

Model kognitywny zdolności językowych

Piotr Konderak

kondorp@bacon.umcs.lublin.pl

1 stycznia 2013

Rozdział 1

Modelowanie poznawcze i architektury poznawcze

1.1 Modelowanie kognitywne

1.1.1 Modele poznawcze a badania nad SI

Istnieją dwa podejścia do badań nad sztuczną inteligencją (SI): pierwsze z nich dziedzinę tę traktuje „inżyniersko”, a wówczas celem badaczy jest tworzenie inteligentnych maszyn; drugie, również rozumie ją jako jedną z nauk empirycznych, której zadaniem jest obliczeniowe modelowanie ludzkiej inteligencji. Początkowo rozwój tych dwu dziedzin badawczych przebiegał razem, ale wraz z osiągnięciami SI coraz wyraźniej rysowało się powyższe rozróżnienie, przy czym pierwsze z wymienionych ujęć stało się domeną badań nad sztuczną inteligencją, podczas gdy drugie z nich przejęły nauki o poznaniu.

Modele poznawcze są teoriami dotyczącymi poznania ludzkiego. Przybierają one postać programów komputerowych. Celem budowy takich modeli/programów jest wyjaśnianie określonych procesów poznawczych – pewnych aspektów poznania w oparciu o proste mechanizmy obliczeniowe. Oczekuje się, iż wynik działania takich modeli/programów może stanowić podstawę do formułowania przewidywań dotyczących zachowania człowieka. Modele poznawcze mają ułatwić poszukiwanie odpowiedzi na problemy formułowane zazwyczaj w ramach psychologii poznawczej, jednak podstaw funkcjonowania takich modeli – elementarnych mechanizmów obliczeniowych i ich zestawień – dostarczyć mają badania w ramach sztucznej inteligencji.

Modele poznawcze przypominają zatem pod wieloma względami programy SI. Stawiane są przed nimi podobne zadania: w ramach badań nad sztuczną inteligencją budowane są inteligentne

maszyny, zaś celem modelowania poznawczego jest tworzenie modeli ludzkich mechanizmów mentalnych. Obydwa projekty różnią się jednak przyjmowanymi kryteriami powodzenia podejmowanego przedsięwzięcia. O ile w wypadku zaimplementowanych programów sztucznej inteligencji badaczy interesuje to, w jakim stopniu maszyna wykonuje poprawnie swoje zadanie, o tyle w przypadku modelowania poznawczego to kryterium nie wystarcza. Oprócz oceny tego, czy program jest w stanie wykonać swoje zadanie ocenia się również stopień, w jakim zachowanie programu przypomina zachowanie człowieka wykonującego to zadanie. Dodatkowym warunkiem nakładanym na programy będące modelami poznawczymi jest wykorzystanie przez nie mechanizmów, które stanowią teoretyczną podstawę wyjaśnienia ludzkich procesów poznawczych.¹

1.1.2 Modelowanie poznawcze w psychologii kognitywnej

Wśród badaczy zajmujących się modelowaniem poznawczym istnieją rozbieżności dotyczące rozumienia statusu tworzonych modeli, a w szczególności relacji pomiędzy modelem poznawczym a „rzeczywistym”, ludzkim poznaniem. Z jednej strony, pojawiają się stwierdzenia (np. Anderson [Anderson 1993]), iż modele poznawcze oraz przyjmowane w ich ramach architektury poznawcze **opisują** strukturę ludzkiego poznania (choć opis ten jest jedynie hipotezą). Z drugiej strony, są badacze (np. [Roth, Woods 1995]) twierdzący, że tworzone przez nich modele poznawcze **nie są** teoriami (ludzkiego) poznania. Od przyjętego podejścia uzależniona jest ocena modeli poznawczych: wynik działania modelu odbiegający od oczekiwań (opartych na. na eksperymentalnych badaniach psychologicznych) można traktować jako świadectwo za odrzuceniem konkretnego modelu poznawczego, bądź jako efekt potraktowania modelu jako użytecznej aproksymacji.

Źródłem modelowania kognitywnego pierwszego z wymienionych wyżej typów jest rozwój psychologii kognitywnej. Jej celem jest bowiem skonstruowanie adekwatnej teorii procesów poznawczych leżących u podstaw ludzkiego zachowania. Teoria taka powinna wyjaśniać obserwowalne zachowanie w kategoriach możliwych (w świetle badań nad poznaniem) procesów wewnętrznych. Wraz z rozwojem psychologii eksperymentalnej pojawiła się potrzeba precyzyjnego formułowania teorii na tyle, by wykazać, że przyjęte pojęcia wystarczają do wyjaśniania ludzkiego zachowania oraz pozwalają formułować (testowalne) przewidywania jego dotyczące. Podstawą większości teorii z zakresu psychologii poznawczej jest tzw. *Standardowa Teoria Poznania* opierająca się na rozczłonkowaniu umysłu ludzkiego na takie podstawowe komponenty jak: pamięci

¹W badaniach SI takie mechanizmy mogą zostać odrzucone jako niewydajne lub zbyt złożone obliczeniowo w stosunku do zadania, jakie mają wykonać.

krótkotrwała i długotrwała, procesy rozpoznawania, procesy pozwalające na przechowywanie w pamięci i przywoływanie z niej oraz procesy przetwarzające zgromadzone w pamięci informacje.

Wraz ze zmianą zainteresowań psychologów kognitywnych z badania prostych procesów uczenia się i percepcji na badanie bardziej złożonych procesów, np. wnioskowania czy pojmowania mowy, rozwijane w latach 50 i 60. w psychologii kognitywnej modele matematyczne okazywały się nieadekwatne, bowiem zbyt statycznie charakteryzowały zachowanie uwzględniając zbyt małą liczbę zmiennych matematycznych i dostarczając jedynie statystycznego opisu zachowania [Kieras 1981]. Ich rolę przejęły realizowalne maszynowo programy komputerowe. Za ich pomocą procesy poznawcze mogły być przedstawiane w sposób dynamiczny, jako przebiegające w czasie procesy obliczeniowe, tj. stopniowo realizujące się procedury obliczeniowe (algorytmiczne lub heurystyczne). Jednym z ważniejszych etapów w historii modelowania poznawczego było wykorzystanie znanego z badań nad sztuczną inteligencją formalizmu sieci semantycznych. Doniosłość tego kroku polegała na powiązaniu pojęć rozwijanych w ramach SI z pojęciami psychologii (w tym przypadku z klasycznym pojęciem asocjacji). [Anderson, Bower 1973] Współcześnie panuje niemal powszechna zgoda, że modele kognitywne najlepiej reprezentują kluczowe pojęcia teoretyczne psychologii kognitywnej, oraz – biorąc pod uwagę zaawansowanie nauk szczegółowych, z których psychologia poznawcza czerpie, w tym neuronauki – realizowalne komputerowo modele poznawcze stanowią najlepszy sposób przedstawienia teorii psychologicznej. Szczególnie wiele uwagi obecnie poświęca się modelowaniu poznawczemu takich dziedzin ludzkiego poznania, jak rozwiązywanie problemów, zadania związane z pamięcią, współdziałanie człowiek-maszyna czy też umiejętności językowe.

1.1.3 Kryteria oceny modeli poznawczych

O wartości modelu poznawczego świadczy zarówno zdolność naśladowania obserwowalnego zachowania ludzi, jak też jakość modelu traktowanego jako fragment teorii. Modele kognitywne – jak wyżej wspominałem – tworzono, by wykazać, iż w oparciu o postulowaną teorię uzyska się pożądane zachowanie podmiotu oraz by oszacować precyzyjność takich teorii. Budowanie modeli poznawczych wymaga odwoływania się do wyników badań kilku nauk: psychologii, lingwistyki, neurobiologii i in. Wyniki obserwacji i eksperymentów uzyskane w ramach tych nauk nakładają pewne ograniczenia na modele kognitywne. Model musi spełnić podstawowy wymóg funkcjonalności, tj. musi w pewnym przybliżeniu realizować czynność poznawczą, jaką ma modelować.

Wiadomo, że nawet jeżeli badacze budując model opierają się na możliwie dużej ilości dostęp-

nych danych empirycznych, zawsze będzie pojawiał się problem niedookreślenia teorii przez te dane. W przypadku modelowania kognitywnego (czy też szerzej: w przypadku nauk o poznaniu) problem ten jest szczególnie odczuwalny. Jak zauważa Lewis [MITECS 2001, s. 141] dzieje się tak z dwóch powodów: po pierwsze, tworząc modele poznawcze trzeba dokonywać licznych szczegółowych wyborów dotyczących przetwarzania, pomimo całkowitego braku danych empirycznych wskazujących na którąkolwiek z alternatyw; po drugie, biorąc pod uwagę uniwersalność schematów obliczeniowych można stworzyć nieskończoną liczbę programów naśladujących dane zachowanie². Można, zatem, modelując umiejętności językowe wybierać pomiędzy systemami reprezentacji wiedzy takimi, jak systemy produkcji, sieci semantyczne czy reprezentacja logiczna. Każde z tych podejść pozwala wyjaśnić pewne aspekty zdolności językowych, każde też zyskało (częściowe) wsparcie ze strony wyników badań psychologicznych. Model przetwarzania języka naturalnego Kierasa z roku 1981 wykorzystujący analizator składniowy ATN i sieci semantyczne wraz z mechanizmami przeszukiwania pamięci opartymi o rozprzestrzeniającą się aktywację [*spreading activation*] przewidywał z dużą dokładnością ilość czasu potrzebną do przeczytania ze zrozumieniem zdań w prostych ustępach. Model z roku 1982, z kolei, pokazywał jak procesy pojmowania wyższego poziomu można reprezentować za pomocą systemów produkcji wyposażonych w mechanizm wnioskowania na podstawie propozycjonalnej reprezentacji zawartości tekstu. [Kieras 1992, s. 180]

Próbując zaradzić problemowi niedookreślenia modelu przez dane obserwacyjne wprowadzono różne poziomy abstrakcji w ramach modelu poznawczego i – w konsekwencji – uznano, że nie wszystkie szczegóły modelu obliczeniowego stanowią część teorii. Powraca tutaj problem relacji między modelem poznawczym a ludzkim poznaniem. Najslabszym rodzajem odpowiedności między nimi jest zgodność na poziomie wejścia-wyjścia. Ocenie podlega zatem zgodność (w pewnym, określonym z góry, przybliżeniu) wyników działania modelu poznawczego z wynikami uzyskanymi w ramach badań nad ludzkim poznaniem, przy tych samych (zbliżonych) danych wejściowych. Innymi słowy, model spełnia jedynie wymagania funkcjonalności. Silniejsza odpowiedność między modelem a ludzkim poznaniem wymaga by model realizował ten sam algorytm, jaki realizuje ludzki system poznawczy.

Drugim kryterium oceny modeli kognitywnych jest integralność architekuralna modelu. Model powinien wykorzystywać zbiór jasno sprecyzowanych mechanizmów obliczeniowych stanowiących podstawę danej architektury kognitywnej, unikając mechanizmów arbitralnych, *ad hoc*.

²Jest to pewna odmiana *paradoksu dopasowania krzywej Hempla*.

Funkcjonowanie modelu powinno być ponadto na tyle przejrzyste, by można było zrozumieć jego działanie, bez specjalistycznej wiedzy z zakresu programowania – dlatego też preferowane są modele wykorzystujące prostą, konsekwentnie stosowaną architekturę.

Oceniając wartość empiryczną modelu zwraca się przede wszystkim uwagę na „realistyczność” zachowania modelu. Rozumie się przez to zgodność parametrów charakteryzujących zachowanie człowieka i modelu kognitywnego. Jeśli zrozumienie wypowiedzianego zdania i reakcja na nie zajmuje człowiekowi określoną ilość czasu, to rezultaty osiągane przez modele nie powinny w znaczący sposób od nich odbiegać. Podobnie etapy funkcjonowania modelu kognitywnego powinny pozostawać w zgodzie z etapami procesów myślowych zachodzących podczas rozwiązywania problemów przez ludzi.³

Modele poznawcze wykorzystywane są do bardzo szczegółowego wyjaśniania danych eksperymentalnych. Oczekuje się zatem, że modele takie wykraczać będą poza etap „realistyczny”, tzn. proste wyjaśnianie dostępnych danych i pozwolą także przewidywać zachowania (np. werbalne) człowieka.

1.1.4 Modelowanie oparte na architekturach

Współczesne badania nad modelowaniem kognitywnym łączą w sobie dwa podejścia:

- wykorzystanie specjalnych języków programowania przeznaczonych do modelowania ludzkiego zachowania takich, jak np. Soar (zaprojektowany przez Rosenbloom, Laird i Newell system wykorzystujący mechanizm systemów produkcji), rodzina systemów ACT, SNePS (sieci semantyczne) czy inne.
- jako uzupełnienie podejścia pierwszego, wprowadza się postulat stosowania się tego samego modelu (z możliwie najmniejszymi zmianami) do jak najszerszego zakresu zadań. Jeśli zatem system ma do czynienia z nowym zadaniem wymagającym operacji na nowych danych, jego twórca powinien wykorzystywać już istniejące w systemie mechanizmy nie tworząc nowych, na potrzeby rozwiązania jedynie tego problemu. Obliczeniowy model pamięci krótkotrwałej wyjaśniający procesy natychmiastowego seryjnego przywoływania informacji z pamięci powinien mieć zastosowanie również do innych procesów np. do procesów związanych z rozpoznawaniem [Anderson, Matessa, 1997]

³Metoda uzyskiwania takich danych w psychologii poznawczej: „myślenie na głos” budzi jednak wątpliwości wśród psychologów [Garnham 1994, s. 176–177].

Obydwa te warunki spełnia modelowanie oparte na architekturze poznawczej. Zakłada się, iż przyjęta dla danej umiejętności poznawczej architektura wyposażona jest w ustalony zbiór mechanizmów obliczeniowych i zasobów leżących u podstaw ludzkich zdolności poznawczych. W zależności od przyjętej architektury poznawczej modelowanie poznawcze może mieć charakter *symboliczny* lub *koneksjonistyczny*. Modelowanie koneksjonistyczne stanowi próbę zrozumienia i wyjaśnienia ludzkiego poznania przy wykorzystaniu symulowanych sieci prostych, przypominających neurony jednostek przetwarzających. Z kolei, model ma charakter symboliczny, jeśli ma własności systemu symbolicznego w znaczeniu zaproponowanym przez Simona i Newella, znajdującym swoje odbicie w *hipotezie fizycznego systemu symboli* [Newell, Simon 1976]. Zgodnie z ich stanowiskiem, system symboliczny powinien być zdolny do tworzenia symboli i złożonych struktur symbolicznych oraz manipulowania nimi. Symbole rozumiane są jako pewne fizyczne wzorce, dające się porównać za względu na kształt, których egzemplarze mogą tworzyć bardziej złożone jednostki – struktury symboliczne. System symboliczny umożliwia tworzenie i interpretowanie nowych struktur symboli.

Podział powyższy, niegdyś uważany za rozłączny, obecnie się zaciera: coraz częściej uznaje się, iż modele symboliczne i koneksjonistyczne są swoim wzajemnym dopełnieniem, powstają też systemy łączące w sobie podejścia koneksjonistyczne i symboliczne⁴.

1.2 Architektura poznawcza

Dowolna teoria obliczeniowa poznania ludzkiego musi – w sposób wyraźny, bądź milczący – przyjąć architekturę podmiotu poznającego⁵. Mówiąc o architekturze poznawczej zazwyczaj mamy na myśli *projekt i organizację umysłu*. Innymi słowy, architektura poznawcza określa organizację leżącą u podstaw procesów poznawczych podmiotu oraz określa mechanizmy obliczeniowe stanowiące podstawę określonych umiejętności podmiotu.

Teorie architektur poznawczych mają za zadanie podać wyczerpujący przegląd systemów poznawczych, opisać ich funkcje i możliwości oraz wskazać sposób integracji takich systemów. Zadaniem architektur poznawczych jest wyliczenie zasad konstruowania modeli poznawczych. [MITECS 2001, s. 124].

⁴Tzw. systemy hybrydowe. Przykładem takiego systemu jest stworzony przez Andersona ACT-R [Anderson 1983]

⁵W literaturze anglojęzycznej przyjęło się używać terminu *cognitive agent*, co często tłumaczone jest na język polski jako *agent* poznający. Ja będę używał określenia *podmiot* poznający, obejmując tym terminem zarówno człowieka (podmiot naturalny) jak i hipotetyczne programy (procesy) poznające – inteligentne (podmioty sztuczne).

Zdaniem Slomana cecha ta odróżnia architektury poznawcze od innego typu teorii kognitywnych, które formułują zbiór hipotez poddawanych testom empirycznym. Anderson [Anderson 1993, s. 3-4.] charakteryzuje architekturę poznawczą następująco:

Architektury poznawcze są stosunkowo pełnymi propozycjami dotyczącymi struktury ludzkiego poznania. Można je, pod tym względem, przeciwstawić teoriom obejmującym tylko pewien aspekt poznania, jak np. teoriom ujmującym rozróżnienie pomiędzy pamięcią krótko- i długotrwałą... [...] Tak, jak architekt próbuje dostarczyć pełnej specyfikacji domu (dla budowniczego), podobnie komputerowy czy poznawczy architekt próbuje przedstawić pełną specyfikację systemu. W specyfikacji podanej przez architekta pozostaje jednak pewien stopień abstrakcji, pozostawiający konkretną realizację budowniczemu. Podobnie istnieje pewien stopień abstrakcji w architekturze komputerowej (poznawczej): nie ma potrzeby dokładnej specyfikacji neuronów w architekturze poznawczej, nie wyszczególnia się też wszystkich elementów obliczanych w architekturze obliczeniowej.

„Stosunkowo pełne” w powyższym fragmencie oznacza, iż są one relatywne do celów określonego aspektu poznania, jaki obejmuje architektura. Można także rozumieć architektury poznawcze jako wskazanie teorii leżącej u podstaw modelowania poznawczego.

Nie istnieje obecnie żadna zadowalająca teoria określająca możliwe architektury systemów poznawczych, wskazująca optymalną architekturę dla danej umiejętności poznawczej. Niektórzy badacze (za Newellem i Simonem) traktują obserwowane rozmaite umiejętności poznawcze jako efekt działania jednej architektury poznawczej: jednolitej, ogólnego zastosowania, nastawionej na rozwiązywanie problemów. Przykładem takiej architektury są systemy produkcji stanowiące podstawę systemu General Problem Solver (GPS) [Roth, Woods 1995, Frazier 2001]. Inni wskazują na funkcjonalny podział architektury poznawczej na moduły odpowiadające za działanie np. percepcji, uczenia się, wnioskowania, komunikacji itd. Często interpretuje się wyniki badań neurobiologii (neuronauki) na korzyść koncepcji modularnej: w ramach mózgu można wyróżnić obszary, z których każdy odpowiedzialny jest za pewną funkcję poznawczą. Funkcje te określa się (nieformalnie) jako umiejętności językowe, rozpoznawanie twarzy, percepcja muzyki i wypowiedzi języka (jako oddzielne funkcje) itp. Istota powiązań między takimi postulowanymi modułami pozostaje jednak niejasna. Dyskusji nad modularnym charakterem umysłu poświęcony jest rozdział 2.

Podział architektur poznawczych można przeprowadzić na wielu poziomach. Jednym z uznawanych do niedawna za podstawowy jest podział na architektury symboliczne, oparte na przetwarzaniu symboli i inspirowane pojęciem komputera cyfrowego oraz architektury koneksjonistyczne, „asocjacyjne” opierające się na sieciach neuronowych. Obecnie, jak już wyżej powiedziano, wielu badaczy próbuje łączyć ze sobą te dwa podejścia, a zatem łączyć w jednym systemie poznawczym obydwie rodzaje architektur. Jedną z propozycji przedstawia sieci neuronowe jako architektury odpowiadające za poznanie niskiego poziomu dostarczające następnie danych architekturom symbolicznym, leżącym u podstaw poznania wysokiego poziomu.

W niniejszej pracy prezentuję model oparty o architekturę symboliczną, w związku z czym pisząc dalej o architekturze poznawczej będę miał na myśli właśnie tę odmianę (o ile nie będzie wyraźnie wskazane, że chodzi o koneksjonistyczną)

Architektury symboliczne wywodzą się od architektury von Neumanna, na którą składa się centralna jednostka przetwarzająca, oraz jednostki: pamięciowa, wejściowa i wyjściowa. System odbiera informacje, które są następnie przechowywane w pamięci, przetwarzane algorytmicznie i wysyłane poza system.

Najniższy poziom architektury poznawczej to język programowania, w którym dana architektura została zaimplementowana. Wybór języka implementacyjnego nie musi w sposób znaczący ograniczać tworzonych modeli. W praktyce może ułatwić lub utrudnić niektóre rozwiązania. Na tym etapie dokonuje się wyboru podejścia symbolicznego, koneksjonistycznego czy hybrydowego.

Na pośrednich poziomach architektur poznawczych podejmowane są pewne strukturalne decyzje dotyczące podstawowych schematów reprezentacji wiedzy i operacji na nich. Można się decydować, na przykład na systemy produkcji, architekturę „blackboard”, sieci semantyczne czy podejście obiektowe (*object-oriented approach*). Jak wspomniano powyżej, realizacja tych schematów w zasadzie jest niezależna od języka programowania, chociaż wybrane języki ułatwiają realizację konkretnych struktur: PROLOG znacznie ułatwia tworzenie systemów produkcji, podczas gdy sieci semantyczne znacznie łatwiej zrealizować w LISP-ie czy Javie. Wybór określonego schematu reprezentacyjnego najczęściej jest konsekwencją wyboru dokonanego na najwyższym poziomie architektury poznawczej.

Na poziomie najwyższym architektur poznawczych przyjmuje się pewne założenia dotyczące celów jakie ma realizować dana architektura poznawcza (np. modelowania rozwiązywania problemów, uczenia się elementów słownika, pojmowanie języka i in.), oraz struktury niezbędnej do realizacji tych celów.

1.3 Charakterystyka architektury poznawczej: ACT-R i SNePS

1.3.1 Hybrydowa architektura poznawcza: ACT-R

Poniżej krótko przedstawię dwa przykłady modelowania poznawczego z wykorzystaniem architektur kognitywnych. ACT-R stanowi rozwinięcie jednego z pierwszych obliczeniowych modeli kognitywnych o nazwie ACT [Anderson 1983].

ACT-R jest ogólną teorią poznania sformułowaną i rozwijaną przez J.R. Andersona wraz z grupą badawczą na Carnegie Mellon University. ACT-R jest zatem również architekturą poznawczą. ACT-R nie jest językiem programowania w sensie ścisłym, bowiem jego konstrukcja odzwierciedla założenia dotyczące ludzkiego poznania. Podobnie jak język programowania ACT-R dostarcza pewnych ram: przy jego wykorzystaniu dla różnych zadań (pojmowanie języka, komunikacja, wieże Hanoi etc.) budowane są modele poznawcze. Pisząc modele w ACT-R przyjmuje się wzmiankowane ogólne założenia dotyczące poznania, zakodowane w architekturze, precyzując je dodając własne założenia. Można je następnie testować drogą porównywania wyników działania modeli z osiągnięciami ludzi wykonującymi te same zadania. Owe wyniki to m.in. czas wykonania zadania, precyzja w wykonywaniu zadania czy (ostatnio) dane neurologiczne otrzymane w wyniku neuroobrazowania, np. metodą fMRI⁶. Za pomocą ACT-R tworzone są modele w ramach następujących dziedzin: uczenie się i pamięć, rozwiązywanie problemów i podejmowanie decyzji, język i komunikacja, percepcja i uwaga, rozwój poznawczy. W miarę rozwoju projektu system będzie w stanie wykonać wszystkie ludzkie zadania poznawcze, szczegółowo wyjaśniając sposób postrzegania świata, myślenia o nim i działania w świecie.

Na najniższym poziomie architektura ACT-R zrealizowana jest w LISP-ie. Na poziomie pośrednim jest to architektura symboliczna wykorzystująca mechanizmy systemów produkcji. Systemy produkcji kontrolowane są przez poziom subsymboliczny reprezentowany przez zbiór równoległych procesów o charakterze koneksjonistycznym. Na przykład, to, czy (lub jak szybko) pewien fakt może zostać przywołany z pamięci deklaratywnej, uzależnione jest od subsymbolicznych równań odpowiadających za odzyskiwanie z pamięci, biorących pod uwagę kontekst i wcześniejsze przywołania tego faktu z pamięci. System ACT-R jest zatem systemem hybrydowym, służącym budowaniu hybrydowych modeli poznawczych odpowiedzialnych za określone zdolności poznawcze, przy wykorzystaniu architektury mieszanej: symbolicznej (systemy pro-

⁶fMRI (functional Magnetic Resonance Imaging) to metoda obrazowania z wykorzystaniem rezonansu magnetycznego. Metody neurobiologii mają coraz większy wpływ na modelowanie kognitywne (w szczególności dotyczy to modelowania zdolności językowych), coraz intensywniej rozwija się w ramach nauk o poznaniu tzw. kognitywna neuronauka. Na temat roli neuroobrazowania w modelowaniu kognitywnym por. np. [Rugg 2000]

dukcji) i koneksjonistycznej.

Na poziomie najwyższym architektura ACT-R charakteryzowana jest jako mechanizm, na który składają się moduły, bufor i program dopasowujący wzorce. Moduły mają dwojaki charakter: perceptualno-motoryczny – te moduły odpowiadają za interakcję z pewnym „światem” zewnętrznym w stosunku do systemu poznawczego (np. symulacją świata rzeczywistego); moduły pamięciowe, na które składa się pamięć deklaratywna oraz pamięć proceduralna w postaci systemów produkcji. Do modułów ACT-R ma dostęp za pośrednictwem *buforów*. Dla każdego modułu istnieje bufor, którego zawartość w danym momencie reprezentuje stan systemu ACT-R w tym momencie. Program dopasowujący wzorce poszukuje produkcji odpowiadających aktualnemu stanowi buforów. Tylko jedna z odnalezionych produkcji może w jednym czasie zostać uruchomiona, zaś efektem jej działania jest modyfikacja zawartości buforów. Modelowaną zdolność poznawczą stanowią sukcesywnie „odpalane” produkcje. Na tym poziomie wyspecyfikowane powinny być również założenia dotyczące celów poznawczych: składają się na nie ogólne założenia dotyczące poznania, jakie realizuje ACT-R oraz specyficzne dla danej dziedziny założenia (por. powyżej).

1.3.2 Architektura sieci semantycznych: SNePS

System SNePS (*Semantic Network Processing System*) stanowi klasyczną, symboliczną architekturę umożliwiającą realizację licznych zdolności poznawczych. Rozwijany jest przez Stuarta C. Shapiro i Williama Rapaporta, wraz ze współpracownikami, od końca lat 70. na Uniwersytecie Stanowym w Buffalo. Pierwotnym zamierzeniem twórców projektu było stworzenie systemu zdolnego do reprezentacji przekonań inteligentnych podmiotów posługujących się językiem naturalnym. SNePS miał zatem reprezentować wszystkie informacje, jakie można przekazać za pomocą języka naturalnego: uzyskiwane drogą komunikacji między podmiotami, ale również z książek czy artykułów pisanych dla ludzi. SNePS miał reprezentować zarówno informacje ogólne, jak i szczegółowe, wnioskować w oparciu o nie, zachowywać się w sposób nie odbiegający od zachowania naturalnych podmiotów poznawczych, nawet mając do czynienia z informacjami niezupełnymi, sprzecznymi czy wykorzystującymi np. definicje kołowe.

Na poziomie najniższym SNePS – jako architektura poznawcza – zrealizowany jest w języku LISP (powstają wersje testowe realizowane w języku Java). Na poziomie pośrednim jest on modularną architekturą symboliczną wykorzystującą sieci semantyczne do reprezentacji wiedzy, wyposażoną w procedury odzyskiwania z informacji z sieci semantycznych oraz przeprowadzanie

inferencji na zgromadzonej wiedzy. Współpracę użytkownika/projektanta ze SNePS (wprowadzanie i odbieranie informacji) umożliwiają: SNePSUL (*SNePS User Language*, język poleceń w notacji przypominającej notację LISP), SNePSLOG (język logiki wyższego rzędu) oraz GINSENG, moduł umożliwiający tworzenie sieci semantycznych (i ich prezentację) w postaci graficznej (tj. w postaci grafów).

Na poziomie najwyższym można SNePS scharakteryzować jako modułarny system, na który składają się:

- moduł reprezentacyjny (baza wiedzy),
- moduł inferencyjny (SNIP, *SNePS Inference Package*) powiązany bezpośrednio z bazą wiedzy, mający za zadanie przeprowadzanie inferencji na tej bazie wiedzy. Interpretuje on niektóre węzły sieci semantycznej jako „węzły reguł”, dzięki którym jest w stanie przeprowadzać wnioskowania oparte na węzłach [*node-based*].
- moduł modyfikacji bazy wiedzy (SNeBR, **SNePS Belief Revision**), odpowiadający za dodawanie nowych informacji do bazy wiedzy i ich usuwanie (w szczególnych przypadkach może „izolować” pewien fragment bazy wiedzy, np. zawierający informacje sprzeczne, gdy system nie ma dość informacji by przesądzić, którą ze sprzecznych informacji uznać).

W zależności od celów stawianych przed modelami kognitywnymi stworzonymi w oparciu o architekturę SNePS, system ten może być poszerzany o dodatkowe procedury i mechanizmy, np. parsery umożliwiające analizę wyrażen języka naturalnego do postaci bazy wiedzy (sieci semantycznej). Elastyczność SNePS potwierdza szeroki zakres zastosowań: od relacyjnych baz danych, poprzez systemy eksperckie (np. NEUREX, system diagnozujący choroby centralnego i obwodowego układu nerwowego), VMES (Versatile Maintenance Expert System, system reprezentacji wiedzy „wizualnej”) aż po systemy rozumienia języka naturalnego (np. system przyswajania znaczeń nowych słów w oparciu o kontekst).

Modelowanie kognitywne w oparciu o architekturę SNePS polega na tworzeniu (modeli) podmiotów poznających⁷ wykazujących określone zdolności poznawcze. W laboratorium Grupy Badawczej SNePS (SNeRG) stworzono sztuczny podmiot poznający CASSIE (**Cognitive Agent of the SNePS System – an Intelligent Entity**) wyposażony w ubogą bazę wiedzy i przyswajający w oparciu o nią nową wiedzę z wypowiedzi wprowadzanych z klawiatury, sformułowanych w

⁷Według twórców systemu SNePS, tworzone w jego ramach sztuczne podmioty poznawcze realizujące odpowiednie programy *poznają* naprawdę, a nie jedynie symulują/modelują poznanie. W niniejszej pracy przyjmuję słabszą interpretację, traktując takie podmioty jako modele, nie rozstrzygając kwestii, czy CASSIE lub OSCAR *naprawdę* poznają.

5.3 Moduł semantyczny. Implementacja

5.3.1 Reprezentacja wiedzy - rozwiązania

Modelowanie kognitywne korzysta – o czym już była mowa – z wypracowanych w ramach SI narzędzi gromadzenia i przetwarzania wiedzy. *Reprezentacja wiedzy* w badaniach nad sztuczną inteligencją, to kombinacja określonych struktur danych i procedur je interpretujących. Systemy rozumienia języka naturalnego wymagają wyspecjalizowanych struktur reprezentacji wiedzy – umożliwiających odzyskiwanie informacji z wyrażen języka naturalnego i wyrażanie ich w tej postaci – a zatem wymagać będą również procedur pośredniczących między systemami analizy/generowania gramatycznego. Ponadto – o czym już była kilkakrotnie mowa – schematy takie winny spełniać warunki nakładane przez badania eksperymentalne w ramach psychologii poznawczej, związane np. z czasem odzyskiwania rozmaitych informacji z pamięci.

Spośród schematów reprezentacji wiedzy, jakie tworzono w historii SI, jedynie nieliczne znalazły zastosowanie w modelowaniu zdolności językowych. Do celów takich nie wykorzystuje się obecnie m.in. reprezentacji wiedzy opartej o logikę, reprezentacji proceduralnych (np. systemów produkcji), czy reprezentacji bezpośrednich (np. analogicznych)²². Często wykorzystywane są natomiast (w przypadku modelowania symbolicznego) reprezentacje oparte o sieci semantyczne oraz ramy i skrypty.

5.3.1.1 Teoria zależności pojęciowej

Teoria zależności pojęciowej (*Conceptual Dependency Theory, CD*) [Schank 1972] wykorzystuje formalizm składający się z wyrażen zwanych „konceptualizacjami” stwierdzającymi zachodzenie pewnego zdarzenia lub stanu rzeczy. Stanowi ona przykład reprezentacji wiedzy opartej o ramy i skrypty. Teoria ta stawia sobie za zadanie reprezentację dowolnej czynności w postaci konceptualizacji, która jest „szkieletem”, ramą na którą składają się jedno lub więcej pierwotnych działań wraz z pośredniczącymi stanami i relacjami kauzalnymi. Schank utrzymuje, iż zadaniem reprezentacji wyrażen języka naturalnego jest ujęcie tkwiącej u ich podstaw struktury pojęciowej. Reprezentacja zatem powinna być jednoznaczna, nawet jeśli wyrażenie języka naturalnego jest wieloznaczne, bowiem osoba wypowiadająca wieloznaczne wyrażenia zazwyczaj ma na myśli określoną jego interpretację. Reprezentacja powinna być również niepowtarzalna: odmienne zdania odzwierciedlające tę samą pojęciową zawartość winny mieć tę samą reprezentację. By uzyskać niepowtarzalną, jednoznaczną reprezentację CD wykorzystuje sześć typów termów:

²²Por. np.: [Barr, Feigenbaum, Cohen 1981/82, s. 143–147, 200–206], [Harris 1985, s. 315–322]

PP - obiekty świata rzeczywistego, ACTy - działania w świecie rzeczywistym, PA - atrybuty obiektów, AA – atrybuty działań, T – czasy i LOC – umiejscowienia. Zbiór ACT jest zamknięty i składa się z pierwotnych działań takich, jak:

PTRANS – zmiana położenia obiektu fizycznego

ATRANS – zmiana relacji abstrakcyjnej, np. zmiana własności

MTRANS – mentalny transfer informacji

MBUILD – tworzenie nowej idei lub wniosku z dostępnych informacji

ATTEND – skupienie na czymś narządu zmysłowego (oka, ucha) i in.

Składniowo konceptualizacja zdarzenia jest strukturą sześćoargumentową, której argumentami są: działający, działanie, obiekt, źródło, cel, oraz instrument. Niektóre z tych argumentów (w zależności od opisywanego zdarzenia) mogą być pominięte (oznaczone są jako opcjonalne, tj. występujące 0 lub więcej razy). Funkcje argumentów mogą pełnić tylko ściśle określone terminy, np. tylko PP może być działającym i tylko ACT może być działaniem. Same konceptualizacje też mogą być terminami: jedynie konceptualizacja może pojawić się jako instrument.

```
[ATRANS
rel: POSIADANIE
działający: JAN
obiekt: PIŁKA
źródło: JAN
cel: ANNA]
```

Rysunek 5.2: Reprezentacja zdania *Jan dał Annie piłkę* w ramach teorii zależności pojęciowej

Pierwotne elementy pojawiające się w konceptualizacjach są – zdaniem Schanka – pojęciami, nie zaś słowami, odzwierciedlają zatem myśl leżącą u podstaw języka, a nie sam język.

Schank miał nadzieję, iż zaproponowany system będzie wystarczał do reprezentacji znaczenia dowolnego czasownika. Okazało się jednak, w trakcie prac nad zastosowaniami systemu w zadaniach związanych z rozumieniem języka potocznego, obejmującego wszelkie dziedziny życia, iż niezbędne są dodatkowe mechanizmy. Problemem w szczególności była konieczność wykorzystania wiedzy tła w komunikacji między podmiotami. Ponadto teoria zależności pojęciowej przedstawiała zdarzenia w postaci ich fizycznej realizacji, pomijając wszelki aspekt społeczny czy emocjonalny.²³

Reprezentacje typu zależności pojęciowej, oparte o ramy, skrypty czy schematy wymagają dużo pracy pojęciowej twórców tych systemów: są oparte o zbiór wcześniej przygotowanych ram pozwalających ująć zdarzenia – trudno zatem realizować za ich pomocą systemy uczące się, rozwi-

²³Np. pocałunek przedstawiany jest jako fizyczne przemieszczenie ust do ust.

jające swoją bazę konceptualną. Są to systemy dające prymat semantyce, podporządkowując jej syntaktykę. Domyślną dziedziną interpretacji tego formalizmu jest świat rzeczywisty, jednak wśród konceptualizacji muszą istnieć obiekty teoretyczne, jak „świadomy procesor” jako element podmiotu poznającego, w którym umieszczone są konceptualizacje; zdarzenia warunkowe czy zdarzenia, które nie zaszły.

5.3.1.2 KL-ONE

System KL-ONE²⁴ [Brachman, Schmolze 1985], to jeden z języków wykorzystywany głównie w badaniach nad pojmowaniem języka naturalnego i reprezentacją wiedzy. Stanowi on realizację sieci semantycznych średniego poziomu (por. poniżej, paragraf 5.3.2) tzn. sieci semantyczne na poziomie epistemologicznym przyjmujące postać ustrukturalizowanych sieci dziedziczenia. Jak pisze Brachman: *Sieć KL-ONE reprezentuje, tym samym, przekonania dotyczące świata (i innych światów możliwych) w taki sposób, jak światy te pojmuje system używający sieci KL-ONE.*

Omawiany system opiera się na dwóch elementach: definicyjnym (terminologicznym) zwanym TBox oraz asertorycznym zwanym ABox. Pierwszy z nich definiuje terminologię wykorzystywaną do modelowania (wybranego aspektu) świata; w jego ramach opisywane są obiekty z danej dziedziny poprzez ustalanie między nimi relacji. W TBox istnieją dwojakiemu rodzaju terminy: pojęcia i relacje (role). Pojęcia mają charakter ogólny lub indywidualny. Pojęcia ogólne definiowane są w postaci sieci. Sieć tą następnie dziedziczą bądź pojęcia indywidualne stanowiące przypadki pojęcia ogólnego, bądź dziedziczona jest ona przez inne pojęcia ogólne, stanowiące podtypy pojęcia definiowanego. Na element asertoryczny składają się zdania w ograniczonym języku logiki pierwszego rzędu, traktowane jako asercje w danej dziedzinie zastosowań. Terminy składnika asercyjnego mają znaczenie egzystencjalne. Stałe reprezentują indywidua istniejące w danej dziedzinie, zdania kwantyfikowane egzystencjalnie postaci $\exists xP(x)$ stwierdzają istnienie indywiduum spełniającego warunek P w danej dziedzinie.

Przykładowe zdania pojawiające się w TBox mogą wyglądać następująco:

Rodzina to pewnego rodzaju **struktura społeczna** z dokładnie jednym *męskim-rodzicem* i ten *męski-rodzic* jest **mężczyzną**; jest to pewnego rodzaju **struktura-społeczna** z dokładnie jednym *żeńskim-rodzicem* i ten *żeński-rodzic* jest **kobietą**; jest to **struktura-społeczna** gdzie wszystkie *dzieci* są **osobami**. Wyrażenia wytłuszczone są to terminy – pojęcia, wyrażenia wyróżnione kursywą są to terminy – relacje. Wszystkie pozostałe wyrażenia stanowią interpunkcję;

²⁴Należałoby mówić o rodzinie systemów KL-ONE. Pierwotną wersję zastąpiły kolejno: KL-TWO, KRYPTON, oraz Loom. Szersze omówienie systemu KL-ONE można znaleźć m.in. w [Harris 1985, s. 305–311.]

zdanie to jako całość traktowane jest jako definicja **rodziny**. Zatem nie przysługuje mu wartość prawda ani fałsz, można mówić co najwyżej o różnych pojęciach rodziny. Zdefiniowanie rodziny w TBox nie mówi nic o istnieniu jej w danej dziedzinie. Jej istnienie stwierdzać mogą dopiero wyrażenia ABox, np. $\exists x \text{rodzina}(x)$ lub $\text{rodzina}(\text{Jan})$.

5.3.2 Charakterystyka sieci semantycznych

Sieć semantyczna jest strukturą pozwalającą na reprezentację wiedzy za pomocą węzłów powiązanych opatrzonymi etykietami, ukierunkowanymi łukami. Sieci semantyczne, traktowane jako system reprezentacji wiedzy, wymagają procedur umożliwiających dodawanie informacji do bazy wiedzy i jej modyfikowanie w oparciu o dochodzące informacje, odzyskiwanie wiedzy z tej bazy, przeprowadzanie wnioskowań w oparciu o zgromadzone informacje. Niezbędna jest zatem pewna liczba systemów pośredniczących między użytkownikiem a bazą wiedzy. Użytkownik może być systemem poznawczym naturalnym lub sztucznym (innym systemem reprezentacji i przetwarzania wiedzy).

W badaniach psychologii poznawczej wykazywano, że pamięć trudno uznać za przechowalnię zdarzeń; należy ją raczej traktować jako zbiór sądów, które muszą być opracowane w celu zrekonstruowania poszukiwanego obrazu²⁵. I właśnie badania nad pamięcią dostarczyły architektury reprezentacji wiedzy, jaką są sieci semantyczne. Strukturę sieciową postulowali m.in. Quillian, Anderson i Bower. Szczególne znaczenie miał model pamięci semantycznej Quilliana, który nie był pierwszym modelem opartym o reprezentacje sieciowe, a jednak miał duży wpływ na realizowane w ramach badań nad sztuczną inteligencją i reprezentacją projekty. Wprowadzony przez Quilliana algorytm przechodzenia znaczników drogą rozprzestrzeniającej się aktywacji [*spreading activation*] zaadaptowany został w późniejszych systemach, jak NETL (Fahlman, 1979), parser równoległy (Waltz, Pollack 1985) i liczne wersje sieci neuronowych. Przyjęta przez Quilliana metoda reprezentacji typów i egzemplarzy miała wpływ na takie systemy taksonomiczne, jak KL-ONE, zaś jego technika generowania języka naturalnego znalazła odbicie w systemach rozwijanych przez Goldmana (1975) czy McNeilla (1979).

Pomimo pierwotnego zastosowania sieci semantycznych jako modelu ludzkiej pamięci, schemat ten szybko został wykorzystany przez informatykę i sztuczną inteligencję (rozumianą jako dziedzinę badań) do innych zadań. Analizując wykorzystanie sieci semantycznych i ich epistemologiczny status Brachman [Brachman 1979] wyróżnił pięć poziomów sieci semantycznych:

²⁵Więcej na ten temat zob. [Lindsay, Norman 1984, s. 379–385].

1. poziom *implementacyjny*: pierwotne elementy sieci to atomy i wskaźniki; sieci na tym poziomie wykorzystywane były jako struktury danych typu sieciowego (grafy, drzewa). Przykładem sieci semantycznej na tym poziomie jest gramatyka ATN opisywana w rozdziale 4, podobnie jak implementacja gramatyki przypadków Fillmore'a;
2. poziom *logiczny*: pierwotne elementy to sądy, predykaty oraz operatory logiczne; konstrukcje i operacje logiczne mają za zadanie reprezentować znaczenie. Przykładami sieci na tym poziomie są rozdzielone sieci semantyczne (Hendrix) oraz rozszerzone sieci semantyczne (Schubert, Cercone);
3. poziom *epistemologiczny*: pierwotne elementy sieci to typy pojęć, składniki pojęciowe oraz relacje dziedziczenia i relacje strukturalizujące. Przykładem jest omawiany powyżej system KL-ONE wraz z następcami;
4. poziom *pojęciowy*: pierwotne elementy sieci to semantyczne (pojęciowe) relacje (np. przypadki semantyczne), pierwotne obiekty i działania. Przykładami mogą być teoria zależności pojęciowej (Schank, por. paragraf 5.3.1.1), struktury gramatyki przypadków (Simmons) czy wreszcie system SNePS (Shapiro, grupa badawcza SNePS);
5. poziom *lingwistyczny*: zrealizowany został w systemie zwanym OWL obejmującym podstawowy schemat strukturalizujący pojęcia wykorzystywany do budowy wyrażeń językowych. Pierwotne w tym systemie są takie zasady, jak specjalizacja, dołączanie czy referencja.

Według Brachmana sieci semantyczne na dowolnym z wymienionych powyżej poziomów powinny spełniać sześć warunków:

- Sieć semantyczna musi mieć jednolitą notację. Wybór notacji jest jednak wyborem projektanta sieci semantycznej. Obecnie istnieje szereg różnych architektur poznawczych opartych o formalizm sieci semantycznych, z których każda wykorzystuje nieco odmienną notację. Można wśród nich wymienić: grafy relacyjne (Ceccato, Quillian, Schank), hierarchia typów i dziedziczenia (KL-ONE: Brachman), propozycjonalne sieci semantyczne (LADDER: Hendrix, grafy pojęciowe: Sowa, SNePS: Shapiro i grupa badawcza SNePS)
- Musi istnieć algorytm pozwalający na kodowanie informacji w postaci sieci. Warunek ten realizuje się dostarczając pośredników (*inteface*) pomiędzy danymi wejściowymi (na przykład w postaci wypowiedzi języka naturalnego²⁶) a siecią semantyczną.

²⁶Szeroki zakres zastosowań sieci semantycznych nie pozwala ograniczać danych wejściowych do wypowiedzi

- Musi istnieć również mechanizm „asymilacji”, umożliwiający budowanie sieci odpowiadającej nowej informacji przy wykorzystaniu elementów sieci już istniejącej. Jeśli system otrzymuje informację o reprezentowanym już w postaci węzła obiekcie (na przykład placu M. Curie-Skłodowskiej w Lublinie), wówczas nowa informacja (np. o przebudowie placu) powinna być dołączona do tego węzła.
- Sieć semantyczna powinna być neutralna w stosunku do formalizmu sieciowego na wyższym poziomie (w hierarchii omówionej powyżej).
- Sieć semantyczna powinna być adekwatna w stosunku do dowolnego formalizmu (na wyższym poziomie), np. system z poziomu „pojęciowego” powinien dać się zaimplementować w systemie poziomu „logicznego”²⁷
- Sieć semantyczna powinna mieć semantykę. W przypadku sieci można mówić jednak przynajmniej o dwojakiego rodzaju semantykach: węzły mają znaczenie *w obrębie sieci*, tj. znaczenie jakie uzyskują dzięki powiązaniom za pomocą łuków z innymi węzłami sieci; węzły mogą mieć również znaczenie *dla użytkowników*, np. dzięki powiązaniom węzłów sieci z elementami leksykonu. Łuki sieci semantycznej mają znaczenie tylko w obrębie sieci dzięki regułom inferencji zarówno opartej na węzłach, jak i na ścieżkach (tj. ciągach węzłów powiązanych łukami). W obydwu przypadkach mamy do czynienia z wewnętrzną, systemową holistyczną semantyką: znaczenie węzłów i łuków dane jest w oparciu o ich miejsce w całej sieci.

Niektórzy badacze zwracają uwagę, iż od neutralności sieci w stosunku do formalizmów sieciowych na wyższych poziomach bardziej istotna jest kwestia doboru konkretnego formalizmu realizującego sieć semantyczną. O wyborze tym decydować może m.in. *wydajność* (w tym zagadnienie tworzenia mechanizmów pośredniczących między innymi modułami a siecią semantyczną), *adekwatność psychologiczna*, ontologia, na jakiej sieć się opiera, lub jaką dopuszcza, *adekwatność logiczna* czy zdolność integracji z innymi modułami przetwarzania języka naturalnego.

Sieci spełniające wyżej wymienione warunki (przynajmniej, jeśli bierze się pod uwagę dotychczas zaproponowane i zrealizowane sieci) łączą następujące cechy:

języka naturalnego. Jednym z ciekawszych projektów jest VMES (Versatile Maintenance Expert System), mający do czynienia z cyfrowymi obwodami reprezentowanymi graficznie. W oparciu o ten system Tsotsos i Shibahara oraz Havens i Mackworth rozwijali system reprezentacji wiedzy wizualnej. Reprezentacja ta nie opierała się na pikselach, ale przybierała postać funkcji LISPa, która odpowiadała za rysowanie obiektu na ekranie. Węzły propozycjonalne wyrażały informację o (relatywnym lub absolutnym) położeniu obiektu i jego własnościach.

²⁷Przykładem mogą być systemy KL-ONE i SNePS - KL-ONE może być implementowany w SNePSie. Por. [Shapiro, Rapaport 1987, s. 312]

- węzły sieci reprezentują pojęcia indywidualów, atrybutów, zdarzeń i stanów;
- odmienne węzły reprezentujące pojęcia tego samego typu odnoszą się do różnych indywidualów tego typu;
- łuki w sieci – zwane relacjami pojęciowymi – reprezentują relacje, jakie zachodzą między węzłami pojęciowymi; etykiety łuków wskazują na typy relacji;
- niektóre relacje pojęciowe reprezentują przypadki lingwistyczne (jak np. agent, pacjens, instrument). Istnieją również łuki reprezentujące przestrzenne, czasowe, przyczynowe oraz reprezentujące logiczne spójniki;
- typy pojęć zorganizowane są w hierarchię zgodnie z poziomem ogólności, jak np. na najwyższym poziomie *obiekt*, podrzędne względem niego *roślina*, *zwierzę*, *minerał*, podrzędne względem typu zwierzę *samiec*, *samica* itd. Hierarchia taka wiąże się z dziedziczeniem własności;

Poza wymienionymi wyżej warunkami i cechami wspólnymi różne implementacje sieci semantycznych różnią się między sobą, m.in: filozoficznymi kwestiami (zwłaszcza związanymi ze znaczeniem i interpretacją formalizmu sieci semantycznych), sposobem implementacji kwantyfikatorów i spójników logicznych, technikami operowania na sieciach: dodawania i odzyskiwania informacji, modyfikacji sieci oraz mechanizmami inferencyjnymi czy ostatecznie sposobem przedstawiania sieci: węzłów i łuków oraz opatrywania tych ostatnich etykietami czy symbolami. Ponadto różne realizacje sieci kładą nacisk na różne ich aspekty: szczególnie widoczny jest tu podział na realizacje reprezentujące sądy i dostarczające mechanizmów rozumowania (np. SNePS) oraz na realizacje skupiające się na procesach definiowania nowych pojęć w ramach istniejącej hierarchii typów (np. KL-ONE).

Niezależnie od implementacji, sieci semantyczne są formalizmem reprezentacji wiedzy wydajnym obliczeniowo, łatwym do pojęcia przez człowieka (co ma szczególne znaczenie w przypadku tworzenia systemów eksperckich) oraz dostatecznie silnym, by reprezentować semantykę języka naturalnego.

5.3.3 Ontologie

Sam formalizm reprezentacji wiedzy, taki jak sieć semantyczna, dostarcza jedynie ram umożliwiających zbudowanie – dostosowanej do potrzeb danego systemu – bazy wiedzy. Badania Collinsa

i Quilliana [Collins, Quillian 1969] wykazują, iż ludzie organizują jednostki mentalne w postaci struktury przypominającej formalną ontologię.²⁸ Ontologia jest to wyraźna specyfikacja konceptualizacji (w sensie przedstawionym w paragrafie 5.1.5). Składa się (w szerszym ujęciu) ze słownika reprezentacyjnego podającego precyzyjne definicje terminów wraz z formalnymi aksjomatami ograniczającymi interpretacje i poprawne użycie tych terminów. W sensie węższym ontologię traktuje się jak słownik reprezentacyjny.²⁹

Ontologia przyjmowana dla danego systemu poznawczego (sztucznego) ustala zatem terminologię, jaką system ten może się posługiwać wraz z relacjami, jakie zachodzą pomiędzy pojęciami odpowiadającymi tym terminom. Przyjmując określoną ontologię podmiot (agent) podejmuje *zobowiązanie ontologiczne*, iż będzie się posługiwał słownikiem, jakiego dostarcza ontologia.

Każdy system mający pełnić funkcje systemu poznawczego - w tym również systemy rozumienia języka naturalnego (SNePS, PENMAN), przetwarzające dane percepcyjne, rozwiązujące problemy itp. muszą się opierać na pewnej ontologii: *Każda baza wiedzy, system oparty na wiedzy, czy agent jest zobowiązany - implicite lub explicite - do pewnej ontologii.* [Gruber 1994] Modele systemów pojęciowych, jakimi są sieci semantyczne, mogą przybierać - w zależności od funkcji, jakie pełnią, dziedzin, jakie reprezentują - różne postaci, tj. opierać się na różnych konceptualizacjach. Przed twórcami systemów poznawczych konstruujących struktury pojęciowe stoją zatem dwa zadania.

1. Zaprojektować dla systemu odpowiednią ontologię, na tyle ogólną, by system mógł reprezentować wszystkie niezbędne do jego działania obiekty, relacje, dziedziny; jednocześnie na tyle elastyczną, by system mógł się uczyć nowych kategorii - w oparciu o kategorie już istniejące.
2. Jeśli system tworzy wiele struktur konceptualnych (np. systemy rozumienia języka naturalnego), muszą dostarczyć środków ich porównywania (dołączania, modyfikowania). By operacje takie były możliwe, sieci pojęciowe nie mogą opierać się na niewspółmiernych ontologiach.

Najczęściej rozwijane są obecnie ontologie dostarczające hierarchicznej organizacji pojęć, gdzie podstawową relacją jest JEST-TO (ISA). Takie ontologie mają zazwyczaj *węzeł szczytowy*, najogólniejsze pojęcie w hierarchii taksonomicznej (np. *rzecz*). Tego typu ontologie wykorzystywane są m.in. w grafach pojęciowych [Sowa 1984], istnieją jednak również ontologie rezygnujące z hierarchii taksonomicznej. W dowolnej ontologii indywidualne pojęcie jest jednoznacznie określone

²⁸Lub: ludzie tworzą formalne ontologie tak, iż odzwierciedlają one ich wewnętrzną organizację pojęć.

²⁹Por. [Gruber 1994] oraz [Konderak 2003, s. 155–157]. Pisząc o ontologii w dalszej części pracy będę miał na myśli szersze znaczenie tego terminu.

przez relacje, w jakich pozostaje w stosunku do innych pojęć. Jeśli wykorzystuje się jedynie hierarchię taksonomiczną, trudno jest wskazać znaczenie dowolnego terminu, bowiem tylko jedna relacja wiąże ze sobą pojęcia. Gdyby, na przykład, wiadomo, że *wszystkie rakany są tyfonami*, *wszystkie tyfony są rzeczami*, oraz istnieje inny rodzaj *tyfonów*, mianowicie *dizory*, jakie można z tego wyciągnąć wnioski? Zarówno *rakany*, jak i *dizory* są rzeczami, różnią się one również między sobą w jakiś sposób, bowiem twórca takiej ontologii nie wprowadzałby odmiennych terminów dla takich samych obiektów. Zatem hierarchia taksonomiczna zazwyczaj nie wystarcza i ontologie są rozbudowywane o inne rodzaje relacji/terminów. Im więcej wprowadzi się takich relacji, tym bardziej precyzyjnie i w bogatszy sposób można wyrazić znaczenia terminów. Decyzja, jednak, jakie relacje i terminy wprowadzić do ontologii jest decyzją projektanta i zależy zazwyczaj od dziedziny zastosowań systemu. W szczególnie trudnej sytuacji są tutaj twórcy ontologii dla systemów rozumienia języka naturalnego, bowiem ontologia musi obejmować dziedzinę codziennego dyskursu, nie może być ograniczona jedynie do pewnych jego aspektów (medycyny, filozofii, mechaniki pojazdowej itp.).

Badania nad ontologiami skupiają się zatem nad stworzeniem ontologii możliwie ogólnej, tak by w jej ramach można było budować ontologie dla dziedzin bardziej szczegółowych – rozbudowując ontologię o dodatkowe terminy/relacje. Najbardziej zaawansowanym projektem jest obecnie *Ontolingua* (<http://ontolingua.stanford.edu>) [Gruber 1993, Gruber 1994], dostępne sieciowo oprogramowanie pozwalające budować ontologie. Umożliwia on projektowanie ontologii w oparciu o ustalony zbiór *pierwotnych jednostek reprezentacyjnych*. Zbiór ten określany jest mianem *ontologii ramowej* (*Frame Ontology*), która jest podstawą tworzenia szczegółowych - charakterystycznych dla danych dziedzin - ontologii. Ontologia ramowa wyszczególnia, w postaci deklaratywnej, wspomniane *pierwotne jednostki reprezentacyjne*, np:

```
klasa relacja (?relacja)
klasa funkcja (?funkcja)
klasa klasa (?klasa)
relacja jest-elementem (?jednostka, ?klasa)
relacja podklasa (?klasa-podrzędna, ?klasa nadrzędna)
funkcja inwersja (?relacja-binarna) → ?relacja
funkcja kompozycja (?relacja-1, ?relacja-2) → ?relacja-binarna
```

Rysunek 5.3: Przykładowa ontologia ramowa (fragment). Na podstawie [Gruber 1993, dodatek 2, s. 218-219]; wyrażenia poprzedzone znakiem „?” są zmiennymi

stanowiące później podstawę dla wszystkich tworzonych w ramach projektu ontologii. Dzięki

temu *Ontolingua* jest narzędziem niezależnym od dziedziny reprezentowanej.³⁰

Narzędzia projektu *Ontolingua* pozwalają użytkownikowi stworzyć własną (choć publicznie dostępną) ontologię, bądź skorzystać z jednej spośród gotowych już struktur. Każda ze stworzonych w ten sposób ontologii zapisywana jest w formie umożliwiającej przekład na strukturę danego systemu reprezentacyjnego³¹. Składnia i semantyka systemu *Ontolingua* oparta jest na notacji i semantyce rozszerzonej wersji rachunku predykatów zwanej KIF (*Knowledge Interchange Format*, Format Wymiany Wiedzy), języka pomyślanego jako środek rozpowszechniania i wymiany wiedzy. KIF dostarcza ram reprezentacji wiedzy, nie wspiera jednak żadnych operacji na reprezentowanej wiedzy: wnioskowania, modyfikowania i in.

W ramach Sztucznej Inteligencji odróżnia się dwa poziomy opisu wiedzy: *poziom wiedzy* (*Knowledge-Level*)³² - opis wiedzy podmiotu niezależny od reprezentacji na poziomie symboli, jakich system używa. Wiedzę przypisuje się systemowi na podstawie jego działania (np. na podstawie zadawanych pytań i reakcji na zadawane z zewnątrz pytania). Drugi poziom opisu wiedzy, to *poziom symboliczny* - opis w kategoriach wewnętrznych struktur wiedzy wykorzystywanych przez system (reprezentacji propozycjonalnej, sieci semantycznych i in.). *Ontolingua* umożliwia komunikację między podmiotami na poziomie wiedzy. Różnice między systemami mogą zachodzić na poziomie symbolicznym. Dzięki temu na wspólnej ontologii mogą opierać się systemy wykorzystujące np. sieci semantyczne i systemy opierające się na reprezentacji propozycjonalnej.³³

Pamiętać należy, iż przyjęcie założenia o wrodzonych *pierwotnych jednostkach pojęciowych* (lub *pierwotnych jednostkach reprezentacyjnych*, gdy mówimy o podmiotach sztucznych) nie daje nam jeszcze odpowiedzi na pytanie **jakie** to są jednostki. Opierając się na *poziomie wiedzy systemu*, obserwując jego zachowania (językowe lub inne) możemy określić *ontologiczne zobowiązania*³⁴ danego podmiotu poznawczego, co z kolei może pomóc zrekonstruować ontologię systemu -

³⁰Ontologie tego typu określa się mianem *ontologii reprezentacyjnych* - ontologii dostarczających pewnych ram, nie wskazujących jednak sposobu reprezentowania świata.

³¹Obecnie *Ontolingua* wspiera przekłady na Loom, Epikit, Algernoon i czysty KIF. Trwają prace nad translatorami na języki reprezentacyjne systemów: SNePS (por. poniżej), Cyc, KEE, EXPRESS. Szerzej na ten temat por. [Gruber 1994, s. 5-6.]

³²Termin wprowadzony przez A. Newella [Newell 1982]. Opis systemu na poziomie wiedzy pozwala racjonalizować zachowanie systemu przez zewnętrznego obserwatora. Obserwator traktuje system jako „czarną skrzynkę”, ale zakłada, że system posiadał określoną wiedzę i ją wykorzystuje do osiągnięcia swoich celów.

³³Konstruowanie (modeli) podmiotów poznających oparte o ontologię ramową w *Ontolingua* i podobne rozwiązania w innych systemach, np. w *WordNet* czy *Cyc* stanowi tylko jedno z możliwych rozwiązań, rozwiązanie „statyczne”. Jeśli jednak podmioty mają się komunikować, specjalizując się jednocześnie w odmiennych dziedzinach, lub gdy bazy wiedzy podmiotów poznających obejmują duże dziedziny – jak to jest w przypadku systemów rozumienia języka naturalnego – wówczas bardziej adekwatne jest podejście „dynamiczne”, oparte np. o ontologiczne mediatory. Więcej na ten temat w kolejnym rozdziale, poświęconym komunikacji, paragraf ??.

³⁴Agent zobowiązuje się do pewnej ontologii, jeśli jego obserwowalne działania są zgodne z definicjami w tej ontologii. Por. [Gruber 1994, s. 2]

wysunąć hipotezę dotyczącą pierwotnych jednostek pojęciowych (reprezentacyjnych). Hipoteza ta podlega - oczywiście - dalszej empirycznej weryfikacji.

5.3.4 Propozycjonalne sieci semantyczne (projekt SNePS)

5.3.4.1 Reprezentacja intensjonalna

Proponowany w niniejszej pracy model kognitywny opiera się o formalizm *propozycjonalnych* sieci semantycznych. Termin ten oznacza, iż *każda* informacja (m.in. sądy), którą można dołączyć do sieci i uzyskać do niej dostęp jest traktowana jako pojęcie, tzn. jest przechowywana w postaci węzła sieci semantycznej.³⁵ Łuki natomiast stanowią jedynie strukturę syntaktyczną organizującą sieć semantyczną. Wyraża się to w ograniczeniu mówiącym, iż nie można dodać (nowego) łuku pomiędzy istniejącymi węzłami. Drugie ograniczenie propozycjonalnych sieci semantycznych wiąże się z relacją między węzłami, a reprezentowanymi przez nie pojęciami. *Zasada jedyności* [*Uniqueness Principle*] wymaga jednojednoznacznej odpowiedniości między węzłami a pojęciami. Zasada ta, zdaniem Rapaporta i Shapiro [Shapiro, Rapaport 1987, s. 267] *gwarantuje, że system będzie współdzielił węzły, gdzie to jest tylko możliwe oraz że węzły reprezentują obiekty intensjonalne.*

Model kognitywny zdolności językowych, a ściślej mówiąc jego moduł semantyczny ma reprezentować struktury wiedzy służące jako interpretacje (znaczenia) wypowiedzi języka naturalnego. Ponieważ – jak przyjęto powyżej – podmioty poznające są w stanie formułować wypowiedzi (i myśleć) nie tylko o obiektach oddziałujących na ich receptory, tj. o obiektach świata zewnętrznego, ale również o wielu innych obiektach o znacznie mniej jasnym statusie ontologicznym, jak przedmioty nieistniejące (złota góra), sprzeczne (kwadratowe koło) niedookreślone lub niezupełne (postaci powieści), czy dwa odmienne, choć mające tę samą ekstensję przedmioty myśli (Gwiazda Poranna i Gwiazda Wieczorna), system poznawczy musi mieć również środki reprezentacji takich obiektów. Ponadto sieć semantyczna w modelu kognitywnym ma reprezentować nie świat zewnętrzny, ale struktury pojęciowe, jakimi dysponuje podmiot poznający, a zatem przyjmuje się, iż węzły sieci reprezentują *obiekty intensjonalne*. Tutaj chcę wyraźnie zaznaczyć, iż sieci semantyczne *mogą* reprezentować obiekty ekstensjonalne. Ich reprezentacja byłaby jednak niezbędna jedynie w przypadku, gdyby przedmiotem zainteresowania była relacja między umysłem a światem. Celem zaś modelowania kognitywnego jest modelowanie umysłu, w szczególności konceptualizacji – modelu świata, jaki umysł tworzy.

³⁵Istnieją bowiem również odmiany sieci, w których asercje reprezentowane są w postaci łuków sieci.

Węzły omawianej sieci semantycznej będą zatem reprezentowały obiekty niezastępowalne w kontekstach nieprzejrzystych, niedookreślone pod pewnymi względami, takie, które nie muszą istnieć, nie muszą nawet być możliwe. Nie jest to jednakże jedyna interpretacja sieci semantycznych. W historii badań nad formalizmami reprezentacji badacze zajmowali różne stanowiska. Jeden z twórców sieci semantycznych jako formalizmu SI, Woods [Woods 1975, s. 50] pisze, że *muszą istnieć dwie jednostki mentalne (pojęcia, węzły czy inne) odpowiadające dwom odmiennym intensjom, Gwieździe Wieczornej i Gwieździe Porannej. Istnieje również asercja dotycząca tych dwóch jednostek, iż denotują one jeden i ten sam zewnętrzny obiekt*. Woods dopuszczał zarówno reprezentacje intensjonalne, jak i ekstensjonalne. Brachman nie widzi potrzeby reprezentacji ekstensjonalnych, choć ich istnienie dopuszcza: *sieci semantyczne są reprezentacjami intensji desygnatorów języka naturalnego* [Brachman 1977, s. 139]. Najbardziej radykalne stanowisko zajmują Maida i Shapiro [Maida, Shapiro 1985, s. 173] twierdząc, iż wszelka informacja w sieci semantycznej ma charakter intensjonalny. Jedyne odniesienie do referencji będzie się pojawiało w sądach stwierdzających, iż dwom intensjom odpowiada te sama ekstensja w pewnym świecie. Sytuacja, w której dwa węzły denotują jeden obiekt ze świata pozapodmiotowego wymagać będzie zatem specjalnego węzła, będącego asercją tego faktu. Przeniesienie asercji dotyczących jednego węzła (np. reprezentującego *Gwiazdę Poranną*) na drugi o tej samej ekstensji (np. *Gwiazdę Wieczorną*) wymagało będzie specyficznych reguł inferencyjnych. Takie podejście wymaga uznania, iż nieprzejrzystość referencyjna jest sytuacją normalną, zaś przejrzystość musi być wyraźnie usankcjonowana przez odpowiednie reguły. Przykłady reguł inferencji odpowiedzialnych za przypisywanie asercji obiektom równoważnym ekstensjonalnie (tj. dwóm węzłom–pojęciom, którym odpowiada jeden obiekt w rzeczywistości pozapodmiotowej) przedstawiam w paragrafie 5.3.6, poniżej.

W propozycjonalnych sieciach semantycznych węzły reprezentują intensje, nie zaś ekstensje, tzn. indywidualne pojęcia, nie zaś desygnaty oraz sądy w zamian wartości logicznych. Jednocześnie sądy traktuje się tutaj jako intensje zdań.

Propozycjonalne sieci semantyczne są modelem *systemu przekonaniowego* podmiotu wyposażonego w kompetencję języka naturalnego. Dowolny podmiot poznający reprezentujący wiedzę w postaci propozycjonalnych sieci semantycznych musi być w stanie reprezentować swoje własne przekonania: zarówno w postaci wyraźnej (świadomej) – *Jestem przekonany (przekonana), że „P”*, jak również w postaci implicytnej (nieświadomej) – gdy podmiot stwierdza po prostu „P”³⁶.

³⁶Wówczas np. współrozmówca danego podmiotu musi być w stanie powiedzieć, iż podmiot ten żywi przekonanie „P”.

Jednocześnie podmiot poznający powinien mieć możliwość reprezentowania przekonania innych podmiotów, zarówno *de re* (*Jan jest przekonany, że Anna ma jakieś zwierzę*), *de dicto* (*Jan jest przekonany, że żona sąsiada ma jakieś zwierzę*), oraz *de se* (*Jan wierzy, że on (sam) ma jakieś zwierzę*, gdzie *on* jest przez Jana identyfikowane z *ja*)³⁷ System rozróżnia sądy, co do których żywi przekonanie [*asserted*] i sądy jedynie reprezentowane w sieci. Jest to istotne rozróżnienie ze względu na przeprowadzane w systemie inferencje. Wiedza dotycząca faktów (łuk wiedzy³⁸) traktowana będzie jako *poprawne* przekonanie, co oznacza iż podmiot poznawczy posiadający wiedzę o pewnym fakcie jest przekonany, iż:

1. ktoś jest przekonany o tym fakcie,
2. fakt ten jest prawdziwy oraz
3. ten ktoś jest przekonany o tym fakcie z poważnych powodów (by wykluczyć np. zwykłe zgadywanie czy przypuszczenie).³⁹

Stosownie do wymienionych wyżej przekonań *de re*, *de dicto* oraz *de se* można również budować węzły reprezentujące wiedzę systemu.

5.3.4.2 Pierwotne jednostki semantyczne i kompozycjonalność w SNePS

Budując reprezentacje wiedzy w postaci sieci semantycznych, opieram się na architekturze poznawczej systemu SNePS. Architektura ta dostarcza pewnych ram, wymaga jednak od projektanta systemu kompetencji języka naturalnego, lub innego systemu poznawczego zobowiązania się do pewnej ontologii (w sensie wyeksplikowanym powyżej). Sieci propozycjonalne takiego typu, jak wykorzystywane w tym systemie mają dwie własności, na które uwagę zwracał już Quillian w swoich modelach pamięci semantycznej:

[N]ie istnieją w sieciach pojęcia „pierwotne”. Wszystko jest po prostu definiowane w

kategoriach pewnej uporządkowanej konfiguracji innych elementów w pamięci [Quillian 1968, s. 239]

³⁷Problematykę reprezentacji przekonań w propozycjonalnych sieciach semantycznych omawiają m.in. [Rapaport 1986, Rapaport, Shapiro, Wiebe 1986]

³⁸Budując bardziej rozwinięty system sieci semantycznych rozróżnić należy dwie relacje „wiedzieć”: Jedna z nich dotyczyć będzie wiedzy o faktach, jak np. *Jan wie, że wiosna rozpoczęła się 20 marca 2005 roku o godzinie 13³³* od wiedzy w sensie bycia zaznajomionym z przedmiotem/osobą, np. *Jan zna numer telefonu Anny*. Rozróżnienie to pojawia się m.in. przy analizie problemu numerów telefonu, przedstawionego przez McCarthy’ego.

W związku z ograniczeniami formalizacyjnymi wszędzie należałoby używać jednej postaci zapisu łuku wiedzy. W poniższych przykładach łuk ten opatrzę etykietą w postaci czasownika w 3 os. l.poj. i stosownie do tego będę dobierał przykłady, aby uniknąć takich wyrażen jak *Jan wiedzieć (żona Pawła równoważność mieszkanka Lublina)*.

³⁹Por. [Maida, Shapiro 1985, s. 176–177].

[P]ełne pojęcie jest definiowane w modelu pamięci jako wszystkie węzły, do których można dotrzeć drogą wyczerpujących procesów śledzących, rozpoczynających swoje działanie w pierwotnym, nadrzędnym węźle [Quillian 1968, s. 238]

Powyższe wypowiedzi można zinterpretować jako stanowisko odrzucające istnienie pierwotnych jednostek semantycznych i zasady kompozycjonalności. W propozycjonalnych sieciach semantycznych nie występują pierwotne jednostki semantyczne. Jedynymi pierwotnymi elementami są etykiety luków wiążących węzły, które nie mają charakteru pojęciowego, pełnią funkcję syntaktyczną i nie mają *znaczenia*. Cecha ta odróżnia SNePS od omówionych na początku sekcji systemów CD i KL-ONE, opierających się na pierwotnych jednostkach semantycznych. Omawiana w paragrafie 5.3.3 ontologia ramowa dostarczać będzie sieciom semantycznym jedynie zestawu etykiet luków, ustalających wzajemne zależności między węzłami – pojęciami w sieci.

Po drugie, w sieciach semantycznych nie obowiązuje zasada kompozycjonalności Fregego. Wyrażenia mające znaczenie w sieciach, węzły, uzyskują swoje znaczenie z wyrażań w ramach których się pojawiają i z wyrażań, które na nie się składają. Tym samym znaczenie węzłów będzie się zmieniało, w zależności od wyrażań w których będą osadzone. W konsekwencji, wraz z dodawaniem nowych struktur do sieci semantycznej zmieniało się będzie również rozumienie węzłów powiązanych z dodawaną strukturą (a ponieważ znaczenie w sieciach ma charakter holistyczny, zmieniać się będzie – w mniejszym lub większym stopniu – znaczenie wszystkich węzłów).