



Tytuł: **Superweniencja semantyczna, koneksjonizm a badania nad SI**

Autor: Piotr Kołodziejczyk ; e-mail: pkolodziejczyk@interia.pl

Źródło: <http://kognitywistyka.net> ; e-mail: mjkasperski@kognitywistyka.net

Data: listopad/grudzień 2003

1. Uwagi wstępne

Artykuł ten jest rozszerzeniem analiz prowadzonych przeze mnie w artykule *Obliczanie, semantyka i superweniencja*. W pracy tej zaproponowałem koncepcję semantyki obliczeniowej opartej na relacji superweniencji. Nie odróżniłem tam jednak logicznego i psychologicznego poziomu prowadzonych analiz. Brak odróżnienia tego rodzaju mógł prowadzić do szeregu nieporozumień, ponieważ

biorąc pod uwagę istnienie wielu wyspecjalizowanych wariantów relacji superweniencji, posługiwanie się jednym, bardzo ogólnym jej pojęciem może nie być wystarczająco instruktywne.¹

Doprecyzowując tezy wygłoszone w *Obliczaniu...*, można zatem stwierdzić, że relacja superweniencji miała, w proponowanym rozumieniu, charakter semantyczny. Znaczy to, że własności syntaktyczne danego systemu traktowałem jako relacje bazowe, natomiast nadbudowane nad nimi własności semantyczne – jako własności superweniencjne. Jest jasne, że relacja syntaktyki i semantyki zawierała się w ogólnym schemacie relacji superweniencji², gdyż spełnia ona warunek jednostronnej współzmienności, własności semantyczne są determinowane przez własności syntaktyczne systemu, oraz własności semantyczne nie są redukowalne do bazowych własności syntaktycznych.

Opierając się na wynikach uzyskanych w artykule *Obliczanie...*, w poniższych analizach podejmę próbę ukazania zależności pomiędzy logicznym i psychologicznym rozumieniem relacji superweniencji. Postaram się także wskazać, że precyzyjne ustalenie rodzaju tych zależności ma fundamentalne skutki dla badań nad AI. Umożliwia, bowiem ono jednoznaczną konstrukcję konekcyjnych systemów przetwarzających informacje.

¹ R. Poczobut, *Superweniencja. Zarys problematyki*, w: „Filozofia Nauki”, Nr 2/2000, s. 20.

² Zob. D. Lewis, *On the Plurality of Worlds*, Oxford, 1984, s. 14.



2. Logiczna charakterystyka relacji superweniencji semantycznej

Poniższe analizy ograniczę wyłącznie do superweniencji semantycznej, ponieważ w sztucznych i naturalnych systemach przetwarzających informacje najbardziej istotne jest dla mnie ukazanie relacji, jakie panują między syntaktycznym a semantycznym poziomem przetwarzania danych dostarczonych do systemu. Wychodząc od badań prowadzonych niezależnie przez Luca Bovensa³ i Roberta Poczobuta⁴, główne założenia idei superweniencji semantycznej można przedstawić w postaci następujących tez wraz z ich rozwinięciami:

1. zasada koincydencji: takie same własności syntaktyczne systemu implikują identyczne własności semantyczne, a zatem: jeżeli dany system A posiada takie same charakterystyki syntaktyczne jak system B , to obydwa systemy posiadają takie same własności semantyczne;
2. zasada różnicy: gdyby system A posiadał inne własności syntaktyczne niż dany system B , to systemy te musiałyby posiadać różne własności semantyczne;
3. zasada determinacji: jeżeli dowolny system przetwarzający informacje posiada własności semantyczne, to musi on również posiadać pewne bazowe własności syntaktyczne.

Stosując formalizację McFertidge'a zasady (1), (2), (3) można przedstawić następująco:

- ◆ zasadę koincydencji: $\forall_x \forall_y \forall_z [\forall A (A_{yx} \equiv A_{zx}) \rightarrow \forall B (B_{yz} \equiv B_{zx})]$,
- ◆ zasadę różnicy: $\forall_x \forall_y \forall_z [\exists A (A_{yx} \wedge \neg A_{zx}) \rightarrow \exists B (B_{yz} \wedge \neg B_{zx})]$,
- ◆ zasadę determinacji: $\forall_x \forall_y \forall A [A_{yx} \rightarrow \exists B (B_{yx} \wedge \forall_y (A_{yx} \rightarrow B_{yx}))]$,

przy czym A oraz B są systemami przetwarzania informacji, w których zachodzą relacje: między własnością y oraz dostarczoną daną x (charakterystyka bazowa) i własnością z oraz daną dostarczoną x (własność superwenienna).

Przenosząc powyższe ustalenia w obszar badań nad AI okazuje się, że w przypadku sztucznych systemów poznawczych relacja superweniencji zawiera się w schemacie słabej wersji relacji superweniencji. Przeformułując stwierdzenie Wiliama Seagera⁵ ideę słabej superweniencji semantycznej można wysłowić następująco: dla każdej własności semantycznej A z logicznej konieczności istnieje własność syntaktyczna B taka, że każdy system egzemplifikujący własność B z konieczności egzemplifikuje także własność A . Superweniencję słabą można za Jaewongiem Kimem⁶ wyrazić w postaci formuły:

$$\forall_x \forall_y \forall_z \{(y \subset x \wedge z \subset x) \rightarrow [\forall A (A_y \rightarrow A_z) \rightarrow \forall B (B_y \rightarrow B_z)]\}$$

Mówiąc inaczej, dla dowolnej własności syntaktycznej y oraz dowolnej własności semantycznej z , jeśli są one składowymi modułu x należącego do systemów A oraz B , to nieodróżnialność własności syntaktycznej y implikuje nieodróżnialność własności semantycznej z w systemie A i B . Zależność tą będę nazywał **superwienienną**

³ Por. L. Bovens, *Principles of Supervenience*, ss. 294-301, w: „Australian Journal of Philosophy”, Nr 3/1994, s. 294.

⁴ Por. R. Poczobut, *Superweniencja...*, ss. 7-9.

⁵ Por. W. Seager, *The Metaphysics of Consciousness*, London, 1991, s. 116.

⁶ Zob. J. Kim, *Concepts of Supervenience*, ss. 153-176, w: „Philosophy and Phenomenological Research”, Nr 2/1984, s. 172.



izomorficznością semantyczną. W przypadku semantyki, jaką zaproponowałem w *Obliczaniu...*, zależność ta przejawia się na poziomie analizy interakcji systemu ze środowiskiem zewnętrznym. Wbrew eksternalistom skrajnym (np. Tylor Burge) twierdzącymi, że

treść stanów intencjonalnych zależy nie tylko od fizycznego stanu mózgu, ale również od relacji do środowiska⁷

można wykazać, że własności semantyczne nie tyle zależą od szerokiej treści dostarczanych do systemu danych, ale od sposobu ich kodowania i przetwarzania. Stwierdzenie to łatwo ująć w postaci następującego algorytmu⁸:

(1). Niech S będzie zbiorem zaimplementowanych w systemie symboli, zaś W – zbiorem własności semantycznych superwenujących na S .

(2). Transponując tezę Kima o korelacji psychofizycznej można stwierdzić, że dla każdej własności W istnieje obliczeniowa konfiguracja S taka, że własność W występuje w danym systemie jeżeli występuje w nim również konfiguracja S :

$$\forall S \exists W (S \rightarrow W).$$

(3). Tezę tą można rozszerzyć poprzez włączenie w jej ramy procedur dekodujących i uzyskać następującą charakterystykę własności semantycznych:

(a). Dana konfiguracja obliczeniowa może być dekodowana przez różne procedury. Stąd też, pewna konfiguracja obliczeniowa S może być zinterpretowana w różnych znaczeniach:

$$\forall S \exists W \exists D \exists P \{P \neq D \rightarrow [(S(P) \rightarrow W_1) \wedge (S(D) \rightarrow W_2)]\}.$$

Superwenienna izomorficzność semantyczna może, więc stanowić uzasadnienie dla funkcjonalistycznego programu badania umysłu. Skoro semantyczne charakterystyki systemu poznawczego są warunkowane istnieniem poziomu syntaktycznego (obliczeniowego) wraz ze zbiorem procedur dekodujących, to można się zgodzić z Jerrym Fodorem, że

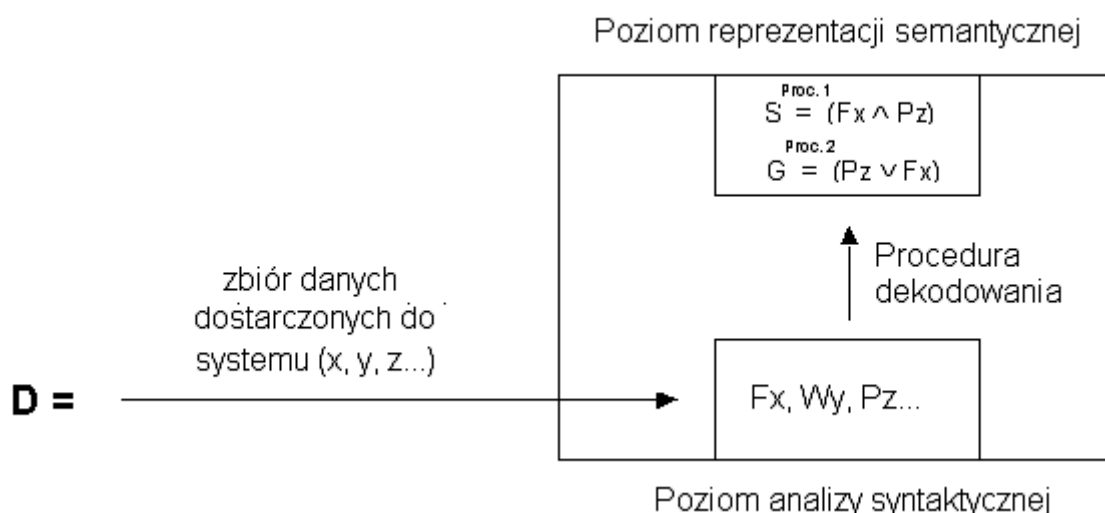
dla egzemplarzy stanów psychicznych (w moim przypadku stanów syntaktycznych – P. K.) mogą być wyznaczone typy stanów psychicznych (stanów semantycznych – P. K.) wyłącznie przez odniesienie do przyczynowych relacji (procedur dekodujących – P. K.), w jakich pozostają one do bodźców proksymalnych i do reakcji proksymalnych oraz przez odniesienie do relacji, jakie zachodzą między tymi stanami.⁹

Podejście to trafnie przedstawia następujący schemat:

⁷ R. Poczobut, *Superweniencja...*, ss. 10-11.

⁸ Por. J. Kim, *Psychophysical supervenience*, ss. 51-70, w: „Philosophical Studies”, Nr 1/1982, ss. 54-58.

⁹ J. Fodor, *Jak grać w reprezentacje umysłowe – poradnik Fodora*, przeł. A. Putko, s. 25, w: *Modele umysłu*, red. Z. Chlewiński, Warszawa 1999, ss. 17-49.



Schemat ten wskazuje na wzajemny związek szerokiej i wąskiej treści w przypadku sztucznych systemów przetwarzających informacje. Własności semantyczne stanowią w tym ujęciu funkcję syntaktycznych stanów systemu. Te zaś są warunkowane danymi dostarczonymi do systemu, które cechują eksternalne charakterystyki. Idea superweniencji semantycznej jest, zatem mocno związana z koncepcją superweniencji psychologicznej (psychofizycznej). Analiza tego związku stanowi tematykę poniższych rozważań.

3. Superweniencja psychologiczna a badania nad AI¹⁰

Podobnie jak w przypadku superweniencji semantycznej, również superweniencja psychologiczna bazuje na idei zmiany zależnej, którą można przedstawić następująco: dany układ nie może zmienić się pod względem własności superweniencyjnych, jeżeli zmiana nie zachodzi na poziomie własności bazowych. W przypadku sztucznych i naturalnych podmiotów poznawczych, o zmianach mówić należy w odniesieniu do problemów intencjonalności i treści postaw propozycjonalnych. Na zasadność ograniczenia rozważań do zasygnalizowanej problematyki wskazuje stwierdzenie Daniela Dennetta:

cała intencjonalność jaką możemy się cieszyć, jest pochodna od bardziej fundamentalnej intencjonalności miliardów elementów systemów intencjonalnych.¹¹

Nie znaczy to oczywiście, że własności psychiczne danego systemu są redukowalne do własności obliczeniowych, fizycznych itd. Pozostając w zgodzie z paradygmatem sztucznej inteligencji oraz zasadą superweniencji można wnosić, że własności psychiczne oraz treści stanów mentalnych bazują na podstawowych kategoriach danego systemu. W przypadku sztucznych podmiotów poznawczych kategoriami tymi są elementy syntaktyczne (obliczeniowe). Takie podejście warunkowane jest przyjęciem **zasady obliczeniowej nieodróżnialności**, stanowiącej trawestację podejścia Kima do zagadnienia superweniencji psychofizycznej¹². Zasadę tą można wyrazić następująco: jeżeli dwa systemy posiadające identyczną strukturę syntaktyczną, mają dostarczone takie same dane ze środowiska

¹⁰ Wiele spośród poniższych ustaleń zawdzięczam dyskusjom prowadzonym z panią Izabelą Sobol. Dziękuję jej za tego rodzaju pomoc.

¹¹ D. Dennett, *Natura umysłów*, tłum. W. Turopolski, Warszawa 1997, s. 70.

¹² Por. J. Kim, *Psychophysical...*, s. 58.



zewnątrznego, to posiadają one (systemy) takie same treści stanów wewnętrznych, nie różnią się zatem pod względem własności semantycznych oraz swoich odniesień intencjonalnych.

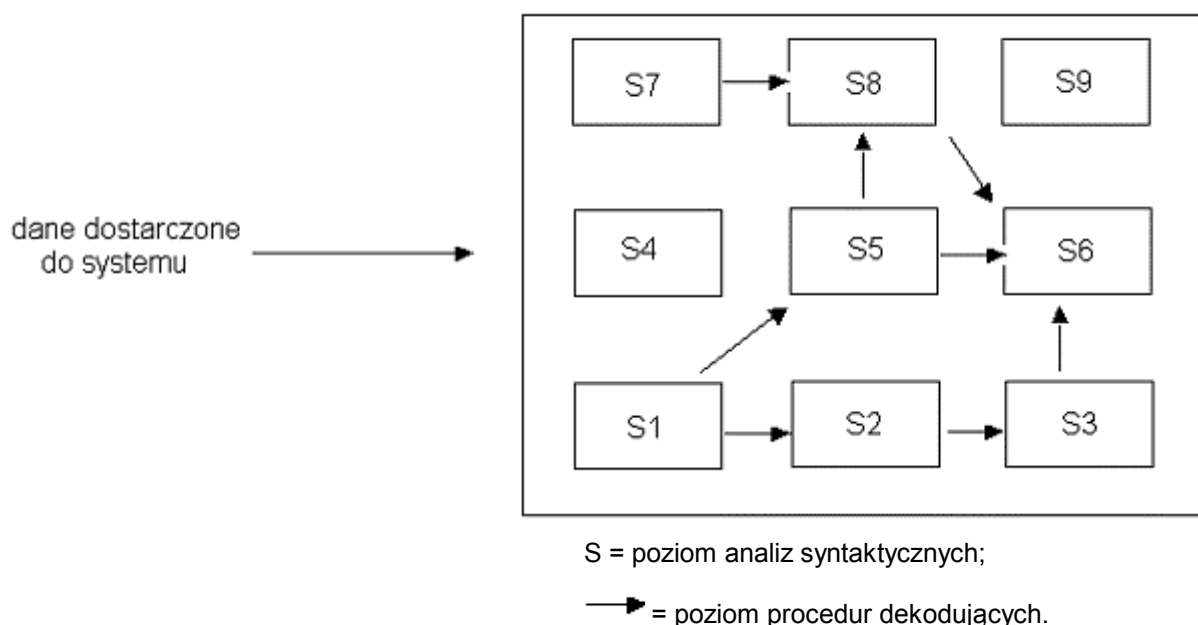
Podejście to potwierdzają analizy przeprowadzone przez Fodora. Na podstawie twierdzeń ogłoszonych przez Fodora w *Psychosemantics*¹³. Dwa systemy posiadają identyczne stany wewnętrzne wtedy i tylko wtedy, gdy stany te są realizowane na podstawie identycznych własności bazowych. Stąd zaś można wnosić, że na poziomie behawioralnym reakcje tych systemów są nieodróżnialne. Zatem wewnętrzne stany tych systemów pełnią rolę procesów pośredniczących pomiędzy danymi wejściowymi a intencjonalnymi reakcjami na te dane. Ukryta w tym stwierdzeniu przesłanka o internalistycznej proveniencji wymusza pytanie o charakter powiązań pomiędzy wewnętrznymi stanami danego systemu. Wnioski uzyskane w powyższych rozważaniach wymuszają tezę o koneksyjnej charakterystyce sztucznych podmiotów poznawczych.

4. Superweniencja a koneksjonizm

Zasadnicza teza koneksyjnego podejścia do badania umysłu może być za Davidem Rumelhartem wysłowiona następująco:

Podstawowa strategia koneksjonizmu jest następująca. Wychodzimy z założenia, że (...) elementy przetwarzania [informacji – P. K.] połączone są przewodami łączącymi. Porozumiewają się ze sobą przesyłając tymi łączami pewne dane liczbowe (obliczeniowe – P. K.).¹⁴

Tezę tą dobrze ilustruje następujący schemat:



¹³ Por. J. Fodor, *Psychosemantics*, Cambridge Mass., 1987, s. 30.

¹⁴ D. Rumelhart, *Architektura umysłu. Podejście koneksyjne*, tłum. H. Grzegółowska-Klerkowska, ss. 240-272, w: *Modele...*, red. Z. Chlewiński, s. 242.



Na podstawie schematu łatwo wykazać związki zasady koneksjonizmu zarówno z superweniencją semantyczną, jak i psychologiczną. Niezależnie od tego, czy stan typu $G = (S1 \wedge S4) \rightarrow S5$ traktuje się jako własność semantyczną, bądź stan mentalny danego systemu, własności wyższego rzędu (w sensie logicznym i przyczynowym) superwenują (są pochodną) bazowych, obliczeniowych stanów systemu. Zatem rysuje się tutaj teza o niemonadycznym charakterze własności semantycznych i mentalnych danego systemu. Logiczny i przyczynowy charakter relacji superweniencji traktuję za Fodorem¹⁵ w sposób izomorficzny. Znaczący to, że relacja superweniencji odzwierciedla zarówno syntaktyczne jak i semantyczne własności swych członów. Ponadto relacja ta charakteryzuje się własnością inferencyjności, która zapewnia automatyzację wnioskowań dokonywanych w ramach sztucznych systemów poznawczych. Dla wyjaśnienia tej tezy posłużę się następującą parą zdań logicznych: $p =$ pada deszcz, $q =$ jest zimno. Widać, że zdania te mogą pozostawać ze sobą w różnych związkach logicznych, na przykład: pada deszcz lub jest zimno; jeśli pada deszcz, to jest zimno, co na poziomie psychologicznym może prowadzić do posiadania przez system zbioru przekonań bezpośrednio lub pośrednio wynikających z tych zdań, np. system jest przekonany, że jeśli jest zimno, to nie jest ciepło, a kiedy nie jest ciepło, to temperatura jest niska, zaś gdy temperatura jest niska, to ludzie nie wychodzą z domu. Przyjmując, że sieć przekonań systemu superwenuje na zdaniach logicznych dostarczonych w jego ramy, należy więc uznać, że relacja superweniencji jest fundamentalna w eksplikacji działania i zachowania systemów AI.

5. Uwagi końcowe

Powyższe analizy nie są pełne. Traktuję je raczej jako wstęp do opracowywanej przeze mnie teorii umysłu, którą nazywam psychologiką. Zasadnicze elementy tej teorii zostały jednak wyłożone w tym artykule. Kwestie szczegółowe – na przykład: problem ujęcia kognitywnych stanów poznawczych przez systemy AI, intencjonalność sztucznych systemów poznawczych, czy problem opozycji internalizm–eksternalizm – będzie, mam nadzieję, tematem książkowego opracowania.

¹⁵ Zob. J. Fodor, *Jak grać...*, s. 31.