissa esitettyjä asioita vuosia niiden julkaisemisen jälkeen. Viittaan siis suoraan tässä tutkielmassa näihin teoksiin, enkä niiden artikkeleihin.

¹⁴ Intentionaalisuus on saanut erilaisen sijan analyyttisessä ja mannermaisessa traditiossa. Analyyttisessä traditiossa sitä on tarkasteltu modernin logiikan ja semantiikan valossa, mannermaisessa traditiossa taas on jatkettu Brentanon esityksen hyödyntämistä fenomenologiassa. Tässä seuran analyyttisen tradition esitystä. Ks. tarkemmin intentionaalisuus-käsitteen historiasta esim. Pierre 2003.

¹⁵ Filosofisessa termistössä uskomus-käsitteellä on erilainen sija kuin arkitermistössämme. Filosofisessa keskustelussa uskomuksella viitataan tietämistä löyhempään tiedon

jotakin (engl. aboutness) tai kykyä suuntautua objekteihin ja tapahtumiin (engl. directedness).¹⁶ Voin esimerkiksi ajatella aurinkoa, jolloin ajatukseni on suuntautunut (tuohon) aurinkoon (tai pikemminkin Aurinkoon). Intentionaalisille mielentiloille on ominaista, että niillä on aina kohde: en voi uskoa uskomatta johonkin, enkä haluta haluamatta jotakin. Filosofian teknisellä intentionaalisuus-termillä ei siis tarkoiteta englanninkielen ”intentiota” tai ”toiminnan tarkoituksellisuutta”.¹⁷

Mentaalisen tilan ja sen kohteen välillä on siis suhde eli *relatio*, mutta se on hyvin kummallinen. Se eroaa tavallisesta relaatiosta joillakin tavoilla. Ensiksi, tavallisen relaation yksilöt ovat samoja, olivat ne missä suhteessa toisiinsa hyvänsä. Yksilö A on sama yksilö, oli hän sitten B:n ystävä tai oli B A:n ystävä. Intentionaalisen relaation tapauksessa näin ei ole. Tietty uskomus ei voi olla yhdellä hetkellä auringosta (että se on häikäisevä) ja toisella hetkellä lamppuvalosta (että se on häikäisevä). Sille, mistä uskomuksen on määrä olla uskomus, on ratkaisevaa kyseiselle uskomukselle. Tavallisen relaation mukaan olisi mieletöntä sanoa, että ”[t]ämä tuoli ei ole sama, jonka näit täällä eilen, mutta se on täsmälleen samanlainen” (Wittgenstein 1999, § 253). Toiseksi, tavallisten relaatioiden objektien täytyy olla olemassa – tai niiden on täytynyt olla olemassa. Intentionaaliset relaatiot eivät tässäkään tapauksessa mukaudu tavallisen relaation muottiin. Voin uskoa esimerkiksi Suomen kuninkaaseen, vaikka Suomella ei mitään ilmeisimmin ole, eikä ole ollut, kuningasta. Näistä ja

muotoon. Jotta joku voisi tietää, että p, täytyy myös p:n olla tosi. Epistemologisista syistä sanotaan mieluummin jonkun uskovan, että p, sillä uskomisen ei edellytä p:n totuutta. Tämän tutkielman kannalta merkitystä on vain sillä, että uskomukset, halut jne. ovat intentionaalisia (mielen)tiloja.

¹⁶ Kuten John Searle (1983, 1) huomauttaa, kaikki mielentilat eivät ilmennä intentionaalisuutta. Esimerkiksi ahdistuksella ei yleensä sanota olevan mitään kohdetta, johon se olisi suuntautunut. Tämä ei kuitenkaan ole tärkeä erottelu tämän työn kannalta. Sekoitan vapaasti keskenään tässä työssä yleiset mielentilat ja intentionaalisuutta ilmentävät mielentilat.

¹⁷ Intentio ja toiminnan tarkoituksellisuus on kuitenkin läheisessä suhteessa intentionaalisuuteen, koska intention ilmaisemisessa viitataan intentionaalisuutta ilmaiseviin idiomeihin (ks. alla). Tästä syystä johtuen nykyään intentionaalisuutta ja toiminnan tarkoituksellisuutta ei juuri erotella toisistaan, enkä itsekään tee tässä tutkielmassa tarkkaa eroa.

muista tässä erittelemättömistä syistä normaali relaatio ei riitä intentionaalisen relaation kuvaamiseen.¹⁸

Niin kutsutun semanttisen nousun (engl. semantic ascent; Quine 1960) johdosta on siirrytty intentionaalisten tilojen itsensä tarkastelemisesta sellaisten lauseiden tarkastelemiseen, jotka ilmentävät intentionaalisuutta (ks. myös Chisholm 1957). Näin vältetään muun muassa epistemologiset häiriötekijät, kuten se, miten voidaan tietää toisen henkilön mentaalinen tila. On siirrytty niin sanottujen *intentionaalisten idiomien* – ”... uskoo, että p”, ”... haluaa, että q” ja niin edelleen – tarkastelemiseen. Intentionaaliset idiomit ilmentävät referentiaalista läpinäkymättömyyttä (engl. referential opacity; Quine 1960).

Lauseen kahden termin sanotaan olevan *referentiaalisesti läpinäkymättömiä*, jos niitä ei voida korvata toisillaan muuttamatta lauseen totuusarvoa, vaikka ne viittaisivat samaan asiaan tai samoihin asioihin. Ajatellaan seuraavaa esimerkkiä. On niin, että Iltatähti ja Aamutähti viittaavat samaan planeetta Venukseen. Kuitenkin, vaikka

1. Otto pohtii, onko Iltatähti Venus,

on epätodennäköistä, että

2. Otto pohtii, onko Venus Venus.

Samaten:

3. Sanna uskoo, että oven takana oleva henkilö on sarjamurhaaja.

Vaikka oven takana oleva henkilö olisi hänen isänsä, emme sanoisi, että

4. Sanna uskoo, että hänen isänsä on sarjamurhaaja.

Näiden lauseiden *ekstension* eli alan sanotaan olevan sama (oven takana oleva henkilö on Sannan isä), jos kohta niiden *intensio* eli merkitys tai sisältö on eri.

¹⁸ Ks. tarkemmin intentionaalisuuden ongelmista esim. Dennett & Haugeland 1987 ja Pierre 2003.

Kun puhutaan lauseen merkityksestä, puhutaan sen *semantiikasta*. Vaikka lauseilla on jokin merkitys tai sisältö, lauseilla itsellään – meistä riippumatta – on vain formaalisia tai *syntaktisia* ominaisuuksia, esimerkiksi tietty sanajärjestys. Semantiikka intentionaalisuuden tavoin viittaa itsensä ulkopuolelle. Tämän yhtäläisyyden vuoksi tulen jatkossa käyttämään molempia toisiaan korvaavina käsitteinä, koska kyseessä on perimmältään samasta asiasta.

2.2 Intentionaalisuuden ongelma

Arkisessa puheessamme mielestä eli *arkipsykologiassa* (engl. folk psychology) ja siihen pohjautuvissa intention tieteissä käytetään alituisen intentionaalisia idiomeja.¹⁹ Selitämme esimerkiksi henkilön avanneen sateenvarjon sateen alettua, koska hän halusi pysyä kuivana ja koska hän uskoi, että sateenvarjoa käyttämällä hän voi välttyä kastumiselta. Yleisemmin sanottuna käytämme tiedostamattamme jotakin praktisen syllogismin tapaista päättelymuotoa ihmisten toiminnan selittämisessä ja ymmärtämisessä. *Praktinen syllogismi* on yleiseltä muodoltaan seuraavankaltainen:²⁰

1. Henkilö haluaa saavuttaa päämäärän p.
2. Hän uskoo, että ellei hän tee w:tä, hän ei voi saada aikaan p:tä.
3. Siis, henkilö ryhtyy tekemään w:tä.

Tässä ensimmäinen ja toinen lause ovat pmissettä, ja kolmas lause on johtopäätös. Johtopäätös eli toiminta seuraa siis kausaalisesti pmissettä, joiden mukaan henkilöllä on haluja ja uskomuksia sekä muita mielentiloja.

¹⁹ Ks. arkipsykologiasta esim. Nichols 2002. Arkipsykologia kuuluu laajempaan kansankulttuurillisesta esiymmärrykseen (ks. Sellars 1963), johon kuuluvat mm. arkiset käsitykset fysiikasta (ks. esim. Hayes 1979).

²⁰ Ks. esim. von Wright 1971. Praktinen syllogismi on Aristoteleelta lähtöisin ihmisen toiminnan ymmärtämiseksi tarkoitettu selitysmalli. Intention tieteissä on ollut paljon tieteenfilosofista kiistaa siitä, mitä selitysmallia niissä itse asiassa käytetään ihmisen toiminnan selittämiseksi. Tässä nojataan kuitenkin yksinkertaisuuden vuoksi praktiseen syllogismiin ja sen tavanomaiseen muotoon.

Intentionaalisuuden ongelma on siinä, että intentionaalista puhetta ei voida esittää standardissa relaatiologiikassa, jota laajalti käytetään fysikaalisissa tieteissä. Syynä on, että relaatiologiikka ei pysty käsittelemään lauseita semantiikan varassa, vaan ainoastaan niiden syntaktisten ominaisuuksien perusteella. Toisin sanoen relaatiologiikka on ainoastaan ekstensionaalinen kieli. Sosiaalitieteiden ja muiden intentionaalisia idiomeja käyttävien tieteiden kieltä ei siis voida palauttaa ekstensionaaliseen kieleen. Siten vaihtoehtoina Quinen (1960, 221) mukaan on, joko (1) tyytyä loogiseen inkoherenssiin, eli hyväksyä täysin intentionaalisen psykologian tai fenomenologian sekä intentionaalisen puheen autonomisuus, tai (2) kerta kaikkiaan hylätä niiden autonomisuus ja intentionaalinen puhe. Ensimmäistä vaihtoehtoa pidetään yleensä poissuljettuna, sillä se johtaa tieteen epäyhtenäisyyteen. Toinen vaihtoehto ei sekään tunnu järkevältä: intentionaalisia idiomeja käyttävät selitykset ovat hyvin toimivia, ja lisäksi voidaan sanoa, että behavioristiset tai fysiologiset teoriat ovat epäonnistuneet pyrkiessään selittämään ja ennustamaan ihmisten ja muiden eläinten älykästä käyttäytymistä.

Kun katsomme ympärillemme näyttää siltä, että intentionaalisuuden ongelmaan on kuin onkin olemassa ratkaisu. Tietokoneet nimittäin näyttävät olevan intentionaalisia olioita, ne tuntuvat muun muassa ”ymmärtävän” ohjeita ja tallentavan tietoa jostakin. Niiden toimintaa voidaan arkisesti selittää intentionaalisin idiomein, ja silti niiden sisäiset tilat ja tapahtumat voidaan kuvata puhtaasti ekstensionaalisella kielellä. (Dennett 1969, 40.) Näyttää siltä, että siten oikea silta intentionaalisen ja ekstensionaalisen kielen välillä on olemassa. On tutkittava millainen tämä silta on, miten se on rakennettu ja luotava yleinen sillanrakennusmalli. On ratkaistava, miten intentionaalinen kuvaus liitetään ekstensionaaliseen kuvaukseen.

Intention tieteet eivät pysty tässä sillan selvitys- ja rakennusurakassa auttamaan, sillä ne edellyttävät intentionaalisuuden eivätkä siten kuvaile edellytyksiään: intention tieteet pohjautuvat arkipsykologiaan, ja myöskään

arkipsykologia ei pysty vastaamaan kysymykseen ekstensionaalisen ja intentionaalisen kuvauksen suhteesta. Sitä varten tarvitaan tieteellisimpiä esiteoreettisia kuvauksia. Konefunktionalismin (esim. Putnam 1967; myös Fodor 1968 ja 1975, vaikka hän ei suoranaisesti konefunktionalismia kannatakaan) kannattajat ovat pyrkineet vastaamaan haasteeseen, ja niin on myös Dennett. Dennett ei kuitenkaan hyväksy konefunktionalismin ehdotusta ja onkin siksi kehittänyt omaa näkemystään. Siirrymme hiljalleen Dennettin ratkaisun pariin.

2.3 Universaali darwinismi Dennettin presuppositioina

On hyödyllisempää aloittaa tuomalla julki Dennettin ajattelu- ja argumentointitapojen taustalla vaikuttavia oletuksia eli presuppositioita, kuin syöksyä suin päin hänen näkemyksensä kimppuun. Dennettin taustaoletuksissa voidaan nähdä kaksi läpäisevää teemaa: varovainen ja epäilevä suhtautuminen asioista (kuten mielestä) puhumiseen ja arkikäsitteistöön sekä evoluutioteoreettisen ajattelun keskeisyys. Näistä ensimmäinen on tuttua jo Wittgensteinilta (1999) ja Rylelta (1949), mutta evoluutioteoreettinen ajattelu on täysin omintakeista juuri Dennetille.

Dennett (1995a) tahtoo laajentaa evoluution sovelluskenttää ja näyttää, miten hedelmällinen ja vakavasti otettava evoluutioajattelu on myös biologian ulkopuolella: esimerkiksi psykologiassa, etiikassa, sosiaalitieteissä, kosmologiassa ja filosofiassa.²¹ Hänen (1995a) mielestään Charles Darwinin (1988) alun perin lajien syntyyn ja kehitykseen tarkoitettua varmallalla pohjalla olevaa ja ihmeisiin vetoamatonta selitysmallia tulisi käyttää yleisenä

²¹ Itse asiassa Dennett on itse samaten kulttuurievoluutioon pohjautuvan memetiikan (engl. memetics) puolestapuhuja. Memetiikassa lainataan biologian selitysmallia, jossa kulttuuriympäristö nähdään erilaisten ideoiden ynnä muiden meemien tai "kulttuurigeenien" runsastumis- ja karsiutumiskenttänä (ks. esim. Aunger 2000b). Meemin idea on lähtöisin Richard Dawkinsin (1993) universaalien kopioitujan ideasta. Ks. memetiikasta esim. Aunger 2000a.

viitekehyksenä kaikelle selittämislle.²² Hänen näkemystään voisi siten luonnehtia *universaaliksi darwinismiksi*.

Darwin osoitti Dennettin (1995a) mukaan, miten ”kaaoksesta” voidaan luoda ”järjestystä” ja miten luonnossa esiintyvä ”järjestys” voidaan selittää vetoamalla minkäänlaiseen tietoiseen suunnitteliin (yliluonnolliseen Jumalaan). Evoluutioprosessia voidaan nimittää dawkinslaisittain ”sokeaksi kellosepäksi” (Dawkins 1989), joka on vailla tarkoitusta, päämäärää ja kaukonäköä. Koska olemme osa luontoa, luontokappaleita siinä missä muutkin biosfäärin eläimet, myös me olemme tämän sokean mekanismin tuotteita. Sen takia ihmisyyttä sinällään ja ihmisen ominaisuuksia on tarpeetonta mystifioida. Luonnossa esiintyy erilaisten ominaisuuksien jatkumo, eikä voida siten määrittellä, mitkä ovat – esimerkiksi lintujen siipien, koiran tai kielen – diskreetit rajat sekä välttämättömät ja riittävät ehdot sille, että jotkin ominaisuudet todella ovat aitoja ominaisuuksia. Mitään eksplisiittistä, oikeaa rajaa ei yksinkertaisesti ole:

Me *haluamme* vetää rajoja; meidän usein *täytyy* vetää rajoja [...] Darwin näyttää meille, että evoluutio ei tarvitse sitä, mitä me tarvitsemme; oikea maailma tulee hyvin toimeen *de facto* –poikkeavuuksien kanssa, jotka ilmaantuvat ajan kuluessa, jättämällä paljon tyhjää tilaa aktualiteettiryhmittymien [engl. clusters of actuality] välille. (Dennett 1995a, 202.)

Tämä tarkoittaa *essentialismin* – jonka mukaan on olemassa välttämättömiä ja riittäviä ehtoja sille, että ominaisuus on jokin tietty ja aito ominaisuus – hylkäämistä. Ajatus on lähtöisin alkujaan Wittgensteinin myöhäisestä tuotannosta (1999), jossa hän näki essentialismin kestävämmäksi.²³ Essentialismin kannattaminen on silti yleistä erilaisissa valeasuissa. ”Filosofin oireyhtymänä” Dennett näkeekin erehtymistä pitämään ”kuvittelun

²² Dennettin (1995a) omin termein voisi sanoa, että universaalissa darwinismissa on kyse ”nostokurkiin” (engl. crane) nojautumisesta eikä kuviteltuihin ”taivaskoukkuihin” (engl. skyhook) pitäytymisestä ”rakennettaessa” selitystä jollekin ilmiölle.

²³ Wittgenstein käytti itse tunnetusti esimerkkiä pelin määrittämisestä essentialismin akilleenkantapäänä.

epäonnistumista välttämättömyyden oivaltamisena” (Dennett 1999, 446).²⁴ (Antiessentialismi johtaa myös Quinen [1951] tavoin analyttisen ja synteettisen totuuden välisen eron hylkäämiseen sekä naturalistisen epistemologian [ks. esim. Kornblith 1994] hyväksymiseen.) Täten on selvää, että Dennett ei näe ihmisten perustavanlaatuisina pidettyjä ominaisuuksia, kuten intentionaalisuutta, täysin ihmisille ominaisina ominaisuuksina. Hänen tunnettu tapansa on lähteä tietystä ilmiöstä käsin ja pohtia, miten se on voinut syntyä ja kehittyä evolutiivisesti. Toisin sanoen hän käyttää menetelmää, jota hän nimittää takaperoiseksi suunnitteluksi (engl. reverse engineering; ks. esim. Dennett 1995a ja Dennett 1999).

Dennettin evoluutionäkemyksessä pohjautuu *adaptationismiin*, ohjelmaan, jonka mukaan evoluution organismit ovat eräässä mielessä optimaalisia. Tai adaptationismin kriitikoiden Stephen Jay Gouldin ja Richard Lewontinin (1979, 581) sanoin ohjelmaan, joka ”perustuu uskoon luonnonvalinnan voimaan optimoivana agenttina”. On huomattava, että optimaalisuusoletusta ei tehdä siksi, että oletetaan evoluution luoneen parhaan mahdollisen maailman. Optimaalisuusoletusta tarvitaan evoluutiota tulkittaessa; optimaalisuusoletuksella pyritään ymmärtämään luonnonvalinnan voimia, jotka ovat muokanneet organismin käyttäytymistä sellaiseksi kuin se on. (Dennett 1987, 279.) Adaptationismista on kiistelty paljon biologian filosofiassa, mitä luultavimmin pikemminkin poliittisista kuin tieteellisistä syistä (Seegerstråle 2000; erityisen tunnettu on Dawkinsin ja Stephen Jay Gouldin välinen kiista), mutta tämän tutkielman tiimoilta asiaan ei paneuduta.²⁵

²⁴ Tämän takia Dennett ei itse esitäkään selkeitä ja tarkkarajaisia määritelmiä, joita puolustaisi, kuten tullaan näkemään.

²⁵ Ks. erityisesti Dennett 1983 ja sen kriittinen kommentaari. Ks. myös Dennett 1995a ja teoksen Dennett 1987 artikkelin Dennett 1983 Reflections-osuus.

3 DENNETTIN INTENTIONAALINEN ASENNE²⁶

Ennen kuin päästään käsiksi varsinaiseen Dennettin ehdottamaan intentionaalisen ja ekstensionaalisen kielen välisen suhteen kuvaukseen, on aluksi jätettävä metafysiikka ja filosofiset intentionaalisuuden analyysit sikseen. Seuraavaksi käydään läpi intentionaalinen strategia ja siihen liittyvät muut strategiat sekä niiden oletukset, koska ne luovat pohjan kielten välisen suhteen ymmärtämiselle Dennettin näkemyksessä.

Ensin (luvussa 3.1 ”Kolme strategiaa olioiden toiminnan tarkastelulle”) käsitellään Dennettin viitekehys olioiden toiminnan selittämiseksi ja ennustamiseksi. Sen jälkeen tarkennetaan viitekehykseen kuuluvaa intentionaalista strategiaa, sen käsitteitä ja oletuksia (luvussa 3.2 ”Intentionaalisen systeemin käsite ja intentionaalinen strategia”). Lopuksi (luvussa 3.3 ”Intentionaalisten systeemien tutkiminen”) tarkastellaan tarkemmin intentionaalisen ja ekstensionaalisen kielen suhdetta ja intentionaalisuutta tieteessä.

3.1 Kolme strategiaa olioiden toiminnan tarkastelulle

Ensimmäinen askel vastauksien löytämiseksi mielen ongelmille on sivuuttaa ontologiset mieltymykset ja sen sijaan tarkastella niiden puhetaipojen eroa, jossa puhumme persoonista ja jossa puhumme kehoista ja muista fysikaalisista kappaleista. Tämä ahkera ontologisten sitoumusten välttely sallii meidän löyhentää mielen ja tieteen kielen välisen lähentymisen vaatimuksia [...] (Dennett 1969, 189.)

Dennettin mukaan olioiden toimintaa voidaan selittää ja ennustaa omaksumalla fysiikan tai fysikaalinen asenne (engl. physical stance), suunnittelun asenne (engl. design stance) tai intentionaalinen asenne (engl. intentional stance).²⁷

²⁶ Tästä luvusta eteenpäin sivunumeroviittaukset viittaavat oletusarvoisesti teokseen Dennett 1987 ja lyhennys ”BS” viittaa teokseen Dennett 1978a.

²⁷ Dennett puhuu myös moraalien ja persoonan asenteesta. Ne ovat kuitenkin intentionaalisen asenteen erityistapauksia, joten tässä työssä niitä ei käsitellä. Ks. moraalien asenteesta luku 12 ”Mechanism and Responsibility” ja persoonan asenteesta luku 14 ”Conditions of Personhood” teoksessa BS.

Näiden jokaisen kolmen asenteen²⁸ omaksuminen on erilainen strategia, jolla olioita – *systemejä* – tarkastellaan.

Fysiikan strategiaa käytetään, kun pyritään selittämään ja ennustamaan olion, kuten tähden tai molekyylin toimintaa suurella tarkkuudella. Sen idea pohjautuu nimensä mukaisesti fysiikan deterministiseen maailmaan, jossa jokaisella tapahtumalla on vaikuttava tai vaikuttavat syynsä. Fysiikan maailma on nomologinen, jossa pätevät luonnonlait. (16 ja BS, 4–5.) Fysiikan strategia operoi suoraan mitattavilla olevilla käsitteillä tai käsitteiden suhteilla, joilla nähdään olevan kutakuinkin suora vastineensa reaali maailmassa. Esimerkiksi mikroaaltouuni on pystytty kehittämään sen pohjalta, että tiedetään, miten vesimolekyylit käyttäytyvät, kun niihin kohdistetaan mikroaaltoja. Vesimolekyylit alkavat liikkua, jolloin ne synnyttävät lämpöenergiaa. Koska useimmissa ruuissa on vettä, ruoka saadaan lämpimäksi mikroaalloilla.

Sen lisäksi, että fysiikan strategialla pystytään ennustamaan olion toimintaa täsmällisesti, sen avulla voidaan myös määritellä olion fysikaalinen koostumus ja rakenne sekä tutkia olioon vaikuttavien tekijöiden fysikaalisia piirteitä. (16.) Esimerkiksi eläin koostuu elimistöistä, joka puolestaan muodostuu elimistä, jotka edelleen muodostuvat soluista, jotka taas koostuvat soluelimistä ja niin edelleen. Samaten tiedetään, että vesi koostuu vesimolekyyleistä, jotka koostuvat hapen ja vedyn yhdisteistä, jotka edelleen koostuvat elektroneista, protoneista ja neutroneista ja niin edespäin.

Fysiikan strategiaa ei kannata aina käyttää. Käyttöön voidaan välillä ottaa tehokkaampia ja taloudellisimpia strategioita, kuten suunnittelun strategia. (16.) Suunnittelun strategiassa hylätään todellinen yksityiskohtainen tieto olion fysikaalisesta koostumuksesta. *Suunnittelun strategia* olettaa systeemin olevan

²⁸ Joku saattaisi puhua "asenteen" sijaan "näkökulmasta" (näin itse asiassa tekee Huttunen [1996, 12]), mutta minusta on syytä puhua juuri "asenteesta". "Asenne" on tässä yhteydessä maailman hahmottamistapa tai ilmiöiden ajattelutapa, joka toisin kuin "näkökulma" on mielestäni täysvaltaisempi käsite.

tietyllä tavalla suunniteltu niin, että kyseisen systeemin elementeillä on funktioita, ja toimivan niin kuin sen on suunniteltu toimivan erilaisissa olosuhteissa. Suunnittelun asennetta käytetään nimensä mukaisesti juuri mekaanisten tuotteiden tarkasteluun (BS, 4), mutta se ei rajoitu pelkästään artefakteihin. Myös biologisten eliöiden, kuten kasvien, eläimien ja sisäelinten toimintaa voidaan ennustaa ja selittää suunnittelun asenteesta, koska ne ovat luonnonvalinnan ”suunnittelema”²⁹.

Usein ihmisillä ei ole pienintäkään aavistusta siitä, mitä herätyskellon tai tietokoneen sisällä tapahtuu, mutta heillä on hyvä käsitys siitä, mitä niiden on suunniteltu tekevän. Herätyskellon käyttäjän ei tarvitse välittää siitä, toimiiko herätyskello pattereilla, ja onko se mekaaninen vai elektroninen toteutukseltaan. Hänen tarvitsee vain olettaa, että herätyskello on suunniteltu päästämään hälytysääni, kun käsillä oleva hetki on yhtäläinen määritellyn herätysajan kanssa. (17.)

Suunnittelun asenteessa on eri tasoja hyvin tarkasta hyvin abstraktiin. Esimerkiksi tietokoneen toiminnan ennustaminen ja selittäminen tietokoneohjelman lähdekoodin tarkastelulla on hyvin abstraktin tason suunnittelun strategia, kun taas tietokoneen toiminnan ennustaminen ja selittäminen laitteiston logiikkapiirien tasolla on hyvin tarkan tason strategia. (Dennett 1988a, 496.)

Viimeisenä on *intentionaalinen asenne*, jonka kautta ennustettavaa systeemiä kohdellaan rationaalisen agenttina. Sille liitetään eli *attribuoidaan* sellaisia uskomuksia ja haluja, joita sillä pitäisi olla ottaen huomioon sen asema ja tarkoitus maailmassa. Tämän pohjalta ennustetaan, että kyseinen systeemi

²⁹ Kuten aiemmin kirjoitin, luonnonvalinta on itsessään päämäärätön ja tarkoitukseton prosessi, joten sen ”suunnittelusta” voidaan puhua (jonkun mielestä ainoastaan) lainausmerkeissä (ks. erityisesti luku 4.4 ”Kiista alkuperäisestä ja johdetusta intentionaalisuudesta”). Lainausmerkkien käyttäminen on joka tapauksessa tarpeetonta, kuten tulemme huomaamaan intentionaalisen asenteen yhteydessä. Evoluutio on nimittäin siinä intentionaalinen systeemi, missä muutkin intentionaaliset systeemit ovat.

toimii päämääräsuuntautuneesti sille attribuoitujen uskomuksien ja halujen mukaisesti. (Dennett 1988a, 496.) Intentionaalisen asenteen läpi oliot näyttäytyvät intentionaalisisina systeemeinä, ja siten niihin suhtaudutaan sellaisina. Intentionaalinen asenne lisää suunnittelun asenteeseen yhden ehdon: sen lisäksi, että suunnittelun asenne olettaa olion toimivan niin kuin sen on suunniteltu toimivan, intentionaalinen asenne olettaa suunnittelun olevan myös optimaalinen. Optimaalisella suunnittelulla tarkoitetaan sitä, että olio toimii mahdollisimman rationaalisti, esimerkiksi valitsee shakissa parhaan mahdollisen siirron. (BS, 5.)

3.1.1 Strategioiden väliset suhteet

Näistä strategioista fysiikan strategia kattaa kaikki oliot, sillä sen ainut ”vaatimus” on luonnonlakien voimassaolo, mikä pätee kaikkialla. Tästä syystä fysiikan strategiaa voidaan pitää luonteeltaan deskriptiivisenä, sillä se ei aseta sinänsä mitään ehtoja tarkasteltaville olioille. Suunnittelun strategia ja intentionaalinen strategia kattavat fysiikan strategian katvealueen osajoukon. Suunnittelun strategia vaatii olion olevan suunniteltu, ja intentionaalinen strategia lisäksi sen, että olio on nimenomaan optimaalisesti suunniteltu ja siten rationaalinen. Siksi suunnittelun ja intentionaalisen strategian kuvaustapaa voidaan luonnehtia normatiivisiksi. Fysiikan strategia on näistä tarkkuudeltaan parhain tai täydellisin, mutta suuri tarkkuus tapahtuu tehokkuuden kustannuksella. Fysiikan strategia on hyvin kustannustehoton ja vaativa, toisin kuin suunnittelun strategia tai intentionaalinen strategia, jotka pystyvät erinomaiseen tarkkuuteen pienillä kustannuksilla. Nämä eri strategioiden tunnuspiirteet on koottu alla olevaan taulukkoon (TAULUKKO 1):

TAULUKKO 1. Strategioiden tunnuspiirteet.³⁰

	Fysiikan strategia	Suunnittelun strategia	Intentionaalinen strategia
Oletukset	(Luonnonlait ovat voimassa.)	Olio on suunniteltu.	Olio on sekä optimaalisesti suunniteltu että rationaalinen.
Kuvaustapa	Deskriptiivinen	Normatiivinen	Normatiivinen
Ala	Kaikki oliot	Suunnitellut oliot: artefaktit ja organismit	Rationaaliset, intentionaaliset oliot: esim. ihmiset ja eläimet
Operoitavat käsitteet	Luonnonlait ja entiteetit (konstituentit)	Artefaktit ja organismit; suunnittelu	Uskomukset, halut ja mielentilat; optimaalisuus ja rationaalisuus
Tarkkuus	Ääretön	Suuri	Suuri
Tehokkuus	Pieni	Suuri	Huomattavan suuri

Esitellyiden strategioiden väliset suhteet ja tunnuspiirteet tulevat selvimmin ilmi esimerkistä. Ajatellaan shakkia pelaavaa tietokonetta. Sen toimintaa eli pelaamista voidaan ennustaa ja selittää kaikista kolmesta käsitellystä asenteesta käsin. Fysiikan asenteen omaksumalla tietokoneen siirtoja ennustetaan esimerkiksi sähköfysiikan tarjoamissa puitteissa. Tällöin tarvitaan tietämystä muun muassa sähkövirran ja elektronisten komponenttien toiminnasta. Vaikka shakkiohjelman siirrot voidaan ennustaa luotettavasti, fysiikan asenteen pohjalta on kuitenkin hyvin kuluttavaa ennustaa tietokoneen toimintaa yleisesti, puhumattakaan tietokoneen shakkiohjelman siirroista. Seuraava askel

³⁰ Dennettin viitekehyksen kenttä on filosofia, mutta esim. David Marr (1982) ja Allen Newell (1982) ovat esitelleet vastaavanlaiset viitekehykset ohjaamaan kognitiivisen teorian teoriaa. Viitekehyksien välillä on selviä eroja, ja kukaan heistä ei ole valmis sitoutumaan toistensa viitekehyksiin sellaisinaan. Esimerkiksi Newell (1982) kritisoi Dennettiä siitä, että hänen viitekehyksessään intentionaalisuus on katsojan silmässä, kun taas Dennett (75) Marria siitä, että tämä ei huomaa, kuinka idealisoitu tämän komputationaalinen taso on, mikä voi ohjata teorian hakoteille. Ks. esim. Newell 1982, 122–123; 74–75; Newell 1988; Dennett 1988a ja Salminen 2004, luku 3 ”Vertailu Newellin ja Marrin viitekehyksiin”. Ks. myös luku **Virhe. Viitteen lähdettä ei löytynyt.** ”Semantiikan luominen syntaksista”.

abstraktioportaassa vie meidät suunnittelun asenteen piiriin. Omaksumalla se tarkastellaankin tietokoneessa pyörivää tietokoneohjelmaa, eikä huomioida koneen fyysistä toteutusta. Erilaiset shakkiohjelmat saattavat olla suunnitellut muodostamaan erilaisia pelistrategioita, jotka voidaan ottaa huomioon suunnittelun asenteesta. Kuten shakkia tietokonetta vastaan pelaava ihminen tietää, ei näistä kumpikaan ole tehokas tapa ennustaa tietokoneen pelisiirtoja. Siksi onkin syytä omaksua intentionaalinen asenne, jolloin pelaaja ennustaa tietokoneen siirtoja sen mukaisesti, mikä olisi paras mahdollinen siirto, jonka shakkipelaaja voi yleensä tehdä, kun otetaan huomioon shakin säännöt ja pelin tavoite. Omaksumalla tämän asenteen pelaaja attribuoi tietokoneelle uskomuksia, haluja ja aikomuksia. Hän olettaa, että tietokone haluaa voittaa, että tietokone uskoo tavoittavansa siirroillaan voittoa ja että tietokone aikoo toimia näiden uskomuksien ja halujen mukaisesti pelin rajoitusten sisällä. Toisin sanoen henkilö kohtelee tietokonetta älykkäänä agenttina tai jopa ihmisenä.³¹ (BS, 4–6.)

Dennett asettaa John Haugelandin (1993) mukaan strategiat ja niiden käsittelemät entiteetit ja tapahtumat samalle viivalle: molemmat ovat ymmärrettävissä tietystä asenteesta.³²

Asenteen valinta tietyssä tapauksessa on, tietysti, rajoitettu siihen, minkä strategioiden odotetaan toimivan; mutta muuten se on vapaaehtoinen – pragmaattinen päätös, joka perustuu intresseille ja tarkoituserille. Siten, henkilöä voidaan kohdella fyysikaalisena objektina, kun lasketaan hissien kuormaa, suunniteltuna systeeminä suoritettaessa kirurgista leikkausta, intentionaalisenä systeeminä kahvia tarjottaessa ja niin edelleen. Ilmiöt, jotka ennustetaan ja selitetään kustakin asenteesta käsin, eivät ole

³¹ Shakkia pelaavaa tietokonetta ei välttämättä tarvitse tulkita juuri uskomusten ja halujen turvin intentionaalisesta asenteesta. Kuten aiemmin mainitsinkin, tietokonetta luonnehditaan intentionaalisin termein myös silloin, kun esimerkiksi sen sanotaan laskevan yrityksen kirjanpidon pohjalta taseen tai kun sen sanotaan tekevän jotakin niinkin yksinkertaista kuin tallentavan informaatiota (vrt. Dennett 1969, 40). Joka tapauksessa, "[s]amoihin ennusteisiin päädytään siitä huolimatta, ajatellaanko tietokonetta suoraan uskomus- ja halu-termeillä tai tietovarasto- ja tavoitemäärittys-termeillä" (BS, 7).

³² Tämä muodostaa Dennettin motiivin intentionaalisen puheen autonomisuudelle. Ks. tästä tarkemmin luku 3.3.4 "Reduktion mahdollisuus".

samoja. Esimerkiksi tarkoituksellinen toiminta ei ole sama kuin fysiologinen liike; ja molemmat ovat erillisiä fyysisestä liikkeestä. (Haugeland 1993, 65.)

3.1.2 Ihmisten arkinen ilmiöiden hahmottaminen

Kaikkia näitä strategioita, erityisesti intentionaalisuuden strategiaa, käytämme jatkuvasti. Emme usein tiedosta käyttävämme niitä, mutta se on lähinnä osoitus niiden olevan implisiittisesti rakentuneita toimintaamme ja asioiden hahmottamiseemme. Saatamme tulla tietoisiksi kasveja tarkastellessamme mielentilojen liittämistä niille, vaikkemme uskoisi niillä olevan oikeasti minkäänlaisia mielentiloja. Huomaamme silti mielentilojen liittämisen olevan tehokas tapa. Oletamme, että kukka haluaa saada auringonvaloa osakseen, ja liitämme sille uskomuksen, että auringonvaloa on saatavilla eniten päiväsaikaan. Näin ollen ennustamme kasvin levittävän terälehtensä laajimmilleen päiväsaikaan.

Tiedostetuimmin mielentilojen attribuominen tulee esille antropomorfismissa. On tavallista, että lasten saduissa on eläimiä, jotka ovat täysin ihmisenkaltaisia, mutta jotka ulkoisten piirteiden avulla tunnistetaan kuitenkin tietyksi eläimiksi. Nämä satujen ihmismäiset eläimet puhuvat, kävelevät kahdella jalalla, pitävät vaatteita yllään ja seurustelevat keskenään. Antropomorfismi ei tietenkään rajoitu sellaisenaan vain lasten satuihin, vaan sitä tavataan laajasti myös eri kulttuureissa ja uskonnoissa, varsinkin varhaiskantaisissa uskonnoissa. Niissä jumalille liitetään ihmismäisiä piirteitä niin ulkomuodon kuin uskomuksien ja halujenkin suhteen. Esimerkiksi Suomen muinaisuskonnoissa tunnettiin muun muassa ukkosen- ja sateenjumala Ukko, jolta pyydettiin sadetta kylvöön ja jota mielisteltiin erilaisilla lahjoilla ja juhlilla. Yleensä ottaen ihmiset ovat kautta aikain antropomorfoineet sellaisia asioita ja ilmiöitä, joita he eivät ole ymmärtäneet. Tieteen edistyessä kaikenlaiset antropomorfistiset selitykset niin kasvien kuin taivaankappaleiden toiminnoille

ovat kariutuneet, ja tiedämme tätä nykyä näiden ilmiöiden taustalla olevan täysin mekaanisia syitä.³³ Siltikin, koska intentionaalinen selittäminen on niin toimivaa, esimerkiksi elävien organismien tapauksessa, se on käyttövoimainen väline.

Dennettin mukaan mielentilojen liittäminen ja arkinen mielestä puhumisemme on meidän osittain synnynnäinen taitomme omaksua intentionaalinen asenne (Dennett 1991c ja 1995a, 237 alaviite).³⁴ Evolutiivisesti tarkasteltuna tämä ei ole lainkaan kummallista: omaksumme erilaisia olioiden tarkastelutapoja tai -strategioita, koska ne ovat ympäristöön sopeuttavia ominaisuuksia. Meidän ja muiden eläinten on ollut ja on hyödyllistä ennustaa muiden ihmisten ja muiden eläinten käyttäytymistä. On selvää, että eläinten on hyvä ennustaa mahdollisten parittelukumppanien, saaliseläinten ja niitä jahtaavien petojen liikkeitä selviytyäkseen. Näin ollen voidaan olettaa, että meidän synnynnäinen taitomme omaksua intentionaalinen asenne on luonnonvalinnan seurauksena kulkenut perimämme mukana.

3.2 Intentionaalisen systeemin käsite ja intentionaalinen strategia

Intentionaalinen systeemi on siis olio, jonka käyttäytymistä ja toimintaa voidaan selittää ja ennustaa attribuomalla sille uskomuksia, haluja ja muita mielentiloja. Voidaan sanoa, että jos ja vain jos intentionaalinen strategia ennustaa

³³ Esimerkiksi Matti Kamppinen on ansiokkaasti esittänyt, miten intentionaalisia selityksiä on ajettu alas viimeisen muutaman vuosisadan aikana. Läpitunkeva mystisyys ja antropomorfismi urbaanin ihmisen maailmassa on vähentynyt huomattavasti sitten uskonnollisten selitysten dominanssin. ”Inhimillistetty maailma on muuttunut mekanistiseksi, takaisin sellaiseksi jollainen se oli ennen ja ilman ihmistä” (Kamppinen 2000). Täysivaltaisten mekanististen selitysten ulkopuolella tuntuu enää olevan ainoastaan itse ihmiset ja kehittyneet nisäkkäät.

Huolimatta mekanististen selitysten voittokulusta, ihmiselle ehkä luontaisin ja helpoin tapa – kun kysymys on vähänkään monimutkaisimmista olioista – on antropomorfisoida olioita eli käyttää intentionaalista strategiaa niiden toiminnan tarkasteluun. Jos intentionaalinen strategia todella on kognitiivisesti vähiten rasittavin, voisi ajatella kehittyneen ja lähes ihmisen kaltaisen – tai sitä asteittain lähestyvän – tekoälyn olevan ihmiselle helpoin tapa olla vuorovaikutuksessa artefaktien kanssa, näin käyttäjäystävällisen tietojenkäsittelyn näkökulmasta.

³⁴ Ks. tarkemmin intentionaalisen asenteen ja arkipsykologian suhteesta Griffin & Cohen 2002 ja teoksen Dennett 1987 luku 3 ”Three Kinds of Intentional Psychology”.

luotettavasti ja laajasti oliion käyttäytymistä – eli strategia toimii –, niin olio on intentionaalinen systeemi (15).

Intentionaalista strategiasta voidaan luoda seuraavanlainen käyttöohje (17; vrt. Huttunen 1996, 13):³⁵

1. Kohtele oliota O, jonka käyttäytymistä tahdot ennustaa, rationaalisena agenttina.³⁶
2. Arvioi, mitä uskomuksia O:lla pitäisi olla ottaen huomioon sen asema ja tarkoitus maailmassa.
3. Arvioi vastaavasti mitä haluja O:lla pitäisi olla.
4. Lopuksi, ennusta mitä O tulee tekemään sen olettamuksen pohjalta, että O pyrkii täyttämään halunsa uskomuksiensa valossa.

Intentionaalisen systeemin tietämiselle ja uskomuksen muodostamiselle normaalisti riittävä ehto on altistuminen yli sopivan ajan jollekin ärsykkeelle, jonka systeemi havaitsee (18). Toisin sanoen, intentionaalinen systeemi tavallisesti muodostaa uskomuksen jostakin asiasta, kun kyseinen ärsyke ylittää systeemin reaktiokynnyksen. Nämä uskomukset useimmiten vastaavat todellisuutta, joten ne ovat tosia uskomuksia. Esimerkiksi kissa näkee edessään kymmenen metrin päässä muutaman sekunnin ajan hiiren, kunnes hiiri kaikkoo nurkan taakse. Kissa muodostaa uskomuksen siitä, että sen edessä oli hiiri, joka on tosi uskomus, koska hiiri todellakin oli sen edessä.

Intentionaaliset systeemit, kuten me, eivät tule kuitenkaan oppimaan tai muistamaan kaikkia totuuksia, joita niiden havaintoaistit niille tarjoavat. Sitä vastoin, ne uskomukset, joita normaalisti tulemme muodostamaan, ovat relevantteja totuuksia (18). En esimerkiksi tiedä, kuinka suurella osalla kadulla vastaantulevista ihmisistä on silmälasit tai minkä värisiä sukia minulla on,

³⁵ Vrt. luvussa 2.2 ”Intentionaalisuuden ongelma” olevaan praktisen syllogismin esitykseen.

³⁶ Rationaalisuutta tarkastellaan tarkemmin luvussa 4.2 ”Rationaalisuusoletus”.

mutta jos olisin sellaisesta tiedosta kiinnostunut, voisin oppia sellaisen tiedonhankkimistaidon. Käyttöohjeen toisen askeleen tarkennukseksi voidaan siis sanoa, että intentionaalista strategiaa käytettäessä attribuoidaan oliolle kaikki relevantit tosiasiat uskomuksiksi. Jos systeemillä on vääriä uskomuksia, ne täytyy perustella jotenkin, yleensä vetoamalla ympäristön läsnä oleviin piirteisiin, jotka ovat harhaanjohtavia suhteessa systeemin havaintokykyihin. (20.) Esimerkiksi kissalla on virheellinen uskomus siitä, että sen edessä on hiiri, joka onkin tosiasiasa vain pehmolelu. Voimme sanoa kissan muodostaneen tämän virheellisen uskomuksen, koska kaukaa katsottuna pehmolelu näytti aivan elävältä, aidolta hiireltä.

Intentionaaliselle systeemille relevantteja uskomuksia on kuitenkin hyvin paljon. Halut onneksi rajoittavat tätä uskomuslistaa ja helpottavat siten ennustamista. Minkälaisia haluja systeemillä sitten pitäisi olla? Perustavanlaatuisimmat halut ovat biologisia, jotka ovat kaikille eliöille yhteisiä. Intentionaaliset systeemit tahtovat selviytyä ja lisääntyä, joten ne haluavat ruokaa, turvaa, seksiä ja niin edelleen. Yleisemmin sanottuna systeemeille attribuoidaan haluja niistä asioista, joiden systeemi uskoo olevan hyväksi sille. Jos systeemille attribuoidaan eriskummallinen halu, sille myös uskomuksen tapaan pitää antaa erityinen perustelu. (20, 49.)

Erotuksena muihin eläimiin ihmiset elävät kulttuurin ja eritoten kielen ympäröimänä. Me voimme hankkia ja muodostaa monia sellaisia uskomuksia, joita eläimet eivät pysty, sillä joitakin uskomuksia on mahdollista esittää ainoastaan kielen avulla (BS, 274). Meillä voi olla uskomuksia viikonpäivistä, taiteesta ynnä muista kulttuurisista ilmiöistä toisin kuin eläimillä. Eläinten käyttäytymistä ennustetaankin pääasiassa biologiasta johdettujen uskomusten ja halujen perusteella.

Kielen mahdollistamana ja pakottamana muodostamme kuitenkin paljon yksityiskohtaisempia ”uskomuksia” ja ”haluja” kuin mitä oikeat uskomukset ja

halumme ovat.³⁷ Esimerkiksi saadakseen jotain täytyy sanoa, mitä haluaa, ja tämä yleensä johtaa siihen, että henkilö ilmaisee jonkin yksityiskohtaisemman halun kuin mitä hän alun perin tarkoitti. Jos olen ravintolassa, en voi vain sanoa tarjoilijalle, että ”minulla on nälkä ja haluan syödä”, vaan minun on tehtävä yksityiskohtainen tilaus, kuten ”härkää punaviinikastikkeessa ja vettä”. (20.)

Intentionaalinen strategia toimii Dennettin mukaan esimerkiksi nisäkkäillä, linnuilla, kaloilla, matelijoilla, hyönteisillä ja simpukoilla – jopa jollain artefakteilla, kasveilla, luonnonilmiöillä ja valtioilla³⁸. Esimerkiksi kun simpukka uskoo jonkin uhkan olevan sen lähettyvillä, se ei hellitä kuorensa kiinnipitämistä, ennen kuin se on vakuuttunut siitä, että uhka on ohi; shakkia pelaava tietokone ei lyö ratsuasi, koska se tietää, että se voi menettää seuraavalla vuorolla torninsa, ja se ei halua niin tapahtuvan; kukka avaa terälehtensä laajimmilleen keskipäivällä, koska se uskoo silloin auringonvaloa olevan eniten saatavilla; salama haluaa kulkea lyhintä ja parasta reittiä maahan, jolloin esimerkiksi maanalainen vesipumppu voidaan suojata salamaniskulta rakentamalla vesipumpun viereen parempi ja ilmeisempi reitti salamalle. (22.)

Sulkeeko intentionaalisen systeemin määritelmä sitten mitään olioita pois? Onko esimerkiksi työpöytä intentionaalinen systeemi? Jos ennustan työpöydän käyttäytymistä täsmällisesti intentionaalista asenteesta, onko se silloin intentionaalinen systeemi? Jos se on intentionaalinen systeemi, mikä tahansa on. Työpöytä ei kuitenkaan ole intentionaalinen systeemi, nimittäin intentionaalista strategiasta ei ole mitään hyötyä työpöytää tarkastellessa, sillä intentionaalinen strategia ei tarjoa meille mitään tietoa, jota meillä ei aikaisemmin ollut.³⁹ Tiesimme jo ennestään mitä työpöytä aikoo tehdä, ei näet

³⁷ Ks. luku 4.1 ”Uskomuksen ja mielipiteen erottelu”.

³⁸ Samassa mielessä kuin muurahais- tai ampiaispesät myös valtiot, hallitukset ja yhtiöt voidaan nähdä yksittäisinä, intentionaalina systeeminä, koska kyseessä on funktionaalinen määritelmä intentionaalille systeemille. Ks. valtioista ja hallituksista Clark 1994 ja yhtiöistä Weaver 1998 intentionaalina systeeminä.

³⁹ Vrt. Dennettin huomioon sisällön liittämistä liikeimpulsseille:

mitään. Toisekseen, uskomukset ja halut sopivat mielestämme hyvin huonosti työpöydälle.⁴⁰ Ihmisten, eläinten ja tietokoneiden tapauksessa asian laita on eri. Niiden ollessa kyseessä ainoastaan intentionaalinen strategia on käytännöllinen, toisin kuin suunnittelun tai fysiikan asenne. (22–23.)

Intentionaalinen strategia on todellakin käytännöllinen, mutta riskialtis ja eräänlainen oikotie verrattuna fysiikan strategiaan. Tietysti on periaatteessa mahdollista ennustaa tietokoneiden – tai jopa ihmisen – käyttäytymistä fysiikan lakien nojalla. Fysiikan strategia on kuitenkin hyvin työläs ja raskas, ja sen käyttäminen elävien organismien ennustamiseen on käytännössä mieletöntä. Esimerkiksi pelkästään veden liikkeen mallintaminen on hyvin hankalaa. Käytännössä nerokkaimmat insinööritkään eivät pysty pitämään selvää fysiikan käsitystä monimutkaisimmista laitteista, kuten hajautetuista tietojärjestelmistä, vaan ajautuvat suunnittelun asenteeseen ja pohtivat, mitä laitteen tulee tehdä ottaen huomioon sen suunnittelun tarkoituksen. Yksinkertaisempien komponenttien, kuten transistorin suunnittelijat kuitenkin voivat käyttää fysiikan näkökulmaa, jolloin he pystyvät ennustamaan komponenttien toimintaa täsmällisesti, toisin kuin tavalliset käyttäjät, jotka turvautuvat vähintään suunnittelun asenteeseen – ja yleensä korkeaan sellaiseen. Toisaalta ihmiset yleensä kohtelevat shakkiohjelmia rationaalisena agenttina, koska heillä ei ole käsitystä ohjelman suunnittelusta, ja myös siksi, koska he olettavat ohjelman olevan rationaalinen; he olettavat ohjelman tahtovan voittaa pelin ja

Asettamalla [engl. assign] sisältö esimerkiksi viimeisille lihassupistusta stimuloiville liikeimpulsseille ei saavuteta mitään. Sellaiselle tapahtumalla käskyn ”supistu nyt, lihas!” antaminen ei juuri selvennä tapahtuman kulkua. (Dennett 1969, 79–80.)

⁴⁰ Sille, että työpöytää ei voida laskea intentionaaliseksi systeemiksi on Dennettiä tulkitessani mielestäni vain se syy, että me emme intentionaalisen asenteen käyttäjinä kerta kaikkiaan näe sitä missään tapauksessa intentionaalisen systeeminä. Loppujen lopuksi ainut kriteeri sille, onko jokin intentionaalinen systeemi vai ei, on se, käytämmekö mielekkäästi intentionaalista asennetta sen suhteen. Näin siksi, koska oliot on nähtävissä intentionaalisen systeeminä ainoastaan intentionaalisen asenteen kautta, sen asenteen kautta joka meillä on sisäänrakennettuna. Tässä mielessä eksplisiittinen määritelmä intentionaalille systeemille ei ole viime kädessä kovin tärkeässä asemassa.

siksi tekevän sellaisen ja sellaisen siirron. Kyse strategian vaihtamisesta on siis ihmisen episteemisistä eli tiedollisista rajoituksista. (23.)

3.2.1 Intentionaalisuuden tulkinnallisuus

Intentionaalisella systeemillä on intentionaalisia tiloja määritelmän mukaan. Siis: jos jonkin oliion toiminta on luotettavasti ennustettavissa intentionaalisesta asenteesta, on oliolla määritelmän mukaan intentionaalisia tiloja.⁴¹ Intentionaaliset tilat ovat Lynne Rudder Bakerin (1989) sanoin ”asenneriippuvaisia” (engl. stance-dependent):

[O]lla tosi uskoja tarkoittaa *ainoastaan* sitä, että on systeemi, jonka käyttäytymistä voidaan luotettavasti ennustaa intentionaalisesta strategiasta, ja siksi se, että todella ja aidosti uskoo, että p (mikä tahansa propositio p) tarkoittaa *ainoastaan* sitä, että on intentionaalinen systeemi, jolle p esiintyy uskomuksena parhaassa (ennustavimmassa) tulkinnassa.⁴² (29.)

Dennettin näkemyksen mukaan intentionaaliset tilat ovat *asenneriippuvaisia*, koska ne ovat ominaisuuksia, joita systeemillä on vain siksi, että se on (mahdollisesti) tietyn asenteen objekti. Niitä ei ole intentionaalisesta strategiasta riippumatta, ne eivät ole ”*asenneriippumattomia*” (engl. stance-independent). (Baker 1989, 305.)

Seuraavaksi onkin luonnollista kysyä, miten Dennettin ehdotus erottaa sellaiset systeemit, kuten ihmiset, joilla on (mielestämme) oikeasti uskomuksia sellaisista, joilla ei mielestämme varmasti ole uskomuksia, esimerkiksi termostaateista? Dennettin mukaan tämänkaltainen erottelu on tulokseton. Hän pyytää meitä tarkastelemaan tarkemmin oletettua eroa esittämällä seuraavanlaisen hupaisan ajatusleikin (30-31; mukailtu):

⁴¹ Tästä syystä edellä puhuttiinkin sekaisin (1) uskomuksien ja halujen attribuomisesta systeemille sekä (2) systeemin uskomuksista ja haluista.

⁴² William Seager (2000, 103–104) huomauttaa onnistuneesti, että Dennettin mielestä ei ole niin, että voimme tulkinta intentionaalisilla systeemeillä olevan mentaalisia ominaisuuksia siksi, koska niillä on mieli, vaan päinvastoin: intentionaalisilla systeemeillä on mieli siksi, koska me voimme tulkita niiden tilat (niiden käyttäytymisen perusteella) mentaalisiksi.

Termostaatille voidaan attribuoida muutamia uskomuksia ja haluja: se voi uskoa, että huoneessa on liian kylmä tai liian kuuma; että boileri on päällä tai pois päältä; että jos se tahtoo huoneen lämpötilaa nostettua, sen pitäisi laittaa boileri päälle ja niin edespäin. Selvästikin tämä on tarpeetonta intentionaalisten tilojen liittämistä termostaatille, sillä eihän sillä ole lämmön tai boilerin käsitteitä. Sen uskomukset ja halut voidaan purkaa tulkinnasta (engl. de-interpret): se voi uskoa, että A on liian F tai G ja että jos se haluaa A:n olevan enemmän F, sen pitäisi tehdä K ja niin edespäin. Loppujen lopuksi termostaatti on vain eräänlainen kontrollimekanismi erilaisten sisääntulo- ja ulostulolaitteiden välillä, joten se voidaan laittaa esimerkiksi säätämään tankin vesimäärää tai junan nopeutta. Oletetaan nyt, että rikastamme näitä termostaatin kiinnitystapoja. Lisäämme sille esimerkiksi useamman kuin yhden tavan saada tietoa lämpötilasta. Annamme sille jonkinlaisen silmän, jolla se voi erottaa palelevat huoneessa olijat, ja korvan, jotta sille voidaan kertoa, kuinka kylmä huoneessa on. Annamme sille muutamia tietoja maantieteestä, jotta se voi päätellä olevansa kylmässä paikassa, jos se oppii, että se on Jyväskylässä tammikuussa. Tietystikin yleiskäyttöisen visuaalisen järjestelmän antaminen termostaatille vaatii laajoja, hankalia muutoksia sen rakenteeseen. Muutamme lisäksi termostaatin käyttäytymistä yleisemmäksi: se valitsee itse boilerin polttoainetyypin, ostaa sen kaupasta, tarkastaa huoneen ikkunoiden tiivisteet ja niin edelleen. Tämä lisää uuden ulottuvuuden sen sisäiseen rakenteeseen ja monimutkaistaa sitä edelleen.

Käyttäytymisen muutos yleisemmäksi antaa termostaatin uskomuksille lisää tehtävää, koska ne pitää johtaa deduktiivisesti muista uskomuksista ja mielentiloista ja koska ne toimivat pmissenä muille päättelyille. Kyseinen prosessi lisää yhteyksiä termostaatin ja maailman välillä, mistä seuraa termostaatin predikaattien, kuten F ja G, merkityksien rikastuminen. Toisin sanoen, sitä vähemmän termostaatti on pelkästään kontrollimekanismi huoneen

lämpötilalle mitä enemmän sen merkityksiä rikastetaan. Muodollisemmin sanottuna tämä tarkoittaa, että järjestelmän tyydytysmallien joukko tulee pienemmäksi ja pienemmäksi, kun lisäämme kuvatun kaltaisia monimutkikkuuksia. Mitä enemmän monimutkikkuuksia lisäämme, sitä rikkaammaksi, vaativammaksi ja yksityiskohtaisemmaksi järjestelmän semantiikka tulee, kunnes lopulta saavutamme järjestelmän, jolle uniikki semanttinen tulkinta on käytännössä määrätty ja sidottu:

Siinä pisteessä sanomme, että laitteella (tai eläimellä tai henkilöllä) on uskomuksia *lämmöstä* tai *tästä tietystä huoneesta* ja niin edelleen, ei vain systeemin tosiasiallisen paikan ja toimintojen maailmassa olemisen takia, vaan sen takia, ettemme pysty kuvittelemaan systeemille toista paikkaa [engl. niche], johon sijoittamalla se toimisi. (31.)

Esimerkkimme alkuperäisellä, yksinkertaisella termostaatilla oli tila, jota kutsuimme uskomukseksi tietysti boilerista, uskomukseksi siitä, oliko kyseinen boileri päällä vai ei. Miksi juuri kyseisestä boilerista? Siksi, koska termostaatti oli kiinnitetty kyseiseen boileriin. Koska kyseessä oli termostaatin hyvin minimaalinen kausaalinen yhteys maailmaan, sen merkitys voitiin muuttaa helposti. Esimerkiksi sama termostaatti olisi voinut säädellä junan nopeutta ilman termostaatin sisäisen rakenteen muuttamista. Mutta mitä rikkaampia havainnollisia kykyjä ja käyttäytymistäipumuksia termostaatille (tai systeemille) annamme, sitä hankalammaksi tulee sen merkityksen eli maailmaan suhteessa olevien kausaalisten yhteyksien muuttaminen kajoamatta systeemin sisäiseen rakenteeseen. Jos systeemin ympäristöön tehdään muutoksia, se huomaa muutoksen ja tekee vastaavan muutoksen sisäiseen rakenteeseensa. Jos laite sidotaan yhteen tiettyyn tilaan, sen käyttölaajuus pienenee (ts. se vaatii yksityiskohtaisemman aseman ympäristössä). Toisaalta, jos systeemiä ei sidota yhteen tilaan, vaan muutetaan pelkästään sen ympäristöä, sen havaintoapparaatti on tarpeeksi tarkka sitoakseen itse oman tilansa toimiakseen tietyssä ympäristössä tehokkaasti. Tällaista tiukkaa suhdetta systeemin sisäisen rakenteen ja sen ympäristön välillä kutsumme

systemin representaatioksi. Systemi ikään kuin peilaa ympäristöä. Dennett toteaaakin painokkaasti, ettei

ole niin, että me attribuoinne (tai meidän täytyisi attribuoida) uskomuksia ja haluja asioille, joista löydämme sisäisiä representaatioita, vaan pikemmin niin, että todetessamme intentionaalisen strategian toimivan jollekin oliolle, me pyrimme tulkitsemaan joitakin sen sisäisiä tiloja tai prosesseja sisäisinä representaatioina. Mikä tekee asian jostakin sisäisestä piirteestä representaation voi olla ainoastaan sen rooli intentionaalisen systeemin käyttäytymisen säätelyssä. (32, kursiivit poistettu.)

Dennettin mukaan ei siis ole mitään taianomaista hetkeä siirtymässä termostaatista systeemiin, jolla todella on sisäinen representaatio sen ympärillä olevasta maailmasta. Termostaatilla on minimaalisen vaativa representaatio maailmasta, hienommilla termostaateilla vaativampi representaatio ja Nasan käyttämällä roboteilla vieläkin vaativampi representaatio. Lopuksi olemme me. Me olemme niin monisyisesti yhteydessä maailmaan, että meidän korvaaminen toisilla systeemeillä on lähes mahdotonta.⁴³ (32.)

Ihmisille attribuoitujen uskomusten tulkinta on hyvin hankalaa, toisin kuin esimerkiksi yksinkertaisen termostaatin, mutta periaate on sama. Erot ovat vain tulkinnallisuuden ongelman suuruudessa. Me olemme samassa asemassa kuin termostaatit: systeemeillä on intentionaalisia tiloja vain intentionaalisesta asenteesta katsottuna. Ero on enintään intentionaalisuuden asteessa, ei laadussa. Tiedämme, miten termostaatin käyttäytymistä voidaan ennustaa alemman tason strategioita käyttämällä. (32–33.) Eläinten ja ihmisten käyttäytymistä ei voida (vielä toistaiseksi) selittää ihmisen episteemisten rajojen takia esimerkiksi fysiikan strategialla.

Entä olisiko intentionaalisia tiloja olemassa, jos olisi olemassa jokin laplacialainen demoni, joka ei olisi sidottu episteemisiin rajoituksiin ja pystyisi ennustamaan kaikkien olioiden, ihmistenkin, käyttäytymisen fysiikan

⁴³ Ks. kuitenkin Putnam 1975 ja luku 4.4.2 ”Intentionaalisuus Rinnakkaismaassa”.

asenteesta käsin? Tällöinhän intentionaalisuus olisi vain episteemisesti rajoittuneen katsojan silmässä. Dennett ei tällaista subjektiivista käsitystä itse allekirjoita, vaan väittää intentionaalisten tilojen olevan objektiivisesti olemassa:

Intentionaalinen asenne on vapaasti valittavissa, mutta tosiasiat asenteen menestyksestä tai epäonnistumisesta [...] ovat täydellisen objektiivisia (24).

Tässä mielessä Dennettin näkemys on realistinen. Ilmeisen ristiriitaisesti hän siltikin väittää, että intentionaaliset tilat ovat näkyviä (olemassa olevia) vain tietystä asenteesta käsin, mikä viittaa enemmän löyhempään instrumentalistiseen ja interpretationistiseen näkemykseen. Tätä päällepäin näkyvää ristiriitaisuutta käsitellään tarkemmin luvussa 4.6 ”Intentionaalisten tilojen ontologinen asema”.

3.2.2 Intentionaalisen strategian epätäydellisyys ja voima

Intentionaalinen strategia ei pysty antamaan yksityiskohtaisia ennustuksia olion käyttäytymisestä juuri sen tulkinnallisuuden elementin takia. Olion toiminnasta voidaan antaa lähes loputtoman monta erilaista toimivaa intentionaalista tulkintaa, eikä voida sanoa, mikä kyseisistä tulkinnoista on oikea tai edes toista oikeampi.

Tulkinnat eroavat toisistaan erityisesti yksityiskohdiltaan, mutta tulkinnan pääsuuntaus on selvä. Esimerkiksi shakkipelaajan siirtoja ei voida ennustaa täsmällisesti, etenkin aloituksessa. Intentionaalinen strategia pystyy karsimaan pois useimpia sääntöjen mukaisia mahdollisia siirtoja. Sen avulla voidaan ennustaa, mitä siirtoja pelaajat eivät ainakaan tee, olettaen heidän olevan rationaalisia. Vain osa kaikista mahdollisista sääntöjen mukaisista siirroista on hyödyllisiä, jos haluna on pelin voittaminen. Sama pätee esimerkiksi pörssimeklarin käyttäytymisen ennustamiseen. On varmaa, että pörssimeklari pyrkii maksimoimaan voiton, jolloin voidaan ennustaa –

yksinkertaistaen – hänen myyvän, kun osakkeiden arvot ovat korkealla ja ostavan, kun niiden arvot ovat matalalla.⁴⁴ (24.)

Intentionaalinen strategia on hyvin tehokas. Joka päivä kulkiessamme liikenteessä moottoritiellä, panemme elämämme likoon luottaen (tiedostamattomasti), että muilla motoristeilla on havaintoihin pohjautuvia uskomuksia, normaaleja haluja ja päätöstaipumuksia (48). Käytämme intentionaalista strategiaa tietämättä lainkaan, mitä oikeastaan tapahtuu ihmisten pään sisällä. Fodoria (1987, 3) mukailleen: kuvitellaan, että joku, jota Fodor ei tunne, soittaa hänelle hänen New Yorkin toimistoonsa Arizonasta. Soittaja kysyy: ”Voisitko tulla luennoimaan tänne ensi tiistaina?” ja Fodor vastaa pyyntöön myöntävästi: ”Totta kai. Saavun lentoasemallenne kello kolmen lennolla.” Tämä on kaikki mitä tapahtuu, ja tämä riittää sen ennustamiseen, että useita päiviä ja tuhansia kilometrejä myöhemmin Fodor on lentoasemalla ja soittaja vastassa häntä.

Intentionaalinen strategia on ennustusvoimaltaan jopa niin suuri, että voimme ennustaa maapallon ulkopuolisen elämän käyttäytymistä jollakin kaukaisella planeetalla, vaikka planeettaa asuttavat oliot olisivat hyvinkin erilaisia kuin ihmiset tai tuntemamme maapallon biologiset organismit. Ainoa oletus on, että planeetan elämä täyttää evoluution yleiset edellytykset. Kun olemme alustavasti tunnistaneeet ympäristön vaarat ja avut suhteessa siinä eläviin olioihin, voimme arvioida, mitkä tavoitteet ja mitkä tavoitteiden painotukset ovat optimaalisia suhteessa olioiden tarpeisiin (selviytymistä ja lisääntymistä varten), minkälainen informaatio ympäristöstä on hyödyllistä tavoitteisiin

⁴⁴ Vrt. Herbert Simonin (1970, 8) huomioon liikeyrityksen ennustamisesta:

Jos tiedämme liikeyrityksestä vain sen, että se on voittoa maksimoiva järjestelmä, voimme usein ennustaa, miten sen käyttäytyminen muuttuu, jos muutamme sen ympäristöä – miten se [liikeyritys] muuttaa sen hintoja, jos sen tuotteista kannetaan arvonnäköveroä.

Lisäksi Simon onnistuneesti huomauttaa, että voimme tehdä tämän ennustuksen – ja ekonomistit tekevätkin – ilman yksityiskohtaistietoa sopeutumismekanismista, päätöksentekojärjestelmästä, joka muodostaa liikeyrityksen sisäisen ympäristön (Simon 1970, 8).

suuntautunutta toimintaa varten ja minkälaiset toiminnot ovat tarkoituksenmukaisia ottaen huomioon ympäristön olosuhteet. Dennettin mukaan tällöin ”[h]alut ’ilmenevät’ [olion] käyttäytymisessä niinä tarkoituksenmukaisina haluina (ottaen huomioon olion tarpeet), joiden suhteen toiminta olisi tarkoituksenmukaista ottaen huomioon olion uskomukset.” (BS, 8.) Sen jälkeen kun nämä mielentilat on attribuoitu olioille, olioiden toimintaa voidaan ennustaa niiden uskomuksien ja halujen pohjalta, ja jos evoluutio on tehnyt työnsä, kuten pitkässä ajassa sen on täytynyt, nämä ennustukset ovat tarpeeksi luotettavia ollakseen hyödyllisiä (BS, 9). Kuten Fodor (1987, 4) asian ilmaisee: ”jos voisimme ennustaa hyvin säätä [kuin intentionaalisten systeemien käyttäytymistä], kukaan ei koskaan kastelisi jalkojaan; ja silti sään etiologia täytyy olla lastenleikkiä verrattuna [monimutkaisten intentionaalisten systeemien] käyttäytymisen syihin.”

Intentionaalinen strategia on täynnä *ceteris paribus* -lakeja. Esimerkiksi Fodorin lausahdus ”saavun lentoasemallenne kello kolmen lennolla” tarkoittaa, että Fodor saapuu lentoasemalle kello kolmen lennolla, jos olosuhteet säilyvät samoina: jos Fodor ei valehtele, jos hän ei ole huumeiden vaikutusten alainen ja niin edelleen. Tämän takia moni filosofi onkin ajatellut, että intentionaalisen strategian (tai arkipsykologian) ennustusvoima on epäaito: ehkä Fodorin lausahdus ei tarkoita viime kädessä mitään muuta kuin ”jos joku sanoo, että hän aikoo olla siellä, sitten hän aikoo olla siellä – ellei hän aio” (Fodor 1987, 4). Mikään ei voi falsifioida sellaista ennustusta. Näin ajatteleva on kuitenkin harhateillä. (Fodor 1987, 4–6.)

Intentionaalinen strategia monen muun erityistieteen tapaan on idealisoiva. Se olettaa, että sen tietyt oletukset ovat voimassa; tapahtuvat poikkeukset eivät kuulu sen piiriin. Jos Fodor on huumeiden vaikutuksen alaisena – jolloin huumeet aiheuttavat poikkeavia biokemiallisia prosesseja ja toimintaa – minkä takia Fodor ei (persoonana) ymmärrä sanomaansa, on livahdettu ulos intentionaalisen strategian piiristä. Koska intentionaalinen strategia ei ota

kantaa olioiden sisäiseen rakenteeseen, se postuloi niiden toimivan niin kuin niiden pitäisi. Oliot eivät kuitenkaan aina toimi, niin kuin niiden pitäisi toimia (tehtäessä optimaalisuuden oletus). Ihmiset eivät ole aktuaalisesti täysin rationaalisia – mitä se sitten onkin – vaan sitä approksimoivia.⁴⁵ Sen tähden intentionaalinen strategiakaan ei ole aina oikeassa.

Yleisemmin ilmaistuna intentionaalisen asenteen omaksuminen kannattaa, kun voidaan tehdä optimaalisen suunnittelun (ja rationaalisuuden) oletus ja epäillä fysiikan ja suunnittelun asenteen käytännöllisyyttä. Optimaalinen suunnittelu on Dennettin mukaan sitä, että suunnitellun systeemin tutkija voi ennustaa systeemin toimintaa tuntematta sen rakennetta ajattelemalla systeemin sisälle homunculuksen eli pikkuihmisen, joka pelkästään tahtoo tehdä systeemin suunnitellun tehtävän ja tietää, miten tehtävä tehdään. Lisäksi systeemin tutkija olettaa, että homunculuksen toiminta on ”täydellistä”. (73.) Otetaan tästä eräs kuvaava esimerkki.

Ensimmäinen shakkia pelaava kone oli Baron Wolfgang von Kempelin 1700-luvun lopulla rakentama puinen mannekiini, joka siirsi shakkinappuloita ja pelasi kunnollista shakkia. Kesti joitakin vuosia ennen kuin huijaus paljastui: pseudomekaanisen koneen sisällä piilotteli salainen pienikokoinen avustaja, joka teki siirrot. Kyseessä oli siis kirjaimellisesti homunculus. Intentionaalinen strategia ei erota Kempelin huijaustapausta nykyisistä shakkia pelaavista tietokoneista, jotka todella pelaavat itsenäisesti shakkia. Hempelin shakkimannekiini ja shakkia pelaavat tietokoneet ovat samassa asemassa intentionaalisesta asenteesta käsin, sillä niiden sisäiseen rakenteeseen ei kiinnitetä huomiota. Intentionaalista asennetta käyttäessä tehdään implisiittisesti oletus, että intentionaaliset systeemit ovat optimaalisesti suunniteltuja (tai ainakin optimaalista suunnittelua approksimoivia) – tässä tapauksessa siis optimaalisesti suunniteltuja shakin pelaamista varten.

⁴⁵ Ks. luku 4.2 ”Rationaalisuusoletus”.

Intentionaalisen strategian ehdoton vahvuus on puuttumattomuus intentionaalisen systeemin aktuaalisen toteutukseen. (74.)

Termostaattia tarkasteltaessa intentionaalinen strategia toimii myös, koska se on hyvin suunniteltu. Sama pätee ihmiseen, koska ihminen on yksi luonnonvalinnan tuotoksista ja vieläpä hyvin myöhäinen sellainen. Termostaatin tapauksessa on vielä helppoa selvittää, miksi se toimii niin optimaalisesti tai miksi intentionaalinen strategia sen suhteen toimii, mutta ihmisen ollessa kyseessä kysymys on hankalampi eikä siihen voida toistaiseksi vastata. (33.)

3.3 Intentionaalisten systeemien tutkiminen

Jos intentionaalisia systeemejä tahdotaan tutkia tarkemmin ja tieteellisesti, on syytä erottaa toisistaan kaksi täysin erilaista lähestymistapaa. Dennett nimittää ensimmäistä Intentionaalisten systeemien teoriaksi (engl. Intentional System Theory) ja toista Alipersonaaliseksi kognitiiviseksi psykologiaksi (engl. Sub-personal Cognitive Psychology). Ensin mainittu vastaa intentionaalista strategiaa. Se on idealisoiva, holistinen ja instrumentalistinen teoria, joka antaa intentionaalisia, käsitteellisiä vastauksia. Jälkimmäinen kertoo, miten erilaiset toteutukset realisoivat intentionaalisen systeemin ja miten intentionaalisten järjestelmien prosessit heijastelevat intentionaalista tulkintaa. Alipersonaalinen psykologia siten muistuttaa suunnittelun strategiaa.⁴⁶

⁴⁶ Tämänkaltainen jako näyttää olevan lähtökohta kaikelle Dennettin puheelle mielestä. Ks. esim. Dennett 1969, luku 11 ”Personal and Sub-personal Levels of Explanation: Pain”. Dennett ei tosin puhu muualla intentionaalisten systeemien teoriasta tai tasosta, vaan persoonan tasosta. Hänen mielestään Rylen ja Wittgensteinin virhe oli se, etteivät he tehneet tätä eroa persoonan (engl. personal level; huom. tämä ei kuitenkaan viittaa siihen persoonan asenteeseen, josta puhuin aiemmin) ja alipersonaalisen (engl. sub-personal level) tason välille (Dennett 1969, 100). Vaikka jaottelu näillä nimenomaisilla termeillä esiintyy pääasiassa Dennettin puheessa tietoisuudesta, on huomattava, että kyseessä on pohjimmiltaan täysin sama jako kuin intentionaalisten systeemien teorian ja alipersonaalisen kognitiivisen psykologian välillä.

Itse asiassa olisi parempi puhua alipersonallisesta tasosta yleisesti, koska sen voidaan katsoa puolestaan koostuvan eri tasoista, esim. fysiikan tasosta ja kognitiivisen psykologian (tai suunnittelun) tasosta (tai eri intentionaalisuuden määrän tasoista). Pitäydyn kuitenkin

3.3.1 Intentionaalisten systeemien teoria

Intentionaalisten systeemien teoria on käsitteellisyydessään, normatiivisuudessaan ja instrumentalismissaan lähisukulainen päätös- ja peliteorialle. Samoin kuin intentionaalisten systeemien teorian lähisukulaiset, se lainaa intentionaalisia idiomeja eli arkipsykologian uskomus- ja halu-termejä, mutta antaa niille uuden, teknisen merkityksen teorian sisällä. (58.)

Intentionaalisten systeemien teoriassa ei voida puhua uskomus- ja halu-termeistä arkipsykologian merkityksin, sillä arkipsykologia sekoittaa intentionaalisten idiomien kausaaliset ja käsitteelliset yhteydet (55). Intentionaalisten systeemien teoria ei voi puhua uskomuksista ja haluista kausaalisina; se on puhtaasti normatiivinen teoria, koska se tekee idealisoidun rationaalisuusoletuksen. Se ei voi tutkia empiirisesti, mitä intentionaalisia tiloja systeemeillä ”todella” on, vaan se tarkastelee systeemeitä suhteessa siihen, mitä intentionaalisia tiloja niillä pitäisi olla.

Dennett (53) vetoaa Hans Reichenbachin (1938) tekemään erotteluun abstractum- ja illatum-termien välille puhuessaan intentionaalisista tiloista. *Abstractum*-luokan termit ovat hyödyllisiä laskennallisia entiteettejä tai loogisia konstruktioita. *Illatum*-luokan termit tarkoittavat postuloituja teoreettisia termejä, joilla toisin kuin abstractum-luokan termeillä voi olla kausaalisia vaikutuksia. Esimerkkejä abstractum-luokan olioista ovat painopiste, päiväntasaaja ja keskivertosuomalainen, ja illatum-luokan olioista geenit, atomit ja elektronit. Päiväntasaaja on looginen konstruktio, jolla ei ole kausaalista vaikutusta, toisin kuin esimerkiksi geeneillä. Tämä tarkoittaa sitä, että illatum-luokan termit ovat olemassa ilman ihmisen käsitteellisiä rakennelmia, sillä ne vaikuttavat kausaalisesti. Abstractum-luokan termit taas eivät vaikuta kausaalisesti.

tässä jaottelussa, koska Dennett on tehnyt niin intentionaalisuudesta kertovassa pääteoksessaan Dennett 1987.

Intentionaalisten systeemien teoria on eräänlaista *holistista loogista behaviorismia*. Looginen behaviorismi sinällään on liitetty Gilbert Rylen (1949) edustamaan näkemykseen, jonka mukaan mentaaliset tilat ovat vain yksi tapa kuvata käyttäytymisdispositioita eli -taipumuksia. Mentaaliset tilat voidaan kääntää sillä tavoin monimutkaisemmiksi väitteiksi, jotka eivät sisällä mentaalisia väitteitä. Dennettin mukaan Rylen virhe oli luonnehtia yksittäisiä mentaalisia tiloja tiettyinä, yksittäisinä käyttäytymisdispositioina. Dennettin mielestä intentionaaliset tilat ovat attribuoitavissa ainoastaan holistisesti suhteessa koko systeemiin ja sen muihin intentionaalisiin tiloihin, ei yksittäisinä intentionaalisinä tiloina – tästä termi ”holistinen”.⁴⁷ Intentionaalisten systeemin teorian tekee loogiseksi behaviorismiksi se, että se kohtelee systeemiä mustana laatikkona. Se ei siis anna kuvauksia systeemin sisäisistä prosesseista. (58.)

Dennett vertaa intentionaalisen systeemin ja alipersonallisen kognitiivisen psykologian välistä eroa fysiikan kinematiikan ja dynamiikan väliseen eroon.⁴⁸ Kinematiikka tarjoaa yksinkertaistetun ja idealisoidun abstraktitason, joka soveltuu moniin käyttötarkoituksiin, kuten lähtöpisteen suunnitteluun vaihteiston suunnittelussa. Silloin kun vaihdelaatikon suunnittelijan tarvitsee tarkempaa yksityiskohtaista tietoa jostakin, kuten kitkasta tai hyötysuhteesta, hän siirtyy käyttämään dynamiikkaa tarkempien ja luotettavampien ennustuksien takia, vaikkakin tämä siirtyminen tapahtuu monimutkaisuuden lisääntymisen ja yleispätevyyden menetyksen kustannuksella. Samalla tavalla voidaan Dennettin mukaan kysyä intentionaalisisista systeemeistä abstraktimman tason kysymyksiä, kuten: ”minkälaisia episteemisiä kykyjä ja taipumuksia systeemillä pitää olla selviytyäkseen ympäristössä A?”, ”mitä

⁴⁷ Idea on peräisin Donald Davidsonin (1967, 1973) radikaalin tulkinnan (engl. radical interpretation) teoriasta, jossa puolustetaan merkitysholismia. Merkitysholismia vastaan on ollut esim. Michael Dummett (1976) sekä Jerry Fodor ja Ernest Lepore (Fodor & Lepore 1992). Tähän keskusteluun ei kuitenkaan puututa tässä tutkielmassa. Ks. merkitysholismia ympäröineestä keskustelusta esim. Pagin tuleva ja Rorty 1993.

⁴⁸ Kinematiikka on mekaniikan osa, joka käsittelee kappaleiden liikettä puuttumatta liikkeen aiheuttajiin. Dynamiikka puolestaan on se mekaniikan osa, joka käsittelee juuri kappaleiden liikkeitä ja liikkeen aiheuttajia eli voimia.

tämän kyseisen systeemin pitää tietää etukäteen, jotta se voisi oppia B:n?” tai ”minkälaisia intentioita tällä systeemillä pitää olla, jotta se voisi tarkoittaa jotakin sanomalla jotakin?” Samalla tavalla kuin kinematiikka ei kerro liikkeiden aiheuttajista mitään, intentionaalisten systeemien teoria pysyy hiljaa kysyttäessä intentionaalisten systeemien toteutuksesta. (58–59.)

Dennett argumentoi seuraavasti sen puolesta, että intentionaalisten systeemien teorian neutraalisuus systeemin toteutukseen nähden onkin itse asiassa teorian vahvuus. Tarkastelemme intentionaalisen luonnehdinnan roolia evoluutiobiologiassa. Jos luonnonvalinnan tuottamia monimutkaisia käyttäytymiskykyjä ja kognitiivisia taitoja tahdotaan selittää, huomataan, että nimenomaan intentionaalisin termein luonnehdituilla kyvyillä (esim. kyvyllä omaksua uskomuksia) on selviytymisarvoa, vaikka ne toteutuvat mutaatioiden tuloksena. Kuvitellaan, että erittäin myrkyllinen hyönteinen ilmestyy ympäristöön. Siinä tapauksessa niillä hyönteisiä syöville linnuilla ja lepakoilla, jotka uskovat kyseisen hyönteislajin olevan huonoa syötävää, on selviytymisetu. Jos huomioidaan lepakkojen ja lintujen suuret erot esimerkiksi hermoston rakenteessa, perimässä ja havaintokyvyissä, voidaan todeta olevan hyvin epätodennäköistä, että tämä piirre on kuvattavissa yhtään alemmalla abstraktiotasolla kuin sillä, millä intentionaalisten systeemien teoria liikkuu. Koska teorian kuvaus on yleisempi tässä tapauksessa kuin esimerkiksi fysiologisen toteutuksen kuvaus, se on suositeltavampi. Sama pätee Dennetin mukaan yleisesti älykkyyteen, tarkoitukseen, referenssiin tai representaatioon. Jos haluamme sanoa näistä asioista jotakin merkityksellistä, tietty yleisyystaso on välttämätöntä. Jos esimerkiksi marsilaiset hyökkäisivät Maahan, kysyisimme onko niillä mielentiloja ja ovatko ne paljolti meidän kaltaisiamme. Intentionaalisen systeemin teorian mukaan vastaus on myönteinen, sillä jos marsilaiset ovat ylipäättään tarpeeksi terävä-älyisiä päästäkseen Maahan, niillä täytyy olla mielentiloja, oli niiden sisäinen rakenne minkälainen tahansa. (59–60.) Kaiken kaikkiaan intentionaalisten systeemien teoria olettaa systeemien,

joilla on samanlaisia uskomuksia ja haluja, toimivan samankaltaisesti, mutta ei oleta niiden olevan samanlaisia toteutukseltaan.

3.3.2 Alipersonallinen kognitiivinen psykologia ja dekompositio- menetelmä

Alipersonallisen kognitiivisen psykologian tutkimuskohteena on intentionaalisten systeemien sisäinen rakenne alijärjestelmien organisaationa. Kun intentionaalisten systeemien teoria tutkii intentionaalisia systeemejä holistisina kokonaisuuksina, alipersonallinen kognitiivinen psykologia hajottaa kokonaisuuden osiinsa. Alipersonallinen kognitiivinen psykologia tutkii kyseisten systeemien alajärjestelmiä ja niiden vuorovaikutuksellisia prosesseja, joiden tuloksena on intentionaalisen systeemin käyttäytyminen ja toiminta. Se on siis täysin funktionalistinen teoria. (BS, 153.)

Tässä tapauksessa alipersonallinen kognitiivinen psykologia siis tutkii mielen prosesseja toiminnan tasolla, joten se käsittelee tuttuun tapaan erillisiksi prosesseiksi nähtyjä järjestelmiä, kuten tarkkaavaisuutta, oppimista, muistia, kieltä ja muita niin sanottuja tiedollisia toimintoja.⁴⁹ Näiden alijärjestelmien sanotaan käsittelevän tietoa, viestittävän sitä ja tätä ja niin edelleen. Alijärjestelmistä puhutaan siis intentionaalisin termein. Alipersonallisen kognitiivisen psykologian tehtävänä on löytää hyviä teoreettisia entiteettejä, illatum-luokan entiteettejä, jotka muistuttavat vain löyhästi arkipsykologian entiteettejä, kuten uskomuksia, haluja ja niin edespäin.

Alipersonallisen kognitiivisen psykologian ottama ”älykkyyslainsäädäntö” (BS, 12) on maksettava kuitenkin takaisin, muussa tapauksessa persoonan alaisissa prosesseissa vaaniva homonculus (tai yleisemmin intentionaalisuus) ei katoa. Tekoälyn tutkimus näyttää, miten homonculus voidaan häivyttää. Tekoälyssä

⁴⁹ Alipersonallinen kognitiivinen psykologia on siis arkisemmin kognitiivinen psykologia. On kuitenkin syytä puhua tässä yhteydessä alipersonallisesta kognitiivisesta psykologiasta (niin kuin Dennett tekee), koska tarkoitus on painostaa ”kokonaisuuden osia”.

käytetään yleisesti menetelmää, jossa intentionaalinen ongelma pilkotaan osiin. Tätä kutsutaan *dekompositioksi*. Siinä aluksi tarkasteltavan järjestelmän toimintoja ajatellaan tekevän joukko homonculuksia, jotka korvataan pienemmillä ja yksinkertaisimmilla homonculuksilla, kunnes päädytään homonculuksiin, jotka ovat niin yksinkertaisia – jolloin niiden ei tarvitse vastata niille esitettyyn kysymykseen kuin kyllä tai ei –, että ne voidaan korvata automaattisilla mekanismeilla (BS, 124).⁵⁰

Milloin intentionaalisuus katoaa dekomposition pilkkomisprosessissa, on vaikeampi kysymys. Dennett kehottaa olemaan kysymättä tällaista kysymystä, koska on turha vaiva selvittää, missä vaiheessa jokin aito ominaisuus ilmaantuu. (Dennett 1994b.) Kyse on nimittäin pelkästään määrällisestä erosta.

3.3.3 Semantiikan luominen syntaksista

Kun siirrytään intentionaalisesta tasosta alaspäin dekompositiossa ajaudutaan lopulta ekstensionaalisen ja intentionaalisen kuvauksen välisen suhteen ongelmaan. Tällöin ongelmana on, toisin termein ilmaistuna, miten syntaksista voi syntyä semantiikkaa.

Koska aivot ovat evoluution tuotos, niillä on jokin tarkoitus, jonka takia ne ovat valikoituneet. Dennettin sanoin aivot ovat *semanttinen kone* (engl. semantic engine), jonka tehtävänä on, karkeasti sanottuna, ”selvittää, mitä sen moninaiset syötteet *merkitsevät*, erottaa ne niiden merkitysten perusteella ja ’toimia sen mukaisesti’” (61).⁵¹ Tätä tehtävää varten aivot ovat valikoituneet, ja

⁵⁰ Kannattaa kuitenkin muistaa, että tällä menetelmällä pilkotaan osiin (ihanteellisella tavalla) ainoastaan ongelma. Biologisten systeemien tapauksessa ei voida olettaa, että vastaavanlainen täydellinen pienempiin osiin pilkkominen toimisi, sillä luonnonvalinta ei voi katsoa kauas tai suunnitella sinällään eliöitä, jolloin eliöt eivät ole ihanteellisia tietyn ongelman ratkaisijoita. (Dennett 1991a.) Lisäksi on huomattava, aivan kuten Marr (1982) toteaa puhuessaan komputationaalisen tason kuvauksesta, että intentionaalisen tason kuvaus on oltava oikea, jotta dekompositio-urakka olisi mielekäs. Mikäli intentionaalisen tason kuvaus on pielessä, siitä käsin ei voida päätyä mielekkäisiin alemman tason kuvauksiin, koska ollaan lähtökohtaisesti väärällä polulla.

⁵¹ Dennett käsittelee tarkoin aivojen semanttisuuden jäljittelyä teoksensa Dennett 1969 kolmannessa osassa ”Evolution in the Brain”.

tältä ne näyttävät intentionaalisten systeemien teoriasta käsin. Kuten luonnontieteet meille kertovat, aivot ovat kuitenkin vain *syntaktinen kone* (engl. syntactic engine), sillä "[aivot] pystyvät ainoastaan erottelemaan niiden syötteet syötteiden rakenteellisten, ajallisten ja fysikaalisten piirteiden perusteella ja antamaan näiden syötteiden 'syntaktisten' piirteiden hallita [aivojen] täysin mekaanisia toimintoja" (61). Tämä on kaikki, mitä aivot voivat tehdä. Miten aivot sitten voivat saada syntaksista aikaan semantiikkaa? Ei mitenkään, vastaa Dennett.

Aivot voivat pelkästään approksimoida semanttista konetta ja jäljitellä semanttisen koneen käyttäytymistä. Jäljittely onnistuu käyttämällä hyväksi tarpeeksi lähelle osuvia onnekkaita vastaavuuksia rakenteellisten säännönmukaisuuksien ja semanttisten tyyppien välillä. Rakenteellisia säännönmukaisuuksia on sekä ympäristössä sekä aivojen sisäisten tiloissa ja operaatioissa. (61.)

Tämä idea ei ole ainutlaatuinen. Eläimen pitää tietää, milloin se on oikeasti saanut ravintoa, kun se on jotain syönyt. Ravinnon saamisen semanttinen tulkinta jää kuitenkin sellaisen syntaktisen mekanismin vastuulle, joka aktivoituu, kun eläimen kurkussa syntyy kitkaa ja kun sitä seuraa vatsalaukun venyminen. Tämän mekanismin aktivoituminen tapahtuu normaalisti samanaikaisesti eläimen "oikean", semanttisen tavoitteen kanssa. (61.) Ottakaamme asiasta toinen esimerkki, tällä kertaa ajankohtainen esimerkki tietokoneiden maailmasta. Kysymys kuuluu: miten roskaposti voidaan erottaa muista sähköpostiviesteistä joillakin automaattisilla menetelmillä? Roskapostiksi luokiteltava viesti vaatii semanttista tulkintaa, jotta luokittelu olisi täydellistä, sillä mitään tiettyä yhteistä säännönmukaisuutta roskaposteilla ei ole. Miten luokittelemme viestit roskaposteiksi jollakin yksinkertaisella mekanismilla?⁵² Kuvitellaan nyt esimerkin vuoksi, että roskaposti suodatetaan

⁵² Voisi sanoa, että roskapostin semantiikka pitäisi operationalisoida jollakin tavalla. Siinä missä tietojenkäsittelyssä ja tekoälyssä pyritään luomaan syntaksista semantiikkaa,

ainoastaan viestin sisällön perusteella.⁵³ Ensiksi on tarkastelumme rajoitettava johonkin kieleen. Olkoon se englannin kieli siitä syystä, että useimmat roskapostiviestit ovat englanninkielisiä. Voisimme analysoida joitakin roskapostiviestejä ja etsiä niistä roskapostiviesteille yhteisiä ominaisuuksia. Oletetaan, että löydämme usein toistuvia sanoja, kuten ”price”, ”sale”, ”hot”, ”sex”, ”adult” ja symboleja, kuten huutomerkkejä ja dollarimerkkejä. Eräs mekanismi voisi roskapostin suodatukseen olla siten seuraavanlainen: kaikki sähköpostiviestit ovat roskapostia, jos niissä esiintyy

- sekä sanat ”price” ja ”sale” että huutomerkki jossain kohtaa viestiä tai
- ennen pistettä esiintyvät sanat ”hot” ja ”online” tai
- ...

Dennettin mukaan tämänkaltaisen perusstrategian variaatiot ovat ainoa mahdollisuus, jos syntaksista tahdotaan luoda semantiikkaa (62). Jotta tällaisia strategioita voitaisiin nimittää intentionaalisiksi, täytyy teoretikon (nimeäjän) tarkastella systeemeitä ulkopuolelta nähdäkseen, mikä tuottaa strategian kaltaisen asetelman, mitä vaikutuksia systeemin vasteilla on normaalisti sen ympäristössä ja mitä hyötyä systeemille karttuu tästä toiminnasta. Dennettin mukaan neurofysiologi ei voi tutkia onnistuneesti intentionaalisen systeemin realisaatiota, jos hän hylkää semanttisen tulkinnan. Neurofysiologi, joka tällöin tutkii ainoastaan ihmisen aktuaalista sisäistä rakennetta hylkäämällä ihmisen ympäristön ja ympäristön yhteydet ihmisen sisäiseen rakenteeseen, tutkii pelkkää sisäistä syntaksia neurofysiologian valossa – ja tästä ei voida saada semanttista tulkintaa. (63–64.)

psykologiassa yritetään tutkia semantiikan muodostumista syntaksista, ts. holistisen intentionaalisen järjestelmän muodostumista sen intentionaalisista ja ei-intentionaalisista alajärjestelmistä. Ks. BS, luku 7 ”Artificial Intelligence as Philosophy and as Psychology”.

⁵³ Esimerkkini on tietysti rajusti yksinkertaistettu. Nykyiset roskapostinsuodattajaohjelmistot luottanevat luokittelussa pääasiassa viestin metatietoihin, kuten otsikkotietoihin.

[J]os he [fysiologit] eivät pysty näkemään neuraalisia tapahtumia *signaaleina, raporteina* tai *viesteinä*, heille ei jää lähes minkäänlaista näkemystä aivojen funktioista. Jos fysiologi kieltäisi kaiken intentionaalisen värityksen kuvauksestaan aivojen funktiosta, hänen teoriansa olisi *parhaassa tapauksessa* seuraavan muotoinen: funktionaalaisella rakenteella A on funktio funktionaalisen rakenteen B ärsykkeenä joka kerta, kun sen [A:n] ärsykkeenä toimii joko C tai D... Minkään suuruinen määrä tämänkaltaista tarinaa ei koskaan vastaa sellaisiin kysymyksiin, kuten miksi rotta A pelkää rotta B:tä tai miten rotta A tietää, mihin suuntaan sen pitäisi mennä saadakseen ruokaa. (Dennett 1969, 79.)

Tämä jälkimmäinen tarina [mitä eläin tekee tai miksi se tekee niin kuin tekee] voidaan kertoa vain intentionaalisin termein, mutta se ei ole tarina maailman ominaisuuksista ekstensionaalisen tarinan ominaisuuksien *lisäksi*; se vain kuvaa, mitä tapahtuu eri tavalla. (Dennett 1969, 78.)

Dennettin mielestä psykologian tulisikin näyttää, miten aivot toteuttavat intentionaalisia tiloja ottamalla huomioon semanttinen tulkinta ja siten eliöt suhteessa niiden ympäristöön.⁵⁴ Intentionaalisia tiloja ei voida kuvata viittaamatta systeemin ulkopuolelle. (63–64.)

Samalla tavalla voidaan ajatella suunnitelluista artefakteista. Tutkittaessa jonkin artefaktin toteutusta on otettava huomioon, mitä tarkoitusta ja tehtävää varten artefakti on luotu palvelemaan. Puhelinjärjestelmää tutkittaessa on hankala ymmärtää, miksi järjestelmän kaistanleveys on 3100 hertsiä, ellei huomioida sen tarkoituspää. Se on rakennettu ihmisten välistä kommunikaatiota ja nimenomaan puheensirtoa varten. Siksi 3100 hertsiä on sopiva kaistanleveys; taajuusalueella 300–3400 hertsiä puhe on vielä ymmärrettävää.

3.3.4 Reduktion mahdollisuus

On selvää, että Dennett tahtoo korvata tieteessä käytetyn arkipsykologian intentionaalisten systeemien teoriallaan. Dennettin mukaan arkipsykologia, tai

⁵⁴ Tässä ympäristöllä ei tarkoiteta pelkästään sosiaalista ympäristöä, sillä se kohdistaa katseen ainoastaan kommunikoiviin intentionaalisiin systeemeihin. Kyse on intentionaalisisista systeemeistä yleensä, joten huomio tulisi kiinnittää yleisemmin organismin ja ympäristön vuorovaikutukseen, kehitykseen ja evoluutioon.

tarkemmin sanottuna ”se osa arkipsykologiasta, josta kannattaa välittää” palautuu eli *redusoituu* käsitteellisesti intentionaalisten systeemien teoriaan (66). Verrataan seuraavaksi keskenään intentionaalisten systeemien teorian ja kahden muun ehdotuksen mukaista intentionaalisuuden käsitteellistä reduktiota ekstensionaaliseen kieleen. Nämä kaksi muuta ehdotusta tulevat tyyppi-identiteettiteorialta ja konefunktionalismilta, jotka ovat paljolti hallinneet perinteistä mielenfilosofiaa.

Tyyppi-identiteettiteorian (engl. type-identity theory; esim. Place 1956 ja Feigl 1958) mukaan jokaiselle mentaalille termille tai predikaatille ”M” on olemassa jokin predikaatti ”P”, joka on ilmaistavissa vain fyysikaalisten tieteiden sanastolla, niin että jokin on M, jos ja vain jos se on P. Symbolisin termein ilmaistuna:

$$(x) (Mx \equiv Px).$$

Toisin sanoen tyyppi-identiteettiteoria väittää jokaisella henkilöllä tai oliolla, jolla on jokin mentaalinen tila M, olevan sama neurofysiologinen tila P.⁵⁵ Tämä väite on nykyään yleisesti hylätty. Ei ole nimittäin mitään syytä olettaa, että kaksi henkilöä (tai intentionaalista systeemiä) olisivat samassa fyysikaalisessa tilassa, jos heillä on sama mielentila, esimerkiksi uskomus, että ulkona sataa.⁵⁶ *Konefunktionalismi* välttää tällaisen naiiviuden väittämällä, että

$$(x) (Mx \equiv x \text{ toteuttaa jonkin Turingin koneen } k, \text{ joka on loogisessa tilassa } A).$$

⁵⁵ Itse asiassa tyyppi-identiteettiteoria ei vastaa intentionaalisuuden ongelmaan, koska se yrittää palauttaa intentionaaliset termit suoraan fyysikaalisiin termeihin. Suoran redusoimisen mahdottomuus nähtiin jo toisessa luvussa puhuttaessa intentionaalisuudesta yleisesti. Se voidaankin nähdä pääasiassa esityksenä mentaalisen ja fyysisen kielen väliselle reduktiolle, kun intentionaalisuutta ei huomioida.

⁵⁶ Selvimmin tyyppi-identiteettiteorian väittämän virheellisyys näkyy yksinkertaisen kelloa käsittelevän analogian kautta. Kaikki kellot ovat varmasti fyysikaalisia, mutta mitä fyysisesti yhteistä digitaalisella kellolla, vedettävällä seinäkellolla ja aurinkokellolla on, kun ne näyttävät kellonaikaa 12:00 (vrt. BS, xv)?

Konefunktionalismin väite on muuten samanlainen kuin tyyppi-identiteettiteorian, mutta se sanoo, että jokaista mentaalista tilaa vastaa jokin Turingin koneen (Turing 1950) tila. Koska tietyn Turingin koneen tilan voi toteuttaa monta erilaista fysikaalista realisaatiota, konefunktionalismia voidaan liittää myös esiintymäidentiteettiteoria (engl. token identity theory). Dennettin mielestä konefunktionalisminkin väite on liian vahva, sillä

ei ole suurempaa syytä olettaa, että Jacquesilla, Sherlockilla, Boriksella ja Tomilla 'on sama ohjelma' *missään* irtonaisessa tai abstraktissa mielessä ottaen huomioon erot heidän perimässään ja kasvatuksessaan, kuin että heidän aivoissaan olisi jokin ratkaisevan identtinen psykokeemiallinen ominaisuus.⁵⁷ (66.)

Dennett ehdottaakin tilalle omaa "reduktionistista" näkemystään, jonka mukaan pätee

(x) (x:llä on intentionaalinen tila $S \equiv x$:lle voidaan ennustavasti attribuoida intentionaalinen tila S).⁵⁸

⁵⁷ Ks. Dennettin esittämistä konefunktionalismin ongelmista tarkemmin luku 2 "Reply to Arbib and Gunderson" teoksessa BS. Kyseisessä luvussa Dennett kyseenalaistaa konefunktionalismia laajemmin ja analyttisemmin. Hän esimerkiksi kyseenalaistaa konefunktionalismin väitteen tarkasteltavan systeemin tilojen tai tapahtumien yksilöinnistä intentionaalisten tilojen suhteen:

Kun liitämme [engl. ascribe] systeemille yksittäisen uskomuksen, että p, meidän täytyy [...] liittää systeemille loputon määrä yhä enemmän muita uskomuksia, haluja jne. Vaikka epäilemättä *jotkin* näistä [uskomuksien] liittämistä tukevat hyvin systeemin suunnittelun keskeisiä ominaisuuksia, jotkut toiset [uskomuksien] liittämiset eivät [tue], vaikka systeemin käyttäytyminen onkin kaiken kaikkiaan niin säänneltyä, että se oikeuttaa nuo [uskomuksien] liittämiset. Jokaiselle äärettömästä joukosta poimitun [uskomuksen] liittämiseksi ei tarvitse olla, eikä voi olla, erikseen erotettavaa mekaanisten osien tilaa, ja siten useissa tapauksissa ei ole mahdollista erottaa mitään systeemin ominaisuutta millään abstraktiotasolla ja sanoa, että 'Tämä ja vain tämä ominaisuus suunnittelussa on vastuussa sen noista käyttäytymisen aspekteista, joiden takia liitämme sille uskomuksen, että p'. Ja siksi, johtuen siitä tosiasista, että molempia systeemejä S ja T luonnehditaan siten, että ne uskovat, että p, ei seuraa, että ne ovat molemmat jossakin muussa kuin sen uskomisen, että p [...] erotettavissa olevassa tilassa. (Siitä syystä, kun S ja T ovat samassa uskomustilassa, ei sen tarvitse merkitä niiden olevan samassa loogisessa tilassa, jos tulkitsemme jälkimmäisen esityksen joksikin Turingin koneen tilaksi jollekin jaetulle Turingin kone -tulkinnaalle, sillä niiden ei tarvitse jakaa mitään relevanttia Turingin kone -tulkinnaa.) (BS, 26.)

⁵⁸ Dennett itse kirjoittaa ehdotuksensa seuraavasti (67): (x) (x uskoo, että $p \equiv x$:lle voidaan ennustavasti attribuoida uskomus, että p). Pääasia kuitenkin on, että puhutaan jostakin termistä, joka määritellään intentionaalisin termein (vrt. Dennett 1995b).

Siis: systeemillä on jokin mielentila (uskomus, halu tms.), jos ja vain jos sille voidaan ennustavasti attribuoida kyseinen mielentila. Epäilemättä tämä vaikuttaa häpeämättömältä tautologialta, kuten lause ”hevonen on eläin, johon termi ’hevonen’ pätee” (BS, xvii). Dennettin mielestä tämä ei kuitenkaan eroa konefunktionalismin väitteestä. Nimittäin, Dennett kysyy: ”mitä yhteistä kahdella Turingin koneen realisaatiolla tai ilmentymällä on, kun ne ovat samassa loogisessa tilassa?” (67) ja vastaa:

Vain tämä: on olemassa kuvausjärjestelmä, jonka mukaan molemmat kuvataan jonkin tietyn Turingin koneen realisaatioina, ja tämän kuvauksen mukaan, joka on ennustava molempien entiteettien toiminnan suhteen, molemmat ovat samassa Turingin koneen konetaulun tilassa. (67.)

Turingin kone –puhetta ei toisin sanoen redusoida toiseen muotokieleeseen, vaan se oikeutetaan varustamalla se attribuutiosäännöillä ja osoittamalla sen ennustusvoima. Samalla tavalla intentionaalisten systeemien teorian tarkoituksena on oikeuttaa mielentiloja koskeva puheemme. (67.)

Dennett yhtäältä siis väittää, että mitään sellaista kuin intentionaalisuus tai semantiikka ei ole olemassa ja toisaalta tahtoo oikeuttaa puheen intentionaalisuudesta. Miksi? Syynä on intentionaalisten idiomien käytön erinomaisuus instrumentalistisena strategiana. Dennett ei siis redusoi intentionaalisten systeemien teoriaa alipersonalliseen kognitiiviseen psykologiaan ja eteenpäin ekstensionaaliseen kieleen. Pikemminkin reduktion suunta on päinvastainen: tarkoitus on ”osoittaa, miten systeemi kuvattuna fysiologisina termeinä voisi oikeuttaa tulkinnan realisoituna intentionaalisena systeeminä” (68). Intentionaalisuutta ei siten naturalisoida eli palauteta ekstensionaalisuuteen. Tämän antireduktionistisuuden takia hänen näkemystään on alettu riepottaa ja syyttää sitä milloin instrumentalistiseksi⁵⁹ ja

⁵⁹ William Lycan (1990, 147) määrittelee instrumentalismin näin:

Instrumentalisti X:ien suhteen on teoreetikko, joka väittää, että vaikka lauseet ”X:istä” ovat usein tosia, ne eivät oikeasti kuvaa mitään erityisiä entiteettejä, vaan ainoastaan auttavat systematisoimaan tuttuja ilmiöitä enemmän.

milloin epifenomenalistiseksi⁶⁰. Tähän paneudutaan tarkemmin seuraavassa luvussa.

⁶⁰ Epifenomenalismi on näkemys, jonka mukaan mentaaliset tapahtumat ovat aivojen fyysikaalisten tapahtumien aiheuttamia, mutta mentaaliset tapahtumat eivät itsessään vaikuta fyysikaalisiin tapahtumiin (ks. esim. Robinson 2003).

4 DENNETTIN NÄKEMYKSEN ERITYISKYSYMYKSET

Edellä esitettiin Dennettin intentionaalisuuden käsityksen ydin, mutta sitä on syytä täydentää tietyiltä osin. Lisäksi on hyvä tarkastella ja pohtia syvemmin tiettyjä, keskustelua herättäneitä erityisteemoja.

Käsitystä täydennetään tekemällä ero uskomuksen ja mielipiteen välillä ihmisistä puhuttaessa (luku 4.1 ”Uskomuksen ja mielipiteen erottelu”). Erottelu on tärkeä, sillä Dennettin mukaan uskomus ja mielipide usein sekoitetaan keskenään, mikä saattaa ihmiset hämmennykseen ja ongelmiin. Seuraavaksi tarkastellaan intentionaalisen asenteen tekemää rationaalisuusoletusta ja rationaalisuuskäsitettä (luku 4.2 ”Rationaalisuusoletus”). Kolmanneksi käsitellään lyhyesti kritiikkiä kolmannen persoonan lähtökohdan ongelmallisuudesta (luku 4.3 ”Kritiikki kolmannen persoonan mukaisesta lähestymistavasta”) ja (luku 4.4 ”Kiista alkuperäisestä ja johdetusta intentionaalisuudesta”) intentionaalisen asenteen käyttämän puheen näennäisyydestä. Lopuksi tutkitaan kohua herättänyttä Dennettin instrumentalismilta vaikuttavaa käsitystä intentionaalisten tilojen ontologisesta asemasta (luku 4.6 ”Intentionaalisten tilojen ontologinen asema”). Ennen intentionaalisten tilojen ontologista tarkastelua käännetään pää silti intentionaalisten tilojen ja funktioiden määrittämättömyyteen (luku 4.5 ”Intentionaalisten tilojen ja funktioiden määrittämättömyys”), koska se liittyy läheisesti intentionaalisten tilojen ontologiseen asemaan.

4.1 Uskomuksen ja mielipiteen erottelu ⁶¹

Dennett väittää, että filosofit ovat kautta aikain kompastelleet uskomuksen ja halujen ongelmien kanssa, mistä esimerkkinä ovat debatit tahdon heikkoudesta eli akrasiasta, koska he eivät ole tehneet eroa uskomuksen ja halun sekä sen välille, jota Dennett kutsuu mielipiteeksi. (BS, 306.) Perinteinen käsitys

⁶¹ Tämä luku pohjautuu pääasiallisesti teoksen BS lukuun 16 ”How to Change Your Mind”.

uskomuksesta, jota usein propositionaaliseksi asenteeksi nimitetään, on Dennettin mukaan mielipide eikä oikea uskomus (BS, 306).

Dennett kertoo, että *mielipiteet* (engl. opinion)

ovat kielellisesti tarttuneita, suhteellisen sofistikoituneita kognitiivisia tiloja – *karkeasti sanoen* tiloja, joissa lyödään vetoa yksittäisen, formuloidun lauseen totuudesta.⁶² (19.)

Näin ollen kaikki kielelliset tilat, joita on usein pidetty uskomuksina ja haluina, ovat ainoastaan mielipiteitä. Esimerkiksi ilmaisu ”tahdon syödä härkää punaviinikastikkeessa” on mielipide eikä halu. Kun nyt ravintolassa ollessani tarjoilija kirjoittaa muistiin tämän ”halun” (eli mielipiteen) tilauksena, tulenkin toisiin ajatuksiin ja tilaan Pekingin ankkaa. Nyt yleensä ajatellaan, että haluni on muuttunut (tai että olen muuttanut mieltäni). Dennettin mukaan näin ei ole; voidaan sanoa ainoastaan, että mielipiteeni on muuttunut. Haluni on sama, joka se on ollut koko ajan, mutta arvioni, oma tulkintani halustani on muuttunut (BS, 303).

Edellisen kaltaiset tahdon ilmaukset ovat pikemminkin valintoja kuin haluja. Valinnat kuitenkin tehdään uskomuksien ja halujen pohjalta, mutta itsessään valinnat eivät ole uskomuksia eivätkä haluja. Valintoja seuraavat teot, joita voidaan kutsua arvioinneiksi tai sitoutumisiksi. Näitä tekojakaan ei voida kutsua uskomuksiksi tai haluiksi. (BS, 303.)

Eläimillä ei ole mielipiteitä, niillä on vain uskomuksia ja haluja. Eläimillä ei ole mielipiteitä siksi, koska niillä ei ole kieltä. (BS, 306.) Meidän kielellinen ympäristömme taas pakottaa meidät uskomuksiemme ja halujemme sanalliseen ilmaisuun eli mielipiteidemme ilmaisuun (21).⁶³

⁶² Dennettin jako uskomuksen ja mielipiteen välillä muistuttaa tietyssä mielessä erottelua niin sanotun propositionaalisen ja hiljaisen (engl. tacit) tiedon välillä. Ks. tästä jaottelusta esim. Polanyi 1966. Gilbert Ryle (1949) esitti vastaavan kuin Michael Polanyi jaon ”tietää että” (engl. know that) ja ”tietää miten” (engl. know how) –tiedon välille.

⁶³ Erilaiset kielet pakottavat meitä muodostamaan eritarkkuuksisia mielipiteitä tai tulkintoja uskomuksistamme. Esimerkiksi suomessa voimme käsitteellistää monia eri lumityyppejä, kun muissa, erityisesti eteläisimmissä kielissä eri lumityyppejä tuskin erotellaan. Olisi siis

Kun arkipsykologiassa puhumme ajattelusta ja omista uskomuksistamme, puhumme eri asiasta kuin mistä Dennett puhuu. Ajattelu on ihmisten sosiaalinen, keinotekoinen prosessi (Dahlbom 1993b ja Dennett 1993, 231). Se ei kuvasta tosiasiallisia prosesseja, eivätkä ajattelun käsittelemät kielelliset lauseet kuvasta uskomuksia ja haluja.⁶⁴

Koska henkilöllä ei ole mitään suoraa pääsyä omiin uskomuksiinsa ja haluihinsa, hänen mielipiteensä ovat vain tulkintoja hänen uskomuksistaan ja haluistaan.⁶⁵ Ne eivät siis voi olla varmaa tietoa, vaan ainoastaan hänen

kovin kummallista olettaa, että vaikkapa perulainen uskoisi Helsingistä marraskuussa otetun valokuvan nähdessään, että kuvassa (suomalaisille) esiintyvä loska on juuri "loskaa" eikä vain "lunta".

⁶⁴ Dennett kritisoi mielen kielen (engl. language of thought; Fodor 1975) kannattajia varoittamalla, että "keskittymällä tämänkaltaisen sosiaalisen pakon *tuloksiin* huomioimatta sen vääristelevää vaikutusta, ajaudutaan helposti ajattelemaan uskomuksien ja halujen olevan *ilmiselvästi* kutakuinkin lauseita tallennettuina päähän" (21). Perinpohjaisesti hän on käsitellyt mielen kieltä teoksessaan BS, luvussa 6 "A Cure for the Common Code?" Yleensä ottaen Dennett kritisoi Fodorin mielen kielen näkemystä erityisesti eksplisiittisistä representaatioista (esim. tiedän, ettei seeproilla ole päälystakkaa; olen varmasti tiennyt tämän jo lapsuudestani asti, mutta ei ole mielekästä olettaa, että sellainen hypoteesi olisi muodostunut lapsuudessani, edes minun sitä tiedostamatta [BS, 104]), jyrkän eron tekemisestä komputationaalisten ja muiden prosessien välille (esim. "[o]nko kalan ja sukelluskellon prosessien välillä tärkeää laadullista eroa? [BS, 99]) sekä alipersonallisen ja persoonan tason sekoittamisesta (esim. Fodor ei pelkää homonculuksia käsitteessään sisäisen koodin viestit itseymmärtävinä representaatioina [BS, 102]).

⁶⁵ Se, että henkilö ei ole suoraan tietoinen omista intentioistaan saattaa hämmästyttää joitakin ihmisiä. Dennett on tähän, jota voisi kutsua introspektiiviseksi harhaksi, puuttunut näkemysensä tietoisuudesta. Teoksen Dennett 1969 luvussa 14 "The Ordinary Words" Dennett on tehnyt seuraavankaltaisen jaon (Dennett 1969, 118–119):

- (1) A on tietoinen₁ [engl. aware], että p ajanhetkenä t, jos ja vain jos p on A:n "puhekeskuksen" syötetilan sisältö ajanhetkenä t.
- (2) A on tietoinen₂, että p ajanhetkellä t, jos ja vain jos p on A:n sisäisen tapahtuman, joka vaikuttaa nykyisen käyttäytymisen ohjaamisessa, sisältö ajanhetkellä t.

Tämän jaottelun mukaisesti ei ole niin, että henkilöt voisivat olla erehtymättömästi tietoisia uskomuksistaan ja haluistaan mielessä 1, koska ihmisillä ei ole Dennettin mukaan pääsyä omiin sisäisiin, käyttäytymistä ohjaaviin tekijöihinsä (esim. BS, 304–305). Arkipuheessa sekoitetaan Dennettin mielestä nämä kaksi tietoisuuskäsitystä keskenään (Dennett 1969, 119). Jako on täysin yhteneväinen Dennettin tekemän erottelun persoonan tason ja alipersonallisen tason selityksiin, jossa tietoisuus₁ on persoonan tason käsite ja tietoisuus₂ alipersonallisen käsite. Hän on kuitenkin myöhemmin hylännyt tämän jaottelun tietoisuuden muotojen välillä, koska se on hänen mielestään liian karteeseinen tai essentialismimainen (ks. luku 2 "Reply to Arbib and Gunderson" teoksessa BS, Dennett 1999 ja Dennett 1994a, 545–551). Tästä syystä hän ei tee jaottelua reflektiivisen tai representationaalisen tietoisuuden ja fenomenaalisen tietoisuuden välille, mikä vaivaa Ned Blockia (1993). Karteesisuuden välttämiseksi ero myöskään alipersonallisen kognitiivisen psykologian ja intentionaalisten systeemien teorian välillä ei ole ehdoton tai edes jyrkkä (esim. 64 ja Dennett 1994a, 530).

arviointejaan omista uskomuksistaan ja haluistaan, joiden totuudenmukaisuudesta henkilö ”lyö vetoa”. (BS, 304–305.) Pahimmassa tapauksessa omien intentioidemme selittäminen onkin puhdasta sepittelyä eli konfabulointia kuten esimerkiksi Korsakovin oireyhtymää kärsivien ihmisten tai aivohalkiopotilaiden tapaukset osoittavat (91). Vain uskomukset ja halut ennustavat henkilön käyttäytymistä suoraan; intentionaalisille systeemeille attribuoidaan vain uskomuksia ja haluja eikä lainkaan mielipiteitä.

Akrasia on mahdollista ihmisillä, mutta ei eläimillä. Uskomukset ovat nimittäin automaattisesti yhdenmukaisia käyttäytymisen kanssa. Mielipiteet ovat kylläkin yleensä yhdenmukaisia uskomuksien ja halujen kanssa, mutta uskomuksia ja haluja on mahdoton määrittää yhtä tarkkaan, kuin mitä kieli mahdollistaa ja vaatii. (21 ja BS, 306–307). Tällöin erehdyksiä omien uskomuksien suhteen sattuu ja ”mieli muuttuu”.

4.2 Rationaalisuusoletus

Uskomusteoria ilman rationaalisuusrajoituksia on vailla ennustussisältöä; sitä käyttämällä, meillä ei voi olla miltei lainkaan odotuksia koskien uskojan käyttäytymistä. (Cherniak 1981, 164.)

Intentionaalisen puhutavan käyttö tekee rationaalisuuden oletuksen toimiakseen. Rationaalisuus on siis epäilemättä intentionaalisen asenteen kulmakivi, joten sen käsitteeseen on syytä paneutua syvemmmälle.

Rationaalisuuskeskustelussa on tavanmukaista puhua normatiivisista ja deskriptiivisistä rationaalisen päättelyn malleista. Normatiivinen rationaalisuuden malli antaa periaatteita, joiden mukaan meidän pitäisi päätellä. *Standardin rationaalisuuskuvan* (engl. standard picture of rationality; Stein 1996) mukaan on rationaalista päätellä niiden periaatteiden mukaisesti, jotka pohjautuvat muun muassa deduktiiviseen logiikkaan ja tilastotieteeseen.⁶⁶

⁶⁶ Rationaalisuuden normatiivisia malleja löytyy myös esimerkiksi toiminnan filosofiasta, peliteoriasta, taloustieteestä ja moraalifilosofiasta. Tietyissä, yhdistävässä mielessä rationaalisuutta on näissä kaikissa normatiivisissa malleissa sanottu (jonkin hyödyn) maksimoinniksi. Keskityn tässä erityisesti rationaalisuuteen päättelyssä.

Näiden tieteiden on nähty, perinteisessä epistemologiassa, olevan tosia sisällään, riippumatta meistä.⁶⁷

Deduktiivinen logiikka pitää sisällään sellaisia sääntöjä kuin modus ponens, joka sanoo:

1. *P*.
2. *Jos P, niin Q*.
3. Siis *Q*.

Tällaiset säännöt antavat aiheen normatiivisille päättelyn periaatteille, kuten modus ponens –periaatteella, joka vastaavasti sanoo, että

1. jos uskon, että *P*, ja
2. uskon, että *P, niin Q*,
3. pitää minun uskoa, että *Q*.

Esimerkiksi jos minulla on uskomus ”jos on aamupäivä, on valoisaa”, ja uskon, että on aamupäivä, pitää minun uskoa myös olevan valoisaa. Logiikan päättelysäännöt vaativat henkilön aina uskovan kaikki uskomuksiansa loogiset seuraukset ja muuttavan uskomuksiaan ristiriitaisuuksien ilmetessä. Kuten sanottu, kaikki päättelyperiaatteet eivät pohjaudu logiikkaan. Tilastotieteen mukaan esimerkiksi tapahtumien A ja B yhdisteen todennäköisyys on pienempi kuin yksinään tapahtuman A tai B todennäköisyys. Myös tällaisista säännöistä on johdettu normatiivisia päättelyperiaatteita. (Stein 1996, 4–7.)

Kuten laaja kirjallisuus (ks. esim. Baron 1998, Kahneman ja muut 1982) osoittaa, normatiivisten ja deskriptiivisten rationaalisuuden mallien välillä on aukko. Deskriptiiviset eli empiiriseen tutkimukseen perustuvat tutkimukset ja mallit

⁶⁷ Perinteinen epistemologia ei ymmärrettävästi ole nykyään enää suosiossa. Naturalistisessa epistemologiassa, johon myös Dennett sitoutuu, logiikka ja muut vastaavat – perinteisen terminologian mukaan aprioriset – tieteet eivät ole missään autonomisessa asemassa suhteessa suoraviivaisiin empiirisiin tieteisiin. Tästä syystä tieteellistä evidenssiä käytetään naturalistisessa epistemologiassa osana sitä prosessia, joka määrää, miten meidän tulisi omaksua uskomuksia (ks. Stein 1996, 16).

näyttävät, että ihmiset eivät suinkaan aina päätele normatiivisten periaatteiden mukaan. Ihmisiä ajaa helposti päättelyvirheisiin erilaiset ennakkoluulot, vinoumat ja heuristiikat (esim. Kahneman ja muut 1982). Näiden selvästi normatiivisesta rationaalisuudesta poikkeavien tapauksien johdosta on ollut paljon keskustelua siitä, onko ihminen irrationaalinen. Kyseisiä poikkeavuuksia ei voida palauttaa normaaliin päättelyyn vaikuttaviin hetkittäisiin häiriötekijöihin tai suoritusvirheisiin (engl. performance errors) niiden ollessa systemaattisia, mikä viittaa perustavanlaatuisen pystyvyyden (engl. competence) riittämättömyyteen (esim. Stein 1996). Näitä tuloksia voidaan tulkita eri tavoin, ja monesti pitäydytäänkin siinä käsityksessä, että kokeet ovat osoittaneet ainoastaan ihmisen suoritusvirheitä, jolloin väite ihmisen irrationaalisuudesta ei saa tukea (ks. esim. Stein 1996).

Rationaalisuusteesin puolesta Dennett ja moni muu nojaa evoluutioon.⁶⁸ Evoluutio takaa, että olemme pääpiirteittäin normatiivisen mallin mukaisesti rationaalisia. Me (ja muut organismit) noudatamme logiikan sääntöjä pääpiirteittäin, ja suurin osa uskomuksistamme on totta luonnonvalinnan seurauksena. Dennettin sanoin: ”[k]yvyyllä uskoa ei olisi selviytymisarvoa, ellei se olisi kyky uskoa totuuksia”, ja ”[v]esiliukoinen kala on evolutiivinen mahdottomuus, mutta systeemille vääriä uskomuksia varten ei voida antaa edes koherenttia kuvausta” (BS, 17). Meidän kykymme olla vuorovaikutuksessa ympäristön kanssa ei onnistuisi, ellei havaintomme ja uskomuksemme vastaisivat totuudellisesti maailman rakennetta ja sen lainalaisuuksia – toisin sanoen ellemme approksimoisi normatiivisia periaatteita.⁶⁹

Tietenkään evoluutio ei takaa, että aivan kaikki uskomuksemme ja päättelymekanismimme ovat totuudenmukaisia. Tähän on ainakin kaksi syytä:

⁶⁸ Edward Steinin (1996, 26) mukaan näin tekee myös Jerry Fodor, Alvin Goldman, William Lycan, Ruth Millikan, David Papineau, Karl Popper, W.V.O. Quine ja Elliot Sober.

⁶⁹ Tätä kuvastaa hyvin Dennettin käyttämä termi ”tosi uskoja” (engl. true believer) intentionaalisen systeemin synonyymina teoksessa Dennett 1987, luvussa 2 ”True Believers: The Intentional Strategy and Why It Works”.

(1) luonnonvalinta ei ole ainut vaikuttava evolutiivinen voima tai mekanismi, ja (2) luonnonvalinta ei takaa optimaalisuutta, vaan ainoastaan sen, että tiettyinä hetkenä valittu piirre on ollut valintahetkellä hyödyllinen. Luonnonvalinnan lisäksi vaikuttaa muun muassa seksuaalivalinta, joka suosii esimerkiksi parinvalintaan vaikuttavia morfologisia ominaisuuksia ja käyttäytymispiirteitä, jotka vetävät puoleensa toista sukupuolta.⁷⁰ Dennettikin myöntää, ettei luonnonvalinta takaa täydellisyyttä: ”[j]os olemme evoluution suunnittelejia, niin emme ole lähes varmasti mitään muuta kuin kassillinen konsteja kyhätty tyydyttämään Luontoa – Herbert Simonin termi (1957) – ja yhtään parempia kuin mitä esi-isämme olivat selviytyäkseen” (51; vrt. Simon 1957). Tällaisia esimerkkejä on eliöiden perimät täynnä: luonnonvalinnan seurauksena syntyneitä ratkaisuja, jotka eivät ole täysin ihanteellisia.

Organismin selviytyminen ei kuitenkaan edellytä täydellistä, vaan ainoastaan riittävää rationaalisuutta. Riittävä rationaalisuus on välttämätöntä, muuten organismi ei pärjäisi ympäristössä. Riittää, että organismilla on tarpeeksi tosia uskomuksia ja oikeita päättelysääntöjä selviytyäkseen ja lisääntyäkseen. Joskin enimmäkseen todet uskomukset ovat paljon hyödyllisempiä kuin virheelliset uskomukset, joskus jopa virheelliset uskomukset voivat olla hyödyllisiä. Dennett antaa esimerkin (51, alaviite), jonka mukaan eläimelle voi olla hyödyllistä, että sillä on virheellisiä uskomuksia siitä, kenet se voi selättää ja kenet ei. On nimittäin parempi eläimelle, ettei se yritä päihittää julmimpia ja hankalimpia vastustajiaan, vaikka se pystyisikin niistä joitakin päihittämään, sillä huonossa tapauksessa tulokset olisivat sen kannalta kammottavat. Liiallinen varovaisuus on kustannustehokkaampi heuristiikka kuin liiallinen varomattomuus.

Luonnonvalinta takaa siis sen, että olemme lähes normatiivisten mallien mukaisesti rationaalisia. Artefaktit, kuten termostaatit ja shakkia pelaavat

⁷⁰ Muita evoluutiomekanismeja ovat esimerkiksi satunnaisajautuminen ja muuttuvat mutaatiotahdit (ks. esim. Gould & Lewontin 1979).

tietokoneet, puolestaan eivät ole luonnonvalinnan muokkaamia. Miten on niiden rationaalisuuden laita? Dennettin mukaan ne ovat rationaalisia, koska ne ovat optimaalisesti suunniteltuja: optimaalinen (kognitiivinen) suunnittelu takaa rationaalisuuden (Dennett 1988a, 497). Miten ihmisten luomat artefaktit voisivat olla optimaalisia?

On selvää, ettei evoluutio takaa sitä, että voisimme olla standardin rationaalisuuskuvan mukaisesti optimaalisia ja siten rationaalisia. Samaten on selvää, etteivät artefaktit ole aivan optimaalisia.⁷¹ Miten vain, mutta Dennettiä seuraten irrationaalisuutta ei voida hyväksyä *suopeuden periaatteen* (engl. principle of charity; Quine 1960) pohjalta: jos jotakuta yritetään ymmärtää, on oletettava, että hän on rationaalinen (ja optimaalisesti suunniteltu), koska irrationaalista (tai ei-optimaalisesti suunniteltua) henkilöä ei voida kerta kaikkiaan ymmärtää (intentionaalisen strategian keinoin). Tämä on intentionaalisen asenteen (optimaalisuus- ja) rationaalisuusoletus. Jos jotain intentionaalista systeemiä yritetään tulkita, täytyy olettaa, että kyseinen systeemi noudattaa normatiivisia periaatteita:

Oletus siitä, että jokin on intentionaalinen systeemi, on oletus siitä, että se on rationaalinen; toisin sanoen oletus, että entiteetillä x on uskomukset p , q , r , ... ei johda mihinkään, ellei oleteta, että x uskoo mitä seuraa p :stä, q :sta, r :stä, ...; muussa tapauksessa ei ole mahdollisuutta sulkea pois ennustusta siitä, että x aikoo, sen uskomuksien p , q , r , ... ollessa kyseessä tehdä jotakin kerrassaan typerää, ja, jos emme pysty sulkemaan pois *tuota* ennustusta, meillä ei ole hankittua ennustusvoimaa lainkaan. Joten sanottiinpa eläimen *uskovan* logiikan *totuuksia* tai ei, täytyy olettaa, että se *noudattaa* logiikan *sääntöjä*. (BS, 10–11.)

Jos huomaamme jotakin ristiriitaisuutta henkilön toiminnassa, oletamme, että hän ei ole kenties kuullut, että K tai havainnut, että H (BS, 10). Esimerkiksi kuvitelkaamme, että henkilö ei halua kastua sateessa työmatkallaan. Hän on lukenut päivän lehden, jonka sääennusteessa luvataan koko päivän olevan

⁷¹ Tästä eteenpäin samaistan rationaalisuuden ja suunnittelun optimaalisuuden, eli puhuessani rationaalisuudesta tarkoitan myös suunnittelun optimaalisuutta.

hyvin sateinen. Hän ei kuitenkaan ulosmennessään ottanut sateenvarjoaan mukaansa. Tämä hämmästyttää meitä, ja koetamme korjata tulkintaamme. Oletamme, että hän on unohtanut tarkistaa sääennusteen, hän menee taksilla (vaikka normaalisti hän menee jalkaisin) tai että hän ei yksinkertaisesti usko sääennusteeseen. Voimme myös vedota päättelyn häiriötekijöihin, kunhan henkilö tai systeemi ei systemaattisesti poikkea normatiivisista periaatteista.

Jos henkilö menisi jatkuvasti ilman sateenvarjoa ulos, vaikka sataisi, ja olisimme vakuuttuneet, ettei hän halua kastua, emme lukisi häntä enää kovin rationaaliseksi. Kun näin käy, saatamme erittäin äärimmäisessä tapauksessa luopua intentionaalisesta asenteesta systeemin suhteen, sillä emme voi ennustaa enää sen toimintaa rationaalisuusoletuksen mukaisesti.⁷² Vastaavasti suunnittelun asenne pitää jättää pois, jos olion ei voida olettaa toimivan niin kuin sen on suunniteltu toimivan (28 ja BS, 6). Esimerkiksi jos osa radiovastaanottimen elektronisista komponenteista tuhoutuu, emme voi enää ennustaa sitä suunnittelun asenteesta. Intentionaalisesta strategiasta luovumme käytännössä todella harvoin ihmisten ja muiden kehittyneiden eläinten tapauksessa, sillä meillä ei intentionaalisen strategian ohella ole niiden suhteen muita tehokkaita ennustusstrategioita: suunnittelun ja fysiikan strategiat ovat meidän episteemisiä rajoituksiamme ajatellen aivan liian työläitä.

Kysymykselle, mitä intentionaalisen asenteen rationaalisuusoletuksen vaatima rationaalisuus on, ei ole niin selvää vastausta. Dennett ei halua rinnastaa tätä rationaalisuutta minkään yksittäisen normatiivisen mallin tai niiden kokoelman (so. standardin rationaalisuuskuvan) kanssa, koska ne vaativat kohdeltavalta organismilta liikaa (94): on selvää, ettei mikään evoluution tuottama organismi

⁷² Tässä mielessä Dennettin rationaalisuuskäsitys muistuttaa sitä, mitä Stein (1996) nimittää heikoksi suopeuden periaatteeksi: ei pidä ajatella ihmisten – tai tässä tapauksessa rationaalisten agenttien – poikkeavan rationaalisuuden normatiivisista periaatteista, ellei sen puolesta ole vahvaa näyttöä.

voi olla standardin rationaalisuuskuvan vaatimuksen mukainen.^{73,74} On myös epäilyttävää, voitaisiinko mitään sellaista sisäisesti ristiriidatonta normatiivista periaatteistoa edes kehittää, joka sanelisi, mitä missäkin tilanteessa olisi kannattavaa tehdä ja mitä pitäisi tehdä. Esimerkiksi deduktiivisen logiikan normimalli ei sovellu kaikkiin tarkoituksiin: jos informaatiosta on puutetta, tai se on epävarmaa, deduktiivisen logiikan päättelyperiaatteet määräävät olla tekemättä mitään epätotuusilta välttymiseksi (97). Käytännön maailmassa, jossa nopeus on tavallisesti valttia epätotuuden riskin sijaan, se ei ole hyvä tai meidän intuitiivisen rationaalisuuskäsityksemme mukainen neuvo. Olisi olion kannalta täysin mieletöntä tehdä loogisia päätelmiä sen kannalta merkityksettömistä asioista, ja kuluttaa sellaisiin koko käsillä oleva aika. Sellainen toiminta estäisi oliota tekemästä päätelmiä sen selviytymisen kannalta merkityksellisistä asioista.⁷⁵ (Cherniak 1981.)

Tietenkään Dennett ei silti rinnasta intentionaalisuustulkinnan vaatimaa rationaalisuutta minkään tosiasiallisen, deskriptiivisen rationaalisuusmallin kanssa, koska rationaalisuusoletuksen oletama rationaalisuus on ihanteellinen ja normatiivinen. Dennett pitäytyy esiteoreettisessa, intuitiivisessa rationaalisuuskäsityksessämme, eikä eksplisiittisesti määrittele sitä. Kyseessä on meidän intuitiivisesta käsityksestämme – mehän rationaalisuusoletuksen teemme. Dennett ei itse kuitenkaan mainitse tai huomaa, että se on mahdollista saada esiin eksplisiittiseksi käsitykseksi empiirisiin, kuten kognitiivisen kulttuuriantropologian, keinoin, oli se kuinka implisiittinen käsitys hyvänsä.

⁷³ Christopher Cherniak (1981, 171) esittää sattuvan huomautuksen siitä, miksi rationaalisuutemme ei voi olla standardin rationaalisuuskuvan mukainen: "[J]os olio [ihminen] tyydyttäisi kelvollisesti [ideaalit ehdot], monet deduktiivisten tieteiden, kuten matematiikan, tehtävät olisivat *triviaaleja* sille".

⁷⁴ Kaiken lisäksi ei ole lainkaan selvää, voiko standardin rationaalisuuskuvan mukaan toimivaa oliota olla edes olemassa. Jos uskomusten ajatellaan olevan propositioita ja päättelyn näistä propositioista deduktiivista johtamista, päädytään väistämättä kombinatoriseen räjähdykseen (ks. esim. BS, 125). Tämä kritiikki pätee myös Fodorin (esim. 1975) komputationaaliseen mielen teoriaan.

⁷⁵ Tämä liittyy läheisesti myös tekoälyn filosofiseen ongelmaan, joka tunnetaan nimellä "frame problem" (mitään vakiintunutta suomennosta kyseiselle termille ei käsittäkseni ole; joku voisi suomentaa sen "kehysongelmaksi"). Ks. "frame problemista" esim. Dennett 1984.

4.3 Kritiikki kolmannen persoonan mukaisesta lähestymistavasta

John Searlen kuin myös Thomas Nagelin (1986) mielestä Dennettin ottama tieteellinen kolmannen persoonan näkökulma on riittämätön mielen tarkasteluun.⁷⁶ Searle (1982, 1988 ja 1992) antaa ymmärtää, ettei intentionaalisen strategian toimivuus takaa intentionaalisuuden olemassaoloa, vaan ainoastaan yksilön oma tietoisuus siitä.^{77,78} Hänen (1980) mukaansa ainoastaan aivot tuottavat intentionaalisuutta. Mitkään koneet eivät pysty vastaavaan edes jäljittelemällä aivojen toimintaa ja tuottamalla aivan vastaavanlaista käyttäytymistä kuin ihminen. Vaikka koneen kausaaliset voimat ja käyttäytyminen olisivat erottamattomia ihmisen käyttäytymisestä ja (aivojen) syöte-tuloste-mekanismeista, ei sillä silti olisi intentionaalisuutta (Searle 1982).

Dennett (1982b) kysyy: mitä intentionaalisuus sitten on, jos sitä ei voida havainnoida henkilön toiminnasta tai edes neurofysiologiasta? Mikä intentionaalisuuden tuottaa? Tämä on Dennettin mielestä mieletöntä intentionaalisuuden mystifointia. Searle ei pysty vastaamaan näihin kysymyksiin ja pitää episteemisiä ongelmia irrelevantteja kannalleen. Hän toteaa, että aivojemme kausaaliset voimat takaavat jollakin tavalla intentionaalisuuden ja sen, ettemme ole tietoisuutta vailla olevia zombeja⁷⁹.

⁷⁶ Koska tämä kritiikki liittyy enimmäkseen tietoisuuden metodologiaan ja funktionalismin yleiseen kritiikkiin, jätän sen käsittelyn lyhyeen pintaraapaisuun (josta kylläkin nähdään peruskysymysasetelma) tässä tutkielmassa. Searlen ja Nagelin näkökulmat ottavat juuri tietoisuuden ja tietoisuuden tunteen ensimmäisen persoonan lähtökohdaksi.

⁷⁷ On huomattava, että edellä esitetyn uskomuksen ja mielipiteen välisen jaottelun nojalla on mahdotonta Dennettin näkemyksen mukaan, että ihminen olisi tietoinen uskomuksistaan ja haluistaan. Siten kyseinen jaottelu asettaa väistämättä Dennettin ja Searlen näkemykset vastakkain.

⁷⁸ Searle tekee naiivin oletuksen, että ensimmäisen persoonan näkökulma olisi väistämättä erehtymätön intentionaalisten tilojen suhteen. Kuten Alison Gopnik (1993) huomauttaa, meidän näkemysemme omasta toiminnastamme ja mielentiloistamme saattaa olla vastaavan prosessin tuottama kuin näkemysemme muiden toiminnasta ja heidän mielentiloistaan. Toisin sanoen käytämme kolmannen persoonan näkökulmaa myös omien intentionaalisten tilojemme suhteen. Tätä tukevat myös Gopnikin tekemät kehityspsykologiset kokeet. Ks. kyseinen artikkeli ja sen kommentaari.

⁷⁹ Zombi on filosofien keskuudessa ihminen (tai ainakin melkein ihminen), joka käyttäytyy aivan normaalisti, kuten tavallinen ihminen ja joka ei ole millään ulkoisilla kriteereillä tarkasteltuna poikkeava, mutta ei ole lainkaan tietoinen. Dennett ei itse usko zombien

(esim. Searle 1980 ja 1988.) Hän olettaa mentaalisten tilojen olemassaolon ja epistemologisen saavutettavuuden, koska ne ovat hänen mielestään kognitiotieteen ennakko-oletuksia (Searle 1980). Dennett ei siedä tällaista salamyhkäisyyttä:

Searle julistaa, että jotenkin – ja hän ei pysty sanomaan mitään yksityiskohdista – ihmisaivojen biokemia takaa, että ihmiset eivät ole zombeja. Tämä on rohkaisevaa, mutta mystifioivaa. Miten aivojen biokemia luo tällaisen onnekkaan vaikutuksen? Joillakin vaikuttavilla kausaalisilla voimilla tosiaankin; kyseessä on sama kausaalinen voima, jonka Descartes luki aineettomien sielujen syyksi, ja Searle ei ole tehnyt siitä yhtään vähemmän vaikuttavaa tai mystistä – tai loppujen lopuksi epäkoherenttia – vakuuttamalla, että kyse on vain biokemiasta. (336.)

Kyse on mielen tutkimisen metodologiasta. Tällaisessa tilanteessa molemmat osapuolet heittelevät puolin ja toisin *reductio ad absurdum* –argumentteja, koska eivät hyväksy toistensa metodologisia lähtökohtia. Thomas Nagel (1986) ja Searle (esim. 1988) pitää Dennettiä skientistina, koska hän hylkää ”epätieteellisen” subjektivistisen lähtökohdan. Dennett syyttää heitä mystifioinnista, koska he sitoutuvat aprioriseen kantaan, essentialismiin ja mystifiointiin eli Dennettin välttämiin kirosanoihin.⁸⁰

4.4 Kiista alkuperäisestä ja johdetusta intentionaalisuudesta

Ikävä kyllä näyttää siltä, ettemme usko tietokoneiden tai termostaattien todella uskovan tai haluavan jotain, toisin kuin Dennett esittää. Puhe niiden intentionaalisuudesta näyttää ainoastaan metaforiselta tai meidän kielenkäytöstämme johdetuilta.

Juuri tähän Searle on tarttunut kolmannen persoonan kritiikin lisäksi. Hänen mielestään on tärkeää tehdä erottelu intrinsisen intentionaalisuuden (engl. *intrinsic intentionality*), johdetun intentionaalisuuden (engl. *derived intentionality*) ja ikään kuin –intentionaalisuden (engl. *as if intentionality*)

olemassaoloon, vaan pitää koko käsitystä epäkoherenttina (esim. Dennett 1982a ja Dennett 1999).

⁸⁰ Ks. näistä lähtökohtaisista eroista esim. Rorty 1993.

välille. Hän havainnollistaa näitä eroja kolmen eri esimerkkilauseen (Searle 1992, 78) avulla:

1. Olen nyt janoinen, hyvin janoinen, koska minulla ei ole ollut mitään juotavaa koko päivänä.
2. Nurmikkoni on janoinen, hyvin janoinen, koska sitä ei ole kasteltu viikkoon.
3. Ranskaksi "j'ai grand soif" tarkoittaa "olen hyvin janoinen".

Searlen mukaan ensimmäisessä esimerkkilauseessa on kyse todellisesta, *intrinsicestä intentionaalisuudesta*, koska sen sanoja on ihminen. Searle ajattelee, ettei kenenkään ole mahdollista kiistää sanojan omaa mielentilaa. Esimerkiksi pelkkä janoisuuden kokemus on riittävä ehto sille, että henkilö on janoinen (Searle 1988, 527–528). Toisen lauseen kuvaama tapaus on puolestaan ainoastaan metafora, jossa nurmikon sanotaan olevan janoinen, hän väittää. Tosiasiassa nurmikolla ei ole intentionaalisuutta, sillä sen intentionaalisuus riippuu meidän kielenkäytöstämme, jossa tietoisesti attribuimme nurmikolle uskomuksen janoisuudesta ja halun veden saamisesta. Koska tällaisissa tapauksissa attribuoinnin kohteena olevilla olioilla ei ole oikeasti intentionaalisia tiloja, vaan ne ovat havaitsijoiden niille attribuomia, Searle käyttää tapauksesta nimitystä *ikään kuin -intentionaalisuus* tai *havaintajasta riippuva -intentionaalisuus* (engl. observer relative intentionality). Kolmannen lauseen ilmaisemassa tapauksessa on kyse *johdetusta intentionaalisuudesta*, sillä lauseen intentionaalisuus on johdettu sanojan mielen ensisijaisesta, *intrinsicestä intentionaalisuudesta*; ainoastaan kielen käyttäjät antavat lauseelle merkityksen. (Searle 1992, 78–79.) Näin ollen niin sanoilla kuin tietokoneillakaan ei ole *intrinsicistä intentionaalisuutta* meistä riippumatta.

Searlen esitys vaikuttaa intuitiiviselta, mutta on kysyttävä, mistä meidän *intrinsicinen intentionaalisuutemme* on peräisin. Searle ei pysty vastaamaan kysymykseen. Dennettin mielestä meidän *intentionaalisuutemme* on aivan samalla tavalla johdettua kuin käyttämiemme sanojenkin. Meidän

intentionaalisuutemme on johdettu evoluutiosta, hän kertoo (299; Dennett 1969, 41 ja Dennett 1995a).

Tarkastelkaamme seuraavaksi Dennettin kantaa tukevia ajatuskokeita – tai kuten Dennett sanoisi: intuitiopumppuja (esim. Dennett 1980).

4.4.1 Virvoitusjuoma-automaatin intentionaalisuus eri ympäristöissä

Ensimmäinen johdettua intentionaalisuutta tukeva ajatuskoe on virvoitusjuoma-automaatin kolikontunnistuslaitteen tapaus (290–293 ja Dennett 1988a, 503; mukailtu):

Yhdysvalloissa rakennettu virvoitusjuoma-automaatti sisältää kolikontunnistuslaitteen, joka erottaa aidot amerikkalaiset 25-senttiset kolikot muista esineistä, kuten pelimarkoista ja muiden maiden kolikoista. Kun automaattiin normaalisti syötetään 25-senttinen, se menee tilaan Q, joka ”tarkoittaa”: ”havaitsen ja hyväksyn nyt aidon amerikkalaisen neljännesdollarin”. Kolikontunnistuslaite ei ole kuitenkaan idioottivarma, joten se voi mennä ”virheellisesti” tilaan Q, vaikka sinne olisi syötetty jokin muu kuin aito amerikkalainen neljännesdollari. Tässä tapauksessa sanotaan tavallisesti, että laite ”erehtyy”. Erehtymiseksi sitä kutsutaan sen takia, että automaatin huijaus tilaan Q syötillä on vastoin automaatin rakentajien ja suunnittelijoiden tarkoituksperiä.

Panamalaiset neljännesbalboat ovat kuvioinniltaan erilaisia kuin amerikkalaiset neljännesdollarit, mutta muilta piirteiltään – painoltaan, paksuudeltaan, läpimitaltaan ja materiaalikoostumukseltaan – täysin samanlaisia. Koska virvoitusjuoma-automaatti toimii sitä muuttamatta Panamassa, niitä aletaan toimittaa myös sinne. Nyt Panamassa virvoitusjuoma-automaatin kolikontunnistuslaitteen sanotaan tekevän virheen, jos se menee tilaan Q, kun automaattiin syötetään neljännesdollari; aivan samoin kuin Yhdysvalloissa neljännesbalboa, on Panamassa neljännesdollari syötti. Sanommeko nyt, että laite on tilassa, jota kutsuimme

aiemmin Q:ksi? Jos sanomme näin, niin silloinhan laite erehtyy toimiessaan Panaman ilmanalassa, koska siellä sen tarkoitus on hyväksyä neljännesbalboa. Kutsukaamme tilaa ”havaitseen ja hyväksyn nyt panamalaisen neljännesbalboan” QB:ksi. Kummassa tilassa laitteen sanotaan nyt olevan, Q:ssa vai QB:ssa?

Dennett sanoo yhteen ääneen Ruth Millikanin (1984) kanssa laitteen tilan riippuvan siitä, mitä tarkoitusta varten se on valittu. Jos Panaman hallitus päättää käyttää valtiossaan amerikkalaisia virvoitusjuoma-automaatteja, jotka on suunniteltu hyväksymään vain neljännesdollareja, Dennett ja Millikan sanovat virvoitusjuoma-automaatin olevan tilassa QB – nimittäin virvoitusjuoma-automaatin tarkoitus on valittu nykyisessä ympäristössään hyväksymään vain neljännesbalboat. Jos taas virvoitusjuoma-automaatti olisi lähetetty vahingossa Yhdysvalloista Panamaan, sen neljännesbalboan hyväksyminen ei tarkoittaisi mitään. On selvää, ettei automaatti vaihda sisäisten tilojensa intrinsistä merkitystä, mutta sen ulkoisen ympäristön muuttuessa sen johdettu sisäisten tilojen merkitys muuttuu. Alkuperäisen intentionaalisuuden kannattajien mukaan tämä muutos on harmiton, sillä artefaktien tilat eivät oikeasti tarkoita yhtään mitään. Niillä on vain johdettua intentionaalisuutta. Edellisen kaltainen puhe on alkuperäisen intentionaalisuuden kannattajien mukaan vain pragmaattinen tapa puhua artefaktin intentionaalisuudesta. (293–294.)

Dennett esittää vastavetona väitteen, jota monikaan ei ole valmis hyväksymään: samat pragmaattiset tulkinnan säännöt pätevät ihmiseen. Ihmisen intentionaalisuus ei eroa millään tavalla virvoitusjuoma-automaatin intentionaalisuudesta. Searlen mielestä ihmisen uskomuksilla on intrinsistä, objektiivista tai aitoa intentionaalisuutta tulkinnasta välittämättä. Hänen mukaansa on olemassa objektiivinen tosiasia siitä, mitä ihmisen mielentila oikeasti tarkoittaa. (294.) Dennettin ajatuskokeen tarkoituksena on vakuuttaa

hänen kriitikkonsa siitä, että merkitykset ovat tulkitsijasta riippuvia; virvoitusjuoma-automaatin tarkoitus on suhteellinen ja tulkinnanvarainen.

4.4.2 Intentionaalisuus Rinnakkaismaassa

Tarkoituksen ja tarkoittamisen tulkinnanvaraisuus selvempi vastaavanlaisessa ajatuskokeessa, jossa virvoitusjuoma-automaatti korvataan ihmisellä (294–295; mukailtu):

Henkilö, nimeltä Jones, katsoo ulos ikkunasta ja ajattelee näkevänsä hevosen. Ajatellaan, että nyt Jones kiidätetään Rinnakkaismaassa, joka on aivan samanlainen planeetta kuin Maa, mutta siellä hevosten tilalla on hemosia⁸¹. Hemoset eivät ole hevosia, mutta vain ammattitaitoinen biologi erityisine apparateineen pystyy erottamaan hemoset hevosista. Kun nyt Jones näkee hemosen Rinnakkaismaassa, onko hän tilassa M, jossa hän uskoo (virheellisesti) näkevänsä hevosen vai tilassa K, jossa hän uskoo (totuudenmukaisesti) näkevänsä hemosen ensimmäistä kertaa elämässään?⁸²

On selvätkin hankalaa sanoa kummassa tilassa, M:ssä vai K:ssä, Jones oikeasti on. Jos Jones viipyisi vain hetken Rinnakkaismaassa ja palaisi sitten takaisin Maahan, tarkoittaisiko hänen uskomuksensa mitään? Olisiko kyse samasta asiasta kuin edellisen virvoitusjuoma-automaatin ajatuskokeessa? Entä jos Jones jäisi Rinnakkaismaahan, olisivatko hänen uskomuksensa (totuudenmukaisesti) hemosista?

Searlen ynnä muiden intrinsistä intentionaalisuutta kannattavien mielestä Jonesin tila on objektiivisesti oltava jompikumpi. Dennettin näkemykseen

⁸¹ Dennett käyttää itse nimitystä "schmorse". Pitäydyn luonnollisessa suomenkielessä kutsumalla "schmorseja" hemosiksi.

⁸² Vrt. Hilary Putnamin (1975) esittämään Rinnakkaismaa-ajatuskokeeseen.

kuuluu käsitys, jonka mukaan intentionaalisia tiloja ei voida determinoida eikä intentionaalinen tila ole objektiivinen, vaan tulkitsijasta riippuva.⁸³

4.4.3 Selviytymiskoneen intentionaalisuus

Dennett esittää vielä yhden ajatuskokeen, ajatuskokeen selviytymisrobotista (295–298). Sen tarkastelun jälkeen Dennettin väite siitä, että myös ihmisen intentionaalisuus on johdettua, on helpompi ymmärtää.

Kuvittele, että haluaisit selvitä 2400-luvulle asti. Oletetaan, että ainoa tapa pitää elimistösi elossa sinne asti, olisi säilyttää se jonkinlaisessa koneessa, joka pitää elimistösi horros- ja tajuttomuustilassa. Valittavanasi on kaksi erilaista strategiaa. Asetu kasvin tapaan johonkin kiinteään ja paikallaan pysyvään otolliseen paikkaan, joka on varustettu vedellä, auringonvalolla ja millä tahansa, mitä kapselisi ja elonjäämistäsi tarkkaileva järjestelmä tarvitsee. Haittapuolena tässä strategiassa tukikohdan liikkumattomuus, joku saattaa esimerkiksi rakentaa moottoritien sen päälle. Toinen strategia onkin parempi ja hienostuneempi: rakenna tukikohtaasi lisäominaisuudet, joiden avulla se pystyy eläimen tapaan havainnoimaan ympäristöönsä ja liikkumaan, siten välttämään vaarat ja hakemaan uusia energianlähteitä niitä tarvitessaan. Rakenna siis robotti, jonka sisuksiin sijoitat itsesi kapselissa. Koska olet itse horros- ja tajuttomuustilassa, sinun täytyy suunnitella robottisi tarpeeksi älykkääksi, jotta se osaa tuottaa omia ratkaisujaan erilaisiin olosuhteisiin ja kohtaamiinsa ongelmiin. Sen täytyy tietää, miten se voi etsiä, tunnistaa ja siten hyödyntää erilaisia energianlähteitä. Lisäksi sen täytyy tietää, miten se voi ennakoida ja sillä tavoin välttää vaaroja. Tehtävä tulee vielä hankalammaksi, jos tulevaisuuden uhkia harkitaan tarkemmin. Et voi nimittäin luottaa siihen, että robottisi tulee olemaan ainut selviytymisrobotti maailmassa. Robottisi

⁸³ Ks. luku 4.5 ”Intentionaalisten tilojen ja funktioiden määrittämättömyys”.

saattaa joutua kamppailemaan muiden selviytymisrobottien kanssa rajoitetuista energianlähteistä. Siis, suunnittele selviytymisrobotillesi lisäksi kommunikointi- ja yhteistyökyvyt.

Tuloksena edellä kuvatussa suunnitteluprosessissa on robotti, joka on itseohjautuva ja itseään hallitseva, sillä et pysty horrostilassa ollessasi sitä ohjaamaan. Oleellista on, että suunnittelet robotin niin, että sen päätavoitteena on suojella sinua. Näin ollen se valitsee sinun intressiesi mukaiset eli selviytymistäsi palvelevat toimintavaihtoehdot. Robotin tulee pystyä johtamaan toissijaiset tavoitteensa tästä päätavoitteestaan.

Yllä esitetty ajatuskoe on hienostuneempi ja monimutkaisempi versio aiemmasta virvoitusjuoma-automaattia käsitelleestä ajatuskokeesta. Kuten virvoitusjuoma-automaatin tapauksessa intrinsistä intentionaalisuutta kannattavat, myös tässä tapauksessa he ovat varmasti sitä mieltä, että intentionaalisuus on loppujen lopuksi aina suunnittelijasta johdettua – molemmissa tapauksissa siis ihmisestä johdettua. Mitä he sanovat, kun paljastamme ajatuskokeen analogian kohteen? Ajatuskoe on muunnelma Richard Dawkinsin (1993) kuuluisasta metaforasta organismeista selviytymiskoneina, joiden tavoitteena on jatkaa geeniensä olemassaoloa. Sen mukaan me olemme artefakteja, jotka on suunniteltu koko elämän historian aikana evoluutioprosessissa. Me palvelemme geeniemme intressejä jokseenkin itsenäisesti, sillä geenit eivät pysty reaaliaikaiseen kontrolliin. Siten intentionaalisuutemme on johdettua, johdettua suunnittelijoistamme eli ”itsekkäistä” geeneistämme.⁸⁴ (298.)

Geenien intentionaalisuus on selvästikin mallitapaus ikään kuin - intentionaalisuudesta. Dennettin ajatuskokeessa suunnittelu oli tietoista ja tarkoituksellista, geenien tapauksessa asia on kumminkin toisin. Geenit suunnittelee pitkä ja hidas evoluutioprosessi, joka ei ole millään tavalla

⁸⁴ Itsekkäät geenit –metaforassa geenit nähdään intentionaalisina systeemeinä.

tietoinen tai tarkoituksellinen. Vaikka luonnonvalinnalla ei ole tarkoitusta tai kaukokatseisuutta, se voi valita tietyn suunnittelun toisen sijaan koskaan ”representoimatta” tietoisesti tai tiedostamattaan kumpaakaan valintaa tai syytä valinnalle.⁸⁵ Samalla tavoin kuin panamalainen virvoitusjuomayhtiö voi valita virvoitusjuoma-automaatin sen neljännesbalboan erottelukyvyn takia, voi evoluutio valita keuhkot sen veren hapettamiskyvyn vuoksi. (299–300.)

Dennettin mielestä ajatusta siitä, että olemme luonnonvalinnan suunnittelemlia artefakteja, on vastustettu sen kahden – monen mielestä vastenmielisen – seuraamuksen vuoksi:

[J]os olemme (vain) artefakteja [niin kuin virvoitusjuoma-automaatit tai selviytymisrobotit ovat], silloin sen suhteen, mitä sisimmät ajatuksemme tarkoittavat – jos ne ylipäättään tarkoittavat mitään – on jotain, johon meillä, näiden ajatusten todellisilla ajattelijoilla, ei ole mitään erityistä auktoriteettia. [...] Jos olemme sellaisia artefakteja, meillä ei ainoastaan ole mitään varmaa etuoikeutettua pääsyä tosiasioihin, jotka sitoisivat ajatuksiemme merkitykset, vaan *sellaisia syvempiä tosiasioita ei ole*. (300.)

Syynä näille seuraamuksille on seuraava. Jos tulkinta jonkin tarkoituksesta (funktioista) tai merkityksestä on epäselvä, emme pysty tarkistamaan mistään evoluution sille suomaatarkoitusta, sillä sellaista totuutta asiasta ei yksinkertaisesti ole olemassa (303). Tätä käsitellään enemmän seuraavassa.

4.5 Intentionaalisten tilojen ja funktioiden määrittämättömyys

Koska intentionaalinen asenne on tulkitsijasta riippuva, myös intentionaaliselle systeemille attribuotitavat intentionaaliset tilat riippuvat niiden attribuotijasta. Tällaisesta interpretationistisesta näkemyksestä seuraa, ettei voida sanoa, mitä intentionaalisia tiloja intentionaalisella systeemillä todella on. Toisin sanoen intentionaalisista tiloista voidaan antaa monta erilaista (kilpailevaa) tulkintaa, eikä ole olemassa yhtä ainuttat kuvausta siitä, missä intentionaalisessa tilassa systeemi oikeasti on.

⁸⁵ Vrt. luku 2.3 ”Universaali darwinismi Dennettin presuppositioina”.

Idea äärettömästä määrästä eri tulkintoja eli loppumattomasta tulkintojen disjunktioiden sarjasta ei ole uusi. Ajatus on lähtöisin Quinen *käännöksen epätäydellisyys* (engl. indeterminacy of [radical] translation; Quine 1960) – teesistä, joka sanoo seuraavaa kielten (esimerkiksi suomen- ja englanninkielen) välisestä merkityksen kääntämisestä:

[K]ielen käännösoppaita voidaan laatia erinäisillä tavoilla, jotka kaikki ovat yhteensopivia puhejäsenyksien kokonaisuuden kanssa ja silti yhteensopimattomia toistensa kanssa. Lukemattomissa tilanteissa ne poikkeavat antaessaan, yhdelle kielelle kuuluvan lauseen käännöksen, toisten kielten lauseita, jotka eivät ole suhteessa toisiinsa missään uskottavassa vastaavuudessa suhteen ollessa kuinka väljä hyvänsä. (Quine 1960, 27.)

Quine väittää, ettei täydellistä vastaavuussuhdetta ole kielten käännösten välillä. Vastaavuussuhteen puuttuminen tarkoittaa, ettei sanoman merkitystä voida sellaisenaan siirtää kielestä toiseen, vaan siitä voidaan siirtää vain erilaisia tulkintoja.

Yhtäläisesti ajatuksillamme tai uskomuksillamme ei ole määrättyä sisältöä tai merkitystä. Ei voida sanoa, esimerkiksi uskooko henkilö ulkona satavan vai ei. Tässä kohdin monet, jotka näkevät intentionaalisten tilojen olevan sellaisenaan aivoissa, ovat eri mieltä. Jos intentionaaliset tilat olisivat sellaisenaan löydettävissä aivoista, paremmalla neurofysiologian tuntemuksella voitaisiin sanoa, missä tilassa henkilö on. Kysymys aivan määrätystä tilasta on mieletön ja tulokseton seuraavasta Quinen (1960, 68–72) esittämästä syystä.⁸⁶ Ensinnäkin henkilön (tai systeemin) toiminnan tulkinta vaatii oletuksia siitä, miten hän hahmottaa maailmaa. Quinen mielestä tällaiset oletukset ovat esiteoreettisia eivätkä empiirisiä, ja siten niitä ei voida konfirmoida eikä diskonfirmoida, koska ne ovat funktioita, jotka muodostetaan aina henkilön ruumiin liikehdinnän arvioimisesta henkilön tekojen arvioimiseen. Siitä syystä sisällön liittäminen henkilön uskomuksiin onnistuu vain suhteessa johonkin tulkintaan.

⁸⁶ Quinen varsinainen esitys on hieman erilainen. Mukailen tässä kohdin Timothy Kenyonin (2000, 79) Quinen käsityksen esittelyä.

Henkilön käyttäytymiselle voidaan antaa monia erilaisia tulkintoja, jotka voivat olla keskenään jopa ristiriitaisia, ja silti henkilön käyttäytyminen ei voi antaa mitään sellaista evidenssiä, joka määräisi, mikä tulkinta on oikeassa. Meidän taipumuksemme päätyä yksittäisiin tulkintoihin intentionaalisista tiloista ei tarkoita sitä, että intentionaaliset tilat todella olisivat yksittäisiä eikä sitä, että maailma olisi määrätty, ainakin mitä tulee intentionaalsiin tiloihin.

Dennettin mukaan intentionaalisia tiloja voidaan identifioida vain suhteessa biologisiin funktioihin tai evoluution valikoimiin tarkoituksiin, koska emme ole missään etuoikeutetussa asemassa intentioidemme suhteen – meillä on vain johdettua intentionaalisuutta. Dennettin mukaan luonnossa kuitenkin vallitsee funktionaalinen määrittämättömyys (engl. functional indeterminacy), jolloin mitään perimmäisiä tarkoituksia ei intentioillemme ole (320).⁸⁷

Luontoäiti ei sitoudu eksplisiittisesti ja objektiivisesti *mihinkään* funktionaaliseen attribuutioon: kaikki sellaiset attribuutiot riippuvat intentionaalinen asenne –ajattelutavasta, jossa oletamme optimaalisuuden, jotta voisimme tulkita löytämämme. (320.)

Esimerkiksi kotkan siivistä samoin kuin pandan ranteen jänneluusta ei voida sanoa, mitä tarkoitusta varten ne todella ovat. Ottakaamme asiasta esimerkki ajatuksen tarkastelua varten. Eläimillä (ihmiset mukaan luettuna) on varustus, joka on hyvin herkkä pystysuuntaisen akselin symmetrialle. Ei ole epäilystäkään siitä, mitä tällaisen varustuksen tuntoelimen signaalit tarkoittavat: ne tarkoittavat yksinkertaisesti, että verkkokalvolla on jotakin, joka on symmetrinen pystyakseliltaan. Mutta mitä varten varustus on olemassa? Dennettin mielestä se on toisten eläinten havaitsemista varten. Tuntoelin

⁸⁷ Dennettin mukaan intentionaalisen tilan määrittävyyden harhaan ovat ajautuneet ainakin John Searle, Jerry Fodor, Fred Dretske, Tyler Burge, Christopher Kripke, Roderick Chisholm, Thomas Nagel, Karl Popper ja John Eccles (295). Millikan (1984) kutsuu heitä yhdistävää käsitystä myös kuvaavalla nimityksellä merkitysrationalismi (engl. meaning rationalism). Yksi nimitys voisi olla myös Dennettin mainitsema merkitysessentialismi (engl. essentialism applied to meaning) (Dennett 1995a, 408).

ilmaisee, että joku katsoo sinua kohti.⁸⁸ Tämä on tietysti tärkeää sen takia, että eläintä vastapäätä saattaa olla potentiaalinen saalistaja, kilpailija tai parittelukumppani. Nyt kysymme, onko tämä varustus saalistaja-ilmais, saalistaja-tai-kilpailukumppani-tai-parittelukumppani-ilmais vai onko se vain ”joku katsoo sinua kohti” -ilmais? (302–304.) Samalla tavalla voisimme kysyä virvoitusjuoma-automaatin tapauksessa onko se neljännesdollarin, neljännesbalboan vai p-painoisen-ja-q-paksuisen-ja-l-läpimittaisen-ja-m-materiaalisen esineen erotin.

Jollei tulkintaa suhteuteta johonkin olettamukseen normaalista funktiosta, päädytään väistämättä tulkinnan raakaan, yksinkertaisimpaan merkitykseen, kuten ”verkkokalvolla on jotain symmetriaa pystyakseliltaan” tai ”havaitseen p-painoisen-ja-q-paksuisen-ja-l-läpimittaisen-ja-m-materiaalisen esineen”. Silloin virheellistä esitystä ei voida kerta kaikkiaan tehdä (302).

4.6 Intentionaalisten tilojen ontologinen asema

Koska Dennett on kieltäytynyt identifioimasta intentionaalisia tiloja minkään toisten tilojen kanssa, on hänen näkemyksensä kimppuun hyökätty monelta suunnalta. Seuraavaksi tarkastelen näiden hyökkäysten motivaatiota.

Arkipsykologian olettamiin intentionaalsiin tiloihin on suhtauduttu perinteisesti realismin tai eliminativismin näkökulmista. *Realismin* (esim. Fodor 1985) mukaan intentionaaliset tilat viittaavat todellisiin ilmiöihin, joilla on fyysiset tai funktionalistiset vastineensa.⁸⁹ *Eliminativismi* (esim. Churchland

⁸⁸ Kohti siinä mielessä, että yleensä eläimien takaosa ei ole niin symmetrinen kuin etuosa. Dennett lisäksi olettaa, että yleensä eläimillä silmät ovat kohdistuneet suoraan eteenpäin samaan suuntaan. Näin ei kuitenkaan käsittääkseni ole asian laita. Eteenpäin kohdistuneet silmät taitavat olla eläinmaailmassa pikemminkin poikkeuksellisia. Esimerkin ideaa tämä ei kuitenkaan turmele.

⁸⁹ Baker (1994, 3) on tiivistänyt lyhyesti Dennettin ja Fodorin välisen perustavanlaatuisen eron:

Fodorin mukaan henkilön uskomisen p:hen riippuu siitä, onko henkilöllä tietty aivotila, joka on tietyssä kausaalisessa roolissa ja joka tarkoittaa, että p. Sitä vastoin Dennettin mukaan henkilön uskomisen p:hen määräytyy täysin siitä, onko uskomus, että p ennustavasti attribuoitavissa henkilölle. [...] Dennettin intentionaalisen asenteen teorian [so. intentionaalisten systeemien teorian] erottaa

1979, 1981, 1986 ja Stich 1983) taas väittää, että intentionaaliset tilat eivät viittaa mihinkään olemassa olevaan, minkä vuoksi niiden postuloiminen on turhaa. Dennett ei näe kuuluvansa kumpaankaan ryhmään. Kuten on nähty, hän ei identifioi missään mielessä intentionaalisia tiloja materiaalistien tilojen kanssa, ei edes samassa mielessä kuin konefunktionalismi. Voitaisiin kuitenkin kysyä, että eikö ole niin, että mitään väliratkaisua ei ole, joko intentionaaliset tilat ovat identifioitavissa funktionalistisiin tai fyysisiin tiloihin tai ne eivät ole? Miten ne muuten voisivat kausaalisesti vaikuttaa mihinkään?

Dennettin näkemystä on pidetty instrumentalistisena (esim. Baker 1989, Bechtel 1985, Fodor 1990 sekä Lycan 1988), ja hän on itsekin siihen aiemmin erheellisesti sitoutunut. Vaikka hän on myöhemmin pyrkinyt irti koko nimikkeestä (esim. 71–72 ja Dennett 1993, 210), on Dennettiä on hänen vanhemman tuotantonsa perusteella pakko pitää instrumentalistina, jos katsotaan esimerkiksi seuraavaa lainausta:

[E]päilykset siitä, onko shakkia pelaavalla tietokoneella *todella* uskomuksia ja haluja, ovat väärin kohdistettuja, sillä antamani intentionaalisten systeemien määritelmä ei sano, että intentionaalisilla systeemeillä *todella* on uskomuksia, vaan että intentionaalisten systeemien käyttäytymistä voidaan selittää ja ennustaa *liittämällä* [engl. ascribe] uskomuksia ja haluja niille. (BS, 7.)

Tällainen selvästi instrumentalistinen kanta väittää, ettei intentionaalisia tiloja ole todella olemassa. Ne ovat kyseisen kannan mukaan vain hyödyllisiä konstruktioita. Tiukan ontologisen jaottelun mukaan Dennettin näkemys siis on pohjimmiltaan eliminativistinen: intentionaalisia ei oikeasti ole olemassa. Dennett ei itse ole samaa mieltä. Hänen mielestään perinteinen ontologinen rajanveto on liian karkea:

Ongelma on siinä, että virheelliset dikotomiat on hankala huomata. Kun pakotat ihmiset hyppäämään aidalta yhdelle tai toiselle puolelle, ennen kuin heillä on ollut mahdollisuus pohtia matkan varrella kunnolla heille

esitettyä kysymystä, he eivät saata päätyä huomaamaan parasta vaihtoehtoa – jotain kolmatta tai neljättä vaihtoehtoa, joita he eivät ajatelleet – koska ”ainoat vaihtoehdot” eivät oikeasti ole täydellisiä vastakohtaisuuksia huolimatta siitä, että ne ensi näkemältä saattavat sellaisilta vaikuttaa. (Dennett 1993, 216.)

Aiemmin hän on yrittänyt näyttää tämän vertaamalla intentionaalisia tiloja ihmisääniin (engl. voice, vastedes pelkkä ”ääni”). Arkipuheessa sanomme: ”kuulen jonkun äänen”, ”hänellä on keski-ikäisen naisen ääni”, ”rasitat ääntäsi” ja ”olen menettänyt ääneni”. Onko ääni olemassa oleva asia? Sanoa, että joku rasittaa ääntään kuulostaa yhtä luontaiselta kuin sen sanominen, että joku rasittaa selkäänsä. Missä ääni sijaitsee? Onko ääni yhtä kuin äänihuulet? Äänihuulia voidaan sanoa rasitettavan, mutta kukaan tuskin silti tarkoittaisi, että jollakin henkilöllä on keski-ikäisen naisen äänihuulet tai että äänihuulet voidaan menettää, samalla tavalla kuin puhuessaan äänestä. Kaiken lisäksi ääni – toisin kuin äänihuulet – voidaan tallentaa digitaaliseen muotoon ja lähettää radioteitse. Mitä ääni sitten on? Onko se identtinen kurkunpään kanssa? Ei. Entä keuhkojen? Ei. Onko se ilmavirtaa? Ei. Vaikuttaa siltä, että ääni ei ole identtinen minkään fysikaalisen entiteetin kanssa. Kysymykset ovat kuitenkin harhaanjohtettuja, sillä ne ovat väärin muotoiltuja ja mieltä vailla, samalla tavalla kuin kehotus ”Tartu tilaisuuteen!” kehotuksena tilaisuuteen fyysisestä tarttumisesta.⁹⁰ (Dennett 1969, 8–12.) Selvästikin äänet ovat silti olemassa olevia asioita.

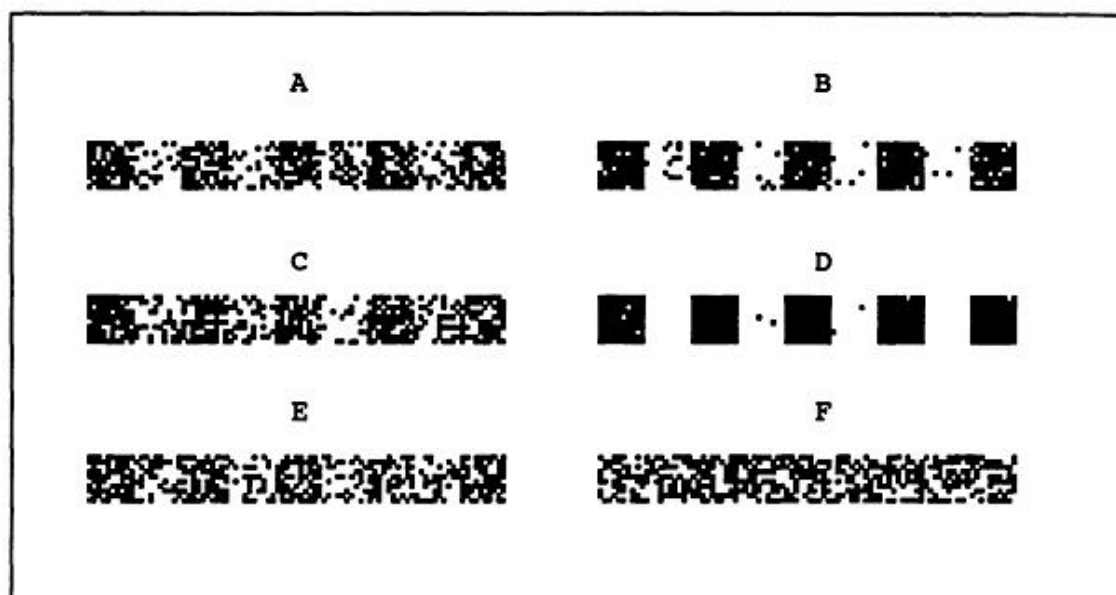
Dennett väittää intentionaalisten tilojen olevan olemassa samalla tavalla kuin äänien ja painopisteiden. Ne ovat olemassa siksi, koska ne ovat hyviä abstrakteja objekteja, hyödyllisiä loogisia konstruktioita eli abstractum-luokan olioita (Dennett 1991b, 29). Intentionaaliset tilat ovat olemassa siksi, koska ne ovat ennustettavissa intentionaalisesta asenteesta. Ennustamisen menestyksellisyys puolestaan riippuu säännöllisestä kuviosta (engl. pattern), joka on erotettavissa satunnaisuuden keskeltä. (39.) Dennettin mielestä kaikki

⁹⁰ Tämä vastaa Rylen (1949) puhetta kategoriavirheestä.

ennustaminen ja selittäminen perustuu johonkin kuvioon, jota voidaan hyödyntää (Dennett 1991b, 30). Jos kuviota ei olisi olemassa, ei voitaisi selittää eikä ennustaa mitään. Käsitelkäämme kuvion olemassaolon kriteerejä tarkemmin siirtämällä tarkastelun kohde kehyksiin ja viivakoodeihin.

4.6.1 Kuvion olemassaolo

Tarkastellaan alla olevaa kuutta kehystä (KUVIO 1). Käytännöllisesti katsoen jokainen kehys on erilainen. Toiselta kannalta katsottuna huomataan, että jokaisessa kehyksessä A:sta F:ään toistuu sama kuvio – eräänlainen viivakoodi – tai ainakin osa siitä. Kehyksissä on satunnaista kohinaa (engl. noise), joka on sotkeutunut viivakoodiin. Viivakoodikuvio ei siksi näy täydellisenä. A:ssa on kohinaa 25 %, B:ssä 10 %, C:ssä 1 %, E:ssä 33 % ja F:ssä 50 %.



KUVIO 1. Kuusi kehystä (Dennett 1991b, 31).

Sanommeko nyt kuvion, joka esiintyy kohinan keskellä jokaisessa kehyksessä, olevan todellinen – olevan oikeasti olemassa? Kuvion sanotaan olevan olemassa, jos se on erotettavissa oleva säännönmukaisuus; erottamaton kuvio

on loogisesti virheellinen samassa mielessä kuin näkymätön kuva. Eikö kuitenkin ole niin, että kuvion erotettavuus riippuu havaintokyvyistämme? Näin onkin, mutta on huomattava, että kuviolle tai satunnaisuudelle voidaan antaa standardi määritelmä. Se näyttäisi pelastavan Dennettin esittämän intentionaalisuuden olemassaolon käsityksen eikä sitoisi sen olemassaoloa havaitsijaan. Dennett tukeutuu tässä Gregory Chaitinin (1975) matemaattiseen satunnaisuuden määrittelyyn. (Dennett 1991b, 32.) Chaitinin (1975, 48) mukaan

[n]umerosarja on satunnainen, jos pienimmässä algoritmissa, joka pystyy kuvaamaan numerosarjan tietokoneelle, on sama määrä bittejä informaatiota kuin numerosarjassa itsessään.

Toisin sanoen ”sarja [...] on satunnainen, jos ja vain jos informaatio sarjan kuvauksesta on pakkaamaton”, jolloin sarjasta ei voida esittää lyhyempää kuvausta kuin mitä sarja itse on (Dennett 1991b, 32; kursiivit poistettu). Jos – ja vain jos – sarja ei ole satunnainen, se on kuvio. Siis sarjassa on kuvio, jos ja vain jos on olemassa tehokkaampi tapa kuvata se kuin sarja itse (Dennett 1991b, 32).

Ajatellaan esimerkiksi kehystä D 900-pikselisenä bittikarttana (jollainen se itse asiassa onkin). Se voidaan kuvata hypoteettisesti seuraavalla tavalla: kehys sisältää kymmenen riviä, jossa jokaisessa on yhdeksänkymmentä pikseliä siten, että jokainen rivi alkaa kymmenellä mustalla pikselillä, ja jokaista kymmentä mustaa pikseliä seuraa kymmenen valkoista pikseliä ja niin edelleen, niinä poikkeuksin, että ensimmäisen rivin pikselit 77 ja 99 ovat valkoisia ja niin edelleen. Tällöin saamme kuvion informaation kuvattua pienempään tietomäärään. Kuviota F ei taas voida kuvata lyhyemmällä tavalla, eli pakata pienempään tilaan, koska siinä on 50 prosenttia satunnaisuutta tai kohinaa ja 50 prosenttia itse kuviota. (Dennett 1991b, 32.)

Viivakoodikuvio kehyksissä A–E on helposti erotettavissa paljaalla silmällä siksi, koska hahmontunnistamisjärjestelmämme (tässä tapauksessa oikeastaan ”kuviontunnistusjärjestelmämme”) on herkkä juuri kulmille, luminanssille ynnä muille ominaisuuksille. Jos kehykset esitettäisiin meille erilaisella tavalla, emme niitä ehkä havaitsisi, mutta kuvio olisi silti olemassa. Nimittäin ainut

kriteeri sille, että kuvio on olemassa, on se, että on olemassa kuvaus kuviosta, joka on tehokkaampi kuvaustapa kuin itse bittikartta (Dennett 1991b, 34).⁹¹

Haugeland (1994) kuitenkin muistuttaa meitä siitä, että Dennett viittaa kuvion näkemisellä sen näkemisenä täydellisenä, eheänä kuviona. Dennett ei tarkoita kuvion näkemisellä kuvion näkyvän osan näkemistä, vaan pikemminkin sen eheän kuvion näkemistä, jonka täydennämme mielessämme näkyvillä olevasta ”rikkonaisesta” kuviosta. Lisäksi,

Dennett ei sano, että loput [eli päätelty osuus] kuviosta on ”oikeasti olemassa”, vaikkakin, jostain syystä, näkymätön – joten meidän täytyy päätellä se siitä osasta [kuviota], jonka näemme. (Haugeland 1994, 116.)

Tässä mielessä kuvion erottaminen vaatii aina havainnoijan, minkä takia kuvio on katsojan silmässä – tai paremminkin päättelijän mielessä (engl. in the sense of the understander; Haugeland 1994). On silti huomattava, että ellei kuvio ole säännönmukainen (ei-satunnainen), emme voi päätellä sitä täydellisenä näkyvän kuvion pisteiden perusteella. Tässä mielessä Dennettin vetoaminen matemaattiseen, ”objektiiviseen” määrittelyyn on perusteltua, vaikkakaan se ei tarkoita – toisin kuin Dennett ehkä antaa ymmärtää –, että eheä, täydellinen kuvio olisi olemassa intentionaalisesta asenteesta riippumatta.

Haugeland (1994) ja William Seager (2000) huomauttavat, että tästä syystä intentionaalinen asenne edellyttää mielen:

Et voi ymmärtää mieltä, ellet jo tiedä mitä mieli on, koska et voi ymmärtää mentaalisuutta ymmärtämättä intentionaalista asennetta, jo[n]ka [käyttäminen] vaatii, että ymmärrät koko joukon pohjimmiltaan mentaalisia käsitteitä (Seager 2000, 104).

Seager (2000) näkee, että juuri tästä syystä intentionaalista asennetta ei voida naturalisoida, sillä selitettävä ilmiö edellyttää itsensä selittäväksi tekijäksi. Toisin sanoen esimerkiksi uskomusta ei voida määritellä muuten kuin

⁹¹ Itse asiassa on olemassa kuvioita, joita vain tietynlaiset ihmiset, kuten harjaantuneet ammattilaisshakkipelaajat havaitsevat: shakin pelaajat voivat vilaukselta tunnistaa ja muistaa erilaisia shakkiasetelmia, joita tavalliset ihmiset eivät millään voi muistaa virheittä. Ks. esim. de Groot 1965 ja Saariluoma 1985.

intentionaalisin termein, mistä seuraa reduktion mahdottomuus. Dennett ei pyrikään naturalisoimaan intentionaalisia tiloja, eikä mielen edeltämisestä siinä mielessä muodostu ongelmaa. Seagerin mielestä on varsinainen arvoitus, että intentionaalinen asenne edellyttää mielen. Dennettin mielestä tämä ei kuitenkaan ole ongelma, vaan ainoastaan kuriositeetti (Dennett 2000, 353–354).⁹²

4.6.2 Epifenomenalismin vaara

Mielentilat ovat Dennettin mukaan abstractum-luokan olioita, jolloin niillä ei ole kausaalista vaikutusta. Sen tähden muun muassa Don Ross (2000) ja Alexander Rosenberg (1988) ovat kehottaneet häntä luopumaan illatum-abstractum –jaottelusta, koska uskomuksilla ei ole mitään reaalista vastinetta, jos ne ovat abstractum-luokan olioita.

⁹² Kuten Marc Slors (1996) panee merkille, ainoa keino selittää intentionaalisen asenteen omaksuminen on tehdä se behavioraalisin termein vetoamalla intentionaaliin termeihin, koska intentionaalinen asenne edeltää intentionaalisia termejä. Hän väittää ettemme kohtele kaikkia intentionaalisia systeemeitä samalla tavalla, mikä vie pohjan intentionaalisen asenteen omaksumisen selittämiseltä. Hän antaa tästä esimerkin (Slors 1996, 96–97; mukailtu):

Jos jokin artefaktinen intentionaalinen systeemi, kuten termostaatti tai tietokone, toimii virheellisesti, omaksumme sen suhteen suunnittelun tai fysiikan asenteen korjausta varten. Jos taas joku henkilö käyttäytyy kummallisesti tai irrationaalisesti tietyistä intentionaalista tulkinnasta käsin, vaihdamme intentionaalista tulkintaa (esim. ”ehkä hän ei usko sinua”, ”ehkä hän on nukkunut huonosti”) emmekä siirry tarkastelemaan henkilöä suunnittelun tai fysiikan asenteen kautta. Tietysti vakavissa tapauksissa esimerkiksi psykiatri saattaa tarkastella henkilöä suunnittelun tai biologian (koska biologiset eliöt ovat [evoluution] suunnittelemlia) asenteesta, mutta hän ei missään vaiheessa luovu intentionaalisen asenteen käytöstä vaan käyttää sitä biologisen asenteen rinnalla. Termostaatin korjaustapauksessa näin ei ole, sitä kohdellaan ainoastaan fysikaalisena systeeminä kunnes se on taas täysin kunnossa, jolloin sitä voidaan jälleen kohdella intentionaalisenä systeeminä.

Slorsin mielestä tämä erilaisten intentionaalisten systeemien kohtelun välinen ero osoittaa intentionaalisen asenteen omaksumisen selittämisen ongelman. Minusta Slors ei huomaa sitä olennaisuutta, että artefaktiset systeemit ovat meidän suunnittelemliamme ja luomiamme, joten tiedämme (tai ainakin systeemien asiantuntijat tietävät) täysin miten ne toimivat, emmekä tarvitse rinnalle intentionaalista tulkintaa. Ihmisten ja monimutkaisempien biologisten organismien tapauksessa näin ei ole asian laita. Me emme tiedä täysin, miten evoluution suunnittelemlat organismit toimivat emmekä siksi voi tarkastella niitä ainoastaan suunnittelun (biologian) tai fysiikan asenteesta. Joka tapauksessa Slors on oikeassa siinä, ettemme voi selittää intentionaalisen asenteen omaksumista, koska kohtelemme joitakin intentionaalisia systeemejä (esim. termostaatteja) eri tavalla kuin toisia (esim. koiria).

Jos abstractum-luokan mukaisia mentaalisia termejä ei voida luonnehtia kausaalisiin termein ja jos mentaalista kausaalisuutta ei voida selittää fysikaalisen kausaation muodossa, herää helposti ajatus epifenomenaalisuudesta.⁹³ Tähän ovat kärkkäästi tarttuneet Jerry Fodor ja Ernest Lepore (1993). He kysyvät (Fodor & Lepore 1993): miten voi olla niin, etteivät uskomukset olisivat todella olemassa, jos kyvyllä uskoa tosiin uskomuksiin on selviytymisarvoa? Miten se, mitä organismilla ei ole, voi kuitenkin auttaa sen selviytymistä luonnossa? Lisäksi voidaan heidän tapaansa kysyä, miten intentionaalisten idiomien pohjalta voidaan hyödyllisesti ennustaa mitään, jos intentionaalisia tiloja ei todella ole olemassa?

Dennett (1993, 126) puolestaan esittää retorisen vastakysymyksen: miten se, mitä purjeveneellä ei ole, esimerkiksi matala painopiste, voi auttaa sen ”taistelussa kaatumista vastaan”? Dennett (2000, 357) vakavoituu ja kysyy: ”[m]itä Connor teki hetkessä *aiheuttaakseen* veneensä nopeamman vauhdin?” Ja vastaa:

Hän alensi sen painopistettä. Totta kai hän teki tämän siirtämällä pyydysvälineitä, laittamalla lyijyharkkoja pilssiin tai korvaamalla maston kevyemmällä mastolla [...] mutta mikä *aiheutti* veneen [vauhdin] parannuksen, oli sen painopisteen alentaminen. (Dennett 2000, 357–358.)

Hänen mukaansa vetoaminen painopisteen alentamiseen ei ole pelkkä oikotie selittämisessä, vaan yleistys siitä, miksi jotkin muutokset – joita on mahdollisesti ääretön määrä – aiheuttavat veneen vauhdin parantumisen eli yleisesti tietyn ilmiön (Dennett 2000, 358). Samaten:

Ajattele kyltin ”ilmainen lounas” väkeä vetävää voimaa, kun se sijoitetaan ravintolan ikkunaan, ja vertaa sen voimaa newyorkilaisessa

⁹³ Huomaa silti, että Dennett on ehdottomasti tieteellinen materialisti (tai naturalisti). Tieteellisten materialistien (tai naturalistien) voidaan yleensä ajatella hyväksyvän epistemologinen reduktio, mutta ei täysivaltaista ontologista reduktiota. Tämän mukaisesti usein ajatellaan mielen ilmiöiden olevan ontologisesti emergenttejä, mutta silti riippuvaisia alemman tason aineellisista tapahtumista ja ”rakennusaineista”. Epistemologinen reduktio ei heillä silti tarkoita sitä, että kaikkia ilmiöitä voisi selittää jännöksettä ”perustavanlaatuisimmalla” selitysmallilla (so. fysiikalla), vaan sitä, että eri selitysmallien välille on rakennettavissa vahva silta. Ks. esim. Bunge 1980.

ravintolassa sen voimaan tokiolaisessa ravintolassa. Intentionaalinen taso on ilmiselvästi oikea taso ennustaa ja selittää sellaisia kausaalisia voimia; kyltti tuottaa luotettavammin tietyn uskomuksen yhdessä havainnoitsijoiden populaatiossa kuin toisessa, ja muutokset kyltin typografian värissä eivät ole yhtä ennustavia muutoksia väkeä vetävässä voimassa kuin muutokset (havaittavissa olevassa) merkityksessä. (Dennett 1991b, 44 alaviite.)

Dennett jatkaa:

Se tosiasia, että säännönmukaisuudet, joihin nämä menestykselliset ennusteet pohjautuvat, ovat tehokkaasti kuvattavissa (vain) intentionaalisiin termein eivätkä ole johdetut ”peittävästä laeista” ei osoita, että säännönmukaisuudet eivät olisivat ”kausaalisia”; se vain osoittaa, että filosofit ovat usein luottaneet riutuneisiin kausaalisuuden käsityksiin, jotka on johdettu valikoidulla keskittymisellä varsin moniin fysiikasta ja kemiasta otettuihin esimerkkeihin. (Dennett 1991b, 44.)

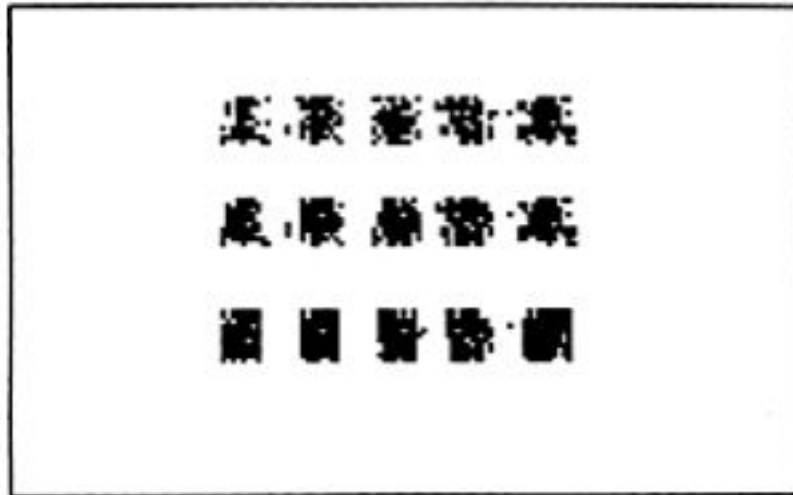
Dennett selvästikin näkee vaikuttavan syyn selitysten olevan riittämättömiä; hänen mukaansa intentionaalisia selityksiä tarvitaan eli hän peräänkuuluttaa Aristoteleen finaalisien syyn tarvetta. Mutta miten kummassa nämä teleologiset intentionaaliset lait voivat olla kausaalisia, jos ne eivät ole todellisia?

Ennen kysymykseen vastaamista on syytä huomata, että intentionaaliset tilat ovat näkyviä ja kausaalisia vain, jos tehdään rationaalisuus- ja optimaalisuusoletukset. Niitä ei voida löytää empiirisiin keinoin, koska ne ovat näkyviä ainoastaan intentionaalisesta asenteesta, muistuttaa Christopher Viger (2000). Tästä syystä ne eivät ole perinteisen ontologisen taksonomian mukaan todellisia ja kausaalisia.

Samalla lailla kuin eheä kuvio on olemassa, ovat myös intentionaaliset lait, vaikka ne eivät ole missään eksplisiittisesti edustettuina. Kuvion säännönmukaisuus (ja intentionaaliset lait) voivat Dennettin mukaan johtua pelkästään pienten, konkreettisten yksityiskohtien tilastollisesta efektistä tai summautumisesta, joka approksimoi ideaalia järjestystä (Dennett 1991b, 43).⁹⁴

⁹⁴ Vrt. luvussa **Virhe. Viitteen lähde ei löytynyt.** ”Semantiikan luominen syntaksista” esitettyyn aivojen kykyyn approksimoida semantiikkaa.

Toisin kuin KUVION 1 kehykset on luotu diskreetillä tavalla (10 mustaa pistettä, 10 valkoista pistettä, 10 mustaa pistettä jne.), joihin on sekoitettu kohinaa, KUVION 2 ylimmällä rivillä olevat kuviot on luotu tilastollisesti niin, että ne pohjautuvat normaalijakaumaan, jonka keskikohdat ovat kohdissa $x = 10, 30, 50, 70$ ja 90 . KUVION 2 loput rivit on luotu yläriivin kuvioiden kontrastin peräkkäisellä lisäämisellä. (Dennett 1991b, 44.)



Kuvio 2. Tilastollisella summautumisella luodut kuviot (Dennett 1991b, 45).

Selvästikin KUVION 2 alimman rivin oikeanpuoleisin kuvio approksimoi KUVIOSSA 1 esitettyjä kuvioita, vaikka toinen on tuotettu eksplisiittisten, diskreettien sääntöjen perusteella ja toinen on pelkän tilastollisen summautumisen aikaansaamaa. Kuvion luomisen sääntöjen ja intentionaalisten lakien ei tarvitse olla yhtään enempää eksplisiittisesti edustettuina kuin lintujen siipien muotoilussa näkyvät aerodynamiikan periaatteet (Dennett 1991b, 43). Meille näkyy kuvion reunojen selkeä särmikkyys, jolloin oletamme kuvion luontiperusteiden olevan diskreettejä. Kyse on kuitenkin meidän havaintokyvyistämme. Samalla lailla kuin kuvio näkyy meille eheänä, myös intentionaaliset lait ovat olemassa. Ne odottavat sitä, että joku ne havaitsee. On kuitenkin hyvin arveluttavaa, että maailmassa esiintyisi puhdasta

diskreettiyttä, puhtaita rajoja tai puhdasta määritettävyyttä. Kieli saa näyttämään (arkipsykologian) intentionaaliset ”lait” laeilta eli ehdottomilta ja tarkoilta (vrt. Dennett 1991b, 45).

Toisin sanoen on turha odottaa, että kuvio tai intentionaaliset lait olisivat olemassa havaitisijasta riippumatta. Me intentionaalisten lakien havaitisijat teemme rationaalisuusoletuksen, jonka takia ne meille näkyvät. Ne ovat idealisoituja.

4.6.3 Suhde pragmatismiin

Dennettin voisi sanoa hylkäävän perinteisen metafysiikan ja pyrkivän irrottautumaan perinteisestä kielipelistä. Richard Rorty (1993) onkin kehottanut Dennettiä radikalisoimaan näkemystään ja luopumaan realismista sekä ontologia-keskustelusta yleensä. Rortyn radikaalin pragmatismien mukaan on kyseenalaista puhua siitä, onko mitään todella olemassa. Dennett ei halua kuitenkaan lähteä mukaan Rortyn kelkkaan. Dennettin mielestä on huomattavasti eroa sillä, sanommeko jonkun olevan oikeasti olemassa vai olemassa vain niin sanotusti pragmaattisena konstruktiona (Dennett 1993, 233–234).

Hän antaa tästä esimerkin puhumalla Fort Knoxin kullasta. Fort Knoxin kullan merkitykseen maailman taloudelle ei vaikuta sen varsinainen sijainti, vaan se, että ihmiset uskovat kullan olevan Fort Knoxissa. Jos Fort Knoxin kaikki kulta siirrettäisiin salaa toiseen sijaintiin, intentionaalinen objekti ”kulta Fort Knoxissa” olisi silti koskematon ja tekisi työtään, vaikka itse kulta olisi siroteltu fyysisesti eri paikkoihin. (Dennett 1993, 235.) Tämä osoittaa Dennettin perinteisessä metafysiikassa pysymisen, minkä hän myöntää itsekkin sanomalla: ”olen oikeastaan silti jokseenkin konservatiivinen Standardin tieteellisen epistemologian ja metafysiikan omaksumisessa” (Dennett 1993, 234).

On loppujen lopuksi hankalaa sanoa, mikä täsmälleen on Dennettin käsitys tai ”ismi”, johon hän lukeutuu. Tällaista sitoutumista hän ei itse asiassa tahdo

tehdä itsekään (Dennett 1993, 212). Bjørn Ramberg (1999) ottaa härkää sarvista ja luokittelee Dennettin pragmatistiksi, koska Dennett samoin kuin muut pragmatistit näkevät pitäytymisen perinteisessä mielen käsityksessä ja filosofisessa traditiossa ongelmana.⁹⁵ Ongelma ei ole heidän mielestään itse ilmiö, vaan se, millä tavalla siitä puhutaan.

Yksi kysymys jää ilmaan: miten tyydyttävä ratkaisu kielipelistä irtisanoutuminen ja uuden tarjoaminen tilalle on? Tätä voidaan pragmaattisesti ja pragmatismien mukaisesti tarkastella uuden kielipelin tarjoaman hyödyn perusteella. Onko Dennettin tarjoama intentionaalisuuden käsitys hyödyllinen? Varmaa on ainakin, että se on hyödyllisempi ja tieteellisempi kuin arkipsykologian tarjoama implisiittinen käsitys intentionaalisuudesta.

⁹⁵ Ramberg luokittelee Dennettin pragmatistiksi käyttäen esimerkkinä Dennettin tietoisuuskäsitystä. Siinä ympäristössä sen sanominen, että ”pragmatistin [Dennettin] filosofisena työnä ei ole teorioiden tarjoaminen” on oikeutettua (Ramberg 1999, 80). Dennettin intentionaalisuuskäsityksen parissa voidaan kylläkin jo puhua lähes teorian rakentamisesta.

5 YHTEENVETO

Tässä tutkielmassa tutkin Daniel C. Dennettin intentionaalisuuden käsitystä. Tavoitteenani oli kuvata Dennettin käsitys seikkaperäisesti. Tätä varten käsittelin ensin lyhyesti intentionaalisuus-käsitteen taustaa ja problematiikkaa. Tämän jälkeen loin katsauksen Dennettin näkemyksen ytimeen eli intentionaaliseen asenteeseen, jolla Dennett on pyrkinyt korvaamaan perinteisen puheen intentionaalisuudesta (tai mielestä). Lopuksi tarkensin esitystä ja paneuduin siihen syvemmälle tarkastelemalla Dennettin näkemykseen kuuluvia erityisteemoja sekä Dennettin näkemyksen kritiikkiä.

Alussa näimme, että intentionaalinen relaatio ei ole kuvattavissa relaatiologiikalla, joka aiheuttaa ongelmia tieteen yhtenäisyydelle ja naturalismille. Huomasimme myös, että arkipsykologia ei sellaisenaan sovi mielen tieteellisen lähestymistavan pohjaksi. Dennett on vastannut näihin haasteisiin omalla näkemyksellään intentionaalisuudesta, joka ei ole perinteinen naturalistinen lähestymistapa. Dennett on esitellyt intentionaalisen asenteen tai strategian, jonka tarkoituksena on tarjota vakaampi pohja tieteelle intentionaalisuudesta puhumiseksi.

Intentionaalisen asenteen läpi oliot näyttäytyvät intentionaalisina systeeminä, joille liitetään relevantteja uskomuksia, haluja ja muita mielentiloja, jos oliot ovat optimaalisesti suunniteltuja ja rationaalisia. Näiden uskomuksien ja halujen mukaan systeemin toimintaa ennustetaan ja selitetään. Dennettin näkemyksen mukaan olio on intentionaalinen, jos sen käyttäytymistä voidaan luotettavasti ennustaa intentionaalisena systeeminä. Intentionaalisia systeemejä voivat olla niin ihmiset kuin termostaatitkin; systeemin koostumuksella ei ole merkitystä. Dennettin antiessentialistinen näkemys painottaa intentionaalisuuden määrällistä aspektia, mitään laadullista eroa ei ole. Ainoa kriteeri systeemin intentionaalisuudelle on, että sen käyttäytymistä voidaan luotettavasti ennustaa tehokkaammin käyttämällä intentionaalisia termejä muiden (esim. fysiikan tai suunnittelun) termien sijaan; joidenkin olioiden

toiminta on ymmärrettävissä (meidän näkökulmastamme) ainoastaan intentionaalisesta asenteesta.

Dennett on korvannut arkipsykologian kahdella tieteellisellä teoriallaan, intentionaalisten systeemien teorialla ja alipersonallisella kognitiivisella psykologialla. Näistä ensimmäinen on niin sanotusti ”heuristinen silaus” (Dennett 1969), joka on eräänlainen oikotie selittämisessä. Sen tarkoituksena on selittää intentionaalisten systeemien toimintaa abstractum-luokan termeillä eli niin sanotuilla loogisilla konstruktiolla, joita tässä tapauksessa ovat uskomukset ja halut. Näillä abstractum-luokan olioilla ei ole (perinteisessä mielessä) kausaalista voimaa. Alipersonallinen kognitiivinen psykologia taas tutkii intentionaalisten systeemien alajärjestelmiä ja sisäisiä prosesseja (eli toteutusta) käyttäen hyväksi illatum-luokan termejä eli postuloituja teoreettisia termejä, joilla puolestaan on kausaalista voimaa. Dennettin mukaan intentionaalisten systeemien teoriaa ei voida redusoida alipersonalliseen kognitiiviseen psykologiaan, vaan intentionaalisella puheella on oma autonominen asemansa alipersonallisen kognitiivisen psykologian ja ekstensionaalisen puheen ohella. Koska Dennett pitää mielentiloja loogisina konstruktiona ja hyväksyy Quinen käännöksen epätäydellisyys -teesin, hänen mukaansa mielentiloja ei voida löytää missään mielessä aivoista tai fyysisestä toteutuksesta eikä siten identifioida niihin.

Eniten keskustelua herättänyt aspekti Dennettin näkemyksessä on juuri mielentilojen yhtäältä instrumentalistinen ja toisaalta antireduktionistinen luonne. Koska Dennettin mukaan intentionaaliset tilat ovat olemassa oliolla vain siksi, että ne sille liitetään, on moni luokitellut hänet instrumentalismin kannattajaksi. Dennett ei ole suostunut tulla luokitelluksi instrumentalistiksi, mutta ei ole myöskään hyväksynyt realistista eikä eliminativistista lähestymistapaa. Sitä vastoin hän on yrittänyt näyttää kriitikoilleen heidän dikotomiansa ja luokittelunsa olevan vajavainen. Ennen kaikkea hän syyttää heitä siitä, että heidän kausaation käsityksensä on liian kapea.

Jossain määrin ongelmallista on se, että intentionaalista asennetta käyttäökseen täytyy itse olla intentionaalinen systeemi. (Ainoastaan me voimme tehdä tämän niin sanotusti tietoisesti; myös muut eläimet kyllä näyttävät käyttävän intentionaalista asennetta, mikä osoittaa intentionaalisen asenteen omaksumisen evolutiivisen jatkumon.) Me voimme attribuoida muille mielentiloja, koska omaamme mentaalisen käsitteistön – mutta mitä loppujen lopuksi tarkoittaa se, että systeemillä on jokin uskomus? Sitä, että systeemin käyttäytymistä voidaan ennustaa attribuomalla sille tämä uskomus. Näitä mentaalisia käsitteitä ei ole kuitenkaan määritelty mitenkään, vaan ne otetaan annettuina. Dennett kieltää epistemologisen reduktion intentionaalisesta kielestä ekstensionaaliseen, mikä jättää mentaalisten tilojen luonteen arvoitukseksi. Toisaalta hän silti väittää, ettei raja intentionaalisen ja ekstensionaalisen kielen tai intentionaalisten sistemien teorian ja alipersonallisen kognitiivisen psykologian välillä ole tarkka, vaan ne lähestyvät toisiaan, mikä on ristiriidassa antireduktionistisen väitteen kanssa.

Dennettin intentionaalisuuden käsitys ei ole kovin intuitiivinen, minkä voisi ajatella johtuvan juurtumisesta arkipsykologian (synnynnäiseen) esiyymmärrykseen. Intuitiivisuus ei takaa totuutta eikä riitä pitkälle, kuten on tieteen historiassa nähty. Voisi ajatella, että Dennett on pragmaattisesti pyrkinyt korvaamaan perinteisen käsityksemme intentionaalisuudesta toisella, jonka ala eli ekstensio on laajempi. Mahdollisena vaarana on kuitenkin se, että uuden käsityksen merkitys eli intensio ei pysy enää samana. Moni syyttäisikin ja syyttääkin Dennettiä siitä, että hänen intentionaalisuus-käsityksensä ei ole se, mistä he ovat puhuneet, vaan jotakin aivan muuta.⁹⁶

⁹⁶ Tähän tulokseen päätyy Huttunen (1996, 84) todetessaan, että [n]äyttää siltä, että puhuessaan intentionaalisuudesta, Dennett puhuikin jostain, josta ei yleensä käytetä termiä intentionaalisuus”. Antti Revonsuo on lähes kirjaimellisesti samaa mieltä Huttusen kanssa Dennettin tietoisuuskäsityksen kohdalla kirjoittaessaan seuraavaa:

Dennettin teoriassa tietoisuus on jotain aivan muuta kuin se tietoisuus, jota me intuitiivisesti ajattelemme tai se, josta empiirisissä tieteissä puhutaan tietoisena prosessointina. Itse asiassa Dennett kyllä ratkaisee tietoisuuden ongelman: ainoa

Tietojenkäsittelyyn voidaan Dennettin intentionaalisuuden käsityksestä johtaa kaksi näkökohtaa. Tekoälyn tutkijoiden kannalta on huojentavaa kuulla, että artefaktinen intentionaalisuus ei eroa missään laadullisessa mielessä ihmisen intentionaalisuudesta. Kyseessä on pelkästään intentionaalisuuden erilaiset ilmentymistavat tai -asteet. Koska tekoäly on insinööritiedettä, voidaan keinotekoista intentionaalisuutta ilmentävän ohjelman tai laitteen rakenne ymmärtää täysin, Dennettin termin, fysiikan asenteesta käsin. Näin ei ole kuitenkaan laita biologisten organismien, kuten ihmisten ja kissojen. Joka tapauksessa "omatekoisen" intentionaalisuuden eli tekoälyn tutkimus avaa ovia myös biologisen intentionaalisuuden ymmärtämiseen. Toinen näkökohta koskee käyttäjäystävällistä tietojenkäsittelyä. Koska ihmisille on luontevinta tarkastella komplekseja, (näennäisen) itseohjautuvia järjestelmiä intentionaalisina systeemeinä, tätä taipumusta tulisi vahvistaa entisestään järjestelmiä suunniteltaessa. Nykyisellään älykkäät järjestelmät monesti vaikuttavat ensisijaisesti intentionaalisilta systeemeiltä, mutta usein ne tekevät niin paljon poikkeuksia rationaalisuusoletuksen suhteen, että intentionaalista strategiaa pyritään paikkaamaan muilla strategioilla. Ihannetapaus käyttäjäystävällisen tietojenkäsittelyn näkökulmasta olisi saada ihmisten käyttöön rationaalisuutta mahdollisimman paljon approksimoivia systeemejä, jolloin järjestelmän aiheuttama kognitiivinen rasitus ihmiselle olisi mahdollisimman pieni. Kognitiivisen rasituksen määrää voitaisiin testata myös empiirisin menetelmin "eriateisilla" älykkäillä järjestelmillä. Ihannetapauksessa ihminen olisi tekemisissä älykkään järjestelmän kanssa samalla tavalla, kun hän on toisten rationaalisten olentojen, kuten ihmisten, kanssa. Tässä nähdään käyttäjäystävällisen tietojenkäsittelyn side tekoälyyn; mahdollisimman intelligentti järjestelmä on Dennettin viitekehyksen pohjalta kognitiivisesti ja luontaisesti ihmiselle helpoin ja tutuin järjestelmä.

Olisi mielenkiintoista jatkon kannalta tarkastella Dennettin intentionaalisuuskäsityksen varaan rakennettuja muita mieleen liittyviä seikkoja. Hän on rakentanut sen varaan käsityksiään moraalisuudesta, tekoälystä ja tietenkin tietoisuudesta. Erityisesti olisi kiinnostavaa asettaa rinta rinnan Dennettin intentionaalisuuden ja tietoisuuden käsitykset. Hän on ne eriyttänyt tuotannossaan, mutta niillä on silti läheinen suhde; edeltäähän nimittäin Dennettin näkemyksessä intentionaalisuus tietoisuutta. Toiseksi, olisi hedelmällistä suhteuttaa Dennettin näkemys hänen näkemystään lähellä oleviin näkemyksiin, kuten Ruth Millikanin ja Quinen perintöä lähellä olevaan Donald Davidsonin näkemykseen, mutta myös loitolla olevaan Jerry Fodorin näkemykseen. Vertailun avulla voitaisiin nähdä, kuinka paljon viime kädessä Dennettin intentionaalisuuden käsitys eroaa muista käsityksistä. Dennett vetoaa paljon evoluutioteoriaan filosofiassaan. Tätä aspektia pitäisi etenkin tarkastella nyt, kun hän on tuotannossaan vahvasti siirtynyt yhä enemmän ja enemmän evoluutioteorian pariin. Tällaiset kysymykset ovat kuitenkin kysymyksiä jatkotutkimukselle.

LÄHDELUETTELO

- Aunger, R. (toim.) 2002a. *Darwinizing Culture: The Status of Memetics As a Science*. Oxford: Oxford University Press.
- Aunger, R. 2002b. Introduction. Teoksessa Aunger 2002a, 1–24.
- Baker, L. R. 1989. Instrumental Intentionality. *Philosophy of Science* 56(2), 303–316.
- Baker, L. R. 1994. Content Meets Consciousness. *Philosophical Topics* 22(1–2), 1–22.
- Baron, J. 1998. *Judgment Misguided: Intuition and Error in Public Decision Making*. Oxford: Oxford University Press.
- Bechtel, W. 1985. Realism, Instrumentalism, and the Intentional Stance. *Cognitive Science* 9(4), 265–292.
- Block, N. 1978. Troubles with Functionalism. Teoksessa C. Savage (toim.) *Perception and Cognition: Issues in the Foundations of Psychology*. Minneapolis: University of Minneapolis Press, 261–325.
- Block, N. 1980. Introduction: What is Functionalism? Teoksessa N. Block (toim.) *Readings in Philosophy of Psychology, Volume 1*. Lontoo: Methuen, 171–184.
- Block, N. 1993. Review of *Consciousness Explained*. *The Journal of Philosophy* 90(4), 181–193.
- Brentano, F. 1973. *Psychology from an Empirical Standpoint*, Lontoo: Routledge & Kegan Paul. (Alkuteos on julkaistu vuonna 1874 nimellä *Psychologie vom Empirischen Standpunkt*.)
- Brook, A. & Ross, D. (toim.) 2002a. *Daniel Dennett*. Cambridge: Cambridge University Press.

- Brook, A. & Ross, D. 2002b. Dennett's Position in the Intellectual World. Teoksessa Brook & Ross 2002a, 3–40.
- Bunge, M. 1980. *The Mind-Body Problem: A Psychobiological Approach*. Oxford: Pergamon Press.
- Chaitin, G. 1975. Randomness and Mathematical Proof. *Scientific American* 232(5), 47–52.
- Cherniak, C. 1981. Minimal Rationality. *Mind* 90(358), 161–183.
- Chisholm, R. 1957. *Perceiving: A Philosophical Study*. Ithaca: Cornell University Press.
- Churchland, P. M. 1979. *Scientific Realism and the Plasticity of Mind*. Cambridge: Cambridge University Press.
- Churchland, P. M. 1981. Eliminative Materialism and the Propositional Attitudes. *Journal of Philosophy* 78(2), 67–90.
- Churchland, P. S. 1986. *Neurophilosophy: Towards a Unified Theory of Mind/Brain*. Cambridge, Massachusetts: The MIT Press.
- Clark, A. 1994. Beliefs and Desires Incorporated. *Journal of Philosophy* 91(8), 404–425.
- Dahlbom, B. (toim.) 1993a. *Dennett and His Critics*. Oxford: Blackwell.
- Dahlbom, B. 1993b. Mind is Artificial. Teoksessa Dahlbom 1993a, 161–183.
- Davidson, D. 1967. Truth and Meaning. *Synthese* 17(3), 304–323.
- Davidson, D. 1973. Radical interpretation. *Dialectica* 27(3–4), 313–328.
- Darwin, C. 1988. Lajien synty. Hämeenlinna: Karisto. (Alkuteos on julkaistu vuonna 1859 nimellä *On the Origin of Species by Means of Natural Selection*.)
- Dawkins, R. 1989. Sokea kelloseppä. Juva: WSOY. (Alkuteos on julkaistu vuonna 1986 nimellä *The Blind Watchmaker*.)

- Dawkins, R. 1993. *Geenin itsekkyyks*. Helsinki: Art House. (Alkuteos on julkaistu vuonna 1976 nimellä *The Selfish Gene*.)
- De Groot, A. 1965. *Thought and Choice in Chess*. Haag: Mouton.
- Dennett, D. 1969. *Content and Consciousness*. Lontoo: Routledge & Kegan Paul.
- Dennett, D. 1978a. *Brainstorms: Philosophical Essays on Mind and Psychology*. Montgomery, Vermont: Bradford Books.
- Dennett, D. 1978b. Current Issues in the Philosophy of Mind. *American Philosophical Quarterly* 15(4), 249–262.
- Dennett, D. 1980. The Milk of Human Intentionality. *Behavioral and Brain Sciences* 3(3), 428–430.
- Dennett, D. 1982a. How to Study Human Consciousness Empirically, or Nothing Comes to Mind. *Synthese* 53(2), 159–180.
- Dennett, D. 1982b. The Myth of the Computer: An Exchange. *New York Review of Books* 29(11), 56.
- Dennett, D. 1983. Intentional Systems in Cognitive Ethology: The “Panglossian Paradigm” Defended. *Behavioral and Brain Sciences* 6(3), 343–355.
- Dennett, D. 1984. *Cognitive Wheels: The Frame Problem of AI*. Teoksessa C. Hookway (toim.) *Minds, Machines and Evolution*. Cambridge: Cambridge University Press, 129–151.
- Dennett, D. 1987. *The Intentional Stance*. Cambridge, Massachusetts: The MIT Press.
- Dennett, D. 1988a. Précis of *The Intentional Stance*. *Behavioral and Brain Sciences* 11(3), 495–504.
- Dennett, D. 1988b. Quining Qualia. Teoksessa A. Marcel & E. Bisiach (toim.) *Consciousness in Contemporary Science*. Oxford: Oxford University Press, 42–77.

- Dennett, D. 1991a. Mother Nature versus the Walking Encyclopedia: A Western Drama. Teoksessa W. Ramsey, S. Stich & D. Rumelhart (toim.) *Philosophy and Connectionist Theory*, Hillsdale, New Jersey: Erlbaum, 21–30.
- Dennett, D. 1991b. Real Patterns. *The Journal of Philosophy* 88(1), 27–51.
- Dennett, D. 1991c. Two Contrasts: Folk Craft versus Folk Science, and Belief versus Opinion. Teoksessa J. Greenwood (toim.) *The Future of Folk Psychology: Intentionality and Cognitive Science*. Cambridge: Cambridge University Press, 135–148.
- Dennett, D. 1993. Back from the Drawing Board. Teoksessa Dahlbom 1993a, 203–235.
- Dennett, D. 1994a. Get Real. *Philosophical Topics* 22(1-2), 505–568.
- Dennett, D. 1994b. Dennett, Daniel C. Teoksessa S. Guttenplan (toim.) *A Companion to the Philosophy of Mind*. Oxford: Blackwell, 236–244.
- Dennett, D. 1995a. *Darwin's Dangerous Idea: Evolution and the Meanings of Life*. New York: Simon and Schuster.
- Dennett, D. 1995b. Do Animals Have Beliefs? Teoksessa H. Roitblat & J.-A. Meyer (toim.) *Comparative Approaches to Cognitive Science*. Cambridge, Massachusetts: The MIT Press, 111–118.
- Dennett, D. 1999. *Tietoisuuden selitys*. Helsinki: Art House. (Alkuteos on julkaistu vuonna 1991 nimellä *Consciousness Explained*.)
- Dennett, D. 2000. With a Little Help from My Friends. Teoksessa Ross ja muut 2000, 327–388.
- Dennett, D. & Haugeland, J. 1987. Intentionality. Teoksessa R. Gregory (toim.) *The Oxford Companion to the Mind*. Oxford: Oxford University Press.
- Dummett, M. 1976. What is a Theory of Meaning? (II). Teoksessa G. Evans & J. McDowell (toim.) *Truth and Meaning: Essays in Semantics*, Oxford: Oxford University Press, 67–137.

- Feigl, H. 1958. The 'Mental' and the 'Physical'. Teoksessa H. Feigl, M. Scriven & G. Maxwell (toim.) *Concepts, Theories and the Mind-Body Problem*. Minnesota Studies in the Philosophy of Science, 2. Minneapolis: University of Minnesota, 370–497.
- Fodor, J. 1968. *Psychological Explanation*. New York: Random House.
- Fodor, J. 1975. *The Language of Thought*. New York: Crowell.
- Fodor, J. 1985. Fodor's Guide to Mental Representation: The Intelligent Auntie's Vade-Mecum. *Mind* 94(373), 76–100.
- Fodor, J. 1987. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, Massachusetts: The MIT Press.
- Fodor, J. 1990. *A Theory of Content and Other Essays*. Cambridge, Massachusetts: The MIT Press.
- Fodor, J. & Block, N. 1972. What Psychological States Are Not. *Philosophical Review* 81(2), 159–181.
- Fodor, J. & Lepore, E. 1992. *Holism: A Shopper's Guide*. Oxford: Blackwell.
- Fodor, J. & Lepore, E. 1993. Is Intentional Ascription Intrinsically Normative? Teoksessa Dahlbom 1993a, 70–82.
- Gopnik, A. 1993. How We Know Our Minds: The Illusion of First-person Knowledge of Intentionality. *Behavioral and Brain Sciences* 16(1), 1–14.
- Gould, S. J. & Lewontin, R. 1979. The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme. *Proceedings of the Royal Society B205(1161)*, 581–598.
- Griffin, R. & Baron-Cohen S. 2002. The Intentional Stance: Developmental and Neurocognitive Perspectives. Teoksessa Brook & Ross 2002a, 83–116.
- Haugeland, J. 1981. *Semantic Engines: An Introduction to Mind Design*. Teoksessa J. Haugeland (toim.) *Mind Design: Philosophy, Psychology, Artificial Intelligence*. Cambridge, Massachusetts: The MIT Press, 1–34.

- Haugeland, J. 1993. Pattern and Being. Teoksessa Dahlbom 1993a, 55–69.
- Haugeland, J. 1994. Understanding: Dennett and Searle. Teoksessa A. Revonsuo & M. Kamppinen (toim.) *Consciousness in Philosophy and Cognitive Neuroscience*. Hillsdale: Lawrence Erlbaum Associates Publishers, 115–128.
- Hayes, P. 1979. Naive Physics Manifesto. Teoksessa D. Michie (toim.) *Expert Systems in the Micro-electronic Age*. Edinburgh: Edinburgh University Press, 242–270.
- Huttunen, H. 1996. Daniel C. Dennettin teoria intentionaalisista systeemeistä. Helsingin yliopisto, Teologisen etiikan ynnä uskonnonfilosofian pro gradu -tutkielma.
- Kahneman, D., Slovic, P. & Tversky, A. (toim.) 1982. *Judgement Under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.
- Kamppinen, M. 2000. Emergenssin asema universumissa [online, viitattu 23.6.2004]. *Tieteessä tapahtuu* 18(2). Saatavilla WWW-muodossa <<http://www.tsv.fi/ttapaht/002/emergenssi.htm>>.
- Kenyon, T. 2000. Indeterminacy and Realism. Teoksessa Ross ja muut 2000, 77–94.
- Kilpinen, E. 2003. Pragmatismista ja sen yhteiskuntatieteellisestä merkityksestä. *Ajatus* 60, 213–239.
- Kornblith, H. (toim.) 1994. *Naturalizing Epistemology*. Cambridge, Massachusetts: The MIT Press.
- Lycan, W. 1981. Form, Function, and Feel. *The Journal of Philosophy* 78(1), 24–50.
- Lycan, W. 1988. Dennett's Instrumentalism. *Behavioral and Brain Sciences* 11(3), 518–519.

- Lycan, W. 1990. Introduction. Teoksessa W. Lycan (toim.) *Mind and Cognition: A Reader*. Oxford: Blackwell, 147–149.
- Marr, D. 1982. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. Cambridge, Massachusetts: The MIT Press.
- Millikan, R. 1984. *Language, Thought, and Other Biological Categories*. Cambridge, Massachusetts: The MIT Press.
- Nagel, T. 1986. *The View From Nowhere*. Oxford: Oxford University Press.
- Newell, A. 1982. The Knowledge Level. *Artificial Intelligence* 18(1), 87–127.
- Newell, A. 1988. The Intentional Stance and the Knowledge Level. *Behavioral and Brain Sciences* 11(3), 520–522.
- Nichols, S. 2002. Folk Psychology. Teoksessa L. Nadel (toim.) *Encyclopedia of Cognitive Science*. Lontoo: Nature Publishing Group, 134–140.
- Pagin, P. tuleva. Meaning Holism [online, viitattu 23.6.2004]. Teoksessa E. Lepore & B. Smith (toim.) *Handbook of Philosophy of Language*. Oxford: Oxford University Press. Saatavilla WWW-muodossa <<http://www.philosophy.su.se/personliga/pagin/mholism.pdf>>.
- Pierre, J. 2003. Intentionality [online, viitattu 27.6.2004]. Teoksessa E. Zalta (toim.) *The Stanford Encyclopedia of Philosophy* (Fall 2003 Edition). Saatavilla WWW-muodossa <<http://plato.stanford.edu/archives/fall2003/entries/intentionality/>>.
- Place, U. T. Is Consciousness a Brain Process. *British Journal of Psychology* 47(1), 44–50.
- Polanyi, M. 1966. *The Tacit Dimension*. Lontoo: Routledge & Kegan.
- Putnam, H. 1967. Psychological Predicates. Teoksessa W. Capitan & D. Merrill (toim.) *Art, Mind and Religion*. Pittsburgh: University of Pittsburgh Press. (Julkaistu uudelleen nimellä *The Nature of Mental States*.)

- Putnam, H. 1975. The Meaning of 'Meaning'. Teoksessa K. Gunderson (toim.) Language, Mind and Knowledge. Minnesota Studies in the Philosophy of Science, 7. Minneapolis: University of Minnesota, 131-193.
- Quine, W. V. O. 1951. Two Dogmas of Empiricism. *Philosophical Review* 60(1), 20-43.
- Quine, W. V. O. 1960. *Word and Object*. Cambridge, Massachusetts: The MIT Press.
- Ramberg, B. 1999. Dennett's Pragmatism. *Revue Internationale de Philosophie* 53(207), 61-86.
- Reichenbach, H. 1938. *Experience and Prediction*. Chicago: University of Chicago Press.
- Revonsuo, A. 1993. Tietoisuus – kognitiotieteen flogiston? Teoksessa T. Vadén (toim.) *Kognitiotiede, tekoäly ja filosofia*. Tampereen yliopisto, Filosofisia tutkimuksia Tampereen yliopistosta, 46, 57-69.
- Robinson, W. 2003. Epiphenomenalism [online, viitattu 27.6.2004]. Teoksessa E. Zalta (toim.) *The Stanford Encyclopedia of Philosophy* (Spring 2003 Edition). Saatavilla WWW-muodossa <<http://plato.stanford.edu/archives/spr2003/entries/epiphenomenalism/>>.
- Rorty, R. 1993. Holism, Intrinsicity, and the Ambition of Transcendence. Teoksessa Dahlbom 1993a, 184-201.
- Rosenberg, A. 1988. Will the Argument for Abstracta Please Stand Up? *Behavioral and Brain Sciences* 11(3), 526-527.
- Ross, D. 2000. Rainforest Realism: A Dennettian Theory of Existence. Teoksessa Ross ja muut 2000, 147-168.
- Ross, D., Brook, A. & Thompson, D. (toim.) 2000. *Dennett's Philosophy: A Comprehensive Assessment*. Cambridge, Massachusetts: The MIT Press.

- Ryle, G. 1949. *The Concept of Mind*. Lontoo: Hutchinson.
- Saariluoma, P. 1985. Chess Players' Intake of Task Relevant Cues. *Memory and Cognition* 13(5), 385–391.
- Salminen, K. 2004. Olioiden käyttäytymisen selittäminen ja ennustaminen Daniel C. Dennettin mukaan. Jyväskylän yliopisto, Tietojärjestelmätieteen kandidaatintutkielma.
- Seager, W. 2000. Real Patterns and Surface Metaphysics. Teoksessa Ross ja muut 2000, 95–132.
- Searle, J. 1980. Minds, Brains, and Programs. *Behavioral and Brain Sciences* 3(3), 417–457.
- Searle, J. 1982. The Myth of the Computer. *New York Review of Books* 29(7), 3–7.
- Searle, J. 1983. *Intentionality: An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press.
- Searle, J. 1988. The Realistic Stance. *Behavioral and Brain Sciences* 11(3), 527–529.
- Searle, J. 1992. *The Rediscovery of the Mind*. Cambridge, Massachusetts: The MIT Press.
- Seegerstråle, U. 2000. *Defenders of the Truth: The Battle for Science in the Sociobiology Debate and Beyond*. Oxford: Oxford University Press.
- Sellars, W. 1963. *Science, Perception and Reality*. Lontoo: Routledge & Kegan Paul.
- Siewert, C. 2003. Consciousness and Intentionality [online, viitattu 27.6.2004]. Teoksessa E. Zalta (toim.) *The Stanford Encyclopedia of Philosophy* (Fall 2003 Edition). Saatavilla WWW-muodossa <<http://plato.stanford.edu/archives/fall2003/entries/consciousness-intentionality/>>.

- Simon, H. 1957. *Models of Man*. New York: Wiley.
- Simon, H. 1970. *The Sciences of the Artificial*. Cambridge, Massachusetts: The MIT Press.
- Slors, M. 1996. Why Dennett Cannot Explain What It Is to Adopt the Intentional Stance. *The Philosophical Quarterly* 46(182), 93–98.
- Stein, E. 1996. *Without Good Reason: The Rationality Debate in Philosophy and Cognitive Science*. Oxford: Clarendon Press.
- Stich, S. 1983. *From Folk Psychology to Cognitive Science*. Cambridge, Massachusetts: The MIT Press.
- Turing, A. 1950. Computing Machinery and Intelligence. *Mind* 59(236), 433–460.
- Viger, C. 2000. Where Do Dennett's Stances Stand? Explaining Our Kind of Mind. *Teoksessa Ross ja muut* 2000, 131–146.
- Von Wright, G. H. 1971. *Explanation and Understanding*. Lontoo: Routledge & Kegan Paul.
- Weaver, W. 1998. Corporations as Intentional Systems. *Journal of Business Ethics* 17(1), 87–97.
- Wittgenstein, L. 1999. *Filosofisia tutkimuksia*. Juva: WSOY. (Alkuteos on julkaistu vuonna 1953 nimellä *Philosophical Investigations*.)