

Stanisław WIDEL, Piotr STERA
Politechnika Śląska, Instytut Informatyki

ZDALNY DOSTĘP DO DANYCH ZA POMOCĄ INTERAKCYJNYCH SYSTEMÓW TELEFONICZNYCH W JEDNOSTKACH ADMINISTRACJI PUBLICZNEJ NA PRZYKŁADZIE URZĘDU MIASTA PSZÓW

Streszczenie. Standardowy aparat telefoniczny może być zastosowany jako urządzenie do interakcji z komputerem. Interakcyjny System Telefoniczny (IST) jest komputerową aplikacją tworzącą interfejs użytkownika dostępny przez telefon. W tym artykule przedstawiono propozycję wprowadzenia IST opartego na koncepcji maszyny stanów do Urzędu Miasta. Projekt ten jest częścią Projektu Celowego nr 8T11C 009 96C/2995 KBN realizowanego w Urzędzie Miasta Pszów.

INTERACTIVE VOICE RESPONSE APPLICATION DEVELOPMENT PROGRAM IN PSZÓW CITY COUNSEL

Summary. Standard telephone can be used as a device for interacting with a computer. Interactive Voice Response (IVR) is a computer application which use standard user interface over a telephone. In this paper was presented approach to introduce IVR system to the City Counsel base on the State Machine Engine. Our development is a part of government program no 8T11C 009 96C/2995 in Pszów City Counsel.

1. Wprowadzenie

Wykorzystanie komputerów do celów multimedialnych od strony użytkownika niewiele się różni od zastosowania do tego celu innych mediów, takich jak kino, telewizja czy wideo i nie daje możliwości interakcji. System, w którym użytkownik biernie obserwuje pokazującą się

na ekranie informację, nawet jeżeli jest skojarzony z ruchomym obrazem i dźwiękiem, nie wykorzystuje w pełni potencjału komputera.

Jednym z wyzwań dla współczesnej technologii informatycznej jest wprowadzenie w języku naturalnym dialogu pomiędzy użytkownikiem a komputerem. Taki sposób porozumiewania się z maszyną jest dla człowieka intuicyjny. Natomiast obecnie nie istnieje powszechnie przyjęty standard rozwiązań sprzętowo-programowych w tym zakresie.

Jeżeli założyć, że z systemem komputerowym będziemy mogli się komunikować w języku naturalnym, to rodzi się wątpliwość, czy system taki musi również być wyposażony w typowe urządzenia wejścia-wyjścia, takie jak monitor, klawiatura, mysz. Należy się spodziewać dla takich systemów zmiany typowej konfiguracji urządzeń wejścia-wyjścia polegającej na uproszczeniu ukierunkowanym w stronę wykorzystania mowy. Oczywiście, zawsze istnieć będą zastosowania, w których inny sposób dialogu będzie efektywniejszy niż mowa (np. administrowanie serwerem sieciowym).

Powszechnie wiadomo, że jednym z podstawowych celów budowy sieci komputerowej jest udostępnienie i współużywanie zasobów tej sieci. Kolejnym zagadnieniem, jakie należy rozważyć wprowadzając nowy interfejs, jest skojarzone z nim rozwiązanie komunikacyjne. Jeżeli założymy, że w wybranych stanowiskach aplikacje będą wykorzystywać język naturalny, to należy również przewidzieć, w jaki sposób aplikacje te mogą być współużytkowane.

Powyższe rozważania stanowią wprowadzenie do przedstawienia założeń, jakie zostały przyjęte dla realizacji projektu systemu obiegu informacji multimedialnej dla Urzędu Miasta Pszów.

Przeprowadzając analizę możliwości zastosowania multimedialnych systemów w jednostkach administracji publicznej, wzięliśmy pod uwagę następujące czynniki:

- koszty (wyposażenie, konserwacja),
- dostęp do stanowiska,
- instalację,
- łatwość obsługi,
- możliwość zdalnego dostępu,
- kontrolę dostępu do danych.

W rezultacie zaproponowaliśmy, jako dodatkowy interfejs użytkownika do komunikacji z systemem komputerowym, standardowy aparat telefoniczny.

2. Interakcyjne systemy telefoniczne

Telefon został skonstruowany w celu komunikowania się i przekazywania informacji na odległość za pomocą głosu, wykorzystując pasmo akustyczne o szerokości ok. 4 kHz. Jego

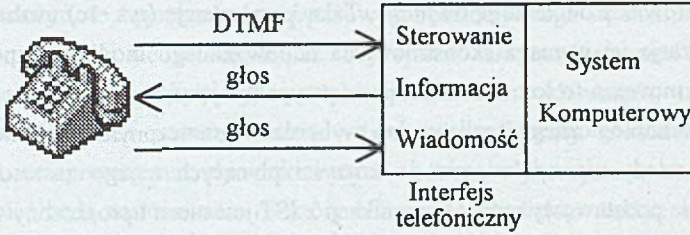
przystosowanie do przekazywania mowy ludzkiej sprawia, że można go wykorzystać jako interfejs głosowy do komputera. Koszty związane z instalacją i konserwacją są wtedy znacznie niższe niż dla stanowiska komputerowego. Takie rozwiązanie nazywane będzie dalej **interakcyjnym systemem telefonicznym (IST)**.

Podstawowym sposobem przekazywania informacji przez telefon jest głos zawierający informację słowną. W taki więc sposób komputer IST powinien przekazywać informację człowiekowi, jak również odbierać ją od niego. Taką komunikację (rys. 1c) można uznać za wzorcową. Realizacja jej wymaga skonstruowania odpowiedniego modułu rozpoznającego i interpretującego mowę, a to jest - szczególnie w przypadku języka polskiego - na obecnym etapie rozwoju technologicznego realizowalne w bardzo ograniczonym zakresie, zaś cena systemu wzrasta wtedy niewspółmiernie do korzyści płynących z jego zastosowania. To sprawia, że obecnie podstawowy model komunikacji z IST jest nieco uproszczony w stosunku do wzorcowego (rys. 1a). Jako dodatkowy sposób przekazywania swoich preferencji użytkownik wykorzystuje wybieranie numerów telefonicznych. Może się ono odbywać w sposób tonowy - za pomocą kodów DTMF (ang. Dual Tone Modulation Frequency - podwójna tonowa modulacja częstotliwości) lub pulsowy (podczas zrealizowanego połączenia odpowiadają im dźwiękowo trzaski). Od strony technicznej zastosowanie pierwszego wariantu do komunikacji z IST jest o wiele prostsze, gdyż kody DTMF jako dźwięki są dobrze zdefiniowane, zaś zastosowanie pulsów wymaga odpowiedniego modułu sprzętowego lub programowego rozpoznającego je i odróżniającego od zakłóceń występujących na linii telefonicznej.

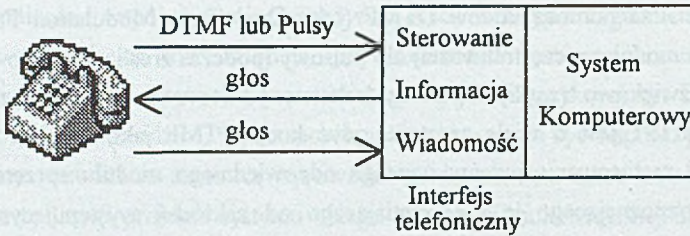
Tonowy system wybierania numerów jest używany w USA, jednak wszystkie nowsze telefony i centrale telefoniczne pozwalają na jego stosowanie również w Europie. Jeżeli centrala telefoniczna obsługuje tylko wybieranie impulsowe, wtedy po nawiązaniu połączenia użytkownik musi przełączyć swój telefon na wybieranie tonowe (przełącznik zainstalowany w telefonie). Zamiast przełączania sposobu wybierania numerów w telefonie można posłużyć się również odpowiednim akustycznym generatorem tonów DTMF, który generuje odpowiednie tony odpowiadające poszczególnym cyfrom, a którego używa się przykładając go do mikrofonu słuchawki. W przypadku gdy użytkownik korzysta z centrali telefonicznej z tonowym wybieraniem numerów, nie musi wykonywać żadnych dodatkowych czynności. Aby umożliwić korzystanie z systemu także osobom nie posiadającym wybierania tonowego w swoim telefonie i nie wymagać od nich również posiadania generatora tonów DTMF, można zainstalować w systemie odpowiednią kartę przekształcającą impulsy na tony, co jednak podnosi kilkakrotnie koszty całego systemu. To ostatnie rozwiązanie nie zawsze będzie poprawnie działać, gdyż niektóre centrale telefoniczne są wyposażone w moduły filtrujące zakłócenia, a impulsowe wybieranie numerów podczas zrealizowanego połączenia telefonicznego jest również w ten sposób traktowane. W związku z tym sygnały takie ulegają

w tym przypadku poważnym zniekształceniom uniemożliwiającym ich prawidłowe rozpoznanie.

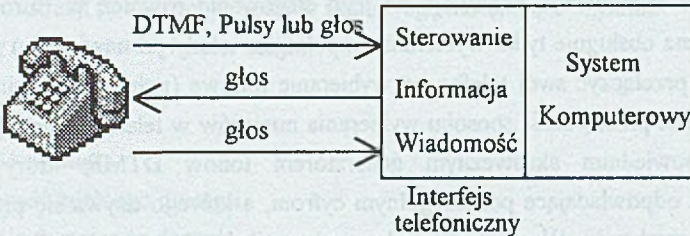
Reakcje ze strony użytkownika podzielono na: sterujące - będące odpowiednikiem zapytania, czyli sprecyzowania, czego chce się dowiedzieć, oraz oznaczone jako wiadomości. Przez tę drugą rozumie się informację głosową, która zostanie zarejestrowana przez komputer w formie pliku dźwiękowego bez jego interpretacji przez system.



a) Podstawowa wersja Interakcyjnego Systemu Telefonicznego



b) Interakcyjny System Telefoniczny z rozpoznawaniem impulsów



c) Interakcyjny System Telefoniczny z rozpoznawaniem mowy

Rys. 1. Interfejs użytkownika do Informacyjnego Systemu Telefonicznego

Fig. 1. User interface to the Interactive Voice Response System

Na rys. 1a przedstawiono podstawowy sposób komunikacji użytkownika z IST: użytkownik przekazuje sterowanie systemowi za pomocą kodów DTMF, zaś odpowiedzi otrzymuje na sposób głosowy i ma możliwość przekazywania wiadomości również głosem.

Problem generacji komunikatów przez system może być rozwiązany dwoma sposobami: informacje mogą być zarejestrowane w postaci danych dźwiękowych lub generowane przez

syntezator mowy na podstawie plików tekstowych (technika text to speech). Podstawowa wersja systemu może zostać rozbudowana o moduł rozpoznawania impulsów pozwalający korzystać z systemu za pomocą aparatów telefonicznych posiadających tylko taki sposób wybierania numerów (rys. 1b). W najbardziej rozwiniętej wersji system może posiadać moduł rozpoznawania mowy (rys. 1c), co umożliwi przekazywanie sterowania systemowi za pomocą języka naturalnego. W tym przypadku jest on zbliżony do wzorcowego modelu.

3. Opis systemu

3.1. Funkcje IST

Do realizacji systemu IST dla miasta Pszów przyjęto następujące założenia:

- system jest autonomiczny w stosunku do istniejącej lokalnej sieci komputerowej,
- podstawowym użytkownikiem systemu jest osoba komunikująca się za pomocą aparatu telefonicznego.

W pierwszym etapie zrealizowano bazę danych z odpowiednimi głosowymi komunikatami informującymi o najczęściej stosowanych i typowych procedurach załatwiania formalności w Urzędzie Miasta. Przykładem może być uzyskanie informacji na temat procedury wymiany dowodu osobistego, zameldowania, pozwolenia na prowadzenie działalności gospodarczej, itp. Informacja ta może być dostępna zarówno dla klientów znajdujących się w Urzędzie Miasta poprzez wydzielone aparaty wewnętrznej centrali Urzędu, jak i z dowolnego aparatu miejskiej sieci telefonicznej.

Kolejnym etapem rozwoju systemu jest uruchomienie poczty głosowej w celu zorganizowania wymiany informacji pomiędzy pracownikami Urzędu Miasta. System ten może również służyć do pozostawienia wiadomości głosowych dla odpowiednich wydziałów Urzędu. Poczta głosowa będzie zrealizowana w oparciu o strukturę skrzynek dla wszystkich pracowników oraz dla odpowiednich wydziałów.

Przez skrzynkę pocztową będziemy rozumieli miejsce w systemie komputerowym oznaczone odpowiednim kodem z możliwością wybrania tego kodu przez telefon. Po zadzwonieniu i wybraniu kodu skrzynki będzie można zostawić w niej wiadomość. Z kolei uprawniona osoba posiadająca specjalne hasło dostępu będzie mogła odsłuchać wszystkie pozostawione komunikaty znajdujące się w danej skrzynce. Z punktu widzenia zorganizowania komunikacji z klientami zewnętrznymi numery skrzynek poszczególnych pracowników nie muszą być opublikowane. Do wiadomości publicznej mogą zostać jedynie podane np. numery skrzynek odpowiadające poszczególnym wydziałom.

3.2. Rozwiązania sprzętowe

Wymagania sprzętowe systemu są w dużej mierze uzależnione od liczby linii telefonicznych, do których ma być on dołączony. W zrealizowanym projekcie liczbę tę ustalono na cztery i dla takiej wartości określono minimalne wymagania. Komputer z procesorem 486 lub wyższym. Minimalną ilość pamięci RAM oszacowano na 8 MB. Wielkość zapotrzebowania na pamięć dyskową jest ściśle uzależniona od ilości informacji, którą ma udostępniać system, gdyż samo oprogramowanie stanowi tylko niewielką część tej wartości. Jeśli system ma być zintegrowany z komputerową siecią lokalną, to konieczna jest odpowiednia karta sieciowa. Do działania samego IST wyposażono komputer w kartę umożliwiającą podłączenie go do sieci telefonicznej. W zrealizowanym projekcie zastosowano w tym celu czterokanałową kartę D41 firmy Dialogic.

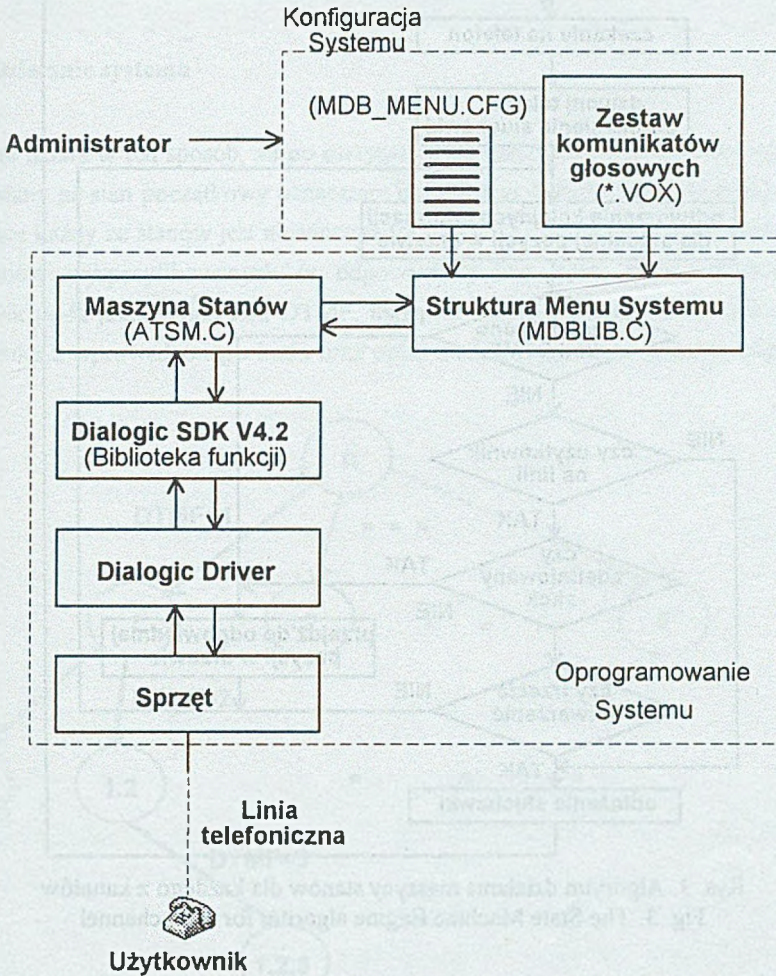
3.3. Oprogramowanie

Zgodnie z założeniami przedstawionymi wcześniej oprogramowanie sterujące pracą systemu musi umożliwiać równoczesną obsługę wielu użytkowników. Ich praca z IST jest asynchroniczna. Drugą ważną cechą zaprojektowanego oprogramowania jest jego skalowalność na dowolną liczbę kanałów. Ograniczenie maksymalnej ich liczby wynika jedynie z możliwości zainstalowanego sprzętu.

Na rys. 2 przedstawiono strukturę systemu. Oprogramowanie stanowi jeden moduł programowy, jednak z punktu widzenia konstrukcji są w nim wydzielone dwie części. Pierwsza, stanowiąca jądro wielowątkowej aplikacji, dalej będzie zwana maszyną stanów systemu informacyjnego (Audio-Text State Machine). Jest ona odpowiedzialna za zapewnienie równoległej obsługi wielu użytkowników. Wyodrębniono w komunikacji użytkownika z systemem kolejne jego fazy i sytuacje, jakie mogą wystąpić, oraz zależności pomiędzy nimi. W ten sposób powstał graf przejść maszyny stanów. Pod wpływem zdarzeń występujących w IST następują przejścia pomiędzy stanami. Aby uwzględnić działania wszystkich użytkowników aktualnie korzystających z systemu, zdarzenia związane z ich obsługą są kolejgowane. Schematyczny algorytm działania maszyny stanów dla każdego z kanałów przedstawia rys. 3.

Druga część oprogramowania realizuje podstawowe operacje na komunikatach głosowych (Message Database Library), takich jak odtwarzanie czy nagrywanie odpowiednich komunikatów. Odpowiada ona również za utworzenie i obsługę całej struktury informacyjnej systemu. Ta część oprogramowania korzysta z biblioteki dostarczanej przez firmę Dialogic (Dialogic SDK V4.2). Program został tak skonstruowany, że jego działanie jest

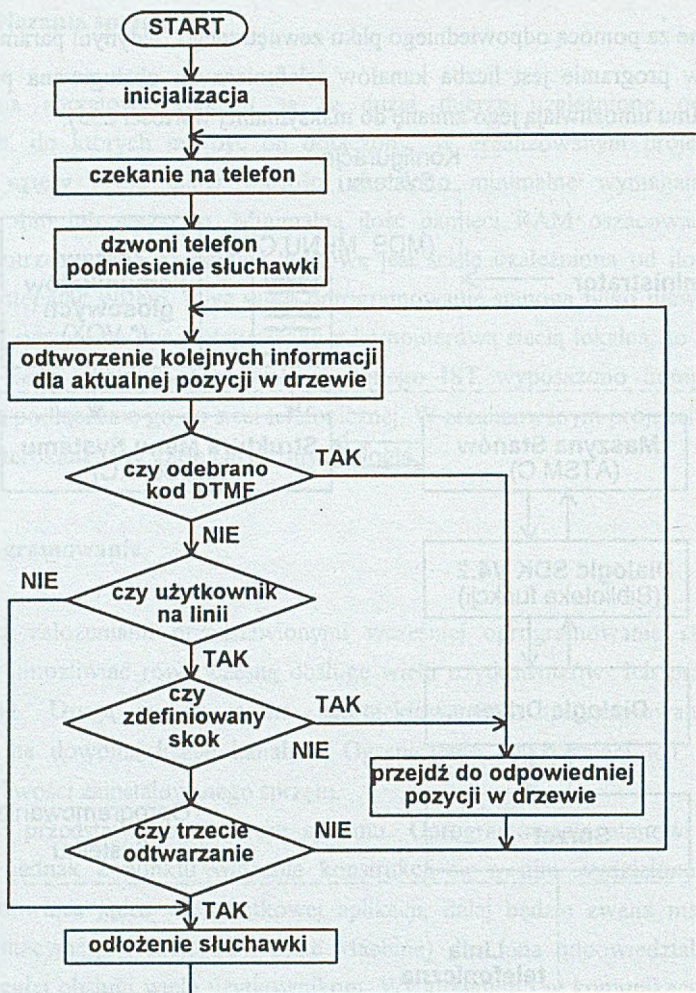
konfigurowalne za pomocą odpowiedniego pliku zewnętrznego. Jedynym parametrem na stałe ustawionym w programie jest liczba kanałów telefonicznych obsługiwana przez program (źródła programu umożliwiają jego zmianę do maksymalnej wartości 255).



Rys. 2. Struktura Telefonicznego Systemu Informacyjnego
Fig. 2. Interactive Voice Response System Architecture

3.4. Konfiguracja systemu

System w zakresie informacji, którą zawiera, oraz struktury jej rozmieszczenia może być dowolnie konfigurowany przez administratora bez konieczności rekompilacji programu.



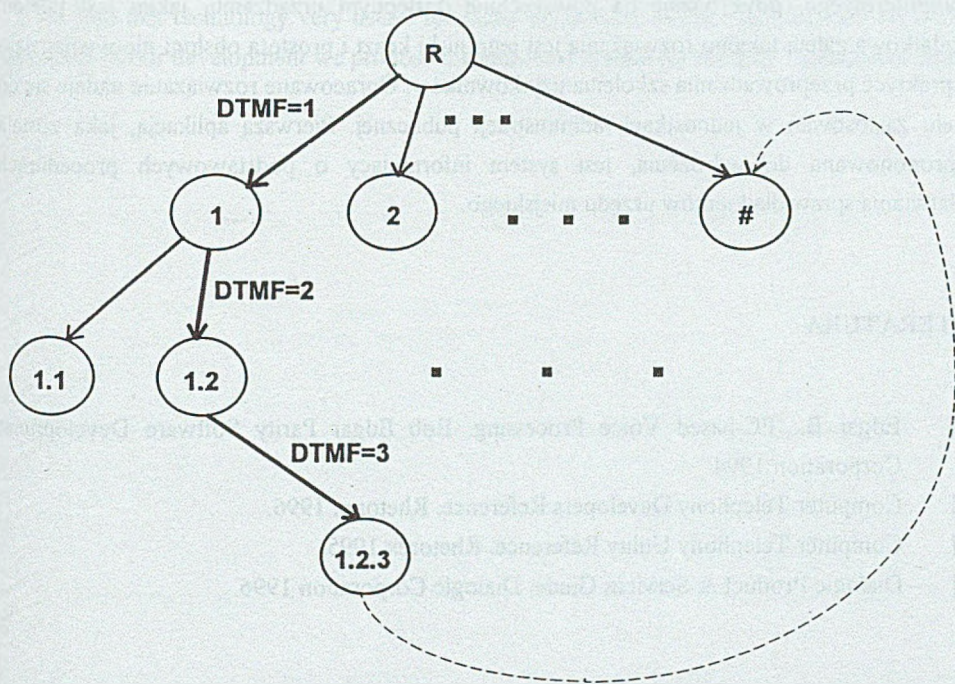
Rys. 3. Algorytm działania maszyny stanów dla każdego z kanałów
Fig. 3. The State Machine Engine algorithm for each channel

W tym celu zaprojektowano plik, w którym można dokonywać odpowiednich zapisów. Założono, że podstawową strukturą, w której są rozmieszczone informacje, będzie drzewo (rys.4). Od struktury tej można jednak wprowadzić odstępstwa. Konfiguracja systemu polega w głównej mierze na pogrupowaniu i odpowiednim hierarchicznym podzieleniu informacji mającej być udostępnianą przez IST, tak by utworzyła ona strukturę drzewiastą. Tak utworzone zależności należy następnie zakodować w pliku konfiguracyjnym posługując się prostym, w tym celu zaimplementowanym językiem opisu IST, samą zaś informację nagrać w postaci plików dźwiękowych (system stosuje format plików *.vox). Ze względu na potrzebę przejść z jednych stanów do innych w sposób niezgodny z tym, jaki wynika ze struktury

drzewiastej, wprowadzono również taką możliwość. Tak więc konfiguracja i rozbudowa systemu informacyjnego polegają na nagrywaniu odpowiednich komunikatów i zmianie zapisów w pliku.

3.5. Działanie systemu

System działa w ten sposób, że po otrzymaniu zgłoszenia od użytkownika jego kontekst jest ustawiany na stan początkowy oznaczony literą R (rys.4; Root - ang. korzeń). Procedura obsługująca każdy ze stanów jest niezmienna (por. rys. 3) i polega na odtwarzaniu kolejnych komunikatów wyspecyfikowanych w odpowiedniej linii pliku konfiguracyjnego. Jeżeli użytkownik poda jakiegokolwiek kod DTMF, maszyna stanów przechodzi do kolejnego stanu, który wynika z wprowadzonego kodu oraz opisu znajdującego się w pliku konfiguracyjnym.



Rys. 4. Przykładowy graf przejść IST z zaznaczoną przykładową ścieżką przejścia
Fig. 4. Example of the IVR state flow with selected transition between states

Dla przykładu (rys. 4), jeżeli użytkownik po zgłoszeniu się systemowi naciśnie klawisz '1', przechodzi do stanu oznaczonego przez '1' i usłyszy informacje, które zostały przypisane temu stanowi. Jeżeli teraz naciśnie klawisz '2', przejdzie do stanu oznaczonego '1.2' i usłyszy

komunikaty określone dla tego węzła drzewa. Z kolei naciśnięcie klawisza '3' spowoduje przejście do stanu '1.2.3'. Jeden z klawiszy ('9') został przewidziany dla funkcji cofania się do stanu bezpośrednio poprzedzającego aktualny. Przewidziano również możliwość automatycznego (bez ingerencji użytkownika) przejścia z jednego stanu do dowolnego innego (na rysunku zaznaczono to linią przerywaną). Przejście takie definiowane jest również w pliku konfiguracyjnym.

4. Podsumowanie

Rozwój technik multimedialnych powoduje również większe zainteresowanie komunikacją z komputerem za pomocą głosu. Zaproponowane w artykule rozwiązanie tego problemu wydaje się być dla wielu zastosowań właściwsze od użycia w tym celu rozbudowanego sprzętu komputerowego, gdyż bazuje na powszechnie dostępnym urządzeniu, jakim jest telefon. Dodatkową zaletą takiego rozwiązania jest jego niski koszt i prostota obsługi nie wymagająca w praktyce przeprowadzania szkolenia użytkowników. Opracowane rozwiązanie nadaje się do wielu zastosowań w jednostkach administracji publicznej. Pierwszą aplikacją, jaka została zaproponowana do wdrożenia, jest system informujący o podstawowych procedurach załatwiania spraw dla kientów urzędu miejskiego.

LITERATURA

- [1] Edgar B.: PC-based Voice Processing. Bob Edgar Parity Software Development Corporation 1994.
- [2] Computer Telephony Developers Reference. Rhetorex 1996.
- [3] Computer Telephony Utility Reference. Rhetorex 1995.
- [4] Dialogic Product & Services Guide. Dialogic Corporation 1996.

Recenzent: Dr inż. Ryszard Winiarczyk

Wpłynęło do Redakcji 7 stycznia 1998 r.

Abstract

Voice processing application has become an important computer development areas.

Voice processing is a combination of telecommunication and data processing allowing callers to use the telephone as an input-output peripheral device for interacting with a computer. This standard telephone interface can also be developed by adding the functional modules to the voice processing application (Fig. 1).

Interactive Voice Response (IVR) is a voice computer which uses remote touch-tone telephones. IVR uses a digitized synthesized voice to the distant caller. In this paper we present our approach to introduce IVR system to the City Counsel. Our development is a part of government program no 8T11C 009 96C/2995 in Pszów City Counsel.

In this paper we present our system design of IVR System Architecture (Fig. 2) base on the State Machine Engine (Fig. 3).

We find this technology very useful for many application in the City Counsel. As a very first level of our development we propose the audiotext system for the City Counsel customer.