

Froese, Tom

Tchnąć nowe życie w kognitywistykę = Breathing new life into cognitive science

Avant 2/1, 95-129

2011

Artykuł został opracowany do udostępnienia w internecie przez Muzeum Historii Polski w ramach prac podejmowanych na rzecz zapewnienia otwartego, powszechnego i trwałego dostępu do polskiego dorobku naukowego i kulturalnego. Artykuł jest umieszczony w kolekcji cyfrowej bazhum.muzhp.pl, gromadzącej zawartość polskich czasopism humanistycznych i społecznych.

Tekst jest udostępniony do wykorzystania w ramach
dozwolonego użytku.

Tchnąć nowe życie w kognitywistykę

Tom Froese

przekład: Piotr Momot

Streszczenie

W artykule tym opowiadam się za zunifikowaną kognitywistyką, przyjmując dla swej argumentacji niecodzienny punkt wyjścia: stanowisko określane czasem jako „teza o kontinuum życia-umysłu”. Zamiast więc traktować jako pewnik powszechnie akceptowane założenia początkowe, a następnie proponować odpowiedzi na pewne dobrze określone pytania, muszę najpierw dowieść, że koncepcja kontinuum życia-umysłu może w ogóle stanowić właściwy punkt startowy. Zacznę zatem od oceny pojęciowych narzędzi, odpowiednich do budowania teorii umysłu na tej podstawie. Czerpiąc spostrzeżenia z wielu różnych dziedzin – szczególnie z połączenia egzystencjalistycznej fenomenologii ze skoncentrowaną na organizmie biologią – dowodzę, że *można pojmować umysł jako zakorzeniony w życiu, ale tylko wtedy, gdy równocześnie zgodzimy się, że interakcja społeczna gra konstytutywną rolę w naszych zdolnościach poznawczych.*

Wstęp

Komputacjonistyczna teoria umysłu jest bez wątpienia jednym z największych naukowych sukcesów dwudziestego wieku. Pojawiła się na tle dominującego behawiorystycznego założenia, które odrzuciło wszelkie badania nad wewnętrznymi działaniami umysłu jako nienaukowe. Mimo tego, udało jej się przekonująco dowieść, że jest wprost przeciwnie: komputacjonizm sprawił, że po raz pierwszy można było pomyśleć o zjawiskach umysłowych jako o potencjalnym przedmiocie nowoczesnych badań naukowych. Pokazał ponadto, że możliwe jest rozwinięcie systematycznego programu badawczego w wyniku zjednoczenia różnorodnych dyscyplin pod jedną wspólną hipotezą, w tym przypadku koncepcją głoszącą, że poznanie jest formą oblicza-

nia. Sztuczka polegała na tym, że rozkwitające pole technologii komputerowych uznano nie jedynie za narzędzie kognitywistyki, a za pełnoprawny obiekt jej dociekań. Daleka od bycia mglistą i nienaukową, komputacjonistyczna teoria umysłu osiągnęła przewagę nad poprzednimi teoriami psychologicznymi dzięki mocnemu ugruntowaniu na formalnych zasadach badań nad sztuczną inteligencją.

Gdy tylko jednak zaczęto powszechniej akceptować założenie, że umysł jest wartościowym przedmiotem poważnych badań naukowych, głosy sprzeciwu wobec fundamentów komputacjonistycznej teorii umysłu stały się coraz lepiej słyszalne. Począwszy od Dreyfusowskiego ataku na symboliczną Sztuczną Inteligencję w latach siedemdziesiątych, za którym podążył koneksjonistyczny ruch lat osiemdziesiątych, kontynuowany przez akcentujące dynamizm i usytuowanie stanowiska w robotyce lat dziewięćdziesiątych – pierwotnie wąskie pole kognitywistyki zaczęło się rozrastać. Obecnie uwagę przyciągają liczne alternatywne wobec komputacjonistycznej kognitywistyki koncepcje, nierzadko łączone przez badaczy w jedną, luźno spójną konstrukcję. Nawet wśród funkcjonalistów stało się modne opisywanie poznania jako ucieleśnionego, zakorzenionego, rozszerzonego i enaktywnego. Moim jednak zdaniem logiczne wnioski z tego postępu wyciąga enaktywistyczna koncepcja zaproponowana przez Francisco Varełę i jego współpracowników¹, traktująca umysł jako fundamentalnie nierozdzielny od istnienia naszego subiektywnego doświadczenia, jego biologicznego ucieleśnienia i usytuowania w społeczno-kulturowym świecie.

Nie powinniśmy jednak zapominać, że również właśnie dzięki historii komputacjonistycznej kognitywistyki zdajemy sobie obecnie dużo lepiej sprawę z tego, w jaki sposób umysł w swej istocie *nie* jest taki, jak komputer. Weźmy jako przykład mój laptop: potrafi przetwarzać dane wejściowe, przechowywać informacje i obliczać dane wyjściowe, a jednak zdarzenia te nie mają dla niego absolutnie żadnego znaczenia. Mówiąc wprost, mój laptop nie rozumie kompletnie niczego, nie jest nawet w stanie przejąć się swoim brakiem zrozumienia. Ta podstawowa kwestia przewija się co jakiś czas w filozofii umysłu i kognitywistyce. Obecnie dyskutowana jest pod wieloma różnymi postaciami, na przykład problemu zdrowego rozsądku, problemu nabierania znaczeń przez symbole, problemu ram odniesienia, trudnego problemu świadomości i – być może najogólniej – problemu znaczenia. Gdyby jednak komputacjonistyczna kognitywistyka nie poniosła porażki w opisie tego aspektu umysłu, nie mielibyśmy tak jasnego pojęcia, do jakiego stopnia umysł jest przede wszystkim istotną częścią naszej egzystencji.

To prawda, że być może wyjaśnienie nurtującego problemu znaczenia uda się ostatecznie osiągnąć przy użyciu odpowiednich komputacjonistycznych algorytmów, na przykład swego rodzaju „metapoznania”. Istnieje jednak ekscytująca możliwość, na którą wskazuje rosnąca popularność alternatywnych wobec kognitywistyki koncepcji,

¹ Np. Varela i in. (1991); Thompson (2007); Stewart, Gapenne i Di Paolo (red.) (2010).

że w celu poczynienia postępów w tej kwestii niezbędna jest bardziej fundamentalna zmiana perspektywy. W dalszej części artykułu chciałbym naszkicować projekt tego, jak taka zmiana mogłaby wyglądać.

Krok 1: Rozwiązując problem umysłu-ciała

Pomyślne ugruntowanie się dziedziny kognitywistyki zależało w dużej mierze od jej zdolności do zjednoczenia psychologii i neuronauki pod patronatem nadrzędnej konstrukcji – informatyki. Jedną z głównych zasług tej konstelacji dyscyplin polegała na tym, że oferowała ona naukowo funkcjonalne rozwiązanie problemu umysłu-ciała: rozróżnienie na to, co personalne i subpersonalne, ujęte zostało w kategoriach rozróżnienia na software i hardware.

Jeśli nasze alternatywne podejście do kognitywistyki ma stać się poważnym kandydatem do roli nowej unifikującej teorii umysłu, musi zarazem być w stanie w satysfakcjonujący sposób zająć się problemem relacji umysłu do ciała. Co ciekawe, w koncepcjach enaktywistycznych rośnie świadomość tego, że problem ten jest zarazem bardziej złożony i bardziej prosty niż tradycyjnie przypuszczano. Po pierwsze, musimy przyznać, że na tak zwanym „poziomie personalnym” możemy rozróżnić kilka różnych perspektyw, przede wszystkim pierwszoosobową (doświadczenie subiektywne), drugoosobową (dialogiczna interakcja Ja-Ty) i trzecioosobową (obiektywna ocena)².

Ponadto, jak dotąd kognitywistyka korzystała z prostego binarnego rozróżnienia na świadome i nieświadome procesy umysłowe (tzn. to, co „poznawczo świadome/nieświadome”). Wnikliwe fenomenologiczne i psychoanalityczne badania wykazały jednak, że musimy uwzględnić przynajmniej potrójne rozróżnienie poziomów psychologicznego opisu, a mianowicie doświadczenie refleksyjne (świadomość przechodnia), doświadczenie przedrefleksyjne (świadomość nieprzechodnia) i procesy nieświadome (brak bezpośredniej świadomości)³. Oczywiście, istnieje jeszcze fizyczny poziom opisu z perspektywy trzecioosobowej.

W jaki sposób ta bardziej zróżnicowana koncepcja ludzkiej egzystencji może pomóc nam w rozstrzygnięciu problemu umysłu-ciała? Proponowane przez nią rozwiązanie skoncentrowane jest na poziomie przedrefleksyjnym, który wcześniej ukryty był za absolutyzmem rozróżnienia na umysł i materię. Faktycznie, jesteśmy zdolni do spędzania znacznych części naszego życia w sposób refleksyjny, jako myśliciele, ale dokładne badanie naszego doświadczenia pokazuje, że momenty te są zaledwie chwilowymi wypadkami. Gdzie egzystujemy przez resztę czasu? Odpowiedź wypracowana przez tradycję fenomenologiczną i przeniesiona przez podejście enaktywistyczne na grunt kognitywistyki jest taka, że zawsze jesteśmy praktycznie zaangażowani w nasze życie. Jest to domena doświadczenia, wypełniona celowymi dążeniami, od-

² Zob. np. Zahavi (2005); Petitmengin (2006); Stawarska (2009).

³ Zob. np. Zahavi (2006); Gallagher i Zahavi (2008, Rozdział 3); Fuchs (w druku).

czuwanymi przepływami czasu, niejasnymi nastrojami, nadawaniem sensu, cielesnymi odczuciami i tak dalej. W skrócie: zjawisko życia odsyła nas do konstytucji naszej „przeżywanej” (od niemieckiego *erlebte*) egzystencji, która jest praktycznie ucieleśniona i znacząco usytuowana w świecie.

Wiemy jednak, że ten doświadczeniowy czy fenomenologiczny aspekt życia nie wystarcza do jego kompletnego opisu. Nowożytna nauka przeszła długą drogę, żeby odkryć wiele złożonych procesów materialnych zachodzących w żyjącym ciele: cykle metaboliczne, przewodzenie energii, cykle wzrostu i obumierania, syntezę chemiczną, aktywność neuronalną i inne. Zauważmy jednak, że w tym przypadku istnieje przynajmniej korzyść z poczucia konceptualnej zgody: zarówno fenomenologia, jak i biologia zainteresowane są badaniem życia. Co ważniejsze, z bezpośredniego doświadczenia wiemy ponadto, że istnieje także jedność ontologiczna. Innymi słowy, moje żyjące ciało istnieje nie tylko jako zewnętrzny przedmiot nauk biologicznych; również Ja istnieję jako to ciało. Doświadczam tego, że *jestem* tym żyjącym ciałem, poprzez moje praktyczne zaangażowanie w świat. Staje się to dla nas jasne w trakcie przerażającej wycieczki do lekarza.

Ten osobisty wgląd, który każdy może sam zweryfikować, inspiruje jedno z podstawowych założeń podejścia enaktywistycznego, a mianowicie, że *fizyczne przejawy mojego ciała jako żyjącego nie mogą być oddzielane od doświadczeniowych przejawów mojego ciała jako przeżywanego*. Zgodnie z tym poglądem, w przeciwieństwie nie tylko do współczesnych wariantów Kartezjańskiego dualizmu ciała i umysłu, ale również i do pozostałych ucieleśnionych koncepcji w kognitywistyce, umysł jest nie tylko ucieleśniony w żyjącym ciele; żyjące ciało jest również „umysłowione” przez przeżywany umysł. Ten splot życia biologicznego i umysłowego jest być może najbardziej radykalną konsekwencją tezy o kontinuum życia-umysłu.

Ważny krok w kierunku rozwiązania tradycyjnego problemu umysłu-ciała może być zatem osiągnięty dzięki zestrojeniu pojęciowych i fenomenologicznych soczewek, których używamy do badania tego zagadnienia. Bezwzględna luka eksplanacyjna może zostać zrelatywizowana poprzez rozważenie zjawiska życia z perspektywy „problemu ciało-ciało”⁴. Co prawda, przeorientowanie to stanowić może zaledwie początek nowej nauki o umyśle. Wczesna kognitywistyka i Sztuczna Inteligencja odniosły taki sukces, ponieważ podjęły wyzwanie i zamieniły filozofię funkcjonalizmu na działający program badań naukowych w kategoriach komputacjonistycznej teorii umysłu. Co może odegrać podobną rolę w tym przypadku? Wydawać by się mogło, że biologia jako nauka o życiu nadaje się do tego zadania idealnie. Niestety, nie jest to do końca prawdą. Podczas gdy potrzebujemy teorii żyjącego ciała jako takiego, to jest teorii organizmu, okazuje się, że współczesna biologia jest przede wszystkim pochłonięta tym, co organizmem nie jest. Z jednej strony koncentruje się na badaniu

⁴ Hanna i Thompson (2003).

procesów z poziomu suborganizmowego, szczególnie związanych z genetyką, a z drugiej strony skupia się na poziomie procesów superorganizmowych, głównie Darwinowskiej ewolucji. Wydaje się zatem, że rozwój nowej nauki o umyśle nie może zostać osiągnięty bez równoczesnego rozwoju nowej nauki o życiu.

Aby wypełnić tę lukę we współczesnej biologii, zaczerpnąć możemy z – rozwijanych w związku z cybernetyką – teorii żyjących organizmów. Jedno z ich głównych założeń głosiło, że żyjącą istotę można ujmować jako autopojetyczny układ, co w istocie oznacza, że jest to układ zorganizowany w taki sposób, że jednym z rezultatów jego czynności jest kontynuacja jego istnienia⁵. Klasycznym przykładem układu autopojetycznego jest organizm jednokomórkowy: gdyby nie zorganizowane istnienie komórki jako całości, procesy metaboliczne nie byłyby kontynuowane i szybko zaniknęłyby. Jednocześnie brak kontynuacji procesów metabolicznych oznaczałby koniec istnienia komórki.

Współzależności takie odnajdziemy na wielu różnych poziomach biologicznego opisu. Jedną z nich jest relacja między metaboliczną siecią komórki a jej składnikami molekularnymi. Dzięki relacji tej sieć wykonana jest z tych samych komponentów, które wytwarza jej własna aktywność. Innym znanym przykładem jest współzależność między błoną komórkową a wewnętrznymi sieciami metabolicznymi, polegająca na tym, że błona reguluje niezbędną wymianę materii i energii oraz ochrania sieci metaboliczne przed rozproszeniem i rozplynięciem się w środowisku, podczas gdy te produkują molekularne komponenty, z których błona się składa. Kolejne współzależności pojawiały się wraz z postępem ewolucji, ale podstawowa idea pozostaje ta sama we wszystkich przypadkach – organizm pojmowany jest jako autopojetyczna sieć procesów.

Istnieje kilka powodów, dla których koncepcja autopojezy jest pomocna w rozwiązywaniu problemu ciało-ciało. Po pierwsze gwarantuje ona, że pojęcia organizmu nie stosuje się wyłącznie jako wygodnej etykiety do określania tego, co w ostateczności chce się zredukować do sumy wyizolowanych części. Możemy być pewni, że organizm istnieje sam w sobie, a nie tylko jako wytwór teoretyków – kiedy rozpoznajemy go jako układ autopojetyczny, dostrzegamy, że posiada *wewnętrzne powiązanie* ze swoją własną tożsamością. Innymi słowy, jest tym, czym jest, z powodu tego, co robi, i robi to, co robi, ponieważ jest taki, jaki jest. Wyrażając to inaczej, w kategoriach systemowych: organizm może zostać skonceptualizowany jako układ autonomiczny, czyli samozachowująca się, i co ważniejsze, samowytwarzająca się – a zatem samowyodrębniająca się – sieć procesów.

Po drugie, istnieją przesłanki zmuszające nas do przyjęcia, że owo wewnętrzne powiązanie organizmu służy jako punkt odniesienia, ze względu na który wewnętrzne

⁵ Zob. np. Maturana i Varela (1987); Varela (1997); Froese i Stewart (2010).

i zewnętrzne zdarzenia mogą jawić się mu jako znaczące. Po trzecie, ponieważ sama egzystencja organizmu jest nieustającym zyskiwaniem w obliczu postępującego rozpadu i prawdopodobnej śmierci, bez przesady można stwierdzić, że to właśnie ta niepewna sytuacja wyposaża organizm w perspektywę zatroskania, zaangażowania. Zgodnie z tym punktem widzenia, egzystencjalne (metaboliczne) przetrwanie jest wartością fundamentalną. Zauważmy też, że w ten sposób porzuciliśmy funkcjonalistyczne podstawy tradycyjnej kognitywistyki: śmierć, jako ustanie wszystkich funkcji, nie może być bez popadania w paradoks postrzegana jako kolejna funkcja⁶.

Rozważania te pokazują możliwość przechodzenia w badaniach od ciała żyjącego do przeżywanego, co zresztą zrobiliśmy powyżej. Nie będąc dokładnie tym samym, dwie strony problemu ciało-ciało nie są też czymś całkowicie sobie obcym.

Kolejną zaletą autopoietycznej koncepcji organizmu jest jej bliski związek z matematyką teorii układów dynamicznych, neutralną wobec ontologicznego statusu modelowanych zjawisk. Zarówno biologiczne, jak i fenomenologiczne zdarzenia rozwijają się w czasie i to właśnie ich czasowa struktura może służyć za formalny pomost pomiędzy nimi. W proponowanej tu alternatywnej unifikacji potencjalna teoria systemów dynamicznych odgrywa rolę podobną do tej, którą w tradycyjnej kognitywistyce odegrała informatyka. Istnieje jednak szereg powodów, aby preferować tę alternatywną: (i) teoria systemów dynamicznych dostarcza bardziej ogólnych matematycznych ram (nawet komputer jest swego rodzaju układem dynamicznym, chociaż szczególnego rodzaju), (ii) pozwala nam na lepszą ocenę ciągłych czasowych zmian zjawisk w różnych skalach czasowych, i (iii) jest już standardowym matematycznym językiem nauk przyrodniczych. Ponadto, co szczególnie ważne w odniesieniu do kognitywistyki, może udaremnić formułowanie wyjaśnień błędnie nadających psychologiczny charakter zjawiskom na nieodpowiednich poziomach opisu, takich jak przedrefleksyjny, metaboliczny czy neuronalny⁷.

Krok 2: Wypełniając lukę poznawczą: teoria

W poprzedniej części zaprezentowałem elementy zunifikowanej na nowo kognitywistyki, jakkolwiek taka interdyscyplinarna konstelacja alternatywnych stanowisk jest całkowicie różna od tej tradycyjnej. Zamiast oscylowania pomiędzy psychologią poznawczą, neuronauką poznawczą i informatyką, wymaga ona rozwoju nowej współpracy między egzystencjalistyczną fenomenologią, skoncentrowaną na organizmie biologią i teorią układów dynamicznych. Zgodnie z takim punktem widzenia pionierska Varełowska propozycja neurofenomenologii⁸ nie jest już wystarczająca; to, czego potrzebujemy, to kompleksowa *bio-fenomenologia*.

⁶ Zob. np. Jonas (1966); Weber i Varela (2002); Thompson (2004); Di Paolo (2009).

⁷ Zob. np. Port i van Gelder (red.) (1995); Roy i in. (1999); Lutz i Thompson (2003).

⁸ Varela (1996).

Podstawowe dyscypliny tradycyjnej kognitywistyki będą oczywiście nadal odgrywały ważną rolę. Psychologia jest niezbędna do dobrego projektowania eksperymentów, neuronauka dysponuje narzędziami potrzebnymi do badania zdarzeń z poziomu subpersonalnego, a wykorzystanie informatyki pozostaje koniecznością na gruncie robotyki czy teorii systemowych w praktyce opracowywania modeli, służących jako robocze dowody danych koncepcji. Jednak podczas gdy metody tych dyscyplin pozostają istotne dla stanowiska enaktywistycznego, sposób ich zastosowania zmienił się dość znacząco.

Przecież nie trzeba każdego koła wymyślać od nowa. Na polu badań nad sztucznym życiem istnieje przykładowo długa tradycja modelowania autopojezy, a studia nad minimalnym poznaniem, bazujące na teorii układów dynamicznych, prowadzone są już od kilku dekad⁹. To, co stanowisko enaktywistyczne wnosi do istniejących już tradycji, to wyrażenie w sposób bardziej sprecyzowany wzajemnie inspirujących się powiązań pomiędzy obszarami zainteresowania owych dziedzin, a także pozostałych dyscyplin, niezależnie od tego, czy są one już częścią kognitywistyki, czy też nie. W mojej własnej pracy badawczej interesuję się przede wszystkim tym, w jaki sposób obserwacje zaczerpnięte z egzystencjalistycznej fenomenologii, skoncentrowanej na organizmie biologii i programowania agentowego, mogą zostać ujęte w relację umożliwiającą wzajemne oświecanie się tych pól badawczych¹⁰.

Rozpoczynając kognitywistyczne badania w nowym punkcie wyjścia, musimy na samym początku zmierzyć się z poważnym wyzwaniem: w jaki sposób spostrzeżenia płynące z rozwijania tezy o kontinuum życia-umysłu mogą stać się bardziej relewantne względem dociekań charakterystycznych dla tradycyjnej kognitywistyki? Pytając inaczej: jak badania nad adaptacyjną egzystencją prostych organizmów mogą zastąpić badania nad abstrakcyjnym poznaniem u istot ludzkich? Określiłem ten problem mianem „luki poznawczej”¹¹, w nawiązaniu do luki eksplanacyjnej, z którą zmagala się teoria funkcjonalistyczna. Należy jednak od razu podkreślić, że mamy tu do czynienia z dwiema zasadniczo różnymi rodzajami luk: luka metafizycznego stosunku zastąpiona została luką stopni złożoności. Musimy teraz udowodnić, że problemem luki poznawczej można się zająć w sposób systematyczny. Innymi słowy: podczas gdy poznawcza luka tezy o kontinuum życia-umysłu dotyczy spójnego punktu wyjścia, który wciąż wymaga dalszych prac, luka eksplanacyjna problemu umysła odnosi się do zupełnego braku punktu wyjścia.

Fakt, że luka poznawcza rzeczywiście może zostać przewyciężona, przejawia się w naturze w dwóch skalach czasowych. W odniesieniu do czasu historycznego mowa oczywiście o ewolucji współczesnego *Homo sapiens* z protokomórkowych załączków. Niestety olbrzymia rozległość czasowa tego historycznego kontinuum życia-

⁹ Beer (2003); McMullin (2004).

¹⁰ Froese i Ziemke (2009); Froese i Gallagher (2010). Zob. też: Rohde (2010).

¹¹ Froese (2009); Froese i Di Paolo (2009); De Jaegher i Froese (2009).

umysłu sprawia, że wykorzystanie zasad ewolucjonizmu w celu bezpośredniego zastosowania naszej wiedzy dotyczącej poziomu życia organizmu do poziomu ludzkiego poznania jest czymś niezwykle trudnym. Większe nadzieje należy wiązać zatem z kontinuum życia-umysłu, które śledzić możemy w bardziej osiągalnej skali naszego osobniczego życia – mianowicie z rozwojem istoty ludzkiej, począwszy od jej jednokomórkowych zaczątków¹².

Jak jednak pokazuje współczesny spór „*natura versus kultura*”, zjawisko rozwoju poznawczego nadal pozostaje poważnym problemem naukowym. Nie będziemy się tu wdawać w szczegóły owej dychotomii¹³. Dla naszych celów wystarczy, jeśli przyjmiemy, że w wielu przypadkach niezwykle trudno – jeśli w ogóle jest to możliwe – pojąć, w jaki sposób *wyzolowana* jednostka może nauczyć się złożonych zachowań, bez uprzedniego posiadania równie złożonych wrodzonych zdolności poznawczych (najbardziej eksponowanym przykładem w tradycyjnej kognitywistyce była kwestia nabywania języka i hipoteza wrodzonej gramatyki). Zarazem wiemy jednak, że możliwa jest inkulturacja małp człekokształtnych, takich jak przedstawiciele obu gatunków szympansa, poprzez wychowywanie ich w naszym kontekście kulturowym¹⁴. Wiemy też, że kiedy młody *Homo sapiens* pozbawiony jest odpowiedniego kontekstu społeczno-kulturowego, nie rozwija zaawansowanych zdolności poznawczych, które zwykliśmy kojarzyć z gatunkiem ludzkim. Innymi słowy: nie można zaprzeczyć, że specyficznie ludzkie zdolności poznawcze zależą w dużym stopniu od naszego zaangażowania w innych. Jeśli zatem chcemy zmniejszyć lukę poznawczą, musimy uwzględnić rolę wymiaru społecznego.

W tradycyjnej kognitywistyce dawno już zauważono, że poszukiwanie sposobu, w jaki dochodzimy do zrozumienia innych, jest ważnym przedmiotem badań naukowych. Od samego początku funkcjonalnej konstrukcji pojęciowej w postaci teorii informacji dostarczała tu informatyka. W teorii tej komunikacja jest zasadniczo pojmowana jako przekazywanie informacji kanałem od nadawcy do odbiorcy. Schemat ten wielu wydawał się atrakcyjny, ponieważ pasował zgrabnie do komputacjonistycznej teorii umysłu. Stosowany był w wielu obszarach kognitywistyki, stał się ponadto standardowym schematem głównego nurtu biologii¹⁵.

Biorąc jednak pod uwagę bliskie pokrewieństwo komputacjonistycznej teorii umysłu oraz tak sformułowanej teorii komunikacji, nietrudno było przewidzieć, że ta druga będzie musiała zmierzyć się z problemami, z którymi borykała się ta pierwsza. I rzeczywiście, obszar psychologii poświęcony badaniom nad poznaniem społecznym doznał do problemu innych umysłów – martwego punktu, analogicznego do problemu

¹² Dodatkową zaletą takiego podejścia jest możliwość łatwego wyrażenie wielu rozwojowych praw przy pomocy schematów teorii układów dynamicznych (np. Thelen i Smith 1994; Smith i Thelen 2003).

¹³ Zob. Oyama (2000).

¹⁴ Savage-Rumbaugh i in. (1998).

¹⁵ Ale zob. też np. Maturana i Varela (1987, Rozdział 8); Di Paolo (1997); Shanker i King (2002).

umysłu-ciała w komputacjonizmie. Zagadnienie to można opisać następująco: skoro umysł jest czymś odrębnym od ciała i skoro otrzymujemy wyłącznie sygnały wysyłane przez zewnętrzne ciała innych – w jaki sposób możemy dowiedzieć się czegokolwiek o ich wewnętrznych umysłach? Nasza wiedza ogranicza się bowiem jedynie do tego – taki wniosek zmuszeni byli przyjąć komputacjonistyczni kognitywiści – że wszyscy poza nami mogą być pozbawionymi umysłu zombie, a my nigdy nie zdołamy nawet poznać różnicy.

W odniesieniu do poznawczego rozwoju człowieka oznacza to, że trudno wyobrazić sobie, w jaki sposób relacje z innymi mogłyby być przydatne, jeżeli niemowlę nie posiadałoby już wrodzonych narzędzi poznawczych, umożliwiających mu zdobycie wiedzy (drogą teoretycznego wnioskowania i/lub symulacji), że bezrozumne automaty poruszające się wokół niego mogą być hipotetycznie takimi samymi intencjonalnymi istotami jak ono. Problem polega jednak na tym, że nawet jeśli niemowlę faktycznie obdarzone jest od urodzenia niektórymi z najbardziej zaawansowanych umiejętności rozumowania, to zgodnie z komputacjonizmem nadal zmuszone jest mierzyć się z prawie niemożliwym do wykonania zadaniem rozszyfrowaniem tego, co musi jawić mu się jako całkowicie arbitralne dźwięki, widoki i ruchy.

Rozważania te są bardzo istotne dla naszego obecnego celu, ponieważ raczej trudno jest wypełnić lukę poznawczą odwołując się do relacji społecznych, kiedy przed naszym rozumieniem relacji z innymi wyrasta problem innych umysłów. W dalszej części tekstu przedstawimy więc rozwiązanie tej kwestii. Będzie ono analogiczne do rozwiązania problemu umysłu-ciała.

Krok 3. Rozwiązując problem innych umysłów

Jak już wcześniej wspomniano, wnikliwe analizy fenomenologiczne wykazały, że kategoryczne rozróżnienie na perspektywę pierwszoosobową (przeżywanie siebie jako „ja”) i trzecioosobową (doświadczenie innego jako „to”), musi zostać odrzucone jako nieadekwatne. Lekceważy ono fakt, że doświadczamy siebie nawzajem również z perspektywy drugoosobowej (doświadczenie innego jako „ty”). Istnieją oczywiście przykre patologiczne przypadki, w których ludzie mają ograniczoną możliwość przyjęcia tej perspektywy (lub nie mają jej w ogóle), na przykład w przebiegu autyzmu czy schizofrenii. W normalnych jednak okolicznościach nie doświadczamy innego człowieka jako zwykle „to”, chociaż może stać się tak w wyniku naszego wyboru. Prawdopodobnie takiego właśnie wyboru dokonuje żołnierz zamierzający zabić swojego wroga, czy też chirurg operujący ciało pacjenta.

W zwyczajnych warunkach, wchodząc w interakcje z moimi bliźnimi, nie postrzegam ich w kategoriach ciał działających na zasadzie zombie, teoretycznie tylko obdarzonych umysłami. Przeciwnie, postrzegam w sposób bezpośredni, że inni wokół mnie są samoistnymi żyjącymi i umysłowymi (living-minded) istotami, oraz że postrzegają

mnie w ten sam sposób. Takie bezpośrednie percepcyjne uświadomienie sobie przeżywanego ucieleśnienia innych poprzedza wszelką teoretyczną refleksję¹⁶.

Powyższa fenomenologiczna analiza ma dwie ważne konsekwencje dla naszych obecnych rozważań. Po pierwsze okazuje się, że większość dotychczasowych teorii poznania społecznego ograniczała się do zbyt wąskiego obrazu społecznego świata. Skupiały się one na poznawczych mechanizmach służących do nabywania refleksyjnej wiedzy o innych, zaniedbując potencjalną rolę poznania o charakterze bezpośrednim, przedrefleksyjnym i intuicyjnym. Kiedy jednak spotykam kogoś, nie muszę teoretycznie zakładać możliwości jego umysłowej egzystencji; w ruchu jego ciała bezpośrednio dostrzegam jego przeżywaną obecność. W tradycji fenomenologicznej ten rodzaj bezpośredniej percepcji został opisany w kategoriach empatii i sympatii. Zauważmy, że kiedy już uznamy istnienie przedrefleksyjnej perspektywy drugoosobowej, nie musimy borykać się z problemem innych umysłów, przynajmniej nie w odniesieniu do poziomu refleksyjnego, stanowiącego tradycyjny obiekt badań psychologii poznawczej. Odsyła nas to do drugiej konsekwencji: musimy się teraz zmierzyć z wyzwaniem, jakie stawia przed nami próba wyjaśnienia podłoża naszego przedrefleksyjnego rozumienia innych.

Obiecującą metodą wydaje się być koncentracja na istnieniu wspólnej podstawy zrozumienia, swego rodzaju wspólnego zmysłu. Różne warianty tej myśli zaproponowane zostały na gruncie egzystencjalistycznej fenomenologii. W kontekście tym mówi się tam zazwyczaj, że podczas interakcji, na przedrefleksyjnym poziomie przeżywanego ciała nie istnieje zdecydowane rozróżnienie na „siebie” i „innego”. Podobne koncepcje były jednak na ogół wyrażane filozoficznie, bez poważniejszych odniesień do współczesnej nauki. Tradycyjnie przedstawiany problem innych umysłów jest z kolei ściśle związany z naukami kognitywnymi, ale za to zazwyczaj ujmowany w kategoriach wyjaśnienia możliwości obustronnego epistemicznego dostępu dwóch metafizycznie wyizolowanych umysłów. Na szczęście teza o kontinuum życia-umysłu otwiera kolejną obiecującą możliwość. Jeśli zaakceptujemy założenie, że żyjące ciało jest zarazem ciałem przeżywanym i jeśli pokażemy, że znajdujące się w interakcji ciała stają się do pewnego stopnia jednym ciałem, będziemy mogli stwierdzić, że odnaleźliśmy podstawę bezpośredniego wzajemnego udziału w umysłowości. Jest to jedna z centralnych koncepcji mojej pracy, określam ją mianem *hipotezy rozszerzonego ciała*.

Nietrudno wykazać, że koncepcja żyjącego organizmu skupiona na pojęciu układu autopojetycznego może zostać, przynajmniej w teoretycznym wymiarze, rozszerzona w odpowiedni sposób. Spora część obecnej dyskusji w ramach stanowiska enaktywistycznego jest w istocie poświęcona właśnie temu zagadnieniu¹⁷.

¹⁶ Zahavi (2001); Stanghellini (2004); Gallagher (2008); Stawarska (2009).

¹⁷ De Jaegher i Di Paolo (2007); Froese i Di Paolo (2009); De Jaegher i Froese (2009).

Przede wszystkim należy podkreślić, że o prawdziwym układzie autopojetycznym nie można myśleć jako o wyizolowanym ze środowiska. I tak na przykład błona komórkowa ochrania komórkę nie przez zwykłe izolowanie, ale raczej poprzez adaptacyjne regulowanie interakcji ze środowiskiem. Po to, by przetrwać, komórka musi oczywiście utrzymywać swą odrębność i zabezpieczać się przed awersyjnymi wpływami otoczenia. Musi być jednak zarazem na tyle otwarta, żeby angażować środowisko w elastyczny sposób i odnajdywać składniki odżywcze niezbędne do podtrzymania metabolizmu. Zgodnie z takim punktem widzenia komórka i środowisko nie są dwoma wyizolowanymi układami, ale są wewnętrznie powiązane w asymetryczny sposób (istnienie pierwszego zależy od drugiego). W efekcie życie nie jest jakąś niezależną substancją, a relacyjnym procesem¹⁸.

Co więcej, w relacji tej środowisko nie ogranicza się wyłącznie do pierwiastków chemicznych, musimy uwzględnić również istnienie innych organizmów. W przypadku komórki możemy założyć, że jej metabolizm zależy bezpośrednio właśnie od ich produktów. Jeżeli ponadto któryś z tych organizmów zależy od substancji wydalanych przez ową komórkę, mamy do czynienia z kolejnym cyklem współzależności, wiążącym te dwa organizmy w większą autonomiczną całość: rozszerzone żyjące ciało. W pewnych przypadkach taka nowa autonomia tworzy złożoną, spójną strukturę, tak że możemy już mówić o organizmie wielokomórkowym. Zastosowanie autopojetycznej teorii organizmu może zostać zatem poszerzone¹⁹. Zauważmy przy okazji, że jeżeli żyjące ciało i przeżywany umysł są zjawiskami komplementarnymi, to mamy częściową odpowiedź na pytanie, jak to jest możliwe, że jesteśmy zbudowani z pojedynczych żyjących ciał (komórek), a mimo to przeżywamy świat z jednej zunifikowanej perspektywy.

Organizmy wielokomórkowe oraz przykłady procesów powstawania metabolicznych współzależności służyć mogą jako ogólne potwierdzenie koncepcji rozszerzonego ciała, nie pomogą nam jednak bezpośrednio w odnalezieniu potencjalnego wspólnego podłoża naszego przedrefleksyjnego zrozumienia innych. Zastanówmy się, czy podobną rolę mogłyby odegrać interakcje społeczne. I znów, potrzebną teoretyczną bazę odnajdziemy w skoncentrowanej na organizmie biologii. Zamiast skupiać się na cyklach współzależności z poziomu chemicznego (takich jak autopojetyczne samowytwarzanie), zwróćmy naszą uwagę na autonomiczne cykle manifestujące się w sposobie, w jaki organizm odnosi się do swojego środowiska. W przypadku zwierząt znajdziemy jasny przykład tego, że to, co zwierzę robi, zależy od tego, czego doznaje, a to, czego doznaje, zależy od tego, co robi²⁰.

¹⁸ Ruiz-Mirazo i Moreno (2004); Di Paolo (2005); Barandiaran i Moreno (2009); Barandiaran i in. (2009); Virgo i in. (w druku).

¹⁹ Maturana i Varela (1987); Froese i Di Paolo (w druku).

²⁰ Ze wszystkich diskutowanych w tym artykule koncepcji, współzależność percepcji i działania stała się najbardziej popularnym przedmiotem badań alternatywnych stanowisk w kognitywistyce. Zob. np. Varela i in. (1991); O'Regan i Noë (2001); Noë (2004, 2009); Barandiaran i Moreno (2009); Thompson (2005).

Więcej: skoro żyjące ciało jest również ciałem umysłu, to każdy jego ruch jest dopełniany nie tylko przez zmianę w percepcji zmysłowej, ale również przez zmianę w ciele przeżywanym. Życie, czucie, działanie są za pośrednictwem środowiska i ciała połączone w jedną sensomotoryczną pętlę. Zgodnie z tym, funkcjonujące w tradycyjnej kognitywistyce bezwzględne rozróżnienie na percepcję, poznanie i działanie (najbardziej znane ze schematycznej ścieżki: odbiór bodźców – odwzorowywanie – planowanie – działanie) okazuje się arbitralne. Stanowisko enaktywistyczne, bardziej niż którekolwiek z alternatywnych koncepcji w kognitywistyce, próbowało oddać sprawiedliwość współzależnej naturze życia, percepcji zmysłowej i działania (a także bardziej ogólnie – doświadczenia i ruchu) poprzez zastosowanie w sposób dosłowny pojęcia *wytwarzania sensu (sense-making)*²¹.

Ponieważ wytwarzanie sensu bazuje na sensomotorycznej pętli organizmu i jako takie zapośredniczone jest przez środowisko, które z kolei obejmuje również inne organizmy z ich własnymi sensomotorycznymi pętlami, istnieje możliwość zajęcia interakcji między aktywnością wytwarzania sensu różnych osobników. Może się ponadto zdarzyć, że intensywność wzajemnego reagowania na siebie organizmów osiągnie taki stopień, że ich – rozszerzone w wyniku tego – sensomotoryczne pętle zaczną tworzyć większy autonomiczny cykl. Zwykłe wytwarzanie sensu zaczyna przybierać tu inny charakter: zachowanie jednego organizmu prowadzi do zmiany w doznaniach drugiego, co skutkuje zmianą jego zachowania, która z kolei zmienia odczucia oraz zachowanie pierwszego organizmu i tak dalej. W ten sposób, modulując wzajemnie swoje sensomotoryczne pętle, modulują zarazem wzajemnie swoją aktywność wytwarzania sensu. Są zaangażowane we *współwytwarzanie sensu (participatory sense-making)*²².

Pojęcie współwytwarzania sensu dostarcza nam teoretycznej bazy niezbędnej do wyjaśnienia podstawy naszego przedrefleksyjnego zrozumienia innych. Kiedy wzajemna interakcja dwojga ludzi zamienia się w samoistny, autonomiczny cykl, ich ciała stają się na pewien czas jednym rozszerzonym ciałem, dzięki dynamicznej mediacji indywidualnych pętli sensomotorycznych. Należy jeszcze raz podkreślić, że zgodnie z podstawowymi założeniami tezy o kontinuum życia-umysłu, *rozszerzone żyjące ciało implikuje rozszerzone ciało przeżywane*. Oznacza to, że w perspektywie drugoosobowej, kiedy nasze ciała stają się dynamicznie związane, możemy wzajemnie partycypować w naszym doświadczeniu.

Podsumowując: pokazaliśmy, że zyskując wiedzę dotyczącą sposobu, w jaki doświadczamy innych w trakcie interakcji, zauważamy, że przynajmniej na poziomie egzystencjalnym problem innych umysłów nie istnieje: bezpośrednio postrzegam innych

²¹ Weber i Varela (2002); Thompson (2004); Di Paolo (2005); Sheets-Johnstone (1999).

²² Pojęcie to zostało wprowadzone przez De Jaegher i Di Paolo (2007) i wzbudziło żywą dyskusję w kręgach enaktywistycznych teoretyków; zob. np. Steiner i Stewart (2009); Gallagher (2009); Torrance i Froese (2011).

jako żyjące i umysłowe (*living-minded*) istoty, widzę też, że one postrzegają mnie w ten sam sposób. Potrafimy ponadto dać spójne wyjaśnienie możliwości dzielenia doświadczenia z innymi, odwołujące się do tezy o kontinuum życia-umysłu i hipotezy rozszerzonego ciała.

Nadal wyjaśnienia wymaga sposób, w jaki zasady rządzące współwytwarzaniem sensu mogą zostać zastosowane do opisanego refleksyjnych poziomów ludzkiego poznania społecznego, nie wykluczając zdolności do teoretyzowania na temat innych umysłów. Co jednak ważniejsze, problem metafizycznego stosunku zastąpiliśmy problemem stopni złożoności i musimy teraz zmierzyć się po prostu z pewną wersją luki poznawczej.

Krok 4: Zmniejszając ponownie lukę poznawczą: praktyka

Czy jednak taka określona droga przez zagadnienie innych umysłów doprowadzi nas do rozwiązania problemu luki poznawczej, czyniąc relacje z innymi zasadniczym elementem rozwoju poznawczego? Jeżeli relacje z innymi również są ograniczone przez lukę poznawczą, wydawać się może, że trafiliśmy w kolejną ślepą uliczkę.

Obawa ta świadczy jednak o niedocenieniu roli, jaką w naszych relacjach z innymi odgrywa przedrefleksyjna, drugoosobowa interakcja, a także o lekceważeniu potencjału procesu interakcji w autonomicznym organizowaniu naszych indywidualnych zdolności. Unikniemy ślepej uliczki, jeżeli zdołamy wykazać, że zachowanie niemowląt może być stosownie regulowane przez autonomiczną dynamikę procesów interakcyjnych, oraz że wewnętrzna budowa ciała niemowlęcia może być spontanicznie konfigurowana tak, by jak najbardziej elastycznie i adaptacyjnie podtrzymywać interakcję. Musimy zatem pokazać, że tak długo, jak niemowlę posiada zdolność do wzajemnego zestrzajania się z innymi w odpowiedni sposób, może dzięki interakcji względnie automatycznie rozwinać bardziej zaawansowane zdolności. Jeżeli niezbędna wrażliwość na innych może zostać również osiągnięta w samej interakcji, cały problem luki poznawczej można rozłożyć na relacyjną matrycę rozciągającą się na osobę, innych i resztę świata.

Czym innym jest oczywiście pokazanie, że scenariusz taki jest teoretycznie możliwy, a całkowicie czym innym wykazanie, że jest możliwy w rzeczywistości. Biorąc pod uwagę wielką zmianę w punkcie wyjścia, mającą znaczący wpływ na dalszy rozwój naszego scenariusza, należy w pierwszym kroku zademonstrować, że coś takiego jest w ogóle możliwe. Skłoniło mnie to do zaprojektowania szeregu agentowych modeli interakcji dowodzących, że:

- Wrażliwość jednostki na innych może być skutkiem interakcji z innymi
- Ruch jednostki może być organizowany przez interakcje z innymi

- Doznania jednostki mogą być organizowane przez interakcje z innymi
- Ciało jednostki może być organizowane przez interakcje z innymi

Bardziej szczegółowy opis wyników tego modelowania, a także jego analiza odwołująca się do układów dynamicznych, wykraczają poza zakres tego artykułu. Zadanie to wykonałem gdzie indziej²³. Dla naszej bieżącej dyskusji ważne jest jednak, że modele te są realnym potwierdzeniem koncepcji, że problemem luki poznawczej można się metodycznie zająć poprzez naturalne dopuszczenie w badaniach dynamiki społecznej interakcji. Kolejnym krokiem musi być kontynuacja tych wyjściowych badań, polegająca na przeprowadzeniu rzeczywistych eksperymentów psychologicznych. Można wiązać duże nadzieje z podjętymi już w tym kierunku pracami²⁴.

Wnioski

Podsumowując: wniosek wypływający z powyższych rozważań jest taki, że komputacjonistyczna teoria umysłu może zostać zastąpiona przez alternatywny model skupiony na pojęciu kontinuum życia-umysłu, pod warunkiem jednak, że uwzględnimy konstytutywną rolę wymiaru społecznego. W artykule tym dowodziłem, że owo alternatywne podejście do kwestii zjednoczenia kognitywistyki ma nad innymi stanowiskami przewagę, której nie jesteśmy w stanie zlekceważyć. Jest funkcjonalnym programem badań naukowych, wolnym od metafizycznych łamigłówek, oferującym większą zgodność z ludzkim doświadczeniem i jednocześnie posiadającym solidne podstawy matematyczne.

Bibliografia

- Auvray, M., Lenay, C. i Stewart, J. 2009. Perceptual interactions in a minimalist virtual environment. *New Ideas in Psychology*, 27(1): 32-47.
- Barandiaran, X., Di Paolo, E. A. i Rohde, M. 2009. Defining Agency: Individuality, Normativity, Asymmetry, and Spatio-temporality in Action. *Adaptive Behavior*, 17(5): 367-386.
- Barandiaran, X. i Moreno, A. 2006. On what makes certain dynamical systems cognitive: A minimally cognitive organization program. *Adaptive Behavior*, 14(2): 171-185.
- Beer, R. D. 2003. The dynamics of active categorical perception in an evolved model agent. *Adaptive Behavior*, 11(4): 209-243.
- De Jaegher, H. i Di Paolo, E. A. 2007. Participatory sense-making: An enactive approach to social cognition. *Phenomenology and the Cognitive Sciences*, 6(4): 485-507.

²³ Jak na razie, najbardziej szczegółowe omówienie tych zagadnień znajduje się w mojej rozprawie doktorskiej (Froese 2009), ale pewne wyniki odnaleźć można również w poszczególnych artykułach (np. Froese i Di Paolo 2008; 2010; w druku).

²⁴ Auvray i in. (2009).

- De Jaegher, H. i Froese, T. 2009. On the Role of Social Interaction in Individual Agency. *Adaptive Behavior*, 17(5): 444-460.
- Di Paolo, E. A. 1997. An investigation into the evolution of communication. *Adaptive Behavior*, 6(2): 285-324.
- Di Paolo, E. A. 2005. Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences*, 4(4): 429-452.
- Di Paolo, E. A. 2009. Extended life, *Topoi*, 28(1): 9-21.
- Froese, T. 2009. *Sociality and the life-mind continuity thesis: A study in evolutionary robotics*, Rozprawa Doktorska, Brighton: University of Sussex.
- Froese, T. i Di Paolo, E. A. 2008. Stability of coordination requires mutuality of interaction in a model of embodied agents: 52-61. Red. Asada, M., Hallam, J. C. T., Meyer, J. - A. i Tani, J. *From Animals to Animats 10: Proc. of the 10th Int. Conf. on Simulation of Adaptive Behavior*. Berlin: Springer-Verlag.
- Froese, T. i Di Paolo, E. A. 2009. Sociality and the life-mind continuity thesis. *Phenomenology and the Cognitive Sciences*, 8(4): 439-463.
- Froese, T. i Di Paolo, E. A. (w druku). The enactive approach: From cell to society. *Pragmatics & Cognition*.
- Froese, T. i Di Paolo, E. A. (w druku). Toward Minimally Social Behavior: Social Psychology Meets Evolutionary Robotics. *Advances in Artificial Life: Proc. of the 10th Euro. Conf. on Artificial Life*. Berlin: Springer-Verlag.
- Froese, T. i Gallagher, S. 2010. Phenomenology and Artificial Life: Toward a Technological Supplementation of Phenomenological Methodology. *Husserl Studies*, 26(2): 83-106.
- Froese, T. i Stewart, J. 2010. Life after Ashby: Ultrastability and the autopoietic foundations of biological individuality. *Cybernetics & Human Knowing*, 17(4): 7-50.
- Froese, T. i Ziemke, T. 2009. Enactive Artificial Intelligence: Investigating the systemic organization of life and mind. *Artificial Intelligence*, 173(3-4): 366-500.
- Fuchs, T. (w druku). Body memory and the unconscious. Red. Lohmar, D. i Brudzinska, J. *Founding Psychoanalysis: Phenomenological Theory of Subjectivity and the Psychoanalytical Experience*. Dordrecht: Kluwer.
- Gallagher, S. 2008. Direct perception in the intersubjective context. *Consciousness and Cognition*, 17(2): 535-543.
- Gallagher, S. 2009. Two problems of intersubjectivity. *Journal of Consciousness Studies*, 16(6-8): 289-308.
- Gallagher, S. i Zahavi, D. 2008. *The Phenomenological Mind: An Introduction to Philosophy of Mind and Cognitive Science*. Londyn: Routledge.

- Hanna, R. i Thompson, E. 2003. The Mind-Body-Body Problem. *Theoria et Historia Scientarum*, 7(1): 24-44.
- Jonas, H. 1966 [2001]. *The Phenomenon of Life: Toward a Philosophical Biology*. Evanston: Northwestern University Press.
- Lutz, A. i Thompson, E. 2003. Neurophenomenology: integrating subjective experience and brain dynamics in the neuroscience of consciousness. *Journal of Consciousness Studies*, 10: 31-52.
- Noë, A. 2004. *Action in Perception*. Cambridge: The MIT Press.
- Noë, A. 2009. *Out of Our Heads: Why You Are Not Your Brain, and Other Lessons from the Biology of Consciousness*. New York: Hill and Wang.
- O'Regan, J.K. i Noë, A. 2001. A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24(5): 939-1031.
- Oyama, S. 2000. *The Ontogeny of Information: Developmental Systems and Evolution*. 2nd edition. Durham: Duke University Press.
- Petitmengin, C. 2006. Describing one's subjective experience in the second person: An interview method for the science of consciousness. *Phenomenology and the Cognitive Sciences*, 5(3-4): 229-269.
- Port, R. F. i van Gelder, T. Red. 1995. *Mind as Motion: Explorations in the Dynamics of Cognition*. Cambridge: MIT Press.
- Rohde, M. 2010. *Enaction, Embodiment, Evolutionary Robotics: Simulation Models for a Post-Cognitivist Science of Mind*. Atlantis Press.
- Roy, J.-M., Petitot, J., Pachoud, B. i Varela, F. J. 1999. Beyond the Gap: An Introduction to Naturalizing Phenomenology: 1-80. Red. Petitot, J., Varela, F. J., Pachoud, B. i Roy, J.-M. *Naturalizing Phenomenology: Issues in Contemporary Phenomenology and Cognitive Science*. Stanford: Stanford University Press.
- Ruiz-Mirazo, K. i Moreno, A. 2004. Basic Autonomy as a Fundamental Step in the Synthesis of Life. *Artificial Life*, 10(3): 235-259.
- Savage-Rumbaugh, S., Shanker, S. G. i Taylor, T. J. 1998. *Apes, Language, and the Human Mind*. New York: Oxford University Press.
- Shanker, S. G. i King, B. J. 2002. The emergence of a new paradigm in ape language research. *Behavioral and Brain Sciences*, 25(5): 605-620.
- Sheets-Johnstone, M. 1999. *The Primacy of Movement*. Amsterdam: John Benjamins Publishing.
- Smith, L. i Thelen, E. 2003. Development as a dynamical system. *Trends in Cognitive Sciences*, 7(8): 343-348.

- Stanghellini, G. 2004. *Disembodied Spirits and Deanimated Bodies: The psychopathology of common sense*. New York: Oxford University Press.
- Stawarska, B. 2009. *Between You and I: Dialogical Phenomenology*. Athens: Ohio University Press.
- Stewart, J., Gapenne, O. i Di Paolo, E. A. Red. 2010. *Enaction: Towards a New Paradigm for Cognitive Science*. Cambridge: The MIT Press.
- Thelen, E. i Smith, L.B. 1994. *A dynamical systems approach to the development of cognition and action*. Cambridge: The MIT Press.
- Thompson, E. 2004. Life and mind: From autopoiesis to neurophenomenology. A tribute to Francisco Varela. *Phenomenology and the Cognitive Sciences*, 3(4): 381-398.
- Thompson, E. 2005. Sensorimotor subjectivity and the enactive approach to experience. *Phenomenology and the Cognitive Sciences*, 4(4): 407-427.
- Thompson, E. 2007. *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Cambridge: The Belknap Press of Harvard University Press.
- Torrance, S. i Froese, T. 2011. An inter-enactive approach to agency: Participatory sense-making, dynamics, and sociality. *Humana.Mente*, 15: 21-53.
- Maturana, H.R., i Varela, F.J. 1987. *The Tree of Knowledge: The Biological Roots of Human Understanding*. Boston: Shambhala Publications.
- McMullin, B. 2004. Thirty Years of Computational Autopoiesis: A Review. *Artificial Life*, 10(3): 277-295.
- Varela, F. J. 1996. Neurophenomenology: A Methodological Remedy for the Hard Problem. *Journal of Consciousness Studies*, 3(4): 330-349.
- Varela, F. J. 1997. Patterns of Life: Intertwining Identity and Cognition. *Brain and Cognition*, 34(1): 72-87.
- Varela, F.J, Thompson, E. i Rosch, E. 1991. *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge: The MIT Press.
- Virgo, N., Egbert, M. i Froese, T. (w druku). The Role of the Spatial Boundary in Autopoiesis. in: *Proc. of the 10th Euro. Conf. on Artificial Life*. Berlin: Springer-Verlag.
- Weber, A. i Varela, F. J. 2002. Life after Kant: Natural purposes and the autopoietic foundations of biological individuality. *Phenomenology and the Cognitive Sciences*, 1: 97-125.
- Zahavi, D. 2001. Beyond Empathy: Phenomenological Approaches to Intersubjectivity. *Journal of Consciousness Studies*, 8(5-7): 151-167.
- Zahavi, D. 2005. *Subjectivity and Selfhood: Investigating the First-Person Perspective*. Cambridge: The MIT Press.
- Zahavi, D. 2006. Two takes on a one-level account of consciousness. *Psyche*, 12(2): 1-9.

Breathing new life into cognitive science

Tom Froese

Abstract

In this article I take an unusual starting point from which to argue for a unified cognitive science, namely a position defined by what is sometimes called the ‘life-mind continuity thesis’. Accordingly, rather than taking a widely accepted starting point for granted and using it in order to propose answers to some well defined questions, I must first establish that the idea of life-mind continuity can amount to a proper starting point at all. To begin with, I therefore assess the conceptual tools which are available to construct a theory of mind on this basis. By drawing on insights from a variety of disciplines, especially from a combination of existential phenomenology and organism-centered biology, I argue that *mind can indeed be conceived as rooted in life, but only if we accept at the same time that social interaction plays a constitutive role for our cognitive capacities.*

Introduction

The computational theory of mind is surely one of the biggest scientific success stories of the 20th century. It first emerged against the backdrop of a behaviorist establishment that rejected all study of the inner workings of the mind as unscientific, and yet it managed to convincingly demonstrate that the opposite was in fact the case. Computationalism made it conceivable for the first time that mental phenomena are a potential target for modern scientific study. In addition, it showed that it is possible to develop a systematic research program by bringing together a variety of disciplines under one shared hypothesis, in this case the idea that cognition is a form of computation. The trick was to conceive of the burgeoning field of computer technology not merely as a means for cognitive science, but as a target for study in its own right as well. Far from being vague and unscientific, the computational theory

of mind had the advantage over previous theories in psychology of being securely founded on the formal principles of artificial intelligence.

However, as soon as the idea that the mind is indeed a valid target for serious scientific study had become more widely accepted, voices of dissent against the foundations of the computational theory of mind began to make themselves heard. Starting with Dreyfus' attack on symbolic AI in the '70s, followed by the connectionist movement of the '80s, and then continuing with the dynamical and situated robotics approach of the '90s, the originally narrow scope of cognitive science has continually been expanded. Today we are in a situation where numerous alternative approaches to computationalist cognitive science are clamoring for attention, and it is not uncommon for researchers to combine them into a loosely knit framework. In fact, even among some of the functionalists it has become fashionable to describe cognition as being embodied, embedded, extended and enactive. On my view, however, it is the enactive approach, in the sense in which it was conceived by Francisco Varela and colleagues¹, that pushes these developments to their logical conclusion, namely by treating mind as fundamentally inseparable from the existence of our subjective experience, its biological embodiment and its situatedness in the socio-cultural world.

Nevertheless, we should not forget that it is also thanks to the history of computationalist cognitive science that we now know much more precisely how the mind is fundamentally *not* like a computer. Take my laptop for instance: it can process inputs, store information and calculate the next outputs, and yet these events have absolutely no meaning for it. To put it simply, it does not understand anything at all, nor can it even care about this lack of understanding. In philosophy of mind and cognitive science this foundational issue makes an appearance every now and again. Today it continues to be debated under various related guises, e.g. the common sense problem, the symbol grounding problem, the frame problem, the hard problem of consciousness, and, most generally perhaps, the problem of meaning. But if it weren't for the failure of computationalist cognitive science to account for this aspect of mind, we would never have had such a clear understanding of how it is an essential part of our existence in the first place.

To be sure, it may be possible that this pervasive problem of how to explain meaning can eventually be solved by using the right kind of computational algorithms, for example by some form of 'meta-cognition'. Yet at the same time it is also an exciting possibility, as indicated by the growing popularity of alternative approaches to cognitive science, that a much more fundamental change in perspective is needed to advance on this issue. In the rest of this article I want to sketch the outlines of what such a change might look like.

¹ E.g. Varela, et al. (1991); Thompson (2007); Stewart, Gapenne and Di Paolo (eds.) (2010).

Step 1: Solving the mind-body problem

The successful establishment of the field of cognitive science depended to a large extent on the field's ability to unite psychology and neuroscience under the umbrella framework of computer science. One of the main contributions of this constellation of disciplines is that it offers a scientifically workable solution to the mind-body problem: the personal / sub-personal distinction is conceived in terms of a software / hardware distinction.

In order for an alternative approach to cognitive science to be a serious contender for a new unifying theory of mind, it must also be able to address the mind-body problem in a satisfactory manner. Interestingly, there is a growing realization in the enactive approach that the mind-body problem is both more complex and simpler than has traditionally been supposed. First, it is important to acknowledge that on the so-called 'personal level' we can actually distinguish between several different perspectives, most crucially between the first-person (subjective experience), the second-person (dialogical I-You interaction), and the third-person (objective measurement) perspective².

Furthermore, so far cognitive science has always been working with a simple binary distinction between conscious and unconscious mental processes (i.e. the 'cognitive unconscious'). But careful phenomenological and psychoanalytical research reveals that we need to consider at least a tertiary distinction on the psychological level of description, namely between reflective experience (transitive awareness), pre-reflective experience (intransitive awareness), and unconscious processes (no immediate awareness)³. And, of course, there is also the physical level of description from the third-person perspective.

How does this more differentiated conception of human existence help us to solve the mind-body problem? The solution is centered on the pre-reflective level which had been previously hidden behind the absolutism of the mind-matter distinction. We may be able to spend significant parts of our existence in the reflective mode of the thinker, yet at the same time careful investigation of our experience shows that these moments are merely temporary excursions. Where do we exist the rest of the time? An answer developed by the phenomenological tradition, and taken up the enactive approach to cognitive science, is that we are always already practically engaged in our lives. This is the experiential realm filled with purposeful striving, temporal flows, ambiguous moods, sense-making, bodily feeling, and so forth. In a word, the phenomenon of life refers to the constitution of our 'lived' (from the German word *erlebte*) existence, which is practically embodied and meaningfully situated in a world.

² See, e.g., Zahavi (2005), Petitmengin (2006) and Stawarska (2009).

³ See, e.g., Zahavi (2006), Gallagher and Zahavi (2008, Chapter 3) and Fuchs (in press).

At the same time we know that this experiential or phenomenological aspect of life does not give us the complete story. Modern science has come a long way to reveal the many intricate material processes at work in the living body. There are metabolic cycles, energy transfers, cycles of growth and decay, chemical synthesis, neural firings, and so forth. In other words, any modern understanding of the phenomenon of life also has to take into account the organization of our biological body. Here we are thus confronted with yet another variant of the famous mind-body problem. But notice that in this case there is at least the advantage of a sense of a conceptual unity: both phenomenology and biology are interested in the study of life. And more importantly, we know from direct experience that there is an ontological unity as well. To put it differently, my living body does not only exist as an external object to biological science; I exist as this body. I experience that I *am* this living body during my practical engagement with the world, as a scary trip to the doctor can easily reveal.

This personal insight, which everyone can verify for themselves, informs one of the core tenets of the enactive approach, namely that *the material manifestation of my body as a living body cannot be separated from the experiential manifestation of my body as a lived body*. On this view, in contrast not only to the modern variants of Cartesian mind-body dualism but to other embodied approaches to cognitive science as well, the mind is not only embodied in a living body; *the living body is also 'minded' by a lived mind*. This intertwining of biological life and mental life is perhaps the most radical implication of the life-mind continuity thesis.

An important step toward dissolving the traditional mind-body problem can therefore be accomplished by fine-tuning the conceptual and phenomenological lenses we use to study the problem. The absolute explanatory gap can be relativized by the phenomenon of life into a 'body-body problem'⁴. To be sure, this refocusing can only be the very beginning of a new science of mind. What made early cognitive science so successful is that AI and computer science stepped up to the challenge and turned the philosophy of functionalism into a workable scientific research program in terms of the computational theory of mind. What could play a similar role in this case? It would seem that biology as the science of life would be naturally suited for this job, yet unfortunately this is not immediately so. While we are in need of a theory of the living body as such, i.e. a theory of the organism, it turns out that modern biology has largely been preoccupied with anything but the organism itself. On the one hand, it has focused on the study of sub-organismic processes, most notably related to genetics, and on the other hand, it has focused on super-organismic processes, especially Darwinian evolution. Unfortunately, it seems that the development of a new science of mind cannot be accomplished without the development of a new science of life as well.

⁴ Hanna and Thompson (2003).

In order to fill this gap in modern biology we can draw on theories of the living organism as they have been developed in relation to cybernetics. One of the central ideas is that a living being can be conceived as an autopoietic system, which essentially means that it is a system that is organized so that one of the results of its operations is the continued existence of that system itself⁵. The classic example of an autopoietic system is a single-cell organism: if it weren't for the organized existence of the cell as a whole, the metabolic processes would not be able to continue and would quickly decay away. And at the same time, if the metabolic processes did not continue to operate, that would spell the end for the cell's existence.

We can find other instances of this co-dependency at several other biological levels of description. For example, we can think of the relationship between the cell's metabolic networks and its molecular components, whereby the networks are made out of the same components which their own activity produces. Another famous example is the co-dependency between the outer cell membrane and the inner metabolic networks, whereby the membrane regulates the necessary material and energetic exchanges and prevents the metabolic networks from diffusing and dissolving into the environment, while the metabolic networks produce the molecular components out of which the membrane is made. Other co-dependencies were added with the major transitions of evolution, but the essential idea remains the same in all cases. An organism is conceived as an autopoietic network of processes.

The concept of autopoiesis is a useful one with regard to the body-body problem for several reasons. First of all, it ensures that the concept of an organism is not merely used as a convenient label for what in the end is a reduction to the mere sum of its isolated parts. We can be certain that an organism exists in its own right, and not merely as a theoretician's product, because when we distinguish it as an autopoietic system we find that it has an *internal relation* to its own identity. In other words, it is what it is because of what it does, and it does what it does because of what it is. Another way of putting this in systemic terms is to say that an organism can be conceptualized as an autonomous system, namely as a self-maintaining, and more importantly, self-producing – and therefore self-distinguishing – network of processes.

Second, there are compelling reasons to assume that this internal relation of the organism serves as a reference point in relation to which internal and external events can show up as meaningful. Third, since the very existence of the organism is an ongoing achievement in the face of continuous decay and possible death, it is not too farfetched to argue that it is precisely this precarious situation which furnishes it with a perspective of concern. On this view, existential (metabolic) survival is regarded as the first of all values. Notice also that in this way we have left the functionalist frame-

⁵ See, e.g., Maturana and Varela (1987), Varela (1997), and Froese and Stewart (2010).

work of traditional cognitive science behind us: death, as the cessation of all function, cannot itself be viewed as another function without falling into paradox⁶.

What these considerations nicely illustrate is that it is possible to work from the living body toward the lived body, just like we examined the living body from the perspective of the lived body. Although they are not immediately the same, the two sides of the body-body problem are not utterly alien to each other, either.

A further advantage of the autopoietic theory of the organism is that it is closely related to the mathematics of dynamical systems theory, which is neutral with regard to the ontological status of the modeled phenomena. Both biological and phenomenological events unfold in time, and this temporal structure can serve as a formal bridge between the two. In this unifying potential dynamical systems theory plays a similar role for this alternative approach as computer science did for traditional cognitive science. And yet there are several reasons for preferring the former: (i) dynamical systems theory provides a more general mathematical framework (even a computer is a kind of dynamical system, although of a peculiar discrete type), (ii) it allows us to do better justice to the continuous temporal changes of the phenomena at multiple timescales, and (iii) it is already the standard mathematical language of the natural sciences. In addition, and this is especially important in relation to cognitive science, it can prevent us from formulating explanations that are mistakenly over-psychologizing phenomena at inappropriate levels of description, such as the pre-reflective, metabolic or neural levels⁷.

Step 2: Bridging the cognitive gap in theory

In the previous section I have introduced the ingredients for a unified cognitive science, although the interdisciplinary constellation of this alternative approach is quite distinct from the traditional one. Instead of triangulating between cognitive psychology, cognitive neuroscience, and computer science, it requires the development of new collaborations between existential phenomenology, organism-centered biology, and dynamical systems theory. On this view, Varela's pioneering proposal of neurophenomenology⁸ is not sufficient; what is needed is a comprehensive *bio-phenomenology*.

Of course, the core disciplines of traditional cognitive science will continue to play an important role. Psychology is indispensable for good experimental design, neuroscience has the tools to investigate events at the sub-personal level, and computer science remains necessary for the practice of devising robotic or systemic models that serve as working proof of concepts. But while the methods of these disciplines

⁶ See, e.g., Jonas (1966), Weber and Varela (2002), Thompson (2004), and Di Paolo (2009).

⁷ See, e.g., Port and van Gelder (eds.) (1995), Roy et al. (1999), and Lutz and Thompson (2003).

⁸ Varela (1996).

continue to be relevant for the enactive approach, the way they are applied is given a different focus.

Similarly, not every wheel needs to be reinvented. For instance, in the field of artificial life there is a long tradition of modeling autopoiesis and the dynamical systems study of minimal cognition has also been ongoing for a couple of decades⁹. What the enactive approach contributes to these existing traditions is to make the mutually informing links between their interests and those of other disciplines, whether they already are a part of cognitive science or not, more explicit. In my own work I have been especially interested in the way in which insights from existential phenomenology, organism-centered biology, and agent-based modeling can be put into a mutually enlightened relationship¹⁰.

But given this new starting point for cognitive science we are immediately confronted by a profound challenge: in what way can the insights of the life-mind continuity thesis be made more relevant to the specific interests of traditional cognitive science? That is, how can the study of adaptive living in the case of simple organisms be scaled up to the study of abstract cognition in the case of human beings? I have labeled this scaling problem, in reference to the explanatory gap faced by the functionalist thesis, the 'cognitive gap'¹¹. However, I immediately hasten to emphasize that we are dealing with two fundamentally different kinds of gaps here: a gap of metaphysical proportions has been replaced by one of levels of complexity. All that needs to be done now is to demonstrate that the cognitive gap can be addressed in a systematic manner. In other words, while the cognitive gap of the life-mind continuity thesis refers to a coherent starting point which still requires further work, the explanatory gap of the mind-body problem refers to the lack of a starting point altogether.

The fact that the cognitive gap can indeed be overcome has been demonstrated by nature on two time scales. In historical time there was of course the evolution of modern *Homo sapiens* from proto-cellular origins. Unfortunately, however, the immense scale of this historical life-mind continuity makes it difficult to use evolutionary principles in order to apply insights gained at the level of organismic life directly to the level of human cognition. A more promising route may be the life-mind continuity which we can witness within the more manageable scale of our lifetimes, namely the development of a human being from its single-cell origins¹².

Yet, as the ever present nature-nurture debate demonstrates, the phenomenon of cognitive development continues to be a deep scientific problem. We will not get into

⁹ Beer (2003); McMullin (2004).

¹⁰ Froese and Ziemke (2009); Froese and Gallagher (2010). See also Rohde (2010).

¹¹ Froese (2009); Froese and Di Paolo (2009); De Jaegher and Froese (2009).

¹² An added advantage of this approach is that many developmental principles are readily expressed in the framework of dynamical systems theory (e.g. Thelen and Smith 1994; Smith and Thelen 2003).

the details of the nature-nurture dichotomy here¹³. For our purposes it is sufficient to accept that in many cases it is indeed difficult if not impossible to conceive how an *isolated* individual could learn complex behaviors without already being endowed with complex innate cognitive capacities (the most prominent example in traditional cognitive science being the case of language acquisition and the hypothesis of innate grammar). And yet at the same time we know that it is possible to enculturate non-human apes such as bonobos and chimpanzees by raising them in our social context¹⁴. And we further know that when a young *Homo sapiens* happens to be deprived of an appropriate socio-cultural context, it will sadly fail to develop the advanced cognitive capacities we have come to associate with the human species. In other words, it cannot be denied that specifically human cognition depends to a large extent on our involvement with others. If we want to bridge the cognitive gap we therefore need to take into account the role of sociality.

Traditional cognitive science has long recognized that a theory of the way in which we make sense of others is an important research goal. Right from the start computer science has provided a workable conceptual framework in the form of information theory. This theory, communication is essentially conceived as the transfer of information over a channel between a sender and a receiver, continues to be attractive to many since it fits neatly with the computational theory of mind. Indeed, the theory has been applied in many areas of cognitive science and has become the standard framework in mainstream biology. We will not enter into the debate about the validity of this theory here¹⁵.

However, given the close relationship between the computational theory of mind and the information theory of communication, we can already expect that the latter will be faced by related problems. And indeed the field of social cognition in psychology has been haunted by an impasse analogous to the mind-body problem, namely the problem of other minds. This problem can be described as follows: given that the mind is distinct from the body, and given that we only receive the signals sent by the external bodies of others, then how do we ever come to know anything about the internal minds of others? For all we know, the computationalist cognitive scientists are forced to conclude, everyone else could be mindless zombies and we would never know the difference.

What this means in relation to the development of human cognition is that it is difficult to conceptualize how relations with others could be helpful unless the infant has already been innately provided with the cognitive tools to know (by theoretical inference and/or simulation) that the mindless automata moving around him could hypo-

¹³ See Oyama (2000).

¹⁴ Savage-Rumbaugh et al. (1998).

¹⁵ But see, e.g., Maturana and Varela (1987, Chapter 8), Di Paolo (1997), and Shanker and King (2002).

thetically be intentional beings like himself. But the problem is that even if the infant was innately endowed with some of the best reasoning abilities, according to the theory of the computationalists it would still be faced with the almost impossible task to unravel the meaning of what must look like entirely arbitrary sounds, sights and movements.

These considerations are important for our current purpose because it is rather difficult to bridge the cognitive gap by appealing to social relations when our understanding of the relations to others is confronted by the problem of other minds. In what follows we will therefore develop a solution to the problem of other minds, and this will be done in an analogous fashion to the dissolution of the mind-body problem.

Step 3: Solving the problem of other minds

As noted before, careful phenomenological analysis reveals that the absolute distinction between the first-person perspective (experiencing myself as an 'I') and the third-person perspective (experiencing the other as an 'it') must be rejected as inadequate. It simply neglects the fact that we can also encounter each other in the second-person perspective (experiencing the other as a 'you'). There are of course distressing pathological cases in which people have limited or no access to enter the second-person perspective, such as in autism and schizophrenia. However, under ordinary circumstances we do not encounter the other merely as an 'it', although we can choose to do so as would perhaps a soldier who is aiming to kill his enemy or a surgeon who is operating on the body of his patient.

But under everyday conditions, while interacting with others, I do not perceive my fellow human beings in terms of zombie-like bodies with the theoretical possibility of having a mind. On the contrary, I directly perceive that the others around me are living-minded beings in their own right, and I perceive that they can perceive me in this way, too. This immediate perceptual realization of the other's lived embodiment precedes any kind of theoretical reflection¹⁶.

There are two important consequences of this phenomenological analysis for our current discussion. First, it follows that the major theories of social cognition have been confined to an overly narrow view of sociality. They have focused on cognitive mechanisms for the acquisition of reflective knowledge about others, while the possible role of direct pre-reflective intuition of others has been neglected. Yet when I meet someone I do not have to theoretically posit the possibility of my interlocutor's mental existence; I immediately perceive the other's lived presence in their gestures. This kind of direct perception has typically been described in terms of empathy or sympathy in the phenomenological tradition. Note that once we acknowledge the existence of the pre-reflective second-person perspective there is no longer any problem

¹⁶ Zahavi (2001), Stanghellini (2004), Gallagher (2008), and Stawarska (2009).

of other minds as such, at least not on the reflective level traditionally targeted by cognitive psychology. And this brings us to the second point, namely that we are now faced with the challenge to explain the basis for our pre-reflective understanding of others.

One promising approach in this regard is to focus on the existence of a shared basis of understanding or a kind of common sense, literally speaking. Variations of this idea have been proposed by existential phenomenology, where it is typically described as the lack of an absolute 'self' and 'other' distinction during interaction at the pre-reflective level of the lived body. But these ideas have typically been expressed philosophically without any serious consideration of modern science. The traditional problem of other minds, on the other hand, is closely aligned to cognitive science, but is instead typically conceived in terms of how to explain the possibility of epistemic access between two metaphysically isolated minds. Fortunately, the life-mind continuity thesis opens up another promising option. If we accept the tenet that a living body is also a lived body, and if we can also show that interacting bodies become one body to some extent, then there is indeed a basis for directly sharing in each other's mindedness. This is one of the central ideas of my work, which I will call the *extended body hypothesis*.

It is not too difficult to show that, in theoretical terms at least, a concept of the living organism centered on the notion of an autopoietic system can be extended in the required manner. In fact, there is currently much discussion in the enactive approach to cognitive science which focuses precisely on this issue¹⁷.

The first thing to emphasize is that a real autopoietic system cannot be thought in isolation from its environment. The membrane of a cell, for instance, preserves the cell not by simply isolating it, but rather by adaptively regulating environmental interactions. Of course, the cell must keep itself distinct and safe from the aversive influences of its environment in order to survive. And yet at the same time it must be open enough to engage the environment in a flexible manner and to find the nutrients which it needs to support its metabolism. On this view, the cell and the environment are not two isolated systems, but are intrinsically related in an asymmetric way (with the former depending on the latter for its existence). The upshot of this is that life is not some kind of independent substance, but a relational process¹⁸.

Furthermore, the environmental side of this relation is not limited to chemical elements alone, but can include other organisms as well. In the case of the cell it is quite possible that its metabolism directly depends on the products of other organisms. And if some of those organisms also depend on the cell's excretions, then we have

¹⁷ De Jaegher and Di Paolo (2007); Froese and Di Paolo (2009); De Jaegher and Froese (2009).

¹⁸ Ruiz-Mirazo and Moreno (2004); Di Paolo (2005); Barandiaran and Moreno (2008); Barandiaran, et al. (2009); Virgo, et al. (in press).

established another cycle of co-dependency which relates these two organisms into a larger autonomous unit, an extended living body. In some cases this new autonomy can form a complex structural coherence so that we start speaking of multi-cellular organisms. The autopoietic theory of the organism can therefore be scaled up¹⁹. As an aside, note that if we accept that a living body and a lived mind are complementary phenomena, then here we have a partial response to the question of how we can be made out of individual living bodies (cells) and yet still experience the world from one unified perspective.

The metabolic formation of co-dependencies and multi-cellular organisms may serve as a general proof of concept for the extended body, but it does not help us directly with the task of finding a potential shared basis for our pre-reflective understanding of others. But can social interactions play a similar role perhaps? Again, organism-centered biology already has the basic theoretical resources we need. Rather than focusing on cycles of co-dependency in the chemical domain, such as autopoietic self-production, we simply turn our attention to the autonomous cycles in the way in which organisms relate to their environment. In the case of animals, for instance, we find a clear example that what the animal does depends on what it senses, and that what it senses depends on what it does²⁰.

And furthermore, given that a living body is also a minded body, each bodily movement is complemented not only by a change in sense perception, but also by a change in the lived body. Living, sensing, and doing are, through the mediation of the environment and the body, linked into one unified sensorimotor loop. On this view, the absolute distinction between perception, cognition, and action in traditional cognitive science (most famously expressed in the sense-model-plan-act framework) appears to be arbitrary. The enactive approach, more than any of the alternative approaches to cognitive science, has tried to do justice to the co-dependent nature of living, sensing and doing (and more generally of experiencing and moving) by employing the notion of *sense-making* in a literal way²¹.

Since sense-making is based on the organism's sensorimotor loop and is therefore mediated through the environment, and the environment includes other organisms with their own sensorimotor loops, there is always the potential that there will be an interaction between the sense-making of distinct individuals. Moreover, it may happen that the organisms interact in a mutually responsive manner such that their sensorimotor loops become extended into a larger autonomous cycle. In addition to the usual sense-making, one organism's behavior leads to a change in the other organis-

¹⁹ Maturana and Varela (1987); Froese and Di Paolo (in press).

²⁰ Out of all the ideas discussed in this article, the co-dependency between perception and action has been the most popular research topic among alternative approaches to cognitive science. See, e.g., Varela, et al. (1991); O'Regan and Noë (2001); Noë (2004; 2009); Barandiaran and Moreno (2006); Thompson (2005).

²¹ Weber and Varela (2002); Thompson (2004); Di Paolo (2005); Sheets-Johnstone (1999).

m's sensations thus resulting in a change of its behavior which in turn modifies the sensations of this first organism and its behavior, and so forth. When the two organisms modulate each other's sensorimotor loop in this manner, they of course also modulate each other's sense-making activities. They are engaging in *participatory sense-making*²².

The concept of participatory sense-making provides us with the theoretical foundation we need in order to explain the basis of our pre-reflective understanding of others. When our mutual interaction with others turns into an autonomous cycle in its own right, our bodies temporarily become one extended body via the dynamical mediation of our respective sensorimotor loops. And, once again, it is important to emphasize that according to the basic tenets of the life-mind continuity thesis *an extended living body entails an extended lived body*. What this means is that in the second-person perspective, when our bodies become dynamically entangled, we can take part in each other's experience.

In sum, we have argued that when we become aware of how we experience others during our interactions with them, we notice that at least on this existential level there is no problem of other minds. I can directly perceive others as living-minded beings like myself, and I can see that they perceive me likewise, too. Moreover, it is possible to give a coherent explanation for the possibility of sharing experience with others in terms of the life-mind continuity thesis and the extended body hypothesis.

Of course, what still needs to be explained is how the principles of participatory sense-making can be scaled up to account for the reflective levels of human social cognition, including the capacity for theorizing about other minds. But, more importantly, we have replaced yet another dilemma of metaphysical proportions by one of levels of complexity, and we are simply faced with a variation of the cognitive gap.

Step 4: Bridging the cognitive gap again, in practice

However, wasn't this detour through the problem of other minds supposed to provide us with a solution to the cognitive gap by making our relations with others an intrinsic element in support of cognitive development? If relations with others are also limited by a cognitive gap, then it may seem that we have hit another dead end.

But this worry underestimates the role played by pre-reflective second-person interaction in our relations with others, and it also underestimates the potential of the interaction process to autonomously organize our individual abilities. We can avoid the dead end if it is possible to demonstrate that, during its mutual interaction with others, a young infant's behavior can become appropriately entrained by the autonomous dy-

²² This notion was introduced by De Jaegher and Di Paolo (2007) and has generated a lively discussion in the enactive community, e.g. Steiner & Stewart (2009); Gallagher (2009); Torrance & Froese (2011).

namics of the interaction process, and its internal bodily organization can be spontaneously configured to further sustain the interaction in a more flexible and adaptive way. That is, we need to show that, as long as the infant has the ability to mutually attune with others in the right kind of way, it can develop more advanced abilities through its interaction with others relatively automatically. And if the required sensitivity to others can also be achieved in the interaction itself, then the entire problem of the cognitive gap can be offloaded into a relational matrix distributed over self, others and the rest of the world.

Of course, it is one thing to show that this scenario is theoretically possible, and it is an entirely different thing to show that it is actually possible. In fact, given the major change in departure point which informs the scenario, an important first step will be to show that something like it is even possible in principle. This has motivated me to design a series of agent-based models of interaction which demonstrate that:

- An individual's sensitivity to others can result from interacting with others
- An individual's movements can be structured through interacting with others
- An individual's sensations can be structured through interacting with others
- An individual's body can be structured through interacting with others

It is beyond the scope of this paper to describe these modeling results and their dynamical systems analysis in more detail here, and I have done so extensively elsewhere²³. What is important for our current discussion is that these models are concrete proof of concepts that it is possible for the cognitive gap to be systematically addressed through the spontaneously enabling dynamics of social interaction. Of course, the next step will be to continue this initial research by conducting actual psychological experiments, and some promising work in this direction is already underway²⁴.

Conclusion

All in all, the upshot of this article is that the computational theory of mind can indeed be replaced with an alternative framework centered on the notion of life-mind continuity, but only by incorporating the constitutive role played by sociality. I have argued that this alternative approach to a unified cognitive science has compelling advantages over other approaches. It is a workable scientific research program that stays

²³ The most detailed discussion so far can be found in my doctoral dissertation (Froese 2009), but some of the results can also be found in individual papers (e.g. Froese & Di Paolo 2008; 2010; in press-b).

²⁴ Auvray, et al. (2009)

clear of metaphysical conundrums, offers a closer alignment to human experience, and at the same time has sound mathematical foundations.

References:

- Auvray, M., Lenay, C. & Stewart, J. 2009. Perceptual interactions in a minimalist virtual environment. *New Ideas in Psychology*, 27(1): 32-47.
- Barandiaran, X., Di Paolo, E. A. & Rohde, M. 2009. Defining Agency: Individuality, Normativity, Asymmetry, and Spatio-temporality in Action, *Adaptive Behavior*, 17(5): 367-386.
- Barandiaran, X. & Moreno, A. 2006. On what makes certain dynamical systems cognitive: a minimally cognitive organization program", *Adaptive Behavior*, 14(2): 171-185.
- Beer, R. D. 2003. The dynamics of active categorical perception in an evolved model agent, *Adaptive Behavior*, 11(4): 209-243.
- De Jaegher, H. & Di Paolo, E. A. 2007. Participatory sense-making: An enactive approach to social cognition. *Phenomenology and the Cognitive Sciences*, 6(4): 485-507.
- De Jaegher, H. & Froese, T. 2009. On the Role of Social Interaction in Individual Agency. *Adaptive Behavior*, 17(5): 444-460.
- Di Paolo, E. A. 1997. An investigation into the evolution of communication. *Adaptive Behavior*, 6(2): 285-324.
- Di Paolo, E. A. 2005. Autopoiesis, adaptivity, teleology, agency. *Phenomenology and the Cognitive Sciences*, 4(4): 429-452.
- Di Paolo, E. A. 2009. Extended life. *Topoi*, 28(1): 9-21.
- Froese, T. 2009. *Sociality and the life-mind continuity thesis: a study in evolutionary robotics*. D.Phil. Thesis, Brighton, UK: University of Sussex.
- Froese, T. & Di Paolo, E. A. 2008. Stability of coordination requires mutuality of interaction in a model of embodied agents: 52-61. Eds. Asada, M., Hallam, J. C. T., Meyer J.-A. & Tani, J. *From Animals to Animals 10: Proc. of the 10th Int. Conf. on Simulation of Adaptive Behavior*. Berlin, Germany: Springer-Verlag.
- Froese, T. & Di Paolo, E. A. 2009. Sociality and the life-mind continuity thesis. *Phenomenology and the Cognitive Sciences*, 8(4): 439-463.
- Froese, T. & Di Paolo, E. A. (in press). The enactive approach: From cell to society. *Pragmatics & Cognition*, forthcoming.
- Froese, T. & Di Paolo, E. A. (in press-b). Toward Minimally Social Behavior: Social Psychology Meets Evolutionary Robotics. *Advances in Artificial Life: Proc. of the 10th Euro. Conf. on Artificial Life*. Berlin, Germany: Springer-Verlag.

- Froese, T. & Gallagher, S. 2010. Phenomenology and Artificial Life: Toward a Technological Supplementation of Phenomenological Methodology. *Husserl Studies*, 26(2): 83-106.
- Froese, T. & Stewart, J. 2010. Life after Ashby: Ultrastability and the autopoietic foundations of biological individuality. *Cybernetics & Human Knowing*, 17(4): 7-50.
- Froese, T. & Ziemke, T. 2009. Enactive Artificial Intelligence: Investigating the systemic organization of life and mind. *Artificial Intelligence*, 173(3-4): 366-500.
- Fuchs, T. (in press). Body memory and the unconscious. Eds. Lohmar, D. & Brudzinska, J. *Founding Psychoanalysis: Phenomenological Theory of Subjectivity and the Psychoanalytical Experience*. Dordrecht: Kluwer.
- Gallagher, S. 2008. Direct perception in the intersubjective context. *Consciousness and Cognition*, 17(2): 535-543.
- Gallagher, S. 2009. Two problems of intersubjectivity: *Journal of Consciousness Studies*, 16(6-8): 289-308.
- Gallagher, S. & Zahavi, D. 2008. *The Phenomenological Mind: An Introduction to Philosophy of Mind and Cognitive Science*. London, UK: Routledge.
- Hanna, R. & Thompson, E. 2003. The Mind-Body-Body Problem. *Theoria et Historia Scientiarum*, 7(1): 24-44.
- Jonas, H. 1966 [2001]. *The Phenomenon of Life: Toward a Philosophical Biology*. Evanston, Illinois: Northwestern University Press.
- Lutz, A. & Thompson, E. 2003. Neurophenomenology: integrating subjective experience and brain dynamics in the neuroscience of consciousness. *Journal of Consciousness Studies*, 10: 31-52.
- Noë, A. 2004. *Action in Perception*. Cambridge, MA: The MIT Press.
- Noë, A. 2009. *Out of Our Heads: Why You Are Not Your Brain, and Other Lessons from the Biology of Consciousness*. New York, NY: Hill and Wang.
- O'Regan, J.K. & Noë, A. 2001. a sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24(5): 939-1031.
- Oyama, S. 2000. *The Ontogeny of Information: Developmental Systems and Evolution*. 2nd edition. Durham: Duke University Press.
- Petitmengin, C. 2006. Describing one's subjective experience in the second person: An interview method for the science of consciousness. *Phenomenology and the Cognitive Sciences*, 5(3-4): 229-269.
- Port, R. F. & van Gelder, T. Eds. 1995. *Mind as Motion: Explorations in the Dynamics of Cognition*. Cambridge, MA: MIT Press.

- Rohde, M. 2010. *Enaction, Embodiment, Evolutionary Robotics: Simulation Models for a Post-Cognitivist Science of Mind*. Atlantis Press.
- Roy, J.-M., Petitot, J., Pachoud, B. & Varela, F. J. 1999. Beyond the Gap: An Introduction to Naturalizing Phenomenology: 1-80. Eds. Petitot, J., Varela, F. J., Pachoud, B. & Roy, J.-M. *Naturalizing Phenomenology: Issues in Contemporary Phenomenology and Cognitive Science*. Stanford, CA: Stanford University Press.
- Ruiz-Mirazo, K. & Moreno, A. 2004. Basic Autonomy as a Fundamental Step in the Synthesis of Life. *Artificial Life*, 10(3): 235-259.
- Savage-Rumbaugh, S., Shanker, S. G. & Taylor, T. J. 1998. *Apes, Language, and the Human Mind*. New York, NY: Oxford University Press.
- Shanker, S. G. & King, B. J. 2002. The emergence of a new paradigm in ape language research. *Behavioral and Brain Sciences*, 25(5): 605-620.
- Sheets-Johnstone, M. 1999. *The Primacy of Movement*. Amsterdam: John Benjamins Publishing.
- Smith, L. & Thelen, E. 2003. Development as a dynamical system. *Trends in Cognitive Sciences*, 7(8): 343-348.
- Stanghellini, G. 2004. *Disembodied Spirits and Deanimated Bodies: The psychopathology of common sense*. New York, NY: Oxford University Press.
- Stawarska, B. 2009. *Between You and I: Dialogical Phenomenology*. Athens: Ohio University Press.
- Stewart, J., Gapenne, O. & Di Paolo, E. A. Eds. 2010. *Enaction: Towards a New Paradigm for Cognitive Science*. Cambridge, MA: The MIT Press.
- Thelen, E. & Smith, L. B. 1994. *a dynamical systems approach to the development of cognition and action*. Cambridge, MA: The MIT Press.
- Thompson, E. 2004. Life and mind: From autopoiesis to neurophenomenology. a tribute to Francisco Varela. *Phenomenology and the Cognitive Sciences*, 3(4): 381-398.
- Thompson, E. 2005. Sensorimotor subjectivity and the enactive approach to experience. *Phenomenology and the Cognitive Sciences*, 4(4): 407-427.
- Thompson, E. 2007. *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Cambridge, MA: The Belknap Press of Harvard University Press.
- Torrance, S. & Froese, T. 2011. An inter-enactive approach to agency: Participatory sense-making, dynamics, and sociality. *Humana.Mente*, 15: 21-53.
- Maturana, H.R., & Varela, F. J. 1987. *The Tree of Knowledge: The Biological Roots of Human Understanding*. Boston, MA: Shambhala Publications.
- McMullin, B. 2004. Thirty Years of Computational Autopoiesis: a Review. *Artificial Life*, 10(3): 277-295.

- Varela, F. J. 1996. Neurophenomenology: a Methodological Remedy for the Hard Problem. *Journal of Consciousness Studies*, 3(4): 330-349.
- Varela, F. J. 1997. Patterns of Life: Intertwining Identity and Cognition. *Brain and Cognition*, 34(1): 72-87.
- Varela, F. J, Thompson, E. & Rosch, E. 1991. *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: The MIT Press.
- Virgo, N., Egbert, M. & Froese, T. (forthcoming). The Role of the Spatial Boundary in Autopoiesis. *Proc. of the 10th Euro. Conf. on Artificial Life*. Berlin, Germany: Springer-Verlag.
- Weber, A. & Varela, F. J. 2002. Life after Kant: Natural purposes and the autopoietic foundations of biological individuality. *Phenomenology and the Cognitive Sciences*, 1: 97-125.
- Zahavi, D. 2001. Beyond Empathy: Phenomenological Approaches to Intersubjectivity. *Journal of Consciousness Studies*, 8(5-7): 151-167.
- Zahavi, D. 2005. *Subjectivity and Selfhood: Investigating the First-Person Perspective*. Cambridge, MA: The MIT Press.
- Zahavi, D. 2006. Two takes on a one-level account of consciousness. *Psyche*, 12(2): 1-9.