



UiO • Universitetet i Oslo

Liv, bevissthet og subjektivitet- et alternativt perspektiv på David Chalmers “hard problem”

Av Morten Audsen

Veileder: Professor Carsten Hansen

Masteroppgave i filosofi (FIL 4090)

60 studiepoeng

Institutt for filosofi, ide- og kunsthistorie og klassiske språk

Humanistisk fakultet

14.11.2020

Forord

Temaet for denne oppgaven har hatt en lang modningstid. Etter jeg sluttet halvveis gjennom hovedfaget for over 20 år siden har jeg alltid tenkt at jeg må komme tilbake til filosofien. Det var to mulige tema jeg kunne grave meg ned i, Kants filosofi eller Chalmers «hard problem». Den utløsende hendelsen for å starte denne reisen var boken *Other Minds: The Octopus and the Evolution of Intelligent Life* som ble lest to ganger fra perm til perm i en høstferie for to år siden. Da jeg i tillegg fikk se en liten blekksprut (cuttlefish) jakte etter reker i en bukt var saken avgjort. Resultatet av disse hendelsene er denne oppgaven. Det skulle også vise seg at jeg fikk litt bruk for Kant sine innsikter i visse deler av denne oppgaven.

Jeg vil takke min veileder Carsten Hansen for å gi meg krevende, men konstruktive tilbakemeldinger. Etter et langt fravær fra filosofi hopper jeg fort til konklusjoner, generaliserer for lett og overser viktige nyanser på jakt etter det jeg mener er viktig innsikt. Takket være Carstens gode råd og grundige gjennomlesninger har dette blitt vesentlig bedre. Alle snarveier, unøyaktigheter og spekulative krumspring som gjenstår er på min kappe. Jeg vil også takke min uformelle referansegruppe – Gunnar Aakvaag, Haakon Aakre og Jan Flæte – for moralsk støtte, ivrige diskusjoner, en skulder å gråte på og tålmodige forsøk på å forstå tanker som utviklet seg fra det uforståelige og forvirrede, til en gradvis mer sammenhengende historie. Sist, men ikke minst, vil jeg takke de som gjør at den jobben jeg har fått fra livet, blir til et lite eventyr hver eneste dag: Mari, Mie, Ole og Thea.

Sammendrag

Et sentralt problem i bevissthetsfilosofi er David Chalmers "hard problem"; hvorfor og hvordan vi har en indre subjektiv opplevelse. Dette problemet drives av det som fremstår som et fundamentalt skille mellom det materielle og det mentale. Motsatsen til det harde problemet er de lette: det å forklare hvilke kapabiliteter og kognitive egenskaper vi kan observere hos bevisste organismer, som hos oss mennesker. Chalmers argumenterer for dette skillet er basert på at de lette problemene er prinsipielt mulig å forklare gjennom en reduktiv forklaring som bygger på en funksjonell analyse, mens det harde problemet ikke er åpen for en tilsvarende analyse.

Jeg vil legge frem et sett argumenter som utfordrer dette skillet mellom lette og harde problemer. Dette vil jeg gjøre ved å utfordre premisset om at de lette problemene lar seg prinsipielt forklare gjennom en reduktiv analyse. Det er visse fenomener der en slik reduktiv analyse ikke er mulig, nemlig hvordan vi kan forklare levende organismer. Denne analysen peker derfor på at det er et sett forutsetninger som må være på plass for å forklare Chalmers lette problemer. Disse forutsetningene er basert på en forståelse av levende organismer som grunnleggende teleologisk. Levende organismer viser en aktiv målrettethet for å opprettholde seg selv i møtet med omgivelsene som vi ikke finner i andre naturlige fenomener. For å forklare de funksjonelle egenskapene Chalmers inkluderer i de lette problemene, må denne aktive målrettetheten forutsettes. Disse positive redegjørelsene vil basere seg på teorier fra Peter Godfrey-Smith, Terrence Deacon og Evan Thompson.

Ved å anerkjenne at en organismes funksjonelle egenskaper ikke bør forstås som *hendelser*, men som en organismes *handlinger*, vil jeg peke på at en slik aktiv realisering av funksjoner kan gi oss en grunn til å anta en tett forbindelse med en tilhørende indre opplevelse. Dette vil jeg først gjøre ved å følge Godfrey-Smith i hans redegjørelse av evolusjonen av kognitive handlinger hos levende organismer, som en konkret tilpasning hos en konkret organisme i en konkret kontekst. Gjennom en slik evolusjonær analyse vil jeg argumentere for at en redegjørelse av proto-kognitive og kognitive prosesser kan forstås som en forlengelse av de grunnleggende teleologiske prosessene som opprettholder livet til en organisme. Kognisjon kan sees på som et spesialtilfelle av slike prosesser som er evolusjonært utviklet for å etablere tilstrekkelig koordinering og kontroll av en komplisert flercellet organisme. Basert på dette vil jeg deretter utforske mulige spekulative teorier som ser på hvordan naturlige prosesser kan danne grunnlag for, og være kausalt integrert med, en organismes indre opplevelser. Til sist vil jeg legge frem et spekulativt argument som peker mot et mulig svar på Chalmers «hard problem».

Innholdsfortegnelse

Forord.....	1
Sammendrag.....	2
1: Innledning	5
1.1 Teoretisk bakgrunn.....	8
1.2 Mekanisme eller teleologi	9
1.3 Reduksjonisme eller struktur.....	12
2: Etablering av problemstillingen	16
2.1 Bevissthetens ontologiske grunnlag.....	18
2.2 Chalmers lette og vanskelige spørsmål om bevisstheten.....	22
2.3 utfordringer med å se på levende organismer som en Turing-robot	24
2.3.1 Komputasjon og Shannon-informasjon.....	24
2.3.2 Evolusjon og semantisk informasjon	31
3: Evolusjon, liv og bevissthet	36
3.1: Fysikk, termodynamikk og liv - livets termodynamiske ubalanse.....	39
3.2: Liv, barrierer, reproduksjon og autopoiesis	42
3.2.1 Gener og liv.....	45
3.2.2 Autogener som den første primitive livsformen- Deacons prinsippforklaring.....	50
3.3: Organismens navigering av omgivelsene	56
3.3.1 Organismens proto-kognitive egenskaper	58
3.3.2 Mål og mening	60
3.3.3 Fra er til bør	65
3.4: Evolusjon fra encellet til flercellede organismer	72
3.4.1 De første subjektive opplevelsene.....	72
3.4.2 Utviklingen av kognisjon.....	75
3.4.3 Tilbake til Chalmers lette og vanskelige spørsmål	81
3.5: Bevissthet og metafysikk	83
3.5.1 Sammenhengen mellom Chalmers lette og harde spørsmål	84

3.5.2 Hvorfor subjektive opplevelser?	89
4: Avslutning	99
Bibliografi.....	101

1: Innledning

David Chalmers "hard problem" om bevissthet gir oss følgende utfordring. Det er mulig å forklare hva mennesker gjør, sier, og til og med tenker - våre kapabiliteter - ved å gjennomføre en analyse av hjernens funksjoner i samspill med resten av kroppen og i relasjon til våre omgivelser. Dette beskriver Chalmers som relativt lette problemer. Men *hvorfor* og *hvordan* disse funksjonene er forbundet med en subjektiv opplevelse, altså en kvalitativ indre opplevelse som er noe for individet, dette er Chalmers "hard problem". Mye har blitt sagt og skrevet om dette temaet uten at det foreligger en tilfredsstillende forklaring.

Hva er en tilfredsstillende forklaring på et fenomen? Vi vet hvorfor døgnet varer i 24 timer (jordens rotasjon rundt sin egen akse), hvorfor det er årstider (jordens helning kombinert med bevegelsen rundt solen) og hvorfor det er tidevann (månen og jorden trekker på hverandre gjensidig og månens gravitasjonsfelt tiltrekker seg havet). Her har vi forklart et fenomen basert på andre etablerte sannheter- det opprinnelige problemet forklares gjennom andre mer grunnleggende forhold. Det vil ikke være mulig å forklare alle deler av "the hard problem" på denne måten. Følgelig kan det hevdes at kjernen i problemet vil bestå. Jeg mener imidlertid at det er mulig å sette problemet inn i en annen kontekst enn det Chalmers gjør, noe som vil gi nye perspektiver og redusere dette gapet som det intuitivt synes å være mellom det mentale og det materielle. I sin kjerne vil dette nye perspektivet være å argumentere for at levende organismer på en grunnleggende måte må forstås som subjekter; som aktive og målrettede. Jeg vil argumentere for at Chalmers "hard problem" bygger på en implisitt antakelse om at mennesker og andre levende vesener er grunnleggende sett objekter, slik at man ikke tar inn over seg dette særskilte fenomenet det er å være levende. Hva som ligger i disse begrepene, vil jeg komme tilbake til.

Det sentrale grepet jeg skal gjøre, et grep jeg mener er nødvendig for å finne en riktig vinkling på analysen av "the hard problem", er derfor å snu fokuset vekk fra et perspektiv som grunnleggende sett baserer seg på fysikk og i stedet se mot biologi; på liv. En sentral del av utfordringen til Chalmers er nemlig å finne en forbindelse mellom den fysiske virkeligheten, det materielle grunnlaget, og hvordan dette er forbundet med både våre objektive kapabiliteter og den subjektive opplevelsen- den kvalitative opplevelsen vi har "fra innsiden". Dette er et kjent og gammelt problem som i hvert fall er tydelig artikulert av Leibniz (Chalmers 2020, s 5). Chalmers knytter nemlig diskusjonen om "the hard problem" til filosofiske diskusjoner om hva som er det metafysiske grunnlaget for det som eksisterer i

verden. Chalmers selv trekker konklusjonen fra “the hard problem” at materialisme, må være feil eller ufullstendig (Chalmers 2002).

Materialisme, eller søskenbarnet fysikalisme, er på mange måter *de rigueur* innen naturvitenskapen i dens forsøk på å forstå den fysiske verden. Man finner et fenomen og forsøker å analysere hvordan dette oppstår, hvilke egenskaper det har og hvilke relasjoner det har med andre fenomener basert på å undersøke de materielle bestanddelene. Dette er i praksis en reduksjonistisk analyse der man ser hvordan deler og helhet henger sammen i fenomenet vi har foran oss, gjerne ved å undersøke årsaksmekanismene til de underliggende materielle byggeklossene for å forklare systemegenskaper til fenomenet. Noen eksempler på slike forklaringsmodeller er at temperatur er molekyler i bevegelse; at vann samler seg i dråper på grunn av van der Waalske krefter som virker mellom molekylene i tillegg til kjemiske forbindelser; eller forklaringen av planetenes bevegelse i solsystemet gjennom gravitasjonskraften som virker gjensidig på hverandre (alternativt krumning av rommet).

Gjennom disse teoriene er det etablert så fantastisk mye kunnskap om hvordan verden er bygget opp. Man kan eksempelvis sende inn en blodprøve til genetisk analyse og finne ut hvordan ens DNA er bygd opp av elementer fra forskjellige folkegrupper. Et annet eksempel er prosjektene med å bygge kvantedatamaskiner og kvantekryptering, der man bygger en maskin som er designet rundt de eksotiske egenskapene man har etablert i teoretisk fysikk. Og disse maskinene virker. Dagens vitenskapelige kunnskap er derfor ikke en ren abstrakt teori om verden. Den er en forutsetning for at alle våre konkrete teknologiske produkter faktisk virker. Å hevde at denne ekstremt vellykkede måten å forstå verden på ikke er korrekt er i seg selv en krevende akademisk posisjon. Å gjøre dette kun basert på, som Chalmers, en intuisjon om at det er mulig å forestille seg kognisjon uten indre opplevelser, en funksjon uten fornemmelser, kan være ekstra krevende.

Fordelen til Chalmers er at han har virkelig satt fokus på det som er et grunnleggende dilemma, kanskje til og med en anomali, i våre teorier om verden¹. De fleste vanskelige problemer blir over tid løst, eller i det minste endret. Vår forståelse av relasjonen mellom det

¹ Strengt tatt peker Chalmers på en grunnleggende utfordring for våre teorier om verden som baserer seg på en gitt måte å forstå disse teoriene, en gitt forståelse av hva som er mulig å forklare i basert på en materialistisk forklaringsmodell. Jeg vil utdype hvilke elementer som inngår i disse forutsetningene i det videre. Jeg mener imidlertid at Chalmers forståelse av materialisme og muligheten for reduktive forklaringer er en utbredt oppfatning som vi finner igjen hos mange tenkere både innen academia og i populære fremstillinger av naturvitenskapelige teorier. Det er derfor nyttig å avdekke eventuelle mangler i en slik fortolkning.

materielle og mentale (den indre opplevelsen) er fortsatt i flux og er under evige diskusjoner, med tilhørende resirkulering av posisjoner. Dette har resultert i den kjente kommentaren fra Stuart Sutherland som han skrev i *The Macmillan Dictionary of Psychology* at "Consciousness is a fascinating but elusive phenomenon; it is impossible to specify what it is, what it does, or why it evolved. Nothing worth reading has been written on it". Det ser ut til at vi ikke har kommet så mye lenger siden Descartes.

Likevel stiller jeg meg i rekken med debattanter som mener at alt håp ikke er ute. Hvilke nye perspektiver bringes til torgs denne gangen? Et viktig element i Chalmers argument for «the hard problem» er våre teorier om den fysiske verden og om de materielle bestanddelene denne består av. Materialisme kan vi strengt tatt beskrive som en metafysisk posisjon: alt som eksisterer grunnleggende sett er fysiske utstrakte ting eller felt, fordelt i tid og rom med et sett fysiske egenskaper (masse, ladning, bevegelse mm). Alle andre ting i verden består av, og er bygget opp av, disse grunnleggende bestanddelene.

En materialistisk og reduksjonistisk forklaringsmodell har sterke koblinger mot fysikk og dens teorier om hva som danner det fysiske grunnlaget de naturlige fenomenene vi kan observere rundt oss. Det er gjennom en slik analyse vi møter "the hard problem". Formulert med Chalmers ord:

Consciousness fits uneasily into our conception of the natural world. On the most common conception of nature, the natural world is the physical world. But on the most common conception of consciousness, it is not easy to see how it could be part of the physical world. So it seems that to find a place for consciousness within the natural order, we must either revise our conception of consciousness, or revise our conception of nature (Chalmers 2002, s1)

Chalmers leter etter et svar på disse spørsmålene ved å analysere samspillet mellom det mentale og dets fysiske grunnlag. Min tilnærming vil i stedet være å søke mer mot biologi og det som lever. Liv består riktig nok av materielle komponenter, men er det noe vi vet om subjektive opplevelser, så er det at dette er noe vi finner hos levende organismer. Det kan være fornuftig å begynne med trygg grunn under føttene nå vi skal begynne å utforske et så omfattende tema som bevissthet.

1.1 Teoretisk bakgrunn

I min analyse vil jeg benytte meg i hovedsak av fire tenkere. Jeg vil benytte Peter Godfrey-Smith som den primære veilederen. Han har jobbet lenge med temaet organismer, biologi og relasjonen til bevissthet og subjektive opplevelser og er en god veileder i feltet. Godfrey-Smith analyserer og belyser sentrale tema innen liv, kognisjon og bevissthet basert på et grunnleggende evolusjonært perspektiv - han har føttene godt plantet i et naturalistisk og naturvitenskapelig perspektiv. Et sentralt poeng for Godfrey-Smith er å jobbe med problemer som ikke omhandler bevissthet som sådan, men hvordan vi kan forstå bevissthet for en bestemt organisme i en bestemt kontekst. Godfrey-Smith beskriver slike former for forklaring som "how-possible" forklaringer (2019) om hvordan indre opplevelser og subjektivitet oppstod. Man kan alltid sette frem den generelle innvendingen om hvordan den indre bevisstheten som sådan oppstår, men hvis det stilles et spørsmål om hvorfor et menneske eller et dyr ser ut til å ha den eller den sensoriske eller kognitive egenskapen så finnes det gode evolusjonære forklaringer. Alle som har blitt veiledet i naturens undre av David Attenborough har lært om dyrenes forskjellige og fantastiske egenskaper; om en bjørn som lukter godt, men ser dårlig; om en fisk som sanser elektrokjemiske signaler i jakt etter bytte som er begravd i havbunnen; om visse farger som oppleves som ekstra intense da dette er et signal om noe spiselig eller giftig.

Jeg vil utdype Godfrey-Smiths tilnærming med analyser av Terrence Deacon. Deacon har utviklet en teori, en slags prinsippforklaring, om hvordan liv kan ha oppstått av de fysiske elementene det består av. Det som er interessant med Deacons analyser er at det viser hvordan liv og levende organismer må forstås som noe mer, noe som fungerer på et annet kausalt nivå, enn ren fysikk. Basert på hans teorier ser vi hvordan vi bør forstå livet som grunnleggende teleologisk. Godfrey-Smith deler synet om at alt liv forholder seg til verden på noe vi kan kalle en proto-kognitiv måte, altså har en rudimentær form for målrettethet. Forskjellen i disse teoriene baserer seg på det teoretiske grunnlaget for hvordan denne målrettetheten realiseres.

Den siste tenkeren jeg vil benytte for å underbygge posisjonen jeg vil presentere er Evan Thompson. Thompson argumenterer for at levende organismer ikke bare må forstås som teleologiske systemer, som at de vil søke å opprettholde seg selv, men at de må forstås som *bevisste* teleologiske systemer; at man kan hevde at alle levende organismer har, med Nietzsche, en *vilje til liv*. Den formålsrettetheten vi ser i livet generelt sett er derfor et uttrykk for en rudimentær form for reell kognisjon- at det på en eller annen måte er et bevisst valg helt ned til den enkleste encellede organisme. Jeg deler ikke Thompsons syn på dette, vi har

ingen grunn til, og trenger heller ikke anta, at enhver levende organisme uttrykker denne formen for bevissthet, med en tilhørende indre opplevelse, der det er bevisstheten som er grunnlaget for de teleologiske egenskapene vi må anta finnes i levende organismer - en "bevissthet" som velger hva som er riktig handling. Thompson har imidlertid gode analyser om hvordan vi må forstå liv som noe som skaper og opprettholder seg selv, og han har relevant kritikk av reduktive forklaringsmodeller. Jeg vil argumentere for at vi trenger en del av perspektivene til Thompsons for å redegjøre for livet som et teleologisk fenomen. I sum gir teoriene til disse filosofene gode redskap til å kaste et forklarende lys over "the hard problem".

1.2 Mekanisme eller teleologi

Som en kontrast til disse tenkerne vil jeg på et overordnet nivå henvise til Daniel Dennett. Dennett sine teorier har utviklet seg med tiden, som han understreker selv. Jeg vil derfor primært henvise til hans siste bok, *From Bacteria to Bach and Back*. Man kan spekulere i om en viktig grunn til disse endringen er Deacon sine teorier. Dennett understreker selv at disse teoriene har fått han til å revurdere en del standpunkter i en anmeldelse av Deacons *Incomplete Nature* (Dennett 2013). Jeg vil følgelig understreke likhetene mellom Dennett og Deacon ved jevne mellomrom. Det er imidlertid en fundamental forskjell, og det er om organiseringen av livet lar seg grunnleggende forklare gjennom en mekanistisk eller teleologisk forklaringsmodell. Jeg benytter Dennett som kontrast til de øvrige filosofiene i denne oppgaven, fordi Dennett deler Chalmers syn på muligheten av reduktive forklaringer av liv og funksjonell bevissthet. Kjernen i en reduksjonistisk forklaringsmodell er å anta at levende organismer kan forstås på samme måte som om de er kombinasjon av en biologisk robot og en datamaskin, en Turing-maskin (Dennett 2017).

Videre er det en tett forbindelse for Dennett mellom en reduktiv forklaringsmodell og hans syn på indre bevissthetstilstander; hans mekanistiske forklaringsmodell av "competence" kan forklare at han likestiller "comprehension" og indre bevissthetstilstander. I motsetning til Godfrey-Smith argumenterer Dennett for at det er noe som fundamentalt skiller menneskelig bevissthet, med tilhørende indre opplevelser, fra andre dyr. Vi kan nemlig forholde oss til omgivelsene ikke bare kompetent, men også med forståelse. En slik forståelse er fundert i vår særskilte evne til å forholde oss til verden gjennom språket². Godfrey-Smith

² Dennett utvikler en teori om hvordan vi kan forholde oss reflekterende til verden basert på Richard Dawkins teori om memer. Jeg vil ikke gå mye inn i denne teorien, men kun fokusere på denne koblingen mellom forståelse og bevissthet med tilhørende indre fornemmelser.

argumenterer for at vi har grunn til å anta at indre opplevelser er et sentralt element for alle eller de fleste kognitive funksjonene til levende organismer og ikke noe som kun kan knyttes til mennesket avanserte kognitive ferdigheter og selvbevissthet. Menneskets spesielle kognitive evner, det man i løst forstand kan beskrive som «bevissthet», bør snarere sees på som et spesialtilfelle av en mye mer grunnleggende egenskap om å ha subjektive opplevelser, om å kunne føle noe (Godfrey-Smith 2016c s 90-92). På dette siste punktet skiller Dennett seg også fundamentalt fra Chalmers. Begge deler synet at det er prinsipielt mulig å forklare liv og bevissthet basert på en reduktiv analyse som baserer seg på et grunnleggende materialistisk utgangspunkt. Dennett argumenterer at det ikke gjenstår noe å forklare (1993, 2017), mens Chalmers nettopp da peker på «the hard problem» som en fundamental utfordring (1996, 1997).

En sentral teori for å argumentere for mulighetene av en prinsipiell reduktiv forklaring av levende organismer er Darwins evolusjonsteori. Denne teorien forklarer hvordan blind naturlig seleksjon kan skape ekstremt kompliserte og tilpassede organismer. Evolusjonsteorien tas naturlig nok for gitt i argumentasjonen til Dennett og Chalmers. Jeg har ikke på noen måte tenkt å utfordre evolusjonsteorien. Det er imidlertid et sentralt element som bør løftes frem. For at evolusjonsteorien skal få sin forklarende kraft forutsetts det at det eksisterer en organisme den kan fungere på; det må være en levende organisme med et sett egenskaper og evne til formering som evolusjonen kan selektere på. Dette er derfor en nødvendig forutsetning for en videre forklaring av hvordan forskjellige livsformer med gradvis økende kompleksitet og spesialtilpasninger utviklet seg fra livets spede begynnelse.

Hva kreves for å opprettholde en levende organisme? Når vi studerer selv de enkleste levende organismer finner vi fenomener vi ellers ikke ser i naturen. Levende organismer må nemlig opprettholde seg selv i møte med en omgivelse som har en tendens til å bryte ned livet. Dette gjelder ikke bare for en mus som ønsker å unngå å bli spist av en rev; det er gjennomgående for alle levende organismer. En bakterie trenger energi for å vokse og leve, den trenger en bestemt type materiale for å opprettholdes. Hvis det er fravær av riktig type energi eller en bestemt type materie, så vil organismen helt enkelt dø. Dette er grunnleggende forskjellig fra andre materielle ting som ikke har en slik aktiv relasjon med sine omgivelser. Spørsmålet er derfor hvordan er en slik opprettholdelse av livet mulig.

Ett mulig svar er å anta at det er bygget inn strukturer i det materielle grunnlaget til levende organismer som prinsipielt kan reduseres til materielle strukturer og prosesser, slik at levende vesener i bunn og grunn er “living robots” (Dennett 2017, s. 85). Dette kan vi forstå som komplekse objekter. Alternativt må vi anta noe mer, nemlig at det er et fundamentalt

skille mellom objekter og levende vesener, at fyllestgjørende forklaring av *subjekter* krever noe mer enn en reduktiv analyse av de materielle bestanddelene med tilhørende fysiske og kjemiske kausalmekanismer.

Det første alternativet søker å forklare organismer som en Turing-robot, som kan utføre komputasjoner på informasjon, gitt som signaler fra omgivelsene slik at nødvendige adferd for å overleve realiseres. Slike komputasjoner er basert på evolusjonært betingede strukturer som muliggjør disse for å sikre organismens overlevelse. Er en slik forklaringsmodell mulig, finnes det en prinsipiell reduksjonistisk forklaringsmodell for livet, og som en forlengelse av det, bevissthetens funksjonelle egenskaper. Vi ville da stå igjen med Chalmers "hard problem" som det eneste fenomenet vi ikke kan forklare.

Jeg vil forsøke å begrunne at det er en del grunnleggende utfordringer ved en slik reduksjonistisk forklaringsmodell. Det kan pekes på forutsetninger og premisser som bare flytter på problemet som skal forklares, uten å forklare fenomenet fullt ut - nettopp at livet søker å opprettholde seg selv. Hvis det ikke er mulig å forklare livets prosesser gjennom en Turing-robot analogi i kombinasjonen med evolusjonsteorien, så må vi finne en annen forklaring av fenomenet. Et alternativ er å anerkjenne at levende organismer kun kan forklares som grunnleggende sett teleologiske; det er et fundamentalt skille mellom objekter som bare blir påvirket av sine omgivelser og subjekter som tilpasser seg og benytter seg av den.

Jeg vil argumentere for at en nærmere analyse av teorier som forklarer levende organismers relasjoner til omgivelsene basert på en «computational» modell, vil *enten* inkludere en nødvendig antakelse som forutsetter noe om en levende organismes egenskaper, *eller* inkluderer man en antakelse om et ubegrunnet kausalt nivå som ikke kan reduseres til rene fysiske årsaks-virknings relasjoner. I begge tilfeller er vi i en situasjon som ikke gir en reduktiv forklaring på fenomenet.

Hva er det primære grunnlaget når vi skal forstå levende organismer? Hva danner grunnlaget for det Dennett beskriver som "competence" (2017)? Dette er grunnleggende spørsmål. Dennett argumenterer som nevnt for at komputasjon (algoritmer, funksjoner) er det grunnleggende for å forstå liv og bevissthet, slik det er realisert gjennom mekanistiske prosedyrer i levende organismer:

More importantly, he [Turing], he showed that if their [computers] instructions included conditional branchingthen these machines could pursue indefinitely complex paths

determined by the instructions, which gave them remarkable competence: they could do *anything* computational....

What Darwin and Turing did was envisage the most extreme version of this point [competence without comprehension]: *all* the brilliance and comprehension in the world arises ultimately out of uncomprehending competences compounded over time into ever more competent - and *hence* comprehending - systems. (Dennett s 55-56)

Dette er et syn som er grunnleggende mekanistisk, som argumenterer for at levende organismer har de egenskapene de har: “ ..thanks to the well-designed machinery they carry with them, thanks to their genes” (Dennett 2017, s 69).

Gallistel (2011) har en omfattende redegjørelse av hvordan vi kan forstå hjernen som en Turing maskin; som “computational”. Jeg vil analysere hans redegjørelser av “the computational mind” nærmere for å peke på de utfordringene jeg mener finnes ved å hevde at en slik forklaringsmodell gir en prinsipiell reduksjonistisk analyse av en organismes evne til å forholde seg til sine omgivelser i kapittel 2.3.1. Jeg vil også se nærmere på den spesielle rollen *gener* ofte tillegges ved reduktive forklaringer av levende organismer.

Jeg vil presentere et syn som anerkjenner disse utfordringene, at vi må anta at livet bør forstås som et grunnleggende teleologisk fenomen. *Det er teleologi som er det grunnleggende elementet, og komputasjon, og dermed også (proto-)kognisjon, som er avledet av teleologi.* Man ikke kan redegjøre for teleologiske hendelser kun gjennom mekanistiske forklaringer, en påstand jeg vil utdype i oppgaven. Følgelig vil jeg argumentere for at det er visse kausale strukturer i verden som er fundamentalt teleologiske. Disse teleologiske strukturene er det som kjennetegner levende organismer. Og det er nettopp i dette forholdet vi vil finne en kobling mot Chalmers “hard problem”.

1.3 Reduksjonisme eller struktur

Som nevnt ovenfor mener Chalmers at materialisme er ufullstendig, og derfor ikke en komplett teori om verden. Materialisme er imidlertid bare problematisk når det gjelder å forstå levende organismer hvis man kombinerer det med en reduksjonistisk forklaring som *kun* henviser til de fysiske prosessene et levende system består av. Dette er den prinsipielle tilnærmingen til Dennett siden han mener dette er en konsekvens av summen av vår vitenskapelige kunnskap, det han kaller “the scientific image”. Kontrasten til dette er “the manifest image”, våre generelle oppfatninger om verden som antas uten en nærmere

analyse begrunnelse eller empirisk belegg; et syn Dennett er dypt skeptisk til (2017, s 61-63). Grunnen til at Chalmers argumenterer mot Dennett er som nevnt ikke fordi han mener at det ikke er mulig å gi en forklaring på bevissthetens ytre kjennetegn, dens funksjoner, basert på materialistiske forklaringer; det er fordi vi da står igjen med bevissthetens «hard problem» (Chalmers 1997)

Ved å fokusere på disse materielle elementene ser man ikke det som er noe *annet*, noe som følger materien, nemlig prosessene som sørger for at liv dannes, opprettholdes og videreføres. Man må være forsiktig med denne påstanden. En sentral, kanskje den viktigste, delen av biologien bygger nettopp på en reduksjonistisk analyse av biologiske fenomener, slik at deres materielle oppbygning og sammensetning klargjøres og sentrale biokjemiske prosesser kan forklares. Dette er grunnlaget for moderne medisin, og er eksempelvis en nødvendig forutsetning i den pågående jakten på en vaksine for koronaviruset (SARS-CoV-2).

En dyp forståelse av biokjemiske prosesser er sentralt og viktig. Jeg vil følgelig ikke argumentere mot en slik redegjørelse, men snarere fokusere på hvordan vi kan forstå liv og levende vesener som sådan. Dette vil jeg gjøre ved å analysere hva som skiller levende organismer fra døde ting, og begrepsapparat som benyttes for å redegjøre for organismer. Gjennom denne analysen ser vi at livet best kan forstås, ikke ved sin materielle oppbygning, men basert på de prosessene som sikrer at en organisme har en gitt oppbygning. Liv, forstått som en generell egenskap, er grunnleggende sett *strukturelt* og ikke materielt. Det kan høres ut som det lefles med gamle teorier som *élan vital*, men jeg vil vise hvordan denne påstanden skal forstås. Det er naturlig nok en variant av emergens, men med en vri. Ved bare å fokusere på det materielle mister man av syne det strukturelle. Det er på denne måten reduktive forklaringsmodeller går glipp av et sentralt bidrag for å forklare levende fenomener. Det er altså ikke materialisme som sådan jeg vil argumentere mot, men materialisme kombinert med reduksjonisme. Konkret knytter dette seg til hvordan Chalmers beskriver materialisme kombinert med reduktive forklaringer for å etablere skillet mellom spørsmål om bevissthet som er «easy» og «hard».

Jeg nevnte over at det er forutsetninger som ikke kan begrunnes basert på en reduktiv analyse av levende organismer og, som en forlengelse, bevissthet. Det ene var en forutsetning om organismen selv. Det andre var å forutsette en ytterligere kausal dimensjon enn rene fysiske virkninger. Dette kan føre til en feilslutning ved at vi analyserer organismer og dens omgivelser basert på en fullt utviklet teori om den verden den må forholde seg til (det vi kan kalle et objektivt perspektiv). Vi ser organismen og dens omgivelser basert på

vårt rasjonelle begrepsapparat og fortolker det gjennom det. Basert på et slikt perspektiv kan det antas at levende organismer kan utfolde seg i verden basert på vår evne til å forstå de oppgaver organismen må løse og de utfordringer den må forholde seg til - vi gir organismen en fortolkende krykke det ikke har. For å forhindre denne feilslutningen må vi i stedet forstå og forklare levende organismer fra et grunnleggende subjektivt perspektiv³. En organisme må skape seg selv, utvikle seg selv, reparere og opprettholde seg selv, reproducere seg selv og fortolke omgivelsene selv. Under en slik antakelse er levende organismer essensielt selv-tilstrekkelig og aktivt skapende. Dette står i en klar kontrast til materielle objekter, som er essensielt ytrestyrt og passiv⁴. En videre begrunnelse og forklaring av disse påstandene vil være et sentralt tema i denne oppgaven.

Det siste sentrale temaet er hvordan vi skal forstå dette skillet mellom livet som et subjekt og materie som et objekt og hvordan dette relaterer seg til de subjektive opplevelsene Chalmers ønsker en forklaring på - nettopp "why and how does conscious experience arise" (2020, s 1). Ved å fokusere på at levende organismer må sikre sin overlevelse ved å forholde seg aktivt til sine omgivelser, i stedet for å analysere de materielle prosessene og strukturene i sentralnervesystemet for å forstå de funksjonelle kapabilitetene til en organisme, peker på to viktige forhold.

For det første peker dette på at det er kun gjennom å se på den *konkrete* organismen i dens *konkrete* situasjon i forhold til dens omgivelser og hvilke *konkrete* utfordringer organismen møter at det gir mening å svare på spørsmålet om hvorfor den eller den bevisste opplevelsen oppstår. Dette gir ikke et svar på Chalmers generelle poeng, men det gir faktisk et konkret svar på Chalmers spørsmål. Vi skal se at vi, sammen med Godfrey-Smith, kan komme et langt stykke på vei i en slik analyse

For det andre, i forhold til Chalmers generelle poeng, ser vi også et nytt perspektiv. Hvis det antas at de prosessene som er nødvendige for å opprettholde og videreføre livet må forstås som grunnleggende aktive og målrettede og realisert av organismen selv, og ikke som en

³ Jeg vil understreke at dette skillet mellom subjektive og objektive perspektiver ikke sier noe om en indre opplevelse, men kun om hvordan vi kan forstå relasjonen mellom en levende organisme og dens omgivelser.

⁴ Det å ta utgangspunkt i livets grunnleggende subjektive karakter, slik at det må i praksis etablere og skape sin relasjon med omgivelsene gjennom disse teleologiske prosessene jeg skal redegjøre for, har likhetstrekk med det Kant gjorde i sine erkjennelsesteoretiske utlegninger basert på hans «kopernikanske vending». Jeg vil komme tilbake til denne sammenhengen. Når det er sagt er det ikke nødvendig å forutsette alle elementer i en kantiansk epistemologi, transcendental idealisme, for redegjørelsen som følger. Jeg vil imidlertid peke på visse sammenfallende elementer gjennom denne oppgaven.

hendelse i de materielle elementene organismen består av, så åpner dette opp for å se en mye tettere sammenheng mellom bevissthetens indre og ytre egenskaper. Mulige konsekvenser av dette perspektivet skal jeg utforske i kapittel 3.5.

Videre i oppgaven vil jeg i kapittel to etablere problemstillingen til Chalmers og analysere utfordringer med reduksjonistiske forklaringsmodeller. Dette vil gi oss en motivasjon til å vurdere alternative forklaringsmodeller. I kapittel 3.1 til 3.3 vil jeg legge frem det alternative perspektivet som ser på livet som grunnleggende teleologisk, slik jeg har skissert det i dette innledende kapittelet. Her vil jeg argumentere for at vi må forstå livet som et grunnleggende teleologisk fenomen. Etter det teleologiske fundamentet er etablert snur jeg oppmerksomheten til spørsmålet om hva bevissthet er og hvordan vi kan forstå dette fenomenet. Jeg vil belyse disse spørsmålene basert på Godfrey-Smiths redegjørelser i kapittel 3.4 og samle trådene i kapittel 3.5 i et sett spekulative analyser som er basert på de forholdene som er etablert frem til da. I kapittel fire vil jeg samle trådene og gjøre noen oppsummerende refleksjoner.

2: Etablering av problemstillingen

Hvorfor kan vi se våre omgivelser? Et konkret svar på dette spørsmålet er at det er nyttig, for eksempel, å se hva som er foran deg. Det kan være en løve som lusker på sletten eller en bil som kommer mot deg; begge er viktig å passe seg for så man ikke blir drept. Videre er det å se viktig for å finne spiselige ting, eksempelvis et rødt eple på en gren. Det å se har en åpenbar evolusjonær nytte. Vi finner derfor at svært mange organismer har utviklet øyne og tilhørende hjerne for å analysere det de ser. På denne måten kan man navigere omgivelsene på en effektiv måte og holde seg i live til man får ført sin genetiske arv videre til neste generasjon.

Det finnes ikke noe bedre generelt svar på det enkle spørsmålet om hvorfor vi kan se våre omgivelser enn forklaringen over. I denne oppgaven vil jeg gå gjennom et sett av resonnement som vil føre oss tilbake til dette utgangspunktet. Den filosofiske relevansen til det innledende spørsmålet har flere sider som jeg i det følgende vil belyse.

Hvorfor holder det ikke at jeg, som en biologisk organisme, har informasjonen om bilen eller løven, slik den er representert i strukturene i mitt sentralnervesystem basert på impulser fra øynene? Hvilken betydning har det at denne informasjonen også er *noe for meg*? Dette temaet knytter seg nettopp til spørsmålene David Chalmers stiller angående “the hard problem of consciousness”, nemlig hvorfor og hvordan vi har en subjektiv opplevelse.

Denne problemstillingen kan man utvide ved å reflektere over *når* i evolusjonens gang denne egenskapen kan ha blitt dannet; når en slik subjektiv bevissthet første gang ble dannet hos en organisme som et resultat av den evolusjonære prosessen. Hvis vi antar at det første livet bestod av komplekse energiske biokjemiske prosesser som sørget for nødvendige strukturer, barrierer og materiale for å opprettholde og reproducere seg selv, så vil det kunne hevdes at en slik biokjemisk struktur ikke har en indre opplevelse- det er bare en kompleks biokjemisk ting. En slik enkel organisme må ha tilpasset seg og respondert på omgivelsene basert på biokjemiske prosesser som i bunn bygger på fysiske og kjemiske lover. Gitt en slik forutsetning, så kan man se for seg en gradvis økning i kompleksitet mellom organismen og omgivelsene og organismen selv, fra encellet til flercellede organismer. Dette er hele tiden basert på det biokjemiske samspillet i organismen og mellom organismene og omgivelsene som sørger for at den best tilpassede versjonen av organismen overlevde og flere avkom enn konkurrerende organismer.

Anta så at det første "lyset" ble tent - hvilket evolusjonært fortrinn ville dette gitt? Hvilken kausal relevans kunne dette ha som ikke allerede ble styrt av de biokjemiske prosessene som er etablert gjennom evolusjonen; hvilke kausale strukturer ble endret som bidro til organismens overlevelse? Hvis den indre opplevelsen fremstår som overflødig for å forklare atferden til organismen, som Chalmers argumenterer for gjennom sitt zombie-argument⁵, så kan det synes som om den indre opplevelsen i beste fall er et evolusjonært biprodukt slik at det fremstår som et epifenomen som følger de materielle kausale prosessene som er etablert gjennom evolusjonen, men som ikke bidrar til organismens overlevelse.

Alternativt kan vi begynne der vi er nå i evolusjonens utvikling, der vi anser det som et faktum at vi har en indre opplevelse som er viktig for å forholde oss til våre omgivelser. Det må derfor være noe ved den subjektive opplevelsen som er kausalt relevant, som bidrar til overlevelsen av organismen. Det kan da være vanskelig å se for seg en organisme som må forholde seg til sine omgivelser uten en eller annen minimum form for indre opplevelse. Det kan da antas at det er gjennom denne indre opplevelsen organismer aktivt forholder seg til omgivelsene. Et slikt resonnement eksemplifiseres i det følgende sitatet:

William Clifford puts the argument thus:

... we cannot suppose that so enormous a jump from one creature to another should have occurred at any point in the process of evolution as the introduction of a fact entirely different and absolutely separate from the physical fact. It is impossible for anybody to point out the particular place in the line of descent where that event can be supposed to have taken place. The only thing that we can come to, if we accept the doctrine of evolution at all, is that even in the very lowest organism, even in the Amoeba which swims about in our own blood, there is something or other, inconceivably simple to us, which is of the same nature with our own consciousness.... (Clifford [1874] 1886: 266) (Goff, et. al "Panpsychism" 2020)

Summen av disse to argumentene fremstår som et klassisk dialektisk problem med en tese og anti-tese der begge konklusjoner synes uunngåelige. Jeg mener ikke at vi finner det riktige svaret gjennom noen av disse argumentene- begge har både elementer av noe riktig og noe galt. Vi må forstå bakgrunnen for argumentene bedre for å finne en ny angrepsvinkel. La oss derfor fortsette analysen.

⁵ Jeg vil beskrive dette argumentet og bakgrunnen for det i mer detalj i kapittel 2.1

Hvis vi ikke kan gi et godt argument for hvordan og hvorfor den subjektive opplevelsen oppstod i den evolusjonære prosessen og hvilken (kausal) betydning dette har for overlevelsen av organismen, så kan det virke som vi enten tvinges til å konkludere:

- at den subjektive opplevelsen i beste fall er et slags epifenomen som følger etter de reelle kausale informasjonsprosessene i hjernen som baseres på evolusjonært utviklede strukturer som hviler på fysikk og kjemi; eller
- at det er en form for indre opplevelse fra livets begynnelse slik at det å kunne ha en subjektiv fornemmelse ikke er et evolusjonært produkt, men heller et premiss for evolusjonen, en ressurs den kan benytte, sammen med organismens materielle sammensetning og tilhørende kausale strukturer, dette for å forbedre sjansene for overlevelse av organismen. Dette kan karakteriseres som en form for panpsykisme, som innebærer at det som legger grunnlaget for den subjektive opplevelsen inngår som en del av naturens grunnleggende byggesteiner, gjerne sett på som en intern egenskap til materien i motsetning til de eksterne egenskapene som er de grunnleggende fysiske elementene i verden⁶.

2.1 Bevissthetens ontologiske grunnlag

Det som danner grunnlaget for disse resonnementene, er det som fremstår som et uoverstigelig skille mellom det *materielle* og det *mentale* (forstått som den indre opplevelsen). Chalmers argumenterer for at dette problemet drives av rådende vitenskapelige teorier som forutsetter og begrunner at det som eksisterer i verden i bunn og grunn er materielle ting og krefter - materialisme eller fysikalisme. Med Chalmers ord:

One has to take materialism seriously in order to take the hard problem seriously as a problem. If one is antecedently a dualist, the hard problem will be unsurprising and not especially worth addressing. The mental and the physical are fundamentally distinct, and that is that..... One might like to know how they interact, but that leads us to other aspects of the mind-problem such as the interaction problem. The problem of explaining the mental in

⁶ Panpsykisme beskrives av Goff et.al som: "Thus, in conjunction with the widely held assumption [...] that fundamental things exist only at the micro-level, panpsychism entails that at least some kinds of micro-level entities have mentality, and that instances of those kinds are found in all things throughout the material universe. So whilst the panpsychist holds that mentality is distributed throughout the natural world—in the sense that all material objects have *parts* with mental properties—she needn't hold that literally everything has a mind, e.g., she needn't hold that a rock has mental properties (just that the rock's fundamental parts do). ("Panpsychism" 2020)

physical terms does not really arise.... Without materialism as a foil, the hard problem did not really get a grip. (Chalmers 2020, side 6)

The “hard problem” tvinger seg derfor frem hvis man tar utgangspunkt i materialisme som den grunnleggende metafysiske tilnærmingen. Siden dette problemet oppstår, trekker Chalmers som konklusjon av “the hard problem” at materialisme ikke kan være en korrekt teori om de grunnleggende elementene i verden (Chalmers 2003). Det må være noe mer; et tillegg til det rent materielle som forklarer bevissthetsfenomener. En naturlig utvidelse av det rent materielle er å postulere, som indikert over, bevissthetsfenomener som en slags grunnleggende egenskap i verden. Dette er en strategi som Galen Strawson forfølger når han argumenterer for panpsykisme som en nødvendig konsekvens av antakelsen om at alt som eksisterer er det fysiske (Strawson i Freeman 2006). Gangen i argumentet er som følger:

1. Det som eksisterer er fysiske materielle ting.
2. Mennesker består i sin helhet kun av disse materielle tingene.
3. Mennesker har som et faktum en subjektiv, indre, opplevelse av verden vi er en del av.
4. En slik subjektiv opplevelse kan ikke oppstå, være en emergent egenskap, av rent materielle egenskaper - egenskaper som utstrekning, masse, elektrisk ladning, plassering i tid og rom og tilsvarende.
5. Det fysiske må derfor inkludere egenskaper som legger grunnlaget for den subjektive opplevelsen - det må være noe “experiential stuff” som inngår i det fysiske som en grunnleggende bestanddel av det som eksisterer i tillegg til rent materielle egenskaper.

Punkt 1 kan man utfordre hvis man argumenterer for en rent idealistisk redegjørelse av verden. I denne oppgaven vil jeg ikke analysere en slik form for metafysisk idealisme nærmere. Antakelsen kan også utfordres ved å argumentere for dualisme. Dualisme forutsetter at det finnes to former for eksistensmodi - den fysiske og det mentale. Disse to elementene må spille sammen på noe vis, men er grunnleggende sett forskjellige. Jeg vil heller ikke analysere dualisme nærmere. Jeg anser begge disse posisjonene som uholdbare. Gitt punkt 1, så fremstår punkt 2 som en analytisk sannhet. Punkt 5 er en konsekvens av premissene 1-4. Det er punkt 3 og 4 som følgelig kan utfordres, gitt premiss 1.

Posisjonen som benekter punkt 3 er eliminativ materialisme, der det argumenteres for at det vi opplever som en subjektiv opplevelse i bunn og grunn er en form for illusjon som bygger på en uholdbar intuitiv psykologi som ikke er konsistent med vår vitenskapelige forståelse av verden (Ramsey 2020). En slik posisjon har store utfordringer med å motivere påstanden om at “that sensations do not actually exist and that they are nothing but brain processes” (Ramsey 2020 siterer Rorty). Det finnes måter å forsvare en slik posisjon på, men jeg velger å fokusere på mindre kontroversielle teorier om mentale tilstander og indre opplevelser.

Alternativt til å bortforklare indre opplevelser, kan man utdype vår forståelse av hva vi mener med «bevissthet» og hvordan vi skal forstå og forklare våre subjektive opplevelse. Følgelig kan punkt 3 analyseres slik at det åpner opp for en mulighet for å vise at premiss 4 er feil. Dette er i hovedsak strategien til Dennett, ved at han fokuserer på hvordan bevissthet bør forstås som sluttproduktet av underliggende kognitive prosesser slik disse prosessene fortolker seg selv og kommer til uttrykk i måter vi tenker om oss selv og i vår forståelse av andre. (1993, 2017). Ved å peke på en det er svært vanskelig å sette fingeren på hva vi akkurat mener med «subjektive opplevelser» og ved vår forståelse av hva innholdet i vår subjektive opplevelse er - som ideen om et enhetlig selv, at vi aldri tar feil av det vi opplever og lignende - kan det redegjøres for at de funksjoner og prosesser som er etablert i evolusjonens gang er tilstrekkelig for å forklare hvordan bevissthet har oppstått og hvordan det skal forstå.

Punkt 4 kan også utfordres. En direkte benektelse av dette punktet er å argumentere for at subjektive opplevelser faktisk oppstår av underliggende biokjemiske prosesser; det komplekse samspillet skaper noe helt nytt som nettopp er det vi kaller subjektive opplevelser, selv om de materielle elementene (altså atomene og molekylene vi er bygget opp av) ikke har noen form for subjektiv opplevelse. Det er nettopp muligheten til å forklare en slik overgang fra det materielle til våre subjektive opplevelser som utfordres av Chalmers “hard problem”. Uten at det er etablert en anerkjent teori om hvordan indre kvalitative fornemmelser kan oppstå fra materiens fysiske egenskaper vil de filosofiske perspektivene stå mot hverandre uten å komme nærmere en løsning. Jeg vil komme tilbake til dette temaet.

En sentral innvending mot punkt 4 bygger på Ockhams barberblad. Denne peker bare på at å godta konsekvensen i punkt 5 gir en mye større omveltning i vår forståelse av verden enn å anta at punkt 4 er korrekt. Gjennom vitenskapshistorien har det vært utallige eksempler på vedtatte sannheter som har måttet vike for den vitenskapelige analysen. Hvis vi gir temaet nok tid vil vi se at dette også vil gjelde for vår forståelse av hvordan bevissthet dannes av

fysiske og biokjemiske prosesser. Dennett bruker dette poenget for alt det er verdt, og ville pekt på at den implisitte antakelsen i punkt 4 - det absolutte skillet mellom det fysiske og det mentale - er det han kaller en "skyhook"; en intellektuell uholdbar snarvei som benyttes der man ikke klarer å forklare et fenomen skikkelig.

Alternativt kan hele resonnementet over vært unngått om man, i stedet for å fokusere på det materielle grunnlaget for bevissthet med en tilhørende indre opplevelser, så ser man heller på dens *funksjoner*. Det sentrale for å forstå bevissthet er ikke mediet den oppstår i (eksempelvis hjernen), men den funksjonelle tilstanden til et system - at bevissthet oppstår som en konsekvens av funksjonen og ikke byggeklossene (se Levine 2018). Chalmers argumenterer for at dette faller inn under bevissthetsforskningens "easy problems" (2003). Funksjonalisme kan redegjøre for det vi gjør, men ikke det vi opplever. Punkt 4 går derfor til kjernen av "the hard problem" og forutsetter i praksis at det er et fundamentalt skille mellom det fysiske og den indre subjektive opplevelsen. Et tilsvarende skille finnes ikke i forhold til det fysiske og våre mentale kapabiliteter, herunder de funksjonene som realiseres av sentralnervesystemet. Dette uttrykkes i Chalmers zombie-argument:

According to this argument, it is conceivable that there be a system that is physically identical to a conscious being, but that lacks at least some of that being's conscious states. Such a system might be a zombie..... These systems will look identical to a normal conscious being from the third-person perspective: in particular, their brain processes will be molecule-for-molecule identical with the original, and their behavior will be indistinguishable. But things will be different from the first-person point of view. What it is like to be an invert or a partial zombie will differ from what it is like to be the original being. And there is nothing it is like to be a zombie. (Chalmers 2002, s. 5)

Når vi har kommet til dette punktet begynner de forskjellige filosofene å grave seg ned i skyttergravene. På den ene siden kan man spørre retorisk: "Skal man virkelig postulere noe så radikalt som panpsykisme bare for å forklare denne siste resten av vår forståelse av det subjektive innholdet i det mentale - som kan være vanskelig å få en god vitenskapelig forståelse av - når vi kan forklare de grunnleggende funksjonene til mennesker og andre organismer fra et objektivt og naturvitenskapelig perspektiv?". Det retoriske svaret er da: "Tillegger du virkelig større vekt til teoretiske vitenskapelige teorier og konsekvensene av disse - teorier som har endret seg over århundrene og som vi har all grunn til å tro vil endre seg igjen - enn den umiddelbare og intuitivt åpenbare opplevelsen du har av verden og deg selv?".

Det kan synes vanskelig å finne en vei ut av dette uføret. Som Chalmers skriver:

I do not claim that idealism is plausible. No position on the mind–body problem is plausible. Materialism: implausible. Dualism: implausible. Idealism: implausible. Neutral monism: implausible. None of the above: implausible. But the probabilities of all of these views get a boost from the fact that one of the views must be true. (Chalmers 2018 s. 28)⁷

For å finne en vei ut av dette må vi lete etter andre argumentative ressurser enn de som er løftet frem så langt. La oss derfor se på problemstillingen på nytt, men med nye argumenter.

2.2 Chalmers lette og vanskelige spørsmål om bevisstheten

Avanserte organismer har et sett objektive kognitive ferdigheter som benyttes for å navigere sine omgivelser. Å forklare denne type ferdigheter er det Chalmers kaller Type I problemer, hans lette problemer. De er lette fordi det ser ut til at vi har metoder og tilnærminger for å forstå dem. I praksis er disse spørsmålene svært krevende å få svar på, men de fremstår som løsbare i prinsippet; de er derfor prinsipielt lette (Chalmers 1998). Chalmers “hard problem” er prinsipielt vanskelige å besvare, det finnes ingen kjent metode eller angrepsmåte å belyse denne problemstillingen på en god måte. Kontrasten mellom disse problemene beskriver Chalmers slik:

Throughout the higher-level sciences, reductive explanation works in just this way. To explain the gene, for instance, we needed to specify the mechanism that stores and transmits hereditary information from one generation to the next. It turns out that DNA performs this function; once we explain how the function is performed, we have explained the gene. To explain life, we ultimately need to explain how a system can reproduce, adapt to its environment, metabolize, and so on. All of these are questions about the performance of functions, and so are well-suited to reductive explanation. The same holds for most problems in cognitive science. To explain learning, we need to explain the way in which a system’s behavioral capacities are modified in light of environmental information, and the way in which new information can be brought to bear in adapting a system’s actions to its environment. If we show how a neural or computational mechanism does the job, we have explained learning. We can say the same for other cognitive phenomena, such as perception, memory, and language. Sometimes the relevant functions need to be characterized quite subtly, but it

⁷ Det finnes en variant til som Chalmers ikke nevner og det er en “dual-aspect” teori, som hevder at det mentale og det fysiske er to sider, en indre og en ytre, av samme fenomen. Som vi skal se i kapittel 3.5 vil jeg argumentere for noe som ligner mest på en slik posisjon.

is clear that insofar as cognitive science explains these phenomena at all, it does so by explaining the performance of functions.

When it comes to conscious experience, this sort of explanation fails. What makes the hard problem hard and almost unique is that it goes beyond problems about the performance of functions. To see this, note that even when we have explained the performance of all the cognitive and behavioral functions in the vicinity of experience—perceptual discrimination, categorization, internal access, verbal report—there may still remain a further unanswered question: *Why is the performance of these functions accompanied by experience?* A simple explanation of the functions leaves this question open (Chalmers 1995 s. 5)

Chalmers nevner mange tema i dette lange sitatet. Jeg inkluderte dette fordi vi skal analyseres disse temaene nærmere for å kaste et nytt lys over sammenhengen mellom det Chalmers betegner som funksjonelle forklaringer og den tilhørende indre opplevelsen. Et eksplisitt premiss i sitatet som jeg vil utfordre i denne oppgaven er antakelsen om at disse funksjonene kan forklares basert på reduktive funksjonelle forklaringer, slik at disse kapabilitetene kan fullt og helt realiseres gjennom rent objektive fysiske prosesser. Jeg vil argumentere for at det som gjør disse funksjonene mulig ikke fullstendig kan forklares basert på fysiske prosesser, at en slik reduktiv forklaring ikke er mulig fordi de stopper opp ved, og forutsetter, en levende organisme som *ikke* kan forklares reduksjonistisk. Hva dette innebærer vil jeg komme tilbake til.

Et sentralt element i denne analysen er komputasjon og, som en forlengelse, kognisjon. Kognisjon er en egenskap levende organismer har, og som er en forutsetning for de avanserte mentale ferdighetene og funksjonene vi finner hos mennesker. Man kan stille seg spørsmålet om når organismer i evolusjonens gang utviklet disse kognitive kapabilitetene som legger grunnlaget for de funksjonelle egenskapene til bevissthet. Det er naturlig å tenke seg at dette henger sammen med utviklingen av et sentralnervesystem, slik at livet før dette ble etablert, i praksis er biokjemiske maskiner styrt av biokjemiske prosesser - rene biologiske roboter.

Jeg vil argumentere mot en slik antakelse i denne oppgaven. Alt liv, selv den enkleste encellede bakterie, må forholde seg aktivt til sine omgivelser for å styre vekst og metabolisme samt for å samhandle med omgivelsene. Sagt med Godfrey-Smiths ord;

all known (metabolically) living systems engage in some cognitive or proto-cognitive processes. The term "proto-cognitive" will be discussed further below, but the activities I see as proto-cognitive include (at least) sensing events and responding to them in a way that helps keep the system alive..... But what we see in basic kinds of metabolic life is something

more specific than computation in that sense, and also something more specific than mere sensitivity to external stimuli. It is the use of sensing and responding, often coordinated with boolean or boole-approximating operations, to maintain the integrity of a system and its activity, seeking and maintaining some states while avoiding others. A collection of *ands* and *if-thens* with no metabolic point to them would be a different sort of thing. When the genome is used to adaptively control the synthesis of metabolically important chemicals by tracking conditions in the external environment, *that* is proto-cognitive in the sense I have in mind. (Godfrey-Smith 2016b. s 8-9)

Chalmers er opptatt av bevissthet og kognisjon. Sitatet til Godfrey-Smith viser imidlertid at før kognisjon så finner vi proto-kognisjon som inngår som et sentralt element i hvordan en organisme forholder seg til sine omgivelser. Vi kan følge Godfrey-Smith i å skille mellom kognitive og proto-kognitive prosesser. For begge prosessene vil vi se en form for målrettethet, en teleologisk årsakssammenheng. Vi kan innføre et skille mellom disse begrepene ved å angi at en kognitiv prosess er noe som inkluderer en *indre bevissthetstilstand som er om noe i verden*, mens en proto-kognitiv prosess vil være en *formålsrettet egenskap som ikke forutsetter en indre bevissthetstilstand*. Proto-kognitive prosesser kan følgelig sees på som en systemegenskap ved organismen. Slike proto-kognitive prosesser kan man enten redegjøre for gjennom en *Turing-robot* analogi, basert på “computational” teorier, eller gjennom å argumentere for at disse prosessene er *et resultat av underliggende teleologiske prosesser*. La oss nå se nærmere på Turing-robot alternativet.

2.3 utfordringer med å se på levende organismer som en Turing-robot

Det er to sentrale elementer som inngår i idéen om en Turing robot. Det første er å redegjøre for kognitive og proto-kognitive egenskaper til levende organismer som om de er en datamaskin, en Turing maskin. Videre må disse komputasjonelle egenskapene inngå i en levende organisme, slik det er forklart gjennom evolusjonsteorien. Jeg skal redegjøre for disse to forholdene etter tur.

2.3.1 Komputasjon og Shannon-informasjon

The Turing machine is a mathematical abstraction rooted in a physical conception. Its importance is twofold. First, it bridges the conceptual gulf between our intuitive conceptions of the physical world and our conception of computation. Intuitively, computation is a quintessentially mental operation, in the Cartesian dualist sense of something that is

intrinsically not physical. Our ability to compute is the sort of thing that led to Descartes' famous assertion, "I think therefore I am." The "I" referred to here is the (supposed) non-physical soul, the seat of thought. In the modern materialist (non-dualist) metaphysics, which is taken more or less for granted by most cognitive scientists and neuroscientists, the material brain is the seat of thought, and its operations are computational in nature. Thus, it is essential to develop a firm physical understanding of computation, how it works physically speaking, how one builds machines that compute. (Gallistel 2011, loc 3546)

Gallistel skriver i rene ord hvordan man skal foreta en reduksjonistisk forklaring av kognisjon basert på en komputasjonell forklaring. Gallistel analyserer ikke eksplisitt proto-kognitivitet og levende organismer som sådan, men vi vil se at gjennom en bedre forståelse av en slik «computational» analyse vil vi også kunne kaste lys over nødvendige antakelser for slike enklere prosesser. Hvor mye kan en komputasjonell modell forklare?

Gangen i å forstå komputasjon som noe fysisk realisert er å identifisere de fysiske symbolene som inneholder relevant informasjon, informasjon som igjen representerer forhold i verden som er relevant for organismen. Ved å identifisere den klare kausale direkte forbindelsen mellom verden, informasjonen i signalene fra denne, sansningen, symbolene og komputasjonen av symbolene med tilhørende tilpasset respons, så har man et representasjonelt system som kan forholde seg aktivt til verden og nettopp oppnå den nødvendige adferden til å tilpasse seg omgivelsene.

Gallistel baserer mye av sin argumentasjon på suksessen til kognisjonsvitenskap og det faktum at vi har laget en Turing maskin som oppfyller disse kriteriene gjennom en datamaskin. Han erkjenner selv at det fortsatt ikke finnes noen konsensus om hva som den fysiske realiseringen av komputasjonen og hva som representerer symbolene i hjernen (Gallistel 2011 loc 898); eksekveringen av prosedyrene og algoritmene som bearbeider symbolene i hjernen. Det spekuleres i at dette kan være spenningsforskjeller (spikes) i nevronene. Gallistel argumenterer for en fortolkning av hjerneaktivitet som er direkte "computational"; aktiviteter i hjernen bør forstås som prosessering av binær informasjon mer eller mindre lik det som skjer i en datamaskin. Som en motsats til en slik forståelse finner vi teorier som antar at komputasjonene som foregår i hjernen er basert på nevralt nettverk, såkalt "connectionism". Dette er tilnærmingen som ligger bak siste års praktiske bruk av lærende systemer og «big data». Disse to teoriene er mulige forklaringer, men vi vet ikke hvordan hjernen faktisk fungerer; det er ingen konsensus.

Videre finnes det ikke en omforent forståelse om hvilke deler av hjernen som inneholder de symbolene som denne komputasjonen benyttes på. Dette er spesielt viktig, fordi i fravær av

slike identifiserte symboler skaper det grunnleggende problemer i å forklare hvordan informasjon fra fortiden, altså minner, kan benyttes for komputasjoner i nåtiden. I fravær av slike identifiserte mekanismer er det ikke mulig å forklare læring på en effektiv måte. Vi vet heller ikke hva som er hjernens fysiske realisering av symboler. Man kunne argumentere for at dette er grunn nok i seg selv til å si at man bør utforske alternative tilnærminger. Utfordringen da er imidlertid appellen til Ockhams barberblad og styrken i appellen til det Dennett kaller «the scientific image»- vi vil snart finne ut av det så det er ingen grunn til å vurdere fundamentalt andre forklaringer.

Det er en mer generell refleksjon som peker på kjernen av problemet med en Turing-robot analogi som en prinsipiell reduktiv forklaring av levende organisme. Det bygger på hvordan organismen får denne relevante informasjonen om omgivelsene som kan danne grunnlaget for den symbolske representasjonen. Først av alt, hva er "informasjon"? Gallistel tar utgangspunkt i Claude Shannon sin definisjon av informasjon:

"Information: The reduction in the entropy of the receiver's probability distribution (over a set of possible messages) effected by a signal. Typically, the signal was received in the indeterminate past, so the reduction it effected must be carried forward in time by a symbol. The information carried by a symbol is the difference between the receiver's uncertainty about the relevant state of the world, given that symbol, and the receiver's uncertainty about that same state of the world, absent that symbol (assuming that the receiver can decode the symbol)." (Gallistel 2011, loc 8204)

Dette er en teknisk beskrivelse som, sammen med teorien om Turing-maskiner, danner grunnlaget for dagens informasjonsvitenskap og -teknologi. Dette har resultert i en praktisk bruk av IKT-teknologi, der smarttelefonene vi har i våre lommer er det ypperste eksempelet på hva som er mulig med et slikt fundament. Det er tre sentrale poeng som må løftes frem.

1. Informasjon er noe som formidler noe fra en avsender (det som danner signalet) til en mottaker- hva som formidles (innholdet i signalet) sier ikke definisjonen noe om
2. Konsekvensen av informasjonsoverføringen er å justere mottakerens forventninger eller oppfattelse av tingenes tilstand (state of the world)
3. Hva er relasjonen mellom signalet (bæreren av informasjonen) og symbolet (lagring av informasjonen) og det informasjonen er om; det som danner grunnlaget for informasjonen som formidles?

Denne tekniske definisjonen ser på formidlingsaspektet ved informasjon, ikke hva som formidles (innholdet). Hvis man tenker på en formidling av informasjon mellom to aktive parter, som mellom to mennesker, så er det klart at nøkkelen til å formidle noe informasjon er at budskapet fra senderen mottas hos mottakeren. Mengden informasjon som kommuniseres er basert på hvilke forventninger mottakeren hadde om verden før kommunikasjonen fant sted. I definisjonen av informasjon over er avsenderen fjernet og erstattet med begrepet "signal". Signalet formidler budskapet som sendes. Et nøkkelspørsmål er hvordan dette signalet fylles med innhold, hvordan dette kan formidle relevant informasjon fra omgivelsene til organismen. Gallistel skriver:

The essential point is that the brain is a receiver of signals that, under the proper conditions, convey to it information about the state of the world. The signals the brain receives are trains of action potentials propagating down sensory axons. (Gallistel 2011, loc902)

Her er det springende punktet hva disse "proper conditions" er. For at signalene i hjernene skal skapes må sanseorganene aktiveres. Et viktig element for å oppnå "proper conditions" er at det etableres et representasjonelt system. For at man skal ha et slikt system må følgende kriterier være oppfylt:

- 1 The mapping from entities in the represented system to their symbols in the representing system is causal (as, for example, when light reflected off an object in the world acts on sensory receptors in an eye causing neural signals that eventuate in a percept of the object...)
- 2 The mapping is structure preserving: The mapping from entities in the represented system to their symbols is such that functions defined on the represented entities are mirrored by functions of the same mathematical form between their corresponding symbols. Structure-preserving mappings are called homomorphisms.
- 3 Symbolic operations (procedures) in the representing systems are (at least sometimes) behaviorally efficacious: they control and direct appropriate behavior within, or with respect to, the represented system. (Gallistel 2011, loc 1716)

Her er vi kommet frem til et sentralt punkt. Hvis man ikke antar hva en gitt «state of the world» er, at det forutsettes hva som er innholdet i budskapet, så følger det av definisjonen av informasjon at en «computational» teori et direkte og nødvendig samspill mellom en organisme og omgivelsene som sikrer at kravet til «structure-preserving» ivaretas⁸. For å sikre denne strukturelle likheten, må det etableres et sett med forutsetninger om at det er forhold i organismen som nettopp har en slik likhet. Det er akkurat slike analyser som gjøres

⁸ Et konkret eksempel på dette er relasjonen mellom en graf i et koordinatsystem og funksjonen som beskriver grafen som en matematisk formel. Innholdsmessig er de forskjellige, men de representerer like strukturer.

av Gallistel; hvordan insekter navigerer sine omgivelser basert på et sett strukturer i hjernen som aktiveres av signalene fra verden og andre insekter. Essensen i en komputasjonell relasjon mellom organismen og omverdenen er at det er etablert nødvendige strukturer i organismen og sentralnervesystemet som disponerer organismen til å utføre en handling. Informasjonstilfanget gjennom signaler benyttes til å justere denne disposisjonen, lik Bayesiske statiske modeller for prediksjon. Det er dette som er kjernen i Shannons begrep om informasjon; mengden informasjon som formidles er endringen i oppfatninger (disposisjoner) før og etter informasjonen er mottatt:

What was essential about a message was not its meaning but rather that it be selected from a set of possible messages. Shannon realized that for a communication system to work efficiently—for it to transmit the maximum amount of information in the minimum amount of time—*both the transmitter and the receiver had to know what* the set of possible messages was and the relative likelihood of the different messages within the set of possible messages. (Gallistel 2011, loc 533, min utheving)

Gjennom evolusjonens gang er det etablert strukturer i hjernen og i organismen som er av en slik karakter at disse speiler strukturelt de faktiske tilstandene i verden. Her er det ikke snakk om abstrakte størrelser som frukt eller ansiktsuttrykk, men helt grunnleggende fysiske fenomener som er konstante elementer i en organismes omverden, som jordens magnetfelt, solens gang over horisonten, månens faser eller, enda mer grunnleggende, rommets tre dimensjoner eller bestemte kjemiske strukturer som eksisterer i stort monn i de relevante omgivelsene til en organisme. Slike grunnleggende, faste fenomener forklares evolusjonært ved at det er etablert strukturer i sentralnervesystemet, men i prinsippet også i andre komputasjonelle prosesser i en organisme (se mer om slike i kapittel 3.3.1), som benyttes for å faktisk fortolke fysiske fenomener som påvirker sanseorganene til organismen. Poenget er at en organismes kapabiliteter og egenskaper er inkludert i en slik analyse for å etablere en nødvendig relasjon mellom organismen og omgivelsene. Teoretisk sett forutsetter man derfor en organisme i en slik modell. Dette er ikke et problem for Gallistels teori, men det holder ikke som en reduktiv forklaring for grunnlaget for hvordan levende organismer kan forholde seg aktivt til sine omgivelser som en ren fysisk prosess, som om det var en datamaskin.

Disse abstrakte resonnementene kan eksemplifiseres. Det som formidles fra omgivelsene kan ikke være informasjon om de fysiske prosessene som påvirker sanseorganene til organismen. De fysiske prosessene er det som skjer, så det må være et tillegg til disse fysiske prosessene. Hvis man antar at jeg ser den nevnte røde eplet, så vil lyset som treffer

frukten som stammer fra solen reflekteres i bestemte bølgelengder før de treffer netthinnen for aktivering av synscellene. Det må være noe med lysets struktur som kan formidle informasjonen om at det nå finnes en moden frukt foran meg - informasjon som for eksempel kan deduseres basert lysbølger i en bestemt sekvens og struktur - slik at hjernen kan foreta en komputasjon basert på denne informasjonen slik den lagres som symboler i hjernen.

Hvis jeg var blind ville jeg ikke kunne tolke denne fysiske hendelsen som kunne vært et signal dersom synet var intakt. Hvis jeg hadde andre typer øyne (som for eksempel kan oppfatte lys i et bredere spektrum), ville jeg da kunne motta mer informasjon da de fysiske egenskapene i mine sanseorganer endrer hvilke fysiske prosesser som kan sende signaler til meg? Kan det hende det finnes svært mange signaler i verden, men jeg har ikke forutsetninger for å motta disse signalene? I så fall er det sannsynlig at det er mye potensiell informasjon som ikke kan formidles til meg⁹.

Hva er informasjonen i å løfte munnvikene mot ørene og vise tennene? For mennesker betyr dette et smil, for sjimpanser betyr det underkastelse. Et mer eller mindre likt fysisk fenomen inneholder forskjellig informasjon. En flaggermus kan forholde seg til omgivelsene like effektivt som en fugl ved å benytte helt andre kommunikasjonskanaler, og derfor forskjellige signaler, for denne navigeringen. Resultatet er den samme informasjonen. Flaggermusen og fuglen vet hvor hindre står, hvor mat befinner seg og hvilke bevegelser den selv må gjøre for å fly i en bestemt retning. Fravær av signaler kan også formidle informasjon (justere mottakerens forventninger), som en flått som venter på et kjemisk signal om at en hjort kommer nær det slik at flåtten kan slippe seg ned på dyret og suge blod for å få næring. Hvis flåtten ikke mottar et signal innen en gitt tid flytter den seg til et annet sted (Sherman 2017).

For at et signal skal mottas, må det både være en relevant «state of the world» og en forutsetning for å motta signalet. Disse to forholdene er derfor gjensidig avhengige, og dette samspillet må begynne med at en organisme utvikler evnen til å plukke opp et signal, slik at relevant informasjon fra omgivelsene kan formidles. Dette er den grunnleggende avhengigheten mellom organismen og verden som forutsettes: For en gitt organisme med en gitt sensorisk kapabilitet kan det antas at signaler fra omgivelsene inneholder informasjon. Dette forutsetter at organismen har nødvendige sensoriske evner til å plukke opp og motta denne informasjonen som er relevant. På denne måten forutsetter en komputasjonell teori en organisme; det er på denne måten relasjonen mellom organismen og omgivelsene må

⁹ I henhold til definisjonen er et signal noe som per definisjon er noe som kan oppfattes, så i henhold til definisjonen så sendes ikke signalet. Poenget med spørsmålene knytter seg til potensielle signaler fra omgivelsene og hvordan disse skal forstås.

skapes. Som jeg skal argumentere for i kapittel 3 er denne gjensidige avhengigheten mellom omgivelsene og organismen med fra livets begynnelse, der dette samspillet utvikles basert på organismens aktive forhold til sine omgivelser. Denne relasjonen må derfor etableres, og den skapes av organismen.

Hvis vi skal følge dette resonnementet til sin ende så må de samme konklusjonene også gjelde for menneskets relasjon til omgivelsene, dersom det ikke antas at mennesket har en særskilt relasjon til sine omgivelser. Følger det da at de teoriene vi har om verden gjennom for eksempel naturvitenskapen ikke er riktige, at molekyler, atomer og naturlige krefter er avhengig av og formet av vårt kognitive apparat? Nettopp disse forholdene finner vi, ved litt refleksjon, igjen i de fleste epistemologiske teorier.

Det grunnleggende spørsmålet her er hvilken status vi gir våre teorier om verden. Jeg har nevnt at det er likheter mellom Kants kopernikanske vending, med den resulterende erkjennelsesteorien han redegjør for i sine Kritikker, og det perspektivet på levende organismer jeg argumenterer for i denne oppgaven. Kants transcendentale idealisme er nettopp konsistent med et syn at vår kunnskap om verden er formet av menneskets erkjennelse, basert på våre teorier og begreper. Thompson, som jeg vil redegjøre for nærmere i kapittel 3.3.3, gjør denne forbindelsen eksplisitt ved å argumentere for et erkjennelsesteoretisk perspektiv han kaller transcendental fenomenologi. En slik erkjennelsesteori kan utfordres på flere områder. Er det nødvendig å anta slike sterke forutsetninger som en konsekvens av denne tette forbindelsen mellom det observerte og den som observerer?

Vi kan gjøre kravet til det epistemologiske rammeverket svakere i forhold å ta høyde for dette samspillet mellom verden og subjektet som fortolker denne. Hva består verden grunnleggende sett av? Er det fysiske materielle objekter, deres deler, molekyler, atomer eller kvarker? Kvarker kan forstås som en idealisering av avanserte matematiske modeller som beskriver samspill mellom felt med forskjellige egenskaper (Carroll 2017). Disse feltene er et uttrykk for både materie og energi, så bak alle fysiske fenomener finner vi dette grunnlaget som ofte er populært eksemplifisert med Einsteins formel $E=mc^2$; dette *noe* som alle fysiske hendelser utfolder seg i. Hvilke av disse forholdene er inkludert i signalene fra verden? Alle, ingen? Vi må ikke anta alle implikasjonene av Kants erkjennelsesteori¹⁰ for å se at det er et sentralt element av fortolkning av empiri i våre teorier om verden. Det er en uendelighet av muligheter av hva slags type informasjon som kan inkluderes i et signal. Det er sannsynligvis dette som skiller mennesker fundamentalt fra andre dyr; evnen til å tilegne

¹⁰ Se Bird (2006) for en oversikt over sentrale tema.

oss svært mye informasjon fra de potensielle signalene om forhold i verden. Selv for oss er vi imidlertid begrenset av våre forutsetninger for å fortolke signalene, forutsetninger som utvikler seg løpende basert på ny kunnskap om verden. Det er konteksten disse signalene fortolkes i, eksempelvis det vitenskapelige begrepsapparatet som benyttes i et gitt eksperiment, eksemplifisert med Kuhns (1970) begrep om paradigmer, som er førende for hvilken informasjon som finnes i signalene fra verden, basert på vår fortolkning av signalene. Den informasjonen vi finner relevant i signalet, er følgelig ikke basert på en iboende egenskap eller struktur i signalene som formidler en bestemt «state of the world».

2.3.2 Evolusjon og semantisk informasjon

La oss vende tilbake til Dennett og hans forsøk på å foreta en grunnleggende reduksjonistisk forklaring av liv og proto-kognisjon. Dennett forutsetter noe mer enn Gallistel i sine analyser av hvordan en organisme forholder seg til sine omgivelser ved at han også tar utgangspunkt i et bestemt innhold i signalene:

The use of the term “information”, which is ubiquitous in cognitive science (and elsewhere), does *not* refer to Shannon information... the concept of information we use in cognitive science is semantic information, that is, information *identified* as being *about* something specific: faces, or places, or glucose, for instance... Some thinkers, perhaps inspired by DNA, think that there *must be* an encoding [sic. symboler] in the nervous system like the DNA code, but I have never seen a pervasive argument for this, and ... there are reasons for skepticism. Semantic information, the concept of information that we must start with, is remarkably independent of encoding (Dennett 2017 s 111-112)

Som argumentert for i forrige kapittel så vil man ikke kunne etablere noe om «states of the world», altså informasjon om noe som Dennett kaller semantisk informasjon, uten å forutsette at en organisme evner å fange opp signaler og fortolke disse. Når Dennett hevder at semantisk informasjon er begrepet som benyttes innen kognisjonsvitenskap så må vi anta at dette settes som et avklart forhold. Dennett løfter frem dette samspillet mellom organismen og omgivelsene ved å hevde at: “Evolution is all about turning ... “noise” into “signal”, and the opportunistic open-endedness of natural selection depends on them” (Dennett 2017, s 33). Det er viktig at et slik samspill forstås på riktig måte. Som nevnt i kapitlet over så er det realistisk å tro at det er organismen som er den aktive parten, det er organismen som må utvikle sanseorganer og fortolke signalene slik at de inneholder informasjon. Hvis det i stedet antas at det er noe i omgivelsene som er styrende for denne

utviklingen, så oppstår denne situasjonen der det kan fremstå som om den semantiske informasjonen (et budskap eller instruks om en tilstand i verden) forårsaker en særskilt respons i organismen.

Ved å sette en slik antakelse beveger man seg imidlertid mot den andre feilslutningen jeg identifiserte i kapittel 2. Ved å anta at det er et bestemt budskap som originerer fra omgivelsene, et budskap som en organisme kan reagere på (en instruksjon) gjennom å utføre en komputasjon, så blir resultatet av en slik antakelse at det vil kunne inntreffe andre virkninger enn rent fysiske. Disse virkningene vil være at det dannes symboler om noe i omgivelsene i organismens hjerne, symboler bestemt av omgivelsens signaler og som har betydning for organismens påfølgende komputasjoner og dermed atferd. Denne adferden vil da være den tilpassede responsen på en gitt hendelse, eksempelvis som en forklaring på hvordan bakterien Godfrey-Smith beskriver i sitatet på side 23-24 kan respondere på hendelser. Resultatet av en slik antakelse blir derfor at det vil kunne oppstå flere effekter enn de rent fysiske, nemlig tilpassede formålsrettede hendelser. Dette blir derfor et nytt kausalt nivå enn et rent fysisk nivå. Vi kan lettere se hva som er uholdbart i en slik antakelse om innholdsmessig (semantisk) informasjon ved å flytte analysen til hvordan man kan forstå gener og idéen om at disse inneholder (semantisk) informasjon:

So we can say that genes contain information about the proteins they make, and also that genes contain information about the whole-organism phenotype. But when we say that, we are saying no more than what we are saying when we say that there is an informational connection between smoke and fire, or between tree rings and a tree's age. The more contentious question becomes whether or not biology needs another, richer concept of information as well. Information in this richer sense is sometimes called "semantic" or "intentional" information.

Why might we think that biology needs to employ a richer concept? One thought is that genes play a special, instructional role in development, telling the embryo how to grow. It is true that genes carry Shannon-information about phenotypes: the genome of a fertilized egg predicts much of the resulting phenotype. In mammals, for example, chromosome structure predicts the sex of the adult animal. But if an informational relationship between gene and phenotype is supposed to involve a distinctive instruction-like mode of causation, then this cannot be information in Shannon's sense. (Godfrey-Smith, Peter and Sterelny, Kim, 2016 «Biological Information»)

Om man ikke er forsiktig med bruken av begrepet informasjon er det lett å gjøre dette hoppet fra det som i sitatet over kalles Shannon-informasjon, som viser til en strengt styrt kausal forbindelse mellom DNA'et og fenotypen til organismen, til det som kalles semantisk eller

intensjonell informasjon. Godfrey-Smith og Sterelny argumentere derfor at DNA'et *ikke* kan antas å inneholde slik semantisk informasjon, i hvert fall ikke hvis man ikke er eksplisitt, og begrunner, en slik antakelse. Hvis det ikke er en slik nødvendig (kausal) forbindelse antas muligheten av en annen form for kausalitet, som er basert på instruksjoner fra informasjonsinnholdet i et symbol (DNA) eller et signal (fra verden). Disse tilleggseffektene er følgelig noe mer enn rene fysiske prosesser. Det kan igjen fremstå som om vi har en reduktiv analyse foran oss, men bak denne ligger det et ubegrunnet premiss. Vi skal se at det er nødvendig å nettopp finne rom til en slik ytterligere kausal dimensjon Denne kausaliteten er av en funksjonell karakter og er basert på de teleologiske prosessene jeg vil redegjøre for i kapittel 3, og ikke på et stipulert informasjonsinnhold i symboler eller signaler. Jeg skal videre se nærmere på relasjonen mellom gener og informasjon i kapittel 3.2.1.

Hvordan forsøker Dennett å løse denne utfordringen uten å forutsette disse ubegrunnede spesielle instruksjonene eller å forutsette det som skal forklares: en levende organisme? Jeg mener at Dennett ikke lykkes med dette. Dennetts bruk av Turing-robot analogien er mer av prinsipiell karakter. Det er mulig å skape kompetente maskiner av blinde algoritmer realisert gjennom et fysisk medium – datamaskiner; har vi mennesker klart dette er det all mulig grunn til å tro at evolusjonen også har funnet en måte å gjøre det på. Selv om Dennett tar utgangspunkt i semantisk informasjon, argumenterer han for at det som muliggjør en organismes aktive relasjon til omgivelsene er styrt av hvordan organismen forholder seg til relevante signaler fra omgivelsene. Som Gallistel, antar Dennett at evolusjonen har etablert disse strukturene for å genere egnede forventninger om atferd, men der atferden må tilpasses og justeres ved kortvarige hendelser - en kvist som knekker - eller ved mer varige endringer – overgang mellom sesonger med tilhørende endring av tilgang til mat basert på endringer i temperatur og lys i omgivelsene. Dennett skriver:

Another virtue of Bayesian models from our vantage point is that an organism can be blessed by natural selection with a high-powered statistical engine These are, if you like, expectation-generating fabrics with a remarkable competence they don't need to understand. The Bayesian fabric is demonstrably not magic. (Dennett 2017, s169-170)

Det som driver disse prediktive maskinene er det Dennett kaller "affordances", ting som betyr noe for organismen i omgivelsene- de tingene i omgivelsene som er viktig å få tak i for å leve og for å unngå å dø. Dette er en klassisk evolusjonær forklaring - evolusjonen har etablert denne "statistiske maskinen" for å skape forventninger som er viktige for organismens overlevelse.

Dette er ikke en holdbar forklaring. Dennett benytter seg her av samme prinsipp som for å forklare livets begynnelse; viktige mekanismer som er en forutsetning for at evolusjonære forklaringsmodeller skal få sin forklaringskraft settet i praksis som et premiss for forklaringen. I dette tilfellet er det Dennetts prinsippforklaring (antakelse) om at livet kan ta i bruk Turing-maskin lignende kalkulerende prosesser for å skape disse disposisjonene. Basert på en slik antakelse argumenterer Dennett for at vi bare skal forstå livets "som om" det er teleologisk og intensjonalt, mens det egentlig er styrt av mekanistiske baserte algoritmer; vi kan innta det Dennett kaller en "intentional stance" (2017, s 37). Det som er en utfordring med slikt argument er, i fravær av en god forklaring og redegjørelse for hvordan dette faktisk skjer basert på mekanistiske prosedyrer og en generell appell til «the scientific image», så kan man stille seg spørsmål om det ikke bare skal forstås «som om» levende organismer er målrettet, men at de *faktisk har* en slik målrettethet. Uten en slik forklaring fremstå det som at Dennett bare antar at dette er mulig, uten egentlig å forklare fenomenet; på samme måte som man tar utgangspunkt i viktige livsprosesser som et premiss for evolusjonære forklaringer.

Dennett begrunner denne antakelsen om at det er en slik statistisk maskin som er etablert av evolusjonen gjennom å anerkjenne at det er en fundamental forskjell mellom hjernens funksjoner og en datamaskin. Her peker Dennett eksplisitt til Deacons teorier (2017 s. 157). Slike disposisjoner kan være fundert i det faktum at nevroner må sees på som levende og derfor, som andre levende organismer, må opprettholde seg selv gjennom å aktivt forholde seg til de lokale forholdene inne i hjernen, i dets miljø:

No adder circuit or flip-flop needs to "worry" about where it's going to get the electric power it needs to execute its duty, and there is no room for "advancement". A neuron in contrast, is always hungry for work; it reaches out exploratory dendritic branches, seeking to network with its neighbours in ways that will be beneficial *to it*. Neurons are thus capable of self-organizing into teams that can take over an important information-handling work, ready and willing to be given new tasks which they master with a modicum of trial-and-error rehearsal. (Dennett 2017 s 163)

Et slikt syn på hjernens funksjoner basere seg på en helt annen tilnærming enn en mer direkte sammenligning mellom hjernen og en datamaskin, noe Gallistel antar i sin redegjørelse. Dette kunne ha dannet grunnlaget for å gi en forklaring for hvordan nødvendige forventninger om omgivelsene er realisert gjennom tidligere utviklede tilpasninger hos enklere celler som nå inngår i, og er tilpasset til, et liv inne i hjernen. Imidlertid kan man fort komme med innvendingen om at man da har flyttet behovet fra å

forklare hvordan hjernen (og dermed organismen) gjør dette, og over til spørsmålet om hvordan dette er mulig for det enkelte, levende og individuelle, nevron. Da er vi tilbake til å måtte forklare hvordan man kan redegjøre for nevroner reduktivt. Igjen skyves det på den grunnleggende utfordringen i stedet for å forklare eller anerkjenne den. Jeg vil skrive mer om disse temaene i neste kapittel.

Det vil alltid være måter å forsvare en posisjon på og man kan alltid sette frem et synspunkt at det argumenteres mot en stråmann ved slike kritiske gjennomganger. Det vil være nyanser som ikke er vurdert grundig nok og mulige svar man ikke har tatt høyde for. Det jeg har forsøkt ved denne kritiske gjennomgangen er å gi en tilstrekkelig motivasjon for å utforske et alternativt perspektiv ved å peke på de utfordringene som finnes i reduktive forklaringsmodeller; forklaringer som er et premiss for Chalmers sondering mellom bevissthetens lette og harde spørsmål. La oss nå endelig se om vi kan kaste et nytt lys på disse spørsmålene.

3: Evolusjon, liv og bevissthet

For å kunne finne en bedre forståelse av Chalmers "hard problem" må vi bevege oss bort fra et grunnleggende fysikalistisk eller materialistisk perspektiv over til et biosentrisk - fra død materie til levende liv.

One link between the two topics [subjective experience and consciousness] is the idea of an organism. An organism is one kind of biological unit, one kind of "individual," and the clear cases of mental phenomena are the mental states of organisms, such as ourselves....Even if you think there is some contingency in that relationship, because of the possibility of minds in AI systems and nonliving robots, the nature of organisms certainly seems relevant to the problem. I think the link is tighter, too. With or without a biological framing, the right approach to the most elusive issues about the mind is by way of the notion of a subject, and subjectivity. If so, part of what we have to understand is a certain kind of unit, a certain kind of constituent of the world: systems that are subjects. Setting aside dualist and panpsychist views, it seems that the special features of subjects are organizational, and this organization will be the product of some sort of evolutionary process. Biological organisms are such important evolutionary products that some connection between the evolution of subjectivity and the evolution of organisms seems inevitable. This does not make it clear how organisms and subjects are related – whether all organisms are subjects, whether subjects are a subset of organisms, whether subjectivity and organismality are matters of degree with some association between them... but biology certainly seems primed to do some of the work. A literature that is especially relevant here is recent discussion of minimal cognition. How does the category of cognition relate to living activity generally? Do plants have cognitive capacities? Is all life cognitive to some degree? (Godfrey-Smith 2016a, s 2)

Dette sitatet dekker mange av de temaene som kan kaste nytt lys på "the hard problem". De sentrale temaene er:

1. Liv som en systemegenskap, drevet av termodynamiske prosesser (omhandlet i kapittelet 3.1: *Fysikk, termodynamikk og liv - livets termodynamisk ubalanse*)
2. Organisering av et subjekt som en selvstendig entitet i verden, med tilhørende subjektivitet (omhandlet i kapittelet 3.2: *Liv, barrierer, reproduksjon og autopoiesis*)
3. Sammenheng mellom (proto-)kognisjon og liv (omhandlet i kapittelet 3.3: *Organismens navigering av omgivelsene*)
4. Evolusjonen av subjektet, subjektive opplevelser og kognisjon (omhandlet i kapittelet 3.4: *Evolusjon fra encellet til flercellede organismer*)

5. Det ontologiske grunnlaget for liv og bevissthet og sammenhengen mellom biologi og fysikk (omhandlet i kapitlet 3.5: *Bevissthet og metafysikk*)

Vi vil gå gjennom disse temaene i tur og orden i de følgende kapitlene. Gjennom denne redegjørelsen vil vi legge grunnlaget for å kunne kaste et nytt lys over “the hard problem”. Det er mange elementer som må spille sammen for at man skal kunne se hvordan analysene til Godfrey-Smith kan kaste lys over sammenhengen mellom bevissthetsfenomener og den underliggende fysiske virkeligheten. Innledningsvis vil jeg derfor sette frem en argumentasjonsrekke som oppsummerer de sentrale elementene som inngår i å belyse spørsmålet “why and how does conscious experience arise” ut fra et biologisk perspektiv. Punktene under er en syntese av argumentasjonen i Godfrey-Smith (2019, 2018, 2016a, 2016b), der jeg har trukket ut de elementene som understøtter min analyse. Tallene i parentes bak hvert ledd i argumentasjonen, henviser til punktene over som igjen henviser til hvilke kapitler jeg vil utdype og begrunne det aktuelle leddet.

1. Det som eksisterer, er den fysiske materielle verden. (1, 2 og 5)
2. I denne fysiske verden finnes det forskjellige materielle strukturer, styrt av rene fysiske og kjemiske prosesser (aggregater, ting) og biologiske prosesser som etablerer og viderefører seg selv gjennom å etablere en klar grense mellom den entiteten prosessen opprettholder og dens omgivelser (organismer). Begge deler er like grunnleggende. (1 og 2)
3. Fysiske strukturer er i termodynamisk balanse med sine omgivelser; biologiske organismer er i termodynamisk ubalanse. (1)
4. For å opprettholde livsprosessene og nødvendige barrierer må den biologiske organismen få en kontinuerlig tilførsel av energi for å opprettholde den termodynamiske ubalansen og sikre tilgang til materie for å vedlikeholde sin materielle struktur (sin individualitet) (1, 2)
5. For å få tilført energi og materie må den biologiske organismen navigere sine omgivelser på en slik måte at den opprettholder nødvendige barrierer og livsprosesser. (2 og 3)
6. For å navigere sine omgivelser må en organisme ha et grunnleggende sett kognitive eller proto-kognitive prosesser som sanser (selvfortolkende) omgivelsene og sikrer (målrettet) opprettholdelse av livsprosessene. (3 og 5)
7. Levende organismer er selvstendige individuelle subjekter som er adskilt fra omgivelsene og som forholder seg kognitivt eller proto-kognitivt til sin omverden (de fysiske objektene). (2, 3 og 5)

8. Gjennom en proto-kognitiv og kognitiv relasjon til sine omgivelser, med mål om å opprettholde sin egen eksistens, realiserer levende organismer selvstendige kausale strukturer som påvirker omgivelsene på en annen måte, og som et tillegg til rene fysiske kausalmekanismer. (1, 2, 3 og 5)
9. Denne selvproduserende, kausalt selvstendige, subjektivt fortolkende prosessen er en grunnleggende del av enhver levende organisme og er en forutsetning for å kunne opprettholde livsprosessene i en omverden den er i termodynamisk ubalanse med. (2, 3 og 5)
10. Over tid utvikles og spesialiseres disse grunnleggende selv-produserende og selvfortolkende kapabilitetene seg basert på en utvikling ved overgangen fra encellet til flercellede livsformer, basert på den evolusjonære historien til organismen og omgivelsen den må forholde seg til (organismen er grunnleggende historisk og kontekstualisert). (4)
11. Bevissthet, forstått som en subjektiv fortolkning og opplevelse av omverdenen basert på indre tilstander med tilhørende kognitive og sensoriske kapabiliteter er en gradvis utvikling fra de grunnleggende livsprosessene og som er basert på behovet en flercellet organisme har for å navigere og tilpasse seg sine omgivelser gjennom koordinering av de delene organismen består av og det evolusjonære mulighetsrommet til organismen. (4, 5)

Dette lange argumentet peker på de sentrale punktene vi må belyse for å svare på Chalmers spørsmål: "why and how does conscious experience arise". Jeg vil nå underbygge påstandene i dette argumentet. Argumentet er ikke en ren logisk deduksjon, da det innføres nye elementer gjennom argumentets gang (det fysiske, subjekter, kognisjon, evolusjon, fortolkning mm), så i det følgende vil jeg underbygge og sannsynliggjøre de forskjellige forutsetningene som ligger til grunn for argumentasjonen. Hensikten med dette oppsummerende argumentet er å løfte frem strukturen i perspektivet som er grunnlaget for å forstå det "hard problem" fra et annet enn et mekanistisk-materialistisk perspektiv.

3.1: Fysikk, termodynamikk og liv - livets termodynamiske ubalanse

Erwin Schrödinger påpekte at levende systemer skiller seg fra andre fysiske fenomener ved at de opprettholder sin struktur og orden over tid (1944). Andre fysiske fenomener vil, som en kontrast, over tid øke sin entropi, gjennom effekten av termodynamikkens andre lov¹¹. Levende organismer bryter ikke denne loven som sådan. De trenger energi for å opprettholde denne strukturen, og ved at de benytter mer energi enn de skaper (forbrenningen er ikke 100% effektiv), følger de også termodynamikkens andre lov. Livet følger imidlertid ikke det vi kan kalle en ytrestyrt statistisk økning i entropi, som er den vanlige tendensen i verden¹², men sørger for å opprettholde og utvikle de etablerte livsprosessene gjennom forbrenning, vekst, formering og reparasjon.

Levende organismer har etablert strukturer som kanaliserer energi slik at denne dynamiske prosessen jobber mot den generelle tendensen til økt entropi i verden. Dette muliggjøres i siste instans av den store mengden energi som til enhver tid er tilgjengelig på grunn av solens energitilførsel til jorden. Å jobbe mot denne naturlige tendensen med økt entropi er et sentralt kjennetegn for livet. Å etablere disse materielle strukturene og dynamiske prosessene, som sørger for at livet opprettholdes og videreføres, er nettopp det som kjennetegner de egne kausale strukturene som er dannet av livet utover naturlige fysiske prosesser. Man kan argumentere for at disse dynamiske prosessene utfører et arbeid som

¹¹ Om termodynamikk (fra Wikipedia: <https://no.wikipedia.org/wiki/Termodynamikk>):

Termodynamikken er aksiomatisk fundert:

- Nullte lov: Dersom system 1 og system 2 begge er i termodynamisk likevekt med system 3, så er også system 1 og system 2 i likevekt med hverandre.
- Første lov: Energi kan verken oppstå eller forsvinne.
- Andre lov: Varme overføres spontant alltid fra et sted med høy temperatur til et sted med lavere temperatur, aldri omvendt.
- Tredje lov: Når temperaturen nærmer seg det absolutte nullpunkt, stanser alle termodynamiske prosesser.

Termodynamikkens andre lov kan også uttrykkes på andre måter. En vanlig formulering er at entropien til et isolert system aldri kan minke. Det betyr at det stadig blir mer uorden, og at et isolert system aldri vil "rydde" seg selv. For eksempel er det aldri observert at en mengde av glassbiter kastes opp i lufta og pusler seg sammen til en tallerken idet skårene lander på bordet. Den motsatte prosessen, at tallerkenen kan bli knust idet den lander, er derimot noe vi tar for en selvfølge. Dette aksiomet har følgende betydning for retningen av naturlige prosesser. Stephen Hawking sier det er termodynamikkens andre lov som fører til at tiden bare går en vei.

Siden aksiomene som termodynamikken bygger på, er svært generelle og ikke avhenger av detaljene i systemet som beskrives, kan teorien som er bygget opp fra dem benyttes i svært varierte problemstillinger. For eksempel benyttes termodynamikk til å analysere svarte hull, prosessene i solen, effektiviteten til forskjellige motorer, kjemiske reaksjoner og evighetsmaskiner.

¹² Standardeksempelet er temperaturen til en gass som jevnes ut og blir lik over tid, eller koppen med kaffe som til slutt får samme temperatur som sine omgivelser

etablerer kausale mekanismer som i høyeste grad påvirker verden direkte, hendelser som ikke ville ha inntruffet hvis livet ikke hadde eksistert:

But subjects also make things happen. Some systems have effects of a certain kind because they are subjects; subjectivity has a causal role. This includes effects on large scales in space and time, including effects on what evolution produces. This is seen especially in the Cambrian. That period saw the evolution of new kinds of bodies, due, at least in part, to the coevolution of bodies, nervous systems, and behavior. New kinds of animal bodies arose because of the subjectrelevant behaviors of other animals – their ability to sense you, track you, and manipulate you, especially as prey. Once actions of those kinds are part of the environment, new evolutionary paths are taken. Such innovations come, in time, to have effects not only on the other organisms they immediately interact with, but on the larger ecology – on how resources are distributed, on what food webs look like, on how carbon and nitrogen are cycled around. (Godfrey-Smith 2016a, s 27-28)

Hvordan etablerer livet så disse strukturene og prosessene som motvirker den generelle tendensen mot entropi og hva skiller disse fra rene fysiske årsaksmekanismer? Dette er en sentral del av Terrence Deacons teori om teleodynamiske systemer (2011), altså systemer med et telos - et mål. For å etablere muligheten for en slik kausal struktur, som ikke vises som en konsekvens av rene fysiske teorier, argumenterer Deacon for at det ikke bare er det rent konkret fysiske (materien) som har en kausal relevans i naturen. Også de rammene som etableres gjennom materielle strukturer og som kanalisere energiutfoldelsen i verden, på en måte som bryter med den generelle tendensen om økt entropi, har kausal påvirkningskraft. Deacon beskriver slike strukturer som begrensninger (constraints) som etableres gjennom samspillet av gradvis økende kompleksitet gjennom naturlige dynamiske prosesser.

Det første nivået av dette samspillet er at et termodynamisk system får tilført energi. Det andre nivået er samspillet mellom slik energitilførsel som resulterer i empirisk kjente dynamiske resultater (det Deacon kaller *morfodynamiske* systemer) som virvler i vann og luft, dannelse av krystaller, autokatalyse og lignende. Til slutt er det et nivå der disse morfodynamiske prosessene spiller sammen i et *teleodynamisk* system. Jeg skal utdype Deacons teorier ved å redegjøre for hans prinsippforklaring for det han kaller et "autogen" i kapittel 3.2.2 der disse abstrakte begrepene konkretiseres. Det som er sentralt for å forstå slike strukturer og effekten de har, er å fokusere på de *begrensningene* som er bygget inn i disse - strukturene forhindrer mulig utfall. Det er gjennom disse begrensningene, det Deacon kaller det som er "absent", at nye resultater skapes. Dette fraværet av utfall er det som setter retningen til systemet. Disse hendelsene som er under kontroll av de teleodynamiske prosessene er styrt og regulert av begrensningene som er etablert av det fysiske systemet. Slike hendelser er virkelige og reelle, og effekten av disse er teleodynamisk kausalitet:

The generic, autonomous, and diaphanous character of autogenic systems makes this a functional property, not a material, chemical, or energetic property. An autogen is a precisely identifiable source of causal influence because it generates and preserves dynamical constraints—the basis for thermodynamic work. But, as we've seen, constraint is the exemplification of something that does not occur. So, in this sense, an autogenic system confronts us finally with an unambiguous absential quality. This is the essence of teleodynamic causality. (Deacon 2011 s 310)

Resultatet av disse strukturene finner vi konkret i hvordan livet benytter materielle strukturer som håndterer energiproduksjonen gjennom ATP-ADP prosessen som finnes i alle celler i dag. Dette er en kjemisk prosess som fungerer som selve motoren for hvordan energi lagres, distribueres og frigjøres i cellene (Carroll 2017).

At livet kanalisere energi basert på strukturer som er etablert gjennom evolusjonære prosesser, og således er i termodynamisk ubalanse med sine omgivelser, viser ikke direkte at disse årsaksmekanismene er teleologiske og ikke mekanistiske. Som nevnt er det å få en bedre forståelse av "the hard problem" å anerkjenne at de årsaksmekanismer som er nødvendig for å forklare en levende organisme er grunnleggende teleologiske, målrettede, og ikke mekanistiske, hendelsesstyrt. Dette vil vi se på snart. Det som imidlertid er etablert, er at livet påvirker hvordan de naturlige energistrømmene i verden utfolder seg og som har en reell og omfattende effekt på omgivelsene. Et tydelig eksempel på denne forskjellen er å sammenligne Mars og jorden. Disse planetene er grunnleggende forskjellige. Selv om begge er fysiske systemer som får tilført energioverskudd fra solen (riktignok mer på jorden), så bidrar åpenbart livets prosesser og dets styring av energi til en ytterligere kausal dimensjon enn det rene fysiske prosesser resulterer i. Vi ser effektene gjennom hvor frodig jorda er basert på disse tilleggseffektene, og hvor gold Mars er der det kun er fysiske prosesser som finner sted (eller Venus som får mer energi fra solen enn jorden).

Oppsummert ser vi her den første bevegelsen mot en anti-reduksjonistisk forklaringsmodell. Selv om alt består av fysiske elementer så finnes det reelle kausale konsekvenser som ikke kan forklares kun gjennom fysiske teorier. Som Deacon sier, så sørger livsprosessene for funksjonelle konsekvenser basert på de strukturelle egenskapene til livet og slike funksjonelle virkninger kommer som et tillegg til materielle, kjemiske eller energiske virkninger som inngår i de fysiske elementene livet består av. Dette er helt konsistent med Godfrey-Smith sin påstand om at "subjects makes thing happen". Her har vi da et alternativ til idéen om at denne kausale tilleggsdimensjonen kommer fra informasjonen som er iboende i de signalene som sendes fra organismens omgivelser.

3.2: Liv, barrierer, reproduksjon og autopoiesis

Levende organismer benytter og styrer energistrømmer for å opprettholde seg selv. For at organismer skal gjøre dette, må de være nettopp det - en organisme. Hva kjennetegner en biologisk organisme?

Konkret vet vi at en levende organisme må ha en måte å opprettholde seg selv på gjennom å få næring, enten direkte fra sollyset gjennom fotosyntese (planter), gjennom å absorbere næring direkte fra omgivelsene (sopp) eller ved å spise andre levende organismer (dyr). Videre må livet videreføres. Dette er realisert gjennom RNA- eller DNA-molekylene, som muliggjør en måte å overføre utvikling og vekst fra en generasjon til en annen, enten gjennom celledeling eller gjennom reproduksjon. Dette er konkrete fysiske forhold som kjennetegner livet på jorden i dag. Mer abstrakt kjennetegnes levende organismer på følgende måte:

In a biological framework, one of the first things that can be explained is the existence of entities that are marked off from their surroundings in relevant ways. This is a consequence of basic features of living systems. Living systems have traditionally been seen as self-maintaining systems, and this is indeed one important feature of them. Living systems are out of thermodynamic equilibrium with their surroundings. Traffic exists between such systems and their surroundings, but this to-and-fro must be limited and controlled, so metabolic systems build and maintain boundaries. Metabolic processes continually recreate both a pattern of organization and the system's distinctness from its surroundings. The term "self-producing" is sometimes used for this feature of living systems (Maturana and Varela 1980), and the term is a good one in emphasizing that what living systems do goes well beyond familiar kinds of maintenance and damage control. They are continually synthesizing the molecules that make them up, preserving a pattern of activity through turnover in materials. (Godfrey-Smith 2019, s 10)

Denne strukturen og disse prosessene som beskrives kalles *autopoiesis*. Dette bygger som Godfrey-Smith skriver på en teori utviklet av Maturana og Varela (se også Thompson 2010 for en utførlig redegjørelse av denne teorien). Dette er de nødvendige egenskapene levende organismer må ha for å kunne eksistere. Som nevnt kreves det energi og materie, men det grunnleggende elementet er å sørge for at den levende organismen er adskilt fra omgivelsene - den er en egen entitet som opprettholder seg selv. Denne separasjonen fra omgivelsene må opprettholdes samtidig som organismen trenger å spille sammen med dens omgivelse for å få nødvendige materie og energi til å opprettholde seg selv. Det må følgelig være et kontrollert samspill mellom organismen og omgivelsene, noe som er kjernen i

teorien om autopoietiske prosesser. En prinsippforklaring på et slik samspill beskriver jeg i kapittel 3.2.2. En levende organisme er grunnleggende sett et selvstendig individ, som av nødvendighet er adskilt fra resten av omgivelsene. Videre må dette individet opprettholde seg selv - det kan ikke hvile på noen ytre mekanismer eller elementer som hjelper det til å eksistere - det skaper seg selv; det opprettholder seg selv; det utvikler seg selv og det reproducerer seg selv - også til et nytt individ som sørger for at livsprosessene videreføres.

En arketype på en organisme som kjennetegnes av egenskapene beskrevet over er encellede organismer som bakterier og encellede alger. Virus, for eksempel, innehar ikke alle kriterier for autopoiesis fordi de er avhengig av andre levende organismer for forbrenning og formering (Thompson 2010). Gitt alle disse egenskapene, denne i bokstavelig forstand "selvtilstrekkeligheten", så er det etablert et sett klare kriterier som differensierer et *subjekt* fra et *objekt*- dette er to grunnleggende forskjellige entiteter i verden. Subjektet er like reelt som fysiske objekter, og like grunnleggende, ikke som et fysisk noe, men som en dynamisk struktur realisert gjennom materien, som et strukturelt overbygg som realiserer subjektet (Deacon 2011). Legg merke til at det å hevde at en levende organisme er et subjekt ikke legger noen føringer om spørsmålet for eventuelle indre bevissthetstilstander til subjektet. Betegnelsen "subjekt" benyttes for å understreke det grunnleggende skillet fra et objekt. Som nevnt må et subjekt karakteriseres som aktivt, indrestyrt og målrettet mens et objekt er passivt, ytrestyrt og statistisk determinert.

Dette reelt eksisterende subjektet har også et annet kjennetegn som viser at det må sees på som en grunnleggende ontologisk realitet. Deacon siterer Kant som skriver at en levende organisme er et "organized natural product ...[where]... every part is reciprocally both ends and means" (Deacon 2011, s 288). Denne abstrakte beskrivelsen av en levende organisme peker mot at funksjonen til delene må forstås gjennom bidraget til den selv-produserende helheten, mens delene kunne ikke eksistere uten helheten og dens funksjoner. For eksempel er barrieren rundt en celle ikke til nytte for seg selv, den er til nytte for cellen som helhet og prosessene i denne. Men uten denne barrieren vil det ikke være en celle. I det abstrakte har vi derfor denne reelt eksisterende strukturen) som vi må anerkjenne som en genuin entitet i verden, et *subjekt* hvis eksistens er ivaretatt av de gjensidig avhengige prosessene og strukturene mellom delene og helheten i en celle (Deacon 2011 s 302, se også Thompson (2010) kapittel 6). Som vi har sett i kapittel 3.1 har denne entiteten en egen kausal kraft, gjennom det Deacon kaller den "teleodynamiske kausaliteten" og danner således egne kausale nettverk utover rene fysiske årsaks-virkningsrelasjoner.

Hvorfor understrekes det at en levende organisme, et subjekt, må sees på som en grunnleggende ontologisk størrelse? Man kan jo hevde at i bunn og grunn er også levende vesener en samling av molekyler. Hva er den prinsipielle forskjellen mellom en levende organisme og et fysisk objekt? Forskjellene er som nevnt flere. En naturlig fysisk ting er grunnleggende sett det vi kan kalle et aggregat av atomene og molekylene det består av. I en slik situasjon kan vi si at nettopp disse materielle byggeklossene er mer fundamentale enn aggregatet. En stein på 2x2 meter er fortsatt en stein om vi deler den i to, men da mindre og med en ny stein ved siden av seg. Som nevnt over kan ikke en levende organisme deles - den må opprettholde sin barriere mot omgivelsene. Poenget er at et levende vesen, er noe reelt eksisterende i kraft av strukturer og prosesser realisert gjennom dets materielle komponenter. Det er ikke de materielle komponentene i seg selv som skaper det "subjektive objektet"- det er de begrensende strukturer realisert gjennom de autopoietiske og dynamiske prosessene som beskrevet over, som styrer, gjennom å forhindre alle andre mulige, utfallet til naturlige prosesser i en gitt retning¹³.

Også Godfrey-Smith argumenterer for et grunnleggende skille mellom et objekt og et subjekt og hevder at det er gode grunner til å tilskrive levende organismer en like grunnleggende ontologisk status som fysiske objekter (2016a). Dette gir en videre styrke til antakelsen om at man ikke kan forstå levende organismer reduksjonistisk. Hadde dette vært en riktig tilnærming ville det ikke være noen grunn til å skille mellom levende og døde ting. Så vi trenger ikke å anerkjenne en *élan vital*, men det er en vesentlig forskjell mellom det levende og det døde ved at livet må sees som en strukturell egenskap som kommer i tillegg til, og er ikke reduserbar til, ren fysisk materie og krefter.

På samme måte er det med sammenhengen mellom helhet og deler. For fysiske objekter så kan man forstå sammenhengen mellom delene og helheten ved at delene konstituerer

¹³ Det er derfor ikke en klassisk supervenient egenskap, som ser på samspillet mellom den materielle helheten og de materielle delene, men en dynamisk emergent egenskap. Det kan synes som en liten forskjell, men det understreker poenget om at det er dynamiske strukturer og ikke materiell komposisjon som kjennetegner liv. Superveniens defineres som: "A set of properties *A* supervenes upon another set *B* just in case no two things can differ with respect to *A*-properties without also differing with respect to their *B*-properties. In slogan form, "there cannot be an *A*-difference without a *B*-difference" (McLaughlin og Bennett 2018). Deacon gjør et stort poeng ut av dette: "But in systems where there is non-linear constraint, propagation with increase in scale (as in processes discussed in the next two chapters), and thus double protection of higher-order attractor features from lower-order fluctuations, both physical and analytical decomposition eliminate the source of this constraint, and hence the source of its causal power. Such cases should therefore be paradigm examples of emergent transitions protected states to describe this sort of insulation between levels of physical dynamics." (Deacon 2011 s 203). Essensen av poenget at "supervenience" er tett knyttet til en materialistisk metafysikk og ikke en dynamisk (strukturell), slik det argumenteres for her der rollen til disse begrensende forholdene i disse dynamiske strukturene er sentralt for å forstå hvilke emergente egenskaper som oppstår.

helheten, helheten er en funksjon eller resultat av delene. Dette er grunnen til at man kan forklare slike fenomener reduktivt. Solen er for eksempel en varm gasskule som sender ut kontinuerlig stråling og energi basert på gravitasjonskraften som virker på hvert enkelt molekyl og atom. Som et resultat av det høye trykket forårsakes en kjernefysisk fusjon av hydrogenatomer til helium siden gravitasjonskreftene overstiger grenseverdien for å starte fusjonsprosessen. Resultatet av denne prosessen er solen slik vi kjenner den som et makroskopisk objekt. For en levende organisme er det annerledes. Delene fungerer på en spesiell måte. Hjertet, for eksempel, slår for at helheten (organismen) skal kunne flytte oksygen og næringsstoffer gjennom blodet til de cellene som trenger dette; delen utgjør en funksjon for helheten og denne delen er slik basert på en evolusjonærhistorisk grunn og ikke en materiell årsak. Relasjonen mellom helhet og del er annerledes enn for fysiske objekter, og det som forklarer hvilke funksjoner disse objektene har er også forskjellig- det ene teleologisk og det andre mekanistisk. Jeg skal redegjøre mer for forskjellen mellom disse forklaringsmodellene i kapittel 3.3.3.

Til et slikt poeng kan man innvende at denne observasjonen er grei nok, men det er jo nettopp dette evolusjonen forklarer. Og det er riktig, men poenget er bare at man må analysere en levende organisme ved å se på organismen som sådan og ikke på delene. Man forstår ikke *hvorfor* en konkret organisme er som den er gjennom en reduksjonistisk analyse, det forklarer bare *hvordan* funksjonene realiseres (og å forbedre denne realiseringen danner eksempelvis grunnlaget for medisinfaget). *Hvorfor* en organisme er som den er og hvordan delene henger sammen med helheten forklares av evolusjonen, men dette er ikke en reduksjonistisk analyse av den konkrete organismen, men av dens evolusjonære historie. Selv om det er en evolusjonær forklaring til en organismes utforming så endrer det ikke det faktum at denne konkrete organismen må forstås som en helhet og der delene står i en annen funksjonell relasjon til helheten enn et fysisk objekt. Derfor hevdes det at levende organismer er grunnleggende ontologiske entiteter; subjekter i stedet for objekter. På denne måten kan man ikke redegjøre for den konkrete organismen basert på en reduksjonistisk analyse.

3.2.1 Gener og liv

Man kan innvende mot argumentet over at man nettopp har en reduktiv forklaringsmodell på hvorfor en organisme er som den er. En organisme skapes og utvikles basert på instruksene slik de ligger i dens arvemateriale, i genene og DNA'et. Det kan hevdes at fenotypen er resultat av en biokjemisk prosess der informasjonen om hvordan organismen skal bli er

skrevet inn i DNA'et, og det som er "skrevet inn" i genene er basert på den evolusjonære historien til organismen. Som jeg skisserte i kapittel 2.3.2 mener jeg denne forklaringsmodellen ikke holder vann. For å forstå denne innvendingen bør vi se nærmere på hva den inneholder.

En slik genosentrisk forklaringsmodell argumenterer for at organismens fenotypiske uttrykk er basert på instruksjoner fra dens genotype, det vil si organismen slik den vokser frem basert på instruksjonene i genene, slik de realiseres gjennom aminosyrer og proteiner som en oppskrift for organismens vekst og funksjoner. Som Thompson skriver:

Much of modern genetics has focused on trying to identify genes as units of heredity with lengths of DNA. For many reasons that are too complex to analyze here, the image of the gene as a molecular repository of information with a replicative agenda of its own has come to exert a great deal of influence in science and society at large, despite the efforts of many scientists to correct this fundamentally biased picture. Because this genocentric doctrine is at odds with the view of life and mind presented in this book, it is important to reveal its flaws and to sketch an alternative to it. (Thompson 2010 loc 3138)

Den genosentriske måten å fremstille levende organismer på, som er popularisert av Richard Dawkins i flere bøker og TV-programmer, gir et helt annet bilde av hvordan livets grunnleggende egenskaper og prosesser er enn den jeg argumenterer for i denne oppgaven (se for eksempel Dawkins 1989). Her finner vi igjen denne antakelsen om at informasjon kan spille en særegen kausal rolle som et tillegg til rene fysiske prosessene, at DNA'et inneholder instruksjoner (basert på symboler i arvematerialet) for hvordan organismen skal bli.

Kjernen i den genosentriske versjonen av evolusjonen er at det finnes et sett gener med en viss variasjon som beskriver oppskriften for liv. Denne variasjonen i oppskriften, som er iboende i genene, resulterer i forskjellige egenskaper. Disse egenskapene har større og mindre sjanse for overlevelse - de organismene som overlever, altså organismene med de mest egnede egenskapene for å leve i et gitt miljø, viderefører sine gener og dermed informasjonen iboende i genene til neste generasjon, slik at evolusjonen i praksis selekterer på denne informasjonen primært, og eventuelt andre forhold sekundært. Denne måten å fremstille evolusjon på forutsetter som nevnt at vi ser på DNA'et som symboler som inneholder informasjon med reelle instruksjoner, en oppskrift (er en materialistisk årsak) til fenotypen som en virkning (resultatet av instruksjonene). En slik antakelse om genes

funksjon er problematisk som jeg argumenterte for i kapittel 2.3.3. La oss gå dypere inn i dette temaet.

Jeg henviste til en generell kritikk av dette perspektivet i et sitat fra Godfrey-Smith og Sterelny i kapittel 2.3.2. Godfrey-Smith utdyper på hvilken måte vi kan forstå informasjon som iboende i genene basert på en analyse med hva som kan betraktes som medfødte egenskaper:

I assume, at the outset, that when people talk of a characteristic as genetically coded (or represented, or programmed for).... they do not mean these phrases as mere synonyms for *genetically caused* or *genetically determined*... So the idea of a genetically coded trait is not supposed to line up with the idea of a genetically determined trait.... the aim here is to assess the possibility of using informational or semantic concepts as an *alternative* to a direct appeal to ordinary causal relations, when understanding the innate as the genetic. (Godfrey Smith 2007, s 6-7, utheving i original)

Denne formen for informasjon er dette rikere begrepet om informasjon, semantisk informasjon og ikke Shannon-informasjon:

So information in the Shannon sense is not something that has an in-principle association with genetic as opposed to environmental mechanisms, and the problem of apportioning of "informational influence" between genes and environment collapses (and messily) to the old problem of apportioning causal influence, which we are trying here to get beyond. (ibid s 8)

Dette er ikke å benekte at det er noe spesielt med DNA'ets funksjon, men dette består i å etablere en særskilt struktur for hvordan gjennom å fungere som "as a template in the specification the linear order of amino acids in the formation of protein molecules. The DNA sequence is, in a sense, "read" by a complex of cellular machinery (including mRNA and tRNA molecules, and associated enzymes) in protein synthesis." (ibid s 9). Dette er helt i tråd med den strukturen og den sekvensen som er etablert i de teleologiske prosessene beskrevet i dette kapitlet og som jeg vil eksemplifisere med Deacons prinsippforklaring av autogen (i kapittel 3.2.2). Uten slik kontroll ville ikke DNA'et fungert som en egnet måte å styre utviklingen over generasjoner av organismer, det ville resultere i for stor variasjon i neste generasjon. Utover denne strengt styrte kontrollen av produksjon av aminosyrer og proteinmolekyler er det ikke noe informasjon som har en særskilt kontroll over fenotypen til organismen:

But my point is that once we get past the production of each protein molecule, the role for a special "coding for" relation is over. The theoretically motivated form of semantic or symbolic description of genes is confined to low-level processes within individual cells. Roughly speaking, it is confined to the process of protein synthesis... Once we are asking questions about levels of aggression, about universal grammar, about sexual orientation or basic numerical cognition, we have far outrun the kinds of entities that can be said, on the basis of biological theory, to be coded for. When understanding those sorts of traits, we are back in the land of ordinary causal explanation (ibid s 10)

Gener spiller derfor en grunnleggende kausal rolle i forhold til utviklingen av fenotypen, men dette skjer basert på særskilt kontrollerte fysiske og kjemiske prosesser og ikke basert på et tillegg om at de faktisk må forstås som symboler som inneholder informasjon som har en preferert kausal rolle utover fysiske prosesser. Med Thompsons ord:

Another aspect of genocentrism we need to examine is the myth of the gene as a unit of pure information. One of the central tenets of genocentrism is that genes have a causally privileged status because they transmit information from one generation to the next, whereas other causes of development are merely material and have no informational status. Nowhere is this conception of the gene as a discrete unit of information proclaimed with more fervor than in Dawkins's writings. For Dawkins, a pronouncement like "Life is just bytes and bytes and bytes of digital information" (Dawkins 1995, p. 19) "is not a metaphor, it is the plain truth" (1986, p. 111). This disavowal of metaphor is indefensible. The plain truth is that DNA is not a program for building organisms, as several authors have shown in detail (Thompson 2010 loc 3320)

Betydningen genene har for utviklingen av en organisme er ikke som en instruksjon for hvordan fenotypen skal være. Hvordan fenotypen, det vil si den konkrete organismen utvikles, består av et komplekst samspill, der deler av genene aktiveres basert på andre omstendigheter til organismen, eller i forhold til omgivelsene gjennom epigenetisk påvirkning. Genene selv skapes i et tilsvarende dynamisk samspill og disse er igjen et produkt av de samme prosessene de inngår i. Alene er DNA'et et avansert organisk molekyl. Det er først når dette molekylet inngår i samspillet med disse prosessene som inngår i livets grunnleggende struktur det får en betydning for organismens utvikling. Igjen, ved at DNA'et inngår i dette styrte samspillet, så dannes funksjonelle virkninger som former fenotypen; de dannes ikke fordi DNA'et inneholder iboende informasjon, symboler om hvordan fenotypen skal dannes. Sherman (2017) har en pedagogisk gjennomgang av denne forskjellen mellom informasjonen i DNA'et som en konsekvens av dens rolle og ikke som en forutsetning for fenotypen. Så igjen, med Thompsons ord:

The deepest fault of the metaphor of DNA as program or information-store is that it implies a dualist framework of matter and information, one homologous to the computationalist and functionalist dualism of the mind as informational software and the brain as hardware. In both cases, processes that are intrinsically dynamic (temporally orchestrated), embodied (somatic and organismic), and embedded (necessarily situated in an environment or milieu)—whether of ontogeny, evolution, or cognition—are projected into the reified abstractions of a genetic program in the cell nucleus or a computer program in the brain. In the one case, to describe DNA as “coding for” phenotypic design reifies the coded content into a kind of mythical “pure information.” Such information “can be encoded, recoded, and decoded, without any degradation or change of meaning” (Dawkins 1995a, p. 19). It is thus conceptually and ontologically distinct from its contingent material expression in the cell, organism, or body. In the other case, to describe the brain as a computer in the head whose function is “information processing” is to reify information into something that preexists “out there,” is “picked up” and “processed” by representational systems in the brain, and is independent in principle of the body, which serves merely as its “vehicle.” In both cases we are handed not simply a dualism of matter versus pure information, but a flight into informational space that is in many ways also a flight from materiality and the body. (Thompson 2010 loc 3438)

Det er et alternativt perspektiv enn det Thompson kritiserer i sitatet over jeg argumenterer for. Det finnes ikke en uttømmende reduksjonistisk og mekanistisk forklaring på de fenomenene vi observerer. Det kan fremstå som om det er mulig å gi en reduksjonistisk analyse, men da vil en nærmere analyse vise at det “smugles inn” begreper som fremstår som objektive og rent faktiske, når de i realiteten forutsetter en form for grunnleggende teleologisk egenskap, semantisk informasjon som kan instruere et fysisk system eller en proto-kognitiv prosess. Dette beskriver Deacon som en *homunculus* i våre teorier om liv og bevissthet, teorier som fremstår som fri fra mening (informasjon) og fortolkning (teleologi), men som egentlig er der, skjult, men likevel til stede, i en eller annen teknisk term som ikke er analysert tilstrekkelig (Deacon 2011). Det er nettopp denne antakelsen Thompson utfordrer i forhold til forståelse av DNA’ets betydning for organismen og som han setter fingeren på i analysen over. Slike bevegelige mål har jeg også pekt på i analysene jeg hadde i kapittel 2.3.

Før vi fortsetter redegjørelsen vil det være nyttig å se på Deacons prinsippforklaring av autogener. Den konkretiserer de abstrakte redegjørelsene jeg har hatt av hans teorier. Redegjørelsen under bygger på Deacons originalverk og populariseringen av hans teorier av Jeremy Sherman i boken *Neither Ghost nor Machine: The Emergence and Nature of Selves*.

3.2.2 Autogener som den første primitive livsformen- Deacons prinsippforklaring

Grunnlaget for Deacons teori om teleodynamiske systemer er å se hvordan livsprosessene inngår i de generelle termodynamiske prosessene som finnes som naturlige fysiske prosesser. Vi ser som nevnt resultatet av at det er noe spesielt som foregår: livet går mot den generelle tendensen mot økt entropi. Hvordan er dette mulig? Som nevnt er utgangspunktet at det bygges opp strukturer i kanaliseringen av denne energien som begrenser mulige utfall. For at dette skal skje må det finnes en "motor" for en slik strukturering. Dette stammer i siste instans fra den kontinuerlige tilførselen av energi jorden får fra solen. Energi må ikke komme fra en stjerne, det kan også komme fra hydrotermiske kilder (som mange mener er der liv oppstod) eller andre egnede former. Men uten en stjerne som sikrer jevn og omfattende tilførsel av energi er det ikke sannsynlig at vi ser liv spredt over en hel planet. Det må videre være riktig mengde energi. Er det for lite energi vil ikke de nødvendige strukturene kunne dannes, for det er ingenting som kan stoppe tendensen mot en høy grad av entropi. Det må heller ikke være for mye energi, for da kan ikke strukturene danne tilstrekkelig sterke barrierer for å motstå nedbrytning fra omgivelsene. Det er derfor man snakker om beboelige soner i et solsystem, hvor forholdene er akkurat passe.

Gitt at det tilføres en riktig mengde energi ser vi at det oppstår dynamiske strukturer som kanalisierer energien på bestemte måter. Det som må etableres basert på en slik styring av energi er å sikre at følgende resultater realiseres:

1: Self-Repair: To overcome second law irregularity, selves constrain or aim energy into work to regenerate their regularities. This aim manifests in a self's capacity to acquire energy and resources, which it converts into work to repair or heal, to regenerate regularities that would otherwise fail through second law degeneration.

2: Self-Protection: Among the features that selves regenerate are protective barriers, constraints that protect and segregate the self's regularities against second law degeneration, for example, cell membranes, seed hulls and shells, skin, and exoskeleton. Repair and protection complement each other. Repair regenerates the regularities that protection locks down, protected from second law tendency

3: Self-Reproduction: Selves proliferate selves, dynamics that inherit the three features of self-regeneration. They do so with evolvable variety—Darwin's heredity and variation.

(Sherman 2017, s 135)

Det finnes fysiske prosesser, prosesser som er en funksjon av termodynamikkens andre lov, som nettopp bidrar til disse sentrale funksjonene. Kjente fysiske fenomener som bidrar til dette er:

1: Auto-katalyserende systemer: Dette er systemer der det foregår en kjemisk kjedereaksjon, der resultatet av en kjemisk prosess gir nye katalyserende molekyler som driver den kjemiske prosessen videre. Denne prosessen foregår så lenge det finnes tilstrekkelig materiale, frem til det tidspunktet materialet er brukt opp av prosessen. Slike prosesser finner vi i svært mange livsprosesser, men da i kontrollerte former. Disse er sentrale for vekst og reparasjon av organismen

2: Formdannende prosesser. Et eksempel er en Bénard celler, som er en konveks mangesidet form som dannes ved oppvarming av olje. Formen dannes fordi dette er den mest effektive måten å bli kvitt energien som tilføres gjennom oppvarmingen. Dette er relevant da cellemembraner benytter fettmolekyler for å danne en barriere mellom seg og omgivelsene, basert på en geometrisk form som dannes gjennom dynamiske effekter.

3: Krystallformasjon. Det finnes mange fenomener der det dannes faste og rimelig varige fysiske former basert på kjemiske prosesser. Snøkrystaller er et eksempel, krystaller i bergarter er et annet. Videre innehar organiske polymerer også disse egenskapene, noe som er sentralt for utviklingen av RNA og DNA.

Dette er de morfodynamiske effektene til Deacon som jeg nevnte tidligere. Sherman kaller disse fenomenene *emergente regulariteter*. Alle disse fenomenene inngår i teorier som forklarer dannelsen av liv; Dennett har en tilsvarende redegjørelse (2017, s. 43-52), men da med en annen underliggende forklaringsmodell. Ifølge Deacon og Sherman er ingen av disse tilstrekkelige hvis fenomenene sees på i isolasjon. Det er bare når det blir etablert strukturer som sørger for (begrenser andre utfall) at disse systemene spiller sammen, man oppnår Deacons teleodynamiske system. Sherman skriver: (2017, side 147)

1: Self-repair instead of metabolism: Metabolism or growth is not enough. Autocatalysis is metabolism, a growing population of chemicals, but it is not a self. Instead, our focus here on self-repair emphasizes what function growth serves, the self's maintenance and restoration of its regularities, not just the growth of a population of molecules.

2: Self-protection instead of containment: Chemicals within a container are not a self. Self-protection replaces the container-first requirement by focusing on the core functional aim of containment. Containment is the source of self-protection, the self's capacity to insulate or segregate itself apart from whatever environmental conditions would result in second law irregularity.

3: Self-reproduction instead of information: Information is an ambiguous term and is too often treated as an intrinsic characteristic of polymers that can replicate their aperiodic sequences the way DNA and RNA do. The fundamental aim served by self-reproduction is not the replication of polymers, but the ability to pass on a capacity to self-regenerate to offspring

For at vi skal få det Sharman kaller et "selv", som er en svært primitiv variant av det jeg kaller et *subjekt*, må disse forholdene fungere i samspill. Det foregår på følgende måte.

En sammensetning av kjemiske komponenter (1)¹⁴ innkapsles i en barriere (2). Disse molekylene viderefører en kjemisk prosess til det er tomt for molekyler som viderefører prosessen. Gitt omgivelsene og den generelle tendensen til økt entropi, vil barrieren brytes opp ved faste mellomrom. Dette åpner bokstavelig talt opp for at nye molekyler kan tilføres til systemet. Denne tilførselen medfører at nye molekyler kan drive den kjemiske prosessen på ny. Videre kan noen av de opprinnelige molekylene fra den første autogenet frigjøres og kapsles inn i en ny barriere, da det antas at det eksistere et relativt homogent miljø med et overskudd av disse kjemikalierne som inngår i punktene 1-3 over. Dette kan forklare en primitiv form for formering (3). Over tid samles kjemiske komponenter som har som resultat, ikke bare å drive den kjemiske prosessen videre (1), men som også resulterer i materiale som raskere skaper en ny barriere (2) som et resultat av autokatalysen. På denne måten effektiviseres samhandlingen med omgivelsene og øker veksttakten til det primitive autogenet.

Membranen kan over tid endres til å brytes raskere ned i møte ved visse kjemikalier. I møte med egnede materielle komponenter vil dette øke tilgangen av nytt kjemisk drivstoff til de kjemiske prosessene som pågår, med tilhørende reparasjon av barrieren (2) og opprettholdelse av den auto-katalytiske prosessen (1). Dette er et eksempel på hvordan en organisme kan etablere affinitet med sine omgivelser og således etablere dette grunnlaget for dens evne til å identifisere hva som er relevant informasjon om noe i omgivelsene. En slik tilpasset reaksjon effektiviserer prosessen ytterligere og kan øke antallet autogener av slike gradvis tilpassede egenskaper og dermed sikre at disse mer effektive autogenene øker i antall raskere enn alternative autogener.

Over tid kan det dannes enda mer kompliserte kjemiske strukturer inne i autogenene som *folder* seg sammen i gitt en fysisk form (3); en primitiv form for dagens DNA. Når et slikt komplisert molekyl får tilført nytt materiale gjennom et brudd i barrieren, vil den kjemiske

¹⁴ Tallene i parentes henviser til det aktuelle punktet i de tre listene over med tilhørende egenskaper, for eksempel autokatalytiske egenskaper for punkt 1.

reaksjonen være styrt av den gradvise frigjøringen av dette relativt kompliserte biologiske polynomet (det er sammenkrøllet slik at ikke alle deler er i kontakt med de nye materialene samtidig), som igjen styrer og effektiviserer reparasjonen av barrieren og samt en mer kontrollert katalytisk prosess. Igjen vil barrierer kunne brytes og disse lange primitive arvemateriale vil frigjøres. Gjennom deres sammensetning og materialene i omgivelsene vil nye barrierer kunne dannes, som et kontrollert kjemisk resultat av den gradvise oppbrytningen av dette kompliserte polymeret. Dette øker effektiviseringen av "autogendeling".

Basert på disse prosessene vil disse autogenen over tid øke i antall og spre seg utover i omgivelsene. Over tid vill de sannsynligvis drive bort fra stedet de ble dannet, kanskje bort fra en hydrotermisk kilde med et relativt stabilt miljø til et område med gradvis endrede ytre forutsetninger for å videreføre prosessene. I møte med slike nye omgivelser vil visse polynomer være mer effektive i å benytte materiale som finnes i omgivelsene og å danne en litt annen barriere som er mer effektiv i disse nye omstendighetene. På denne måten begynner muligheten for å videreføre egenskaper mellom generasjoner, som tilpasses til den konkrete situasjonen autogenet befinner seg i. Dette er et eksempel på hvordan grunnlaget for hvordan et DNA kan etablere hva som er relevant for organismens overlevelse i et gitt miljø ved å etablere informasjon i arvematerialet basert på et funksjonelt resultat som øker sjansen for at nettopp dette kompliserte molekylet å videreføres. Igjen, dette kan ikke sees på informasjon som er iboende i dette primitive molekylet og senere i DNA'et, men i stedet at det får sin funksjonelle relevans basert på hvordan sjansen for samhandling med omgivelsene, reparasjon og videreføring av autogenet økes

Dette er bare en grov skisse til hvordan et slikt autogen kan tenkes oppstått. Fremstillingen kombinerer fysiske prosesser som opprettholdes gitt riktig tilførsel av energi og materie og evolusjonær forklaringsmodell. Vi kan se overgangen mellom hvordan tilfeldige fysiske regulære hendelser begynner å samspille og basert på et sett sammentreff så begynner dannelsen av en struktur, autogenet, som gradvis etablerer egenskaper som gir det bedre eller dårligere evne til å respondere på en egnet måte gitt omgivelsene. Poenget er at alle disse egenskapene må spille sammen for at man kan bevege seg fra død materie til levende liv. Og det er ikke i de konkrete prosessene vi finner de teleologiske effektene. Det er i *samspillet* mellom dem, i strukturene; det er i disse et sett egnede hendelser bygges opp og bevares og så bidrar til at dette samspillet styrer mulige utfall mot det som er nyttig for organismen.

Dette samspillet mellom strukturene som er etablert i autogenet, og levende organismer, er nettopp dette som er inneholdt Kant-sitatet over- det er en *gjensidig vekselvirkning* mellom *deler* og *helheten* som danner en organisk enhet. Fokuserer vi på disse rammene, denne kanaliseringen av energi og materie, og ikke den konkrete fysiske prosessen som utspiller seg i det materielle grunnlaget, så ser vi dette fenomenet: *livet søker å opprettholde seg selv*.

Igjen, disse fenomenene som er beskrevet finner man i mange teorier om livets begynnelse, men det er den teoretiske konteksten til Deacon jeg mener gir en virkelig forklaring på hvordan disse fenomenene skal forstås, og ikke som en appell til et dette er noe evolusjonen har fått til “..thanks to the well-designed machinery they carry with them, thanks to their genes” (Dennett 2017, s 69¹⁵). Da finner vi grunnlaget til å sette dette teleologiske elementet som er grunnleggende for å forklare livet som fenomen og som en forklaring på hvordan organismers målrettethet kan oppstå fra tilfeldigheter. Nøkkelen til det er nettopp å ikke forstå disse fenomenene reduksjonistisk, men som strukturer som styrer utfallet, funksjonelle virkninger, basert på de begrensninger i mulige utfall (resultater) som er bygget inn i samspillet mellom dem. *Dette fenomenet forklarer Deacons fraværende (absent) som man ikke kan se direkte, men som er et reelt fenomen i verden, og resulterer i funksjonelle virkninger som representerer et nytt kausalt nivå*. Videre, denne prinsippforklaringen er et nyttig eksempel på hvordan dette samspillet skal forstås og som en eksemplifisering av sentrale elementer av Deacons teori. Det kan tenkes at dette samspillet oppstod på en annen måte, men det illustrerer fortolkningspoenget, selv om den ikke er empirisk nøyaktig.

Vi skal nå gå videre i analysen og se nærmere på organismens relasjon til sine omgivelser. Oppsummert så har jeg argumentert for at det ikke er mulig å forklare proto-kognitive egenskaper som en Turing-robot analogi slik at vi må anta noe mer, nemlig at en organisme har en grunn for sin handling, forstått som en funksjonell redegjørelse som sikrer opprettholdelse av livet uten å forutsette bevissthet som sådan. Den informasjonen som

¹⁵ Dennetts forklaring av livet og livets opprinnelse er svært lik Deacons, bortsett fra nettopp dette grunnleggende, at levende organismers kompetanse (competance) er basert på “machinery” realisert av evolusjonen (Dennett) og ikke teleologiske prosesser som en forutsetning for evolusjonen (Deacon). Redegjørelsen av hvordan vi skal forstå etableringen av semantisk informasjon er også svært lik mellom Dennett og Deacon. Innenfor dette spørsmålet er det derfor i praksis ikke så stor forskjell mellom dem, selv om grunngrunnelsen er fundamental annerledes. Som en konsekvens av denne forskjellen finner vi imidlertid en praktisk forskjell. Dennett er skeptisk til om hvor mange organismer som har en indre opplevelse. Dennett likestiller en indre opplevelser med vår evne til ikke bare å ha kompetanse (competence) men også forståelse (comprehension) basert på språket vårt med tilhørende idéer (memer) som dette muliggjør. Dennett argumenterer for at det er her vi finner dette særskilte menneskelige fenomenet om et rikt indre liv, noe vi ikke trenger å anta hos andre levende organismer. Som jeg vil redegjøre for i kapittel 3.5 vil jeg utforske en alternativ kobling mellom kompetanse og indre opplevelser, en som er inspirert av Godfrey-Smith.

muliggjør denne opprettholdelsen, finnes ikke som signaler i verden før den etableres basert på hva som er av betydning for organismen. Det som gjør dette samspillet mulig er basert på de teleodynamiske prosessene som er eksemplifisert i forklaringen av autogenet. Det er dette som er de grunnleggende teleologiske effektene som finnes i verden, årsaksstrukturer som baseres på de teleodynamiske prosessene beskrevet av Deacon og de autopoietiske prosessene beskrevet av Thompson. Disse fenomenene etableres og opprettholdes av nødvendigheten av at livet må være i termodynamisk ubalanse med sine omgivelser. For er ikke livet i slik ubalanse vil den konkrete organismen dø, livsprosessene vil stoppe, fordi den blir brutt ned av den naturlige tendensen til økt entropi i verden.

3.3: Organismens navigering av omgivelsene

Kan det virkelig være slik at en encellet organisme forholder seg teleologisk til omgivelsene? Bryter ikke en slik idé grunnleggende med det som er god latin innen vitenskapen, også innen biologien? Selv om en celle er forbausende komplisert (det er lett å tenke, helt feilaktig, at en celle er et slags “biologisk molekyl”, en byggekloss for livet), så må det da være slik at den i bunn og grunn er en avansert biokjemisk maskin? Til dette kan man svare at det kommer an på hva man mener med “maskin”. Det er et sentralt poeng for den nødvendige endringen i perspektiv jeg argumenterer for, det å virkelig innse at en celle umulig kan være en maskin i vanlig forstand.

Hvis vi sammenligner en celle med en topp moderne bil, så er begge deler kompliserte materielle strukturer drevet av komplekse prosesser for styring og kontroll av energi og materie. En bil kontrollerer energistrømmene som frigjøres av forbrenningen av drivstoff gjennom svært kompliserte mekanismer (forbrenningskammer, stempler, turboer, kamakslinger, drivverk, girsystem mm). Denne kanaliseringen gjør fremdriften av bilen mulig. Bilen har avanserte kjemiske prosesser for å rense eksosen for nitrogenmonoksid og andre skadelige avgasser. Bilen har også kontrollsystemer for å overvåke den interne tilstanden (temperatur, oljetrykk, drivstoff, mm) og eksterne forhold (retning og utforming av veien, hastigheten til andre biler, sjåførens oppmerksomhet mm). Bilen har en klar barriere med sine omgivelser, som holder den indre tilstanden i både motoren og kupéen tilnærmet lik uansett vær og vind. Det finnes flere likheter, men poenget er etablert. I hovedsak gjør også cellen dette i sitt miljø, men den kan vurderes som mer kompleks og fullendt siden den også kan videreføre seg selv i et nytt eksemplar og reparere seg selv (en slik bil hadde nok solgt godt).

Poenget her er imidlertid ikke at en celle er mer komplisert enn en bil. Poenget er at bilen forutsetter at den er designet av en ytre skaper, av oss mennesker, dens funksjoner og strukturer er avhengig av en slik ytre referanse. Designet og formålet med funksjonene er for å ivareta målsettinger definert av skaperen, og ikke for bilen. For å parafrasere Kant, så er en designet gjenstand grunnleggende sett et middel og ikke et mål i seg selv. En celles funksjoner er nettopp designet, av evolusjonen, for å ivareta cellens formål, nemlig å sikre sin egen eksistens. Det er dette grunnleggende skillet mellom et subjekt og objekt (igjen uavhengig av spørsmålet om bevissthet) som skiller skapte maskiner fra biologiske, ikke maskiner, men organismer.

Det samme poenget kan også gjøres i forhold til datamaskiner. Som vi kan sammenligne en celledes prosesser og strukturer med funksjonene til en bil, kan vi sammenligne en organismes navigering av og relasjon til omgivelsene med en datamaskin. Det jobbes med å utvikle selvkjørende biler som kan navigere og kjøre seg selv uten menneskelig kontroll. Dette gjøres ved å benytte selv-lærende algoritmer som analyserer svært store datasett. Mye av disse dataene analyseres i sanntid på kraftig spesialdesignede AI-optimalisert hardware i bilene. Det er mulig at datasettet blir så stort at de fanger så mange relevante forhold i verden slik at algoritmene vil bli mer effektive og sikrere til å kjøre en bil enn et menneske. Men den grunnleggende forskjellen mellom subjekt og objekt består. For datamaskinen i bilen har ikke prosesseringen av dataene i maskinkode noen betydning - de er ikke noe for den (Deacon 2011).

Resultatene og funksjonene til slike selvkjørende biler er avhengig av et design utført av og mål satt av oss mennesker. Moderne AI og big-data baserte selvkjørende biler er riktignok basert på et indirekte design gjennom maskinlæring og ikke bare et sett eksplisitte regler programmert av mennesker (direkte design). Begge varianter forutsetter en designer som kan sette retning, vurdere resultater og gjøre nødvendige justeringer ved behov.

Funksjonene som designes inn i selvkjørende biler, og i prinsippet alle dataprogrammer, også inkludert de som simulerer kunstig liv, er avhengig av et slikt ytre satt mål som definerer hvilke funksjoner som skal realiseres. Hvis man helt hypotetisk hadde tenkt at disse funksjonene ville ha en subjektivt opplevd side, så kan man forestille seg en fullstendig splittet opplevelse uten noe indre sammenheng fordi funksjonene¹⁶ retter seg mot noe eksternt og mangefasettert og ikke noe internt og integrert (subjektet). Om det overhode er mulig at datamaskiner har forutsetninger for å danne subjektive opplevelser, noe jeg tviler sterkt på¹⁷, ville det derfor være en god grunn til å anta at dette ville være en indre opplevelse som ikke på noen måte kan sammenlignes med en levende organisme¹⁸.

¹⁶ Jeg antar ikke her en funksjonalistisk forklaring på bevissthet og indre opplevelser. Se kapittel 3.5 for en nærmere redegjørelse.

¹⁷ Jeg vil redegjøre denne skepsisen i kapittel 3.5. Den sentrale innvendingen er en antakelse at subjektive opplevelser er basert i noe grunnleggende fysisk som må ha tilstrekkelig kompleksitet, der denne kompleksiteten etableres i samspillet mellom organismen og miljøet. Det fysiske grunnlaget for datamaskiner er noe helt annet, det er en ekstremt hurtig sekvensiell og parallell lineær prosess. Når det er sagt er moderne databrikker svært kompliserte.

¹⁸ Det følger ikke av dette resonnementet at det er umulig å utvikle selvkjørende biler. Det følger heller ikke at det er umulig å utvikle intelligente systemer basert på datamaskiner, hverken kunstig intelligens (AI) eller kunstig generell intelligens (AGI). Det er ikke intelligens som er det sentrale kjennetegnet ved levende organismer - det er den selv-skapende og selv-tilstrekkelige relasjonen den har med sine omgivelser som er det sentrale. Det er fullt mulig at en slik levende organisme bare blir et første skritt mot en økt kompleksitet i verden ved at det (dvs vi mennesker) utvikler et nytt nivå av dynamiske systemer som har en høyere grad av kompleksitet og intelligens (hva nå enn kriteriet for

3.3.1 Organismens proto-kognitive egenskaper

Hvis vi nå går bort fra menneskeskapt ting og vender tilbake til organismen. Hvordan forholder de seg til verden? En ting er at levende organismer ikke er maskiner på samme måte som maskiner laget av mennesker. Å hevde at alle organismer forholder seg proto-kognitivt til verden, selv encellede organismer, kan for mange fremstå som intuitivt lite plausibelt. Jeg har tidligere hevdet at dette er tilfelle, så la oss nå gi empirisk belegg for denne påstanden.

František Baluška og Michael Levin (2016) går gjennom litteraturen om kognitive prosesser i vid forstand som eksisterer gjennom andre biologiske strukturer enn sentralnervesystemet (deres redegjørelse skiller ikke mellom proto-kognitivt og kognitivt, altså teleologiske relasjoner til verden med eller uten en tilhørende indre bevissthet). En innledende kommentar er at det ikke er noe grunnleggende forskjell mellom en nevron og andre celler; et nevronalt nettverk er optimalisert for å styre og kontrollere en komplisert organisme, men er selv utviklet fra andre biologiske funksjoner. Som de skriver:

Importantly, CNS [Central Nervous System] neurons do not embody cognition due to any magical, unique property. In fact, these basic properties are present in biological systems at many complexity scales (from subcellular protein networks to coupled tissues). Might they too underlie some aspects of cognitive-like information processing? Indeed, neurons did not invent their special tricks – they merely optimized them for speed to drive adaptive behavior. These functions, and the molecular mechanisms that implement them – ion channels, electrical synapses (gap junctions), and neurotransmitter molecules are all ancient (Goldsworthy, 1983; Baluška, 2010; Brunet and Arendt, 2016; Moroz and Kohn, 2016). Neural networks evolved from far older signaling pathways that orchestrated development, physiology, and other cellular functions long before the CNS arrived on the evolutionary scene (Buznikov et al., 1996; Levin et al., 2006; Keijzer et al., 2013). Already simple cells of bacteria enjoy sensory systems feeding into cognitive-behavioral circuits and showing many other neural features (Miller and Koshland, 1977; Koshland, 1980; Lyon, 2015). Electrical long-distance signaling and information exchange via spatially propagating waves of potassium is synchronizing bacterial biofilms (Beagle and Lockless, 2015; Nunes-Alves, 2015; Prindle et al., 2015). Integrated bacteria within the biofilm community appear to act as

det er) enn hva livet selv har. Dette vil være analogt med hvordan livet har utviklet seg fra encellet til flercellede strukturer. Så det er ingen *a priori* spådommer om fremtiden, men det er en påminnelse om at dagens AI systemer er noe fundamentalt annerledes enn intelligensen til levende organismer. Og det kan være en nyttig innsikt for en del fremtidsforskere.

some kind of 'microbial brain'. Obviously, the neuronal communication has bacterial origins (Baluška and Mancuso, 2009). (Baluška og Levin 2016)

Artikkelen går i dybden på mange temaer om ikke-nevronal (proto-)kognisjon. Under følger noen hovedpunkter hvor vi finner slike proto-kognitive prosesser realisert gjennom andre biologiske prosesser:

- Cytoplasmaen, det vil si den væsken som er på innsiden av en celle
- Kjemiske nettverk som benyttes for utvikling av embryoet
- Regulatoriske nettverk som styrer utviklingen og aktiveringen av gener
- Bioelektrisk aktivitet mellom celler som ikke er nevroner, inkludert i planter
- Minnefunksjoner som resulterer i modifisert adferd basert på tidligere hendelser i bakterier (en lik hendelse vil altså resultere i forskjellig respons basert på "erfaringen" til bakterien)
- Læring og kognitive prosesser i store encellede organismer som slimsopper¹⁹
- Tilpassede vekststrukturer i benceller og muskelceller basert på tidligere hendelser
- Feedbackmekanismer i vekstprosesser som sikrer korrekt utvikling av organismen

Dette er bare oppsummerende punkter av den inngående gjennomgangen til Baluška og Levin. Til dem som er interessert i flere detaljer anbefales det å lese artikkelen. De oppsummerer slik:

We have avoided here the thorny issues of philosophy of mind that arise from trying to define exactly under what conditions words like "knowledge" are appropriate, in favor of an intentional stance-like pragmatic, engineering approach grounded in cybernetics. The coverage of cognitive terms across biology can expand to the extent that information-centered approaches are shown to be effective in predicting and controlling the behavior of biological systems. The practical implications for biotechnology, unconventional computation, and regenerative medicine are enormous. Equally likely, the lessons we learn from unconventional cognitive systems will deeply impact our most basic understanding of how mind emerges from the brain. (Baluška og Levin 2016)

¹⁹ Se følgende linker for en kjent variant av slike <https://www.nrk.no/urix/viser-frem-mystisk-organisme-med-720-kjonn-1.14748973> og <https://www.dagbladet.no/nyheter/hjernelos-klatt-har-720-kjonn/71721079>

Vi har en berettiget grunn, basert på empiriske analyser, å anta at selv enkle organismer forholder seg proto-kognitivt til sine omgivelser²⁰. I fraværet av muligheten til å kunne forklare denne egenskapen ved levende organismer utelukkende basert på rene fysiske kausale strukturer har vi grunn til å anta at disse prosessene er grunnleggende teleologiske. Det er ikke teleologisk i en menneskelig rasjonell forstand (at den er et uttrykk for en rasjonell beslutningsprosess). I stedet kan vi se på dette som et nødvendig alternativt perspektiv for å forstå livet på en måte som ser livet slik det er, og ikke som ren fysikk. For å supplere Deacons teori og Godfrey-Smiths analyser kan vi vende blikket mot Stuart Kauffman's innledning til *Beyond Mechanism: Putting Life Back Into Biology* (Henning og Scarfe 2015) for å få et bredere perspektiv som gir ytterligere motivasjon til å anerkjenne disse teleologiske prosessene.

3.3.2 Mål og mening

Et første poeng er at det vi ser rundt oss i den biologiske verden ikke kan ha oppstått av rene tilfeldigheter, noe som ville ha vært nødvendig dersom biologi «bare var fysikk». Hva menes med utsagnet «bare var fysikk»? Et slikt argument løftes ofte frem av forfattere av Intelligent Design, som argumenterer for at det står en skaper bak verden, altså Gud. Det pekes på at sjansene for at alt det vi ser rundt oss skal oppstå basert på rene tilfeldigheter er mikroskopiske. Hadde biologiske systemer bare blitt dannet av rene statistiske tilfeldigheter som ble selektert, ville det i beste fall tatt “10 raised to the 39th power times the lifetime of our universe to make all possible proteins of length 200, just *once*” (Kauffman i Henning og Scarfe 2015, s 5).

Å bare appellere til at det er ren tilfeldighet som setter i gang og styrer evolusjonen faller derfor på sin egen urimelighet. Det må være noe annet. Det er nettopp å forklare dette andre som er gjort over. Det er noe i verden som begrenser mulighetene, som styrer utviklingen i en *retning* og ikke kun basert på rene tilfeldigheter. Dette er i tråd med det Deacon argumenterer for med hans begrep om “absent”; nemlig som fravær av mulige utfall. Det som styrer denne retningen er den grunnleggende teleologiske egenskapen til livet, i generell forstand. Livet styrer utviklingen av verden med disse dynamiske prosessene som er beskrevet. Denne styrte utviklingen har et fundamentalt normativt eller evaluerende element; det sikrer at et utfall (overlevelse) går foran et annet (død). Det som bidrar til denne

²⁰ Se også Godfrey-Smith (2016b) for en diskusjon om samme tema og der han argumenterer for tilsvarende proto-kognitive funksjoner som finnes i organismer uten sentralnervesystem, inkludert encellede organismer som bakterier. Dennett argumenterer også for det samme (2017), men da med en annen forklaringsmodell for hvordan dette er mulig.

retningen er det samme fenomenet som også setter retning for alle andre hendelser i verden, termodynamikkens andre lov som sier at entropien i verden, samlet sett, bare øker. Fenomenet liv har tatt kontroll over denne grunnleggende motoren og benytter den til sitt eget formål: å opprettholde seg selv.

Livsprosesser som sådan har derfor etablert strukturer som bidrar til å opprettholde og videreføre den konkret levende organisme. Igjen, dette er ikke et bevisst valg, men en styrt prosess og grunnleggende teleologisk:

Thus, as Kull points out, the set of states of the different molecules outside the cell that can bind to the outside parts of these transmembrane proteins and unleash intracellular signalling and a coordinated cellular response, constitute a *semiotic code* by which the cell navigates its “known” world, “known” - without positing consciousness - via the code and, in general, probably evolved by selecting encoding of the world as “seen” by the organism. Change the molecule species binding the outside of the transmembrane protein, and the world the cell “knows” and evaluates changes. Biosemiosis is real in the universe²¹. (Henning og Scarfe 2015, s 8)

Deacon, Godfrey-Smith og Thompson understreker det samme poenget. Organismer responderer ikke mekanistisk (gjennom årsak-virkning eller logiske beslutningsalgoritmer), de *sanser* sine omgivelser (normativt) og responderer deretter. Responsen *søker å opprettholde* organismen selv (som i bunn og grunn er disse autopoietiske prosessen som er beskrevet), ikke bare en ren mekanisk virkning på en forutgående hendelse (årsak):

So far we have bounded self-producing units engaged in traffic with their environments, where this traffic includes sensing and adaptive response. Some of their capacities can be seen as a simple form of cognition, and as units that *sense* and *act*, they are also subjectlike, or subjects in a first, minimal sense. In the previous section I discussed mostly prokaryotic cells. Richer forms of sensing are seen in eukaryotic protists, along with a more elaborate internal skeleton, enabling manipulation of objects and new lifestyles. (Godfrey-Smith 2016a s 8, min utheving)

²¹ Biosemiose omhandler hvordan liv fortolker signaler, altså forholder seg til informasjon med utgangspunkt i sin egen relasjon til sine omgivelser. En gruppe tenkere har laget en Prolegoma til en teoretisk biologi der man må forholde seg til denne realiteten. Se <http://www.zbi.ee/~kalevi/Saka.pdf> for en relativt utførlig beskrivelse av hva biosemiose innebærer. Hovedpoengene er i tråd med temaene i denne oppgaven. Deacon var med å skrive denne prolegomaen og det kan virke som hans *Incomplete Nature* var et direkte svar på alle utfordringene som ble satt opp i denne akademiske pamfletten.

Basert på dette kan vi ha en berettiget grunn til å hevde at det er to forskjellige hendelsesmodi i verden. Vi har kausalt statistisk bestemte mekanistiske som er beskrevet av fysikken og som omhandler objekter (ting) og vi har den aktive, teleologiske som tilhører det biologiske domene og som omhandler subjekter (individer)²². Begge har kompatible, men utvidende kausale nettverk og begge er grunnleggende ontologiske enheter, objekter forankret i dets materielle komponenter, subjekter forankret i dynamiske strukturer, realisert gjennom objekter. Fysikkens verden er (statistisk) forutsigbar uavhengig av hvilke elementer (atom A eller B av et hydrogenatom) som inngår i en gitt hendelse mens naturens verden er kontekstbasert. En gitt organisme i et gitt miljø må forstås i et gitt samspill, altså avhengig av de konkrete forutsetningene for å forklare utfallet etter hendelsen har inntruffet. Videre fremstår fysikkens verden som prinsipielt ahistorisk, men gjennom termodynamikkens andre lov drives fysiske hendelser i en gitt retning (det ville krevd energi for å endre tilbake til den opprinnelige tilstanden) slik at den i praksis følger et forløp som styres av termodynamikkens andre lov (Carroll 2017). Det er nettopp å etablere strukturer som kanaliserer energien til sine mål og ikke følge den generelle tendensen om økt entropi at livet sikrer et utfall som opprettholder livet til en gitt organisme. Av denne grunn kan vi si at naturens verden er både prinsipielt og praktisk historisk i sin utfoldelse siden utfoldelsen av energi er styrt av organismen for å bidra til dens egen overlevelse.

Å opprettholde et prinsipielt skille mellom levende og død natur kan virke som et skritt tilbake fra Darwins teori om livets utvikling; han viste at det ikke er behov for å anta at det er noe spesielt med livet som et fenomen for å forklare de fantastiske livsformene vi ser i dag. Dette er riktig når det gjelder hvorfor vi har de livsformene vi har. Det kreves ikke noen henvisning til en designer for å redegjøre for hvilke levende organismer som finnes og hvorfor de har de egenskapene de har. Dette er et resultat av organismens evolusjonshistorie. Men for den enkelte levende organisme er det annerledes. Der kan vi ikke redegjøre for livet som fenomen uten å benytte teleologiske begreper for å forklare

²² Et vanlig tema når man skriver om sammenhengen med det materielle og det mentale er om verden er kausalt lukket. Det er kan tenkes at en bevisst handling (jeg skriver disse ordene) bryter inn i de rene fysiske kausale prosessene som ville skjedd (eller snarere ikke skjedd) hvis jeg ikke hadde skrevet noe. Tilhørende dette temaet så har man spørsmålet om "downward causation", som er en nødvendighet hvis jeg kun er bygd opp av de materielle komponentene som er en del av kroppen min og disse styres av fysiske prosesser og krefter. Hvis mine "holistiske" tanker og handlinger ikke påvirker disse materielle komponentene vil "jeg" ikke ha en kausal kraft, da alt som skjer er styrt av de grunnleggende fysiske prosessene. Dette er et omfattende tema. I forhold til denne problemstillingen argumenterer jeg verken for «downward causation» eller at verden ikke er kausalt lukket. Det jeg argumenterer for er at som en del av den kausalt lukkede verden så finnes det et *funksjonelt* nivå over de rene fysiske prosessene, et nivå som realiseres av de (teleo-)dynamiske prosessene som er beskrevet i denne oppgaven. Det er disse prosessene som setter en retning på hva livet gjør, som har det grunnleggende elementet av formål.

organismens grunnleggende funksjoner. Det er dette jeg har forsøkt å etablere så langt i denne oppgaven.

Deacon beskriver en metafor som illustrerer teorier som forsøker å forklare teleologiske *fenomener* uten å benytte teleologiske *begreper*. Han kaller det en *homunculus* i våre teorier. Dette må ikke forstås som en idé om hvordan vi kan forklare bevisste fenomener, men hvordan vi forsøker å forklare et fenomen gjennom objektive beskrivelser uten å benytte begreper som er grunnleggende sett intensjonelle. Et konkret eksempel der slike forutsetninger avdekkes er analysen i kapittel 2.3. Begreper kan fremstå som objektive, men der man står i fare for å la dem i praksis fungere som noe intensjonalt, eksempelvis at verden har noe å formidle oss. Intensjonalitet har ofte en assosiasjon med noe mentalt, noe bevisst. Derfor etablerer Deacon et nytt begrep som kan kalle "ententional" som dekker alle disse funksjonelle områdene han analyserer:

Ententional phenomena include functions that have satisfaction conditions, adaptations that have environmental correlates, thoughts that have contents, purposes that have goals, subjective experiences that have a self/other perspective, and values that have a self that benefits or is harmed. Although functions, adaptations, thoughts, purposes, subjective experiences, and values each have distinct attributes that distinguish them, they all also have an orientation to a specific constitutive absence, a particular and precise missing something that is their critical defining attribute. When talking about cognitive and semiotic topics (sign processes), I will, however, continue to use the colloquial terminology of teleology, purpose, meaning, intention, interpretation, and sentience. And when talking about living processes that do not involve any obvious mental features, I will likewise continue to use the standard terminology of function, information, receptor, regulation, and adaptation. But when referring to all such phenomena in general, I will use entention to characterize their internal relationship to a telos—an end, or otherwise displaced and thus non-present something, or possible something. (Deacon 2011 s 27)

Fortolkes en teori som objektiv når den egentlig forutsetter slike intensjonelle fenomener, så er Deacons homunculus smuglet inn:

Homunculi are stand-ins for incomplete explanations, and since, according to current scientific canon, good explanations should take the form of mechanistic analysis, all references to the teleological properties attributed to homuncular loci need to be replaced by mechanisms. But often, efforts to explain away blatant homunculi lead to their unwitting replacement with many less obvious homunculi. Instead of a little man in the head, there are sensory maps; instead of an élan vital animating our bodies, there are genes containing information, signalling molecules, receptor sites, and so on, to do the teleological work. The process is thus

implicitly incomplete, and possibly not even able to be completed, requiring a similar effort at some later point in the analysis to deal with these new creations. The effort to deny the special character of intentional processes at that locus, and thus eliminate altogether any reference to teleological phenomena, only serves to displace these functional roles onto other loci in the explanatory apparatus. Instead of one homunculus problem, we end up creating many. And in the end the original homunculus problem remains. Removed from the body of scientific theory but unable to be silenced, it can only be hidden away, not finally eliminated. The analogy is striking, and cautionary (Deacon 2011 s 82-83)

Skal man derfor lykkes med å etablere grunnleggende mekanistiske forklaringsmodeller (årsak- virkning-relasjoner) må det utarbeides en teori som viser hvordan kognitive og proto-kognitive prosesser i vid forstand er mulig som noe selvfortolkende, noe som virkelig drar seg opp etter halen og ikke forutsetter implisitt en rasjonell fortolker eller designer. Det er nettopp et slikt forsøk som jeg har problematisert i kapittel 2.3. Spørsmålet til Deacon er om dette vil overhode vil være mulig.

Tilsvarende, hvis man vil redegjøre for levende organismer, så må man finne en måte å redegjøre for adferden til organismer uten å ta i bruk normative eller evaluerende begreper (så bort med instinkter). Opererer man på det rent fysiske nivået, forsvinner muligheten til å se livet, man ser bare materiens kausalmekanismer. Her finner vi et annet interessant fenomen. Kauffman skriver:

I note also that Humes famous “one cannot deduce an ‘ought’ from an ‘is’”, namely, the famous Naturalistic Fallacy, rests on the critical fallacy that Hume, like Descartes, thought of a mind “knowing” its world. Hume did not think of an agent “acting” in its world. Given “action” and “doing”, doing “it” well or poorly enters inevitably, and with it, “ought”. With ought, the need for evaluation, the rudiments of emotion without positing consciousness, enter. (Henning og Scarfe 2015, side 6)

I skiftet fra en mekanistisk forklaringsmodell til en teleologisk, innføres det en sentral tilleggsdimensjon, nemlig et normativt element; om noe som er egnet eller ikke egnet, om det som er bra eller dårlig. Er det mulig å gi en forklaring på hvordan et slikt normativt element kommer inn i forklaringsmodellene? Deacon mener at han kan ha en forklaring i hans redegjørelse for hvordan en organisme etablerer hva som er relevant informasjon som den må forholde seg til i omgivelsene. La oss se på denne redegjørelsen.

3.3.3 Fra er til bør

Deacon utvider sin teori om teleodynamiske prosesser, som beskrevet over, ved å koble disse sammen med hvordan organismer kan utvikle en relasjon til omgivelsene som muliggjør etableringen av en fortolkning om verden, hvordan semantisk (referensiell) informasjon kan utvikles. Dette er en videre utdypning av hvordan livet oppstod gjennom de teleodynamiske prosessene, først gjennom at det etableres autogener som et nødvendig skritt til en komplett celle.

Løsningen for å gå fra disse dynamiske prosessene til et holdbart begrep om informasjon er ifølge Deacon å knytte fortolkning av informasjon med den kausale relevansen til de teleodynamiske prosessene; det arbeidet de utfører:

This is why an interpretive process is more than a mere causal process. It organizes work in response to the state of a sign medium and with respect to some normative consequence—a general type of consequence that is in some way valued over others. This characterization of creating/ making a difference in both senses suggests that the sort of work that Bateson has in mind is not merely thermodynamic work. To “make a difference” in the normative sense of this phrase is (in the terms we have developed) to support some teleodynamic process. It must contribute to the potential to initiate, support, or inhibit teleodynamic work, because only teleodynamic processes can have normative consequences. So to explain the basis of an interpretation process is to trace the way that teleodynamic work transforms mere physical work into semiotic relationships, and back again. (Deacon 2011, side 393)

Deacon redegjør for hvordan samspillet mellom overføringen av informasjon fra omgivelsene til organismen skjer gjennom disse dynamiske prosessene. Oppsummert skriver han:

We can conclude that a representational relationship cannot be vested in any object or structure or sign vehicle. It is not reducible to any specific physical distinction, nor is it fully constituted by a correspondence relationship. But neither is it a primitive unanalyzable property of minds. Instead, even simple functional and representational relationships emerge from a nested interdependence of generative processes that are distinctive only insofar as they embody specific absences in their dynamics and their relationships to one another. These absences embody, in the negative, the constraints imposed on the physical substrates of signals, thoughts, and communications which can be transferred from one substrate to another, and which thereby play efficacious roles in the world as inherited constraints on what tends to occur, rather than acting as pushes or pulls forcing events in one direction or another. Constraints don't do work, but they are the scaffolding upon which the capacity to do work depends. (Deacon 2011, s 418-419)

Deacon forsøker her å forklare to ting. For det første hvordan en organisme kan bli responsiv til visse elementer i sin omgivelse slik at den evner å forholde seg aktivt til sine omgivelser ved å plukke opp signaler som er relevant for organismen. En skisse til en slik prosess er det jeg har beskrevet i denne oppgaven.

Videre argumenterer han for at dette også danner grunnlaget for å etablere hvilken verdi informasjonen kan ha for organismen, at det er en genuin overgang fra *er* til *bør* basert på fysiske og teleodynamiske prosesser. Hadde han lyktes med dette ville det virkelig vært et gjennombrudd. Jeg tror ikke han lykkes med dette siste leddet. Grunnen er det samme som Hume påpekte, man kan rett og slett ikke avlede *bør* fra *er*, fra hendelser til handlinger, fra det som skjer til det som søkes (Hume 1985). Disse kategoriene tilhører to separate begrepsapparat; dette er en *logisk*, og ikke en *empirisk* sannhet.

Teleologiske forklaringer tar utgangspunkt i et resultat og forklarer hvorfor dette oppstod basert på noe som vil gi et positivt resultat fremover i tid, mens en mekanistisk forklaring, uansett om den er basert på Newtons mekanikk eller kvantemekanikkens matematiske formler for felt, tar utgangspunkt i inngående betingelser for å beregne påfølgende effekter. Empirisk kan det virke som tidsdimensjonen er snudd, men det logiske poenget er at man forklarer resultatet på to forskjellige måter:

The general prima facie problem with teleological explanations is often said to be that they are 'forward-looking'. Teleological explanations explain the means by the ends; a development or trait is explained by reference to goals, purposes or functions, and so the explanans refers to something that is an effect of the explanandum, something that is forward in time relative to the thing explained. Of course this is quite unlike ordinary causal explanations in which the explanans refer to prior causes of the explananda. Indeed, because teleological explanations seem to refer to effects, rather than prior causes, it looks at first sight as though backward causation is invoked. (Neander 1991 s 455-56)

Det sentrale poenget ved Deacons innsikt, og som jeg har argumentert for her eksemplifisert med prinsippforklaringen til autogenet, er at behovet for teleologiske forklaringer aldri stopper. Man må anta teleologi fra første stund for å forklare fenomenet som levende organismer er. Det Deacon har gjort er å peke på hvilke naturlige elementer og prosesser som gir oss en grunn til å anvende et teleologisk begrepsapparat, som beskriver noe annet enn de konkrete fysiske prosessene og materielle elementene; denne teleodynamiske strukturen. Det er i dette skiftet av begrepsapparat, i *hvordan* vi forklarer et gitt fenomen, vi

finner grunnlaget i overgangen fra det deskriptive til det normative. La oss se nærmere på dette.

Ett alternativ til å akseptere Deacons teoretiske redegjørelse er å si: slik er det. Naturen er slik at liv forholder seg normativt til omgivelse. Vi kan redegjøre for de proto-kognitive prosessene som inngår i denne teleologiske relasjonen mellom organismen og omgivelsene, men det er så langt vi kommer i vår forklaringsmodell. Som Kauffman skriver i sitatet over: "Biosemioses is real in the universe". Hvis vi ikke benytter teleologiske og Deacons "ententional" begreper forstår vi ikke hva liv er; da ser vi ikke fenomenet vi ønsker å undersøke. Vi ser kun de fysiske prosessene som ikke kan forklare hva liv er. Opererer vi kun i begrepsapparatet til fysikken vil vi nemlig ikke kunne etablere de nødvendige egenskapene som skal til for å forklare de fenomenene som inngår i livet med tilhørende autopoietiske prosesser.

Vi kan imidlertid øke forståelsen av dette skifte fra hendelser til handling, fra statistisk-mekaniske årsaker til teleologiske mål. Deacon forsøker å vise hvordan hans begreper om "the absent", "constraints", morfodynamiske og teleodynamiske prosesser, arbeid, informasjon og rekursive lag av teleodynamiske prosesser er en slags tilleggsdimensjon til det materielle, som er like reell og sann, men ikke umiddelbart observerbar. Den viser seg gjennom dens funksjonelle virkninger i den naturlige verden, basert på en visuelt usynlig ramme som må avledes intellektuelt; en avstøpning vi kan lese ut av verden. Man kan metaforisk si at Deacon begynner med det materielle og foreslår en teori for hvordan dette må utvides for å gjøre plass for formål og mening.

Thompson begynner i den andre retningen med det han kaller "transcendental fenomenologi". Dette perspektivet bygger på Kants transcendentale idealisme slik det er videreført av den kontinentale fenomenologiske tradisjonen (særskilt Merleau-Ponty). Et slikt perspektiv forutsetter at erkjennelsen begynner med det subjektivt opplevde (bevisstheten) og fra denne etablerer vi kunnskap om objektet:

The point here would have to be something along the following lines: because of the transcendental status of consciousness (that consciousness is always already presupposed as an invariant condition of possibility for the disclosure of any object), there is no way to step outside, as it were, of experiencing subjectivity, so as to effect a one-one mapping of it onto an external reality purged of any and all subjectivity. It is in this transcendental sense that the phenomenal world is richer than any region of scientific objects—even the presumed "universal" region of physics. And it is only from this transcendental standpoint that Merleau-

Ponty's reversal of the order "physical → vital → human" or "matter → life → mind" makes sense. Mind emerges from matter and life at an empirical level, but at a transcendental level every form or structure is necessarily also a form or structure disclosed by consciousness. With this reversal one passes from the natural attitude of the scientist to the transcendental phenomenological attitude (which, according to phenomenology, is the properly philosophical attitude). (Thompson 2010 loc 1691)

Thompson begynner med bevisstheten (subjekter) og utarbeider en teori for hvordan man skal redegjøre for materielle objekter. Dette er ikke den idealismen jeg avviste i begynnelsen av denne oppgaven, som er empirisk idealisme. Dette er den transcendentale varianten, der man begynner med den menneskelige (organismens) erfaring som et utgangspunkt, uten dermed å hevde at det er slik verden er, i-seg-selv. Det er i stedet vårt epistemiske utgangspunkt vi må begynne med og fra dette benytte vårt konseptuelle begrepsapparat for å redegjøre for de empiriske fenomenene, objektene i den empirisk (materielle) virkelige verden. Jeg nevnte i innledningen at det er sterke likhetstrekk mellom en slik epistemologisk tradisjon og det synet på liv og kognisjon jeg har argumentert for i denne oppgaven, ved at begge tilnærminger tar utgangspunkt i subjektet (det erkjennende, aktivt handlende) i stedet for objektet (det subjektet har en relasjon til og søker å forholde seg til).

Som jeg skisserte i kapittel 2.3.1 trenger vi ikke å anta alle implikasjoner av transcendental idealisme, men vi kan bevare denne sentrale innsikten: Vi må fortolke fenomenene basert på signaler fra omgivelsene (altså empiri), og disse signalene fortolkes innenfor det begrepsmessige rammeverket som benyttes til å fortolke fenomenene; kunnskap er ikke utelukkende en kausal prosess. Det Deacon har vist er at vi har en berettiget grunn til å anta dette strukturelle elementet som realiseres gjennom livets materielle elementer. Verden består også av strukturer som er dannet og der disse strukturene er organisert slik at de søker mot et mål. For å beskrive disse fenomenene og det de søker, og ikke deres materielle komponenter, må det benyttes et teleologisk begrepsapparat. Det er dette skiftet av begreper, anvendt på de empiriske fenomenene, som danner grunnlaget for skillet mellom det som *er* og det som *bør* være (relativt til et mål). Dette siste skiftet mener jeg derfor ikke kan forklares empirisk, med Deacon, som en ytterligere teleodynamisk effekt, men basert på vår erkjennelsesteoretiske kontekst; i tråd med Thompsons analyser som bygger på Kants erkjennelsesteori. På denne måten, epistemisk, er subjektet mer grunnleggende enn objektet.

En slik forståelse legger også grunnlaget for en reflektert vurdering av hvordan en organisme sikrer sin egen opprettholdelse som ikke er tilgjengelig for organismen selv.

Organismen er underlagt de teleologiske prosessene som er etablert; den kan ikke forstå seg selv fra utsiden. Bare mennesker, så langt vi kjenner til, kan ha en begrepsmessig forståelse av hva som er godt eller dårlig for en organismes overlevelse, slik at vi kan forstå at og vurdere prosessenes egnethet. Vi kan for eksempel vurdere om det som tidligere var en vellykket overlevelsesstrategi, å fange mat i bevegelse; å unngå å ødelegge synet ved at det er utviklet store øyenbryn som fungerer som skylapper for solen; å navigere etter lyset fra månen eller søke korteste vei til havet, nå ikke er en god strategi. Disse overlevelsesstrategiene virker nå mot sin hensikt, ved at hvalen spiser en plastbit som beveger seg i vannet; ørnen som ikke ser vindmølleblader over seg; insekter som flyr mot lamper eller insektsfangere og skilpadden som krysser en trafikkert vei på vei til havet.

I sitatet til Kauffman over ser vi at det som er bra eller dårlig for en organisme bare er noe vi mennesker kan vurdere, og inngår således ikke i de teleologiske prosessene som sikrer overlevelse. Dennett gjør også et poeng av dette (2017 s 55-59). Han understreker at det er derfor det er så vanskelig å forstå bevissthet og dens funksjoner, da vi så lett forklarer "comprehension" ved å anta og tilskrive "competence" til organismen som utfører handlingen. Der Dennett trår feil er å koble denne evnen til å innta en reflektert relasjon til sine omgivelser til spørsmålet om indre opplevelser. Vi skal nå gradvis bevege oss over til dette temaet.

Oppsummert har vi etablert at verden har disse egenskapene som muliggjør mekanistiske og materielle prosesser og hendelser (les: fysiske) og at den samtidig muliggjør aktive, skapende teleologiske prosesser og hendelser. Denne tosidigheten kan vi ikke erkjenne samtidig fordi vi må benytte forskjellige begrepsapparat for å erkjenne disse forskjellige sidene av fenomenene. Dette er noe som har vært kjent siden Aristoteles, men der behovet for andre årsaker enn *causa efficiens*, årsak-virkningsrelasjonen, er blitt strekt utfordret i naturvitenskapen. I Store norske leksikon står det (<https://snl.no/%C3%A5rsak>):

Ikke minst synes det umulig å forstå språkhandlinger og deres virkninger (kommunikasjon) bare på basis av et mekanisk og statistisk årsaksbegrep.

Varianter av det gamle teleologiske årsaksbegrepet (*causa finalis*), som er bannlyst i naturvitenskapene, synes å komme inn igjen på en sentral plass i form av begreper som motiv og intensjon.

Men det hevdes også at motiver og intensjoner er grunner for handling og ikke årsaker. Skillet mellom grunn og årsak er viktig, men det er uklart hvordan de to begrepene forholder seg til hverandre.

I denne oppgaven er flere av disse påstandene utfordret, og en utdypet forståelse, og begrunnelse for, dette «bannlyste» begrepet om teleologi er pekt på som helt sentralt for ikke bare å forstå avanserte kognitive egenskaper som språkhandlinger, men også for den enkleste levende organisme. Det å peke på hvor man skal benytte hvilket begrepsapparat, er en sentral del av denne oppgaven, samt hvilken rolle de spiller og sammenhengene mellom dem. Hvis vi kun benytter det ene, så går vi glipp av sentrale elementer av virkeligheten. Som man ikke skal se på levende organismer som ren fysikk, må man heller ikke se på fysiske systemer som organismer; der animisme og Gaia-teorien kan være eksempler på en slik feil bruk av det teleologisk begrepsapparat.

Mange biologiske fenomener kan forstås gjennom en reduksjonistisk analyse av de materielle komponentene, men livets grunnleggende egenskaper ser man ikke fra et slikt perspektiv (begrepsapparat). Enten blir fenomenet borte eller så smugler man inn Deacons homunculus, altså disse restene av teleologi og mening som skjules i våre teoretiske redegjørelser. Uten dette doble perspektivet mellom statistisk-mekaniske og de teleologiske forklaringsmodellene jeg redegjorde for i denne oppgaven, så ser vi ikke hele bildet.

Jeg henviste til likheten mellom Thompson og Kant over. Kant selv pekte i samme retning som jeg har argumentert for, ved å peke på hvordan man skal forstå sammenhengen mellom levende organismer og den fysiske materielle verden:

But it is just as certain and beyond doubt that, in regard to our cognitive faculties, the mere mechanism of nature can furnish no ground of explanation of the production of organised beings. *For the reflective Judgement* it is therefore a quite correct fundamental proposition, that for that connexion of things according to final causes which is so plain, there must be thought a causality distinct from that of mechanism, viz. that of an (intelligent) cause of the world acting in accordance with purposes; but *for the determinant Judgement* this would be a hasty and unprovable proposition. In the first case it is a mere maxim of the Judgement, wherein the concept of that causality is a mere Idea, to which we by no means undertake to concede reality, but which we use as a guide to reflection, which remains thereby always open to all mechanical grounds of explanation and does not withdraw out of the world of Sense. In the second case the proposition would be an objective principle prescribed by Reason, to which the determinant Judgement must subject itself, whereby however it withdraws beyond the world of Sense into the transcendent and perhaps is led into error. (Kant 1914 s 297)

Motivasjonen for denne måten å skille mellom mekanistiske og teleologiske forklaringer, mellom (den empiriske) dømmekraften og den (reflekterende, abstrakte) fornuften, bygger på Kants transcendentale idealisme og hans skille mellom fenomen og noumen, mellom

tingen-for-oss og *tingen-i-seg-selv*. Kant kan hevde dette fordi, selv om vi kun kan ha kausale forklaringsmodeller på naturlige (materielle) fenomener, så er det rom for å se på den samme tingen som underlagt andre lover siden vår erkjennelse ikke sier noe hvordan tingene egentlig er (i-seg-selv) (Kant 2000). Kant indikerer i *Kritikk av Dømmekraften* at vi kan ha grunn til å stipulere, men ikke vite gjennom en empirisk dom, at levende organismer (ting) bør sees på som skapt av en fornuftig skaper som en forklaring på dette iboende formålsrettetheten.

Dette er en imidlertid en appell til en *ren mulighet*. Det er ikke etablert at det er *riktig*, at vi har et empirisk grunnlag for et slikt perspektiv, å beskrive naturlige fenomener som teleologiske. Det er å gi belegg til denne muligheten Deacon har åpnet opp for. Ved å gi en forklaring på hvordan vi kan forstå subjekter som ontologisk ureduserbare, ved å løfte frem disse ikke direkte observerbare rammene, "constraints", og hvordan disse styrer mulige utfall med det som ikke lenger er tilstede gitt slike strukturer, "absent", så har *han gitt et empirisk fundament for nettopp å anvende et teleologisk begrepsapparat på den naturlige verden*. Dette er en usedvanlig god innsikt og forklaring, og det er synd hans teorier ikke er løftet frem som et sentralt bidrag til teorier om liv og bevissthet. Jeg håper denne oppgaven er et lite bidrag til dette.

3.4: Evolusjon fra encellet til flercellede organismer

Vi har i hovedsak forholdt oss til de grunnleggende prosessene som styrer enkle organismer; encellede skapninger. La oss bevege oss fremover i evolusjonshistorien for å se når denne kompleksiteten øker og hva det kan ha å si for utviklingen av bevissthet.

La oss begynne med det enkleste: bakterien. Har så en bakterie noen form for subjektive opplevelser? Vi har nå etablert at alle levende organismer har en form for grunnleggende kognitiv kapasitet: de proto-kognitive prosessene er basert på at organismen forholder seg målrettet, teleologisk, overfor sine omgivelser. Måltrettetheten består i å søke å opprettholde og videreføre seg selv som noe levende; med nødvendighet adskilt fra sine omgivelser. Vi har sett at levende organismer er basert på strukturer som er bygget over dens materielle komponenter. Dette strukturelle nivået sikrer at livet kan styre energiutfoldelsen i verden, og det er denne kontrollen over energistrømmen, og de materielle komponentene som inngår i disse, som er den fysiske drivkraften til livet. Hvordan er relasjonen mellom denne proto-kognitive kapasiteten og den subjektive opplevelsen og, som en konsekvens, genuin kognisjon?

Er vi ikke tilbake der vi startet? Hvis det nå postuleres at det er en forbindelse, så fremstår det som om vi ikke har kommet noe nærmere en forståelse av “the hard problem”. Det vi har tjent augmentativt med vår analyse er en dypere forståelse av det levende som noe grunnleggende subjektivt og ikke som et objekt. Det er *nettopp disse grunnleggende subjektive livsprosessene som muliggjør at organismen har en teleologisk, og dermed i forlengelsen av det, en proto-kognitiv relasjon til omgivelsene*. Gjennom denne innsikten kan vi nå se en tettere forbindelse mellom kognisjon og subjektive opplevelser. Kan disse livsprosessene som resulterer i en proto-kognitiv relasjon mellom organismen også være relevant for å forklare bevissthetens indre fornemmelser? Alternativt kan det antas at den indre opplevelsen dukker opp på et senere tidspunkt. La oss se nærmere på dette.

3.4.1 De første subjektive opplevelsene

Godfrey-Smith forsøker å svare på dette direkte. Han benytter metaforen “white noise” (2016c, s 95) når man kan forestille seg den enkleste formen for indre subjektive opplevelse som kan følge disse livsprosessene vi har beskrevet. Hvis man forsøker å danne et bilde av hva den enkleste formen for opplevelse vi kan forestille oss er, så må det i sin kjerne bare være *et noe*, og denne idéen om hvit støy fungerer godt for vårt formål. Hva det faktisk er vil vi aldri kunne vite.

Hvis “white noise” er den enkleste formen for indre opplevelse, når kan dette ha oppstått? Godfrey-Smith spekulerer i at en slik udifferensiert indre opplevelse kan være en form for fornemmelse som følger “a crackle of metabolic electricity” (2016c s. 95-96). Dette kan sees på som er en slags udifferensiert fornemmelse som følger nettopp disse dynamiske prosessene som livet har kontroll over for å skape seg selv og opprettholde seg selv. Dette er ikke om noe, det bare er der. Dette kan da være basisen som den videre utviklingen av subjektiv bevissthet gjennom evolusjonen. Vi må huske at det er ikke noe fundamentalt skille mellom nerveceller og andre celler. Nevroner og nerveceller fyller en spesiell funksjonell rolle (sensitivitet, kontroll og koordinering over, relativt sett, store distanser, raskt og effektivt), men de besitter ikke grunnleggende egenskaper som andre celler ikke har.

Hvorfor gnistrer det ikke i hele kroppen vår da? Hvorfor er ikke en subjektiv opplevelse forbundet med alle levende aktiviteter? Og her kommer det sentrale poenget til Godfrey-Smith. En slik “white noise” kan sies å bare følge disse prosessene- den endrer ikke noe. For at vi skal komme fra dette hypotetiske grunnlaget for indre fornemmelser, til der vi kan begynne å snakke om reelle bevissthetsfenomener (sentience) som følger de kognitive prosessene livet utfører, så må de være *om noe*, de må ha en evolusjonær funksjon:

Why isn't [subjective experience] brimful of basic bodily rhythms, the division of cells, life itself? Some people might say it *is* full of those things- more than we realize anyway. I don't think so, and suspect there is a clue here. Subjective experience does not arise from running the system, but from the modulation of its state, from registering things that matter. These need not be external events; they might arise internally. But they are tracked because they matter and require response. Sentience has some *point* to it. It's not just a bathing in living activity. (Godfrey-Smith 2016c s 96)

Her har vi to viktige ideer. Den ene er at det må være et *noe* - eksempelvis “white noise” - som knytter seg til livsprosessene, for ellers ville vi ikke endt opp der vi er i dag. Den andre idéen relaterer seg til hvorfor vi har den eller den andre formen for subjektive opplevelse, så knytter dette seg til *noe som er viktig* for organismen, noe som hjelper den å overleve og å holde livet gående. Jeg vil utdype og motivere disse poengene i neste kapittel. For den følgende diskusjonen kan vi derfor anta dette som en mulig hypotese. Nå vil jeg derfor fokusere på den videre reisen fra “white noise” til “sentience” og hva vi kan utlede om subjektive opplevelser.

Godfrey-Smith generaliserer denne konkrete spekulasjonen. Som nevnt har vi etablert at alle levende organismer er kognitive i en grunnleggende forstand. For å illustrere relasjonen mellom (proto-)kognisjon (C på illustrasjonen) og subjektive opplevelser (sentience, S på illustrasjonen) har han skissert følgende modell (hentet fra Godfrey-Smith 2016a, s 25).

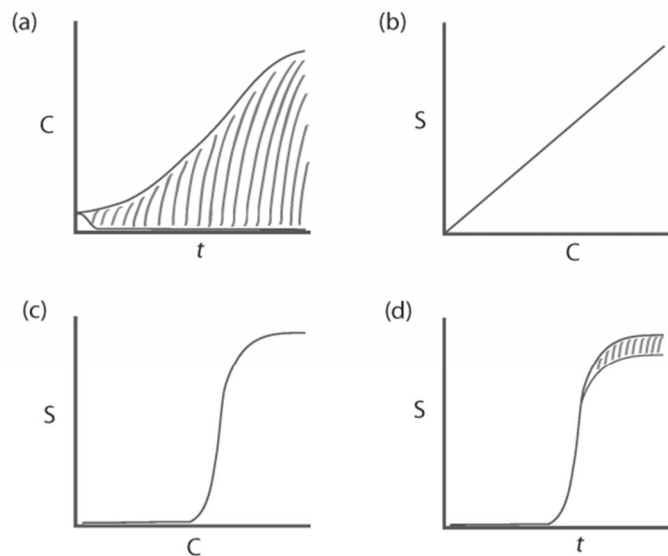


Figure 2: Possible functions relating evolutionary time to complexity of cognition (2a), complexity of cognition to sentience (2b and 2c), and time to sentience (2d, which assumes the applicability of 2c). See the main text for details.

Godfrey-Smith forklarer illustrasjonen på følgende måte:

Suppose that what we refer to as subjective experience, or the qualitative side of the mind, is just the cognitive side "from the inside." Then we can picture this with a simple linear function from cognitive complexity to the complexity of subjective experience (Figure 2b). For brevity I will use the term sentience (symbolized S) in the charts and the text below, to refer to this second side of the mind, and that term also encourages thinking in terms of degrees. (ibid s 26)

Her ser vi to alternative varianter. Figur 2a er det vi har argumentert for så langt. Kognisjon i vid forstand er med oss fra starten av livets opprinnelse (t på x aksen). C går ned til 0 etter livets opprinnelse på grunn av utvikling av virus og andre kvasi-livsformer som ikke kan opprettholde seg selv uten å basere seg på andre levende organismer. For illustrasjonsformål kan man forestille seg en direkte og nærmest lineær forbindelse mellom

kognisjon og “sentience”. Da får man graf 2b. Alternativt kan man anta at “the white noise” ligger som en grunntilstand uten noen særlig utvikling av den indre subjektive opplevelsen, og heller ikke endrer på det livsprosessene som danner grunnlaget for de proto-kognitive funksjonene som vi har beskrevet, før man fikk en eksplosjon av liv og kompleksitet i Cambrium. I denne perioden ble store og kompliserte kroppar utviklet med tilhørende økning av kompleksitet i sentralnervesystemet (2c). Legg merke til at “sentience”-streken også her ligger lavt på y-aksen fra starten. Da får vi konklusjonen, som er i tråd med rådende oppfatningen om at det kun er noen få dyr som har bevissthet verdt å snakke om, gjerne bare kompliserte pattedyr. Godfrey-Smith velger som vanlig den forsiktige veien og foreslår følgende konklusjon:

The two scenarios I sketched here are extreme options, and I'd expect the truth to be somewhere in between: only some cognitive capacities are associated with subjective experience, but these include some old inventions, such as the refference compensation I mentioned above. Another good candidate is the internal processing of valence, seen in internal reward systems and instrumental learning. Various other kinds of cognitive capacity are probably less subjectivity-relevant. (Godfrey-Smith 2016a s 26-27)

Godfrey-Smith argumenterer her i en retning som kan oppfattes som kontraintuitiv. Subjektive opplevelser har vært med oss lenge, veldig lenge. Med disse betraktningene har vi grunn til å tro at subjektive opplevelser har lite med hjernens kompleksitet å gjøre, eller Dennetts *memer*, men er mye mer integrert i livets grunnleggende prosesser. La oss se litt mer konkret på hvordan livet, kognisjon og bevissthet har utviklet seg over tid.

3.4.2 Utviklingen av kognisjon

Ideelt sett burde denne oppgaven nå legges bort for heller å lese Godfrey-Smiths *Other minds: The octopus and the Evolution of Intelligent Life*. I denne boken tar Godfrey-Smith oss med på en reise fra det første enkle livet til intelligent liv slik vi kjenner det i dag. Grunnen til at Godfrey-Smith løfter frem blekkspruter (cephalopoder) er at dette er det beste eksemplet på en fremmed intelligens vi kjenner til. Blekksprutfamilien skilte lag fra det som senere ville bli reptiler, fugler og pattedyr for 600 millioner år siden, altså vår gren på livets tre. Blekkspruten er nemlig den arten med klare tegn til intelligens og med kompleks atferd som samtidig er svært langt fra oss evolusjonært sett. Det finnes derfor svært få felles kognitive og somatiske prosesser mellom oss og blekkspruten, likheten er i så fall utviklet separat. Det er rett og slett det nærmeste vi kommer en intelligent “alien” her på jorden sammenlignet med oss mennesker.

Gangen i boken, og tilsvarende artikler fra Godfrey-Smith, er at det er utviklet proto-kognitive og kognitive ferdigheter som styrer samspelet i organismen selv og dens relasjon til omgivelsene som er tett forbundet med både miljøet og mulighetene som ligger i organismen. En blekksprut har for eksempel en mye mer distribuert intelligens enn mennesker. Den har svært mange nevroner i hver enkelt arm, armer som kan fungere semi-autonomt. Dette er tilpasset blekksprutens anatomi. Dens intelligente adferd er ofte et resultat av nødvendigheten av å forstå og forholde seg til omgivelsene slik at den hele tiden kan unngå farer. Som et bløtdyr har den få fysiske forsvarsmekanismer. For å klare seg har den utviklet andre forsvarsmekanismer, som blekket; evnen til kamuflasje gjennom tilpasning av fargene i huden; evnen til å etterligne andre dyr for avskrekking eller kamuflasje; og evnen til å beskytte seg med skjell som et heldekkende kroppskjold.

For å forstå hvordan livet har utviklet seg må vi forstå hvilke konkrete utfordringer det må løse, hva som skal til for å opprettholde og videreføre seg selv. Dette finnes mange måter å strukturere flercellede organismer på og hvor de tilhørende kognitive prosessene og somatiske strukturene vil være tilpasset hvilke utfordringer organismen må håndtere for sin egen del og i samspill med sine omgivelser:

For a large category of animals, including ourselves, there is definite individuality, achieved not so much through demarcation of boundaries, but through a coordination of parts that connect a whole system. We are readily countable things; there is a fairly definite number of humans in the world at each time, even though the boundary between each human and environment is more indefinite. There is not, in contrast, a definite number of corals or grasses. Those organisms have a modular design, made of small units, like coral polyps, that are organism like to various degrees in their own right. In familiar plants like oak trees, there is also partial independence of each branch, especially in reproduction. A coral or tree is in some respects more like a social entity than a single organism. Vertebrate animals like us, and also many invertebrates, are unitary organisms with respect to their organization and development. There is tighter integration across the whole and less autonomy of parts. (Cancer, often deadly for us, routinely arises in trees but is usually of little consequence.) Rapid and targeted interaction between parts in such organisms is ubiquitous, especially via the elaborate organ of coordination and integration seen in nearly all animals: the nervous system (Godfrey-Smith 2019, s 11)

Man kan derfor forvente å se helt forskjellige kontroll og koordineringsfunksjoner mellom de nevnte eksemplene som koraller, trær og komplekse dyr. Dyr har dette særskilte fenomenet at det fremstår som en enhet. Det er en kropp som må koordineres i riktig sekvens for at

man skal kunne bevege seg i en gitt retning, navigasjonen av omgivelsene basert på sanseorganene må sikre at relevant handlingsmønster gjennomføres og, sannsynligvis, en eller annen form for felles bevissthet slik at dyret velger det ene utfallet i stedet for det andre. Det finnes imidlertid mange måter å sikre på at en slik "enhet" blir realisert på, som eksempelet under viser:

The special kind of mental fragmentation seen in split-brain humans seems to be routine part of many animals' life.

Animals seem to have a range of ways of dealing with this situation. In the case of birds, incoming visual information can be... fragmented ...In birds like pigeons. Each retina has two different "fields", the yellow field and the red field. The red field sees a small area in front of the bird where there is binocular vision, and the yellow field sees a larger area that the other eye cannot access. Pigeons not only failed to transfer information between eyes; they also did quite badly at transfer between different regions of the same eye. This might explain some distinctive bird behaviours. Marian Dawkins did a simple experiment showing a novel object (a red toy hammer) to hens, who were allowed to approach and inspect it. She found that hens approached such an object in a weaving way that seemed designed to give the different parts of each eye access to it. That, apparently, is the way the whole brain gets access to the object. The weaving gaze of a bird is a technique designed to slosh the incoming information around²³. (Godfrey-Smith 2016c s 86-87)

Adferden til dyret kompenserer for "manglene" i det kognitive apparatet, slik at man får en samlet enhet som er godt nok for fuglen. Dette ser man igjen og igjen i evolusjonære prosesser, godt nok er tilstrekkelig for å videreføre livet. Det er ikke bare basert på behovet for å navigere omgivelsene vi finner hvordan kognitive prosesser er utviklet evolusjonært. Som nevnt må organismer forholde seg til omgivelsene, slik at dette kan legge grunnlag for en intern subjektiv evaluering av hvordan dyrets "indre tilstand" er for nettopp dette formålet:

While many discussions of the evolution of consciousness focus on perception, another tradition views a particular kind of feeling as a plausible candidate for a primitive, basic form of subjective experience. What is important, on this view, is not outward-directedness and perceptual engagement with the world, but the valenced registration of events—the distinction between good and bad, welcome and unwelcome. The folk notion of "feeling" identifies this phenomenon quite well; these are transitory internal events with a motivational role. (Godfrey-Smith 2019, s 13)

²³ Dette sitatet viser, som et apropos, hvor fullstendig integrert teleologiske og semantiske begreper er fullstendig integrert i vår beskrivelse av levende organismer, ref Deacons homunculus.

Dette begrepet om “valence”, som en slags indre fornemmelse om indre tilstander, er en sannsynlig kandidat for en kognitiv relasjon mellom organismen og omgivelsene som man har en god grunn til å tro inkludere en indre fornemmelse, altså en genuin kognitiv egenskap om organismens tilstand.

Det er mange fascinerende historier Godfrey-Smith beskriver. Under følger noen interessante elementer. Jeg søker ikke å begrunne disse, men heller illustrere de mulige perspektivene en evolusjonær forklaringsmodell av konkrete spørsmål om “hvorfors og hvordan” bestemte bevissthetsfenomener er mulig.

- Nødvendigheten for en organisme å opprettholde en barriere med omgivelsene reduseres fra overgangen til encellet til en flercellet organisme. Flercellede organismer er mer “avslappet” i forhold til sine ytre barrierer, fordi den enkelte celle i kroppen opprettholder sin. På denne måten kan man få på plass tette symbiotiske bånd mellom forskjellige livsformer, som relasjonen mellom sopp og planter og som bakteriene i magen og tarmen vår som hjelper til å bryte ned maten vi spiser.
 - Et spesielt eksempel på et slikt samspill mellom en flercellet organisme og enklere organismer finner vi hos blekkspruten. Hos bakterier er det utviklet en måte å kommunisere på: “quorum sensing”. Dette gjør at de kan koordinere sine aktiviteter, for et bedre resultat. En bakterie, som bor i huden til en blekksprut på Hawaii, kan produsere lys. Det lyset bakteriene produserer gjennom den nevnte kommunikasjonsformen er akkurat nok til å kansellere skyggen som blekkspruten kaster på bunnen, slik at den i praksis blir usynlig for de dyrene den jakter på om natten. Blekkspruten kan styre dette lyset ved å frigjøre næringsstoffer til bakteriene²⁴.
- Encellede organismer kan i de fleste tilfeller bevege seg for å navigere omgivelsene, ved hjelp av flageller (som halen til en sædcelle). Cellene i en flercellet organisme har i hovedsak mistet denne egenskapen. For flercellede organismer er det viktig å koordinere aktiviteten til mange celler i koordinasjon. Derfor oppstår behovet for “kontroll og koordinerings-celler”, altså nevroner. Når man må koordinere en kompleks og stor kropp må det være mulig å koordinere aktivitetene over et stort område raskt. Dette er nettopp funksjonen til nevroner, de er lange og kommuniserer raskt gjennom elektriske signaler. Deretter er de videre spesialisert i hjernen for å

²⁴ Se følgende link for mer informasjon om dette fascinerende samspillet.
<https://itsnotwhatitswhy.wordpress.com/2019/10/14/the-squid-without-a-shadow-how-hawaiian-bobtail-squid-use-light-producing-bacteria-to-sneak-up-on-prey/>

håndtere flere og flere kognitive oppgaver, herunder å motta informasjon fra omgivelsene gjennom forskjellige sanseorganer.

- Når en flercellet organisme beveger seg er det nødvendig for den å vite hva som er en endring, både forårsaket av organismens egne bevegelser og hva som er en ytre, ekstern, endring slik at den kan vite at det neste skrittet, vingeslaget eller svømmetaket går i riktig retning. Dette gir viktige feedbackmekanismer. Denne formen for kontroll og tilbakemelding er en annen mulig kandidat for en kognitiv relasjon mellom organismen og omgivelsene som sannsynligvis inkluderer et subjektiv opplevd element. Tenk bare på en fugl som flyr gjennom en skog, hvor perfekt tilpasset den er, og hvor lite tilpasset en fugl er til et transparent vindu. Den har ikke utviklet sanseorganer for dette formålet.
- Blekkspruter mangler reseptorer i øynene for å oppfatte hele fargespekteret, men allikevel klarer de å tilpasse fargene sine til omgivelsene. Dette får de til gjennom fotoreseptorer i huden (som man mener har en sammenheng med å oppfatte kinetisk energi- fysisk kontakt). Så det finnes mange måter å løse kontroll- og koordineringsoppgavene på og det er alltid den samlede funksjonen til organismen som helhet som betyr noe (deler og helhet står i gjensidig vekselvirkning)
- Med en gang man får skille mellom rovdyr og bytte oppstår det et evolusjonært press for å etablere en "teori" om andre sinn: Hva tenker det dyret jeg ser foran meg nå å gjøre? Dette "våpenkappløpet" gir en sterk evolusjonær tendens for å utvikle intelligens gitt at det er en vellykket strategi for overlevelse- som hos blekkspruten og hos nakne aper som er uten et effektivt naturlig forsvar eller våpen.
- En viktig del av den komplekse menneskelige bevisstheten er det såkalte "global workspace", som gjør omfattende integrasjon av minner, opplevelser og talehandlinger mulig. Dette kan også danne grunnlaget for vår spesielle selvbevissthet. Bakgrunnen for dette kan være drevet av vårt behov for omfattende sosial koordinering gjennom språk. Men dette er som nevnt ikke den biologiske forbindelsen med subjektive opplevelser det argumenteres for her, selv om dette er en populær filosofisk teori.²⁵

²⁵ "Driven by a variety of allegedly counter-intuitive consequences of first-order theories of consciousness, including skepticism about the range of organisms it spans, a number of philosophers have offered a variety of higher-order accounts of phenomenal consciousness. Such accounts invoke mental states directed towards other mental states to explain phenomenal consciousness. Carruthers' "higher order thought" (HOT) theory is that a mental state is phenomenally conscious for a subject just in case it is available to be thought about directly by that subject (Carruthers 1998a,b, 2000). The term "available" here makes this a "dispositionalist" account. The contrast is an "actualist" account, which requires the actual occurrence of the 2nd order thought for subject to be conscious in the relevant sense. According to Carruthers, such higher-order thoughts are not possible unless a creature has a "theory of mind" to provide it with the concepts necessary for thought about mental states. Carruthers' view is of particular interest in the current context because he has used it explicitly to deny phenomenal consciousness to (almost) all nonhuman animals" (Allen og Trestman 2017).

- Selv om blekkspruter er intelligente, så får de ikke mulighet til å lære så mye. Blekkspruter dør etter 1-2 år fordi det er krevende for et bløtdyr å beskytte seg tilstrekkelig i et fiendtlig miljø. De er derfor tilpasset korte, men intense liv. Unntaket er kjempeblekkspruten (squid) som lever i dyphavet.

Som nevnt innledningsvis så må vi besvare “why and how does conscious experience arise” for å besvare “the hard” problem. Problemet i sin rene form er bare dette generelle spørsmålet om hvorfor og hvordan vi har en indre subjektiv opplevelse? Dette skal vi se nærmere på i neste kapittel. Imidlertid er denne måten å stille spørsmålet på en generalisering. Spørsmålet ser bort fra den konkrete opplevelsen og stiller bare spørsmålet som sådan. Opplevelser finnes ikke i det abstrakte, så en “conscious experience” er alltid en opplevelse av noe og for noen. Og ser vi på den konkrete forklaringen på hvorfor man kan anta at både dyr og mennesker har den og den atferden med tilhørende konkret opplevelse, så kan vi faktisk svare på spørsmålet i sin konkrete form: “hvorfor og hvordan finnes denne opplevelsen?” basert på en analyse som skissert over. Med “opplevelse” menes den informasjonen en levende organisme kan tilegne seg fra omgivelsene basert på dens evolusjonært betingede forutsetninger og hvilken måte organismen opprettholder og viderefører seg i et gitt miljø basert på sine kognitive kapabiliteter.

Gjennom en analyse om de evolusjonære prosessene som har formet det vi har grunn til å anta er de subjektive opplevelsene til forskjellige organismer, så legger vi grunnlaget til å få svarene på det spørsmålet Chalmers stilte; ikke i det abstrakte, men i det konkrete. Og dette er en begynnelse for å redusere brodden i “the hard problem”. Vi har nå et bedre grunnlag til å svare på hva vi vet om den subjektive opplevelsen til andre mennesker og dyr. Godfrey-Smith forklarer:

If our treatment of subjective experience is guided by the biology, that suggests a graded treatment of this feature itself—of the fact of it feeling like something to be a particular organism. Philosophically, this outcome seems awkward and hard to think about, and if this is where we end up, we will want to look for new ways to describe the graded character of the situation on the mental side. We might say the goings on in various organisms are more experience-like, more experiential. The position that would result from accepting a gradualist view here need not be one in which the gradient runs down to encompass bacteria and the like. It might do so, but the account sketched above has been one in which some distinctively animal features—features of the relation between animal sensing and action—have special importance. Then the unicellular proto-subjects discussed early in this section would provide a frame or foundation for later stages that engender subjective experience.

Especially with respect to its details, the evolutionary sketch in this section should be seen as a how-possibly explanation. But it's a little more than that; the aim has been to walk through a series of steps such that once we reach one stage, we can see the next ones waiting. The result could then be called a how-possibly-necessarily explanation. Here the first modality—"possibly"—is epistemic. The second modality is not a strong sense of necessity, but something that contrasts with mere accident—a kind of robustness or predictability. It's a story about how things might have trodden a predictable path, on a planet like ours, a path by which subjectivity reliably arose. A story like this begins to make the origin of subjects intelligible, as Nagel and Levine asked. (Godfrey-Smith 2019, side 15-16)

3.4.3 Tilbake til Chalmers lette og vanskelige spørsmål

Vi er nå bedre filosofisk skodd til å se hvilken feilaktig antakelse som ligger bak Chalmers "hard problem". Den forutsetter en feilaktig dikotomi mellom det de lette og vanskelige (hard) problemene som ikke tar høyde for den formålsrettetheten vi finner i levende organismer. Etter denne gjennomgangen mener jeg at vi kan få en justert forståelse av Chalmers argumentasjon om "the hard problem". La oss derfor vende tilbake til det lange sitatet fra Chalmers i kapittel 2.2 og se hvor vi står nå.

Jeg har uthevet i **fet** skrift alle de stedene der jeg har utfordret Chalmers beskrivelser for hva som er en riktig forklaring på disse fenomenene. Det er virkelig ikke rett frem at disse fenomenene kan forklares reduktivt slik Chalmers hevder i innledningen av sitatet. Disse fenomenene er sterkt forbundet med teleologi og det Deacon kaller "entensjonalitet", som kan forstås som intensjonalitet uten en antakelse om bevissthet, disse proto-kognitive ferdighetene som det er gjort rede for. Det som er uthevet i *kursiv* er de stedene som, dersom Chalmers hadde sett alle de forholdene som er uthevet i fet, forutsetter dynamiske prosesser *av et subjekt*. Videre må disse kapabilitetene forstås *i en evolusjonær og historisk kontekst* (slik Godfrey-Smith redegjør for). Med et slikt perspektiv kan man stille spørsmålet om Chalmers ville sett på skillet mellom de lette og de harde spørsmålene like absolutt.

Throughout the higher-level sciences, reductive explanation works in just this way. To explain the gene, for instance, we needed to specify the mechanism that **stores and transmits hereditary information** from one generation to the next. It turns out that **DNA performs this function**; once we explain how the function is performed, we have explained the gene. To explain life, we ultimately need to explain how a system can reproduce, adapt to its environment, metabolize, and so on. All of these are questions about **the performance of functions**, and so are well-suited to **reductive explanation**. The same holds for most problems in cognitive science. To explain learning, we need to explain the way in which a

system's behavioral capacities **are modified** in light of **environmental information**, and the way in which **new information can be brought to bear in adapting a system's actions** to its environment. If we show how a **neural or computational mechanism does the job**, we have explained learning. We can say the same for other **cognitive phenomena**, such as perception, memory, and language. Sometimes the **relevant functions** need to be characterized quite subtly, but it is clear that insofar as cognitive science explains these phenomena at all, it does so by **explaining the performance of functions**. When it comes to conscious experience, this sort of explanation fails. What makes the hard problem hard and almost unique is that it goes *beyond problems about the performance of functions*. To see this, note that even when we have *explained the performance of all the cognitive and behavioral functions* in the vicinity of experience—*perceptual discrimination, categorization, internal access, verbal report*—there may still remain a further unanswered question: *Why is the performance of these functions accompanied by experience?* A simple explanation of the functions leaves this question open (Chalmers 1995 side 5)

La oss nå gripe tyren ved hornene og se på den eksplisitte linken mellom bevisstheten som en "indre opplevd" og som noe "ytre" eksternt funksjonelt. Det må understrekes at i kapitlet som følger vil jeg forlate en grunnleggende forankring i konkrete og empirisk funderte fysiske og biologiske fenomener. Neste kapittel vil være spekulativ metafysikk, men spekulasjoner basert på de teoretiske rammene som jeg har argumentert for og begrunnet så langt i oppgaven. Denne teoretiske rammen kan gi oss noen nye perspektiver på det metafysiske fundamentet til bevissthet og indre opplevelser. Slike perspektiver kan være nyttige intellektuelle verktøy for å få et bedre grep om dette vanskelige fenomenet bevissthet. Det er imidlertid mulig at noen resonnement må revideres eller forkastes. Jeg får leve med risikoen da enhver form for metafysisk spekulasjon medfører høy risiko for å ta feil i fravær av empirisk forankring. Leseren er herved advart. La oss gå i gang.

3.5: Bevissthet og metafysikk

Man kan fortsatt spørre: Hvordan er denne første fornemmelse, av en indre gnist, “white noise” mulig? Er ikke dette bare en variant av panspsykisme? Det er en grunnleggende forskjell å tenke på fornemmelsen som følger “the crackle of metabolic electricity”, at de finnes et indre “noe” som følger livets teleologiske prosesser, og panspsyksime. Det første er en konsekvens av en dynamisk prosess, et resultat av livets kontroll over energistrømmene. Denne kontrollen forutsetter overraskende kompliserte og sammensatte biokjemiske strukturer. Standard panspsykisme derimot, kan hevdes å tilskrive mulighetene for subjektive opplevelser som noe som ligger som basis i, som en indre egenskap, til det fysisk utstrakte og eksternt synlige materielle grunnlaget i verden.

Man kan si at panspsyksime er et forslag til en løsning på “mind-matter” problemet formulert i et materialistisk utgangspunkt: hvis vi antar at materialisme er riktig, og at ren materie ikke kan produsere subjektive opplevelser, så er det eneste stedet “sinnet” kan plasseres er inne i materien. Det finnes ikke noe annet sted hvis man antar Chalmers zombie argument og vi samtidig tar utgangspunkt i Chalmers redegjørelse av materialisme. Den redegjørelsen jeg har argumentert for her er at bevissthet er forbundet med livet, og livet er bygget over materien, gjennom de dynamiske strukturene og prosessene som er beskrevet. Det er disse prosessene som er fundamentet for subjektet, ikke materien som inngår i disse. Dette passer godt med Godfrey-Smiths kobling mellom subjektive, indre, opplevelser og den teleologiske “crackle of metabolic electricity”.

Men gitt at bevissthet er noe i verden, så må vi anta at det er basert på noe, noe som muliggjør det. Det følger med nødvendighet, at det verden er bygd opp av må tilby en slik mulighet. Det er i praksis det Galen Strawson argumenterer for. Leser vi hans panspsykiske teorier i konteksten av livets dynamiske strukturer kan vi finne en bedre fortolkning av panspsykisme enn noe som er universelt i materien:

So now I say that physicalism, i.e. real physicalism, entail panexperientialism or panspsychism. All physical stuff is energy, in one form or another, and all energy, I throw, is an experience-involving phenomenon. This sounded crazy to me for a long time, but I am quite used to it now that I know that there is no alternative short of ‘substance dualism’, a view for which (as Arnauld saw) there never has been any good argument. (Strawson i Freeman side 25-26).²⁶

²⁶ Se Barrett (2016) for en mer raffinert redegjørelse av denne ideen om energi som “experience-involving phenomenon”. Barrett peker på elektromagnetisme som den sannsynlige kandidaten, noe som passer godt med livets dynamiske prosesser, som nettopp styrer de elektromagnetiske kreftene;

Hvis alt som eksisterer er det fysiske og alt det fysiske er dette fundamentet for materie-energi som skissert i kapittel 2.3.1, hvor vi i tillegg hevder at subjektive opplevelser eksisterer, så, ved å følge resonnementet innledningsvis i kapittel 2.1, følger det at dette noe også må danne grunnlaget for indre opplevelser; altså er et “experience-involving phenomenon”. Om vi husker tilbake til kapittel 3.1, så er det som driver disse teleologiske prosessene som jeg har beskrevet nettopp en kontroll av energistrømmer, disse strukturene som setter rammer for hva som er mulige utfall av energi i de materielle komponentene. Gjennom å følge Godfrey-Smith og postulere det spede grunnlaget for subjektive opplevelser som “white noise” som følger med på disse energistrømmene, og at disse energistrømmene er slik dette grunnleggende fysiske noe utfolder seg, så har vi inkludert den grunnleggende idéen til panpsykisme, men uten dermed å godta eller måtte ta høyde for implikasjonene av en slik idé fortolket innenfor en materialistisk metafysikk²⁷.

Det er ikke slik at alle materielle elementer, selv den minste gravitasjonsbølge, det enkelte elektronet eller fotonet som har bevissthet og en indre opplevelse. Det må over en gitt terskel eller kompleksitet før det blir noe. Denne intuisjonen synes uunngåelig. Og det er nettopp dette som det argumenteres for. Det er ikke energi i seg selv som er bevisst, i betydningen av nevnte Einsteins $E=mc^2$. Det er først når den er strukturert og kanalisert at vi har grunnlaget for å anta at bevissthet oppstår. Den eksplisitte forbindelsen kan vi finne i livets grunnleggende teleologiske egenskap.

3.5.1 Sammenhengen mellom Chalmers lette og harde spørsmål

Tilbake til Godfrey-Smiths “crackle”. Vi har nå beskrevet disse grunnleggende livsprosessene utførlig. Benytter vi innsikten til Strawson så kan vi anta at det Godfrey-Smith beskriver som en “crackle” nettopp er det tilhørende “insideperspektivet” som følger disse livsprosessene. La oss derfor anta dette opplevde indre som følger livsprosessene til den enkleste bakterie. En slik opplevelse trenger imidlertid ikke å endre de teleologiske

som i bunn og grunn er kreftene som legger grunnlaget for molekyler og biokjemiske reaksjoner. <https://doi.org/10.3389/fpsyg.2014.00063>

²⁷ Tilnærmingen jeg argumenterer for unngår dette særskilte kombinasjonsproblemet som oppstår i kombinasjonen med panpsykisme fortolket i en materialistisk kontekst (se f.eks. Chalmers 2017). Påstanden er nettopp at subjektive opplevelser ikke dannes basert på mentale egenskaper i materien, men som et resultat av de energistrømmene livet har kontroll over, som ikke har noen særskilte (fenomenologiske) bestanddeler som må kombineres, men at den subjektive opplevelsen er en systemegenskap av de dynamiske prosessene som er realisert gjennom de materielle komponentene. Dette passer godt med den empiriske observasjonen at det er først ved aktivering av nevroner opplevelser dannes.

resultatene og den proto-kognitive prosessene vi finner i en celle, og mellom celler. Vi mangler ikke noe i forklaringsmodellen, nettopp som Dennett også understreker. Vi kan derfor forstå dette som en slags “dual-aspect” teori. De proto-kognitive funksjonene og formålsrettetheten til en organisme er det eksterne elementet, og “the white noise” er den interne egenskapen. Det er to sider av samme sak, de kan derfor sies å være et uttrykk for, eller et resultat av, de samme underliggende prosessene; det som opprettholder livet. Det er her Thompsons teorier går for langt. Vi har ikke noe god grunn til å anta at det må være et bevisst valg som styrer utfallet som sikrer organismens overlevelse.

Hvis våre indre fornemmelser bare følger etter livets grunnleggende prosesser, hvordan har det seg da at vi har dette rike indre livet som definitivt representerer den verden jeg som levende organisme forholder meg til? Hvordan etableres en genuin kognitiv relasjon til omgivelsene og ikke bare proto-kognitive teleologiske prosesser? For å få et grep om dette må vi tenke på livsprosesser, med en tilhørende indre tilstand, på to nivåer. Vi må skille mellom en forstått som en *førsteordens* funksjon og en *andreordens* funksjon. Det følgende argumentet er en generalisering og spissing av Godfey-Smiths redegjørelse om det samme tema (2016a).

Som en *førsteordens* funksjon er det snakk om metabolske (teleologiske) prosesser som sørger for å opprettholde livet som sådan. Dette er de teleologiske prosessene som er beskrevet i kapittel 3.1-3-3. I en slik relasjon kan det stipuleres at den indre fornemmelsen bare er en udifferensiert fornemmelse som følger disse førsteordens prosessene. Selv om dette kan sies å være en indre opplevelse, så er det som jeg sier ikke *om* noe for organismen, det bare *er*.

For at de indre opplevelsene kan sies å være *om* noe og ikke kun være direkte styrt av livets grunnleggende prosesser, så må det være snakk om en *andreordens* funksjon. Dette kaller Godfrey-Smith en “meta-metabolic” prosess. En slik *andreordens* funksjon vil være et resultat av en direkte evolusjonær funksjon som ikke kun har som funksjon å opprettholde organismen, men også å styre og kontrollere organismen. Her er vi tilbake påstanden om at det konkrete bevissthetsinnholdet må være *om* noe. For at dette skal være mulig må det finnes prosesser som ikke kun omhandler en direkte kontroll over livets teleologiske prosesser, men også indirekte for å styre organismen. Dette forutsetter at vi har en utvikling fra encellede til flercellede organismer. Flercellede organismer har nemlig helt andre utfordringer med kontroll, navigering og koordinering av det komplekse samspillet mellom og i seg selv og omgivelsene, som skissert i kapittel 3.4.2.

Det er her vi finner en genuin frikobling mellom metabolisme og en kognitiv representasjon som er *om* noe som er av betydning for organismen. Her ser vi en *andreordens* funksjon for livsprosessene, som danner grunnlag for en form for integrert indre opplevelse og en avansert kognitiv relasjon til omgivelsene, akkurat slik vi finner i flercellede organismer med et sentralnervesystem. Vi ser at denne overgangen i kompleksitet fra encellet, med en tilhørende udifferensiert indre opplevelse (hvor opplevelsen ikke er om noe, men prosessene som opplevelsen «følger» nettopp er dette gjennom organismens teleologiske relasjon til omgivelsene), til der *både* den indre opplevelsen og de teleologiske prosessene i de forskjellige spesialiserte nevronene er om noe; det de forholder seg til i omgivelsene eller kroppen til organismen. Her finner vi derfor en grunnleggende form for referanse mellom organismen og omgivelsene. Denne relasjonen mellom organismen og omgivelsene er *om* noe ved at det er utviklet funksjonelle elementer i en organisme hvis funksjon ikke direkte er styrt for å opprettholde organismen som sådan, men for å utvikle en funksjonell kognitiv relasjon med omgivelsene. Dette er den nødvendige relasjonen som skal sikre at en komplisert flercellet organisme skal kunne overleve.

Eksempler på slike grunnleggende kognitive prosesser er det nevnte “valence” og feedbackmekanismer for å navigere omgivelsene (Godfrey-Smith 2016a). Slike kontroll og koordineringsfunksjoner er som nevnt nettopp realisert av spesialiserte celler: nevroner og nerveceller. Noen eksempler på utviklingen av slike funksjoner er beskrevet i kapittel 3.4.2 over. Dette er også inkludert i det nødvendige samspillet mellom organismen og omgivelsene «states of the world» som beskrevet i kapittel 2.3.1, siden utviklingen av et sentralnervesystem og tilhørende sanseorganer er en forutsetning for å etablere muligheten for organismen til å motta informasjon fra omgivelsene gjennom signaler. Det er nettopp i dette tette samspillet vi finner våre subjektive opplevelser; at de er *om* noe, at de er viktig for organismens overlevelse.

I redegjørelsen for livets grunnleggende prosesser har jeg argumentert for at vi har to grunnleggende forskjellige hendelsesmodi i verden, en mekanistisk styrt av fysiske lover som omhandler objekter og en teleologisk som omhandler subjekter. Har vi grunn til å anta at det er snakk om et fundamentalt nytt nivå av teleologiske eller andre holistiske årsaksmekanismer når disse kognitive prosessene utvikles i flercellede organismer: alternativ 1? Alternativt er det snakk om en gradforskjell mellom de samme fenomenene som vi finner i encellede organismer, der “white noise” blir frikoblet fra rene metabolske prosesser og speiler utviklingen til en integrert indre opplevelse som følger de mer komplekse strukturer som blir utviklet i flercellede organismer: alternativ 2? Det alternativet vi velger vil ha en direkte kobling til vår analyse av “the hard problem”.

Hvis vi velger alternativ 1, må det antas at det er snakk om et ytterligere lag med og eventuell type av kausale nettverk, og som skiller seg vesentlig fra de underliggende teleologiske prosessene i den enkelte celle. Ved et slikt alternativ vil den indre opplevelsen, vår bevissthet, ha, eller danner grunnlaget for, en egen kausal kraft. Gitt en slik forståelse er vi tilbake til et skille mellom funksjon og opplevelse som er like sterkt som ved en materialistisk fortolkning. For å komme nærmere en løsning må vi velge alternativ 2.

Det er bare en økning i kompleksitet og spesialisering av funksjonen, en spesialisering som er drevet frem av evolusjonære prosesser, som skiller de proto-kognitive fenomener vi finner i alle livsprosesser og de fantastisk komplekse strukturene vi finner blant flercellede organismer, inkludert oss mennesker. Det holder å etablere teleologi og koblingen til indre fornemmelser en gang. Men er vi ikke tilbake til konklusjonen om at bevissthet da bare er et epifenomen? Nei, for det analysen har vist er at bevissthet og liv er to sider av samme sak, og livsprosessen som etablerer levende organismer og avledede proto-kognitive og kognitive egenskaper, er definitivt ikke et epifenomen av de materielle komponentene det består av.

Bevissthet inngår i en bestemt evolusjonær funksjon som er utviklet for organismens overlevelse og på denne måten er bevissthetens funksjoner og egenskaper direkte kausalt integrert i disse teleodynamiske prosessene med tilhørende funksjonelle virkninger, med en tilhørende innside som resulterer i en gradvis rikere indre opplevelse. Her har vi en konkret måte bevissthet, våre indre opplevelser, samspiller kausalt med resten av den fysiske verden, hvordan det mentale og det materielle er direkte kausalt forbundet og tar del i et felles kausalt nettverk. Bak det hele ligger livets kontroll av energistrømmene.

Mellom bevissthet og det materielle grunnlaget finner vi nettopp disse teleologiske livsprosessen som danner egne kausale nettverk, og der det vi kaller kognisjon – med en tilknyttet indre subjektive opplevelsen - er et spesialtilfelle av disse grunnleggende prosessene. Disse strukturene er evolusjonært utviklet for å kunne styre en aktiv målrettet organisme som ivaretar organismens behov for strukturell og relasjonell enhet, der de funksjonelle egenskapene legger til rette for å styre kroppen, fortolke omgivelsene gjennom sansene, kommunisere med andre likesinnede og lære abstrakt symbolsk manipulering for å tilpasse seg et komplisert sosialt miljø. Når disse funksjonene realiseres følger det med et rikt indre liv som en del av disse prosessene. Og det er nettopp dette vi søkte å forklare når vi begynte vår analyse av Chalmers "hard problem". Mange detaljer om hvordan dette

foregår i praksis er ikke redegjort for, men her er et konkret forslag på en prinsippforklaring på samspillet mellom bevissthetens «easy» og «hard» problem.

Med en konklusjon som beskrevet i dette kapitlet følger jeg Dennett i analysen av hvordan vi kommer fra noen grunnleggende enkle elementer til svært kompliserte fenomener. For å gjennomføre sine forklaringer hevder Dennett å trenge bare en naturlig realisert Turing-maskin samt Darwinistisk evolusjon (Dennett 2017), det jeg har kalt en Turing-robot. Som jeg har argumentert for så må vi nekte Dennett en slik Turing-robot. I stedet må vi forstå livets grunnleggende prosesser som teleologiske og som styrende for hva som er relevant informasjon i omgivelsene for organismen. Når disse elementene er etablert så har vi det vi trenger for å analysere livet og bevisstheten. Den videre reisen er, med Chalmers idé om “enkle” spørsmål, “bare” lag på lag med teleologiske prosesser, slik de kan forstås gjennom evolusjonære forklaringsmodeller.

En mulig innvendig som bør belyses før vi går videre er at det kan synes som det finnes empiriske funn som direkte motbeviser en form for «dual-aspect» teori som jeg argumenterer for, hvor den subjektive opplevelsen er det som følger med sensoriske kapabiliteter og tilpasset adferd realisert gjennom evolusjonært utviklede tilpasninger for å sanse, styre, koordinere og kontrollere en komplisert organisme. Fenomenet «blindsight»²⁸ - blindsyn – kan synes å falsifisere en slik teori fordi det her kan synes som om man kan «se» omgivelsene på en grunnleggende måte, uten å være bevisst dette.

Blindsight-fenomenet oppstår når en person får skader i den delen av hjernen (visual cortex) eller i tilknyttet signalsystem som benyttes for å prosessere synsinformasjon og som følgelig gjør at man kan se sine omgivelser. Personer som har fått en slik skade rapporterer om at de ikke kan se ting foran seg- at de i praksis er blinde i større eller mindre grad. Imidlertid kan de fortsatt reagere på objekter som beveger seg mot dem og samhandle med objekter i omgivelsene som de ikke ser visuelt. Hvis påstanden er at det følger med en subjektiv opplevelse med evnen til å prosessere informasjon fra omgivelsene, at dette er to sider av samme sak, så kan en slik påstand umulig stemme.

Nøkkelen her er hva vi mener med «subjektive opplevelser» og bevissthet. Som nevnt har mennesker en avansert og integrert subjektiv opplevelse – bevissthet om – våre omgivelser. Dette er en bestemt type for subjektive opplevelser som forutsetter et komplisert kognitivt

²⁸ Beskrivelsen av dette fenomenet baserer seg på Godfrey-Smith (2016c s 87-90) og <https://en.wikipedia.org/wiki/Blindsight>

repertoar som er svært integrert på tvers av funksjoner. Ved skader i hjernen så kan denne integrasjonen brytes opp, slik at en gitt subjektiv opplevelse ikke lenger er integrert med evnen til å avlegge verbale rapporter (av avart av split-brain fenomenet), eller mulighet til å danne minner. Videre er det realistisk å tro at visse grunnleggende evner til å respondere til omgivelsene fungerer i andre, og mer primitive, feed-back nettverk enn den delen av hjernen som benyttes til å analysere og forstå sensorisk informasjon (den delen som er skadet), med tilhørende forskjellig subjektiv opplevelse. Dette kan forklare hvorfor man kan tilpasse seg omgivelsene uten at det føles som om man ser; at man opplever seg selv som blind. Med Godfrey-Smiths ord:

A simple interpretation of these cases holds that you need the ventral stream [performing categorisation, recognition] to have any experience of what's coming in through your eyes at all. That is probably too simple. It's likely that dorsal stream [performing real-time navigation through space] vision feels like something, though it doesn't feel much like seeing. The details of these two "streams" are less important than the larger picture coming out of this work. This is the fact that quite complicated processing of visual information – processing that runs all the way from eyes, through brain, to legs or hands – can take place without the subject experiencing any of this *as seeing*. (Godfrey-Smith 2016c s 88).

Å redegjøre for slike fenomener krever en grundig analyse og jeg mener ikke at den korte redegjørelsen over gir et utdypende svar på fenomenet. Imidlertid kan man ikke avfeie den foreslåtte forklaringsmodellen med å peke på at det er mulig å gjøre noe uten at man er bevisst hvordan dette ble gjort, som anekdoter at man kan kjøre bil på «mental autopilot» eller at hjernen tar en «beslutning» noen millisekunder før «man selv» er bevisst det. Begrepet «bevissthet» kan benyttes i mange kontekster og har mange betydninger og det krever en grundig analyse for å forstå hva som skjer i hvert av de enkelte tilfellene.

3.5.2 Hvorfor subjektive opplevelser?

Gitt analysen så lang gjenstår det å spørre: Hvorfor er det slik at det følger med en indre opplevelse som en del av de grunnleggende teleologiske livsprosessene? Godfrey-Smith behandler dette temaet eksplisitt og kommer med følgende konklusjon:

I said earlier that the heart of my view, which I see as materialist, is monism and the construction of mind from things that are not mental. Some kinds of neutral monism would seem to pass this test for being materialist. The mental is not derivative on the paradigmatically physical, but it is derivative, in a sense, on the neutral, and hence on

something non-mental. (The paradigmatically physical is derivative in the same way.) This blurring of the distinction between materialism and neutral monism might be OK, or we might look for something stronger to mark out materialism, such as the non-derivative nature of the physical. (Godfrey-Smith 2018 s 16).

Dette er en posisjon som minner sterkt om denne idéen jeg presenterte om at dette noe (materie-energi, kvantemekaniske felt) som alle naturlige hendelser folder seg ut i, er det som er det grunnleggende elementet for alle naturlige fenomener. Gitt en slik antakelse kan det hevdes at både det fysiske og det mentale er “derivative” fra et grunnleggende noe som er det som er basisen for alt som eksisterer. Dette er en posisjon som ofte beskrives som en form for monisme. I praksis må alle som anerkjenner subjektive bevisste fenomener som noe som eksisterer godta en slik antakelse. Intuisjonen til Godfrey-Smith om at det materielle på en måte er mer grunnleggende enn det mentale er også ivaretatt i redegjørelsen jeg har argumentert for, siden livet er realisert gjennom strukturer som forutsetter de materielle komponentene disse inngår i. Den sentrale innsikten er at man ikke kan forstå disse fenomenene reduksjonistisk.

Alternativet er å anta at subjektive opplevelser oppstår basert på “ren kompleksitet” på et eller annet tidspunkt i evolusjonens gang. Hvis vi ser bort fra de prinsipielle betraktningene rundt sterk og svak emergens, tilsier empirien:

Alternatively, we might argue that a full-fledged version of experienced mentality comes in at some specific stage. Why would this happen? It would help if there were some *sui generis* features of human brains, or perhaps mammal brains. It is not impossible to argue for such a thing, and many sorts of empirical work might be relevant. Recent work, though, has gone the other way, and has also tended to find apparent marks of some sort of experience in an ever-wider range of animals, however (crabs, bees, octopuses...). The view that looks plausible in the light of recent biology is one that recognizes a gradual evolution of the capacity to experience events. A break between "higher" and "lower" animals is looking less and less likely. (Godfrey-Smith 2018 s 15)

På en måte slipper vi ikke unna kjernen i “the hard problem”. Vi vet simpelthen ikke hvorfor vi er bevisste og hva som er den fundamentale egenskapen for dette fenomenet som både er mystisk og uforståelig, men også samtidig umiddelbart intuitivt åpenbart. Vi har i praksis den samme utfordringen med materie, som er det andre grunnleggende metafysiske elementet som inngår i verden, drevet frem av energiutfoldelsen. Vi vet at den er der og vi kan se hvilke deler det er bygget opp av, materie. Den ytre verden generelt sett har denne klare fordelingen at den kan observeres direkte gjennom alle sansene våre. Med våre indre

opplevelser er det annerledes. Disse kan vi bare oppleve subjektivt og ikke observere intersubjektivt. I fravær av muligheten til at våre indre opplevelser kan observeres av en tredjepart, kan de virke som noe uforklarlig eller upålitelig. Poenget til Thompson står seg imidlertid. Vi ville ikke kunne observert den ytre verden om ikke det var for våre subjektive opplevelser. Som Thompson skrev, så er den fysiske materielle verden epistemisk avhengig av, og forutsetter, den indre opplevde, mens den indre opplevde er fysisk avhengig og er betinget av at den kan realiseres av dens fysiske grunnlag.

Hvorfor er det slik? Hva er det som "egentlig" er grunnlaget for materie, for elektriske felt, for feltene beskrevet i kvantefysikken, dette noe som alle fysiske fenomener utfoldes i? Hvorfor eksisterer noe i det hele tatt, hvorfor finnes universet? Vi har kommet til et punkt der vi ikke kommer lenger. Vi kan bare si: Sånn er verden. Verden er materialistisk, men den er også grunnleggende teologisk, og sammen med disse teleologiske livsprosessene følger det en "innside" som legger grunnlaget for subjekter; for organismer med indre opplevelser av sine omgivelser bestemt av evolusjonært utviklede kapabiliteter. Når en avansert organisme, mennesker, reflekterer om sin plass i denne verden og hvordan den skal forstås, så finner vi disse to forholdene med en gjensidig, men forskjellig, avhengighet til hverandre. Disse forholdene danner den filosofiske bakgrunnen for Chalmers lette og vanskelige spørsmål. Det er mulig å reflektere rundt dette temaet og komme til sannsynlige svar, slik jeg har gjort i denne oppgaven med veiledning fra Godfrey-Smith, Deacon og Thompson.

Redegjørelsen over er ikke den eneste måten å argumentere for en kobling mellom en indre opplevelse og kognitive prosesser som vi finner hos levende organismer. Som nevnt har Dennett et annet perspektiv hvor han argumenterer for at menneskelig form for bevissthet, vår selvrefleksjon og vår måte å redegjøre for vår introspeksjon på, er et resultat av vår evne til språk og kommunikasjon. Denne ligner på teorien om "global workspace" jeg nevnte i kapittel 3.4.2. Dennett er skeptisk til det han mener er en mangelfullt ekspliserte og forklart forståelse av hva vår indre, subjektive opplevelse innebærer. Dennett beskriver derfor vår common-sense basert forståelse av indre opplevelser og «selvet» som en «user illusion» (Dennett 2017, s 335). Implisitt i et slikt perspektiv går Dennett mot Godfrey-Smiths konklusjon om at «the view that looks plausible in the light of recent biology» er at bevisstheten følger livet langt tilbake: Dennett ser helt annerledes på det:

Temporary agnosticism about consciousness is fine—I have just advocated it—but not agnosticism saddled with the proviso that of course other animals are conscious even if we can't say what that means. That is, at best, the expression of confidence in the eventual triumph of the manifest image over the scientific image, in the face of a long history of defeats.

It used to be obvious that the sun circled the earth, after all. People who won't let themselves even think about whether grizzly bears are "conscious like us" (whatever that means) are succumbing to ideology, not common sense. (Dennett 2017 s. 337)

Dennett ønsker observerbare fenomener som gir grunnlag for å anta bevissthet i vid forstand²⁹. Dette finner Dennett bare i forbindelse med kommunikasjon og språk.

There might be some other evolutionary path - genetic, not cultural- to a somewhat similar user-illusion in other animals, but I have not been able to conceive of one in convincing detail, and according to the arguments advanced by the ethologist and roboticist David McFarland ... , "Communication is the only behaviour that requires an organism to self-monitor its own control system." (Dennett 2017, s. 342)

Jeg har ingen ambisjoner om å analysere Dennetts forklaring på hvordan vi skal forstå subjektive opplevelser, men mer å understreke muligheten for flere perspektiver. Det vi ser som en virkelig vanskelig posisjon er å opprettholde at bevissthet og indre opplevelser i en vid forstand dukket opp på et gitt tidspunkt i evolusjonens gang. Det er med en slik antakelse svært vanskelig å peke på hva som kan forklare denne overgangen. Dette er både Godfrey-Smith og Dennett enige om. Da blir vurderingen om hva som gir den beste forklaringen, enten svært tidlig eller svært sent i utviklingen av livet. Jeg har argumentert for at det første alternativet er mer plausibelt.

Nettopp en slik mellomløsning løftes frem når man intuitivt prøver å forstå hvordan bevisstheten utviklet seg i levende organismer. En mulig tanke er at denne intuitive oppfatningen er motivert, ikke bare av idéer om bevissthet, men også at liv er et spesielt fenomen. Hvis man antar en grunnleggende materialistisk forklaringsmodell i tråd med Chalmers så må denne idéen om at levende vesener må være bevisste «som oss» også ta høyde for en intuisjon om at mindre komplisert levende organismer ikke bare kan være rene biokjemiske maskiner. En celle, ja; trær, trolig; insekter, kanskje; men fisk, fugler, reptiler og andre pattedyr? Man kan umulig se på dem som rene biokjemiske maskiner!

²⁹ Et sentralt poeng i forhold til Dennetts teorier om bevissthet er at han er uklar på hvordan vi skal forstå begrepet om «consciousness». Han argumenterer som nevnt for at menneskets evne til å forstå (comprehension) stammer fra vår evne til å ta til oss idéer, memorer, som gir inntrykk av at vi er bevisste selv om vi egentlig ikke er det, altså den nevnte «user-illusion». Det kan virke som om at han sier at organismer som ikke har denne evnen til forståelse følgelig ikke er bevisste, og derfor ikke har en indre opplevelse, uten at han er veldig eksplisitt på dette temaet om subjektive opplevelser. Godfrey-Smith ser imidlertid på denne vide forståelsen til bevissthet som et spesialtilfelle av den mer grunnleggende evnen til å føle noe som sådan, til å ha subjektive opplevelser (2016c s 92). Det er denne grunnleggende evnen til å føle, «sentience» og ikke «consciousness» som Godfrey-Smith kaller det (2016c s. 79), som er det vi må forklare for å forstå bevissthetsfenomener og ikke motsatt.

Det må være *noe* som skiller levende vesener fra «ren materie». Dette noe som sikrer at vi får etablert det ønskede skillet kan derfor være «bevissthet». Dette er den siste rest av noe uforklart, idéen om bevissthet, som støtter denne intuisjonen om at levende organismer skiller seg fra død materie. Hvis vi imidlertid har sikret skillet mellom død materie og levende organismer ved livets opprinnelse, så trenger vi ikke denne appellen. Det er grunn til å tro at bevissthet kan komme i alle fasonger og varianter, og det er ikke behov for å hevde at andre organismer er bevisste *som oss*. Men der vi ser tegn på bevissthet, så har vi heller ingen grunn til ikke å anta at de også har en indre opplevelse som speiler de kognitive ferdighetene de viser gjennom sin adferd. Dette er mulig siden vi ikke trenger å anta denne særskilte formen for reflektert og begrepsmessige bevisstheten som vi mennesker har utviklet, for å forklare de valgene og den atferden levede organismer viser. Det som fremstår som et ekstra lag med rasjonalitet, det Dennett kaller “comprehension” er noe som kun er tilgjengelig for oss mennesker når vi har en reflektert, begrepsmessig relasjon til de teleologiske fenomenene. Organismen selv har ikke forutsetning til å forstå hvorfor den gjør som den gjør, hvis ikke den har et språk og begreper som oss mennesker.

Før jeg runder av denne siste spekulative delen av oppgaven kan det pekes på et siste argument. Det går fra å sannsynliggjøre en forbindelse mellom subjektive opplevelser og livets teleologiske prosesser, til å hevde at en organisme med avanserte kognitive kapabiliteter faktisk *må ha* en indre opplevelse, at vi har grunn til å tro at det er en form for bevissthet hos en kognitiv organisme. Et slikt argument kan vi se ved å vurdere levende organismer på en kontrafaktisk måte. I sin mest generelle form knytter slike kontrafaktiske vurderinger seg til de grunnleggende egenskapene til universet, som universelle konstanter som er styrende for hvordan de forskjellige fysiske kreftene virker. Eksempler på dette er Placks konstant, Newtons gravitasjonskonstant, lysets hastighet, styrken i elementærladninger og lignende.

Hadde disse konstantene vært litt annerledes ville ikke universet slik vi kjenner det i dag vært mulig. Hadde f.eks. gravitasjonen vært vesentlig sterkere ville ikke ekspansjonen vi ser i vårt univers vært mulig eller at en massiv ansamling av materie ville kollapset til et svart hull og forhindret den kjernefysiske fusjonen som muliggjør stjerner slik vi ser i vårt univers. Hadde elektronbindingene mellom atomene vært vesentlig svakere ville ikke kompliserte molekyler vært mulig. I begge tilfeller ville ikke universet slik vi kjenner det vært mulig og følgelig ville ikke levende organismer kunne eksistert, inkludert mennesket. En slik finjustering av universelle konstanter er blitt løftet frem som et mulig argument for en intelligent skaper. Alternativt kan dette settes frem som et argument for at det kan finnes uendelig mange univers, med tilhørende varierende universelle konstanter, slik at bare noen

få har forutsetninger for å danne de kompliserte fysiske systemene vi finner i vårt (Carroll 2017).

Hvis vi går nærmere denne oppgavens tema, så finner vi det samme fenomenet. Levende organismer må være relativt konstante. De må derfor basere sin materielle struktur på et materiale som er kjemisk stabilt. Organiske molekyler basert på karbon har denne nødvendige stabiliteten. For å sikre energiproduksjonen i en organisme med en slik kjemisk sammensetning kreves oksygen. Oksygen må derfor fraktes til cellene som trenger oksygenet. Dette kan gjøres av et materiale som binder seg lett til oksygen. Jernioner har denne egenskapen, og dette benyttes av pattedyr og fugler, derfor er blodet rødt. Kobberioner har også denne egenskapen og dette er et design som benyttes av blekkspruten, derfor er blodet blått³⁰. For begge løsningene er det en forutsetning at det finnes materialer, fysiske komponenter, som binder seg med oksygen, eller generelt det materialet som inngår i energiproduksjonen til en organisme. Hos trær ser vi også dette fenomenet ved at vann, som benyttes for energiproduksjonen og tilførsel av næring til planter gjennom fotosyntese, gjøres tilgjengelig for cellene i bladene. For at et slikt drivstoff for organismen skal gjøres tilgjengelig, må dette materialet transporteres til de cellene som trenger energien. Det er følgelig noe som frakter materialet rundt i organismen. Dette vil være en form for pumpe, som et hjerte hos dyr, eller kapillæreffekter og fordamping i trær, som sørger for transport av det som dekker energibehovet i organismen.

For at sentrale funksjoner i en organisme skal utvikles så er det en forutsetning at visse naturlige fenomener har visse egenskaper, slik at dette mulighetsrommet kan benyttes gjennom seleksjon, som en ressurs eller redskap for evolusjonen. Oksygen må kunne binde seg til jern eller kobber, vann må ha en høy kapillæreffekt for at det skal kunne suges opp fra røttene til bladene. Hadde ikke disse naturlige egenskapene vært som de er, ville ikke slike prosesser blitt etablert. Tilsvarende ser vi også i Deacon sin teori. Termodynamiske effekter etablerer de emergente regularitetene (morfodynamiske prosesser) som er en forutsetning for at prosessen mot teleodynamiske systemer skal kunne etablere seg. Hadde det vært annerledes ville ikke det resultatet vi observerer vært mulig; levende organismer ville ikke blitt realisert.

³⁰ Lite i naturen skjer uten en grunn. Grunnen til at blekksprutens blod er basert på kobber og ikke jern er at det er mindre tilgang på oksygen. Opptaket av oksygen basert på kobber er mer effektivt enn jern, men kan knytte seg til mindre oksygen. Dette er derfor en mer effektiv løsning for dyr som lever i vannet. Se <https://forskning.no/blod-menneskekroppen-spor-en-forsker/er-blod-alltid-rodt/387769> for mer informasjon om forskjellige måter å absorbere oksygen fra omgivelsene på.

Det er tenkelig at noe tilsvarende kunne skje, men da basert på andre effekter og egenskaper enn de som er de faktiske forholdene vi finner i vår verden. Det er mange veier til mål. Blod kan som nevnt benytte både kobber og jern som materialet som binder oksygen. Et annet eksempel på slik konvergent evolusjon finner vi hos dyr som lever under vann. Morfologien til fisker, fugler og pattedyr er i utgangspunktet svært forskjellige. Disse forskjellige kroppsfasongene blir utsatt for et sterkt evolusjonært press ved at organismen må ta høyde for de grunnleggende utfordringene som må løses, et må sikres en effektiv bevegelse i vann. Vannet har visse fysiske egenskaper som tetthet og trykk som må håndteres av organismens kroppsfasong og muskulatur for å bidra til organismens overlevelse. De naturlige omgivelsene setter derfor et krav til hvilke funksjoner som må etableres og evolusjonen vil da kunne «utforske» alternative løsninger basert på de fysiske mulighetene og egenskapene som kan testes ut gjennom en konkret organisme. Vi ser derfor at selv om utgangspunktet er forskjellig, så konvergerer makrellstørje, pingvin og delfin til lignende kroppsfasong fra svært forskjellig utgangspunkt basert på ytre rammer i miljøet. Videre ser vi at løsninger for transport av næring og energi konvergerer funksjonelt, men der forskjellige konkrete måter å realisere det funksjonelle designet på er basert på egenskapene ved de fysiske komponentene som kan benyttes av evolusjonen for å «bygge» organismen.

De grunnleggende utfordringene som må løses for at en organisme skal kunne forholde seg aktivt til sine omgivelser for å opprettholde seg selv er beskrevet av Deacons teori. I sin mest grunnleggende form er dette de førsteordens funksjonene som knytter seg til metabolsk aktivitet som jeg beskrev i kapittel 3.5.1. Slike funksjoner kan realiseres på samme måte som prinsippforklaringen til Deacons autogener, der forklaringen tar utgangspunkt i de faktiske fysiske egenskapene som kan inngå i disse strukturene.

Spørsmålet er hvordan det er mulig å etablere denne andreordens funksjonen, der de teleologiske prosessene knytter seg til og begynner å motta signaler fra omgivelsene samt fortolke disse. Hvis vi ser på dette som en ren komputasjonell prosess, i tråd med teorien til Gallistel som beskrevet i kapittel 2.3.1, så har vi grunn til å anta de subjektive opplevelsene er «innsiden» som følger parallelt med slike eksterne komputasjonelle strukturer og prosesser, begge realisert av livets teleologiske prosesser. Da får vi den dual-aspekt teorien som skissert i dette kapitlet, hvor vi har grunn til å anta at subjektive opplevelser følger med, og er tett forbundet med, de teleologiske prosessene som inngår i, og danner grunnlaget for, levende organismer.

Selv om det å fokusere på livets teleologiske prosesser kan gi grunn for å anerkjenne en mye tettere forbindelse mellom funksjon og fornemmelse, mellom ytre kapabiliteter og indre opplevelser, så kan man fortsatt hevde at dette er en løs kobling. Det må ikke være slik, det kan sies å være basert på en tilfeldighet ved hvordan vår verden er. Vi har grunn til å anta denne koblingen, basert på det faktum at vi er bevisste, men det *må* ikke være slik. I Chalmers zombie-verden, der disse teleologiske prosessene eksplisitt ikke har en innside, så er det ikke lagt frem en grunn til at en (teleologisk) zombie er umulig. Det kan innvendes mot en slik tenkelig zombie-verden at den rett og slett er empirisk feil; det er mye vi kan tenke oss som ikke er tilfelle. Dette har heller ikke vært ambisjonen å argumentere for at en slik forestilling er umulig, men i stedet å begrunne at det er en *sannsynlig* forbindelse mellom kognitive kapabiliteter og en indre opplevelse. Denne argumentasjonen har gitt oss en grunn til å anerkjenne hvordan vi bør forstå denne forbindelsen ved å redegjøre for hvordan sensoriske og kognitive kapabiliteter kan oppstå, og hvordan det mentale er kausalt integrert med det materielle gjennom livets teleologiske prosesser. Det finnes imidlertid et siste mulig argument som tar oss hele veien og som gir oss en grunn til å tro at en kognitiv zombie rett og slett er en dyp selvmotsigelse; at det er en *nødvendig* forbindelse mellom kognitive kapabiliteter og en indre opplevelse.

Som jeg nevnte i redegjørelsen om en komputasjonell teori om bevissthet, så er det ikke avklart hvilke prosesser som faktisk realiserer komputasjonen; det som realiserer en organismes kognitive evner. Basert på fravær av en slik forklaring, kan følgende idé legges frem, analogt med de kontrafaktiske resonnementene over. *Hadde det ikke vært for at det finnes en «innside» som følger de komputasjonelle prosessene, så ville ikke avanserte komputasjonelle prosesser blitt dannet, og som en konsekvens resulterende kognitive kapabiliteter hos levende organismer.* I fravær av en sentral fortolker eller ekstern designer av det komputasjonelle systemet, så er dette egenskapen som kan bygge et komputasjonelt og representasjonelt system fra bunnen. Organismens subjektive opplevelser er det som kobler kognitive kapabiliteter sammen med omverdenen og den flercellede kroppen til organismen. Uten en slik innside ville ikke avanserte kognitive livsformer vært mulig.

Dette er kun en spekulativ idé som det ikke er nødvendig å anta, men som det er mulig å vurdere. Den kan motiveres som følger. Den har det fortrinnet at en slik indre tilstand ikke forutsetter et annet eller nytt kausalt nivå. Den kognitive funksjonen med tilhørende subjektive opplevelse vil spille sammen slik som det er beskrevet i dette kapittelet. Idéen får imidlertid en forklarende kraft gjennom potensielt å gjøre disse kognitive andreordens funksjonene som er beskrevet i dette kapittelet mye mer sannsynlig og derfor *mulig*; det er bevissthetens *sine qua non*. Dette er derfor et forslag til en empirisk mulighetsbetingelse for

å bygge noe kognitivt fra noe ikke-kognitivt og som kan forklare hvorfor og hvordan vi har en indre opplevelse som følger våre kognitive kapabiliteter. Idéen kan sees på som en mulig forklaring ved at den øker sannsynligheten for hvordan kognitive kapabiliteter kan utvikle seg, den reduserer behovet for en appell til å bare peke på evolusjonens kraft til å utforske mulige utfall.

For at det skal dannes et representasjonelt system mellom en organisme og omgivelsene, så er det et krav fra en komputasjonell teori at det er en form for strukturell likhet. Skal en slik strukturell likhet etableres, må alle mulige variasjoner i en organismes sentralnervesystem utforskes helt til det blir tilstrekkelig kongruens mellom strukturen og det denne representerer. Gitt en slik likhet kan relevante «states of the world» bli formidlet gjennom signaler som fortolkes av organismen. En slik teori kan være egnet til å forklare mange reaktive og umiddelbare prosesser, som for eksempel reflekser. Det kan være mer krevende å få etablert et virkelig representasjonelt system. Hva hvis det er mulig at dette kravet om strukturell likhet reduseres basert på at representasjonen blir realisert av denne innsiden, denne grunnleggende egenskapen ved verden som vi har grunn til å tro følger teleologiske prosesser. En måte å skissere hvordan dette kan forstås er forslaget under, som er en omskriving av Gallistels krav til et representasjonelt system, som angitt i kapittel 2.3.1. Som et tillegg kan vi anta Dennetts idé om hvordan nevronale strukturer kan aktivt representere det som er av interesse for organismen, dens «affordances». Det tas videre utgangspunkt i den grunnleggende aktive og målrettede relasjonen til omgivelsene som er en konsekvens av organismens egenskaper.

- 1 Receptivity: The mapping from entities in the represented system to their symbols (structures used in computation) in the representing system requires a receptiveness based on the sensory capacity of the organism (as, for example, when light reflected off an object in the world stimulates the sensory receptors in an eye resulting in neural signals that eventuate in a percept of the object through a determined activity of a set of neurons)
- 2 : Representation: The mapping from entities in the represented system to their symbols is such it is represented in neuronal activity in the organism through its internal sense of the «affordances» of the organism. This determines what signal to react to, based on the information in the signal available to the sensory receptors and the overall internal state of the organism, again represented through the configuration of the nervous system,
- 3 Response: Symbolic operations (procedures realised through the activities in a neuronal structure) in the representing systems are (at least sometimes) behaviorally efficacious through the metabolism and surplus energy of the organism: they control and direct appropriate behavior within, or with respect to, the represented system.

Denne mulige alternative forståelsen av representasjonen reduserer behovet for at variasjoner som må testes ut blindt av evolusjonen, siden representasjonen kan ivaretas av disse subjektive opplevelsene, og ikke kun gjennom en strukturell likhet mellom organismen og omgivelsene. Følgelig kan de komputasjonelle strukturene i de materielle nevronene realiseres på mange måter, gjennom forskjellige strukturer og nevronale nettverk, på samme måte som forskjellige dyrs kroppsfasong kan tilpasses til et liv i havet basert på helt forskjellige utgangspunkt. En slik idé forutsetter på ingen måte at en hendelse må «gjøres tilgjengelig for en bevissthet» for «beslutning» for videre hendelser- det er i stedet den indre opplevelsen som gjør en eller annen form for fysisk realisert komputasjon for levende organismer mulig, der den indre opplevelsen ikke har noen ytterligere kausal kraft utover å gjøre etableringen av kognisjon mulig, med påfølgende utvikling av det funksjonelle kausale nivået som er et spesialtilfelle av livets grunnleggende teleologiske prosesser.

Her er da en idé som kan utforskes videre som et mulig svar på Chalmers «hard problem». Det er mulig den bommer på målet og er feil. Er den det, står vi igjen med en sannsynlig forbindelse mellom våre kognitive kapabiliteter og tilhørende indre opplevelse, noe som har vært den primære ambisjonen i denne oppgaven. Jeg ønsket å legge frem dette siste alternativet som en idé som kan utforskes. Det kan hevdes at denne oppgaven burde begynt med en slik idé og utforsket den, men tatt ut av løse luften ville forslaget ikke vært forståelig. Det å legge frem et perspektiv på liv og bevissthet slik at den kan vurderes basert på et egnet teoretisk rammeverk er en forutsetning for et mulig videre arbeid med idéen. Når alt dette er sagt, så gjenstår dette mysteriet: hvorfor det er slik at det eksisterer en innside sammen med den eksterne fysiske materien.

4: Avslutning

Hvorfor kan vi se våre omgivelser? Fordi dette gir et evolusjonært fortrinn for organismer som må navigere omgivelsene. Dette er en egenskap en komplisert flercellet organisme trenger for å overleve i en fiendtlig verden, og der hvor min indre opplevelse inngår som en del av denne evnen. Gjennom denne analysen har jeg argumentert for et perspektiv som jeg mener setter Chalmers "hard problem" i et nytt lys. Jeg har argumentert for at styrken i dette problemet ble basert på en mer eller mindre eksplisitt materialistisk og reduksjonistisk forståelse av problemet. Hva er sammenhengen mellom meg som et objekt med visse funksjoner og jeg som et subjekt med visse indre opplevde fornemmelser?

Når vi heller forstår dette basert på redegjørelsen om liv og bevissthet, som jeg har gjort i denne oppgaven ser vi at begge deler kan sies å være et uttrykk for en grunnleggende teleologisk prosess som er dannet av evolusjonen og opprettholdt av de autopoietiske livsprosessene som er beskrevet. Problemet til Chalmers får dermed en helt annen innretning. Det som gjenstår, er spørsmålet om hvorfor det er slik at det følger med disse indre opplevelsene sammen med disse livsprosessene. For å svare på dette måtte vi over i spekulative betraktninger. Disse betraktningene peker imidlertid på at det er mye som taler for at den indre opplevelsen som følger med våre kognitive kapabiliteter, enten er unik for mennesket, eller så er det basert på et grunnlag som har vært med oss fra livets begynnelse. Jeg har argumentert for at det har vært med oss fra livets begynnelse. Som en siste mulighet satt jeg frem en idé som faktisk gir en begrunnelse for hvorfor og hvordan det følger en subjektiv opplevelse med våre kognitive ferdigheter.

Motivasjonen for denne oppgaven har vært å komme nærmere et svar på «the hard problem». Som jeg skrev i innledningen så tror jeg ikke et direkte svar er mulig å begrunne, men å spekulere i, men jeg mener redegjørelsen i denne oppgaven belyser sentrale elementer som gjør at vi kan redusere det som fremstår som den mest uønskede konsekvensen av Chalmers problem: Denne distansen mellom mine indre opplevelser og jeg som en fysisk skapning. Jeg mener at vi kan ha mindre grunn til bekymring om dette temaet.

Selve «the hard problem» og Chalmers argumenter for dette har ikke vært direkte behandlet i denne oppgaven. I stedet har jeg i hovedsak redegjort for et alternativt perspektiv for hvordan vi skal forstå levende prosesser, der det sentrale grepet var å se på livet, ikke som en biologisk maskin, mekanistisk, men som grunnleggende levende; teleologiske og aktivt

utførende drevet frem av en særskilt kontroll av energi vi ikke finner i andre fysiske systemer. Et sentralt poeng har imidlertid vært å vise sammenhengen fra den enkleste organisme til de fullverdige kognitive kapabilitetene vi mennesker har. Denne analysen har vist at disse kan forklares gjennom de samme prosessene der kognisjon og en rik bevissthet er et spesialtilfelle og som er et resultat av en evolusjonær prosess som har ført oss mennesker til der vi står i dag i vår evolusjonære historie.

Et slikt perspektiv bryter med deler av det som kan ansees som en for sterk trend innen populariserte og forenklete fremstillinger av dagens naturvitenskap som fokuserer på at alt egentlig bare er materie i bevegelse, at reduktive forklaringsmodeller er prinsipielt mulig. Hvor mange som forfekter et slik syn kan debatteres, men jeg vil påstå at det ligger som et implisitt perspektiv hos både en del filosofer og vitenskapsmenn. Det er et slik syn Thomas Nagel argumenterer kraftig mot i sin siste bok; *Mind & Cosmos: Why the Materialist Neo-Darwinian Conception of Nature is Almost Certainly False*. Tittelen sier det meste. Her ser vi også et poeng med å benytte Chalmers «hard problem» som en inngang til denne redegjørelsen. Anta at dette er et genuint og viktig problem. Da vil det synes at dette problemet er en grunnleggende anomali i en tilnærming til vår kunnskap om verden når man forutsetter både materialisme og reduksjonisme. Ved å redegjøre for et alternativt perspektiv der brodden i «the hard problem» blir mye svakere, om ikke helt borte, så har vi et slags meta-argument for det perspektivet jeg argumenterer for. Dette kan gi en motivasjon til å virkelig vurdere om dette alternativet har noe for seg; at Nagel hadde rett i sin kritikk av en slik overforenkling av det metafysiske grunnlag for naturvitenskapen.

Hvis vi nå godtar dette som et faktum, endrer dette noe på hvordan vi skal forholde oss til verden? Hvis denne redegjørelsen hadde hatt konkrete konsekvenser som hadde gått mot fysikk, biolog og informasjonsvitenskap så ville denne redegjørelsen falt død til jorden. Dette er vitenskapsgrener som er så utrolig godt fundert og som det er meningsløst å kritisere fra utsiden uten dyp kunnskap om dem som en aktiv utøver. Jeg mener at den alternative redegjørelsen som er presentert ikke har noen slike konsekvenser. Det som endrer seg er kun vår forståelse av disse fenomenene, det metafysiske bakteppet som er de filosofiske implikasjonene for fortolkningen av disse teoriene. Når det gjelder en slik fortolkning så er det virkelig rom for en ny forståelse om oss selv, bevissthet, naturen og det materielle grunnlaget for dette. Det er med fortvilelse man ser hvordan naturen blir behandlet av mennesker- enorme ødeleggelse og massivt misbruk. En filosofisk oppgave vil ikke endre noe, men hvis vi kan se bedre hvor fundamentalt livet er og hvor lite spesielle vi mennesker egentlig er i den store sammenhengen, så kan man håpe at vi behandler livets fantastiske resultater med mer respekt og omtanke.

Bibliografi

Allen, Colin and Trestman, Michael, "Animal Consciousness", *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2017/entries/consciousness-animal/>.

Alter, Torin and Pereboom, Derk, "Russellian Monism", *The Stanford Encyclopedia of Philosophy* (Fall 2019 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2019/entries/russellian-monism/>.

Baluška, František, and Michael Levin. "On Having No Head: Cognition throughout Biological Systems." *Frontiers in psychology* vol. 7 902. (21 Jun. 2016), doi:10.3389/fpsyg.2016.00902

Barrett, Adam B. *An integration of integrated information theory with fundamental physics*. Front. Psychol., (04 February 2014) | <https://doi.org/10.3389/fpsyg.2014.00063>

Bird, Graham, ed. *A companion to Kant*. Oxford: Blackwell, (2006).

Blackmore, Susan. *State of the Art – The Psychology of Consciousness*. The Psychologist, vol.14 p.522-525 (2001)

Boghossian, Paul, and Christopher Peacocke, eds. *New essays on the a priori*. Clarendon Press, (2000).

Carroll, Sean. *The big picture: on the origins of life, meaning, and the universe itself*. Penguin, (2017).

Chalmers, David J: "Is the Hard Problem of Consciousness Universal?" *Journal of Consciousness Studies*, (2020) (hentet fra <http://consc.net/consciousness/>)

Chalmers, David J. "Idealism and the Mind-Body Problem". *Panpsychism*. The Routledge Handbook to Panpsychism. Routledge, (2018) (hentet fra <http://consc.net/consciousness/>)

Chalmers, David J. "The Combination Problem for Panpsychism". *Panpsychism*. Oxford University Press, (2017). (hentet fra <http://consc.net/consciousness/>)

Chalmers, David J. "Consciousness and its Place in Nature." *Blackwell Guide to the Philosophy of Mind*. Blackwell, (2003) (hentet fra <http://consc.net/consciousness/>)

Chalmers, David J. "What is a neural correlate of consciousness." *Neural correlates of consciousness: Empirical and conceptual questions* (2000): 17-40. (hentet fra <http://consc.net/consciousness/>)

Chalmers, David J. *Panpsychism and Panprotopsychism*. Amherst Lecture in Philosophy, (2013) (hentet fra <http://consc.net/consciousness/>)

Chalmers, David J. "The problem of Consciousness". In (H. Jasper, L. Descarries, V. Castellucci, & S. Rossignol, eds) *Consciousness: At the Frontiers of Neuroscience (Advances in Neurology, Vol. 77)*. Lippincott-Raven Press, 1998.

Chalmers, David J. «Moving forward on the Problem of Consciousness». *Journal of Consciousness Studies* vol. 4, pp. 3-46, (1997) (hentet fra <http://consc.net/consciousness/>)

Chalmers, David J. *The conscious mind: In search of a fundamental theory*. Oxford university press, (1996).

Chalmers, David J. "Facing up to the problem of consciousness." *Journal of consciousness studies* 2.3 (1995): 200-219. (hentet fra <http://consc.net/consciousness/>)

Cole, David, "The Chinese Room Argument", *The Stanford Encyclopedia of Philosophy* (Spring 2019 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2019/entries/chinese-room/>

Collins, Arthur. *Possible experience: understanding Kant's Critique of pure reason*. Univ of California Press, (1999).

Darwin, Charles. *On the origin of species by means of natural selection*. (1859) (hentet fra <http://darwin-online.org.uk/contents.html>).

Dawkins, Richard,. *The Selfish Gene*. Oxford ; New York :Oxford University Press, (1989).

Deacon, Terrence W. *Incomplete nature: How mind emerged from matter*. WW Norton & Company, (2011).

Dennett, Daniel C. *Consciousness explained*. Penguin uk, (1993).

Dennett, Daniel C. *Aching Voids and Making Voids A review of Incomplete Nature: How Mind Emerged from Matter by Terrence W. Deacon* *The Quarterly Review of Biology*, Vol. 88, No. 4 (December 2013), pp. 321-324

Dennett, Daniel C. *Intuition pumps and other tools for thinking*. WW Norton & Company, (2013).

Dennett, Daniel C. *From bacteria to Bach and back: The evolution of minds*. WW Norton & Company, (2017).

Freeman, Anthony. *Consciousness and its place in nature: does physicalism entail panpsychism?* Imprint Academic. (2006).

Gallistel, Charles R., and Adam Philip King. *Memory and the computational brain: Why cognitive science will transform neuroscience*. Vol. 6. John Wiley & Sons. Kindle edition, (2011).

Giulio Tononi, Scholarpedia, 10(1):4164. (2015)
(http://www.scholarpedia.org/article/Integrated_information_theory)

Giulio Tononi, Melanie Boly, Marcello Massimini & Christof Koch. "Integrated information theory: from consciousness to its physical substrate" *Nature Reviews Neuroscience volume 17*, pages 450–461 (2016)

Goff, Philip, Seager, William and Allen-Hermanson, Sean, "Panpsychism", *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/win2017/entries/panpsychism/>>.

Godfrey-Smith, Peter. "Evolving Across the Explanatory Gap." *Philosophy, Theory, and Practice in Biology* 11 (2019). (Hentet fra <https://petergodfreysmith.com/publications>)

Godfrey-Smith, Peter: "Materialism, then and now" (2018a) On webpage: <https://petergodfreysmith.com/wp-content/uploads/2020/07/Armstrong-and-Materialism-PGS-2020-K.pdf>

Godfrey-Smith, Peter: Varieties of Subjectivity. Forthcoming in Philosophy of Science (PSA 2018b symposia volume) (Hentet fra <https://petergodfreysmith.com/publications>)

Godfrey-Smith, Peter. "Individuality, subjectivity, and minimal cognition." *Biology & Philosophy* 31.6 (2016a): 775-796. (Hentet fra <https://petergodfreysmith.com/publications>)

Godfrey-Smith, Peter: "Mind, Matter, and Metabolism," *Journal of Philosophy* 113 (2016b): 481-506. (Hentet fra <https://petergodfreysmith.com/publications>)

Godfrey-Smith, Peter. *Other minds: The octopus and the Evolution of Intelligent Life*. London: William Collins, (2016c).

Godfrey-Smith, Peter.: "Triviality Arguments Against Functionalism," *Philosophical Studies* 145 (2009): 273-295 (Hentet fra <https://petergodfreysmith.com/publications>)

Godfrey-Smith, Peter.: "Innateness and Genetic Information," in P. Carruthers, S. Lawrence, and S. Stich (eds.), *The Innate Mind, Volume 3: Foundations and the Future*. Oxford University Press, (2007), pp. 55-68. (Hentet fra <https://petergodfreysmith.com/publications>)

Godfrey-Smith, Peter and Sterelny, Kim, "Biological Information", *The Stanford Encyclopedia of Philosophy* (Summer 2016a Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/sum2016/entries/information-biological/>>.

Henning, Brian G and Scarfe, Adam: *Beyond Mechanism: Putting Life Back Into Biology*, Lexington Books, (2013).

Hume, David. *A treatise of human nature*. Penguin Classics, (1985).

Kant, Immanuel, *Critique of pure reason*. Cambridge University Press, (2000).

Kant, Immanuel, *Critique of Judgment*. Translated with Introduction and Notes by J.H. Bernard (2nd ed. revised). London: Macmillan, (1914). 31.8.2020. <<https://oll.libertyfund.org/titles/1217>>

Kuhn, Thomas S. *The structure of scientific revolutions. Second edition, Enlarged* University of Chicago press, (1970).

Jolley, Nicholas, ed. *The Cambridge Companion to Leibniz*. Cambridge University Press, (1995).

Levin, Janet, "Functionalism", *The Stanford Encyclopedia of Philosophy* (Fall 2018 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2018/entries/functionalism/>.

Levine, Joseph. "Materialism and qualia: The explanatory gap." *Pacific philosophical quarterly* 64.4 (1983): 354-361.

Luisi, Pier L. *Life and consciousness. A simultaneous origin?* Wall Street International Magazine 23. mars (2017) (<https://wsimag.com/science-and-technology/24425-life-and-consciousness>)

McFaddon, Johnjoe: "The Conscious Electromagnetic Information (Cemi) Field Theory" *Journal of Consciousness Studies*, 9, No. 8, , pp. 45–60. (2002)

McLaughlin, Brian and Bennett, Karen, "Supervenience", *The Stanford Encyclopedia of Philosophy* (Winter 2018 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2018/entries/supervenience/>.

Mørch, Hedda Hassel. "Panpsychism and Causation: A New Argument and a Solution to the Combination Problem (Doctoral Dissertation)." *Department of Philosophy, Classics, History of Art and Ideas, University of Oslo, Oslo* (2014).

Mørch, Hedda Hassel. "The evolutionary argument for phenomenal powers." (2018).

Nagel, Thomas. "What is it like to be a bat?." *The philosophical review* 83.4 (1974): 435-450.

Nagel, Thomas. *Mind and cosmos: why the materialist neo-Darwinian conception of nature is almost certainly false*. Oxford University Press, (2012).

Neander, Karen. «The teleological notion of 'function'», *Australasian Journal of Philosophy*, 69:4, 454-468, (1991)

Ramsey, William, "Eliminative Materialism", *The Stanford Encyclopedia of Philosophy* (Summer 2020 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2020/entries/materialism-eliminative/>.

Ravenscroft, Ian, "Folk Psychology as a Theory", *The Stanford Encyclopedia of Philosophy* (Summer 2019 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2019/entries/folkpsych-theory/>

Schrödinger, Erwin. *What is life? The physical aspect of the living cell and mind*. Cambridge: Cambridge University Press, (1944).

Searle, John R. "Minds, brains, and programs." *Behavioral and brain sciences* 3.3 (1980): 417-424.

Seibt, Johanna, "Process Philosophy", *The Stanford Encyclopedia of Philosophy* (Winter 2018 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2018/entries/process-philosophy/>.

Sherman, Jeremy, *Neither Ghost nor Machine: The Emergence and Nature of Selves*. Columbia University Press, (2017). Kindle edition.

Stang, Nicholas F., "Kant's Transcendental Idealism", *The Stanford Encyclopedia of Philosophy* (Winter 2018 Edition), Edward N. Zalta (ed.), URL = [<https://plato.stanford.edu/archives/win2018/entries/kant-transcendental-idealism/>](https://plato.stanford.edu/archives/win2018/entries/kant-transcendental-idealism/).

Stoljar, Daniel, "Physicalism", *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), Edward N. Zalta (ed.), URL = [<https://plato.stanford.edu/archives/win2017/entries/physicalism/>](https://plato.stanford.edu/archives/win2017/entries/physicalism/).

Strawson, Galen. "The consciousness myth (revised)." *The Times Literary Supplement* 5839 (2015).

Strawson, Galen. "Consciousness isn't a mystery. It's matter." *New York Times* 16.5 (2016): 16.

Stubenberg, Leopold, "Neutral Monism", *The Stanford Encyclopedia of Philosophy* (Fall 2018 Edition), Edward N. Zalta (ed.), URL = [<https://plato.stanford.edu/archives/fall2018/entries/neutral-monism/>](https://plato.stanford.edu/archives/fall2018/entries/neutral-monism/).

Thompson, Evan. *Mind in life*. Harvard University Press, (2010).

Van Gulick, Robert, "Consciousness", *The Stanford Encyclopedia of Philosophy* (Spring 2018 Edition), Edward N. Zalta (ed.), URL = [<https://plato.stanford.edu/archives/spr2018/entries/consciousness/>](https://plato.stanford.edu/archives/spr2018/entries/consciousness/).

Weber, Andreas, and Francisco J. Varela. "Life after Kant: Natural purposes and the autopoietic foundations of biological individuality." *Phenomenology and the cognitive sciences* 1.2 (2002): 97-125.

Weber, Bruce, "Life", *The Stanford Encyclopedia of Philosophy* (Summer 2018 Edition), Edward N. Zalta (ed.), URL = [<https://plato.stanford.edu/archives/sum2018/entries/life/>](https://plato.stanford.edu/archives/sum2018/entries/life/).

Wohlleben, Peter. *The Hidden Life of Trees*. Greystone books, (2016)

Wu, Wayne, "The Neuroscience of Consciousness", *The Stanford Encyclopedia of Philosophy* (Winter 2018 Edition), Edward N. Zalta (ed.), URL = [<https://plato.stanford.edu/archives/win2018/entries/consciousness-neuroscience/>](https://plato.stanford.edu/archives/win2018/entries/consciousness-neuroscience/)