ZESZYTY NAUKOWE POLITECHNIKI ŚLĄSKIEJ 1999
Seria: INFORMATYKA z. 36 Nr kol. 1414

Stanisław WIDEŁ, Halina RYCHLICKA
Politechnika Śląska, Instytut Informatyki

ALGORYTMY TRANSLACJI PODSTAWOWYCH FORMATÓW DANYCH NA POSTAĆ GŁOSOWĄ DLA INTERAKCYJNYCH SYSTEMÓW TELEFONICZNYCH

Streszczenie. Standardowy aparat telefoniczny może być zastosowany w systemach komputerowych jako dodatkowe urządzenie wejścia-wyjścia dla informacji fonicznej. W artykule autorzy podają przykłady zastosowania systemów głosowych. Interakcyjna Odpowiedź Głosowa to aplikacja, która poprzez interfejs foniczny udostępnia użytkownikowi dane zawarte w Bazach Danych. Autorzy podają klasyfikację algorytmów mogących służyć do translacji na postać foniczną danych różnych formatów przechowywanych w Bazach Danych.

POLISH GRAMMAR ALGORITHMS IMPLEMENTATION FOR TRANSLATION FILEDS OF THE DATA BASE TO THE INTERACTIVE VOICE RESPONSE

Summary. Standard telephone can be used as a device for interacting with a computer. Interactive Voice Response System (IVR) is a computer application that provides the voice user interface for Data Base access over a phone. For developing IVR system, it is a very important to understand the language grammar role. In this paper we present classification of the algorithm to follow the Polish grammar rules. This algorithms set can be used in IVR system for translation well know fields format of the Data Base to the Interactive Voice Response.

1. Wstęp

Rozwój współczesnej technologii informatycznej zmierza do coraz bardziej przyjaznej człowiekowi komunikacji z komputerem m.in. za pośrednictwem głosu. Jest to jeden z istotnych kierunków rozwijającej się techniki komputerowej. Wprawdzie prace w tej dziedzinie są prowadzone intensywnie - jednak ciągle nie istnieją jeszcze powszechnie przyjęte

rozwiązania programowo-sprzętowe. Jednym z rozwiązań jest wykorzystanie standardowego aparatu telefonicznego jako urządzenia zewnętrznego komputera.

Systemy, w których użytkownik komunikuje się z komputerem poprzez mowę, będziemy dalej nazywać *informatycznymi systemami głosowymi*. Wśród tych systemów można wyróżnić rozwiązania oparte na standardowym aparacie telefonicznym – wówczas systemy te będziemy nazywać *Interakcyjnymi Systemami Telefonicznymi* (IST).

2. Klasyfikacja Interakcyjnych Systemów Telefonicznych ze względu na zastosowanie

Zastanawiając się, w jaki sposób można wykorzystać IST lub w jakiego rodzaju aplikacjach mogą być stosowane – przeprowadziliśmy następujący podział: *Audiotekst*, *Automatyczny Operator*, *Poczta Głosowa*, *Interakcyjna Odpowiedź Głosowa*.

Audiotekst. Aplikacja tego typu korzysta z predefiniowanych, wcześniej nagranych komunikatów głosowych, których użytkownik wysłuchuje. Integralną częścią systemu audiotekstu jest funkcja wyboru informacji, jaką użytkownik chce usłyszeć. Zwykle jest ona realizowana przez wybranie określonej cyfry w systemie tonowym. Oczywiście, możliwe jest korzystanie z technik bardziej zaawansowanych technologicznie, np. rozpoznawanie głosu, jednak ze względu na koszty implementacyjne takiej techniki, przy jednoczesnym, prawie powszechnym dostępie do aparatu telefonicznego z wybieraniem tonowym – rozwiązania oparte na rozpoznawaniu mowy są jak na razie mało popularne.

Automatyczny Operator. Taką aplikację instaluje się przeważnie w dużych centralach wewnętrznych. Jej zadaniem jest bezpośrednie odbieranie przychodzących rozmów po to, by dzwoniącemu użytkownikowi podać informację, w jaki sposób może się połączyć z odpowiednią osobą. Dla przykładu: automatyczny operator może połączyć dzwoniącego z pierwszą wolną osobą z danego działu. Możliwe jest również zadanie pewnych pytań użytkownikowi w celu wstępnego sklasyfikowania typu rozmowy i połączenia z kompetentną osobą. Użytkownik dzwoniący do systemu *Automatyczny Operator* nie musi znać (ani pamiętać) numerów wewnętrznych.

Poczta Głosowa. Jest to aplikacja pozwalająca zdefiniować skrzynkę, w której można pozostawić wiadomość w postaci głosowej. Właściciel skrzynki może, również przez telefon, wysłuchać pozostawionych dla niego komunikatów i zarządzać nimi poprzez np. rozsyłanie ich do skrzynek innych użytkowników, kasowanie niepotrzebnych komunikatów itp.

Automatyczny Operator i Poczta Głosowa jako funkcja wbudowana w wewnętrzną centralę telefoniczną. Aplikacje te są ściśle powiązane ze sposobem działania centrali telefonicznej. Można zatem powiedzieć, że naturalnym środowiskiem ich funkcjonowania jest centrala telefoniczna danego typu. Z punktu widzenia rozwoju techniki należy się spodziewać

implementacji aplikacji typu *Automatyczny Operator i Poczta Głosowa* bezpośrednio przez centralę telefoniczną.

Realizacja tych funkcji w postaci odrębnego systemu komputerowego, niepowiązanego z centralą, ma te wady, że system jako zewnętrzny w stosunku do centrali może uzyskać tylko takie same informacje, które otrzymuje użytkownik podczas telefonowania, czyli – dla przykładu – system dysponuje tylko informacjami typu:

- sygnał „podniesienia" linii,
- sygnał dzwonienia,
- sygnał zajętości linii,
- sygnał zwolnienia linii.

System zintegrowany z centralą może dysponować większą liczbą danych o przebiegu połączenia i przez to może być lepiej dostosowany do jej specyficznych potrzeb.

W chwili obecnej rozszerzenie funkcjonalności danej centrali telefonicznej można zrealizować w oparciu o zewnętrzny autonomiczny system. W przyszłości należy się spodziewać jednak central telefonicznych, które będą zawierały w sobie funkcje *Automatycznego Operatora i Poczty Głosowej*.

Interakcyjna Odpowiedź Głosowa. Jest to rodzaj aplikacji, w której komputer udziela informacji głosowej, korzystając z zewnętrznych baz danych. Przykładem aplikacji tego rodzaju jest, np.:

- sprawdzanie stanu konta w systemie bankowym,
- informacja o zaległych i zapłaconych podatkach,
- informacja o numerach rejestracyjnych pojazdów i o ich właścicielach,
- informacja o aktualnej stopie procentowej w funduszu powierniczym,
- informacja o stanie konta emerytalnego w systemie ZUS.

Ogólnie można przyjąć, że interfejs głosowy jest jednym z możliwych interfejsów dostępu do specyficznej bazy danych. W zależności od rodzaju bazy danych możemy mówić o specyficznej aplikacji, jednak ogólne narzędzia programistyczne pozostają uniwersalne.

Aplikację, która udziela informacji głosowej z bazy danych, można sklasyfikować jako typ, który dalej będziemy nazywać *Interakcyjną Odpowiedzią Głosową*.

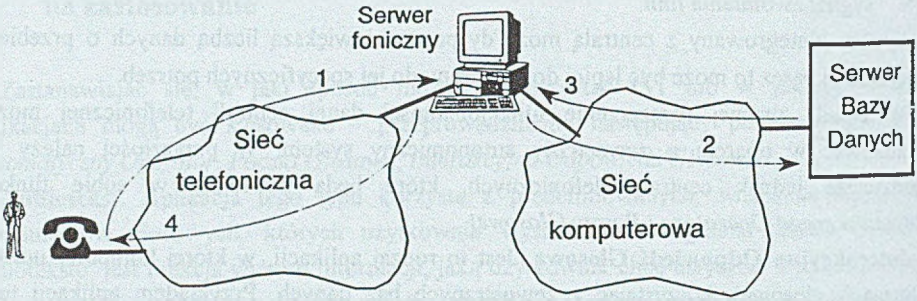
Zdaniem autorów przyszły rozwój systemów głosowych będzie się opierał głównie na aplikacjach typu *Interakcyjna Odpowiedź Głosowa*.

W związku z powyższym, jednym z kluczowych zagadnień, jakie trzeba rozwiązać, jest translacja danych przechowywanych w bazach danych na postać głosową. W dalszej części artykułu będziemy się zajmować tym zagadnieniem.

3. Integracja Informatycznych Systemów Telefonicznych z sieciami komputerowymi

W chwili obecnej sieci telefoniczne i komputerowe nie są zintegrowane z punktu widzenia usług, jakich dostarczają użytkownikowi.

Informatyczne Systemy Telefoniczne mogą tworzyć tzw. serwery foniczne, które będą dostarczać usług dostępnych poprzez standardowy aparat telefoniczny. Na rys. 1. przedstawiono przykładową architekturę systemu dostępu do bazy danych za pośrednictwem telefonu.



Rys. 1. Przykładowa Architektura Systemu dostępu do bazy danych za pośrednictwem telefonu

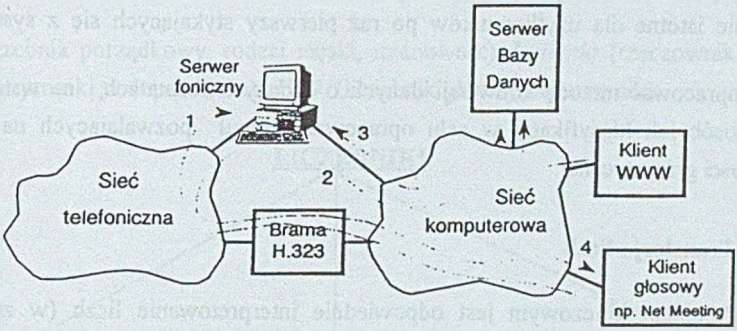
Fig. 1. Interactive Voice Response architecture

Z punktu widzenia niezawodności systemu jest bardzo wskazane rozbicie funkcji realizacji dostępu fonicznego poprzez standardowy aparat telefoniczny i obsługi bazy danych. Serwer pokazany na rys. 1. korzysta z baz danych poprzez sieć lokalną. Architektura tego typu jest preferowana przez autorów w implementowanych przez nich systemach.

Należy stwierdzić, że wyodrębnienie serwera fonicznego w sensie osobnej platformy sprzętowo-programowej połączonej z siecią telefoniczną i siecią komputerową ma szereg zalet, z których wymienimy następujące:

- dedykowane środowisko sprzętowo-programowe jest łatwiejsze do testowania i wykrywania ewentualnych nieprawidłowości w działaniu oprogramowania,
- zostają zmniejszone wymagania zarówno sprzętowe, jak i programowe dla serwera fonicznego (dla przykładu – w serwerze fonicznym można zrezygnować ze środowiska graficznego, np. WINDOWS, gdyż użytkownik, korzystając z interfejsu fonicznego, nie wykorzystuje w żaden sposób funkcji graficznych, a zrezygnowanie ze środowiska graficznego pozwoli lepiej wykorzystać moc obliczeniową do obsługi w czasie rzeczywistym wielu interfejsów fonicznych),
- wyodrębnienie serwera fonicznego pozwala na większą niezależność i instalowanie w środowiskach heterogenicznych sieci komputerowych dla różnych systemów operacyjnych i systemów z bazami danych.

Na rys. 2. przedstawiono przykładową architekturę systemu umożliwiającego dostęp do bazy danych za pomocą interfejsu fonicznego ze stacji roboczej pracującej w sieci komputerowej.



Rys. 2. Przykładowa Architektura Systemu dostępu do bazy danych za pomocą oprogramowania telekonferencyjnego
Fig. 2. Network Client Access to the IVR Systems Architecture

Do serwera bazy danych użytkownik komputera może mieć dostęp poprzez interfejs przeglądarki WWW, jak również za pomocą standardowego oprogramowania. Możliwy jest również dostęp z sieci komputerowej do serwera fonicznego. Wykorzystując bramę H.323 można połączyć się z siecią telefoniczną i następnie skorzystać z serwera fonicznego, który udostępni w postaci głosowej dane przechowywane w serwerze bazy danych.

Rozwiązania tego rodzaju, przedstawione na rys. 2., wydają się bardzo skomplikowane, jednak zdaniem autorów staną się bardzo popularne w miarę upowszechniania się transportu głosu przez sieć komputerową. Szczególnie dynamicznie rozwijającą się dziedziną w tym zakresie jest transport głosu poprzez sieć INTERNET (ang. VoIP Voice over IP).

4. Translacja danych w postaci liczb na postać głosową

Przystępując do opracowania zbioru algorytmów, których zadaniem byłoby konwertowanie danych alfanumerycznych o znanym predefiniowanym formacie na postać głosową należy opracować klasyfikację i metody określenia i przydziału odpowiednich algorytmów, którym na wejściu podaje się dane, aby na wyjściu otrzymać poprawnie gramatyczny ciąg słów fonicznych.

Przystępując do zagadnienia syntezy mowy możemy rozróżnić dwa podejścia: pierwsze polega na rozpoznawaniu formatów danych i ich zamianie na postać foniczną, przy wykorzystaniu predefiniowanego zbioru słów. Podejście takie pozwala na lepsze zrozumienie przez słuchacza syntezowanej mowy. Niestety, podejście takie nie jest tak uniwersalne jak druga metoda, a mianowicie - metoda syntezy mowy bezpośrednio z tekstu na podstawie predefiniowania fonemów. Metoda konwersji *tekst-mowa* generuje fonemy na podstawie

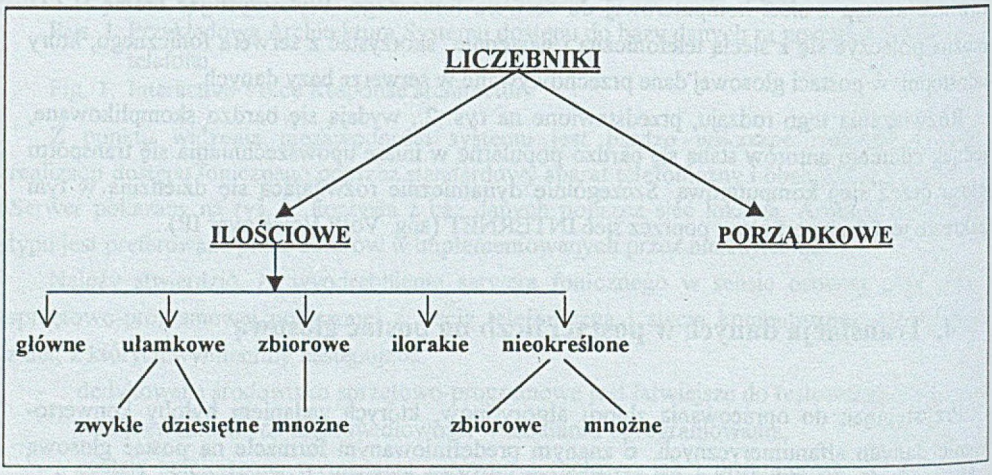
tekstu. Dla autorów w niniejszym artykule interesującym podejściem jest tworzenie postaci fonicznej z predefiniowanych, nagranych wcześniej, słów. Zgodnie z aktualnie dostępną technologią metoda ta zapewnia lepszą jakość i zrozumienie generowanej mowy, co jest szczególnie istotne dla użytkowników po raz pierwszy stykających się z systemami tego rodzaju.

Aby opracować metodę konwersji danych o zadanych formatach, na wstępie musimy podać sposób ich klasyfikacji w celu opracowania reguł pozwalających na zachowanie poprawności gramatycznej.

1.1. Translacja liczb

Zagadnieniem kluczowym jest odpowiednie interpretowanie liczb (w zależności od kontekstu) w celu ich zamiany na postać głosową. Częścią mowy określającą w języku liczby są *liczebniki*. Zgodnie z literaturą możemy przeprowadzić następującą klasyfikacje liczebników:

- ogólną,
- słowotwórczą.



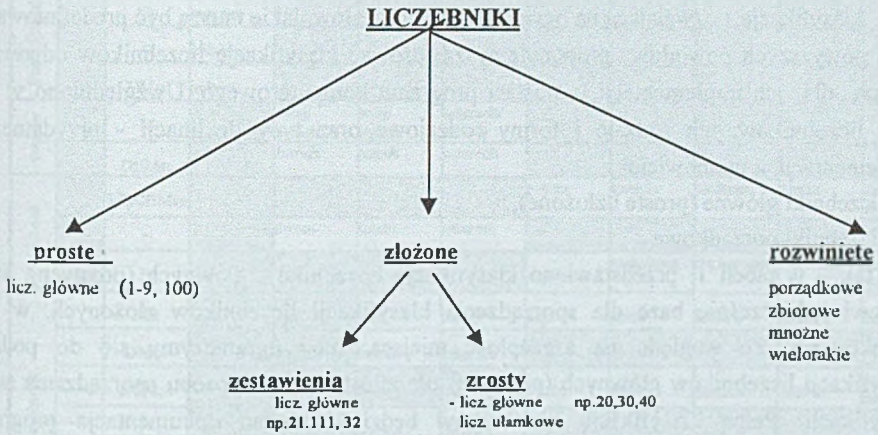
Rys. 3. Ogólna klasyfikacja liczebników
Fig. 3. Polish Cardinal numbers general classification

Klasyfikacja przedstawiona na rys. 3. powstała na podstawie ogólnego, przyjętego w gramatyce, podziału liczebników na liczebniki określające konkretną lub przybliżoną liczbę osób, przedmiotów [np. *dwie* panie, *cztery* krzesła (liczebniki główne), *pół* kilograma, *jedna dziesiąta* sekundy (liczebniki ułamkowe), *troje* dzieci (liczebnik zbiorowy), *potrójny* zamek (liczebnik zbiorowy – mnożny), *dwojaka* rada (liczebnik iloraki), *kilka* dni (liczebnik nieokreślony), *kilkoro* dzieci (liczebnik nieokreślony – zbiorowy), *wielokrotny* zabójca

(liczebnik nieokreślony – mnożny)] oraz ich kolejność – czyli miejsce w szeregu, np. *pierwszy* dzień miesiąca (liczebnik porządkowy).

Dla przykładu - data: 8 kwietnia 1999 według rozbioru gramatycznego typu 1 (rys. 3.) – to:

ósmy (liczebnik porządkowy, rodzaj męski, mianownik), *kwietnia* (rzeczownik – nazwa własna – rodzaj męski w dopełniaczu), *1999* (liczebnik główny, mianownik).



Rys. 4. Klasyfikacja liczebników pod względem słowotwórczym
Fig. 4. Numbers generation classification

Ten sam przykład dla klasyfikacji słowotwórczej - data: **8 kwietnia 1999** - rozbiór gramatyczny typu 2 – to: *ósmy* (liczebnik rozwinięty), *kwietnia* (rzeczownik – nazwa własna, rodzaj męski w dopełniaczu), *1999* (liczebnik złożony – zestawienie).

Przedstawione powyżej klasyfikacje liczebników z punktu widzenia potrzeb ich konwersji na postać foniczną nie są przydatne, gdyż bezpośrednio nie definiują algorytmów, jakie należy zaimplementować i stosować przy zamianie, np. daty na postać foniczną. Nie mogą być one podstawą do budowy programu komputerowego, gdyż:

- klasyfikacja typu 1 (rys. 3) dzieli liczebniki pod względem gramatycznym na określające liczbę lub miejsce w szeregu, zaś klasyfikacja 2 typu (rys. 4.) bierze pod uwagę budowę słowotwórczą liczebników – natomiast liczebniki niezależnie od typu wymienionych klasyfikacji mogą być konwertowane na postać foniczną przy użyciu tego samego algorytmu i z tego względu nie ma potrzeby w systemie komputerowym odróżniania liczebnika, np. 1 [jeden – liczebnik ilościowy - główny (prosty)] od 111. [(sto jedenaście – liczebnik główny (złożony-zestawienie)],

- rożne formy gramatyczne liczebników, jak np. rodzaje i przypadki, które wymuszają konieczność zastosowania końcówek właściwych dla danego rodzaju lub końcówek fleksyjnych, nie są rozróżniane z punktu widzenia klasyfikacji gramatycznej lub słowotwórczej, gdyż zakłada się, że liczebniki zakwalifikowane do danej kategorii podlegają odmianie,
- z punktu widzenia komputerowej translacji rzeczowników - nazw własnych (np. nazwa miesiąca - *kwiecień*) na postać foniczną należy przeprowadzić dużo dokładniejszą klasyfikację pozwalającą na ograniczenie zbioru słów, jakie muszą być predefiniowane.

Z powyższych powodów proponujemy zastosować klasyfikację liczebników odpowiedniejszą dla ich implementacji w postaci programu komputerowego. Uwzględniono w niej typy liczebników, ich rodzaje i formy rodzajowe oraz typy deklinacji - przydatne do implementacji, a mianowicie:

- liczebniki główne (proste i złożone),
- liczebniki porządkowe.

I tak – w tabeli 1. przedstawiono klasyfikację liczebników głównych (prostych), która stanowi jednocześnie bazę dla sporządzenia klasyfikacji liczebników złożonych, w tym ułamkowych. Ze względu na szczupłość miejsca, tutaj ograniczymy się do podania klasyfikacji liczebników głównych (prostych) dla zilustrowania sposobu sporządzania takiej klasyfikacji. Pełną klasyfikację liczebników będzie zawierać dokumentacja programu komputerowego, który będzie opracowany m.in. na tej podstawie.

Tabela 1

Klasyfikacja liczebników głównych

Liczebniki główne				1	2	3	4	5	6	7	8	9
			Przypadki ⇩									
	Rodzaj Męski	żywot.	M	jeden								
			DB	jednego								
		nież.	MB	jeden								
			D	jednego								
			C	jednemu								
			NMsc	jednym								
	Rodz. Żeński	żywot	M	jedna								
			DCBMsc	jednej								
		nież.	MB	jedna								
			DCMsc	jednej								
			N	jedną								
			MB		dwie							
			DMsc		dwóch							

Liczebniki główne				1	2	3	4	5	6	7	8	9
			Przypadki ⇩									
			C		dwóm							
			N		dwiema							
	Rodzaj Nijaki		MB	jedno								
			D	jednego								
			C	jednemu								
			NMsc	jednym								
	FMęskos		M		dwaj/ dwóch	trzej/ trzech	czterech/ czterej					
			DBMsc		dwóch	trzech	czterech					
	FMęskoos.		MDCBMsc					pięciu	sześciu	siedmiu	ośmiu	dziewięciu
			C		dwom	trzem	czterem					
			N					pięcioma	sześcioma	siedmioma	ośmioma	dziewięcioma
	FNiemęskoosobowa		MB		dwa	trzy	cztery	pięć	sześć	siedem	osiem	dziewięć
			DMsc		dwóch	trzech	czterech					
			DCMsc					pięciu	sześciu	siedmiu	ośmiu	dziewięciu
			C		dwom	trzem	czterem					
			N		dwoma	trzema	czterema	pięcioma	sześcioma	siedmio-ma	ośmioma	dziewięcioma

Tabela 2

Algorytmy czytania liczebników głównych (prostych)

Lp.	Liczebniki główne ⇨		1	2	3	4	5	6
	Nazwa ⇩	Przy-padki	Nazwa ⇩	Nazwa ⇩				
			LGRM	LGFMos				
1.	CzLGdoRzŻywRM	M	jeden	dwaj/dwóch	trzejtrzech	czterej/ czterech	pięciu	sześciu
2.		D	jednego	dwóch	trzech	czterech	pięciu	sześciu
3.		C	jednemu	dwóm	trzem	czterem	pięciu	sześciu
4		B	jednego	dwóch	trzech	czterech	pięciu	sześciu
5.		N	jednym	dwoma	trzema	czterema	pięcioma	sześcioma
6.		Msc	jednym	dwóch	trzech	czterech	pięciu	sześciu
			LGRM	LGFNmos				
7.	CzLGdoRzNżywRM	M	jeden	dwa	trzy	cztery	pięć	sześć

Lp.	Liczebniki główne		1	2	3	4	5	6
	Nazwa	Przy-padki	Nazwa	Nazwa				
8.		D	jednego	dwóch	trzech	czterech	pięciu	sześciu
9.		C	jednemu	dwom	trzem	czterem	pięciu	sześciu
10.		B	jeden	dwa	trzy	cztery	pięć	sześć
11.		N	jednym	dwoma	trzema	czterema	pięcioma	sześcioma
12.		Msc	jednym	dwoma	trzema	czterema	pięcioma	sześcioma
			LGRŻń	LGRŻń	LGFMos			
13.	CzLGdoRzŻywRŻń	M	jedna	dwie	trzy	cztery	pięć	sześć
14.		D	jednej	dwóch	trzech	czterech	pięciu	sześciu
15.		C	jednej	dwóm	trzem	czterem	pięciu	sześciu
16		B	jednej	dwie	trzy	cztery	pięć	sześć
17		N	jedną	dwiema	trzema	czterema	pięcioma	sześcioma
18.		Msc	jednej	dwóch	trzech	czterech	pięciu	sześciu
			LGRŻń	LGRŻń	LGFNmos			
19.	CzLGdoRzNżywRŻń	M	jedna	dwie	trzy	cztery	pięć	sześć
20.		D	jednej	dwóch	trzech	czterech	pięciu	sześciu
21.		C	jednej	dwóm	trzem	czterem	pięciu	szesciu
22.		B	jedna	dwie	trzy	cztery	pięć	sześć
23.		N	jedną	dwiema	trzema	czterema	piecioma	sześcioma
24.		Msc	jednej	dwóch	trzech	czterech	pięciu	sześciu
			LGRNi	LGFNmos				
25.	CzLGdoRzNżywRNi	M	jedno	dwa	trzy	cztery	pięć	sześć
26.		D	jednego	dwóch	trzech	czterech	pięciu	sześciu
27		C	jednemu	dwóm	trzem	czterem	pieciu	sześciu
28.		B	jedno	dwa	trzy	cztery	pięć	sześć
29.		N	jednym	dwoma	trzema	czterema	pięcioma	sześcioma
30.		Msc	jednym	dwóch	trzech	czterech	pięciu	sześciu

Objaśnienia: Cz – czytanie LG – liczebnika głównego do RzŻyw – rzeczownika żywotnego RM – rodzaju męskiego, RŻń – rodzaju żeńskiego, RNi – rodzaju nijakiego; RzNżyw – rzeczownik nieżywotny; LGRM – liczebnik główny rodzaj męski;

LGRŻń – liczebnik główny rodzaj żeński; LGRNi – liczebnik główny rodzaj nijaki; LGFMos – liczebnik główny forma męsko-osobowa, LGFNmos – liczebnik główny forma niemęsko-osobowa

Przykłady:

CzLGdoRzŻywRM – M (Czytanie Liczebnika Głównego do Rzeczownika Żywotnego – Rodzaju Męskiego w Mianowniku)

Na połączenie czeka *jeden* (LGRM-Mianownik) *użytkownik* (rzeczownik żywotny, rodzaj męski, mianownik)

CzLGdoRzNżywRM – M (Czytanie Liczebnika Głównego do Rzeczownika Nieżywotnego Rodzaju Męskiego) – Mianownik

W skrzynce są *dwa* (LGRM-Mianownik) *komunikaty* (rzeczownik nieżywotny, rodzaj męski, mianownik).

CzLGdo RzŻywRŻń – M (Czytanie Liczebnika Głównego do Rzeczownika Żywotnego Rodzaju Żeńskiego – Mianownik)

Dwie (LGRŻń-Mian) *abonentki* (rzeczownik żywotny, rodzaj żeński – mianownik) nie odpowiadają na listy.

CzLGdoRzNżywRŻń – M (Czytanie Liczebnika Głównego do Rzeczownika Nieżywotnego Rodzaju Żeńskiego)

Trzy (LGRŻń –Mian) *wiadomości* (rzeczownik nieżywotny, rodzaj żeński, mianownik) nie zostały odebrane.

CzLGdoRzNżywRNi – M (Czytanie Liczebnika Głównego do Rzeczownika Nieżywotnego Rodzaj Nijaki – Mianownik)

Jedno (LGRNi – M) *łącze* (rzeczownik nieżywotny, rodzaj nijaki, mianownik) jest niesprawne.

Jak widać, tabela 2. zawiera 30 pozycji. Jest to pięć rodzajów i form rzeczowników odmienianych przez 6 przypadków. Dało to razem 30 algorytmów czytania liczebników głównych.

W tabeli 3 podano reguły doboru końcówek fleksyjnych rzeczowników w powiązaniu z liczebnikami.

Tabela 3

Dobór końcówek fleksyjnych liczebników głównych i rzeczowników

Przypadki	Licz. Gł.	Liczba poj.	Licz. Główne	Liczba mnoga	Liczba mnoga
		Rzeczownik		Rzeczownik	Rzeczownik
Mianownik	jeden	komunikat	dwa, trzy, cztery pięć, sześć, itd.	komunikaty	komunikatów (Dop.)
Dopełniacz	jednego	komunikatu	dwóch, trzech, czterech, pięciu	komunikatów	
Celownik	jednemu	komunikatowi	dwóm, trzem, czterem, pięciu	komunikatom	
Biernik	jeden	komunikat	dwa, trzy, cztery pięć, sześć, itd.	komunikaty	komunikatów (Dop.)

Przypadki	Licz. Gł.	Liczba poj.	Licz. Główne	Liczba mnoga	Liczba mnoga
		Rzeczownik		Rzeczownik	Rzeczownik
Narzędnik	jednym	komunikatem	dwoma, trzema, czterema, pięcioma	komunikatami	
Miejscownik	jednym	komunikacie	dwóch, trzech, czterech, pięciu	komunikatach	

Tabela 4

Klasyfikacja liczebników porządkowych

Liczebniki porządkowe				1.	2	3.	4.	5.	6.	7.
Rodzaje			Przypadki ⇩							
	RMęski	żyw.	M	pierwszy	drugi	trzeci	czwarty	piąty	szósty	siódmy
			DB	pierwszego	drugiego	trzeciego	czwartego	piątego	szóstego	siódmego
		nieżyw	MB	pierwszy	drugi	trzeci	czwarty	piąty	szósty	siódmy
			D	pierwszego	drugiego	trzeciego	czwartego	piątego	szóstego	siódmego
			C	pierwszemu	drugiemu	trzeciemu	czwartemu	piątemu	szóstemu	siódmemu
			NMsc	pierwszym	drugim	trzecim	czwartym	piątym	szóstym	siódmym
	RŻeński		M	pierwsza	druga	trzecia	czwarta	piąta	szósta	siódma
			DCMsc	pierwszej	drugiej	trzeciej	czwartej	piątej	szóstej	siódmej
			BN	pierwszą	drugą	trzecią	czwartą	piątą	szóstą	siódmą
	RNijaki		MB	pierwsze	drugie	trzecie	czwarte	piąte	szóste	siódme
			D	pierwszego	drugiego	trzeciego	czwartego	piątego	szóstego	siódmego
			C	pierwszemu	drugiemu	trzeciemu	czwartemu	piątemu	szóstemu	siódmemu
			NMsc	pierwszym	drugim	trzecim	czwartym	piątym	szóstym	siódmym

Tabela 5

Algorytmy czytania liczebników porządkowych

Liczebnik porządkowy ⇨		1.	2.	3.	4.	5.
Nazwa ⇩	Przy-	Nazwa ⇩				
	padki	Liczebnik Porządkowy Rodzaj Męski				
CzLPdoRzŻywRM	M	pierwszy	drugi	trzeci	czwarty	piąty
	D	pierwszego	drugiego	trzeciego	czwartego	piątego
	C	Pierwszemu	drugiemu	trzeciemu	czwartemu	piątemu
	B	Pierwszego	drugiego	trzeciego	czwartego	piątego
	N	Pierwszym	drugim	trzecim	czwartym	piątym
	Msc	Pierwszym	drugim	trzecim	czwartym	piątym

Liczebnik porządkowy ⇨		1.	2.	3.	4.	5.
Nazwa ⇩		Nazwa ⇩				
		Liczebnik Porządkowy Rodzaj Męski				
CzLPdoRzNżywRM.	M	Pierwszy	drugi	trzeci	czwarty	piąty
CzLPdoRzNżywRM.	D	Pierwszego	drugiego	trzeciego	czwartego	piątego
	C	Pierwszemu	drugiemu	trzeciemu	czwartemu	piątemu
	B	Pierwszy	drugi	trzeci	czwarty	piąty
	N	Pierwszym	drugim	trzecim	czwartym	piątym
	Msc	Pierwszym	drugim	trzecim	czwartym	piątym
		Liczebnik Porządkowy Rodzaj Żeński				
CzLPdoRzRŻń	M	pierwsza	druga	trzecia	czwarta	piąta
	D	pierwszej	drugiej	trzeciej	czwartej	piątej
	C	Pierwszej	drugiej	trzeciej	czwartej	piątej
	B	pierwszą	drugą	trzecią	czwartą	piątą
	N	pierwszą	drugą	trzecią	czwarta	piąta
	Msc	pierwszej	drugiej	trzeciej	czwartej	piątej
		Liczebnik Porządkowy Rodzaj Nijaki				
CzLPdoRzRNi	M	pierwsze	drugie	trzecie	czwarte	piąte
	D	pierwszego	drugiego	trzeciego	czwartego	piątego
	C	pierwszemu	drugiemu	trzeciemu	czwartemu	piątemu
	B	pierwsze	drugie	trzecie	czwarte	piąte
	N	pierwszym	drugim	trzecim	czwartym	piątym
	Msc	pierwszym	drugim	trzecim	czwartym	piątym

Objaśnienia: Cz – czytanie LP – liczebników porządkowych do RzŻyw – rzeczownika żywotnego RM – rodzaj męski; do RzNżyw – rzeczownika nieżywotnego; RŻń – rodzaj żeński, RNi – rodzaj nijaki; M – mianownik, D – dopełniacz, C – celownik, B – biernik, N – narzędnik, Msc - miejscownik

Przykłady:

1. CzLPdoRzŻywRM – M (Czytanie Liczebnika Porządkowego do Rzeczownika Żywotnego Rodzaju Męskiego – Mianownik)

   Zgłosił się *pierwszy* (LPRM – M) *użytkownik* (rzeczownik żywotny, rodzaj męski, mianownik).

2. CzLPdoRzNżywRM – M (Czytanie Liczebnika Porządkowego do Rzeczownika Nieżywotnego Rodzaju Męskiego – Mianownik)

   Nie doszedł *drugi* (LPRM – M) *komunikat* (rzeczownik nieżywotny, rodzaj męski, mianownik).

3. CzLPdoRzRŻń – M (Czytanie Liczebnika Porządkowego do Rzeczownika Rodzaju Żeńskiego – Mianownik)

   Druga (LPRŻń – M) *wiadomość* (rzeczownik rodzaju żeńskiego – mianownik) jest bardzo ważna.

4. CzLPdo RzRNi – M (Czytanie Liczebnika Porządkowego do Rzeczownika Rodzaju Nijakiego w mianowniku)

   Piąte (LPRNi – Mianownik) *łącze* (rzeczownik rodzaju nijakiego, mianownik) jest uszkodzone.

Z punktu widzenia klasyfikacji algorytmów zamiany liczebników z postaci alfanumerycznej na postać foniczną klasyfikacja liczebników przedstawiona w tabeli 1. nie nadawała się do uporządkowania i wprowadzenia jasnej klasyfikacji dla potrzeb przygotowania koniecznych algorytmów. Przedstawiając inaczej tę sytuację – nie można sporządzić klasyfikacji algorytmów odmiany liczebników w oderwaniu od rzeczownika z nim związanego. W tym celu sporządzono tabelę porządkującą odmiany liczebników w ich związku z rzeczownikami (tab. 3). Wynika z niej, że liczebniki główne 1, 2, 3, 4 zgadzają się z towarzyszącym im rzeczownikiem pod względem przypadku, natomiast liczebnikom od pięciu wzwyż w Mianowniku i Bierniku odpowiada rzeczownik w Dopełniaczu.

Znacznie prostsza jest klasyfikacja liczebników porządkowych z uwagi na to, że mają jednolitą formę rodzajową – taką jak przymiotniki i odmianę taką jak przymiotniki.

Na podstawie algorytmów sporządzonych dla liczebników głównych można ustalić algorytm czytania liczebników ułamkowych (ułamków dziesiętnych). Czytanie tych ułamków odbywa się w następujący sposób:

- ułamek zostaje podzielony na część całkowitą i część dziesiętną oddzieloną przecinkiem;
- część całkowitą stanowi liczebnik główny poddający się algorytmom ustalonym dla tych liczebników, natomiast część dziesiętna składa się z liczebnika głównego rodzaju żeńskiego i liczebnika porządkowego, np.:

<u>część całkowita</u>	<u>przecinek</u>	<u>część dziesiętna</u>		
CzLGdoRzNżywRMMian *jeden*	,	CzLGdoRzNżywRŻńMian	*jedna*	CzLPdoRzRŻńMian *dziesiąta*

Wykorzystując algorytmy opracowane dla liczebników głównych można podać algorytm czytania, np. daty i godziny:

- data: **1 maja 1999**

  algorytm: LPRMNżywMian. <u>*Pierwszy*</u> - RzNżywRMDop. - *maja* – LGRMNżywMian - <u>*tysiąc*</u>- LGFNmosMian *dziewięćset* – LGFNmosMian *dziewięćdziesiąt dziewięć*

- czas: godzina 13:30 minut:26 sekund

algorytm: RzNżywRŻńMian – *godzina* – LPRŻńMian. *trzynasta* – LGFNmosMian - *trzydzieści* RzLMnRŻńDop. *minut* –LGFNmosMian – *trzydzieści* – RzLmnRŻńDop. – *sekund.*

Podsumowując, należy stwierdzić, iż nasze podejście do translacji z postaci alfanumerycznej na postać foniczną opierało się głównie na założeniu, że badany system będzie posiadał skończoną liczbę słów lub fraz, które można przygotować w nagraniu studyjnym. Takie podejście w chwili obecnej gwarantuje najlepszą jakość generowanej mowy z punktu widzenia płynności i zrozumienia przez użytkowników, którzy po raz pierwszy spotykają się z systemami komputerowymi.

Po przeprowadzonych klasyfikacjach algorytmy translacji z postaci alfanumerycznej do postaci fonicznej sprowadzają się jedynie do wyszukiwania w tabeli odpowiedniej formy przygotowanej jako niepodzielny plik zawierający zdigitalizowaną mowę.

2. Wnioski

Przystępując do rozwiązania założonych zadań, autorzy założyli również praktyczny aspekt wykonania translacji danych z postaci alfanumerycznej na postać głosową: tak przeprowadzić klasyfikację, aby implementacja algorytmów była najprostsza. I można powiedzieć, że cel ten został osiągnięty.

Pojawiło się jednak wiele problemów przy opracowywaniu założonych algorytmów, które wykraczają poza ramy tego referatu. I tak – w dalszych pracach będzie niezbędna głęboka analiza związków składniowych, m.in. liczebników z rzeczownikami. Stworzy to możliwości opracowania bardziej zaawansowanych algorytmów translacji mowy z predefiniowanych słowników.

Dalszy postęp w tej dziedzinie będzie wymagać ścisłej współpracy informatyków z językoznawcami, zwłaszcza że problem generacji mowy nabiera coraz większego znaczenia w bardzo szybko rozwijającej się dziedzinie, jaką są sieci komputerowe.

LITERATURA

1. Wideł S., Stera P., Werner A.: Założenia wstępne do projektu komputerowego systemu telefonicznego w projekcie *Pszów*, Analiza literaturowa, Gliwice, 22.01.1997.
2. Wideł S., Stera P., Piasecki J.: Interakcyjne Systemy Telefoniczne, Tom VI, Gliwice, marzec 1997.

3. Wideł S., Stera P.: Zdalny dostęp do danych za pomocą interakcyjnych systemów telefonicznych w jednostkach administracji publicznej na przykładzie Urzędu Miasta Pszów, V Seminarium Sieci Komputerowe,Informatyka,zeszyt 34,Gliwice 1998.
4. Grzywak A., Wideł S.: Multimedialny rozproszony system zarządzania Urzędem Miasta Pszów, Komputerowe Systemy Wielodostępne KSW'98, ISBN 83-908599-4-7 Bydgoszcz 1998.
5. Dialogic, "Voice Hardware Reference" Dialogic Corporation 050147-002 1994.
6. Dialogic, "Voice Software Reference for MS-DOS" Volume 1, 2 Dialogic Corporation 05-004-006 1992-1993.
7. Dialogic, "Voice Programmer's Guide for MS-DOS" Dialogic Corporation 05-004-006 1994.
8. Dialogic, "Voice Features Guide for MS-DOS" Dialogic Corporation 05-004-006 1994.
9. Rhetorex, "RealCT User Guide for Windows NT and Windows 95", rev 1, 00 00-00-3528, October 1996.
10. Rhetorex, "Programming Guide for DOS", rev 5.01, 00-00-3360, March 1995.
11. Rhetorex, "Function Reference Manual for DOS", rev 5.01, 00-00-3060, March 1995.
12. Rhetorex, "API Reference for Windows 95 and Windows NT", rev 1.01, 00-00-3037, March 1996.
13. Edgar B.: "PC-Based Voice Processing" Second Edition Flatiron Publishing Inc..ISBN 0-936648-45-7 1994r.
14. Szlam A., Thatcher K., "Predictive Dialing" Flatiron Publishing IncISBN 0-936648-80-5 1996r.
15. Szober S.: Gramatyka języka polskiego, wyd. VI, Warszawa 1968.
16. Klemensiewicz Z.: Zarys składni polskiej, PWN, Warszawa 1961.
17. Klemensiewicz Z., Lehr-Spławiński T., Urbańczyk S.: Gramatyka historyczna języka polskiego, PWN, Warszawa 1964.
18. Bąk P.: Gramatyka języka polskiego, wyd. X Wiedza Powszechna, 1997.

Recenzja: Prof. dr inż. Stefan Węgrzyn

Wpłynęło do Redakcji 4 maja 1999 r.

Abstract

Interactive Voice Response System (IVR) is a computer application that provides the user interface for Data Base access over a phone. The Interactive Voice Response Systems will

become more popularity not only into a PBX environment. In figure 1 you can find the conventional IVR Architecture. The applications like Automatic Attendants and Voice Mail can be easily build in the next generations of the PBX's (as a standard or an option)

Voice Response System can be the parts of the more general service available of the future Data and Voice integrated network. Figure 2 is an example that voice response service can be available to phone and to the network client.

For developing IVR system it is a very important to understand the language grammar role. In this paper we present classification of the numbers to follow the Polish grammar rules. From this point of view the Polish grimmer are extremely difficult. Cardinal number one (1) has 20, number two (2) 7, number three and up (3) 7 different case endings.

This we develop the classification and algorithms set that can be used in IVR system for translation well know filed format of the Data Base filed to the Interactive Voice Response.