

Rozmowa nie jest tekstem takim jak np. książka czy artykuł gazetowy.  
Cechy szczególne rozmowy to:

- zamiana ról mówcy i słuchacza,
- potwierdzanie,
- przekazywanie informacji nie wprost

W rozmowie komputera z człowiekiem, komputer musi wiedzieć, kiedy może zacząć mówić. Jest to przedmiotem **analizy rozmowy** (ang. *conversation analysis*, CA). Określa się **miejsca odpowiednie do przekazania głosu** (ang. *transition-relevance place*, TRP) – miejsca, w których struktura rozmowy pozwala na przekazanie głosu rozmówcy. Pojawiają się na końcu wypowiedzi, które nie są jednoznaczne ze zdaniem tekstu pisanego.

Do segmentacji tekstu na wypowiedzi stosuje się:

- **słowa-wskazówki,**
- **ciągi słów lub ich oznaczeń,**
- **prozodię.**

W przypadku rozmowy więcej niż dwóch osób, zabieraniem głosu rządzi reguła mówiąca, że w każdym TRP:

- 1 Jeżeli bieżący mówca wybrał A jako następnego mówcę, to A musi mówić jako następny.
- 2 Jeżeli bieżący mówca nie wybrał następnego, dowolny mówca może kontynuować rozmowę.
- 3 Jeżeli nikt nie podejmuje rozmowy, bieżący mówca może mówić dalej.

Są wypowiedzi, które bezpośrednio wyznaczają następnego mówcę. Razem z reakcjami na nie tworzą **przyległe pary** (ang. *adjacent pairs*), np. pytanie-odpowieź, powitanie-powitanie, pochwała-zbycie, prośba-odpowieź. Odpowiedzią może być cisza i wtedy jest to **cisza znacząca**. Może oznaczać odmowę udzielenia odpowiedzi lub odpowiedź, której pytający nie chciał otrzymać.

Szczególnie gdy mówca nie widzi twarzy słuchacza (trudne do osiągnięcia w przypadku rozmowy z komputerem), musi wiedzieć, że słuchacz go słyszy i rozumie. Może tego dokonywać przez wypowiedzi typu *hmm...*, *super!*, *co ty powiesz?* itp., przez wypowiedzenie (być może innymi słowami) całości lub części wypowiedzi mówcy, przez dokończenie wypowiedzi mówcy. Mogą też występować prośby to powtórzenie lub wyjaśnienie: *eee...*, *co?*, *nie wierzę!* itp.

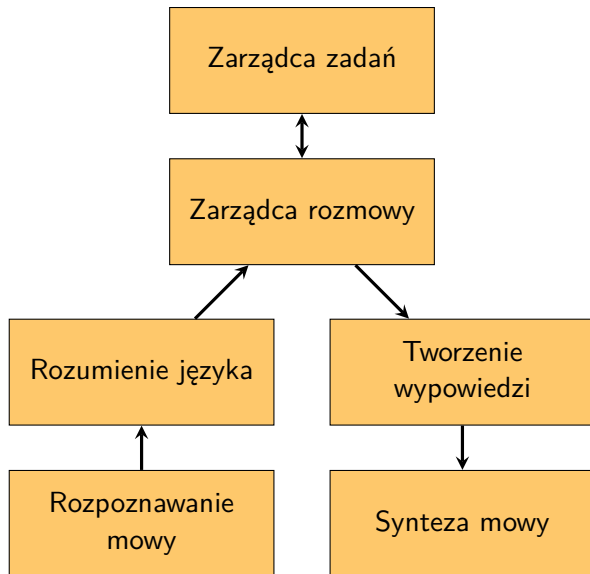
**Informacja:** O której godzinie chciałby pan jechać?

**Klient:** Mam spotkanie o 17:15.

:

**Informacja:** O której godzinie chciałaby pani jechać?

**Kientka:** Chyba nie ma wielu pociągów bezpośrednich.





Wykorzystuje się fakt, że dziedzina rozmowy jest zwykle ograniczona.  
Stosuje się modele:

- z wykorzystaniem automatów skończonych
- z wykorzystaniem gramatyk bezkontekstowych
- z wykorzystaniem statystyk (N-Grams)

Stosowane są:

- gramatyki semantyczne
- probabilistyczne gramatyki semantyczne
- Niejawne procesy Markowa
- semi-HMMs

Generowanie wypowiedzi dzieli się na:

- planowanie treści (zwykle włączane do zarządcy rozmowy)
- moduł generowania wypowiedzi

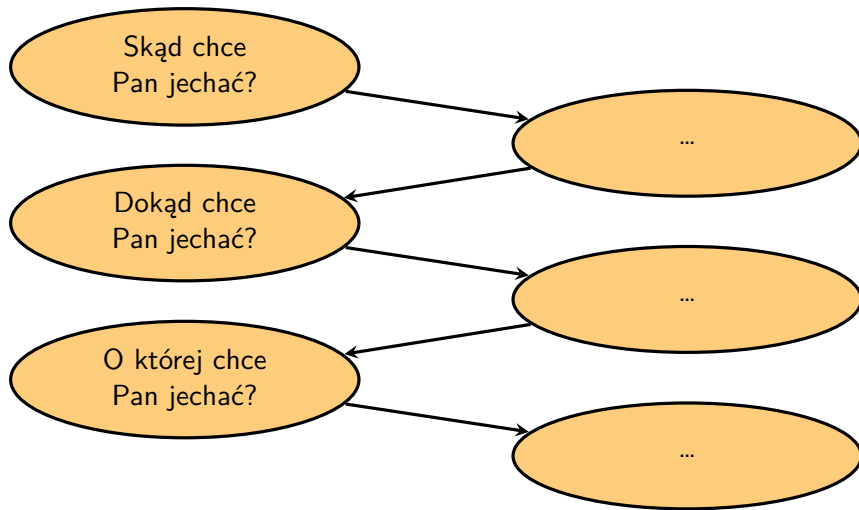
Moduł generowania wypowiedzi jest realizowany z wykorzystaniem szablonów lub z wykorzystaniem reprezentacji znaczenia przekazywanej z zarządcy rozmowy. W tej drugiej metodzie składa się z:

- 1 modułu planowania zdania
- 2 modułu realizacji powierzchniowej
- 3 modułu ustalania prozodii

Możliwe są różne sposoby realizacji zarządcy rozmowy:

- oparte o automaty skończone,
- oparte o szablony / reguły produkcji
- oparte o modele BDI (belief-desire-intention)

# Zarządca rozmowy oparty o automaty skończone



Tego typu zarządca rozmowy nazywany jest systemem z jednostronna inicjatywą (ang. *single initiative*) lub z inicjatywą systemu (ang. *system initiative*). Wystarczają jedynie do bardzo prostych zadań, typu obsługa konta w telefonii komórkowej. Zawodzą, kiedy użytkownik podaje więcej informacji w jednym zdaniu, podaje je w różnej kolejności itp.

# Zarządca rozmowy oparty o szablony

Ten rodzaj zarządcy (ang. *template-based*) nazywany jest także systemem opartym o ramy (ang. *frame-based*) lub opartym o reguły produkcji (ang. *production-rule*). Posiada szablony lub ramki z miejscami na różnego typu informację, np.:

Miejsce	Dodatkowe pytanie
Ze_stacji	„Skąd chce Pan jechać?”
Do_stacji	„Dokąd chce Pan jechać?”
O_godzinie	„O której chce Pan jechać?”
Klasa	...

Każda szczelina w ramce ma przyporządkowane dodatkowe pytanie, które może być zadane użytkownikowi, gdy brakuje kojarzonej z nim informacji.

Różne wypowiedzi użytkownika mogą powodować wypełnianie różnych miejsc w szablonie. Nie muszą być wypełniane w określonej kolejności. Jeżeli systemowi brakuje informacji, zapyta o nią użytkownika.

Tego typu zarządca rozmowy jest używany wtedy, gdy zakres możliwych wypowiedzi użytkownika jest ograniczony, ale użytkownik może chcieć je podawać w dowolnej kolejności, grupami itp. Ponieważ inicjatywa może przenosić się wielokrotnie pomiędzy systemem a użytkownikiem, ten typ zarządcy rozmowy jest nazywany typem z mieszaną inicjatywą (ang. *mixed initiative*).



System może dosłownie potwierdzać zrozumienie wypowiedzi użytkownika pytając, czy dobrze je zrozumiał. Zaletą jest to, że użytkownik może po prostu odpowiedzieć „nie”. W potwierdzaniu nie wprost system może powtórzyć informację jako część innego pytania.

Użytkownicy postrzegają jako naturalne szybkie prośby o powtórzenie („Słucham?”, „Przepraszam, nie zrozumiałem”). Gdy wypowiedź użytkownika jest odrzucona po raz drugi, system, zamiast powtórzyć pytanie, może podać więcej informacji, np. zasugerować składnię.

Wypowiedzi użytkownika można interpretować jako podające fakty, jako pytania, prośby/żądania, sprawdzenia itp. Jak było to widać na przykładach kilka folii wcześniej, czasami zdanie oznajmujące oznacza pytanie lub żądanie.

Możliwe są dwa podejścia do tego zagadnienia: wykorzystujące wnioskowanie (ang. *plan inference model*) i wykorzystujące wskazówki (ang. *cue model*).

W systemie tego typu wypowiedź użytkownika jest interpretowana za pomocą całego zestawu reguł nazywanego w oparciu o model BDI (ang. *belief, desire, and intention* - osąd, pragnienie i zamiar). Pojęcia te definiowane są w sposób formalny za pomocą predykatów pierwszego rzędu. Mówca  $S$  wierzy w prawdziwość  $P$  jest wyrażane jako  $B(S, P)$ . Pewna liczba aksjomatów pozwala wnioskowanie o osądach, np.  $B(A, P) \wedge B(A, Q) \Rightarrow B(A, P \wedge Q)$ . Wiedza jest reprezentowana jako prawdziwa wiara:

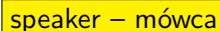
$$\text{KNOW}(S, P) \equiv P \wedge B(S, P)$$

believes – wierzy, sądzi

W systemie tego typu wypowiedź użytkownika jest interpretowana za pomocą całego zestawu reguł nazywanego w oparciu o model BDI (ang. *belief, desire, and intention* - osąd, pragnienie i zamiar). Pojęcia te definiowane są w sposób formalny za pomocą predykatów pierwszego rzędu. Mówca  $S$  wierzy w prawdziwość  $P$  jest wyrażane jako  $B(S, P)$ . Pewna liczba aksjomatów pozwala wnioskowanie o osądach, np.  $B(A, P) \wedge B(A, Q) \Rightarrow B(A, P \wedge Q)$ . Wiedza jest reprezentowana jako prawdziwa wiara:

$$\text{KNOW}(S, P) \equiv P \wedge B(S, P)$$

speaker – mówca



W systemie tego typu wypowiedź użytkownika jest interpretowana za pomocą całego zestawu reguł nazywanego w oparciu o model BDI (ang. *belief, desire, and intention* - osąd, pragnienie i zamiar). Pojęcia te definiowane są w sposób formalny za pomocą predykatów pierwszego rzędu. Mówca  $S$  wierzy w prawdziwość  $P$  jest wyrażane jako  $B(S, P)$ . Pewna liczba aksjomatów pozwala wnioskowanie o osądach, np.  $B(A, P) \wedge B(A, Q) \Rightarrow B(A, P \wedge Q)$ . Wiedza jest reprezentowana jako prawdziwa wiara:

$$\text{KNOW}(S, P) \equiv P \wedge B(S, P)$$

## REZERWACJA-MIEJSCÓWKI:

**Ograniczenia:**  $\text{Agent}(A) \wedge \text{Pociąg}(P) \wedge \text{Klient}(K)$

**Warunki wstępne:**  $\text{Know}(A, \text{data-odj}(P)) \wedge$   
 $\text{Know}(A, \text{czas-odj}(P)) \wedge \dots$

**Wynik:**  $\text{Miejscówka-zarezerwowana}(A, K, P)$

**Ciało:**  $\text{dokonaj-rezerwacji}(A, K, P)$

**INFORM(S,H,P):**

**Ograniczenia:**  $Mówca(S) \wedge Słuchacz(H) \wedge$   
 $Twierdzenie(P)$

**Warunki:**  $Know(S, P) \wedge W(S, INFORM(S, H, P))$

**Wynik:**  $Know(H, P)$

**Ciało:**  $Know(H, W(S, Know(H, P)))$

wants – chce

**INFORM(S,H,P):**

**Ograniczenia:**  $Mówca(S) \wedge Słuchacz(H) \wedge$   
 $Twierdzenie(P)$

**Warunki:**  $Know(S, P) \wedge W(S, INFORM(S, H, P))$

**Wynik:**  $Know(H, P)$

**Ciało:**  $Know(H, W(S, Know(H, P)))$



## Reguła “Wynik-działanie”:

Dla wszystkich uczestników  $S$  i  $H$  jeśli  $Y$  jest wynikiem działania  $X$  i jeśli  $H$  uważa, że  $S$  pragnie wykonania  $X$ , to prawdopodobnym jest, że  $H$  uważa, że  $S$  pragnie uzyskać  $Y$ .

## Reguła “Wiedza-Pragnienie”:

Dla wszystkich uczestników  $S$  i  $H$  jeśli  $H$  sądzi, że  $S$  chce KNOWIF( $P$ ), to  $H$  uważa, że  $S$  chce spełnienia  $P$ :

$$B(H, W(S, \text{KNOWIF}(S, P))) \stackrel{\text{ppod.}}{\implies} B(H, W(S, P))$$