

Czym jest to, co zwiemy ucieleśnieniem?

Tom Ziemke

School of Informatics,
Univeristy of Skövde
tom.ziemke@his.se⁶¹

Przekład: Aleksandra Derra (wersja ostateczna) oraz uczestnicy translatorium⁶²

(Przekład zmodyfikowanej przez autora wersji tekstu pt. „What's that thing called embodiment?”, który ukazał się w 2003 r. w: R. Alterman, D. Kirsh, red. *Proceedings of the 25th Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum: 1134–1139⁶³)

Abstrakt

Ucieleśnienie stało się ważnym pojęciem wielu obszarów kognitywistyki. Jednak różnie określa się, czym ono dokładnie jest i jakiego rodzaju ciała wymaga się dla określonego typu poznania ucieleśnionego. Stąd chociaż wiele osób zgodziłoby się dzisiaj, że ludzie są ucieleśnionymi podmiotami poznającymi, nie ma pełnej zgody co do tego, jakiego rodzaju artefakt można uznać za ucieleśniony. W tym artykule wyróżniamy i zestawiamy sześć różnych pojęć ucieleśnienia, które z grubsza można scharakteryzować jako: (1) *sprzężenie strukturalne* między podmiotem [*agent*] a środowiskiem, (2) *ucieleśnienie historyczne* jako coś, co wynika z historii sprzężenia strukturalnego, (3) *ucieleśnienie fizyczne*, (4) *ucieleśnienie organizmoidalne*, czyli dotyczące organizmopodobnych form cielesnych (na przykład robotów humanoidalnych), (5) *ucieleśnienie organizmowe* autopojetycznych, żywych systemów oraz (6) *ucieleśnienie społeczne*.

Słowa kluczowe: kognitywistyka; poznanie; ucieleśnienie; podmiot; środowisko; sprzężenie strukturalne; fizyczny; organizm; AI; robot; społeczne.

⁶¹ W porównaniu z oryginałem dane kontaktowe zostały uaktualnione (przyp. red.).

⁶² Pierwotna, robocza wersja tłumaczenia powstała na zajęciach translatorskich z języka angielskiego prowadzonych przez Aleksandrę Derra w roku akademickim 2014/2015 w Instytucie Filozofii UMK, w którym wzięły udział następujące osoby: Emilia Cheładze, Magdalena Józefowicz, Jan Nikadon, Magdalena Szymtke, Olga Walaszczyk, Jakub Wojciechowski. Przekład przejrzała, poprawiła i przygotowała do publikacji Aleksandra Derra.

⁶³ Przekład publikowany za zgodą właścicieli praw do tekstu.

Wprowadzenie

Pojęcie ucieleśnienia jest obszernie używane od połowy lat osiemdziesiątych XX wieku w literaturze z zakresu kognitywistyki i sztucznej inteligencji (AI). Znajdziemy w niej takie terminy jak *umysł ucieleśniony* (np. Lakoff i Johnson 1999; Varela i in. 1991), *inteligencja ucieleśniona* (np. Brooks 1991), *działanie ucieleśnione* (np. Varela i in. 1991), *poznanie ucieleśnione* (np. Clark 1997), *ucieleśniona sztuczna inteligencja* (np. Chrisley 2003; Franklin 1997) i *ucieleśniona kognitywistyka* (np. Clark 1999; Pfeifer i Scheier 1999). Co więcej, istnieją rozmaite typy i pojęcia ucieleśnienia, co widać po liczbie sposobów ich określania, takich jak: *ucieleśnienie usytuowane* (Zlatev 1997), *ucieleśnienie mechaniczne* (Sharkey i Ziemke 2001), *ucieleśnienie fenomenalne* (Sharkey i Ziemke 2001), *ucieleśnienie naturalne* (Ziemke 1999), *ucieleśnienie naturalistyczne* (Zlatev 2001), *ucieleśnienie społeczne* (Barsalou i in. 2003) oraz, w tym artykule, *ucieleśnienie historyczne, fizyczne, organizmoidalne i organizmowe*.

Współcześnie wielu badaczy uważa ucieleśnienie za *conditio sine qua non* jakiegokolwiek formy naturalnej lub sztucznej inteligencji. Na przykład Pfeifer i Scheier (1999) przekonywali, że „inteligencja nie może istnieć jedynie pod postacią abstrakcyjnego algorytmu, ale wymaga fizycznej konkretyzacji – ciała”. Ponadto ucieleśnienie powszechnie uważa się za jedną z kluczowych idei, które odróżniają najnowsze teorie ucieleśnionego, usytuowanego i rozproszonego poznania od klasycznej kognitywistyki, która, opierając się na funkcjonalizmie, skupiała się na „odcieleśnionych” obliczeniach (por. Chrisley i Ziemke 2002; Clark 1997; Pfeifer i Scheier 1999; Varela i in. 1991; Ziemke 2002). Mimo że obecnie wielu badaczy zgadza się, że poznanie musi być ucieleśnione, jak dotąd nie jest jasne, co to dokładnie znaczy. Wilson (2002) zwróciła niedawno uwagę, że różnorodność poglądów w tej dziedzinie jest problematyczna:

Pomimo że ogólne podejście [ucieleśnionego poznania lub ucieleśnionej kognitywistyki] cieszy się coraz szerszym poparciem, mamy tu w istocie do czynienia z ogromną różnorodnością wchodzących w grę twierdzeń, które są w różnym stopniu kontrowersyjne. Jeżeli chcemy sensownie posługiwać się terminem „ucieleśnione poznanie”, musimy rozwikłać i ocenić tę różnorodność.

W szczególności niejasne jest jakiego rodzaju *ciała* (jeśli jakiegokolwiek) potrzebujemy do ucieleśnionego poznania. Stąd, o ile można uznać, że ludzie są ucieleśnionymi podmiotami poznającymi, nie ma wyraźnej zgody co do tego, w jakiego typu ciało należałoby wyposażyć sztuczną inteligencję. Na przykład pogląd Pfeifera i Scheiera (1999), że inteligencja wymaga ciała *fizycznego*, nie jest wcale tak powszechnie akceptowany, jak mogłoby się wydawać.

W tym artykule wyróżniamy i omawiamy kilka różnych pojęć ucieleśniania z literatury z zakresu kognitywistyki i sztucznej inteligencji. Ze względu na objętość tekstu żadne z tych pojęć nie jest tu poddane szczegółowej ocenie, chociaż trzeba przyznać, że więcej uwagi niż pozostałym poświęcono pojęciu „ucieleśnienia organizmowego”, ponieważ pojawiło się ono w naszych wcze-

śniejszych pracach (Sharkey i Ziemke 1998, 2001; Ziemke 1999, 2001a; Ziemke i Sharkey 2001).

Dalsza część artykułu ma następującą strukturę. W kolejnym fragmencie pokrótce opisujemy różne poglądy na temat ucieleśnionego poznania, trzymając się rozróżnień wprowadzonych przez Wilson (2002). W następnej części wyróżniamy różne pojęcia ucieleśnienia, koncentrując się na pytaniu o rodzaj ciała, jakie potrzebne jest do ucieleśnionego poznania. Należy zauważyć, że wiele omawianych tutaj pojęć i przykładów pochodzi z rozważań na temat systemów AI po prostu dlatego, że problem ucieleśnienia jest zazwyczaj najbardziej konkretny i naglący przy tworzeniu systemów sztucznych. Jednak przewodnie pytanie naszej rozprawy dotyczy tego, na ile użyteczne dla kognitywistyki są różne pojęcia ucieleśnienia, w szczególności do rozpoznania i zrozumienia przypadków ucieleśnionego poznania. W ostatniej części przedstawiamy krótkie podsumowanie.

Różne ujęcia ucieleśnionego poznania

W niedawnym artykule Wilson (2002) wyróżniła sześć różnych ujęć ucieleśnionego poznania, jednak tylko jedno z nich wyraźnie odnosi się do roli ciała.

„Poznanie jest usytuowane”. Twierdzenie to, rzecz jasna, jest powszechnie przyjmowane w literaturze na temat ucieleśnionego poznania⁶⁴. Sama Wilson wyróżniła poznanie usytuowane, które przebiega „w kontekście istotnych dla zadania danych wejściowych i wyjściowych” i „poznanie *offline*”, które przebiega bez ich udziału.

„Poznanie zachodzi pod presją czasu”. Oznacza to, że poznanie jest ograniczone przez wymagania zachodzącej w czasie rzeczywistym interakcji z otoczeniem. Jako przykład może posłużyć „reprezentacyjny efekt wąskiego gardła” [*representational bottleneck*] (Clark 1997).

„Ciężar pracy poznawczej przenosimy na środowisko”. Brooks (1991) sformułował podobną tezę, zgodnie z którą „świat jest swoim najlepszym modelem”. Dobrze znany przykład stanowią badania Kirsh i Maglio (1994) „czynności epistemicznych” podczas gry w „Tetris”, czyli ruchów przygotowujących do podejmowania decyzji, które wykonywane są raczej w świecie niż w głowie.

„Środowisko jest częścią systemu poznawczego”. Takie podejście może ilustrować praca Hutchinsa (1995) o rozproszonym poznaniu, w której na przykład przyrządy w kokpicie traktuje się jako część systemu poznawczego. Jednak według Wilson „stosunkowo niewielu teoretyków zdaje się konsekwent-

⁶⁴ Należy zauważyć, że samo pojęcie usytuowania nie jest zbyt dobrze zdefiniowane (por. np. Ziemke 2001a), a to przedstawiane przez Wilson nie jest pozbawione kontrowersji. Zob. także pracę Clancey'a (2001), który wprowadził użyteczne rozróżnienie na funkcjonalną, strukturalną i behawioralną perspektywę usytuowanego poznania.

nie bronić to stanowisko w jego skrajnej postaci” (bardziej szczegółowe omówienie na ten temat można znaleźć w: Susi, Lindblom i Ziemke 2003).

„Celem poznania jest działanie”. Twierdzenie to zostało wysunięte przykładowo przez Franklina (1997), który wykazywał, że umysły są strukturami sterującymi w podmiotach autonomicznych.

„Poznanie *offline* jest ugruntowane w ciele”. Według Wilson temu twierdzeniu nie poświęcono dotąd wystarczającej uwagi w literaturze z zakresu kognitywistyki, chociaż „może ono być najlepiej udokumentowane i najmocniejsze z wszystkich sześciu”. Być może jego najsłynniejszą ilustracją jest praca Lakoffa i Johnsona (1980), którzy przekonywali, że pojęcia abstrakcyjne opierają się na metaforach ugruntowanych w doświadczeniu/działaniu cielesnym. Twierdzenie to zostanie dokładniej omówione w kolejnej części.

Różne pojęcia ucieleśnienia

Jak zauważono w poprzedniej części, co może zaskakiwać, w wielu opisach czy ujęciach ucieleśnionego poznania właściwie niewiele (jeśli w ogóle) uwagi poświęca się naturze i roli samego ciała. Jedynie w szóstym ujęciu, „poznania *offline*” jako ugruntowanego w ciele, Wilson bezpośrednio mówi o ciele jako czymś, co odgrywa kluczową rolę. Pozostaje jednak otwarte pytanie o to, czy na przykład robot humanoidalny, a więc mniej lub bardziej podobny do człowieka, mógłby posiadać ten sam typ poznania co jego żywy odpowiednik.

Chcielibyśmy tu wyodrębnić sześć następujących pojęć odnoszących się do rodzaju ciała/ucieleśniania, jakiego wymaga (ucieleśnione) poznanie: (1) *sprzężenie strukturalne* pomiędzy podmiotem a środowiskiem, (2) *ucieleśnienie historyczne* jako to, co wynika z historii sprzężenia strukturalnego, (3) *ucieleśnienie fizyczne*, (4) *ucieleśnienie organizmoidalne*, czyli dotyczące organizmopodobnych form cielesnych (na przykład robotów humanoidalnych), (5) *ucieleśnienie organizmowe* autopojetycznych, żywych systemów i (6) *ucieleśnienie społeczne*.

Każde z powyższych pojęć ucieleśnienia omawiamy poniżej w osobnych podsekcjach. Warto od razu zauważyć, że przynajmniej część z nich tworzy raczej grupy mniej lub bardziej powiązanych koncepcji niż pojedyncze, dobrze zdefiniowane stanowiska.

Ucieleśnienie jako sprzężenie strukturalne

Prawdopodobnie najszersza koncepcja ucieleśnienia głosi, że systemy są ucieleśnione, jeśli są „strukturalnie sprzężone” ze swoim środowiskiem. Warto zauważyć, że niekoniecznie wymaga to ciała. Franklin (1997) na przykład przekonywał, że:

Systemy operacyjne, które nie posiadają ciała w zwykłym fizycznym sensie, mogą być inteligentne. Muszą jednak być ucieleśnione w sensie usytuowania jako autonomiczne podmioty strukturalnie sprzężone ze swoim środowiskiem.

Pojęcie *sprzężenia strukturalnego* pochodzi z prac Maturany i Vareli (1980, 1987) na temat biologii poznania i będzie przedmiotem rozważań w dalszej części. Quick i inni (1999)⁶⁵ zainspirowani tą koncepcją próbowali podać „dokładną definicję” ucieleśnienia, która brzmi:

System X jest ucieleśniony w pewnym środowisku E, jeśli istnieją pomiędzy nimi kanały zakłócające. To znaczy, X jest ucieleśniony w E, jeśli w każdym czasie t, w którym zarówno X jak i E istnieją, pewien podzbiór możliwych stanów E powiązanych z X-em ma zdolność zakłócania stanu X-a oraz pewien podzbiór możliwych stanów X-a powiązanych z E ma zdolność zakłócania stanów E.

Powyższe pojęcie ucieleśnienia ma tę przewagę, że powstało w wyniku jednej z niewielu prób, aby faktycznie je zdefiniować. Jednak można uznać, że ma ograniczoną użyteczność w kognitywistyce, ponieważ nie jest szczególnie restrykcyjne. Riegler (2002) na przykład dowodził, że „(...) chociaż ta próba uściślenia pojęcia ucieleśnienia stanowi ważny pierwszy krok, jednocześnie jest to niewystarczająca charakterystyka” ze względu na to, że „każdy system jest w tym czy innym sensie strukturalnie sprzężony ze swoim środowiskiem”. Oznacza to, że ta definicja ucieleśnienia nie odróżnia systemów poznawczych od niepoznawczych, co można zilustrować przykładem skały granitowej (X) z tundry na Antarktydzie (E) z publikacji Quicka i innych (1999). Skała jest nieustannie zakłócana przez wiatr, a z kolei przepływ powietrza jest zakłócany przez skałę. Zatem zgodnie z powyższą definicją jest to system ucieleśniony, mimo że z pewnością niewielu kognitywistów naprawdę uznałoby to za przykład ucieleśnionego poznania.

Ucieleśnienie historyczne

Kilku badaczy podkreślało, że systemy poznawcze są nie tylko strukturalnie sprzężone ze swoim środowiskiem w czasie rzeczywistym, ale też ich ucieleśnienie jest w istocie wynikiem lub odzwierciedleniem *historii* interakcji pomiędzy podmiotem a środowiskiem i, w wielu wypadkach, obustronnej adaptacji. Varela, Thompson i Rosch (1991) na przykład dowodzili, że „wiedza zależy od bycia w świecie, co nie daje się oddzielić od naszych ciał, naszego języka i naszej historii społecznej – w skrócie, od naszego ucieleśnienia”. W podobnym tonie Ziemke (1999) przekonywał, że „(...) naturalne ucieleśnienie [systemów żywych] ... odzwierciedla/ucieleśnia historię strukturalnego sprzężenia pomiędzy podmiotem a środowiskiem i ich wzajemnego dookreślenia, w trakcie którego ciało zostało skonstruowane.” Podobnie Riegler (2002)

⁶⁵ Zob. także Dautenhahn, Ogden, Quick 2002.

w swoją definicję ucieleśnienia włącza adaptację podmiotu do środowiska: „System jest ucieleśniony, jeśli wypracował umiejętności w obrębie środowiska, w którym się rozwinął”.

Mimo że jest to bardziej restrykcyjne pojęcie niż poprzednie oraz prawdopodobnie użyteczne w dookreślaniu ewolucyjnych i/lub rozwojowych korzeni ucieleśnionego poznania, rzecz jasna nawet ono nie jest szczególnie restrykcyjne, ponieważ wydaje się stosować do wielu, jeśli nie wszystkich, systemów żywych, obejmując także systemy niefizyczne (por. poniżej).

Ucieleśnienie fizyczne

Pojęcie ucieleśnienia, które wyklucza takie podmioty jak systemy operacyjne, wiąże się z poglądem, że systemy ucieleśnione potrzebują „fizycznej realizacji” w rozumieniu Pfeifera i Scheiera (por. wprowadzenie), czyli fizycznego ciała. Samo w sobie nie jest to bardzo restrykcyjne i wciąż obejmuje skalę granitową z przykładu Quicka i innych (1999), jak również krzesła, stoły i wiele innych obiektów, które najwyraźniej nie są przedmiotem szczególnego zainteresowania kognitywistyki.

Warto zwrócić uwagę, że chociaż zarówno ucieleśnienie historyczne, jak i ucieleśnienie fizyczne można potraktować jako szczególne przypadki sprzężenia strukturalnego, to ani się wzajemnie nie zawierają, ani nie wykluczają. Riegler (2002) na przykład zauważył, że jego definicja ucieleśnienia (zob. powyżej) „nie wyklucza sfer innych niż fizyczna”; w szczególności „(...) programy komputerowe także mogą stać się ucieleśnione”, jeśli powstały raczej wskutek samoorganizacji niż jawnego zaprojektowania. Podobnie systemy żywe są systemami fizycznie ucieleśnionymi, które są również ucieleśnione historycznie, co nie ma miejsca w przypadku wielu innych systemów fizycznych.

Nieco bardziej restrykcyjne rozumienie pojęcia ucieleśnienia fizycznego wiąże się z poglądem, że systemy ucieleśnione powinny być połączone ze swoim środowiskiem nie tylko za pomocą sił fizycznych, ale też sensorycznych i motorycznych. Brooks (1990) na przykład z perspektywy badań nad sztuczną inteligencją określił to jako *zakotwiczenie fizyczne*, przekonując, że „aby zbudować inteligentny system, konieczne jest zakotwiczenie jego reprezentacji w świecie fizycznym ... [to jest] konieczne jest połączenie go ze światem za pomocą zestawu sensorów i urządzeń wykonawczych”.

Prawdopodobnie można potraktować to jako niezależne pojęcie i określić mianem *ucielesnienia sensomotorycznego*. Niemniej ze względu na to, że raczej trudno jest zdefiniować, czym dokładnie jest „to, co sensoryczne” i „to, co motoryczne” i czym różnią się one od innych „kanałów zakłócających”, nie czynimy tego w naszym artykule.

Ucieleśnienie organizmoidalne

Zgodnie z kolejnym, jeszcze bardziej restrykcyjnym rozumieniem ucieleśnienia fizycznego przynajmniej niektóre typy poznania organizmopodobnego mogą mieć miejsce jedynie w ciałach organizmoidalnych, czyli w ciałach fizycznych, które przynajmniej do pewnego stopnia mają identyczną lub podobną formę i zdolności sensomotoryczne jak żywe ciała. Należy zauważyć, że pojęcie ucieleśnienia „organizmoidalnego” ma odnosić się tutaj zarówno do organizmów żywych, jak i ich sztucznych odpowiedników⁶⁶.

Za jeden z najprostszych przykładów ucieleśnienia organizmoidalnego może uchodzić robot Khepera, którego użyli Lund, Webb oraz Hallam (1998). Miał on naśladować fonotaksję świerszcza, dlatego posiadał układ dźwiękowy i dwa mikrofony, które dzieliła taka sama odległość jak „usz” owada. W tym przypadku rozmieszczenie czujników zarówno u świerszcza, jak i u robota zredukowało ilość przetwarzania wewnętrznego niezbędnego do wybiórczego reagowania na dźwięki o określonej częstotliwości. Zauważmy, że w tym przypadku ciało świerszcza i ciało robota wyposażonego w koła w rzeczywistości *bardzo* się różniły za wyjątkiem jednego, kluczowego szczegółu, odległości między „uszami”.

Najważniejszym i prawdopodobnie najbardziej złożonym przykładem sztucznego ucieleśnienia organizmoidalnego są roboty człękoksztaltne takie jak słynny Cog (Brooks i in. 1998). Przyjmujemy tutaj założenie, że badania z zakresu AI są w stanie odnieść się do poznania typu ludzkiego i je zgłębić, jeśli dotyczą artefaktów podobnych do człowieka. Dreyfus (1996) na przykład wskazywał, że „[odcieleśnione] sieci neuronowe różnią się na wiele istotnych sposobów od ucieleśnionych mózgów”. Przekonywał, że sieci neuronowe trzeba by „umieścić w [humanoidalnych] robotach, ponieważ brak ciała i środowiska

... stawia odcieleśnione sieci neuronalne w bardzo niekorzystnej sytuacji, jeśli chodzi o uczenie się, jak radzić sobie w ludzkim świecie. Nic nie jest bardziej obce naszej formie życia niż sieć bez orientacji góra/dół i przód/tył, bez rozróżnienia na wewnątrz/zewnątrz,... Szanse, że taka sieć nie będzie w stanie uogólniać tak jak my ... są przytłaczające.

Taki rodzaj argumentacji jest ściśle powiązany z szóstym prezentowanym przez Wilson ujęciem ucieleśnionego poznania, a także na przykład ze wspomnianymi już pracami Lakoffa i Johnsona (1980, 1999) o cieleśnych/metaforycznych podstawach pojęć abstrakcyjnych. Lakoff (1988) streścił podstawową ideę tej teorii następująco:

⁶⁶ Warto zauważyć, że chociaż żywe i sztuczne „byty organizmoidalne” można uznać za ucieleśnione fizycznie, tylko te pierwsze są z konieczności ucieleśnione historycznie, podczas gdy te drugie zazwyczaj zaprojektowali ludzie.

Znaczeniowe struktury pojęciowe pochodzą z dwóch źródeł: (1) ustrukturyzowanej natury doświadczenia cielesnego i społecznego oraz (2) naszej wrodzonej umiejętności wyobraźniowego rzutowania z pewnych dobrze ustrukturyzowanych aspektów cielesnego i interakcyjnego doświadczenia na abstrakcyjne struktury pojęciowe.

Jeśli na przykład pojęcie „uchwytywania idei” jest ugruntowane w cielesnym doświadczeniu/czynności chwytania obiektów fizycznych, to trudno spodziewać się, że robot pozbawiony chwytających ramion/rąk będzie w stanie to pojęcie zrozumieć. Powiązany z tym argument przedstawił Keijzer (1998), który podważył stosowność używania robotów na kółkach jako modelu organizmów, które poruszają się za pomocą zupełnie innych środków.

Z oczywistych powodów sztuczne ucieleśnienie wzorowane na ludzkim lub ucieleśnienie faktycznie ludzkie można uznać za szczególny przypadek o dużym znaczeniu dla kognitywistyki. Należy zauważyć jednak, że nie daje to odpowiedzi na pytanie o to, jakie dokładnie cielesne różnice pomiędzy ludźmi i innymi organizmami byłyby istotne dla poznania. To znaczy, równie dobrze może być tak, że posiadanie rąk, ramion, nóg itd. jest kluczowe dla ludzkiego (człękopodobnego) ucieleśnionego poznania, ale samo w sobie to pojęcie nie wyjaśnia nam, dlaczego tak miałyby być.

Ucieleśnienie organizmowe

Koncepcja ucieleśnienia organizmowego głosi, że poznanie nie tylko ogranicza się do fizycznych ciał organizmopodobnych, ale właściwie do organizmów, czyli do *żywych ciał*. To pojęcie ma swoje korzenie w pracach biologa teoretycznego von Uexküll (1928, 1982) i odpowiadających im współczesnych pracach Maturany i Vareli (1980, 1987) na temat biologii poznania, które głoszą, z grubsza rzecz ujmując, że poznanie jest to, co systemy żywe robią w interakcji ze swoim otoczeniem⁶⁷. Zgodnie z tym poglądem istnieją zasadnicze różnice pomiędzy organizmami żywymi, które są *autonomiczne* i *autopojetyczne*, a stworzonymi przez człowieka maszynami, które są heteronomiczne i allopojetyczne (por. Sharkey i Ziemke 2001, Ziemke i Sharkey 2001).

Von Uexküll (1928) dowodził, że zachowanie organizmu zależy od historycznie wytworzonych podstaw reakcji, to znaczy zależnych od kontekstu skłonności behawioralnych. Maszyny natomiast, przynajmniej w czasach Uexküll (1864-1944), nie miały takich podstaw reakcji, te bowiem mogły powstać jedynie w wyniku rozwoju, a maszyny nie rozwijają się. Wyjaśniał dalej, że zasady, które maszyny stosują, nie mogą się zmienić ze względu na to, że maszyny są sztywnymi strukturami. To znaczy, zasady, które kierują ich operacjami, nie są ich „własnymi”, lecz ludzkimi zasadami, które zostały w nie wbudowa-

⁶⁷ Zob. także Lemmen 1996, Zlatev 2003 oraz Stewart 1996, którzy streszczają ten pogląd jako „Poznanie = Życie”.

ne, stąd mogą je zmienić (bądź naprawić w przypadku uszkodzenia) tylko ludzie, co znaczy, że mechanizmy są *heteronomiczne*. Podsumowując, można powiedzieć, że *maszyny działają zgodnie z planem* (ludzi, którzy je zaprojektowali), podczas gdy *żywe organizmy są działającymi planami*. Dodatkowo ściśle wiąże się to z tym, co von Uexküll (1982) nazywa „najważniejszą różnicą między konstrukcją mechanizmu a konstrukcją żywego organizmu”. Polega ona mianowicie na tym, że maszyny są konstruowane dośrodkowo [*centripetally*], to znaczy z części, które trzeba wytworzyć jako pierwsze, podczas gdy organizmy „konstruuja” się odśrodkowo [*centrifugally*], to znaczy, że najpierw mamy całość, z której „rozwijają się” części.

Podobnie Maturana i Varela (1980, 1987) odróżnili organizację systemu od jego struktury. *Organizacja*, podobnie jak pojęcie planu budowy von Uexkülla (*Bauplan*), oznacza „te relacje, które muszą zachodzić między składowymi systemu, aby mógł on należeć do określonej klasy” (Maturana i Varela 1987). Systemy żywe charakteryzują się organizacją autopojetyczną. System *autopojetyczny* to szczególny rodzaj maszyny homeostatycznej, dla której podstawową zmienną wymagającą utrzymania stałej wartości jest jej własna organizacja. Natomiast *struktura* systemu oznacza „składowe i relacje, które faktycznie tworzą określoną całość i sprawiają, że jej organizacja jest rzeczywista” (Maturana i Varela 1987). Zatem struktura systemu autopojetycznego to konkretne urzeczywistnienie faktycznych składowych i faktycznych relacji pomiędzy nimi. Jego organizację konstytuują relacje pomiędzy tymi składowymi, co definiuje go jako całość określonego rodzaju. Relacje te to sieć procesów wytwarzania, które same poprzez transformacje i niszczenie wytwarzają części składowe. To właśnie interakcje i transformacje składowych stale odnawiają i realizują sieć procesów, które wytworzyły te składowe.

Zatem według Maturany i Vareli systemy żywe zupełnie nie są takie same jak wytworzone przez ludzi maszyny, takie jak samochody czy roboty, które są *allopjetyczne*. W odróżnieniu od maszyn autopojetycznych organizacja maszyn allopjetycznych jest dana poprzez konkatencję procesów. Nie są one procesami wytwarzania składowych, które określałyby maszynę jako całość. Raczej jest tak, że jej części składowe są wytwarzane w innych procesach, niezależnych od organizacji maszyny. Zatem zmiany, przez które przechodzi maszyna allopjetyczna bez utraty określającej ją organizacji, są z konieczności podporządkowane wytwarzaniu czegoś innego niż ona sama. Innymi słowy, nie jest ona prawdziwie autonomiczna, lecz jest heteronomiczna. System żywy przeciwnie, jest naprawdę autonomiczny w tym sensie, że jest maszyną autopojetyczną, której funkcją jest wytwarzanie i podtrzymywanie całości, która odróżnia ją od medium, w którym istnieje. Warto podkreślić, że pomimo odmiennej terminologii rozróżnienie Maturany i Vareli na maszyny autopojetyczne i allopjetyczne jest bardzo podobne do podziału Uexkülla (1928) na skonstruowane dośrodkowo mechanizmy i na organizmy, które są „żywymi planami” „tworzącymi” się na sposób odśrodkowy.

Jak omówiono to szczegółowo gdzie indziej (Ziemke 2001a, Ziemke i Sharkey 2001), w najnowszych badaniach nad sztuczną inteligencją i sztucznym życiem dokonał się duży postęp prac nad robotami samoorganizującymi się. Inaczej niż maszyny z czasów Uexküll'a dzisiejsze roboty adaptacyjne mogą „rozwijać się” w interakcji ze środowiskiem poprzez użycie sztucznych technik ewolucyjnych i technik uczenia się. Co więcej, ciała robotów w pewnym sensie mogą zostać przekształcone odśrodkowo z pokolenia na pokolenie (np. Lipson i Pollack 2000). Wyposażanie artefaktów w zdolność do samoorganizowania się można rozumieć jako próbę nadania im *sztucznej ontogenezy*. Niemniej próba zapewnienia im w ten sposób autonomii w powyższym silnym sensie wydaje się skazana na porażkę, ponieważ z powyżej przedstawionej argumentacji wynika, że autonomii nie można „włożyć” z zewnątrz do systemu, który już by jej nie „zawierał”. Ontogeneza utrzymuje autonomię organizacji, nie „stwarza” jej. Jednak próba wprowadzenia artefaktu w jakąś formę strukturalnej kongruencji z jego środowiskiem może się „powieść”, ale jedynie w tym sensie, w jakim kryterium kongruencji nie leży w samym heteronomicznym artefakcie, lecz musi znajdować się w perspektywie obserwatora. Zatem biorąc pod uwagę ograniczenia obecnej technologii, ucieleśnienie organizmowe jak dotąd właściwie ogranicza się do biologicznych systemów żywych, co nie wyklucza wcale możliwości powstania sztucznych systemów autopoietycznych w przyszłości, na przykład opartych na tak zwanych materiałach samonaprawiających się.

Koncepcja ucieleśnienia organizmowego ma tę zaletę, że ogranicza poznanie do naturalnych systemów, czyli do organizmów żywych. Jednak jej przydatność dla tradycyjnie rozumianej kognitywistyki jest, rzecz jasna, ograniczona w tym sensie, że poprzez zrównywanie życia z poznaniem w pewnym stopniu zwyczajnie uchyla się od odpowiedzi na pytanie o to, co sprawia, że poznanie u ludzi i zwierząt wyższych jest szczególnie interesujące.

Ucieleśnienie społeczne

Barsalou i inni (2003) posługują się pojęciem *ucieleśnienia społecznego*, przez które rozumieją:

... stany ciała, takie jak postawy, ruchy ramion i wyrazy twarzy, powstają podczas interakcji społecznych i odgrywają kluczową rolę w przetwarzaniu informacji społecznych.

Do jakiegoś stopnia omówienie ucieleśnienia społecznego ich autorstwa jest rozbieżne z poprzednimi pięcioma koncepcjami w tym sensie, że raczej wskazują oni na rolę ucieleśnienia w interakcjach społecznych, niż pytają o rodzaj ciała wymagany dla określonego typu poznania. Skupiając się na ludzkim poznaniu i interakcjach społecznych, Barsalou i inni po prostu nie zajmują się szczegółowo różnymi typami ucieleśnienia czy różnymi pojęciami ucieleśnie-

nia w powyższym znaczeniu. Niemniej jednak aspekt ten pojawia się pośrednio w ich omówieniu czterech typów „efektów ucieleśnienia”:

Po pierwsze, dostrzeżone bodźce społeczne nie tylko wytwarzają stany poznawcze, wytwarzają także stany cielesne. Po drugie, postrzeganie stanów cielesnych u innych wytwarza cielesną mimikrę u danej osoby. Po trzecie, stany cielesne danej osoby wytwarzają stany afektywne. Po czwarte, dopasowanie stanów cielesnych i stanów poznawczych modyfikuje efektywność działania [*performance effectiveness*].

Ograniczenia w objętości tekstu nie pozwalają nam odnieść się szczegółowo do kwestii powiązań i (nie)kompatybilności pojęcia ucieleśnienia społecznego z pięcioma wcześniej omówionymi pojęciami. Niemniej należy przynajmniej wspomnieć o tym, że pojęcie wiedzy, jakim posługują się Barsalou i inni, czyli wiedzy jako „częściowego modelowania [*simulations*] stanów zmysłowych, motorycznych i introspekcyjnych”, w dużym stopniu odpowiada pojęciu organizmoidalnego ucieleśnienia i wiedzy (*off-line*) jako ugruntowanej w ciele. Pytanie o to, czy ucieleśnienie społeczne wymaga żywego ciała, czy mogłoby równie dobrze urzeczywistniać się w systemach fizycznych bądź obliczeniowych, jest pytaniem, które należy postawić w przyszłych badaniach (por. również Dautenhahn 1997).

Podsumowanie i wnioski

W artykule omówiono wiele różnych pojęć ucieleśnienia, jak również odpowiednio ich wady i zalety. Towarzyszyła nam motywacja podobna do tej w artykule Wilson (2002), to jest motywacja rozwikłania różnych twierdzeń i pojęć w badanej dziedzinie (zob. także Chirsley i Ziemke 2002). Niemniej, inaczej niż u Wilson, skupiliśmy się na różnych pojęciach ucieleśnienia i na pytaniu o to, jakie dokładnie ciało uważa się za zdolne do ucieleśnionego poznania. Pojęcia, jakie wyróżniliśmy w literaturze są następujące: (1) *sprzężenie strukturalne* pomiędzy podmiotem a środowiskiem, (2) *ucieleśnienie historyczne* jako coś, co wynika z historii sprzężenia strukturalnego, (3) *ucieleśnienie fizyczne*, (4) *ucieleśnienie „organizmoidalne”* czyli dotyczące organizmopodobnych form cielesnych, (5) *ucieleśnienie organizmowe* autopojetycznych, żywych systemów i wreszcie (6) *ucieleśnienie społeczne*.

Rzecz jasna, ten krótki artykuł ma poważne ograniczenia: pojęcie ucieleśnienia organizmowego, najbardziej restrykcyjne ze wszystkich omawianych tutaj pojęć (a przynajmniej z pierwszych pięciu), w rzeczywistości stosuje się do wszystkich systemów żywych, co wcale nie jest szczególnie restrykcyjne. Humanoidalne i ludzkie ucieleśnienie można by rozumieć w bardziej restrykcyjny sposób, jako szczególne przypadki odpowiednio ucieleśnienia organizmoidalnego i organizmowego. Zwłaszcza w połączeniu z pojęciem ucieleśnienia społecznego mogłoby to być szczególnie istotne dla kognitywistyki, ale

nie zostały tutaj przedstawione żadne argumenty za tym, że specyficzne przypadki mogłyby czy powinny pozwolić na wyróżnienie znacząco odmiennych typów ucieleśnionego poznania bardziej niż przedstawiciele kategorii ogólniejszych.

Niemniej jednak mamy nadzieję, że przedstawione rozróżnienia pomogą rozwikłać ogromną różnorodność twierdzeń, pojęć i teorii, którymi obecnie odznaczają się badania nad ucieleśnionym poznaniem.

Podziękowania

Autor chciałby podziękować Jordanowi Zlatevowi, Noelowi Sharkey'owi, Ronowi Chrisley'owi, Jessice Lindblom, Henrikowi Svenssonowi i Stefanowi Berglundowi za interesujące dyskusje, które bardzo wzbogaciły ten artykuł oraz anonimowym recenzentom za wiele użytecznych sugestii, chociaż nie wszystkie z nich dało się uwzględnić w tym krótkim artykule.

Literatura

- Barsalou, L.; Niedenthal, P, Barbey, A. & Ruppert, J. 2003. *Social embodiment*. B. Ross, ed. *The Psychology of Learning and Motivation*, vol. 43. San Diego, CA: Academic Press.
- Brooks, R. 1990. Elephants don't play chess. *Robotics and Autonomous Systems*, 6(1-2): 1-16.
- Brooks, R. 1991. Intelligence Without Reason. *Proc. of the Twelfth Intl. Joint Conf. on Artificial Intelligence*. San Mateo, CA: Morgan Kaufmann.
- Brooks, R.; Breazeal, C.; Marjanovic, M.; Scasselati, B. & Williamson, M. 1998. The Cog Project: Building a Humanoid Robot. Ch. Nehaniv, ed. *Computation for metaphors, analogy, and agents*. New York: Springer.
- Chrisley, R. 2003. Embodied Artificial Intelligence. *Artificial Intelligence* 149: 131-150.
- Chrisley, R. i Ziemke, T. 2002. Embodiment. In *Encyclopedia of Cognitive Science*. Macmillan.
- Clancey, W. 1997. *Situated Cognition: On Human Knowledge and Computer Representations*. New York: Cambridge University Press.
- Clark, A. 1997. *Being There*. Cambridge: MIT Press.
- Clark, A. 1999. An embodied cognitive science? *Trends in Cognitive Science* 9: 345-351.
- Dautenhahn, K. 1997. I could be you – The phenomenological dimension of social understanding. *Cybernetics and Systems*, 25(8): 417-453.
- Dautenhahn, K.; Ogden, B. & Quick, T. 2002. From embodied to socially embedded agents. *Cognitive Systems Research*, 3(3): 397-428.
- Dreyfus, H. 1996. The Current Relevance of Merleau-Ponty's Phenomenology of Embodiment. *The Electronic Journal of Analytic Philosophy* 4.

- Franklin, S. 1997. Autonomous agents as embodied AI. *Cybernetics and Systems*, 25(8): 499-520.
- Hutchins, E. 1995. *Cognition in the Wild*. Cambridge, MA: MIT Press.
- Keijzer, F. 1998. Armchair considerations about wheeled behavior. R. Pfeifer *et al.* red. *From Animals to Animats 5*. Cambridge, MA: MIT Press.
- Kirsh, D. i Maglio, P. 1994. On distinguishing epistemic from pragmatic action. *Cognitive Science*, 18: 513-549.
- Lakoff, G. 1988. Cognitive Semantics. U. Eco *et al.* red. *Meaning and Mental Representations*. Bloomington: Indiana University Press.
- Lakoff, G. & Johnson, M. 1980. *Metaphors we live by*. Chicago: University of Chicago Press. [wydanie polskie *Metafory w naszym życiu*. Przeł.. T. Krzeszkowski. Warszawa: PIW 1988]
- Lakoff, G. i Johnson, M. 1999. *Philosophy in the flesh: The embodied mind and its challenge to western thought*. New York: Basic Books.
- Lemmen, R. 1996. Embodiment for Non-Cartesians. *Conf. of the Europ. Soc. of Philosophy and Psychology*. CSRP 441, COGS, Univ. of Sussex, UK.
- Lipson, H. & Pollack, J. B. 2000. Automatic design and manufacture of robotic lifeforms. *Nature*, 406: 974-978 .
- Lund, H.; Webb, B. & Hallam, J. 1998. Physical and temporal scaling considerations in a robot model of cricket calling song preference. *Art. Life* , 4: 95-107.
- Maturana, H. i Varela, F. 1980. *Autopoiesis and Cognition*. Dordrecht, The Netherlands: Reidel.
- Maturana, H. i Varela, F. 1987. *The Tree of Knowledge*. Boston, MA: Shambhala.
- Quick, T.; Dautenhahn, K.; Nehaniv, C. & Roberts, G. 1999. On Bots and Bacteria: Ontology Independent Embodiment. In *Proceedings of the Fifth European Conference on Artificial Life*. Heidelberg: Springer.
- Pfeifer, R. i Scheier, C. 1999. *Understanding Intelligence*. Cambridge, MA: MIT Press.
- Riegler, A. 2002. When is a Cognitive System Embodied? *Cognitive Systems Res.*, 3(3): 339-348.
- Sharkey, N. i Ziemke, T. 1998. A consideration of the biological and psychological foundations of autonomous robotics. *Connection Sci.*, 10: 361-391.
- Sharkey, N. i Ziemke, T. 2001. Mechanistic vs. phenomenal embodiment - Can robot embodiment lead to strong AI? *Cogn. Systems Res.*, 2(4): 251-262.
- Stewart, J. 1996. Cognition = Life. *Behavioral Processes*, 35: 311-326.
- Susi, T.; Lindblom, J. & Ziemke, T. 2003. Beyond the Bounds of Cognition. K. Forbus, D. Gentner, T. Regier, eds. *Proceedings of the 25th Annual Conference of the Cognitive Science Society*: 1305-1310. Mahwah, NJ: Lawrence Erlbaum.
- Varela, F.; Thompson, E. Rosch, E. 1991. *The Embodied Mind*. Cambridge, MA: MIT Press.
- von Uexküll, J. 1928. *Theoretische Biologie*. Berlin: Springer Verlag.

- von Uexküll, J. 1982. The Theory of Meaning. *Semiotica*, 42(1), 25-82.
- Wilson, M. 2002. Six views of embodied cognition. *Psychological Bulletin and Review*, 9(4): 625-636.
- Ziemke, T. 1999. Rethinking Grounding. A. Riegler et al. eds. *Understanding Representation in the Cognitive Science*. Plenum Press: New York.
- Ziemke, T. 2001a. The Construction of 'Reality' in the Robot. *Foundations of Science*, 6(1): 163-233.
- Ziemke, T. 2001b. Disentangling Notions of Embodiment. *Workshop on Developmental Embodied Cognition*, Edinburgh, UK, July 2001.
- Ziemke, T. 2001c. Are Robots Embodied?. *Proc. Of the First Intl. Workshop on Epigenetic Robotic*. Lund University Cognitive Studies, vol. 85, Lund, Sweden.
- Ziemke, T. (ed.) 2002. Situated and Embodied Cognition (special issue). *Cognitive Systems Research*, 3(3).
- Ziemke, T. i Sharkey, N. E. 2001. A stroll through the worlds of robots and animals. *Semiotica*, 134(1-4): 701-746.
- Zlatev, J. 1997. *Situated Embodiment. Studies in the Emergence of Spatial Meaning*. Stockholm: Gotab.
- Zlatev, J. 2001. The epigenesis of meaning in human beings, and possibly in robots. *Minds and Machines*, 11: 155-195.
- Zlatev, J. 2003. Meaning = Life (+ Culture): An Outline of a Unified Biocultural Theory of meaning. *Evolution of Communication* 4(2): 253-296.

What's that Thing Called Embodiment?

Abstract: Embodiment has become an important concept in many areas of cognitive science. There are, however, very different notions of exactly what embodiment is and what kind of body is required for what type of embodied cognition. Hence, while many nowadays would agree that humans are embodied cognizers, there is much less agreement on what kind of artifact could be considered embodied. This paper identifies and contrasts six different notions of embodiment which can roughly be characterized as (1) structural coupling between agent and environment, (2) historical embodiment as the result of a history of structural coupling, (3) physical embodiment, (4) organismoid embodiment, i.e. organism-like bodily form (e.g., humanoid robots), (5) organismic embodiment of autopoietic, living systems, and (6) social embodiment.

Keywords: cognitive science; cognition; embodiment; agent; environment; structural coupling; physical; organism; AI; robot; social.