

elektroniczna
technika
obliczeniowa

P. 3054/74

NOWOŚCI
NR 9
1979

ZJEDNOCZENIE
PRZEMYSŁU
AUTOMATYKI
I APARATURY
POMIAROWEJ „MERA”

●
INSTYTUT MASZYN
MATEMATYCZNYCH
BRANŻOWY
OŚRODEK INTE



Z przykrością zawiadamiamy, że z przyczyn od nas niezależnych - zmuszeni jesteśmy zakończyć na niniejszym numerze rocznik XIII/1974 Elektronicznej Techniki Obliczeniowej NOWOŚCI. Również w 1975 r. ukazać się tylko 4 numery ETO Nowości.

Komitet Redakcyjny

Wszystkie wyliczenia, które zostały
dokonane w tym celu, zostały
dokonane w oparciu o dane
z lat 1975-1977. W tym celu
zostały wykorzystane dane
z lat 1975-1977.

Wszystkie wyliczenia

ELEKTRONICZNA TECHNIKA OBLICZENIOWA
NOWOŚCI
DWUMIESIĘCZNIK

Rok XIII

Nr 4

1974

Komunikacja człowiek-maszyna

S p i s t r e ś c i

	str.
OD REDAKCJI	3
mgr Maria ŁĄCKA, mgr Ewa ZABOROWSKA: Wprowadzenie do systemów konwersacyjnych	7
dr inż. Teresa BUCZKOWSKA: Stan obecny i kierunki rozwojowe gra- ficznych monitorów ekranowych	59
mgr inż. Ryszard PATRYN: Akustyczne urządzenia wyjściowe	71
mgr inż. Ryszard PATRYN: Eksperymentalny syntezytor fonematyczny w IMM	81
mgr inż. Eugeniusz SAWICKI: Automatyczne rozpoznawanie mowy . .	83
mgr inż. Zbigniew KĘDZIOR: Urządzenia rozpoznające znaki alfanu- meryczne zapisane na nośniku papierowym i filmowym	95
mgr inż. Andrzej KOJEMSKI, mgr inż. Jerzy KASPRZYK: Wykorzystanie małych maszyn cyfrowych w sieciach teleinformatycznych . . .	111

Wydaje

INSTYTUT MASZYN MATEMATYCZNYCH

B r a n ż o w y O ś r o d e k I n f o r m a c j i
N a u k o w e j T e c h n i c z n e j i E k o n o m i c z n e j

KOMITET REDAKCYJNY

Jerzy Dańda (red. nacz.), Hanna Drozdowska (sekr. red.),

Antoni Kwiatkowski, Ryszard Patryn

Dorota Prawdzic (zast. red. nacz.), Zbigniew Świątkowski

Redaktor techniczny - Maria Kozłowska

Adres redakcji: 02-078 Warszawa, ul. Krzywickiego 34
tel. 28-37-29 lub 21-84-41 w. 431

OD REDAKCJI

Prezentowany zeszyt przedstawia osiem¹ artykułów związanych wspólnie tematyką dotyczącą najogólniej problemu komunikacji człowiek-maszyna.

Problem komunikacji człowiek-maszyna w odniesieniu do komputerów istniał i narastał wraz z rozwojem komputerów. Początkowo nie był on jednak czynnikiem w istotny sposób wpływającym na rozwój komputerów. Postęp w dziedzinie komputerów był związany głównie ze zwiększaniem efektywności przetwarzania wewnątrz komputera, łącznie ze sprawą szybkiego wprowadzania danych w sposób jak najmniej zakłócający przetwarzanie danych. Typowy sposób porozumienia się człowieka z maszyną polegał na wykorzystywaniu jako pośredniego ogniwa nośników informacji w postaci kart lub taśm perforowanych.

Dopiero wprowadzenie na początku lat sześćdziesiątych pracy z podziałem czasu, umożliwiające bezpośredni dostęp wielu użytkownikom do dużego komputera przy efektywnym jego wykorzystywaniu, oraz rozwój minikomputerów, będących narzędziami pracy użytkownika w bezpośrednim zasięgu jego ręki, spowodowały intensywne zainteresowanie sprawami łączności człowieka z maszyną. Na szerszą skalę zaczęto stosować konwersacyjne metody współpracy. W ten sposób stało się możliwe udostępnienie człowiekowi do bezpośredniej pracy (współpracy) na bieżąco komputera - jako jednego z najbardziej rozwiniętych współczesnych narzędzi, a równocześnie umożliwiło wystąpienie człowieka w roli jednego ze składników systemu komputerowego, w istotny sposób wpływającego na efektywność działania tego systemu. Ta nowa jakość przejawiała się m.in. w rozwoju wielu nowych dziedzin wykorzystania komputerów, noszących nazwę działań wspomaganych przez komputer (np. projektowanie, nauczanie i in.).

Doskonalenie kontaktu człowiek-maszyna związane jest przede wszystkim z ułatwieniem sposobu przekazywania informacji przez człowieka do komputera oraz otrzymywania tej informacji z komputera. Stąd dążenie do wprowadzania i wyprowadzania informacji w sposób bardziej przystosowany do człowieka i w pełnym zakresie jego typowych możliwości i upodobań. Można wyróżnić następujące formy informacji, którymi posługuje się człowiek, zarówno w życiu codziennym, jak również coraz szerzej przy wykorzystywaniu komputerów:

¹ Artykułu mgr inż. Zbigniewa Kędziora pt. "Urządzenia odczytujące informację graficzną zapisaną na nośnikach naturalnych" ze względów technicznych nie udało się zamieścić w bieżącym numerze ETO NOWOŚCI. Będzie on wydrukowany w następnym numerze.

- o informacje w postaci symboli (znakach alfanumerycznych),
- o informacja graficzna (rysunki, zdjęcia),
- o informacja foniczna.

Niemalą rolę w ułatwianiu współpracy człowieka z komputerami odgrywa też przestrzenne rozmieszczenie znaków, która przez monokomputerowe systemy wielokodowe praktycznie z podziałem czasu prowadzi w toku realizacji zadań w złożonych sieci teleinformatycznych. Rozszerza to zakres wykorzystania komputerów, a jednocześnie wyznacza pewne nowe formy współpracy człowieka z maszyną, uwidaczniające się np. w systemach zdalnej współpracy z komputerami za pośrednictwem przyrządnie dostępnego aparatu telefonicznego.

Wyżej omawiane sprawy są tematem prac grupy pracowników IMM.

Zeszyt otwiera "Wprowadzenie do systemów konwersacyjnych" Marii Łackiej i Ewy Zaborowskiej, omawiające ogólne zagadnienia systemów konwersacyjnych i ich zastosowań, ze szczególnym zwróceniem uwagi na sprawy rozwoju języków programowania służących do tych celów.

Teresa Buczkowska w pracy "Stan obecny i kierunki rozwojowe graficznych monitorów ekranowych" przedstawia rozwiązania funkcjonalne i techniczne oraz perspektywy dalszego rozwoju tych podstawowych urządzeń wejścia i wyjścia graficznego w systemach konwersacyjnych.

Następne trzy prace związane są z zagadnieniami komunikacji fonicznej z komputerem. Ryszard Patryn w artykule "Akustyczne urządzenie wyjściowe" daje przegląd metod rozwiązań urządzeń pozwalających na uzyskiwanie wypowiedzi słownych w urządzeniach wyjściowych komputera. W krótkim opracowaniu o charakterze komunikatu, pt. "Eksperymentalny syntezytor fonematyczny w IMM", ten sam autor prezentuje wyniki własnej pracy nad zbudowaniem laboratoryjnego syntezytora fonematycznego. W pracy Eugeniusza Sawickiego "Automatyczne rozpoznawanie mowy" przedstawione są metody analizy sygnałów mowy, z punktu widzenia możliwości wykorzystania ich do sterowania pracą komputera za pomocą głosu.

W kolejnych dwóch pracach Zbigniew Kędzior zajmuje się urządzeniami do wprowadzania do komputera informacji zawartej w dokumentach pisanych i rysowanych. Artykuł "Urządzenia rozpoznające znaki alfanumeryczne zapisane na nośniku papierowym i filmowym" omawia optyczne czytniki znaków, zwłaszcza zagadnienia dotyczące ich stosowania i rozwiązań technicznych. W drugim artykule "Urządzenia odczytujące informacje graficzne zapisane na nośnikach naturalnych" przedstawiono zastosowanie i rozwią-

zanie urządzeń najczęściej zwanych koordynatometrami, służących do wprowadzania do komputera wartości współrzędnych dla rysunków.

Praca Andrzeja Kojemskiego i Jerzego Kasprzyka "Wykorzystanie małych maszyn cyfrowych w sieciach teleinformatycznych" wprowadza w zagadnienia związane z budową sieci teleinformatycznych, szczegółowiej omawiając sposoby wykorzystania w tych sieciach minikomputerów.

Przedstawione prace nie wyczerpują oczywiście tak rozległego problemu jakim jest komunikacja człowiek-maszyna, być może jednak częściowo uzupełniają stosunkowo ubogie piśmiennictwo dotyczące tej tematyki w języku polskim¹.

Na zakończenie pragnę podziękować autorom za trud włożony w napisanie i współpracę przy przygotowaniu do wydania prac, a mgr Jerzemu Dańdzie, Redaktorowi Naczelnemu ETO NOWOŚCI, za inicjatywę i zachętę do przygotowania tego zeszytu tematycznego.

Andrzej Kojemski

¹ Czytelników zainteresowanych szerzej tą tematyką zachęcam do zwrócenia uwagi na książkę J. Martina, Design of Man-Computer Dialogues, Prentice-Hall, Inc. 1937, której tłumaczenie na język polski jest obecnie w przygotowaniu do wydania przez WNT.

mgr Maria ŁACKA

007.51:681.322.004.14

mgr Ewa ZABOROWSKA

800.92

Instytut Maszyn Matematycznych

WPROWADZENIE DO SYSTEMÓW KONWERSACYJNYCH

1. Wstęp

Terminem "komunikacja w trybie konwersacyjnym" określamy bezpośrednią współpracę człowieka z maszyną cyfrową, współpracę opartą na zasadzie pytań i odpowiedzi. Systemy przetwarzania, w których stosuje się głównie taką metodę komunikacji nazywamy systemami konwersacyjnymi.

W ostatnim dziesięcioleciu w systemach przetwarzania nastąpił wyraźny zwrot w kierunku stosowania współpracy konwersacyjnej.

Niektórzy z takim podejściem do wykorzystania maszyn wiążą nadzieję wyjścia z "kryzysu softwarowego", do którego doprowadziło tworzenie potężnie rozbudowanych i przez to trudnych do opanowania dla przeciętnego programisty i operatora - systemów oprogramowania. Dzięki łatwej komunikacji z maszyną w systemach konwersacyjnych (realizowanej np. przez prowadzenie dialogu w języku naturalnym, instruowanie użytkownika przez maszynę itp.) stają się one dostępne również dla ludzi nie związanych profesjonalnie z komputerami.

Powstają więc, w kręgu coraz to nowych zastosowań, coraz funkcjonalniejsze systemy konwersacyjne, powstaje nowa baza sprzętowa dla tych systemów i narzędzia programowe do ich realizacji. Przegląd nawet tylko ważniejszych systemów konwersacyjnych z punktu widzenia ich kilkunastoletniej historii przekraczałby znacznie ramy tej pracy.

Celem jej jest tylko wskazanie, na podstawie najbardziej reprezentatywnych przykładów, ogólnej idei działania takich systemów oraz rozpatrzenie podstawowych problemów pojawiających się przy ich realizacji i sposobów rozwiązywania tych problemów.

W pracy omówiono zagadnienie systemów konwersacyjnych na elementarnym poziomie. Chcemy, żeby była ona zrozumiała również dla ludzi nie zajmujących się profesjonalnie tym zagadnieniem i aby stanowiła wstęp do kolejnych opracowań, bardziej szczegółowych, ale dotyczących już węższego zakresu problemów z omawianej dziedziny, np. systemów konwersacyjnych o konkretnych zastosowaniach. W związku z takim ujęciem tematu niektóre cytowane przykłady są celowo uproszczone.

Pewne elementy systemu konwersacyjnego nie zależą od tego, czy stosuje się go np. do nauczania przy pomocy komputera, czy w medycynie, czy w projektowaniu wspomaganym przez maszynę itd. Przy każdym typie zastosowań pojawiają się jednak specyficzne problemy, których rozwiązanie wymaga indywidualnych metod. W pkt 2 zostały przeanalizowane zagadnienia ogólne, niezależne od zastosowań, jak dobranie odpowiedniej bazy sprzętowo-programowej koniecznej do realizacji systemu konwersacyjnego i określenie języka, w którym człowiek będzie prowadził dialog z maszyną. To ostatnie zagadnienie wiąże się ściśle z problemem stosowania języka naturalnego jako języka dialogu. Problemy związane z poszczególnymi zastosowaniami systemów konwersacyjnych omówione są w pkt 3.

2. Ogólna charakterystyka systemów konwersacyjnych

2.1. Właściwości funkcjonalne

W tradycyjnych metodach użytkowania maszyn człowiek musiał stosować środki przekazu zrozumiałe dla maszyny (przyciski, klawisze, kod maszyny, języki programowania itp.) a nienaturalne w komunikacji między ludźmi. Ideą przewodnią współpracy konwersacyjnej jest przystosowanie maszyny do rozumienia ludzkich metod przekazu informacji i ludzkiego procesu myślenia. Aby komunikacja z maszyną była rzeczywiście zbliżona do porozumiewania się z człowiekiem, system umożli-

wiający taką komunikację powinien mieć następujące właściwości:

- możliwość przyswojenia informacji podawanych w języku naturalnym (lub przynajmniej zbliżonym do naturalnego),
- umiejętność przyswojenia informacji wizualnej (rysunków, tablic, wykresów itp.),
- odpowiedni (tzn. zbliżony do ludzkiego) czas reakcji,
- łatwą dostępność, czyli możliwość reagowania na informacje podawane w dowolnym momencie i z dowolnego, wygodnego dla użytkownika miejsca (np. przez telefon),
- możliwość przekazywania wyników w formie bezpośrednio zrozumiałej dla człowieka.

Wymienione cechy wpływają na wybór sprzętu i oprogramowania podstawowego stanowiącego bazę systemu (pkt 2.2) oraz na kierunek rozwijania oprogramowania na wyższym poziomie (pkt 2.3).

2.2. Sprzęt i oprogramowanie podstawowe

Najnaturalniejszą bazą dla programów konwersacyjnych byłaby maszyna z wejściem i wyjściem graficznym (umożliwiającym przekazywanie napisów i rysunków) i fonicznym (umożliwiającym przekazywanie tekstu mówionego), zaopatrzona w dostępny w każdej chwili system operacyjny, o odpowiednim czasie reakcji. Względny natury technicznej i ekonomicznej stanowią jednak przeszkodę w powszechnym stosowaniu takiego zestawu. Monitory ekranowe i inne urządzenia pozwalające wprowadzać i wyprowadzać informację graficzną są kosztowne. Urządzenia foniczne, zwłaszcza wejściowe, analizujące mowę - kosztowne i jeszcze w fazie eksperymentu. Dokładniejsze informacje na temat budowy takich urządzeń i ich stosowania w kraju i na świecie znajdują się w pracach [6], [12], [20] i [23].

Z wyżej wymienionych względów podstawowym urządzeniem wejściowo-wyjściowym, na którym opiera się większość działających systemów konwersacyjnych jest monitor ekranowy alfanumeryczny lub monitor dalekopisowy, a podstawową formą informacji wprowadzanej i wyprowadzanej - teksty, mające charakter zbliżony (mniej lub więcej) do wypowiedzi w języku naturalnym.

W dziedzinie systemów operacyjnych umożliwiających konwersację również mają znaczenie względy ekonomiczne. Z powodu strat wynikających z oczekiwania maszyny na odpowiedź człowieka, łączność konwersacyjna stała się możliwa dopiero w systemach z podziałem czasu, a ze względu na koszt całego systemu konwersacyjnego, w pełni opłacalna dopiero w systemach wielodostępnych, umożliwiających kontakt z maszyną przez linie transmisyjne (np. z mieszkania, uczelni, czy zakładu pracy użytkownika) i przez to zwiększających liczbę użytkowników.

Istotny z punktu widzenia konwersacji jest również problem gospodarki pamięcią w systemie operacyjnym. Przetwarzając wypowiedź człowieka lub wykonując jego polecenie system korzysta z dużych obszarów pamięci, które po zakończeniu żądanej akcji mogą być zwolnione, aż do chwili gdy człowiek ponownie wyrazi chęć współpracy z systemem. Najlepiej rozwiązuje ten problem dynamiczny przydział pamięci. Poza tym informacje tekstowe wprowadzone do maszyny, zwłaszcza w języku naturalnym i informacje graficzne nie są na ogół związane. Postać wewnętrzna tych informacji i aparat umożliwiający ich zrozumienie przez maszynę wymagają (jak to będzie widoczne w następnych punktach) wiele miejsca w pamięci i szybkiego dostępu. Zatem pojemność i czas dostępu do pamięci operacyjnej oraz możliwość dołączania do systemu szybkich pamięci pomocniczych o bezpośrednim dostępie mają znaczny wpływ na jakość programu konwersacyjnego. Systemy operacyjne o wymienionych właściwościach są już dzisiaj powszechnie stosowane. Podstawę dla programów konwersacyjnych mogą stanowić np. systemy: VM/370, OS lub TSS dla maszyn IBM, GEORGE dla ICL, SCOPE dla CDC, MULTICS dla GE i inne.

Ważnym problemem przy tworzeniu programów konwersacyjnych jest stosowanie odpowiednich języków programowania.

Znane są przykłady systemów konwersacyjnych kodowanych w PL1, FORTRAN-ie lub innych powszechnie używanych językach programowania. W ten sposób powstawały liczne systemy projektowania wspomaganego przez maszynę, ponieważ języki te były najlepiej dostosowane do pewnych zagadnień projektowania. Rozbudowywano je tylko o standardowe biblioteki podprogramów przeznaczonych do manipulacji danymi graficznymi. Tak np. FORTRAN dla OS IBM/360 wyposażono w GSP (Graphic

Subroutine Package) . Jednak przy realizacji systemów, w których podstawowym problemem jest analiza tekstów, wygodniej jest stosować języki wyspecjalizowane w tej właśnie dziedzinie. Na początku lat sześćdziesiątych, osyli w pierwszych latach rozwoju takich systemów, przy ich realizacji wykorzystywano np. języki: IPL, SLIP, COMIT¹. Najczęściej stosowany COMIT pomyślany był początkowo jako narzędzie ułatwiające maszynowy przekład tekstów w języku naturalnym. W latach 1963-64 karierę w dziedzinie analizy napisów zaczyna robić LISP 1.5, stosowany zresztą do manipulacji na napisach do dzisiaj. W drugiej połowie lat sześćdziesiątych stał się popularny SNOBOL 4 [11], ze względu na bogaty aparat do analizowania wprowadzanej wypowiedzi na podstawie napisów wzorcowych. Rozwój zastosowań komputerów w dziedzinie zwanej "sztuczną inteligencją", obejmujących rozumienie języka naturalnego, automatyczne rozwiązywanie problemów, automatyczne dowodzenie twierdzeń, podejmowanie decyzji i w ogóle modelowanie procesów psychicznych i intelektualnych człowieka w komputerze - zainspirował powstanie nowych języków², lepiej dostosowanych do programowania wyżej wymienionych zagadnień. Jako najważniejsze należy tu wymienić języki: PLANNER, CONNIVER, QA4, QLISP. Proces opracowywania i ulepszania tego typu języków rozpoczęty w końcu lat sześćdziesiątych trwa do tej pory.

Wszystkie języki wymienione w tym punkcie ułatwiają kodowanie fragmentów systemów konwersacyjnych niezależnych od zastosowań tych systemów. O językach zaprojektowanych dla specyficznych zastosowań będzie mowa w następnych punktach.

2.3. Język dialogu

Komunikacja z maszyną w tradycyjnym trybie współpracy odbywa się za pośrednictwem języków sformalizowanych zwanych językami programowania. W miarę rozwoju systemów konwersacyjnych używane języki mody-

¹ Zestawienia własności języków do przetwarzania struktur listowych opracowanych na początku lat sześćdziesiątych (w tym także LISP) były podane w polskiej literaturze w pracach [13] i [29].

² Informacje o językach przeznaczonych do rozwiązywania problemów sztucznej inteligencji można znaleźć w pracach [4] i [10].

fikowano w kierunku dostosowania ich do współpracy konwersacyjnej i w kierunku zbliżenia do języka naturalnego. Na czym polega dostosowanie sformalizowanego języka programowania do współpracy konwersacyjnej, wyjaśnia punkt 3.4, w którym omawiamy języki konwersacyjne do wykonywania obliczeń. Poniżej zajmiemy się zagadnieniem zbliżania języka dialogu do języka naturalnego. Na wstępie omawiamy celowość wprowadzania naturalnego dialogu z maszyną. Dalej podajemy przykłady ilustrujące kolejne etapy rozwoju konwersacji w języku naturalnym:

- tzw. bierne stosowanie języka naturalnego, przy którym wypowiedź człowieka jest sformalizowana, a tylko wypowiedzi maszyny są naturalne,
- stosowanie naturalnych wypowiedzi człowieka, z których tylko pewne słowa (tzw. słowa kluczowe) są rozpoznawane przez maszynę,
- próby rozpoznawania przez maszynę pełnej naturalnej wypowiedzi człowieka.

Czy językiem dialogu z maszyną w systemach konwersacyjnych powinien być język naturalny? Problem jest kontrowersyjny. J.S. Bień pisze np. w pracy [3] : "Przez system konwersacyjny rozumiemy program lub zespół programów, który wykorzystuje język naturalny do komunikacji z człowiekiem", a w książce J. Martina [14], która ukazała się również w 1973 r. można znaleźć taką np. wypowiedź: "... literatura fachowa pełna jest egzotycznych artykułów twierdzących, że najlepszą metodą porozumiewania się z maszyną powinien być dla człowieka jego własny język".

Na poparcie swojego stanowiska Martin dołącza liczne przykłady systemów działających w trybie konwersacyjnym (np. systemy rezerwacji miejsc lotniczych), w których w dialogu z maszyną operator stosuje skróty mnemotechniczne, klawisze funkcyjne lub języki programowania nie mające nic wspólnego z językiem naturalnym. Ponieważ operatorzy tych systemów pracujący na ogół długo na jednym stanowisku przyswoją sobie do takiej metody przekazywania informacji, zaś ze względów ekonomicznych i czasowych dąży się do minimalizacji danych przesyłanych liniami transmisyjnymi, wprowadzanie do takich systemów naturalnego dialogu jest zdaniem autora książki niecelowe. Tym bardziej, że prace nad stosowaniem języka naturalnego w konwersacji z maszyną są w

fazie eksperymentalnej, dalekiej jeszcze od możliwości praktycznych zastosowań. Pogląd ten ilustrują w książce Martina dwa przykłady. Pierwszym jest system KLIZA (opracowany w 1966 r.), który odpowiada pozornie sensownie na zdania w języku angielskim rozpoznając w nich tylko pojedyncze słowa, bez wnikania w treść wypowiedzi człowieka i ma bardzo ograniczone znaczenie praktyczne. W systemie DEACON, podanym jako drugi przykład (opracowanym w 1964 r.) są już pewne próby rzeczywistego rozumienia tekstu, ale tekstu o ubogim zakresie tematycznym, którego rozszerzenie prowadziłoby do istotnych zmian w aparacie realizacyjnym i wydłużało i tak niezadowolający czas reakcji systemu.

Nie przekonuje nas argument, że do skrótów mnemotechnicznych można się przyzwyczaić. Fakt, że system rezerwacji miejsc muszą obsługiwać przeszkoleni i wprawni agenci, jest przeszkodą w bezpośrednim udostępnieniu go np. pasażerom linii lotniczych. A dołączone przykłady ilustrujące stan badań nad stosowaniem dialogu naturalnego wydają nam się nieco przestarzałe. Nie negujemy, oczywiście, że obecny stan wiedzy na ten temat nie pozwala jeszcze na pełne i ekonomiczne stosowanie naturalnego dialogu i że nie wszędzie to jest celowe, uważamy jednak, że stosowanie języka naturalnego w konwersacji z maszyną będzie wypierać w wielu dziedzinach inne metody współpracy, ponieważ:

- jest to najłatwiejszy sposób przekazywania informacji,
- istnieją już proste metody realizacji systemów spełniających częściowo postulat naturalnego dialogu,
- daje się zauważyć znaczny rozwój metod prowadzących do pełnego rozumienia języka naturalnego przez maszynę.

Do powyższych konkluzji można dojść też na podstawie artykułu S. Waligórskiego "Problemy komunikacji człowiek-maszyna w języku naturalnym" [30]. Tutaj chcemy poprzeć je tylko kilkoma dodatkowymi przykładami.

Są dziedziny, np. wykonywanie obliczeń, w których przyzwyczajaliśmy się używać zwięzłego języka matematycznego i gdzie zbyt gadałli-

wosc bylaby nawet niepozadana. Dlatego w systemach konwersacyjnych przeznaczonych do wykonywania obliczen stosuje sie konwersacyjne wersje znanych jezykow programowania, lub specjalne jezyki konwersacyjne nie majace czasem nic wspolnego z jezykiem naturalnym. Np. dialog w APL moze wygladac tak:

$X \leftarrow (2,5,7)$	}	wypowiedz czlowieka
$B \leftarrow 2 (3) < (3,1,5)$		
X, B		
$(2,5,7,1,0,1)$		odpowiedz maszyny

Komus, kto nie zna jezyka APL, trudno byloby sie domyslec, ze zadaniem maszyny bylo obliczenie wspolrzędnych wektora bedacego konkatenacja¹ wektora o wspolrzędnych dziesietnych 2,5,7 nazwanego X i wektora (1,0,1) nazwanego B. Wektor B powstaje jako wynik porownania (relacja mniejszosci) odpowiednich wspolrzędnych wektorow (2,2,2) i (3,1,5). Gdy relacja jest spezniona odpowiednia wspolrzędna wektora B jest jedynka, w przeciwnym przypadku - zerem.

O wiele czytelniejsze i latwiejsze do zapamietania sa jezyki stosowane do obliczen, do ktorych wprowadzono juz pewne elementy jezyka naturalnego. Przykladem moze byc JOSS (Johnnieac Open Shop System, Rand Corporation, 1964) [24]. W zamieszczonym w tabeli 1 dialogu w jezyku JOSS male litery oznaczaja wypowiedz czlowieka, duze - maszyny.

Wypowiedzi maszyny sprawiaja tu wrazenie bardziej naturalnych niz wypowiedzi czlowieka. Jest tak dlatego, ze latwiej przygotowac w pamieci maszyny gotowy "naturalny" tekst niz zbudowac aparat analizujacy zwobodna wypowiedz czlowieka. Wypowiedzi czlowieka w ramach JOSS, pozornie naturalne, w rzeczywistosci sa sformalizowane i jest to przyklad tzw. biernego stosowania jezyka naturalnego w systemie konwersacyjnym (tzn. tylko maszyna "mowi" w jezyku naturalnym). Do powyzsze-go przykladu wrócimy jeszcze w nastepnych punktach.

1 Konkatenacja wektorow A i B nazywamy wektor powstajacy w wyniku dopisania wspolrzędnych wektora B za wspolrzędnyimi wektora A.

Tabela 1. Przykład dialogu w języku JOSS

dialog w JOSS	objaśnienia
<pre> 1.1 set x (1) = [-b+sqrt (b*2-4.a.c)]/[2.a] 1.2 set x (2) = [-b-sqrt(b*2-4.a.c)]/[2.a] 1.3 type x (1), x (2) in form 1. set b=-1. set c=-6. do part 1. ERROR AT STEP 1.1: A IS UNDEFINED. set a=1. go. ERROR AT STEP 1.3: I CAN'T FIND THE REQUIRED FORM. form 1: a= ___ b= ___ c= ___ roots: ___ . ___ . ___ go. ERROR AT STEP 1.3.: I HAVE TOO FEW VALUES FOR THE FORM. 1.3 type a,b,c,x (1) , x (2) in form 1. go. A=1 B=-1 C=-6 ROOTS: 3.0000 -2.0000 do part 1 for c=-1 (1) -4. A=1 B=-1 C=-1 ROOTS: 1.6180 -.6180 . . . </pre>	<p>niech $x (1) = \dots$</p> <p>wydrukuj $x (1)$, $x (2)$ wg.wzorca 1</p> <p>wykonaj część 1</p> <p>błąd w zdaniu 1.1: niezdefiniowane a</p> <p>licz dalej</p> <p>błąd w zdaniu 1.3: nie mogę znaleźć żdanego wzorca.</p> <p>\dots mam za mało war- tości dla tego wzorca.</p> <p>wykonaj część 1 dla $c = \dots$</p>

Innym przykładem biernego stosowania języka naturalnego jest tzw. metoda "jadłospisowa". W systemie opracowanym według tej metody maszyna po zadaniu pytania przedstawia kilka propozycji odpowiedzi, a człowiek tylko wybiera z nich właściwe. Ze względu na łatwość realizacji systemu opartego na tej metodzie, mimo pojawienia się innych metod, pozwalających na czynną rozmowę, "jadłospisu" używa się do dzisiaj np. w systemach stosowanych w medycynie, przeprowadzających wstępny wywiad z pacjentem lub przy nauczaniu wspomaganym przez maszyny. Metodę jadłospisową ilustruje przykład zamieszczony w tabeli 2. Są to fragmenty dialogu z systemem uczącym programowania w języku BASIC (system opracowany przez B. Beckera, General Electric Company, Phoenix, Arizona, 1966) [1]. Dialog przetłumaczono z języka angielskiego. Odpowiedzi człowieka są podkreślone.

Program realizujący dialog oparty na zasadzie jadłospisu lub na zbliżonych zasadach (jak np. wpisywanie pojedynczych wyrazów w miejsce kropek, podobne do wypełniania kwestionariusza) nie musi "rozumieć" skomplikowanych wypowiedzi człowieka; wystarcza tu rozpoznawanie kilku cyfr lub przewidzianych słów. Przygotowanie takiego programu nie jest zatem trudne z punktu widzenia analizy języka naturalnego; pojawiające się przy tym problemy będą omówione w następnych punktach.

ELIZA

Inny, stosunkowo prosty aparat analizy języka naturalnego, dopuszczający dialog czynny, opiera się na zasadzie rozpoznawania w wypowiedzi człowieka tzw. słów kluczowych. Zastosowany w programie ELIZA (J. Weizenbaum, Massachusetts Institute of Technology, 1965) [31], [32] przyczynił się do światowej kariery tego programu. Prostymi środkami osiągnięto wrażenie naturalności dialogu i stąd taka popularność systemu ELIZA. Ze względu na tę popularność uważamy, że i tutaj należy poświęcić temu systemowi parę słów. Ponieważ jednak przykłady dialogu z programem ELIZA i zasadę jej działania podawano już w polskiej literaturze [19], [30] potraktujemy to zagadnienie skrótowo.

Tabela 2. Przykład zastosowania metody jadłospisowej

LEKCJA 1 SAMOUCZEK JĘZYKA BASIC

PROGRAM TEN JEST KURSEM PODSTAW PROGRAMOWANIA W JĘZYKU BASIC. OPIERA SIĘ NA ZASADZIE SAMOKSZTAŁCENIA. JAK ZOBACZYSZ DALEJ, BĘDZIE PRZEDSTAWIAŁ CI INFORMACJE OPRAZ ZADAWAŁ PYTANIA DAJĄC DO WYBORU KILKA ODPOWIEDZI. ŻEBY ODPOWIEDZIEĆ NA PYTANIE NAPISZ, PO PROSTU, NUMER WYBRANEJ ODPOWIEDZI A POTEM NACIŚNIJ KLAWISZ "RETURN" PO PRAWEJ STRONIE KLAWIATURY... CZY JESTEŚ GOTOWY?

1 - TAK, 0 - NIE TWOJA ODPOWIEDŹ? 1

DOBRZE. ZACZYNAJEMY.

INFORMACJA:

PROGRAM W JĘZYKU BASIC JEST UPORZĄDKOWANYM ZBIOREM ZDAŃ PRZEZNACZONYCH DO STEROWANIA KOMPUTEREM PODCZAS ROZWIĄZYWANIA PROBLEMU.

PYTANIE:

KTÓRA METODA OKREŚLANIA PORZĄDKU ZDAŃ JEST WEDŁUG CIEBIE LEPSZA

1 - USTAWIENIE ZDAŃ WE WŁAŚCIWEJ KOLEJNOŚCI

2 - POPRZEDZANIE KAŻDEGO ZDANIA NUMEREM I USTAWIANIE ICH W DOWOLNYM PORZĄDKU.

TWOJA ODPOWIEDŹ? 1

WTEDEY WYNIKIŁYBY PROBLEMY GDYBYŚMY COŚ OPUŚCILI NA POCZĄTKU LUB CHCIELI ZMIENIĆ WEWNĄTRZ PROGRAMU. SPRÓBUJ JESZCZE RAZ: 3

DOPUSZCZALNĄ ODPOWIEDZIĄ JEST TYLKO 1 LUB 2.

TWOJA ODPOWIEDŹ? 2

DOBRZE. BĘDZIEMY ZATEM NUMEROWAĆ KAŻDE ZDANIE I USTAWIAĆ JE W DOWOLNEJ KOLEJNOŚCI.

ZAPAMIĘTAJ: ZDANIE MUSI SIĘ ZMIĘŚCIĆ W JEDNYM WIERSZU MONITORA.

PYTANIE:

CZY MYŚLISZ, ŻE ZDANIA POWINNY BYĆ NUMEROWANE KOLEJNO: 1,2,3,4 ITD.?

1 - TAK, 0 - NIE. TWOJA ODPOWIEDŹ?

⋮
⋮
⋮

KONIEC LEKCJI 1

PO 6 PRÓBACH ODPOWIEDZIAŁEŚ PRAWIDŁOWO NA 4 PYTANIA OSIĄGAJĄC 66,667 PROCENT EFEKTYWNOŚCI. GRATULUJĘ!

JEŻELI JESTEŚ GOTOWY DO LEKCJI 2 WYWOŁAJ PROGRAM BASIC 2.

CZAS: 12 SEKUND.

Program ELIZA opiera się na tzw. scenariuszach, związanych z tematem rozmowy, w których zawarte są słowa kluczowe dla wybranego tematu i reguły budowania odpowiedzi, zależne od danego słowa kluczowego i kontekstu, w jakim się ono znajduje. Umieszczone w tabeli 3 fragmenty scenariusza (tematem jest rozmowa w psychiatry, ELIZA odgrywa rolę doktora) dotyczące słów kluczowych ALWAYS, ALIKE, SAME i YOU wyjaśniają zasadę generowania odpowiedzi programu ELIZA po wykryciu w wypowiedzi człowieka określonego słowa kluczowego. Fragmenty podane są w oryginalnej wersji angielskiej, ponieważ znaczenie słów nie jest w systemie ELIZA istotne, a realizacja jej jest dostosowana do składni i fleksji języka angielskiego. Na zdanie "you and my brother are quite alike" (ty i mój brat jesteście całkiem podobni) ELIZA odpowie: IN WHAT WAY? (w jaki sposób), ponieważ słowo ALIKE ma najwyższy priorytet wśród występujących w tym zdaniu słów kluczowych. Przy następnym pojawieniu się tej samej wypowiedzi, lub wypowiedzi tego samego typu, ELIZA odpowie: WHAT RESEMBLANCE DO YOU SEE? (jaki widzisz podobieństwo?). W wypowiedzi: "you are stupid" (jesteś głupi) wybranym słowem kluczowym będzie YOU i po zastosowaniu odpowiedniego schematu ELIZA odpowie: WHAT MAKES YOU THINK I AM STUPID? (na jakiej podstawie sądzisz, że jestem głupi?). Czasem przeniesienie do odpowiedzi fragmentu wprowadzanego tekstu wymaga jeszcze dodatkowych czynności. Gdyby zamiast "stupid" w powyższym zdaniu użyto zwrotu "my friend" odpowiedź brzmiałaby: WHAT MAKES YOU THINK I AM YOUR FRIEND. Słowo "my" (mój) zamieniono tu na YOUR (twój).

Tabela 3. Fragment scenariusza systemu ELIZA

słowo kluczowe (priory- tet)	kontekst	odpowiedź	forma po przeniesie- niu do odpo- wiedzi
ALWAYS (1)	dowolny	<ul style="list-style-type: none"> - CAN YOU THINK OF A SPECIFIC EXAMPLE ? - WHEN ? - WHAT INCIDENT ARE YOU THINKING OF ? - REALLY, ALWAYS ?! 	ALWAYS
ALIKE, SAME (10)	dowolny	<ul style="list-style-type: none"> - IN WHAT WAY ? - WHAT RESEMBLANCE DO YOU SEE ? - WHAT DOES THAT SIMILARITY SUGGEST TO YOU ? - WHAT OTHER CONNECTIONS DO YOU SEE ? - WHAT DO YOU SUPPOSE THAT RESEMBLANCE MEANS ? - WHAT IS THE CONNECTION, DO YOU SUPPOSE ? - COULD THERE REALLY BE SOME CONNECTION ? - HOW ? 	ALIKE, SAME
YOU (0)	<p>... YOU REMIND</p> <p>1 2 3 4 ...YOU ARE...</p> <p>1 2 3 4 ...YOU...ME</p> <p>1 2 3 ...YOU...</p>	<p>jak przy ALIKE</p> <ul style="list-style-type: none"> - WHAT MAKES YOU THINK I AM 4 ? - DOES IT PLEASE YOU TO BELIEVE I AM 4 ? - DO YOU SOMETIMES WISH YOU WERE 4 ? - PERHAPS YOU WOULD LIKE TO BE 4 . - WHY DO YOU THINK I 3 YOU ? - YOU LIKE TO THINK I 3 YOU - DON'T YOU ? - WE WERE DISCUSSING YOU NOT ME . - OH, I 3 ?! 	I

Jak widać wszystkie te operacje nie wymagają pełnego "rozumienia" opracowywanego zdania ani zdań poprzednich. W efekcie rozumienia pojedynczych, wybranych słów system może jednak wykonywać pewne czynności, np. wybierać odpowiedni tekst, który przekazuje w odpowiedzi. Nawet takie prymitywne rozumienie wypowiedzi człowieka może mieć, jak to będzie pokazane dalej, praktyczne zastosowanie.

Oczywiście o wiele bardziej frapujące jest przystosowanie maszyny do rozumienia pełnej wypowiedzi człowieka - rozumienia polegającego na kojarzeniu nowo wprowadzanych informacji z tymi, które są już zapisane w pamięci maszyny i formułowaniu odpowiednich wniosków.

Eksperymentalne programy, które w tym sensie próbują rozumieć wypowiedź człowieka, analizują ją w trzech głównych płaszczyznach: syntaktycznej, semantycznej i pragmatycznej. Analiza syntaktyczna polega na sprawdzeniu czy wypowiedź jest poprawna gramatycznie i na określeniu, jakie funkcje w zdaniu pełnią jej poszczególne elementy. W wyniku analizy syntaktycznej program stwierdza np., że zdanie "Biała dama straszy nocą" jest poprawne i że "dama" jest podmiotem, "straszy" orzeczeniem itd. Takie wstępne rozpoznanie ułatwia potem analizę semantyczną, której celem jest określenie, czy dana wypowiedź ma znaczenie i czy tylko jedno, czy wiele. Np. zdanie "Straszne były mordy tych Tatarów" ma dwa znaczenia, a zdanie "Bezbarwne zielone myśli śpią wściekle" nie ma w ogóle sensu, mimo że jest poprawne gramatycznie. Ustalenie, które z możliwych znaczeń należy przyporządkować aktualnej wypowiedzi jest już zadaniem analizy pragmatycznej. Zależy to od okoliczności, w których zdanie zostało wypowiedziane, a więc od poprzednich wypowiedzi, od wiedzy człowieka prowadzącego dialog, od tematu dialogu itd. Na przykład zdanie "Chłopiec ma świnkę" w diagnozie lekarskiej będzie brzmiało jednoznacznie, chociaż ogólnie ma kilka znaczeń.

W opracowywanych obecnie systemach uwzględnia się zwykle powyższe trzy aspekty badania wypowiedzi człowieka, jednak nacisk na nie może nie być jednakowy. W opisanym niżej dla ilustracji systemie MUSB nałożono np. pewne ograniczenia na język angielski, w którym wprowadza

się wypowiedzi: wyrazy muszą być w formie nieodmiennej, spójniki są dopuszczalne tylko między zdaniami współrzędnymi i podrzędnymi a nie między rzeczownikami, czasownikami itd., niedozwolona jest negacja i wprowadzanie kwantyfikatorów, zabronione pewne typy zdań podrzędnych. W wyniku tych ograniczeń aparat analizy syntaktycznej jest bardzo uproszczony i np. pominięto zupełnie problem analizy fleksyjnej.

Dla języków fleksyjnych, takich jak język polski, problem ten jest bardzo istotny i dlatego grupa z Uniwersytetu Warszawskiego zajmująca się badaniami nad konwersacją w języku polskim zaczęła pracę właśnie od tego zagadnienia (zob. system MARYSIA, [2], [3]).

MUSE

System MUSE (A Model to Understand Simple English) był opracowany na University of Alberta w Kanadzie. Implementowano go w SNOBOL-u [15]. Na zdanie wprowadzone w ograniczonym języku angielskim MUSE odpowiada wypisując swoją interpretację tego zdania.

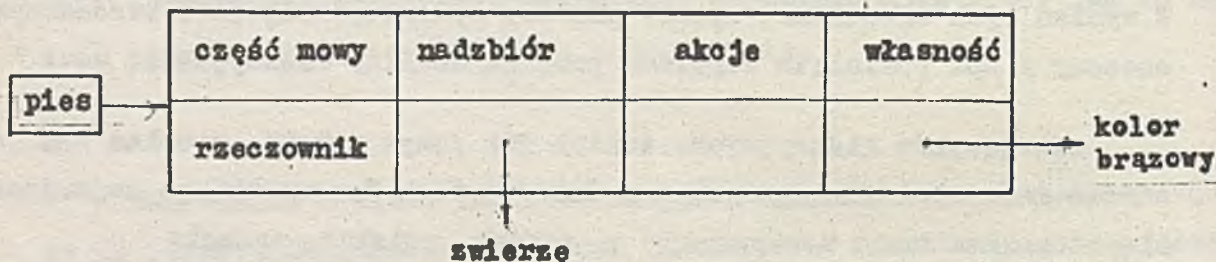
Celem poniższego opisu jest wyjaśnienie zasad interpretowania przez system wprowadzanego tekstu. Podanie przykładów tekstów w "uproszczonym" języku polskim wystarczy do zrozumienia zasady działania systemu.

Interpretowanie podanego tekstu przez MUSE polega na kojarzeniu nowych informacji z tymi, które są już zapisane w pamięci systemu. Pamięć MUSE dzieli się na: słownik, pamięć semantyczną i plik syntaktyczny. Słownik jest spisem słów znanych systemowi. Pamięć semantyczna zawiera informacje o każdym słowie zapisanym w słowniku i jego powiązaniach z innymi słowami. Pamięć ta dzieli się na plik jednostek semantycznych i plik własności. Jednemu słowu ze słownika odpowiada jedna lub więcej jednostek semantycznych, zależnie od liczby znaczeń tego słowa i odwrotnie, kilku synonimom odpowiada ta sama jednostka. Każda jednostka semantyczna zawiera następujące informacje:

- część mowy,
- numer (w pliku jednostek) jednostki opisującej pojęcie nadrzędne czyli tzw. nadzbiór,
- akcje MUSE związane z daną jednostką,

- 4) numer lub numery (w pliku własności) własności przypisanych danej jednostce.

Ilustruje to rysunek 1 opisujący jednostkę "pies".



Rys. 1. Schemat jednostki semantycznej: pies

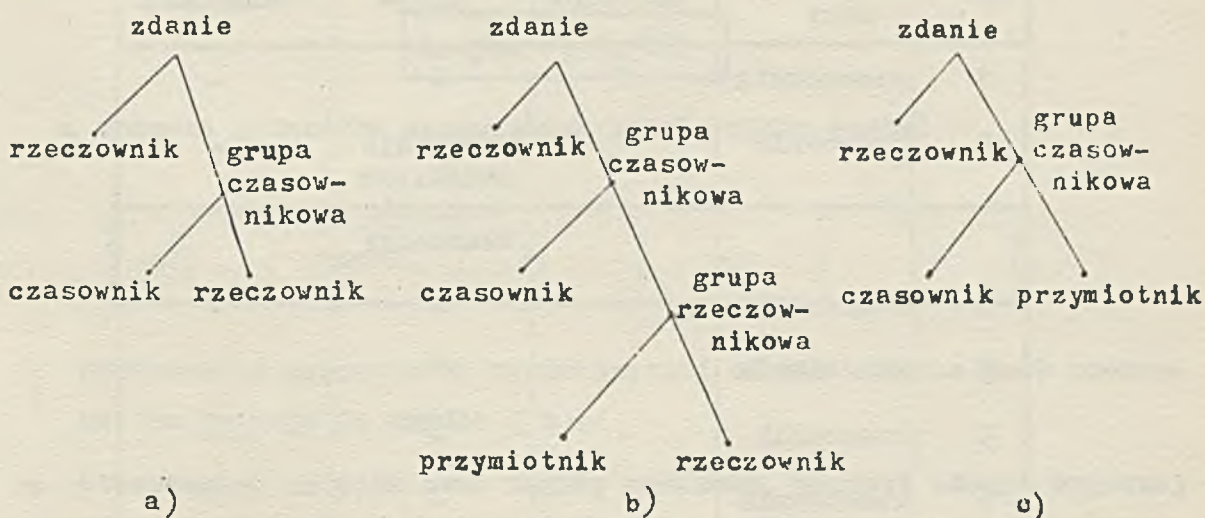
Informacje o części mowy i ewentualnych akcjach MUSE związanych z daną jednostką muszą towarzyszyć nowym słowom przy wprowadzaniu, pozostałe informacje mogą być dopisywane do jednostki semantycznej w wyniku późniejszego dialogu z MUSE. Ze słowem "pies" nie wiąże się żadna akcja systemu, ale np. ze słowem "równoważne" wiąże się akcja utódszania jednostek występujących po obydwu stronach tego słowa, a ze słowem "być" akcja tworzenia nadzbioru, gdy "być" występuje przed rzeczownikiem (pies być zwierzę) albo akcja tworzenia własności, gdy "być" występuje przed przymiotnikiem (pies być brązowy). Własności zapisywane są w pliku własności w postaci (cecha - wartość) np. (kolor brązowy) lub (prędkość wolno). Własnością jednostki "pies" mogłaby być też para (gonić samochód) gdyby wcześniej wprowadzono zdanie "pies gonić samochód".

Plik syntaktyczny służy do sprawdzania poprawności syntaktycznej wprowadzonych wypowiedzi. Zapisane są tam reguły rozkładu typów zdań akceptowanych przez MUSE, np. zdanie "koń jeść siano" będzie akceptowane przez MUSE, jeżeli w jej pliku syntaktycznym znajduje się reguła rozkładu zilustrowana rysunkiem 2a. Będą wówczas akceptowane też wszystkie zdania typu: rzeczownik - czasownik - rzeczownik.

Mechanizm interpretowania przez MUSE tekstów wejściowych pokazemy na poniższym przykładzie.

Założmy, że w pamięci MUSE zapisane są już pojęcia i relacje między tymi pojęciami, uwidocznione w tabelach 4, 5, 6.

Oznacza to, że system przyswoił sobie już takie informacje: (kropka NADZB. znak), (fiat NADZB. samochód), (pies NADZB. zwierzę), (pies gonić samochód), (zielony NADZB. kolor). Załóżmy, że plik syntaktyczny zawiera reguły rozkładu przedstawione na rysunku 2.



Rys. 2. Reguły rozkładu zdania

Tabela 4.

Słownik systemu MUSE

słowo	n-ry jednostek
być	2
fiat	3
gonić	5
kolor	7
kropka	1
pies	8
samochód	4
zielony	9
znak	6
zwierzę	10

Tabela 5.

Plik jednostek semantycznych systemu MUSE

nr	część mowy	nadzbior	akcje	własności
1	rzeczownik	6	-	-
2	czasownik	-	tworzenie nadzbioru tworzenie własności	-
3	rzeczownik	4	-	-
4	rzeczownik	-	-	-
5	czasownik	-	-	-
6	rzeczownik	-	-	-
7	rzeczownik	-	-	-
8	rzeczownik	10	-	1
9	przymiotnik	7	-	-
10	rzeczownik	-	-	-

Tabela 6.

Plik własności systemu MUSE

nr	cecha	wartość
1	gonić	samochód

Przypuśćmy teraz, że przedstawiamy systemowi do interpretacji kolejne zdania:

- kropka (rzeczownik) być pies,
- fiat być zielony,
- kropka gonić zielony fiat.

Pierwsze zdanie spowoduje następujące działanie MUSE.

- Zaznaczenie w słowniku, że pojawiło się nowe znaczenie słowa "kropka" i że przyporządkowuje się mu roboczo nową jednostkę semantyczną. W słowniku na piątej pozycji będzie teraz zapisane:

kropka	1, (11)
--------	---------

a robocza jednostka semantyczna będzie miała postać:

(11)	rzecz.	-	-	-
------	--------	---	---	---

- Sprawdzenie poprawności syntaktycznej zdania. Zdanie jest poprawne, bo opisuje je reguła 2 a).
- Utworzenie, na podstawie reguły rozkładu, postaci zdania wygodnej do analizy semantycznej, tzw. postaci trójkowej. W tym wypadku po prostu: (kropka-2 być pies)¹.
- Wykonanie akcji tworzenia nadzbioru. Jednostka semantyczna "kropka-2" będzie miała teraz postać

(11)	rzecz.	8	-	-
------	--------	---	---	---

co odpowiada trójce (kropka-2 NADZB. pies).

- Porównanie z trójkami typu (jednostka cecha wartość) i z trójkami typu (jednostka NADZB. jednostka) zapisanymi już na stałe w pamięci MUSE. Porównanie kończy się sukcesem, jeżeli MUSE znajdzie w pamięci trójkę, której elementy leżą na tych samych łańcuchach nadzbiorów, co odpowiednie elementy trójki porównywanej. Gdyby wcześniej wprowadzono np. zdania: "kropka być jamnik" i "jamnik być pies" - zdanie "kropka być pies" byłoby już dla MUSE zrozumiałe (porównanie skończyłoby się sukcesem).

¹ Różne znaczenia tego samego słowa będziemy wyróżniać kolejnymi liczbami, np. kropka-1, kropka-2.

W tym przykładzie kropka-2 jest nowym słowem i nie ma jeszcze w pamięci MUSE stałych powiązań. Ponieważ jednak trójka (kropka-2 NADZIE pies) jest jedyną możliwą interpretacją wprowadzonego zdania MUSE odpowie:

PRAWDOPODOBNIIE WEJŚCIE NALEŻY INTERPRETOWAĆ TAK:

KROPKA-2 BYĆ PIES (PIES POJĘCIE NADRZĘDNE WZGLĘDEM KROPKA-2)

CZY INTERPRETACJA JEST PRAWIDŁOWA ?

Jeżeli człowiek zaakceptuje taką interpretację, MUSE zapyta, czy dołączyć ją na stałe do pamięci. Zgoda człowieka spowoduje przepisanie jednostki kropka-2 do pliku semantycznego. Założmy, że mamy ją już w pliku semantycznym.

Działanie MUSE przy interpretacji drugiego wprowadzonego zdania (fiat być zielony) jest podobne, z tym, że nie tworzy się nowej jednostki semantycznej, bo wszystkie słowa są znane i że słowo "być" prowadzi w tym przypadku do akcji tworzenia własności. W efekcie MUSE przedstawi poniższą interpretację wejścia:

FIAT KOLOR ZIEŁONY

Założmy, że interpretację zaakceptowano i że informacje w pamięci semantycznej MUSE związane ze słowem fiat mają teraz następującą postać:

plik jednostek semantycznych

plik własności

nr	cz.mowy	nadzb.	akcja	własność
3	rzeczownik	4	-	2

nr	cecha	wartość
2	kolor	zielony

Trzecie wprowadzone zdanie (kropka gonić zielony fiat) będzie już zrozumiałe dla MUSE. Ponieważ jednak słowo "kropka" ma dwa znaczenia system będzie rozpatrywać kolejno dwie możliwe interpretacje:

- kropka-1 gonić zielony fiat
- kropka-2 gonić zielony fiat

Gdyby inne słowa tego zdania też miały kilka znaczeń, liczba kombinacji odpowiednio by wzrosła, ale niektóre mogłyby być odrzucone już w trakcie analizy syntaktycznej. W naszym przykładzie obie interpretacje są poprawne syntaktycznie (odpowiada im reguła na rys. 2 c) a więc dojdą do analizy semantycznej, każda w postaci dwóch trójek:

- I. (kropka-1 gonić fiat) (fiat kolor zielony)
- II. (kropka-2 gonić fiat) (fiat kolor zielony).

Zajmijmy się możliwością I. Druga trójka jest dla MUSE zrozumiała, bo dokładnie taką samą ma już zapisaną w pamięci, ale pierwsza nie, bo kropka-znak nie ma własności "gonić fiat", ani "gonić samochód". Ogólniej - MUSE nie znajdzie w pamięci trójki, której elementy leżałyby na odpowiednich łańcuchach nadzbiorów. Interpretacja ta zostanie więc odrzucona i MUSE zajmie się drugą możliwością. Trójka (kropka-2 gonić fiat) będzie zrozumiała w wyniku skojarzenia z trójką (pies gonić samochód). Zatem odpowiedź MUSE w związku z trzecim wprowadzonym zdaniem będzie brzmiała:

PRAWDOPODOBNIIE WEJŚCIE NALEŻY INTERPRETOWAĆ TAK:
KROPKA-2 GONIĆ FIAT. FIAT KOLOR ZIELONY
CZY INTERPRETACJA JEST PRAWIDŁOWA ?

Przedstawiony powyżej dosyć szczegółowo proces myślenia MUSE daje pojęcie o złożoności problemu i objętości informacji, które mogą być zaangażowane przy interpretacji nawet prymitywnych wypowiedzi człowieka. Rozszerzanie zakresu tematycznego wypowiedzi prowadzi do lawinowego powiększania liczby potrzebnych informacji i powiązań. Problem organizacji ogromnych zbiorów tego typu danych nie jest jeszcze rozwiązany. Dlatego lepszy aparat analizujący wypowiedź człowieka osiąga się w systemach eksperymentalnych kosztem ograniczania zakresu pojęć rozumianych przez system.

ROBOT WINOGRADA

Na przykład system opracowany przez T. Winograda (Massachusetts Institute of Technology, 1972) [33] symuluje robota, którego świat zewnętrzny ogranicza się do kolorowych klocków w formie sześciątów, ostrosłupów, prostopadkościanów itp. przedstawianych przez robota lub ochowanych do pudełka w wyniku dyrektyw podawanych mu przez człowieka. Dyrektywy podawane są w języku naturalnym a inspirowane nimi działania robota widoczne jest na ekranie monitora graficznego. Robot rozumie o wiele bogatszy język niż MUSE. Potrafi np. odpowiedzieć na pytanie typu: "Czy miałeś już w rękę jakiś sześciąt?" lub wykonać polecenie: "Włóż do pudełka wszystkie czerwone klocki", ponieważ do aparatu analizującego wypowiedzi dołączono rachunek kwantyfikatorów i elementy automatycznego dowodzenia twierdzeń.

Tabela 7. Przykład dialogu z robotem

Ile prostopadkościanów leży obok pudełka ?

CZTERY

Czy co najmniej jeden z nich jest węższy niż ten, który kazałem ci podnieść ?

TAK, CZERWONY SZEŚCIĄT

Czy stół może podnosić klocki ?

NIE

Czy ostrosłup może się opierać na prostopadkościanie ?

TAK

Który sześciąt leży na stole ?

DUŻY ZIELONY, KTÓRY PODPIERA CZERWONY OSTROSŁUP

Czy dotykałeś jakiegoś ostrosłupa zanim położyłeś zielony na małym sześciącie ?

TAK, ZIELONEGO

Kiedy go podniosłeś ?

PODCZAS GDY USTAWIAŁEM NA STOSIE CZERWONY SZEŚCIĄT, DUŻY CZERWONY PROSTOPADKOŚCIĄT I DUŻY ZIELONY SZEŚCIĄT

Dlaczego ?

ŻEBY SIĘ GO POZBYĆ

•
•
•

Załączony jako tabela 7 fragment dialogu z robotem (wypowiedzi człowieka pisane są małymi literami) ilustruje stopień jego "inteligencji" i jednocześnie poziom osiągnięć w dziedzinie rozumienia przez maszynę języka naturalnego. Problem nie jest jeszcze ostatecznie rozwiązany i prace badawcze nad tym zagadnieniem prowadzone są w dwóch zasadniczych kierunkach:

- opracowanie takiej organizacji danych, przy której można by dowolnie uwzględnić zakres tematyczny dialogu zapewniając jednocześnie właściwe powiązania wszystkich nowowprowadzonych pojęć z już istniejącymi i łatwy dostęp do tych pojęć i powiązań;
- uelastycznienie aparatu analizy wypowiedzi człowieka, przez zerwanie z tradycyjnym ścisłym rozgraniczeniem działań aparatu analizującego na dziedziny syntaktyki, semantyki i pragmatyki.

W rzeczywistości te dziedziny się zazębiają, np. pozornie jednoznaczne słowo "czerwony" znaczy co innego w zwrocie "czerwony klocek" niż w zwrocie "czerwona oberża", a jeszcze co innego gdy mówimy, "czerwony odłam partii". Pozornie niepoprawne gramatycznie zdanie: "Bardzo i chcę go kupić" jest sensowną odpowiedzią np. na pytanie "Czy podoba ci się obraz?". Współdziałanie programów analizujących wypowiedzi na wielu różnych poziomach wymaga jednak specjalnej organizacji danych. Musi to być postać wygodna do analizy na danym poziomie, która jednocześnie daje się łatwo przetransponować na formę z innego poziomu.

3. Stosowanie systemów konwersacyjnych

3.1. Uwagi ogólne

Omówimy tu różne dziedziny stosowania współpracy konwersacyjnej: nauczanie wspomagane przez maszyny, komunikacja z systemem operacyjnym, projektowanie za pomocą maszyn, wykonywanie obliczeń, wyszukiwanie informacji.

Nie są to wszystkie możliwe dziedziny zastosowań systemów konwersacyjnych. Można wymienić również medycynę (np. przeprowadzanie wstępnego wywiadu z pacjentem), zastosowania administracyjne (np.

prowadzenie ewidencji ludności), zastosowania komercyjne (np. systemy rezerwacji miejsc) i inne dziedziny. Przeprowadzanie wywiadu z pacjentem podobne jest do lekcji w ramach nauczania wspomaganego przez maszynę. Systemy ewidencji i rezerwacji to pewnego typu systemy wyszukiwania informacji. Większość typowych problemów pojawiających się przy wymienionych zastosowaniach będzie więc omówiona w tym rozdziale.

Przy analizie poszczególnych zastosowań będzie położony nacisk na rodzaje współpracy konwersacyjnej właściwe dla omawianych systemów, a nie na problemy związane z tematyką konkretnych zastosowań.

3.2. Nauczanie wspomagane przez maszynę

Dyscyplina ta u nas jest również nazywana "dydaktyką komputerową" i "nauczaniem skomputeryzowanym" lub oznaczana skrótem CAI pochodzącym od określeń angielskich "Computer-Assisted Instruction" lub "Computer-Aided Instruction". Charakterystyczne dla tej dziedziny zastosowań maszyn matematycznych jest odwrócenie roli człowieka i maszyny podczas współdziałania. W tym przypadku człowiek nie podaje maszynie informacji i instrukcji, ale sam jest przez nią (a dokładniej przez odpowiedni program) informowany, pytany, instruowany. Poniżej omówimy różne typy programów uczących, wskażemy problemy związane z realizacją programów danego typu oraz podamy przykłady tzw. języków autorskich, będących narzędziem do realizacji tych programów.

Systemy uczące, niezależnie od przedmiotu nauczania, mogą stosować metody dopuszczające mniej lub bardziej czynne zachowanie się ucznia. Wyróżnia się tu dwa główne schematy. W pierwszym, materiał jest prezentowany uczniowi małymi porcjami, po każdej porcji następują pytania i ewentualne dodatkowe wyjaśnienia w przypadku nieprawidłowej odpowiedzi. Proces uczenia przebiega z inicjatywy systemu - instruktora, a nie ucznia. Nazwijmy ten schemat instruktorskim (ang. "tutorial"). W drugim schemacie system stawia przed uczniem większe zadanie i pozwala mu (oczywiście w ograniczonym zakresie) wybierać metody rozwiązywania i odwoływać się do systemu po potrzebne informacje. Nazwijmy taki schemat sokratycznym (ang. "socratic"). System uczący programowania w języku BASIC, wspomniany w pkt 2.3 jest typowy

wym przykładem systemu instruktorskiego, tzn. systemu opartego na schemacie instruktorskim.

Oprócz zasadniczych lekcji, prowadzonych według jednego z powyżej wspomnianych schematów, przewidziane są też lekcje, w których stosuje się pomocnicze schematy pozwalające:

- organizować utrwalanie materiału przez serię pytań i ćwiczeń (schemat ćwiczeniowy),
- pomagać przy rozwiązywaniu zadanych problemów - np. matematycznych lub inżynierskich (schemat problemowy),
- ułatwiać symulowanie doświadczeń - np. chemicznych lub fizycznych (schemat symulacyjny),
- ułatwiać dostęp do informacji encyklopedycznej zapisanej w pamięci maszyny (schemat informacyjny).

Stosowanie różnych schematów uczenia powinno być elastyczne. System powinien móc zmieniać metody uczenia nawet w obrębie jednej lekcji, zależnie od reakcji ucznia. Na razie są to jednak tendencje przyszłościowe a istniejące systemy CAI budowane są na ogół według jednego z wymienionych wyżej schematów, głównych lub pomocniczych.

Trudności realizacyjne są różne przy różnych typach schematów. W systemach informacyjnych występuje problem organizacji banku danych i dostępu do nich, omawiany dokładniej w pkt 3.5. Systemy problemowe muszą uwzględniać jak najwięcej możliwych metod rozwiązywania problemów, dla których są przeznaczone i możliwych błędów ucznia, powinny więc być opracowywane iteracyjnie, tzn. modyfikowane po zanotowaniu w czasie działania nowych reakcji i błędów ucznia. Z symulowaniem doświadczeń wiążą się zagadnienia operowania danymi graficznymi. Aparat rozumiejący wypowiedzi ucznia w systemach instruktorskich może opierać się nawet na metodzie jadłospisowej, ale przy dialogu sokratycznym lub w systemach informacyjnych konieczne jest już przynajmniej rozumienie słów kluczowych występujących w wypowiedzi. Wspomniana ELIZA również była wykorzystywana do celów CAI [27] i potrafiła prowadzić dialog sokratyczny na temat zadanego problemu. Przykład

Tabela 8. Przykład dialogu sokratycznego z systemem ELIZA

ZASTANÓW SIĘ TERAZ JAK MOŻNA ZSYNCHRONIZOWAĆ DWA ZEGARY, TAK ABY WSKAZYWAŁY TEN SAM CZAS, GDY JEDEN Z NICH JEST W SAN FRANCISCO A DRUGI W NOWYM JORKU. ROZWAŻ PROBLEM A JEŚLI PRZYJ- DA CI NA MYSL JAKIEŚ WATPLIWOŚCI, ZADAJ MI PYTANIA. GDY BĘDZIESZ GOTOWY ZAPROPONOWAĆ METODĘ SYNCHRONIZACJI NAPISZ "GOTÓW". WTEDY PRZEDYSKUTUJĘ Z TOBĄ WYBRANĄ METODĘ.

Jak daleko jest z Nowego Jorku do San Francisco?
ODLEGŁOŚĆ Z NOWEGO JORKU DO SAN FRANCISCO WYNOŚI OKOŁO 3000 MIL LUB 5000 KM.

Powiedz coś o różnych strefach czasu.
STREFY CZASU, CZAS LETNI - SĄ WYNIKIEM USTALEŃ WPROWADZONYCH DLA LUDZKIEJ WYGODY, A NIE PRAW FIZYCZNYCH. ZAUWAŻ, ŻE GDYBY NASZĄ METODĘ SYNCHRONIZACJI ROZSZERZYĆ NA CAŁĄ ZIEMIĘ WSZYSCY JEJ MIESZKAŃCY OPEROWALIBY TYM SAMYM, STANDARDOWYM CZASEM.

Czy mogę użyć sygnałów świetlnych?
OCZYWIŚCIE, SWOBODNIE UŻYWAJ SYGNAŁÓW ŚWIETLNYCH.

Czy mogę stosować zegary?
DOBRZE, SPRÓBUJ I ZOBACZ CO SIĘ STANIE.

Teraz jestem gotów
JEŻELI JESTEŚ GOTÓW NAPISZ POJEDYŃCZE SŁOWO "GOTÓW". W INNYM PRZYPADKU NIE UŻYWAJ TEGO SŁOWA.

Gotów
TERAZ BĘDĘ ZADAWAŁ PYTANIA DOTYCZĄCE TWOJEJ METODY SYNCHRONIZACJI. CZY TA METODA PRZEWIDUJE PRZEKAZYWANIE SYGNAŁÓW LUB PRZEDMIOTÓW MATERIALNYCH Z JEDNEGO MIEJSCA NA DRUGIE?

Oczywiście
JAKIEGO RODZAJU SYGNAŁY LUB PRZEDMIOTY PROPONUJESZ STOSOWAĆ?

Sygnały świetlne
ALE SYGNAŁY ŚWIETLNE NIE DOSTOSOWUJĄ SIĘ DO KRZYWIZNY ZIEMI, ZATEM TRZEBA ZAŁOŻYĆ UŻYCIĘ JAKIEGOS SYSTEMU WIEŻ PRZEKAŹNIKOWYCH. POZWÓL MI TERAZ ZAPYTAĆ CZY SYGNAŁ ŚWIETLNY BĘDZIE BIEGŁ WPROST MIĘDZY NOWYM JORKIEM I SAN FRANCISCO CZY OD LUB DO JAKIEGOS PUNKTU POŚREDNIEGO.

Odbije go od punktu pośredniego na księżycu.
CZY TWÓJ PUNKT POŚREDNI JEST RÓWNOODLEGŁY OD NOWEGO JORKU I SAN FRANCISCO?

•
•
•
•
•

takiego dialogu jest przytoczony w tabeli 8 (małe litery - wypowiedzi człowieka, duże - maszyny).

Mimo tak dużej różnorodności we wszystkich typach systemów CAI występuje jednak wspólny, ważny problem: przygotowanie tekstów lekcji, tzn. pytań, wyjaśnień, ćwiczeń itd. - tak, aby zawierały odpowiednią treść i aby ta treść była podana w odpowiedniej formie dydaktycznej. Nawet w przypadku uczenia programowania na ogół kto inny opracowywało ramy systemu konwersacyjnego, a kto inny lekcje, które system prezentuje. Konieczne jest więc stworzenie pomostu między programistą budującym system a nauczycielem odpowiedzialnym za podawany uczniowi materiał.

Pomostem tym są tzw. języki autorskie, w których nauczyciel - autor kursu nie będący na ogół programistą, może łatwo napisać program CAI. Przykładem bardzo prostego języka autorskiego może być PILOT (Programmed Inquiry Learning or Teaching) - omawiany w pracy [14]. Poniżej przytoczono jego podstawowe instrukcje:

- T:..... (Type) - Wypisz następujący (wpisany w miejscach kropek) tekst
- R:..... (Recognize) - Rozpoznaj następujący element w odpowiedzi
- R2:...,... Rozpoznaj jeden z dwóch następujących elementów w odpowiedzi (R3, R4 itd. oznaczają to samo dla większej liczby elementów)
- G:..... (Good) - Gdy odpowiedź dobra (rozpoznano w niej odpowiedni element) wypisz następujący tekst
- B:..... (Bad) - Gdy odpowiedź zła wypisz następujący tekst
- C:..... (Control) - Skocz do fragmentu programu oznaczonego następującą etykietą

Język PILOT jest prosty, więc łatwy do opanowania. Nie jest on jednak z powodu tej prostoty tak dobrym narzędziem do tworzenia lekcji, jak np. stosowany na maszynach IBM język COURSEWRITER, pozwalający prezentować informację dydaktyczną na różnych urządzeniach i z różnych urządzeń oczekiwać odpowiedzi, uwzględniać różne formy prawidłowych i nieprawidłowych odpowiedzi, prowadzić statystykę błędów studenta, wywoływać funkcje i podprogramy. COURSEWRITER ma około 60 instrukcji.

Języków autorskich opracowano już bardzo dużo, zestawienie ich i klasyfikacje znajdują się np. w pracach [5] i [9].

Dając pedagogom, specjalistom od geografii, chemii czy języków obcych tak wygodne narzędzie do tworzenia systemów uczących programiści mogą nie kłopotać się już problemami dydaktycznymi związanymi z daną dziedziną nauczania.

A jak jest z uczeniem programowania? Jest to dziedzina wyróżniona w dydaktyce komputerowej ze względu na bliskość "laboratorium", w którym studenci, po przyswojeniu sobie pewnej wiedzy teoretycznej mogą prowadzić ćwiczenia praktyczne (niesymulowane). Najbardziej naturalną metodą egzaminowania ucznia z programowania byłoby sprawdzenie przez system CAI poprawności zadanego uczniowi programu. Wiąże się to jednak z zagadnieniem równoważności programów - zagadnieniem niedopracowanym jeszcze nawet teoretycznie. Rozwikłanie tego problemu dla potrzeb CAI nie jest więc łatwe.

3.3. Współpraca z systemem operacyjnym

Przy współpracy z systemem operacyjnym stosowane są dialogi dwójki typy: takie, w których operator przekazuje informacje systemowi (np. podaje bieżącą datę, godzinę lub jakieś parametry przetwarzanych programów) oraz takie, w których operator otrzymuje od systemu potrzebne informacje zadając odpowiednie pytania. Dialogi pierwszego typu stosowane są powszechnie również w tradycyjnych systemach, nie będziemy więc poświęcać im tu specjalnej uwagi. Zajmiemy się problemami realizacji dialogów drugiego typu.

HELP-QAS

Przyjrzyjmy się systemowi HELP-QAS (Question Answering System - SDS 930 - University of California, Berkeley, 1970, [22]). Można go zaklasyfikować również do grupy zastosowań opisanej w pkt 3.2, ponieważ jest to aparat używany przy współpracy z systemem operacyjnym, który uczy jak posługiwać się tym systemem. Umieszczono go tutaj ze względu na metodę realizacji dostosowaną do obsługi systemu operacyjnego.

Pierwsza wersja systemu HELP powstała w 1966 r. i działała pod nadzorem Berkeley Time Sharing System - był to HELP-BTSS. Odpowiadał na pytania, jak posługiwać się systemem operacyjnym BTSS. Forma pytań była dowolna. Stosowano dość skomplikowaną analizę pytań, co powodowało długą reakcję systemu. Ze względu na charakter zastosowania czas reakcji systemu powinien w 75% nie przekraczać 1 sekundy. BTSS nie miał tej własności, nie mógł więc zdystansować typowych podręczników obsługi i zdecydowało to o modyfikacji systemu. Miała ona ponadto umożliwić zadawanie pytań dotyczących innych składowych systemu oprogramowania maszyny SDS 930 (np. języków programowania) oraz zapewnić łatwe rozszerzanie zakresu tematycznego.

Postulaty te uwzględniono w HELP-QAS, który jest właściwie aparatem do budowania programów typu "help" zróżnicowanych tematycznie.

Analizę pytań w QAS uproszczono, stosując zasadę wybierania słów kluczowych, jednak nieco odmienną niż we wspomnianym wyżej systemie ELIZA. Rozpatrzmy dialog z QAS przytoczony w tabeli 9. (Małe litery - wypowiedź człowieka, duże - maszyny).

Tabela 9. Przykład dialogu z systemem HELP-QAS

? what is a file?
A FILE IS A COLLECTION OF DATA
? how can I open a file?
USE BRS 15 TO OPEN A FILE FOR INPUT. USE BRS 16 TO OPEN A FILE FOR OUTPUT
? brs 15?
BRS 15 OPENS FILE FOR INPUT. A-CONTROL WORD, X-DUAL FILE NUMBER
? tell me about a dual file number?
A DUAL FILE NUMBER HAS THE COMMAND INPUT FILE IN THE LOWER 12 BITS AND THE COMMAND OUTPUT FILE IN THE TOP 12 BITS
⋮

Dla każdego pytania QAS tworzy listy słów kluczowych zawartych w pytaniu i wybiera odpowiedź przyporządkowaną najdłuższej z nich, np. dla pytania: "How can I open a file?" najdłuższą listą słów kluczowych jest [open, file], a dla pytania: "How can I open a file for input?" - [open, file, input]. Tym dwu listom słów kluczowych przyporządkowane są różne odpowiedzi. Ilustruje to schemat wybierania odpowiedzi zależnie od listy słów kluczowych zamieszczony w tabeli 10.

Tabela 10. Schemat wybierania odpowiedzi w systemie HELP-QAS

lista słów kluczowych	odpowiedź
[file]	M1: A FILE IS A COLLECTION OF DATA
[open, input, file]	M2: USE BRS 15 TO OPEN A FILE FOR INPUT
[open, output, file]	M3: USE BRS 16 TO OPEN A FILE FOR OUTPUT
[open, file]	M4: (M2).(M3)
[close, file]	M5: USE BRS 17 TO CLOSE A FILE
[brs 15]	M6: BRS 15 OPENS FILE FOR INPUT (M8)
[brs 16]	M7: BRS 16 OPENS FILE FOR OUTPUT (M8)
	M8: A-CONTROL WORD, X-DUAL FILE NUMBER
[dual, file, number]	M9: A DUAL FILE NUMBER HAS THE COMMAND INPUT FILE IN THE LOWER 12 BITS AND THE COMMAND OUTPUT FILE IN THE TOP 12 BITS

Gdy pytanie nie zawiera żadnego słowa kluczowego, QAS odpowie: "I DON'T KNOW" (nie wiem) . Gdy w pytaniu zawartych jest kilka list tej samej długości, odpowiedź będzie sumą odpowiedzi na każdą z list, np. pytanie: "How can one open and close a file?" zawiera listy [open, file] i [close, file], obie maksymalnej długości, ponieważ lista [open, close, file] nie jest przewidziana. Odpowiedź na to pytanie będzie więc sumą odpowiedzi M4 i M5, wypisanych w tabeli 10.

Przy opracowaniu pojedynczej odpowiedzi program HELP-QAS wykonuje następujące czynności typowe:

- standaryzuje pytanie, tzn. oddziela końcówki, tworzy bezkolizyjniki itp.,
- przeszukuje tablicę słów kluczowych,
- tworzy listy słów kluczowych,
- generuje odpowiedź na podstawie list słów kluczowych,
- notuje pytania, na które nie znalazł odpowiedzi.

W niektórych przypadkach, tzn. po napotkaniu specjalnych słów wykonuje dodatkowe czynności, takie jak:

- zakończenie pracy programu HELP,
- wywołanie innego programu HELP (o innym zakresie tematycznym),
- zastąpienie słowa jego synonimem.

Wszystkie te czynności nie są zależne od tematyki zadawanych pytań, czyli od bazy danych, na której opiera się system udzielający odpowiedzi. QAS zapewnia obsługę dowolnej bazy danych i ułatwia jej wprowadzanie do pamięci maszyny. Autor nowego, specjalistycznego programu typu "help" podaje tylko systemowi QAS swoją bazę danych tzn. tablicę słów kluczowych, listy słów kluczowych i związane z nimi odpowiedzi, synonimy i inne słowa wyróżnione oraz związane z nimi akcje, a system QAS już sam troszczy się o ich wykorzystanie. Na przykład następujący tekst może być fragmentem wejścia dla QAS:

SYNONIMS. INPUT READ, NUMBER NO., NUMBER NO, OUTPUT WRITE.

ANSWERS. [BRS 15]

BRS 15 OPENS FILE FOR INPUT

[FILE, CLOSE]

USE BRS 17 TO CLOSE A FILE

•
•
•

Tekst ten oznacza, że w utworzonej bazie danych słowo INPUT jest synonimem słowa READ, słowo NUMBER synonimem NO. itd. oraz, że liście słów kluczowych BRS 15 ma być przyporządkowana odpowiedź: BRS 15 OPENS FILE FOR INPUT, a liście [FILE, CLOSE] odpowiedź: USE BRS 17 TO CLOSE A FILE. Słowa SYNONIMS i ANSWERS należą do języka zrozumiałego dla QAS (jest to rodzaj języka autorskiego), pozostały tekst stanowi bazę danych związaną z tematyką budowanego systemu HELP.

Ogólnie - w dialogach z systemem operacyjnym podobnych do przytoczonego w tabeli 9, w których operator otrzymuje od systemu potrzebne informacje, jak również w dialogach, w których operator przekazuje informacje systemowi - istotny jest czas reakcji systemu i tempo przekazywania informacji. Aby ograniczyć czas systemu zajęty na analizę pytań operatora, stosuje się proste metody prowadzenia dialogu, często opiera się go na metodzie jądłospisowej. Informacje podawane systemowi przez operatora są na ogół w standardowej formie. Podawanie ich przypomina też często wypełnianie kwestionariusza.

3.4. Projektowanie wspomagane przez maszynę

Projektowanie wspomagane przez komputer (termin ang. Computer-Aided Design) jest jednym z tych typów zastosowań, w których konwersacyjny tryb współpracy człowieka z maszyną pozwala na najefektywniejsze wykorzystanie możliwości zespołu złożonego ze specjalisty w zakresie jakiejś dziedziny techniki oraz maszyny reprezentującej szybkość i bezbłądność wykonywania standardowych operacji. Wynika to przede wszystkim z iteracyjnego na ogół charakteru procesu projekto-

wania, w którym to procesie można wyróżnić następujące stadia:

- określenie modelu odpowiadającego realizowanemu projektowi, przez przyjęcie założeń dotyczących parametrów opisujących np. właściwości funkcjonalne, konstrukcyjne, eksploatacyjne itp.,
- wykonanie analizy tego modelu w różnych przekrojach tematycznych łącznie z porównywaniem otrzymanych charakterystyk z normami przemysłowymi obowiązującymi w danej branży albo z zamierzonymi celami projektu; jeśli w wyniku przeprowadzonej analizy projektant stwierdza zgodność przeanalizowanego modelu z przyjętymi założeniami, to proces projektowania uważa się za zakończony, inaczej następuje korekta,
- wprowadzenie zmian w modelu w celu zbliżenia go do pierwotnych założeń i przejście do ponownego wykonania analizy.

Z opisanego powyżej toku postępowania widać bezpośrednio, że systemy konwersacyjne są wygodniejszym aparatem wspomagającym projektanta niż odpowiednie fragmentaryczne programy przetwarzane w sposób tradycyjny, ponieważ tryb konwersacyjny pozwala użytkownikowi:

- poprawiać na bieżąco zauważone błędy we wprowadzonych rysunkach bądź instrukcjach,
- kontrolować na bieżąco kolejne etapy rozwiązywania problemu, a w razie potrzeby przerwać bieżąco realizowane operacje projektowe i zmienić pewne parametry,
- powziąć w punktach krytycznych subiektywne decyzje wpływające na wybór dalszego sposobu postępowania komputera.

Zwykle z opracowywanym projektem związane są rysunki. W zależności od przedmiotu projektu rysunki te mogą być obrazami projektowanych obiektów (np. modeli architektonicznych, karoserii samochodowych itp.) bądź mogą one przedstawiać projektowane obiekty w sposób symboliczny za pomocą schematów, na których występują oznaczenia ogólnie przyjęte w danej dziedzinie techniki (np. przy projektowaniu układów, elektronicznych, systemów optycznych itp.). Dlatego też między innymi praktyczny rozwój systemów projektowania wspomaganego przez komputer nastąpił dopiero po skonstruowaniu urządzeń wejścia i wyjścia umożli-

wiających bezpośrednią wymianę informacji graficznej między użytkownikiem a maszyną. Na marginesie warto wspomnieć, że właśnie z inicjatywy i na zlecenie zespołu pracowników opracowujących jeden z chronologicznie pierwszych systemów projektowania wspomaganego przez komputer¹, skonstruowano w firmie IBM graficzny monitor ekranowy będący prototypem powszechnie dzisiaj stosowanego monitora - IBM 2250.

W systemach, w których tworzone rysunki są schematami, projektantowi trzeba udostępnić zbiór symboli reprezentujących elementy składowe projektowanego obiektu oraz reguły łączenia tych symboli. W ścisłej zależności od przedmiotu projektowania pozostaje zarówno zbiór stosowanych symboli (postać graficzna tych symboli powinna być identyczna z postacią oznaczeń używanych tradycyjnie w danej dziedzinie), jak też programy realizujące konkretne obliczenia oraz programy analizujące zaprojektowany układ.

W każdym takim systemie projektant musi mieć możliwość wybierania podstawowych trybów współpracy z systemem, tj. trybu projektowania bądź trybu testowania. Podczas działania w trybie projektowania wykonuje się np. operacje:

- wybierania potrzebnego symbolu,
- umieszczania wybranego symbolu w żądanym miejscu tworzonego schematu bądź przesuwania go z jednego miejsca na drugie,
- łączenia liniami wskazanych symboli,
- usuwania ze schematu wskazanych symboli bądź linii.

Również i w trybie testowania można wypunktować operacje niezależne od konkretnego zastosowania, np. ustalanie wartości wejściowych czy też przejście do symulowania pracy układu.

Opisany dialog realizuje się zwykle metodą jaskłospisową a urządzeniem zapewniającym komunikację projektanta z systemem jest graficzny monitor ekranowy wyposażony w pióro świetlne. Ekran jest zwykle dzielony na dwie części: pole funkcyjne i pole robocze. W polu

¹ Dotyczy to zespołu pracowników firmy General Motors, którzy w latach 1960-65 opracowali system DAC-I (Design Augmented by Computer) przeznaczony dla projektanta karoserii samochodowych [8].

funkcyjnym są wyświetlane przez system teksty bądź symbole będące odpowiednikami dopuszczalnych operacji a w polu roboczym jest sukcesywnie tworzony schemat.

Spośród opracowanych do tej pory systemów prawie każdy był związany z konkretnym zastosowaniem. Przykładem takiego systemu jest program GAZELLE stosowany w firmie IBM przez konstruktorów maszyn matematycznych przy projektowaniu sieci logicznych. Zasady współpracy z tym programem są szczegółowo omówione w książce J. Martina [14], rozdz.15.

W ostatnich pracach z tego zakresu realizatorzy idą w kierunku opracowania szkieletu systemu projektowania, uniwersalnego dla kilku dziedzin. W ramach tego szkieletu są opracowywane następnie programy rozwiązujące problemy swoiste dla danej dziedziny. Przykładem takiego systemu jest np. system SNAP [16], za pomocą którego można analizować układy elektryczne, symulować sieci logiczne oraz redukować wyrażenia boolewskie.

W systemach operujących obrazami projektowanych obiektów na pierwszy plan wysuwają się problemy związane z przetwarzaniem informacji graficznej. Są to problemy dotyczące sporządzania obrazu modelu trójwymiarowego, szanowania niewidocznych krawędzi a także przeszerzania reguł perspektywy. Takie systemy realizuje się z reguły dla wysokiej klasy graficznych monitorów ekranowych dających możliwości wykonywania obrotów figur wokół wskazanej osi, rozsuwania lub łączenia elementów obrazu itp. (patrz [6] i [21]).

Wśród tych systemów są takie, które tworzą obraz modelu na podstawie wprowadzanych przez projektanta danych analitycznych bądź liczbowych oraz takie, które interpretują wprowadzane rysunki.

HUNCH

Przedstawimy teraz system, który próbuje interpretować wprowadzane informacje graficzne jako rysunek o znanej tematyce.

Przykładem takiego systemu jest eksperymentalny system HUNCH opracowywany od

1970 r. na Wydziale Architektury Massachusetts Institute of Technology [18]. HUNCH rozpoznaje odrębne ry-

sunki, o tematyce architektonicznej, wykonywane na urządzeniu graficznym Sylvania Tablet¹, a po "zrozumieniu" rysunku tworzy obraz trójwymiarowego obiektu przedstawionego na tym szkicu. Rysunki wprowadzane za pośrednictwem Sylvania Tablet, z założenia, przedstawiają rzuty prostokątne bądź perspektywę jakiegos obiektu.

Przy identyfikowaniu narysowanych figur system odwołuje się do pojęć obowiązujących w architekturze. Intencje graficzne użytkownika rozstrzyga się między innymi także na podstawie szybkości rysowania poszczególnych fragmentów (np. kwadrat rysowany szybko ma zwykle zaokrąglone rogi) oraz nacisku rylca na pulpit².

W opisywanym systemie można wyróżnić dwa typy konwersacji projektanta z systemem. W pierwszym z nich system zadaje użytkownikowi pytania (w języku angielskim) w celu wyjaśnienia nierozstrzygniętych problemów bądź otrzymania potwierdzenia prawidłowości interpretacji. W drugim - przeprowadzany dialog jest ściśle związany z pracami projektowymi i dotyczy bądź redagowania rysunku, bądź analizowania gotowego projektu. Ten dialog jest prowadzony za pomocą pióra świetlnego, klawiszy funkcyjnych lub alfanumerycznych i nie ma charakteru dialogu naturalnego. Wspomniane dwa typy konwersacji przeplatają się nawet w obrębie poszczególnych faz, na które dzieli się opracowywanie wprowadzonego rysunku. Wyróżniamy cztery podstawowe fazy:

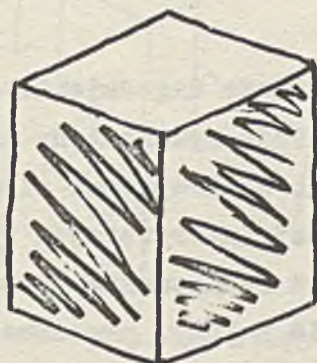
- wyrównanie rysunku,
- redagowanie rysunku,

¹ Sylvania Tablet jest wizualnym urządzeniem wejścia typu dynamicznego tzn., że informacje o rysunku są wprowadzane do pamięci maszyny na bieżąco podczas tworzenia tego rysunku. Zasada działania Sylvania Tablet jest analogiczna do zasady działania Grafaconu opisanego w literaturze polskiej w [26] i [28]. Różnica polega na innej technicznie zasadzie produkowania wartości współrzędnych punktu rysunku. W obu przypadkach podczas rysowania jest generowana zawsze stała liczba par współrzędnych (X,Y) w jednostce czasu, co pozwala na obliczanie prędkości tworzenia poszczególnych partii rysunku.

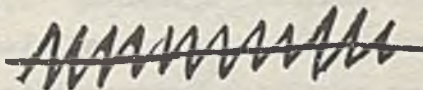
² W MIT zmodyfikowano Sylvania Tablet wyposażając rylce dodatkowo w czujnik mierzący siłę nacisku, rozróżniający trzy stany nacisku. W ten sposób każdy opisywany punkt jest reprezentowany przez 3 współrzędne (X,Y,Z), z których Z określa nacisk.

- utworzenie obrazu trójwymiarowego modelu,
- analizowanie modelu.

Podczas wyrównywania rysunku usuwa się zniekształcenia spowodowane odręcznym wykonaniem szkicu a więc drganiami nadgarstka, nierównościami papieru bądź też szybkością rysowania. W tym celu, z rysunku wyodrębnia się najpierw wszystkie elementy nie będące integralnymi składowymi rzutów, takie jak: symbole opisujące rzuty (np. litery, strzałki), zygzaki zastosowane do wycieniowania płaszczyzny (np. rys. 3a) czy też zygzaki oznaczające wymazanie uprzednio poprowadzonej linii (np. rys. 3b). Znaczenie zygzaków ustala się określając położenie środka ciężkości zygzaków oraz kolejność umieszczenia ich na rysunku. Jeśli ten środek ciężkości leży na uprzednio poprowadzonej linii, to zygzaki oznaczają wymazanie jej z rysunku.



a)



b)

Rys. 3. Stosowanie zygzaków

Dalsze operacje odnoszą się już do "właściwego" rysunku. Analizując rysunek - a ściślej mówiąc - jego opis, wykonuje się kolejno:

- prostowanie linii, tj. zastąpienie linii, wszędzie gdzie jest to uzasadnione, odcinkiem prostej,

- domykanie linii, tj. doprowadzenie dwóch linii, wszędzie gdzie jest to uzasadnione do wspólnego punktu końcowego,
- wyznaczanie przecięć, tj. ustalenie wszystkich odcinków powstających w wyniku przecięcia danej linii innymi liniami,
- korygowanie położenia linii, tj. uwzględnienie poprawek wynikających z powziętej decyzji o tym, że:
 - linia jest pozioma bądź pionowa,
 - dwie linie są równoległe bądź prostopadłe,
 - linia jest ciągła,
 - linia jest podkreśleniem tj. uwydatnieniem uprzednio narysowanej linii.

W czasie realizowania powyżej opisanych operacji jest tworzony opis "wyrównanego" rysunku. Ma on formę bardziej zwartą niż opis rysunku pierwotnego, powstający podczas rysowania, ponieważ np. każda linia będąca odcinkiem prostej jest w nim reprezentowana przez dwa punkty końcowe.

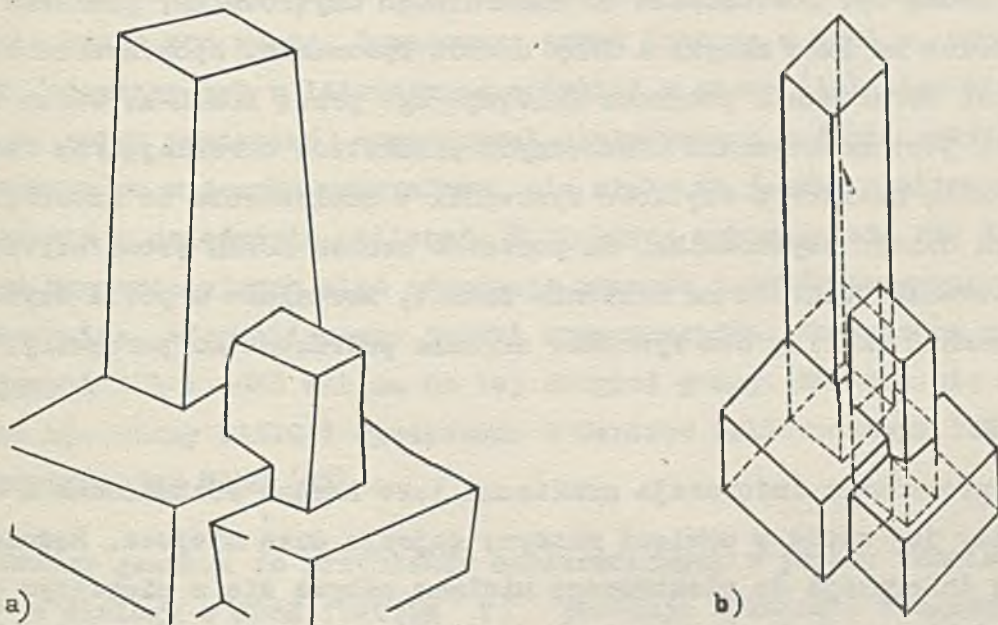
W fazie redagowania rysunku następuje uzgodnienie ostatecznej wersji szkicu. "Wyrównany" rysunek zostaje wyświetlony na ekranie monitora graficznego a projektant może go zaakceptować albo wprowadzić zmiany powodujące poprawienie maszynowej interpretacji rysunku bądź modyfikację rysunku pierwotnego. Zakres redagowania jest ograniczony¹. Zezwala się dodawać lub odejmować linie i kąty, przesuwać segmenty obrazów oraz wywoływać poszczególne programy korygujące położenie linii. Redagowanie odbywa się w trybie konwersacyjnym.

Po uzgodnieniu rysunku, rozpoczyna się faza tworzenia obrazu trójwymiarowego modelu projektowanego obiektu. W tym celu trzeba najpierw określić jaki rzut przedstawia uzgodniony rysunek, tj. rzut perspektywiczny czy prostokątny (a jeśli prostokątny - to czy przekrój pionowy, poziomy czy też rzut aksonometryczny). Za pomocą odpowiednich programów systemu HUNCH identyfikuje się rzut perspektywiczny (ustalając np. pozycję obserwatora) oraz aksonometryczny (perspektywa dla obser-

¹ Przytoczony tutaj opis odpowiada jednej z pierwszych wersji systemu HUNCH.

watora w nieskończoności). W przypadkach pozostałych rodzaj przekroju określa projektant odpowiadając na pytanie zadane przez system. Następnie przystępuje się do tworzenia obrazu trójwymiarowego modelu.

W prezentowanej wersji systemu HUNCH taki obraz jest budowany tylko w przypadku, gdy uzgodniony rysunek przedstawiał perspektywę. Na obrazie zaznacza się wszystkie niewidoczne linie odpowiadające przecięciom płaszczyzn reprezentowanej bryły oraz wprowadza się dodatkowe symbole dla oznaczenia różnych przekrojów tej bryły. Rys. 4 a ilustruje formę szkicu uzgodnioną w fazie redagowania, a rys. 4 b - utworzony



Rys. 4. Etapy opracowywania szkicu

obraz modelu trójwymiarowego wraz z zaznaczonymi ukrytymi krawędziami oraz niektórymi dodatkowymi przekrojami (w rzeczywistości jest ich dużo więcej, ale zaznaczenie wszystkich uczyniłoby rysunek nieczytelnym).

Prace nad sporządzaniem obrazu trójwymiarowego modelu w przypadku, gdy rysunek przedstawia rzuty prostokątne obiektu, odłożono na później. Autorzy uważali, że rozwiązanie tego problemu będzie skomplikowane, ponieważ odrębnie szkicowane rzuty mogą opisywać geometryczne "niemożliwości" lub tworzenie obrazu trójwymiarowego modelu może prowadzić do alternatywnych rozwiązań.

Ostatnią fazą opracowywania rysunku jest wyświetlenie utworzonego obrazu trójwymiarowego modelu i udostępnienie go projektantowi dla przeprowadzenia oceny ostatecznego wyglądu modelu. W trybie konwersacyjnym architekt może np. wykonywać obroty przedstawionego obiektu w celu obejrzenia go z różnych stron, bądź też zmieniać proporcje bryły ale nie może zmieniać jej topologii.

Istotną informacją zawartą w opisie tego systemu są uwagi sformułowane na podstawie doświadczeń nabytych przy uruchamianiu systemu HUNCH. Okazało się na przykład, że zdolności rozpoznawania elementów szkicu muszą być dostosowane do konkretnego użytkownika, ponieważ każdy człowiek ma inne nawyki i inny sposób rysowania. Spowodowało to konieczność opracowania programu inicjującego pracę HUNCH-a. Celem tego programu jest zanotowanie niezbędnych parametrów określających indywidualną siłę nacisku i szybkość rysowania w odniesieniu do intencji graficznych danego użytkownika. Na początku seansu HUNCH prosi użytkownika o narysowanie kwadratu na Sylvania Tablet, normalnie a potem szybko i po dokonaniu analizy obu rysunków określa potrzebne mu parametry.

Wprowadzana informacja graficzna jest bardzo rozbudowana i odpowiadający jej zapis w pamięci maszyny zajmuje dużo miejsca. Redukowanie tej informacji do niezbędnego minimum odbywa się w pierwszym etapie projektowania. Na ogół więc ze względów ekonomicznych rozdziela się proces projektowania na etapy wykonując wstępną analizę rysunku za pośrednictwem minikomputera, do którego podłącza się graficzne urządzenie a zredukowaną informację przekazuje się do właściwego systemu liczącego realizującego dalsze prace projektowe. W systemie HUNCH prace wstępne (tj. aż do uzgodnienia postaci rysunku) wykonuje się za pomocą minikomputera Interdata, a prace pozostałe za pomocą komputera GE 645 wyposażonego w system MULTICS.

3.5. Wykonywanie obliczeń

W tej dziedzinie zastosowań systemów konwersacyjnych, problemów jest bez porównania mniej niż np. w dziedzinie dydaktyki komputerowej. Wynika to z faktu, że nauczanie "nie leży w naturze maszyny matematy-

cznej" i trzeba ją do tego specjalnie zaadaptować, podczas gdy wykonywanie obliczeń jest jej właściwą domeną. W związku z tym istniały już od dawna metody definiowania dla maszyny problemów obliczeniowych. Służyły do tego celu tradycyjne języki programowania przeznaczone do wykonywania obliczeń.

Pojawienie się możliwości obliczania w trybie konwersacyjnym wymagało określenia języka dla tego typu dialogu. Poniżej zajmiemy się właśnie zagadnieniem języków do obliczeń konwersacyjnych.

Formułowanie problemów obliczeniowych w języku naturalnym również było brane pod uwagę. Opracowany przez Bobrowa w 1964 r. program STUDENT (omawiany był w literaturze polskiej w pracy [19] więc nie będziemy tu opisu powtarzać) rozwiązywał arytmetyczne zadania tekstowe przekazywane mu w języku naturalnym, nie miało to jednak praktycznego zastosowania w dziedzinie obliczeń. W praktyce wykorzystuje się do obliczeń konwersacyjnych albo adaptacje znanych języków programowania, albo specjalne, sformalizowane języki konwersacyjne. Wspominane w pkt 2 języki APL i JOSS należą do tej drugiej grupy. Zalicza się do niej też np. znany BASIC i opracowany w Centrum Obliczeniowym PAN język konwersacyjny MARO [17] .

Wróćmy jeszcze do przykładów zamieszczonych w pkt 2. Zdanie "do part 1" z dialogu w JOSS (tabela 1) powoduje wykonanie fragmentu programu wcześniej wprowadzonego do maszyny (zdania oznaczone numerami 1.1, 1.2, 1.3). Podobnie jest w języku APL. Operacja " \leftarrow " (np. $B \leftarrow 2(3) < (3,1,5)$) powoduje przyporządkowanie zmiennej występującej po lewej stronie strzałki fragmentu obliczenia zapisanego po prawej stronie, a dopiero odwołanie się do tej zmiennej (w przykładzie instrukcja X,B) powoduje wykonanie tego fragmentu obliczenia i wypisanie wyników.

Podział na dwa "tryby" współpracy tj. przygotowanie programu i wykonanie programu, jest charakterystyczny dla tego typu języków. Pewne rodzaje błędów mogą być wykrywane na etapie przygotowania programu, inne dopiero przy jego wykonywaniu. W obu przypadkach możliwa jest natychmiastowa korekta programu. W załączonym przykładzie dialogu w JOSS przygotowany program był wykonywany najpierw dla pojedynczych danych,

a dopiero po usunięciu błędów zamknięty w pętli w celu wykonania pełnego obliczenia.

Adaptacja znanego języka programowania do celów konwersacyjnych polega zwykle na wykreśleniu z niego instrukcji nieprzydatnych w trybie konwersacyjnym i dodaniu instrukcji sterujących, przełączających z reżimu przygotowania programu na reżim wykonania i odwrotnie, oraz instrukcji ułatwiających uruchomienie programu. Na tej zasadzie budowane są konwersacyjne wersje PL1 (np. OPS, RUSH, CALL 360 PL1) i konwersacyjne wersje FORTRAN-u, np. QUIKTRAN.

3.6. Wyszukiwanie informacji

Powyżej była już mowa o kłopotach pojawiających się przy tworzeniu ogromnych banków informacji, ze względu na liczbę powiązań między pojęciami - potrzebnych do właściwego interpretowania wyszukiwanej informacji. Mając na uwadze fakt, że istniejące systemy wyszukiwania informacji nie mogą być jeszcze w pełni ogólne i nie wchodząc w szczegóły organizacji banków danych, spójrzmy na systemy przechowujące i wyszukujące dokumenty od strony dialogu, który prowadzi użytkownik z systemem podczas wybierania potrzebnego dokumentu. Wypowiedzi użytkownika w tym wypadku są trojakiego typu:

- sformalizowane zdania określające parametry potrzebnego dokumentu,
- niesformalizowane pytania lub polecenia dla systemu, w wyniku których system przekazuje w odpowiedzi potrzebną informację,
- informacje uzupełniające, ułatwiające systemowi wyszukanie odpowiedniego dokumentu lub udzielenie odpowiedzi.

Poniżej zilustrujemy przykładami wspomniane 3 typy wypowiedzi.

Zacznijmy od przykładu dosyć prostego systemu wyszukiwania informacji, w którym:

- zakres tematyczny przechowywanych informacji jest bardzo ograniczony, są to bowiem informacje o krajowych rejsach lotniczych (każdy dokument zawiera: numer rejsu, punkt startu, punkt docelowy, czas odlotu, czas lotu, typ samolotu, nazwisko kapitana),

- typ zadawanych pytań jest znany - na ogół potrzebne są informacje o locie lub lotach z X do Y, gdzie X i Y są odpowiednio punktem startowym i docelowym,
- zakłada się sformalizowany język dialogu.

Realizacja takiego systemu (oparta na systemie GPAS Query System [7]) mogłaby przebiegać według opisanego niżej schematu.

- Bank danych porządkuje się hierarchicznie, np. względem punktów docelowych lotów (poziom I), następnie punktów startowych (poziom II), następnie numerów rejsów (poziom III) - rys. 5. Takie uporządkowanie danych ułatwia wyszukanie odpowiedzi na typowe zapytanie np. "Wypisz rejsy z Poznania do Krakowa". Pytanie "Kiedy kapitan Guzik będzie w Krakowie" byłoby przy takim układzie hierarchicznym trudniejsze dla systemu, a łatwe gdyby hierarchię dokumentów oprzeć np. na nazwiskach kapitanów.

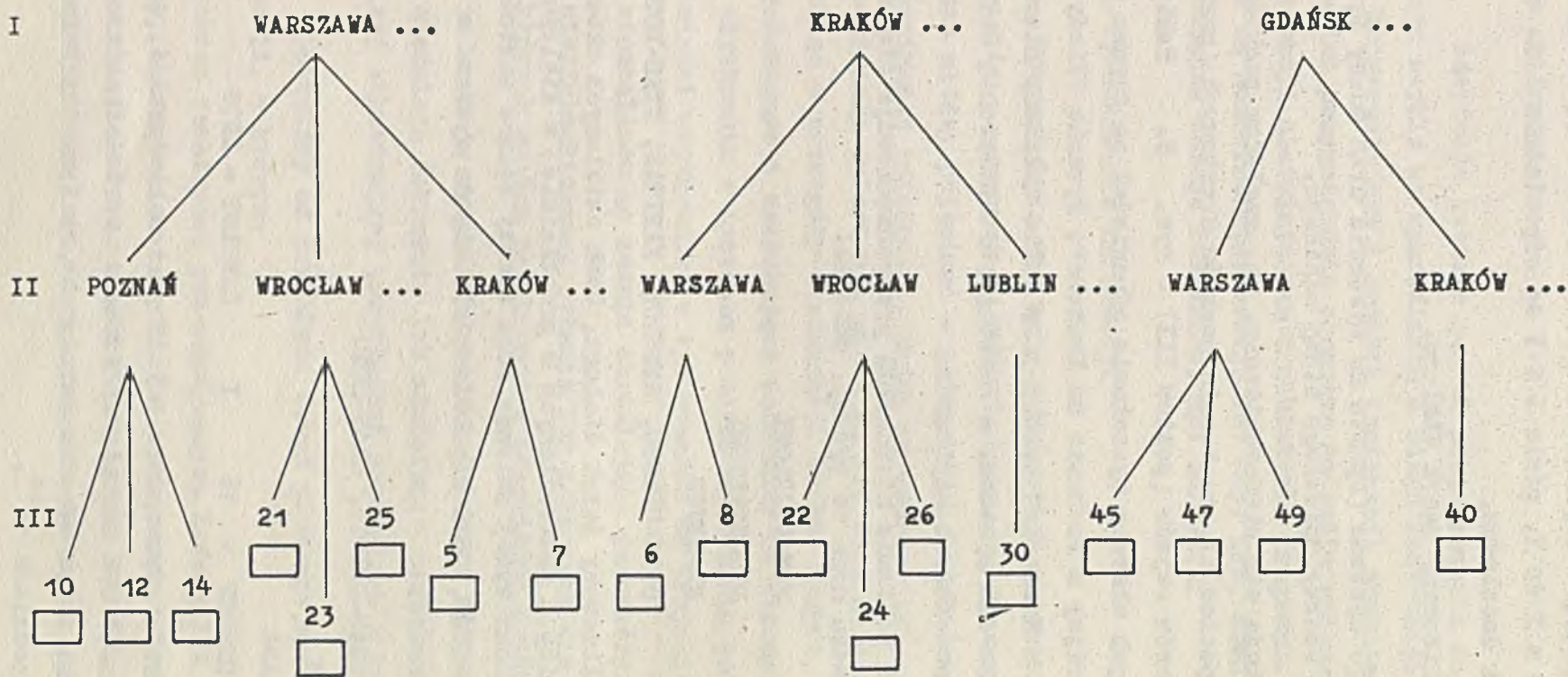
- Elementom na każdym poziomie hierarchii przyporządkowuje się zmienne. W naszym przypadku mogą to być np. zmienne:

na I poziomie	PKT-DOCELOWY
na II poziomie	PKT-STARTOWY
na III poziomie	NR-REJSU
na IV poziomie	CZAS-ODLOTU, SAMOLOT, KAPITAN, CZAS-LOTU

- Określa się sformalizowany język dialogu. Obok ustalonych słów (np. WYDRUKUJ, GDY itp. - podkreślonych w przykładzie) w dialogu występują też wymienione wyżej zmienne oraz teksty będące wartościami zmiennych np. KRAKÓW, IL18 itp. Odwołanie się do systemu w tym języku może mieć postać:

<u>WYDRUKUJ</u>	CZAS-ODLOTU, CZAS-LOTU
<u>DO</u>	KRAKOW
<u>Z</u>	WARSZAWA
<u>GDY</u>	CZAS-ODLOTU > 16 <u>I</u> SAMOLOT = IL18

Słowa DO i Z określają odpowiednio I i II poziom hierarchii, za słowem GDY mogą występować warunki nałożone na wartości zmiennych z dowolnego poziomu, w tym wypadku zależy nam na tym, aby samolot IL18 startował po godzinie 16⁰⁰.



Rys. 5. Hierarchiczne uporządkowanie danych

Taki język dialogu (szerzej opisany w [7]) ułatwia dostęp do informacji na określonych poziomach hierarchii. Rola systemu sprowadza się do zbadania wartości podanych zmiennych i nałożonych na nie warunków i wykonaniu czynności określonej przez ustalone słowa (np. wydrukowanie lub wyświetlenie wartości zmiennych).

SYNTHEX

Aparat wyszukiwania informacji związanej z pytaniem o dowolnej formie, w systemie o dużym zakresie tematycznym jest znacznie bardziej skomplikowany. Przyjrzyjmy się mu na przykładzie systemu SYNTHEX [25], który udziela odpowiedzi na pytania z zakresu małej encyklopedii (Golden Book Encyclopedia). Kolejnymi etapami realizacji tego systemu były opracowywane, począwszy od 1963 r., systemy: Protosynthex I, Protosynthex II i Protosynthex III.

Dane dla systemu SYNTHEX uporządkowane są jednopoziomowo. Każde zdanie encyklopedii stanowi oddzielny dokument i ma swój adres zwany liczbą TCPZ (określającą tom, część, paragraf i zdanie).

Działanie systemu można rozbić na następujące etapy:

- wybranie z pytania słów kluczowych na podstawie słownika, po uprzednim zestandaryzowaniu słów występujących w pytaniu (tzn. oddzieleniu końcówek, utworzeniu bezokoliczników itd.), np. w pytaniu "W którym stanie znajduje się Wielki Kanion?" słowami kluczowymi są: stan, wielki, kanion;
- utożsamienie słów o zbliżonym znaczeniu np. rządzić, rząd, rządzenie;
- określenie dla każdego słowa kluczowego zbioru liczb TCPZ, czyli adresów zdań, w których występuje dane słowo kluczowe lub jego synonim lub słowo pokrewne; SYNTHEX robi to na podstawie słownika synonimów i słownika słów pokrewnych; np. przy pytaniu "Które zwierzęta żyją dłużej niż człowiek?" będzie zastosowany następujący fragment słownika słów pokrewnych:

zwierzęta - ssaki, ptaki, gady ...

żyć
dłużej } - wiek, lata, starszy

człowiek - osoba, ludzie, kobieta, mężczyzna.....;

jeśliby zatem w encyklopedii występowało np. zdanie: "średni wiek kobiet wynosi 60 lat" liczba TCPZ tego zdania również będzie wzięta pod uwagę;

- wybranie przecięcia, tzn. najmniejszej części tekstu, w której występuje najwięcej wyróżnionych słów, np. w związku z pytaniem "Czym żywią się robaki?" mogą być wybrane zdania:

"robaki żywią się trawą",

"trawa jest zjadana przez robaki",

"ptaki żywią się robakami",

a w związku z pytaniem "W którym stanie znajduje się Wielki Kanion?" zdania:

"zdumiewający wielki kanion znajduje się w stanie arizona",

"od niego stan ten otrzymał nazwę stan wielkiego kanionu",

"wielki dom obok kanionu jest w bardzo złym stanie";

- przeanalizowanie, które z wybranych zdań mają istotny związek z pytaniem i nadają się do sformułowania odpowiedzi; np. zdanie "ptaki żywią się robakami" wybrane przy pytaniu "Czym żywią się robaki?" będzie odrzucone już na etapie analizy syntaktycznej, w wyniku której będzie wiadomo, co w tym zdaniu jest podmiotem a co przedmiotem związanym z czynnością "żywić się". Natomiast do odrzucenia zdania "wielki dom obok kanionu jest w bardzo złym stanie" wybranego przy pytaniu "W którym stanie znajduje się Wielki Kanion?" potrzebny jest już aparat analizy semantycznej rozpoznający właściwe znaczenie słowa "stan", podobny jak w opisanym w pkt 2.3 systemie MUSE.

Zajmijmy się teraz systemami wyszukiwania informacji, w których konwersacja z człowiekiem jest narzędziem ułatwiającym proces wybierania właściwego dokumentu. Zasada wyszukiwania dokumentów jest w nich taka, jak w systemie opisanym bezpośrednio powyżej, prześledźmy więc

tylko etapy, na których wykorzystanie informacji uzupełniających podawanych przez użytkownika systemu w trybie konwersacyjnym ma istotne znaczenie.

- Wybranie słów kluczowych poprzedzone jest zwykle standaryzacją (tzn. oddzieleniem końcówek, zamianą na bezokoliczniki i liczby pojedyncze itp.), zlepianiem synonimów, przypisaniem słowom tzw. "wag". Synonimy mogą być zlepione na podstawie słownika, ale również metodą statystyczną - jeżeli przeglądając dużą liczbę dokumentów system napotka dwa różne słowa powtarzające się w tych samych kontekstach, to uzna je za synonimy. "Waga" słowa jest to przypisana mu liczba odwrotnie proporcjonalna do częstości występowania tego słowa w dużych tekstach; w ten sposób słowa rzadkie mają duże "wagi" i są w pierwszej kolejności wybierane jako słowa kluczowe. Jeżeli więc wybieranie słów kluczowych opiera się w systemie na metodach statystycznych, które mogą być w niektórych przypadkach zawodne - system przekazuje wybrany zbiór słów do skorygowania człowiekowi i dopiero po odpowiedniej korekcie przystępuje do dalszego działania.
- Pomocą przy wyszukiwaniu dokumentów związanych z danym zbiorem słów kluczowych jest tzw. tworzenie fraz, tj. często spotykanych i mających określone znaczenie połączeń słów np. równania różniczkowe, pierwiastek chemiczny. Połączenie słów, w obrębie zbioru słów kluczowych, w odpowiednie frazy też może być oparte na metodzie statystycznej i również wymaga potwierdzenia w drodze konwersacji z użytkownikiem.
- Jak widzieliśmy poprzednio, między wybranymi dokumentami pojawiają się często takie, które wprawdzie wiążą się z tymi samymi słowami kluczowymi, ale nie interesują użytkownika. Ingerencja człowieka może znacznie przyspieszyć proces wyszukania właściwego dokumentu, jeżeli spośród przedstawionych przez system opisów dokumentów wybierze tylko, jego zdaniem, interesujące. W ten sposób np. dokument wybrany na hasło: "fauna lasów tropikalnych" i zatytułowany: "Malowidła na szkle przedstawiające zwierzęta lasów tropikalnych" może być wydrukowany lub nie, zależnie od zainteresowań użytkownika współpracującego z systemem.

4. Podsumowanie

Zapoczątkowany w końcu lat pięćdziesiątych proces przystosowywania komputerów do komunikacji z człowiekiem opartej na metodach naturalnych dla człowieka nie jest jeszcze zakończony, ale można już wypunktować istotne osiągnięcia w tej dziedzinie.

- Opracowano wiele metod przekazywania informacji między człowiekiem i maszyną w formie wygodnej dla człowieka, jak np. komunikacja w języku zbliżonym do naturalnego, komunikacja za pomocą głosu, przekazywanie rysunków i innej informacji graficznej.
- Systemy oparte na konwersacyjnych metodach współpracy działają już niemal w każdej dziedzinie tradycyjnych zastosowań komputerów. Pomogły one rozszerzyć zakres tych zastosowań i zwiększyć krąg użytkowników.
- Powstały nowe dziedziny zastosowań komputerów - (np. nauczanie wspomagane przez maszynę) nieznane przy tradycyjnych metodach współpracy z maszyną.
- Powstały narzędzia ułatwiające opracowywanie systemów konwersacyjnych, np. specjalne języki programowania.

Wiele istotnych zagadnień z dziedziny systemów konwersacyjnych pozostało jeszcze do rozwiązania:

- trzeba udoskonalić bazę techniczną dla systemów konwersacyjnych wprowadzając na rynek tanie i niezawodne urządzenia graficzne i foniczne;
- trzeba dopracować metody rozpoznawania mowy przekazywanej komputerowi przez urządzenia foniczne;
- proces "rozumienia" przez maszynę wypowiedzi w języku naturalnym wymaga jeszcze znacznych udoskonalień;
- konieczne jest opracowanie metod takiego zapisu dużych zbiorów informacji tekstowej i graficznej w pamięci maszyny, który zapewniłby właściwe powiązanie znaczeniowe między występującymi pojęciami i łatwy dostęp do każdej informacji;

- należy, na koniec, znacznie rozpowszechnić wykorzystanie systemów konwersacyjnych. W wielu dziedzinach stosuje się je dopiero eksperymentalnie lub na małą skalę.

Literatura

- [1] BECKER H.B.: Programming self - taught. Computers and Automation 1967, t.16, nr 10, s.30.
- [2] BIEŃ J.S., ŁAKUSZEWICZ W., SZPAKOWICZ S.: Wprowadzenie do systemu MARYSIA. W: Sprawozdania IMM i ZON UW, nr 39. Warszawa 1973, UW.
- [3] BIEŃ J.S.: Założenia polskiego systemu konwersacyjnego "Marysia". W: Zastosowanie maszyn matematycznych do badań nad językiem naturalnym. Warszawa 1973, UW.
- [4] BOBROW D.G., RAPHAEL B.: New Programming Languages for AI Research. Third International Joint Conference on Artificial Intelligence. Stanford Univ. Stanford, California 1973.
- [5] BRAHAN J.W.: Language Standards for Computer-Aided Learning. Int. J. Man - Machine Studies 1973, t.5, nr 3, s.337.
- [6] BUCZKOWSKA T.: Stan obecny i kierunki rozwojowe graficznych monitorów ekranowych. ETO Nowości 1974, nr 4.
- [7] CHAI D.T.: An Information Retrieval System Using Keyword Dialog. Inform. Stor. Retr. 1973, t. 9, nr 7, s.373.
- [8] DEVER G.S., HARGREAVES B., WALKER D.M.: The DAC-I System. Datamation 1966, t.12, nr 6, s. 37.
- [9] EMPACHER A.B.: Nowy aspekt dydaktyki komputerowej: D-języki. Maszyny Matematyczne 1970, t. 6, nr 7-8, s. 7.
- [10] Fall Joint Computer Conference, Proceedings 1972, t.41. AFIPS Press Montvale. Sesja na temat: Programming Languages for Artificial Intelligence.
- [11] GRISWOLD R.E., POAGE J.F., POLONSKY I.P.: The SNOBOL 4 Programming Language. New Jersey 1971, Prentice-Hall, Inc. Englewood Cliffs.
- [12] KĘDZIOR Z.: Urządzenia rozpoznające znaki alfanumeryczne zapisane na nośniku papierowym i filmowym. ETO Nowości 1974, nr 4.

- [13] KONCEWICZ J.: Języki do przetwarzania struktur listowych. Maszyny Matematyczne 1970, t. 6, nr 5, s. 3.
- [14] MARTIN J.: Design of Man - Computer Dialogues. New Jersey 1973, Prentice Hall, Inc. Englewood Cliffs.
- [15] McCALLA G.J., SAMPSON J.R.: MUSE: a Model to Understand Simple English. Comm.ACM 1972, t. 15, nr 1, s. 29.
- [16] MORSE S.P.: Interactive Graphic System for Modeling Symbolic Networks. Computer Design 1973, t. 12, nr 2, s. 75.
- [17] MUCHLADO-MAROŃSKA B.: Język konwersacyjny MARO. Warszawa 1973, PWN.
- [18] NEGROPONTE N., GROSSIER L.B., TAGGART J.: HUNCH: An Experiment in Sketch Recognition. W: Symposium über Computer Graphics. Berlin 1971.
- [19] NIEVERGELT J., FARRAR J.C.: Co maszyny mogą, a czego nie mogą robić ?. Informatyka 1971, t. 7, nr 10, s. 17.
- [20] PATRYN R.: Akustyczne urządzenia wyjściowe. ETO Nowości 1974, nr 4.
- [21] PRINCE M.D.: Interactive Graphics for Computer-Aided Design. Addison 1971, Wesley Publishing Company, Inc.
- [22] ROBERTS R.: Help - A Question Answering System. Proc. Fall Joint Comp. Conf. 1970, t. 37.
- [23] SAWICKI E.: Automatyczne rozpoznawanie mowy. ETO Nowości 1974, nr 4.
- [24] SHAW J.C.: JOSS: Experience with an Experimental Computing Service for Users at Remote Consoles. W: Conversational Computers. New York 1968, John Wiley and Sons, Inc.
- [25] SIMMONS R.F.: SYNTEX. W: Conversational Computers. New York 1968, John Wiley and Sons, Inc.
- [26] STOLARSKI M.: Nowoczesna komunikacja graficzna człowieka z maszyną cyfrową. W: Problemy Przetwarzania Informacji. Warszawa 1970, WNT.
- [27] TAYLOR E.F.: The ELIZA program: conversational tutorial. IEEE International Convention Record 1967, część 10, s. 8.
- [28] TURSKI W.M.: Podstawy użytkowania maszyn cyfrowych, Warszawa 1968, PWN.
- [29] TURSKI W.M.: Struktury danych. Warszawa 1971, WNT.

- [30] WALIGÓRSKI S.: Problemy komunikacji człowiek - komputer w języku naturalnym. W: Zastosowanie maszyn matematycznych do badań nad językiem naturalnym. Warszawa 1973, UW.
- [31] WEIZENBAUM J.: ELIZA - A Computer Program for the Study of Natural Language Communication Between Man and Machine. Comm. ACM 1966, t. 9, nr 1, s. 36.
- [32] WEIZENBAUM J.: Contextual Understanding by Computers. Comm. ACM 1967, t. 10, nr 8, s. 474.
- [33] WINOGRAD T.: Research in Natural Language. Computer 1973, t. 6, nr 5, s. 25.

dr inż. Teresa BUCZKOWSKA
Instytut Maszyn Matematycznych

681.327.13:621.385.832

STAN OBECNY I KIERUNKI ROZWOJOWE GRAFICZNYCH MONITORÓW EKRANOWYCH

1. Wstęp

Graficzne monitory ekranowe są obok monitorów alfanumerycznych podstawowym urządzeniem wejściowo-wyjściowym w systemach informacyjnych pracujących na bieżąco (on-line), uwarunkowanych czasowo (real-time) oraz we wszystkich systemach, w których wymagana jest szybka, efektywna wymiana informacji między użytkownikiem maszyny cyfrowej i maszyną. Stosuje się je np. w systemach projektowania wspomaganego przez maszynę (Computer Aided Design, CAD) obejmujących obecnie praktycznie wszystkie dziedziny projektowania technicznego. Coraz częściej również znajdują zastosowanie w niektórych dziedzinach sztuki np. w tworzeniu obrazów, wzorów tkanin konfekcyjnych i dekoracyjnych, w produkcji filmów animowanych itp.

2. Krótki zarys rozwoju

Historia graficznych monitorów ekranowych rozpoczęła się na przełomie lat 1950/60, a zapoczątkowały one swój bardzo szybki rozwój w 1962 r., kiedy to I.E. Sutherland z Massachusetts Institute of Technology (USA) przedstawił system Sketchpad, otwierający nowe i duże możliwości wykorzystania monitora graficznego jako wygodnego i efektywnego środka komunikacji człowiek-maszyna. W rok później, podczas Fall Joint International Computer Conference w San Francisco, firma

International Business Machines (IBM) zaprezentowała system IBM Alpine, głównie przeznaczony do inżynierskiego projektowania. Wzbudził on ogromne zainteresowanie i wywołał kontrowersyjne dyskusje, w których wyrażano zarówno głęboki sceptycyzm jak również prawie fanatyczny entuzjazm [1].

Dalszy rozwój monitorów graficznych to rozszerzanie i ulepszanie ich oprogramowania: zmiany w systemach operacyjnych dużych komputerów współpracujących z konwersacyjnymi urządzeniami graficznymi w celu ich efektywniejszego wykorzystania, np. zmodyfikowanie systemu operacyjnego dla Systemu IBM 360 w 1966 r., w celu umożliwienia współpracy z dwoma monitorami typu 2250, powstawanie nowych języków programowania znanych ogólnie pod angielską nazwą "Graphic languages", do których należą między innymi L⁶, LEAP, ASP, IPL, PGS, rozszerzenie języków wyższego rzędu np. języka FORTRAN o GRAPH SYSTEM, powstanie pakietów FRKD i LUISA itp [1].

Jednocześnie z rozwojem systemów wielodostępnych zrodziła się idea systemów monitorów graficznych, w których wiele pulpیتów graficznych mogłoby pracować jednocześnie z jedną dużą elektroniczną maszyną cyfrową. Wiąże się to z nowymi wymaganiami dotyczącymi pamięci operacyjnych, buforowych i masowych oraz kanałów przysyłania informacji graficznej. Coraz szersze zastosowanie monitorów graficznych w różnych dziedzinach stwarza konieczność powiększania repertuaru funkcji, jakie monitor graficzny ma spełniać i wymaga zmian i rozszerzenia parametrów charakteryzujących monitor, co wobec równoczesnego rozwoju elektroniki pociąga za sobą zmiany konstrukcyjne i technologiczne.

3. Stan obecny

Obecnie, tj. w roku 1974, po 12 latach rozwoju, liczba typów produkowanych monitorów obejmuje ponad 50 pozycji [2], a ceny ich wahają się od dziewięciu do kilkuset tysięcy dolarów, zależnie od ich możliwości funkcjonalnych i stosowanych rozwiązań technicznych.

W celu porównania monitorów graficznych różnych typów między sobą pod względem możliwości ich zastosowań, zostaną one podzielone na

dwie grupy. Za kryterium podziału będzie przyjęty rodzaj wskaźnika używanego do zobrazowania graficznej informacji.

Pierwsza grupa monitorów graficznych to urządzenia, w których wskaźnikiem jest lampa oscyloskopowa z odchyłaniem swobodnym lub rastrowym; ta grupa nosi często w literaturze angielskiej nazwę "CRT displays". Przy odchyłaniu rastrowym stosowana jest niekiedy w literaturze angielskiej dla monitorów nazwa "video displays". W monitorach tej grupy, aby uzyskać stabilny obraz na ekranie przez dłuższy okres czasu, konieczne jest cykliczne wyświetlanie obrazu przechowywanego w pamięci buforowej pamięci monitora lub też pamięci maszyny cyfrowej współpracującej z nim, z pewną określoną częstotliwością (30 - 60 Hz) [3].

Druga grupa to monitory, w których wskaźnikiem jest specjalna lampa oscyloskopowa mająca zdolność pamiętania (storage tube). Monitory te noszą w literaturze angielską nazwę "DVST Displays (Direct View Storage Tube Displays)" lub "Low-cost Graphic Displays" (tanie monitory graficzne). Ta druga nazwa nadana została ze względu na niską ich cenę (ok. 10.000 - 20.000 dolarów) w porównaniu z monitorami pierwszej grupy. Monitory DVST najczęściej wyposażone są w lampy firmy Tektronix (np. Tektronix 611 o wymiarach ekranu 160 x 205 mm).

Do pierwszej grupy zaliczyć można graficzne monitory produkowane we Francji przez firmę SINTRA (Cociété Industrielle des Nouvelles Techniques Radioelectriques et de l'Electronique Française) - typy GIDI i VU 2000 oraz firmę CSEE (Compagnie de Signaux et d'entreprises Electriques) - typ AFIGRAF, węgierski GD 71, radziecki 7064, monitory firmy ITT (International Telephone and Telegraph), DEC 340 (Digital Equipment Corporation) i IBM 2250.

Monitory graficzne pierwszej grupy charakteryzują się wieloma zaletami, których nie mają monitory z lampami pamiętającymi (typu DVST), a mianowicie:

- możliwości selektywnego kasowania (selective erase) fragmentu obrazu wskazanego przez operatora,

- możliwość zmian w intensywności świecenia elementów obrazu wybranych przez operatora
- "rozsuwanie" lub "zsuwanie" tekstu lub elementów obrazu,
- wprowadzenie "migotania" elementu obrazu wybranego przez operatora,
- przesuwanie wektorów lub figur w dowolne miejsce ekranu i obroty wokół dowolnej osi wybranej przez operatora,
- szybkie kasowanie informacji zobrazowanej na ekranie.

Monitory graficzne typu DVST Displays to przede wszystkim monitory T-4002, T-4005, T-4014/4015 firmy Tektronix, monitory serii 400, modele 10,12,15,20 firmy Computer, Inc., monitory systemu ARDS (Advanced Remote Displays Station) firmy Computer Displays, Inc. oraz niektóre monitory firmy DEC.

Monitor Tektronix T-4002 wyróżnia się wśród innych monitorów tej grupy posiadaniem dodatkowego sposobu wprowadzania danych z klawiatury. Sposób ten znany w literaturze pod nazwą "compose mode" charakteryzuje się tym, że pewna liczba znaków (odpowiadająca jednej linii) wprowadzanych przez operatora do komputera jest przechowywana w pamięci roboczej monitora (scratchpad memory) i równocześnie wyświetlana na części ekranu bez właściwości pamiętania; informacja ta może być więc sprawdzana i poprawiana przez operatora na bieżąco. Przy wprowadzaniu informacji sposobem bezpośrednim (direct mode) znaki czy symbole przesyłane są do komputera tak, jak są wprowadzane przez operatora. W przypadku błędu lub zmian, wobec niemożliwości skasowania pojedynczego znaku, wymazana musi zostać z ekranu cała zapisana informacja.

Wadą urządzeń z lampami typu DVST jest długi czas kasowania informacji tekstowej i graficznej wyświetlonej na ekranie (ok. 0,5 s. dla lampy Tektronix 611), przy czym kasowanie jest z reguły tym trudniejsze im dłużej informacja jest wyświetlana ponad 15 minut.

Konwersacyjne urządzenia graficzne są obecnie najwygodniejszymi urządzeniami zapewniającymi komunikację człowiek-maszyna. Ten dwustronny kontakt człowieka z maszyną zapewnia możliwość wyprowadzania

dźwiękowymi, przy czym doświadczenia wykazują, że wzrośnie zapotrzebowanie na pulpity o dużych wymiarach (mniej więcej odpowiadających średnim deskom kreślarskim). Jako urządzenia interakcyjne identyfikujące element obrazu wyświetlonego na obrazie monitora pozostaną pióra świetlne.

W oprogramowaniu przede wszystkim należy spodziewać się standaryzacji języka graficznego w celu zapewnienia całkowitej wymienności programów. Już zarysowują się tendencje zmiany organizacji listy obrazowej (display file) z wykorzystaniem właściwości pamięci asocjacyjnych. Rozwiną się programy umożliwiające częściowe transformowanie zobrazowań, tj. zmianę typu symboli czy oznaczeń symbolicznych na mapach, rysunkach konstrukcyjnych i schematach elektrycznych przy równoczesnym zachowaniu podstawowych relacji topologicznych czy geometrycznych.

Za lat 5 do 10 nastąpi prawdopodobnie odwrót od zobrazowań wykorzystujących lampy oscyloskopowe. Wykorzystując nowe zjawiska fizyczne (np. cienkie kryształy) można będzie przekazywać cyfrową informację celem zobrazowania wprost do wskaźnika informacji graficznej, bez potrzeby przejścia przez konwertery cyfrowo-analogowe. Wydaje się również, że monitory wzbogacone zostaną o liczne układy elektroniczne wykonywujące funkcje, które obecnie są realizowane za pomocą specjalnych programów. Być może wykorzystana zostanie technika światła koherentnego (lasery) do przetwarzania danych graficznych.

Prawdopodobnie wzrośnie liczba urządzeń dodatkowych do uzyskania trwałych kopii obrazu. Należy się również spodziewać, że ceny monitorów graficznych nie będą wiele wyższe od cen monitorów alfanumerycznych.

Należy również spodziewać się nowej koncepcji organizacji współpracy monitora graficznego z maszyną cyfrową. Wydaje się, że w następnym etapie zmian w konfiguracji maszyna cyfrowa - urządzenie graficzne nastąpi dalszy proces odciążania maszyny głównej. Prowadzi to do koncepcji wykorzystania standardowego minikomputera jako komputera pomocniczego, który spełniałby również wiele funkcji obecnie realizo-

puterami bezpośrednio przez kanał selektorowy lub multipleksorowy, a przy połączeniu zdalnym współpraca odbywa się dodatkowo przez modemy i linie telekomunikacyjne.

Wczesne rozwiązania graficznych systemów monitorowych (początek lat sześćdziesiątych) charakteryzowały się bezpośrednim dołączaniem monitorów do procesora głównego komputera, który realizował niektóre funkcje i interpretował wszystkie działania operatora. Pamięć głównego komputera służyła przy tym często do przechowywania danych obrazu, w celu jego odświeżania na ekranie.

W następnej generacji graficznych systemów monitorowych (połowa lat sześćdziesiątych) wprowadzono podział funkcji między główny komputer i dołączony monitor, który w tym celu został uzupełniony o dodatkowe układy elektroniczne i pamięć buforową (np. system CDC 3344/274) [6] .

Późniejsze doświadczenia wykazały, że dla systemów monitorowych obciążenie procesora głównego komputera rośnie niewspółmiernie szybko do wzrostu liczby podłączonych monitorów, w związku z czym dla lepszego i elastyczniejszego wykorzystania monitorów wprowadzono dodatkowy komputer, który przejął od głównego komputera realizowanie wielu funkcji, między innymi takich jak obsługa przerw, kolejkowanie, interpretacja instrukcji sterujących dla komunikacji maszyna-monitor, proces śledzenia pióra świetlnego itp. Przykładem takiego systemu jest system CDC 6000 IGS (INTERACTIVE GRAPHIC SYSTEM) zaprezentowany w 1965 r. [5] . W systemie tym do dodatkowego komputera pomocniczego (terminal computer), którym jest komputer CDC 1700, może być podłączone 6 monitorów graficznych za pośrednictwem jednostki sterującej.

Dalszy krok pod względem liczby monitorów podłączonych do jednego centralnego komputera uczyniła firma ITT (Szwecja). Zaprezentowała ona system monitorowy pod nazwą GRAFOSKOP DISPLAY SYSTEM, w którym dzięki wprowadzeniu procesora CENSOR 908 i pamięci buforowej (obie te jednostki tworzą tzw. Censor 908 Data Processor) z jednym komputerem może współpracować zdalnie do 80 monitorów graficznych. Mogą one być wyposażone oprócz klawiatury alfanumerycznej i funkcyjnej dodatkowo w pióro świetlne i nastawnik kulowy. System ten jest w pewnym sensie

systemem modułowym; podstawowym modułem pamięci buforowej Censor 908 Data Processor jest pamięć o pojemności 16 kbajtów, przy czym maksymalną liczbą jest 8 modułów, tj. maksymalna pojemność pamięci wynosi 128 kbajtów. Censor 908 pozwala na podłączenie przez kanał multipleksorowy do 5 generatorów obrazu zwanych Display Generator, a do każdego z generatorów obrazu można dołączyć do szesnastu monitorów typu GRAFOSKOP. Generatory znaków, wektorów, okręgów jak również układy logiczne pióra świetlnego i układy kontrolno-sterujące zawarte są w jednostkach Display Generator, przy czym maksymalna odległość pomiędzy nimi a monitorami wynosi 160 m. Oprogramowanie systemu Censor 908 jest bardzo bogate i spełnia wiele funkcji [6] .

W skład systemu proponowanego przez firmę SINTRA, opartego na monitorze VU 2000, wchodzi mały komputer typu ETM 1116, do którego można podłączyć do 4 jednostek sterujących. Do każdej takiej jednostki z kolei można dołączyć do 4 monitorów graficznych. Urządzeniami wejściowymi (interakcyjnymi) monitora VU 2000 są: pióro świetlne, nastawnik kulowy, klawiatura alfanumeryczna i klawiatura funkcyjna złożona z 16 kluczy, która ma 16 wymiennych nakładek (overlays). VU 2000 jest obecnie jednym z nielicznych jeszcze monitorów graficznych, który może być wykonany na życzenie klienta w wersji kolorowej i daje do dyspozycji trzy kolory: zielony, żółty i czerwony. Ekran jest prostokątny o przekątnej 21 cali, a użyteczna powierzchnia ekranu to kwadrat o wymiarach 13 x 13 cali. Luminofor lampy obrazowej wymaga częstotliwości odświeżania równej 40 Hz. Produkowana jest również wersja monitora VU 2000 bez pamięci buforowej, dla połączeń o bezpośrednim dostępie do pamięci komputera współpracującego.

Wszystkie omawiane monitory graficzne różnią się znacznie między sobą nie tylko pod względem technologii wykonania, konstrukcji, a także środkami interakcyjnymi i funkcjami, oraz listą rozkazów, długością słowa rozkazowego i oprogramowaniem.

Nie ma chyba dwóch podobnych do siebie ze względu na oprogramowanie systemów graficznych opracowanych dla współpracy monitora z komputerem. Np. znane systemy graficzne Sketchpad, Graphic-2 i APL, róż-

nią się między sobą przede wszystkim filozofią podejścia do struktury danych graficznych, która wpływa na pojemność pamięci i szybkość przeszukiwania danych np. dla uaktualniania czy modyfikacji danych. Firmy oferują na ogół po kilka wersji rozwiązań monitora graficznego, a mianowicie: z pamięcią buforową i bez pamięci buforowej, tylko z generatorami znaków lub z generatorami znaków, wektorów, łuków i okręgów, z jednym lub kilkoma urządzeniami interakcyjnymi (prócz klawiatur), z drukarkami i innymi urządzeniami umożliwiającym uzyskanie trwałych kopii. Np. firma Tektronix skonstruowała urządzenie (hard copy unit) typ 4601, które umożliwia otrzymanie trwałej kopii informacji zobrazonej na ekranie o wymiarach $8\frac{1}{2} \times 11$ cali w ciągu kilku sekund. Stosuje się do tego urządzenia specjalny papier (3M Type 777 Dry Silver Paper) dostarczany przez tę firmę. Oprócz tego firma Tektronix proponuje jako wyposażenie dodatkowe kamerę fotograficzną typu Polaroid C10 dla monitorów jedenasto-calowych.

Monitory na życzenie klienta mogą być wyposażone w różne układy współpracy z innymi urządzeniami. Na przykład, firma Tektronix oferuje wyposażenie w układy szeregowego interface dla transmisji w systemie duplex i półduplex (wg EIA, norma RS-232-B), układy interface dla szybkiej transmisji między monitorem graficznym, np. T-4002 i jednym z minikomputerów: Nova i Supernova firmy Data General; Hewlett - Packard 2114, 2115 i 2116, PDP - 8/I i PDP - 8/L firmy DEC.

Coraz bogatsze jest również oprogramowanie proponowane przez różne firmy. Wynika to z coraz szerszych zastosowań monitorów graficznych.

5. Tendencje rozwojowe

Bez wahania można obecnie powiedzieć, że nie ma dziedziny technicznej, w której zastosowanie monitora graficznego współpracującego z komputerem nie przyniosłoby dodatkowych korzyści i nie stworzyło nowych możliwości naukowo-badawczych. "Graficzne urządzenia zwrotne stanowią, jak dotychczas, najwyższe osiągnięcie w dziedzinie dostosowania trybu współpracy człowieka z maszyną do wymagań pierwszego partnera" pisze w swojej książce prof. W. Turski [5]. We wstępie do książki Advanced Computer Graphics [1] i [6] wydawcy E.D. Parslow i R.E. Green piszą: "Komputerowe monitory graficzne przestały być

ciekawostką techniczną. W ciągu ostatnich kilku lat monitory graficzne wyszły ze stanu, w którym stosowanie ich ograniczone było niemal wyłącznie do wielkich przemysłów jak lotniczy i samochodowy. Obecna technika dysponuje dodatkowymi możliwościami wykraczającymi poza koncepcję szybkiego wykonywania funkcji monitora; proponuje zupełnie nowe sposoby ich wykonywania, w rezultacie czego uzyskuje się większą sprawność i dokładność, a jednocześnie umożliwia wykorzystanie monitorów w nowych dziedzinach i wykonywanie czynności w sposób uprzednio niemożliwy". Powyższe zdania napisane zostały w 1970 r. Od tego czasu wzrosła jeszcze liczba typów monitorów graficznych i rozszerzył się krąg ich zastosowań. Publikacje z konferencji organizowanych przez IEE (Institution of Electrical Engineers) pod hasłem "Computer Aided Design" [12] wykazują niezbicie, że monitory graficzne stają się nieodzownym narzędziem w pracach naukowo-badawczych nie tylko dziedzin technicznych, lecz również w medycynie, biologii, szkolnictwie. Można więc postawić pytanie: w jakim kierunku rozwijać się będzie technika monitorów graficznych? Jak będzie wyglądać monitor graficzny za kilka lat i lat kilkanaście? Spróbujmy znaleźć odpowiedzi, biorąc pod uwagę wszystkie te parametry i właściwości określające przydatność monitora graficznego, które były uprzednio analizowane.

Przewidując kierunki rozwojowe przede wszystkim należy rozpatrzyć dwa aspekty:

- sprzęt
- oprogramowanie.

Powiedzmy, za lat trzy do pięciu można się spodziewać, że w zakresie rozwiązań sprzętowych zobrazowanie nie ulegnie zmianie, tj. pozostaniemy przy lampach oscyloskopowych, a więc przy odchyłaniu analogowym. Prawdopodobnie zwiększy się liczba różnych generatorów, tj. oprócz generatorów znaków jako standard monitory wyposażone będą w generatory wektorów, krzywych stożkowych i inne. W budowie tych generatorów będą wykorzystane specjalne układy scalone, szczególnie wobec potaniaenia produkcji układów o dużym stopniu integracji. Jako urządzenie interakcyjne pozostaną klawiatury, a oprócz tego rozwiną się wszelkiego rodzaju pulpity, najprawdopodobniej z ołówkami ultra-

dźwiękowymi, przy czym doświadczenia wykazują, że wzrośnie zapotrzebowanie na pulpity o dużych wymiarach (mniej więcej odpowiadających średnim deskom kreślarskim). Jako urządzenia interakcyjne identyfikujące element obrazu wyświetlonego na obrazie monitora pozostaną pióra świetlne.

W oprogramowaniu przede wszystkim należy spodziewać się standaryzacji języka graficznego w celu zapewnienia całkowitej wymienności programów. Już zarysowują się tendencje zmiany organizacji listy obrazowej (display file) z wykorzystaniem właściwości pamięci asocjacyjnych. Rozwiną się programy umożliwiające częściowe transformowanie zobrazowań, tj. zmianę typu symboli czy oznaczeń symbolicznych na mapach, rysunkach konstrukcyjnych i schematach elektrycznych przy równoczesnym zachowaniu podstawowych relacji topologicznych czy geometrycznych.

Za lat 5 do 10 nastąpi prawdopodobnie odwrót od zobrazowań wykorzystujących lampy oscyloskopowe. Wykorzystując nowe zjawiska fizyczne (np. ciekłe kryształy) można będzie przekazywać cyfrową informację celem zobrazowania wprost do wskaźnika informacji graficznej, bez potrzeby przejścia przez konwertery cyfrowo-analogowe. Wydaje się również, że monitory wzbogacone zostaną o liczne układy elektroniczne wykonywujące funkcje, które obecnie są realizowane za pomocą specjalnych programów. Być może wykorzystana zostanie technika światła koherentnego (lasery) do przetwarzania danych graficznych.

Prawdopodobnie wzrośnie liczba urządzeń dodatkowych do uzyskania trwałych kopii obrazu. Należy się również spodziewać, że ceny monitorów graficznych nie będą wiele wyższe od cen monitorów alfanumerycznych.

Należy również spodziewać się nowej koncepcji organizacji współpracy monitora graficznego z maszyną cyfrową. Wydaje się, że w następnym etapie zmian w konfiguracji maszyna cyfrowa - urządzenie graficzne nastąpi dalszy proces odciążania maszyny głównej. Prowadzi to do koncepcji wykorzystania standardowego minikomputera jako komputera pomocniczego, który spełniałby również wiele funkcji obecnie realizo-

wanych przez jednostkę sterującą. W przypadku opracowania języka pozwalającego na dokonanie formalnego opisu własności monitora graficznego z uwzględnieniem jego wyposażenia elektronicznego, formatów danych i listy rozkazów oraz po wyposażeniu komputera pomocniczego w program pozwalający na przetworzenie każdej instrukcji graficznej napisanej w ogólnie przyjętym standardowym języku graficznym powstanie możliwość dołączania dowolnego monitora graficznego do komputera pomocniczego, jeśli zapewnione są właściwe warunki współpracy na poziomie fizycznych rozwiązań układów interface'u. Zapewniłoby to możliwość współpracy standardowej jednostki sterującej z szerokim zbiorem różnych graficznych monitorów ekranowych.

Literatura

- [1] LUNDAY P.A.: The Lessons of the 60's, Advanced Computer Graphics. 1971 Wyd. R.D. Parslow, R.E. Green.
- [2] AUERBACH: Computer Technology Reports.
- [3] Fourth Generation - Infotech, State of the Art Report 1.
- [4] Latest developments in electronic data processing, Auerbach Publishers Inc. 1971.
- [5] TURSKI W.: Podstawy użytkowania maszyn cyfrowych w ośrodkach naukowo-technicznych, Warszawa 1973, PWN.
- [6] DOWE R.J.: ASPECTS OF INTERACTIVE GRAPHIC SYSTEMS, Advanced Computer Graphics. R.E.Green. R.D. Parslow 1971.
- [7] ITT, Grafoskop Graphic Data Display System - opis systemu, sprzętu i oprogramowania.
- [8] STILLMAN N.I., DEFIORE C.R.: Associative Processing of Line Drawings. Spring Joint Computer Conference. 1971.
- [9] SUTHERLAND I.E.: Sketchpad - A man - machine graphical communication system. Proceedings Spring Joint Computer Conference 1963.
- [10] DAŃDA J.: Pióra świetlne i strumieniowe. Maszyny Matematyczne 1968, nr 9.
- [11] Software for Control. IEK Conference Publication. 1973. nr 102.
- [12] STOLARSKI M.: Nowoczesna komunikacja graficzna człowieka z maszyną cyfrową. Problemy Przetwarzania Informacji, Warszawa 1970, WNT.

mgr. inż. Ryszard PATRYN
Instytut Maszyn Matematycznych

007.51:681.327.22:534.78

AKUSTYCZNE URZĄDZENIA WYJŚCIOWE

1. Wstęp

Obecnie w komunikacji maszyna cyfrowa-człowiek wykorzystywany jest głównie i powszechnie zmysł wzroku. Podstawowe urządzenia wyjściowe, które umożliwiają człowiekowi bezpośredni dostęp do wyników, to wszelkiego rodzaju urządzenia drukujące lub wyświetlające. Zmysł słuchu wykorzystywany jest w bardzo małym stopniu, głównie do przekazywania sygnałów zwracających uwagę, np. brzęczyk, dzwonek, jeśli nie brać pod uwagę nielicznych dotychczas akustycznych urządzeń wyjściowych.

Zastosowanie sygnałów mowy pozwala na przekazywanie bardziej złożonych informacji, przy czym przyjmowanie takich sygnałów nie wymaga dalece natężonej uwagi, jak to ma miejsce w przypadku sygnałów odbieranych wzrokowo.

Zastosowanie sygnałów mowy umożliwia poza tym bezpośrednie korzystanie z linii telefonicznych, a aparat telefoniczny, sprzęt powszechnie stosowany i ogólnie dostępny, staje się końcowym urządzeniem wyjściowym maszyny cyfrowej.

Stosowanie wyjścia akustycznego może być uzasadnione w następujących przypadkach:

- do przekazania informacji dodatkowych w systemach wykorzystujących tradycyjne urządzenia wyjściowe, do polepszenia warunków pracy i podwyższenia jakości systemu,

- jako jedyna możliwość komunikacji maszyna-człowiek w tych warunkach, gdy inne środki komunikacji nie spełniają należycie swego zadania tzn.
 - gdy człowiek jest w ruchu,
 - gdy człowiek jest zajęty wykonywaniem czynności uniemożliwiających lub utrudniających korzystanie z sygnałów odbieranych wzrokowo,
 - gdy dostęp do tradycyjnych urządzeń wyjściowych jest utrudniony (rolę urządzenia wyjściowego może spełniać aparat telefoniczny).

Obecnie akustyczne urządzenia wyjściowe znajdują się jeszcze w fazie prób i badań, chociaż pierwsze ich zastosowanie miało miejsce w 1964 r. na giełdzie nowojorskiej. Od tego czasu powstało wiele koncepcji realizacji wyjścia akustycznego, z których część została praktycznie wykonana, lecz nie znalazła szerszego zastosowania.

2. Krótka charakterystyka koncepcji i sposobów realizacji akustycznego wyjścia maszyny cyfrowej

a) Zapis analogowy

W metodzie tej zapisuje się wyrazy na bębnie z nośnikiem magnetycznym lub na filmie fotooptycznie. Ze względu na jednakowy czas obrotu bębna wszystkie wyrazy zapisane na nim powinny mieć długość równą krotkości czasu obrotu. Każdy wyraz zapisany jest na oddzielnej ścieżce, krótkie wyrazy na jednej, dłuższe na dwóch lub trzech ścieżkach. W urządzeniu - dla każdej ścieżki wymagane jest stosowanie oddzielnego układu odczytującego, tj. głowicy magnetycznej lub fotodiody.

Wyrazy zapisane na bębnie tworzą określony zbiór, słownik, który umożliwia wytworzenie wielu wypowiedzi zaspokajających potrzeby systemu.

Wyodrębnienie dziedziny, w której ma pracować system z wyjściem akustycznym pozwala na ograniczenie objętości słownika. Przyjmuje się, że objętość słownika w urządzeniach tego typu wynosi na ogół kilkadziesiąt wyrazów, a może dochodzić do tysiąca wyrazów, liczba kanałów akustycznych, czyli oddzielnych wyjść sterowanych jednocześnie przez ma-

szynę cyfrową nie przekracza 100. Jakość uzyskiwanej mowy może być bardzo dobra, gdyż wyrazy są zapisane w brzmieniu naturalnym przy pewnej kompresji czasowej, której celem jest normalizacja czasu trwania.

b) Zapis cyfrowy

W metodzie tej sposób tworzenia wypowiedzi jest taki sam, jak w metodzie poprzednio opisanej, ale wyrazy są zapisywane w pamięci w postaci cyfrowej. Z tego względu pobieranie z pamięci może odbywać się w taki sposób, jak to ma miejsce przy operowaniu innymi danymi przez komputer. Ze względu na jednolitą formę danych sposób taki byłby wygodny, ale wymaga przechowywania dosyć dużej ilości informacji cyfrowych.

Zakładając dynamikę sygnału mowy ok. 30 dB i częstotliwość próbkowania przy przetwarzaniu analogowo-cyfrowym ok. 7 kHz co teoretycznie zapewnia pasmo akustyczne do 3,5 kHz, dla zapisu mowy o czasie trwania 1 s wymagana jest pamięć o objętości ok. 40 000 bitów. Przy zwiększeniu częstotliwości próbkowania wartość powyższa zwiększyłaby się i praktycznie może osiągnąć 100 000 bitów/s.

Do przechowywania wyrazów słownika nadawałyby się pamięci dyskowe, które mogłyby przechowywać kilka tysięcy wyrazów, ale w przypadku pamięci dyskowych z ruchomymi głowicami możliwość obsłużenia wielu kanałów akustycznych jest ograniczona do kilku. Dla każdego kanału akustycznego należy ponadto przewidzieć obszar pamięci operacyjnej ok. 10000 bitów jako pamięci buforowej w celu wyeliminowania przerw przy czytaniu sygnałów poszczególnych wyrazów, ze względu na znaczny czas dostępu.

c) Kompansja parametryczna i zapis cyfrowy

W celu zmniejszenia pojemności pamięci wymaganej w urządzeniach z wyjściem akustycznym, sygnał mowy poddaje się kompansji parametrycznej, tj. przekształceniu polegającemu w pierwszym etapie na wydzielaniu ograniczonej liczby parametrów, z których każdy jest wolnozmienną funkcją czasu, a w drugim etapie wytworzenie na podstawie uzyskanych informacji zrozumiałej mowy.

Urządzeniem służącym do parametrycznej kompensacji mowy jest między innymi wokoder. Wokoder składa się z części analizującej, w której następuje wydzielenie niezbędnej liczby parametrów określających sygnał mowy oraz z części syntetyzującej, w której następuje odtworzenie z większą lub mniejszą dokładnością sygnału mowy. Głównym celem stosowania wokodera jest znaczne zmniejszenie ilości informacji potrzebnej do przekazania sygnału mowy.

Wykorzystując technikę stosowaną w wokoderach zapisuje się w pamięci określony zbiór wyrazów podobnie jak w metodach poprzednich. Przy zastosowaniu kompensacji parametrycznej do zapamiętania mowy o czasie trwania 1 s wymagane jest 2400 - 4000 bitów.

Najbardziej znane typy wokoderów to wokoder kanałowy i wokoder formantowy. Formant - obszar koncentracji energii w widmie dźwięku mowy (fonemu, głoski). W wokoderze kanałowym sygnał mowy zostaje rozdzielony w układzie filtrów środkowo-przepustowych, pokrywających zakres częstotliwości widma sygnału, na kilka do kilkunastu przylegających do siebie częstotliwości i rejestruje się zmianę w czasie wartości amplitud dla każdego pasma częstotliwości (kanału). Otrzymane sygnały poddaje się przetwarzaniu analogowo-cyfrowemu przy czym stosuje się na ogół częstotliwość próbkowania $40 \div 100$ Hz i $8 \div 15$ poziomów kwantyzacji amplitudy.

W wokoderze formantowym wykorzystuje się mniejszą liczbę sygnałów składowych, np. sygnał tonu krtaniowego, sygnały zmienne (amplituda i częstotliwość) pierwszego i drugiego formantu oraz sygnały trzeciego i czwartego formantu. Otrzymane sygnały są określone przez wartości częstotliwości i amplitudy, które zapamiętuje się w postaci cyfrowej. Liczba bitów potrzebna do zapamiętania sygnałów sterujących jest mniejsza niż w wokoderze kanałowym, kształtuje się poniżej 2000 bitów/s.

d) Synteza fonematyczna

Dla każdego języka można wydzielić z mowy najmniejsze elementy fonetyczne - fonemy. Zależnie od języka ich liczba wynosi od 40 do 50. W mowie średniej prędkości występuje około 10 fonemów na sekundę.

System syntezy wykorzystujący małą liczbę elementów mowy byłby pożądany, gdyż byłby to system bardzo elastyczny, umożliwiający tworzenie dowolnych wyrazów, bez żadnych ograniczeń.

W praktyce synteza fonematyczna przybiera postać bardziej skomplikowaną ze względu na zniekształcenia mowy powstające na styku dwóch sąsiednich fonemów syntetycznych. W mowie naturalnej między poszczególnymi fonemami tworzą się przebiegi przejściowe, które przyczyniają się do zachowania ciągłości funkcji wartości częstotliwości poszczególnych formantów w czasie.

Dla każdego języka można ustalić zasady powstawania przebiegów przejściowych (transientów) przy styku dwóch fonemów. Na podstawie tych zasad powstała metoda syntezy fonematycznej, w której tworzy się zbiór fonemów i zbiór zasad (reguł) ich łączenia. W metodzie tej, zwanej syntezą, według reguł, wytwarza się przebiegi przejściowe wg zasad zawartych w programach sterujących syntezą. Oczywiście informacje określające fonemy i przebiegi przejściowe przechowuje się w postaci cyfrowej, a sygnał elektroakustyczny uzyskuje się po przetworzeniu cyfrowo-analogowym.

Odmianą powyższej metody jest wytworzenie gotowych zestawów fonemów w takich połączeniach, w jakich występują w danym języku, za pomocą poprzednio wspomnianych reguł - i zapisanie tych zestawów w pamięci komputera. Liczba takich zestawów zależy od właściwości danego języka i wynosi od 500 do 1000. W celu zmniejszenia zapamiętanych ilości informacji przechowuje się w tym przypadku te fragmenty, które tworzą charakterystyczny okres podstawowy. Do zapamiętania kilkuset zestawów fonemów wystarczy objętość pamięci ok. 100 000 bitów. Sygnał cyfrowy odpowiadający danemu fragmentowi mowy tworzy się za pomocą specjalnych programów, które powodują wielokrotne odtworzenie okresu podstawowego danego fonemu i uzupełnienie otrzymanej sekwencji przebiegami przejściowymi między sąsiednimi fonemami. Metoda taka umożliwia składanie wypowiedzi na podstawie tekstu pisanego, czyli następuje składanie wyrazów z głosek a z wyrazów tworzenie zdań.

3. Porównanie metod wytwarzania mowy w urządzeniach wyjściowych maszyny cyfrowej

Porównanie metod opisanych w pkt 2 można przeprowadzać na podstawie różnych kryteriów:

- łatwość realizacji technicznej,
- zrozumiałość i wyrazistość mowy,
- objętość słownika,
- koszt urządzenia,
- trwałość i częstotliwość zabiegów konserwacyjnych,
- gabaryt urządzenia.

W przypadku ograniczonego słownika i małej liczby wyjściowych kanałów akustycznych najdogodniejsza wydaje się metoda zapisu analogowego na bębnie. W tym przypadku można uzyskać bardzo dobrą zrozumiałość zapisanych wyrazów, a urządzenie jest proste i stosunkowo tanie. Wadą tej metody jest duża objętość urządzenia (ze względu na duże wymiary bębna) oraz ograniczona liczba możliwych do uzyskania wypowiedzi.

Metoda opisana w pkt 2 b charakteryzuje się podobnymi właściwościami jak metoda wg pkt 2 a, ale ogólnie biorąc układ pamięciowy jest o wiele kosztowniejszy. Zaletą jest stosowanie cyfrowych metod operowania danymi.

Metoda wg pkt 2 c może być zaliczona do tej samej grupy urządzeń co metody omówione poprzednio, a różnica polega na znacznym zmniejszeniu ilości przechowywanych informacji dla wytworzenia wyrazu w urządzeniach wyjściowych maszyny cyfrowej, przy czym długości wyrazów nie muszą być normalizowane. Jeżeli w metodzie omówionej w pkt 2 b na jeden wyraz 0,5 s wymagana jest co najmniej 20 000 bitów to w metodzie wg pkt 2.3 ok. 300+500 bitów. Jakkolwiek słownik w takiej metodzie również zawiera skończoną liczbę wyrazów, to ze względu na znacznie zmniejszone wymagania co do objętości pamięci przypadającej na jeden wyraz, można wykorzystywać słownik o większej liczbie wyrazów. Jakość mowy uzyskanej w wyniku przekształcenia jest nieco niższa niż mowy zapisa-

nej bez przekształceń, ale praktycznie jest wystarczająca w opinii autorów opisujących takie urządzenia. Z punktu widzenia składania wypowiedzi, powyższe metody charakteryzują się podobnym działaniem.

Zupełnie inny charakter ma synteza mowy wg metody opisanej w pkt. 2 d, w której sylaby tworzy się z fonemów, z sylab wyrazy, z wyrazów pełną wypowiedź. W tym przypadku można tworzyć wypowiedzi na podstawie tekstu pisanego, oczywiście po przekształceniu formy ortograficznej na fonetyczną.

4. Przegląd urządzeń opracowanych na świecie

W pierwszych urządzeniach mówiących stosowanych praktycznie (1964r. na giełdzie nowojorskiej) wykorzystywano metodę zapisu analogowego na bębnie. Przykładem takiego rozwiązania może być urządzenie 7770 firmy IBM zawierające 32 wyrazy (z możliwością powiększenia do 128) o czasie trwania 0,5 s. Firma Cognitronics Co. wyprodukowała urządzenia pod nazwą The Speechmaker, w którym wyrazy zapisano na bębnie metodą fotooptyczną (na filmie). Słownik zawierał 69 lub 189 wyrazów o czasie trwania 1,6 s.

Przykładem zastosowania metody opisanej w pkt. 2.3 może być urządzenie 7772 firmy IBM, w którym rolę syntezy spełnia wkoder kanałowy. Sygnały sterujące syntezą przechowywane są w pamięci dyskowej, a na ich podstawie można wytworzyć zbiór obejmujący kilka tysięcy wyrazów.

Urządzenia wg metody opisanej w pkt. 2 d nie są reprezentowane w grupie urządzeń produkowanych, ale prasa techniczna donosi o próbach budowy takich urządzeń. Opis układu syntezy fonematycznej (odmiana syntezy wg reguł) dla mowy japońskiej podali w 1971 r. T. Sakai, K. Ohtami, S. Tomita [6]. W układzie tym przechowuje się w pamięci bębnowej informacje cyfrowe o zestawach fonemów. Autorzy przewidują dla języka japońskiego około 550 zestawów fonemów. Większość zestawów fonemów utworzona była syntetycznie za pomocą komputera. Osiągnięto zrozumiałość około 75%.

W 1969 r. firma Culler Harrison Co. zapowiadała akustyczne urządzenie wyjściowe Hal 1 [9], pozwalające na wybór rodzaju głosu (męski, żeński, dziecienny) umożliwiające wprowadzenie intonacji i akcentu. Urządzenie miało mieć wymiary 2x3x1 stóp, cena miała wynosić 40 000 \$, ale nie natrafiono w okresie późniejszym na wzmianki w prasie o takim urządzeniu.

5. Kierunki prowadzonych prac badawczych na świecie

Pierwsze akustyczne urządzenia wyjściowe wykorzystywały techniki znane w danym momencie. Można to powiedzieć o metodach przechowywania w pamięci gotowych wyrazów i również o synteźatorach typu wokoderowego. Techniki kompensacji parametrycznej mowy opracowywano, ogólnie biorąc, dla telekomunikacji, gdzie przy odtwarzaniu mowy wyjściowej co najmniej pożądanym jest zachowanie pewnych cech mowy pierwotnej. W przypadku wyjścia akustycznego maszyny cyfrowej właściwości pierwotne, w sensie jak wyżej, nie istnieją.

Ostatnio można zauważyć tendencję do budowy akustycznych urządzeń wyjściowych metodą syntezy fonematycznej. W przypadku syntezy według reguł konieczne jest przechowywanie w pamięci dosyć dużej ilości informacji i układ jest raczej skomplikowany. Z drugiej strony sprzęt (elementy) układów cyfrowych tanieją, a wzrasta zainteresowanie bezpośrednią komunikacją maszyna-człowiek.

6. Prace prowadzone w PRL

W Polsce prace związane z akustycznym wyjściem maszyny cyfrowej prowadzone są w Instytucie Podstawowych Problemów Techniki w Zakładzie Akustyki Cybernetycznej. Zbudowano tam synteźator formantowy Synfor II sterowany sygnałami analogowymi. Prowadzi się prace w celu przejścia na sterowanie sygnałami cyfrowymi. W Instytucie Maszyn Matematycznych Uniwersytetu Warszawskiego prowadzone są prace związane z akustycznym wyjściem i wejściem maszyny cyfrowej. W układzie wyjściowym badane są możliwości zastosowania syntezy fonematycznej.

W Instytucie Maszyn Matematycznych (ZPA i AP "MERA") prowadzone są badania możliwości wyposażenia maszyn cyfrowych w akustyczne urządzenia wyjściowe oparte na zasadzie syntezy fonematycznej.

7. Wnioski

W przyszłości największe możliwości rozwoju wydają się mieć akustyczne układy wyjściowe oparte na metodzie syntezy fonematycznej, która jest najbardziej elastyczna i wymaga mało rozbudowanego oprogramowania. Mankamentem tej metody jest niewątpliwie monotoność mowy, jej charakter "maszynowy". Z jednej strony nie wydaje się, aby można zmienić łatwo ten charakter, ale z drugiej strony nie wiadomo, czy to jest potrzebne, jeśli mowa syntetyczna będzie dostatecznie zrozumiała. Pozbawienie jej elementów emocjonalnych lub innych subiektywnych cech mowy naturalnej nie powinno stanowić przeszkody w jej stosowaniu. W pewnych przypadkach możnaby dopuścić gorszą jakość mowy syntetycznej, jeśli odbiorcy tej mowy przyzwyczajają się do jej rozumienia, jak to ma miejsce w telekomunikacji.

Oceniając poszczególne metody, należy brać również pod rozwagę również właściwości języka. Ogólnie można stwierdzić, że nie wszystkie doświadczenia dotyczące syntezy mowy prowadzone dla jednego języka (np. angielskiego) mogą być wykorzystane bezpośrednio dla innego języka (np. polskiego). Uwzględnienie specyficznych właściwości języka zobowiązuje do prowadzenia w pewnym sensie odrębnych badań w różnych krajach.

Literatura

- [1] SAPOŹKOV M.A.: Sygnał mowy w telekomunikacji i cybernetyce. Warszawa 1966, WNT.
- [2] WELICZKO B.G.: Multiplikativnyj metod fonemnego sinteza reči s pomoščju ECVM. Seminar - Raspoznavanie obrazov i konstruirovanie čitajuščich avtomatov. Kijev 1967.
- [3] TATCHAM M.A.: Speech Synthesis - A Critical Review of the State of the Art. International Journal of Man Machine Studies 1970, nr 7.
- [4] AINSWORTH W.W.: Estimation of Speech Synthesiser Parameters from Acoustic Wave forms. International Journal of Man Machine Studies 1970, nr 7.
- [5] DIBBERN U.: Rechnergescheuerte Sprachausgaben. Internationale Elektronische Rundschau 1971, nr 10.
- [6] SAKAI T., OHTANI K., TOMITA S.: On Line Real Time Multiple - Speech Output Systems. IFIP Congress 1971.
- [7] KACPROWSKI J.: Akustyczny sygnał mowy w komunikacji człowiek-komputer. Warszawa 1972, IPPT-PAN.
- [8] HILL D.R.: An Abbreviated Guide to Planning for Speech Interaction with Machines: the State of the Art. International Journal of Man Machine Studies 1972, nr 4.
- [9] Electronics. September 69 U.S. Reports, s. 54.

mgr inż. Ryszard PATRYN
Instytut Maszyn Matematycznych

534.78

EKSPERYMENTALNY SYNTEZATOR FONEMATYCZNY W IMM

W czerwcu 74 r. w IMM zakończono badania, których celem było sprawdzenie możliwości realizacji fonematycznej syntezy mowy bez uwzględniania stanów przejściowych.

W wyniku tych badań powstał eksperymentalny syntezaator fonematyczny, za pomocą którego można uzyskiwać dowolne wypowiedzi. Syntezaator sterowany jest sygnałami cyfrowymi z klawiatury elektrycznej lub z czytnika taśmy dziurkowanej albo z innego urządzenia, z którego można wyprowadzać zakodowane cyfrowo znaki alfanumeryczne.

Eksperymentalny syntezaator składa się z pamięci dyskowej, minikomputera, specjalizowanej jednostki sterującej i wzmacniacza akustycznego. W pamięci dyskowej przechowuje się sygnały wyodrębnionych z mowy naturalnej głosek, które są zapisane w postaci cyfrowej. W minikomputerze realizuje się program sterujący wybieraniem fonemów z pamięci dyskowej, na podstawie uprzednio wprowadzonego tekstu. Sygnały cyfrowe z pamięci dyskowej kierowane są do układów pośrednich, między innymi w celu uzyskania właściwej prędkości wyjściowej oraz przekształcenia na sygnał analogowy. Tym sygnałem jest sterowany wzmacniacz akustyczny.

W syntezaatorze wykorzystuje się 42 fonemy, z których można złożyć dowolną wypowiedź w języku polskim. Wszystkie fonemy charakteryzują się jednakową wysokością tonu krtaniowego, co jest powodem monotoności uzyskiwanej mowy. W celu polepszenia zrozumiałości wprowadzo-

no akcentowanie, polegające na przedłużeniu czasu trwania akcentowanych samogłosek.

Tekst, który jest podstawą wypowiedzi jest pisany fonetycznie, ale rozpoczęto już prace nad translatozem, który umożliwi korzystanie z tekstów pisanych zgodnie z zasadami ortografii.

Pozytywne wyniki badań syntezy upoważniają do podjęcia prac, w celu opracowania fonicznego układu wyjściowego, którego działanie oparte będzie na zasadzie syntezy fonematycznej, a budowa prosta i ekonomiczna, co umożliwi łatwe wprowadzenie tego typu urządzeń do produkcji.

mgr inż. Eugeniusz SAWICKI
Instytut Maszyn Matematycznych

007.51:534.78

AUTOMATYCZNE ROZPOZNAWANIE MOWY

1. Wstęp

Proces fonicznej komunikacji człowieka z maszyną (urządzeniem) składa się z przesyłania informacji od człowieka do urządzenia (analiza) i od urządzenia do człowieka (synteza). Niniejsze opracowanie dotyczy przekazywania informacji od człowieka do urządzenia za pomocą głosu.

Zespół urządzeń i środków przeznaczonych do rozpoznawania mowy przyjęto nazywać systemami automatycznego rozpoznawania mowy. W początkowym okresie urządzenia rozpoznawały bardzo małą liczbę sygnałów akustycznych. Wraz z pojawieniem i rozpowszechnieniem się maszyn cyfrowych prace nad zagadnieniem automatycznego rozpoznawania mowy - ARM automatic speech recognition (ASR) zostały znacznie zintensyfikowane.

Zadaniem systemu ARM jest rozpoznanie informacji zawartej w sygnale mowy i wykonanie określonej czynności zgodnie z treścią przesyłanej informacji. Jakkolwiek uniwersalne systemy ARM powinny w zasadzie rozpoznawać mowę ludzką przekazywaną do urządzenia przez różne osoby bez ograniczeń dotyczących liczby rozpoznawanych wyrazów, sposobu wymawiania itp., to z użytkowego punktu widzenia można zrezygnować z pewnych wymagań. W konkretnych zastosowaniach mogą być przydatne systemy rozpoznające ograniczoną liczbę wyrazów wypowiedzianych przez większą liczbę osób lub większą liczbę wyrazów wypowiedzianych przez wybrane grono użytkowników.

Zaletą stosowania mowy jako środka komunikacji z maszyną jest naturalność (z punktu widzenia człowieka) i ogólna dostępność tej formy porozumiewania się. Komunikacja głosem może być użyta w warunkach, w których inne sposoby zawodzą (np. w ciemności) oraz dostępna jest dla osób częściowo upośledzonych (inwalidzi, niewidomi itp.). Środkiem technicznym zapewniającym łączność z maszyną może być np. zwykła sieć telefoniczna.

Jak dotychczas główną przeszkodą w wykorzystaniu mowy do ionicznej komunikacji człowieka z urządzeniem jest niedoskonałość systemów ARM. Dotychczas znane systemy są w większości urządzeniami laboratoryjnymi przeznaczonymi do badań, rozpoznającymi ograniczoną liczbę wyrazów (kilkadziesiąt do kilkuset) z dokładnością znacznie mniejszą od 100 %.

2. Ogólna charakterystyka automatycznego rozpoznawania mowy

Automatyczne rozpoznawanie mowy jest procesem bardzo skomplikowanym i do dnia dzisiejszego niecałkowicie rozwiązanym. Szczególne trudności stwarza określenie i wybór cech dystynktywnych umożliwiających jednoznaczne rozpoznanie sygnału mowy, tym bardziej, że nie można stworzyć modelu opartego na rozpoznawaniu mowy przez mózg ludzki, gdyż procesy percepcji mowy u człowieka nie zostały jeszcze dokładnie poznane. Znaczne rozproszenie prac (w sensie geograficznym i tematycznym) oraz odmienność poszczególnych języków utrudniają wzajemną wymianę doświadczeń. Z tych względów rozwój ARM postępuje stosunkowo powoli.

Do automatycznego rozpoznawania mowy używa się powszechnie maszyn cyfrowych. W procesie tym można wyróżnić następujące operacje:

- a) przetworzenie ciśnienia akustycznego na przebiegi elektryczne,
- b) normalizacja przebiegów elektrycznych,
- c) wydzielenie z sygnału wejściowego cech dystynktywnych odróżniających rozpoznawany element mowy od innych,
- d) określenie przez analizę cech dystynktywnych do jakiego podzbioru spośród pełnego zbioru elementów mowy (rozpoznawanych przez urządzenie) należy lub może należeć rozpoznawany element mowy,

- e) rozpoznanie dłuższych segmentów mowy,
- f) analiza zawartości semantycznej i syntaktycznej rozpoznawanej informacji,
- g) wykonanie określonej czynności na podstawie rozpoznanej informacji wejściowej.

Niektóre spośród wymienionych wyżej operacji dokonywane są poza maszyną cyfrową. Dotyczy to w szczególności pkt. a. Normalizacja przebiegów elektrycznych (b) oraz wydzielenie cech dystynktywnych (c) może być dokonane przez zewnętrzne urządzenia analogowe lub też przez maszynę cyfrową. Pozostałe operacje (d, e f) wykonywane są z reguły przez maszynę cyfrową. Po rozpoznaniu przekazywanej informacji element wykonawczy systemu ARM wykonuje określone czynności zgodnie z treścią przekazywanej informacji (g).

Większość dotychczas przeprowadzanych badań związanych z automatycznym rozpoznawaniem mowy odbywała się w wyciszonych i bezpogłosowych pomieszczeniach tzn. w warunkach zbliżonych do studyjnych (dotyczy to oczywiście tylko pkt. a, tj. przetwarzania ciśnienia akustycznego na przebiegi elektryczne). W celu wyeliminowania zakłóceń utrudniających prawidłowe rozpoznanie używano wysokiej jakości aparatury przetwarzająco-zapisującej: mikrofonów, wzmacniaczy, magnetofonów.

W zależności od długości czasu potrzebnego maszynie cyfrowej na analizę i podjęcie decyzji sygnał elektryczny może być przesyłany do maszyny na bieżąco w trakcie wypowiedzania badanego elementu mowy lub też może być zarejestrowany (na taśmie magnetycznej lub perforowanej) i przesłany do maszyny w późniejszym okresie czasu (gdy rozpoznanie trwa dłużej od czasu wypowiedzania). W celu przynajmniej częściowego zmniejszenia różnorodności sygnałów przeprowadza się normalizację niektórych jego parametrów. Stosuje się zwykle automatyczną regulację wzmocnienia dla osiągnięcia zbliżonego poziomu głośności i ograniczenie czasu trwania sygnału oraz pasma przenoszonych częstotliwości.

Normalizacja przebiegów elektrycznych (b) nie powinna spowodować zniekształcenia cech dystynktywnych, tj. tych parametrów sygnału, które umożliwiają odróżnienie danego elementu mowy od innych. Znorma-

lizowany sygnał elektryczny poddawany jest analizie w celu wydzielenia z niego cech dystynktywnych (o) . Cechami dystynktywnymi sygnału może być przebieg ciśnienia akustycznego w czasie, rozkład widmowy energii lub inne parametry sygnału w zależności od przyjętego sposobu rozpoznawania. Przykłady sposobów rozpoznawania oraz wybierania związanych z tym cech dystynktywnych podane są w dalszej części opracowania.

Maszyna cyfrowa analizuje cechy dystynktywne sygnału badanego (d) porównując je z przechowywanymi w jej pamięci cechami dystynktywnymi wzorców wszystkich elementów mowy rozpoznawanych przez urządzenie. Decyzja o przynależności badanego elementu mowy do określonej podgrupy polega na badaniu zgodności cech dystynktywnych rozpoznawanego elementu mowy i wzorca. Jako kryterium zgodności przyjmuje się różne parametry w zależności od sposobu rozpoznawania używając do analizy różnych algorytmów porównania i podejmowania decyzji (przykłady w dalszej części opracowania) . Jako najmniejsze rozpoznawane elementy mowy przyjmuje się fonemy, sylaby lub wyrazy. Dla prawidłowego rozpoznania treści przekazywanej informacji analizuje się dłuższe segmenty mowy (e) oraz przeprowadza się analizę zawartości semantycznej (f) . Niekiedy dopiero te analizy pozwalają prawidłowo rozpoznać mniejsze elementy mowy (np. analiza całego zdania umożliwia rozpoznanie pojedynczego wyrazu).

Jak wspomniano, najtrudniejszym i dotychczas niecałkowicie rozwiązany problemem w komunikacji człowiek-maszyna jest wybór cech dystynktywnych oraz opracowanie najbardziej efektywnych algorytmów porównania i podejmowania decyzji. Wybór cech dystynktywnych ma decydujący wpływ na jakość rozpoznawania i zależy od przyjętej metody rozpoznawania.

3. Krótki przegląd stosowanych metod rozpoznawania

a) Porównywanie wartości chwilowych sygnałów

W tym przypadku w pamięci maszyny przechowuje się zapisane analogowo lub cyfrowo wzorcowe sygnały wszystkich rozpoznawanych przez urządzenie elementów umowy. Analiza polega na porównaniu wartości

chwilowych sygnału wejściowego i sygnałów wzorcowych. Metoda ta jest mało efektywna ze względu na dużą różnorodność sygnałów mowy, np. ten sam wyraz wypowiedziany przez jedną i tę samą osobę daje za każdym razem różniące się od siebie przebiegi. Wymaga to przechowywania w pamięci maszyny dużej liczby sygnałów wzorcowych (co najmniej kilka na jeden rozpoznawany element mowy), co powoduje, że rozpoznawanie jest bardzo czasochłonne, a uzyskiwane wyniki nie są w pełni zadowalające.

b) Porównywanie rozkładu widmowego energii sygnałów

Jako cechy dystynktywne przyjmuje się w tej metodzie rozkład energii sygnału w funkcji częstotliwości. Maszyna cyfrowa porównuje widmo badanego sygnału z przechowywanym w swej pamięci widmem sygnałów wzorcowych. Rozkład widmowy energii sygnału (spektrogram) otrzymuje się przez podanie analizowanego sygnału do urządzenia wejściowego układu składającego się z kilku do kilkunastu filtrów pasmowych (tzw. analizatora kanałowego) pokrywających w sposób ciągły całe pasmo częstotliwości mowy ludzkiej. Ponieważ widmo energii zmienia się w czasie, do analizy pobiera się uśrednione wartości sygnałów w przedziałach czasowych, określonych częstotliwością próbkowania, którą przyjmuje się zwykle od 40 do 100 Hz. Analizator kanałowy pokrywa pasmo częstotliwości od ok. 100 do 7000 Hz przy liczbie kanałów wynoszącej od 10 do 20.

Metoda analizy widmowej daje lepsze wyniki niż metoda porównywania wartości chwilowych sygnałów i nie wymaga tak dużej pojemności pamięci, ze względu na mniejszą liczbę informacji potrzebnych do zdefiniowania sygnału.

c) Metoda analizy przez syntezę

Metoda ta jest pewną odmianą rozpoznawania przez porównywanie rozkładu widmowego energii sygnałów. Polega ona na porównaniu widma sygnału badanego z widmem syntetycznym wytworzonym przez maszynę cyfrową. Układ detekcji błędu przesyła do maszyny sygnał powodujący powstanie nowego widma, które jest ponownie porównywane z widmem sygnału wejściowego. Cykl ten jest powtarzany aż do osiągnięcia wymaganej zgodności obydwu sygnałów. Metoda ta jest również czasochłonna

i jest jeszcze słabo opanowana. Autorzy niektórych publikacji [12] wróżą jej jednak powodzenie, szczególnie w systemach samouczących się, tj. takich, w których algorytmy porównywania i podejmowania decyzji modyfikowane są automatycznie w trakcie rozpoznawania coraz to nowych elementów mowy.

d) Analiza formantowa

Metoda ta stosowana jest w rozpoznawaniu elementów mowy, charakteryzujących się obecnością w swoim widmie kilku niezbyt szerokich przedziałów częstotliwości, w których zawarta jest większość energii sygnału, tzw. formantów. Stosuje się ją do rozpoznawania fonemów o strukturze formantowej, np. samogłosek. W pamięci maszyny przechowuje się wzorce rozpoznawanych fonemów, przy czym ze względu na małą liczbę formantów (2 do 4) charakteryzujących fonemy, pojemność pamięci może być znacznie mniejsza niż w innych metodach rozpoznawania. Podobnie jak w metodach wg pkt. b) i c) sygnał wejściowy poddawany jest próbkowaniu z częstotliwością $40 \div 100$ Hz, a porównanie dotyczy wartości uśrednionych.

Metoda ta daje dobre wyniki przy rozpoznawaniu fonemów o zdecydowanie formantowej strukturze, szczególnie w tym przypadku, gdy formanty jednego fonemu są zdecydowanie inne niż formanty innych fonemów. Nie nadaje się ona natomiast do rozpoznawania fonemów nie mających struktury formantowej.

e) Metoda analizy przejść przez zero

Oprócz wyżej wymienionych metod wykorzystuje się pomiar ilości przejść przez zero sygnału elektrycznego odpowiadającego przebiegowi ciśnienia akustycznego w jednostce czasu (np. w przeliczeniu na 1 s). Analiza ilości przejść przez zero może być również użyteczna np. przy wyznaczaniu częstotliwości formantów w metodzie analizy formantowej oraz rozkładu widmowego energii w metodzie analizy widmowej, przy czym może ona dotyczyć sygnału oryginalnego lub poddanego odpowiednim przekształceniom (np. przebieg scałkowany, zróżniczkowany, przepuszczony przez filtr pasmowy itp.). Metoda analizy ilości przejść przez zero może być stosowana samodzielnie lub w połączeniu z jedną z innych me-

tod. W tym ostatnim przypadku umożliwia ona zróżnicowanie elementów mowy mających podobne inne parametry np. dwa fonemy nierozróżnialne metodą formantową mogą mieć inną ilość przejść przez zero .

4. Przykłady rozwiązań

W początkowym okresie badań nad ARM usiłowano skonstruować proste urządzenia sterowane głosem. Rozpoznawanie przeprowadzano dostępnymi w tym okresie środkami. Do porównania sygnału analizowanego i wzorcowego używano układów zawierających duże liczby elementów, za pomocą których realizowano urządzenia rozpoznające, działające wg sztywno ustalonych algorytmów (potocznie - realizacja hardwareowa). Rozpoznawano przede wszystkim liczby od 0 do 9 oraz kilka prostych wyrazów przydatnych w konkretnym eksperymencie. Jedne z pierwszych prac dotyczyły automatycznego wybierania numeru w telefonii [13] i prowadzone były między innymi w firmie Bell oraz w Massachusetts Institut of Technology - USA. Dokładność rozpoznawania liczb wynosiła 97 do 99% po przystosowaniu urządzenia do konkretnego użytkownika i spadała do 50÷60 % dla nieznanymi abonentów. W następnej kolejności pojawiły się maszyny piszące pod dyktando, rozpoznające do 10 sylab [13]. W 1961 roku ukazał się opis maszyny piszącej pod dyktando zawierającej słownik wejściowy złożony ze 100 sylab.

Zastosowanie elektronicznych maszyn cyfrowych do automatycznego rozpoznawania mowy umożliwiło zwiększenie liczby rozpoznawanych sygnałów akustycznych oraz osiągnięcie większej dokładności rozpoznawania.

Większa pojemność pamięci maszyn cyfrowych pozwala przechowywać większą ilość informacji o cechach dystynktywnych rozpoznawanych sygnałów, a większa szybkość liczenia umożliwia stosowanie bardziej złożonych lecz efektywniejszych algorytmów porównania i podejmowania decyzji. Dodatkową zaletą maszyn cyfrowych jest wspomniana wyżej możliwość samouczenia się.

W literaturze spotyka się wiele opisów prób automatycznego rozpoznawania mowy za pomocą maszyn cyfrowych.

W 1961 r. ogłoszono wyniki badań przeprowadzonych w Nowosybirskim oddziale AN Związku Radzieckiego [14]. Eksperyment dotyczył rozpoznawania 203 słów zawierających terminy języka ALGOL-60 oraz innych, wybranych wyrazów rosyjskich. Rozpoznawanie przeprowadzono w wyciszonym pomieszczeniu używając do tego celu komputera BESM-6. Analizowany sygnał podawany był do urządzeń wejściowych pięciu filtrów pasmowych RLC o częstotliwościach środkowych 225, 450, 900, 1800, 7200Hz i dobiegci 2,45. Okres próbkowania wynosił 1,4 ms. W każdym okresie próbkowania obliczano energię sygnału w całym paśmie częstotliwości E_0 oraz sygnałów otrzymywanych przez urządzenia wyjściowe każdego z ww filtrów E_1 do E_5 . Wartości $\ln E_0/E_1$ ($i = 1$ do 5) przyjęto jako parametry charakteryzujące badany segment wyrazu. Cały wyraz przedstawiał więc zbiór punktów w przestrzeni pięciowymiarowej odpowiadających wszystkim segmentom wyrazu. Za pomocą specjalnego algorytmu obliczano najmniejszą odległość we wspomnianej przestrzeni między analizowanym wyrazem i przechowywanymi w pamięci maszyny wzorcami wszystkich 203 wyrazów. Średnia dokładność rozpoznawania w eksperymencie obejmującym 5000 prób wynosiła ok 95% dla specjalnie dobranego lektora. Czas potrzebny na analizę wynosił $2 \div 4$ s.

W Istituto Electronico Nazionale G.Ferraris and Politecnico di Torino [4] zbudowano urządzenie rozpoznające 15 słów włoskich (liczby i pięć innych wyrazów) oparte na zasadzie analizy widmowej. Analityzator składał się z dziesięciu filtrów pasmowych o częstotliwościach środkowych 400, 500, 630, 800, 1000, 1250, 1600, 2000, 2500 i 4000Hz. Okres próbkowania wynosił 15 ms. Maszyna cyfrowa obliczała odległość w tzw. przestrzeni Hamminga [4] między wyrazem rozpoznawanym i wzorcowym. Rozpoznawanie przeprowadzane było na bieżąco. Dokładność rozpoznawania po dostosowaniu urządzenia do konkretnej osoby wynosiła ok. 99% oraz spadała do ok. 90% w eksperymencie z pięcioma różnymi osobami.

W laboratoriach firmy IBM [1] przeprowadzono eksperyment z rozpoznawaniem 33 wybranych wyrazów używając do tego celu komputera IBM 1130. Sygnał wejściowy poddawano próbkowaniu z częstotliwością 100Hz. W każdym okresie próbkowania był on charakteryzowany trzema wielkościami: parametrem proporcjonalnym do średniej wartości napięcia sy-

gnału, ilością przejść przez zero sygnału oraz ilością przejść przez zero pierwszej pochodnej sygnału otrzymanego na wyjściu filtra dolno-przepustowego o częstotliwości granicznej 2600Hz. Czas potrzebny na wyliczenie wynosił ok. 250 ms. Uzyskano dokładność rozpoznawania ok. 98% w przypadku tzw. przyuczonego lektora oraz ok. 85% dla lektorów dowolnych (przypadkowych).

Przykłady innych urządzeń opisane są przez: D.R.Hilla [7], I.B.Millara [11] i innych.

Prace dotyczące pośrednio lub bezpośrednio automatycznego rozpoznawania mowy rozpoczęto w Polsce już w latach pięćdziesiątych w Zakładzie Akustyki Cybernetycznej IPPT-PAN [8]. Celem ich było uzyskanie opisu głosek polskich za pomocą parametrów fonetyczno-akustycznych. Wynikiem prac jest prawie kompletna klasyfikacja polskiego zestawu fonemów. Badania te dotyczyły zagadnień uważanych powszechnie za podstawowe w dziedzinie automatycznego rozpoznawania. W latach sześćdziesiątych w IPPT-PAN przeprowadzono liczne badania eksperymentalne oraz prace teoretyczne związane z ARM. Dotyczyły one między innymi rozpoznawania samogłosek polskich metodą analizy widmowej [9] na podstawie geometrycznego modelu rozpoznawania. Wyniki tych i innych badań opublikowane zostały w Pracach IPPT [5],[6],[8],[9].

W chwili obecnej prace nad ARM prowadzone są w IPPT-PAN, na Uniwersytecie Warszawskim oraz w większości ośrodków naukowo-badawczych związanych z akustyką i informatyką.

5. Zakończenie

W chwili obecnej możliwe jest skonstruowanie urządzenia rozpoznającego ograniczoną liczbę słów (kilkadziesiąt do kilkuset) z dokładnością zbliżoną do 100%. Ze względu jednak na duże koszty budowy takich urządzeń oraz ich małą elastyczność (przystosowanie urządzenia do jednej osoby mówiącej, mała liczba rozpoznawanych słów) nie są one dotychczas rozpowszechnione. Wśród specjalistów z dziedziny ARM panuje pogląd, że w niedalekiej przyszłości możliwe będzie zbudowanie systemu ARM o znacznie większych możliwościach.

Niektórzy specjaliści zachodni przewidują skonstruowanie pod koniec lat siedemdziesiątych systemu zdolnego do rozpoznawania prawie wszystkich słów języka angielskiego z dokładnością zbliżoną do 100%.

Dalszy postęp w dziedzinie ARM uzależniony jest od rozwiązania konkretnych problemów, takich jak: wybór cech dystynktywnych, normalizacja i segmentacja sygnału, opracowanie optymalnych wzorców rozpoznawanych sygnałów, wybór najbardziej efektywnych algorytmów porównania i podejmowania decyzji. Większość działających obecnie systemów ARM to urządzenia laboratoryjne przeznaczone do badań ww zagadnień. Dalszy rozwój maszyn cyfrowych będzie miał z pewnością bardzo duży wpływ na postęp w dziedzinie automatycznego rozpoznawania mowy. Bardziej pojemne pamięci, większa skala integracji oraz związana z tym większa szybkość działania maszyn cyfrowych w połączeniu z jeszcze doskonalszym oprogramowaniem umożliwi z pewnością osiągnięcie znacznie lepszych rezultatów.

6. Literatura

- [1] BAGLEY J.D.: A Voice Input Facility for Digital Computers. IFIP Congres, 1971.
- [2] BEZDEL W.: Some Problems in Man-Machine Communication Using Speech. Intern.Journ.of Man-Mach.Stud.1970, nr 4.
- [3] BROAD D.J.: Formants in Automatic Speech Recognition. Intern. Journ.of Man-Mach.Stud.2972, nr 4.
- [4] De Mori: A. Flexible Real-time Recognizer of Spoken Words for Man-Machine Communication. Intern.Journ.of Man Machine Studies. 1970, nr 4.
- [5] GUBRYNOWICZ R.: Ocena parametrów widmowych samogłosek metodą przejść przez zero. Prace IPPT, 1971.
- [6] GUBRYNOWICZ R.: Analiza przejść przez zero sygnału mowy o harmoniczej strukturze widmowej. Prace IPPT, 1972.
- [7] MILL D.R. : An Abbreviated Guide to Planning for Spech Interaction with Machines: the state of the Art. Intern.Journ. of Man-Mach.Stud. 1972, nr 4.
- [8] KACPROWSKI J.: Akustyczny sygnał mowy w komunikacji człowiek-komputer. Prace IPPT, 1972.

- [9] KACPROWSKI J., GUBRYNOWICZ R.: Automacyjne rozpoznawanie samogłosek polskich metodą segmentacji widma. Prace IPPT, 1967.
- [10] KING I.H.: Some Experiments in spoken Word Recognition. IBM Journ. of Res. and Dev. 2966, nr 1.
- [11] MILLAR I.B.: An Interactive Speech Processing System Using a Large Computer. Intern. Journ. of Man-Mach. Stud. 1972, nr 3.
- [12] REDDY D.R.: Speech Recognition. Prospects for the Seventies IFIP Congress, 1971.
- [13] SAPOŹKOV M.A.: Sygnał mowy w telekomunikacji i cybernetyce. Warszawa, 1966 WNT.
- [14] VELICHKO V.M., ZAGORUYKO N.G.: Automatic Recognition of 200 Words. Intern. Journ. of Man-Mach Stud. 1970, nr 3.
- [15] VINCIUK T.K.: Raspoznavanie slov ustnoj reči metodami dinamičeskogo programirovanija. Kibernetika A.N. SSRR.

mgr inż. Zbigniew KĘDZIOR
Instytut Maszyn Matematycznych

681.327.12:681.327.54:681.327.57

URZĄDZENIA ROZPOZNAJĄCE ZNAKI ALFANUMERYCZNE ZAPISANE
NA NOŚNIKU PAPIEROWYM I FILMOWYM

1. WSTĘP

Informatyka jest szczególnie zainteresowana urządzeniami, które potrafią odczytywać znaki alfanumeryczne bezpośrednio z dokumentów pierwotnych. Do olbrzymich ilości danych przetwarzanych przez maszyny cyfrowe konieczne są automaty czytające, które zastąpiłyby człowieka w procesie przygotowywania danych, tzn. skróciły czas wprowadzania danych, w wyniku czego uzyskuje się dodatkowe efekty ekonomiczne.

Z punktu widzenia zasady pracy, urządzenia odczytujące znaki alfanumeryczne można podzielić na trzy następujące grupy [1]:

- I. Optyczne czytniki znaków
- II. Czytniki znaków zapisanych atramentem magnetycznym
- III. Czytniki znaków zapisanych na nośniku filmowym

Największym zainteresowaniem cieszą się czytniki grupy I, zwane również urządzeniami optycznego rozpoznawania znaków [ORZ]. Jest to najliczniejsza grupa z punktu widzenia liczby typów. Na rynkach krajów zachodnioeuropejskich, Stanów Zjednoczonych i Japonii jest obecnie dostępnych ponad 50 typów tego rodzaju czytników. Produkcją tego rodzaju urządzeń zajmuje się około trzydziestu firm. Urządzenia te odczytują znaki pisma maszynowego stylizowanego oraz zwykłych maszyn do pisania i drukarek wierszowych. Wiele z nich odczytuje znaki numeryczne ręczne pisane wg odpowiednich reguł (tzw. ręczne drukowanie - handprinting). W tabeli 1 podano parametry kilku wybranych typów

urządzeń ORZ produkcji firm amerykańskich (dane z lat 1972 - 1973). W ZSRR również wyprodukowano kilka typów czytników znaków alfanumerycznych, jak np. czytnik RUTA 701 i Sever 3.

Oprócz wyżej wymienionych, do grupy I można zaliczyć czytniki znaków uzupełnionych kreskami kodującymi oraz czytniki kresiek. Urządzenia te rozwinęły się w okresie kiedy technika ORZ była na niskim poziomie. Czytniki kresiek są rozpowszechnione w krajach zachodnioeuropejskich, w Stanach Zjednoczonych i Kanadzie, jednak należy się spodziewać, że w przyszłości zmniejszy się popyt na tego rodzaju urządzenia ze względu na dynamiczny rozwój techniki optycznego rozpoznawania znaków alfanumerycznych.

Czytniki grupy II powstały również mniej więcej w tym samym czasie co czytniki kresiek i znaków z kreskami kodującymi. W tych czytnikach technika zapisu zapewnia większą niezawodność odczytu, ale z drugiej strony zakłada bardzo wysokie wymagania dotyczące jakości druku, co utrudnia i powiększa koszty eksploatacji. Ceny urządzeń komercyjnych są stosunkowo wysokie i wahają się w granicach od 34000 do 123000 dolarów. Znaki pism magnetycznych są trudno czytelne dla człowieka. Urządzenia te obecnie nie cieszą się zainteresowaniem. Tradycyjnym zastosowaniem ich jest odczyt czeków w bankowości. W ostatnich latach nie spotyka się nowych typów tego rodzaju urządzeń.

Najbardziej nowoczesnymi urządzeniami są czytniki grupy III. Oznaczają się one wysoką dokładnością odczytu znaków alfanumerycznych oraz możliwością rozpoznawania informacji graficznej. W zakresie rozpoznawania znaków stosuje się tu taką samą technikę jak w urządzeniach grupy I. Liczba rodzajów czytanych pism jest bardzo duża; prędkość odczytu stosunkowo duża. Pewnym utrudnieniem jest tu przenoszenie na nośnik filmowy, tzn. filmowanie dokumentów, ale z drugiej strony w ten sposób zapisana informacja jest łatwa do przechowywania. Czynnikiem utrudniającym ich rozpowszechnienie są przede wszystkim bardzo wysokie ceny. Znane są tylko dwie firmy, które produkują te urządzenia: Compuscan (pierwszy egzemplarz wyprodukowano w 1970 r.) i Information International (pierwszy egzemplarz z 1971 r.).

Reasumując powyższe można powiedzieć, że do nowoczesnych i jednocześnie perspektywicznych należy zaliczyć tylko optyczne czytniki znaków alfanumerycznych zapisanych na nośnikach papierowych i filmowych. Dlatego w dalszym ciągu niniejszego opracowania będzie mowa tylko o tego rodzaju urządzeniach.

2. PRZYCZYNY MAŁEGO ROZPOWSZECHNIANIA CZYTNIKÓW

Czytnikami utrudniającymi rozpowszechnienie urządzeń ORZ są wysokie ceny oraz wymagania dotyczące jakości druku. W 1970 r. znajdowało się w eksploatacji około 1000 urządzeń rozpoznających znaki numeryczne i alfanumeryczne. Ceny ich wahają się w granicach od 25 tysięcy do 1 mln dolarów. Natomiast ceny wspomnianych dwóch typów urządzeń mikrofilmowych (bez opcji) wynoszą odpowiednio: 900 tys. dolarów-Compuscan 370 OCR System i 1,5 mln dolarów-Grafix Reader firmy Information International. Wymagania dotyczące wysokiej jakości pisma zniechęcają użytkowników. Dla poprawienia parametrów eksploatacyjnych producenci stosują pisma stylizowane. Podobnie, jak dla czytników znaków zapisanych atramentem magnetycznym lub czytników znaków z kreskami kodującymi, pisma stylizowane utrudniają rozpowszechnianie urządzeń rozpoznających znaki.

Użytkownicy muszą odczytywać dokumenty wypełniane na specjalnych maszynach do pisania. Tymczasem producenci stosują do dziś różne rodzaje pism stylizowanych [2] . W większości współcześnie produkowanych czytników należy stosować specjalne formularze (zawierające cyfry kontrolne, pola odczytu, symbole sterujące lub kreski kasujące). Przy zmianie rodzaju odczytywanych danych należy zaprojektować nowy typ formularza. Wymagania dotyczące jakości papieru również są wysokie.

3. PODZIAŁ URZĄDZEŃ I ICH ZASTOSOWANIE

Z użytkowego punktu widzenia optyczne czytniki znaków zapisanych na nośniku papierowym dzielą się na:

- a) czytniki dokumentów
- b) czytniki stron

- c) czytniki stron i dokumentów
- d) czytniki taśmy rejestracyjnej.

Czytniki dokumentów są urządzeniami, które odczytują informację z wybranych miejsc (tzw. pól) formularza. Informacja odczytywana obejmuje przeważnie nie więcej niż 200 znaków; pola odczytu znajdują się w jednym lub dwu wierszach. Parametrem charakteryzującym szybkość działania jest prędkość czytanych dokumentów, która w zależności od typu urządzeń waha się w granicach od 60 do 1200 dok/min. W urządzeniach tych stosuje się przeważnie odczyt znaków numerycznych.

Natomiast czytniki stron są urządzeniami odczytującymi wszystkie znaki występujące na dokumentach. Istotnym parametrem jest tu prędkość czytania znaków ($400 \div 3600$ zn/s). Czytniki stron rozpoznają znaki alfanumeryczne. W grupie tej spotyka się urządzenia uniwersalne odznaczające się zdolnością rozpoznawania wielu rodzajów pisma maszynowego i znaków numerycznych pisma ręcznego, możliwością odczytu dokumentów o różnych formach oraz wieloma innymi zaletami.

Powyższe dwa rodzaje urządzeń różnią się zasadniczo konstrukcją mechaniczną. Wykorzystuje się je w różnych dziedzinach.

Dla użytkowników o dużym wachlarzu zastosowań produkuje się urządzenia uniwersalne (grupa C), które albo mają dwa mechanizmy transportujące albo bardzo szybkie mechanizmy przystosowane do transportu małych i dużych dokumentów. Czytniki grupy d odczytują znaki zapisane na taśmie kas rejestracyjnych lub arytmometrów; rozpoznają tylko znaki numeryczne maszynowe i kilka symboli specjalnych. Mają ograniczone zastosowanie. Stąd liczba typów dostępnych na rynku jest mała (natomiast wiele urządzeń z grup a, b, c jest wyposażonych - jako opcja - w mechanizm do transportu taśmy).

Czytniki znaków zapisanych na nośnikach filmowych są czytnikami stron. Urządzenia ORZ stosowane są najliczniej (I dziedzina zastosowań) do przetwarzania dowodów opłaty (turnaround applications) takich jak:

- czeki kredytowe
- dowody opłaty za energię elektryczną i gaz

- dowody uiszczenia podatków, składek ubezpieczeniowych, subskrypcji
- bilety lotnicze.

Drugą dziedziną zastosowań jest przetwarzanie dokumentów powstających w ramach jednej instytucji (in-house applications), w której jest ośrodek obliczeniowy (banki, przedsiębiorstwa przemysłowe, instytucje użyteczności publicznej).

W Stanach Zjednoczonych, w Wielkiej Brytanii i w RFN istnieją biura usług ORZ. Są to ośrodki przygotowywania danych z dokumentów dostarczanych przez klientów. Dane odczytane są rejestrowane na taśmach magnetycznych, taśmach papierowych lub kartach dziurkowanych.

Trzecią dziedziną są zastosowania terenowe (field applications). Tutaj przedstawiciele firm (agenci, komiwojaźerowie) wypełniają określone formularze, które są odczytywane maszynowo w centrali.

4. POTENCJALNI NABYWCY W POLSCE

Niektóre instytucje państwowe (urzędy statystyczne, przemysł poligraficzny, poczta) interesują się nabyciem urządzeń odczytujących w sposób automatyczny informacje zapisane na różnego rodzaju dokumentach i pismach. Inne instytucje i urzędy ze względu na znaczne ilości przechowywanych i przetwarzanych u nich informacji można uważać za potencjalnych nabywców i użytkowników czytników znaków.

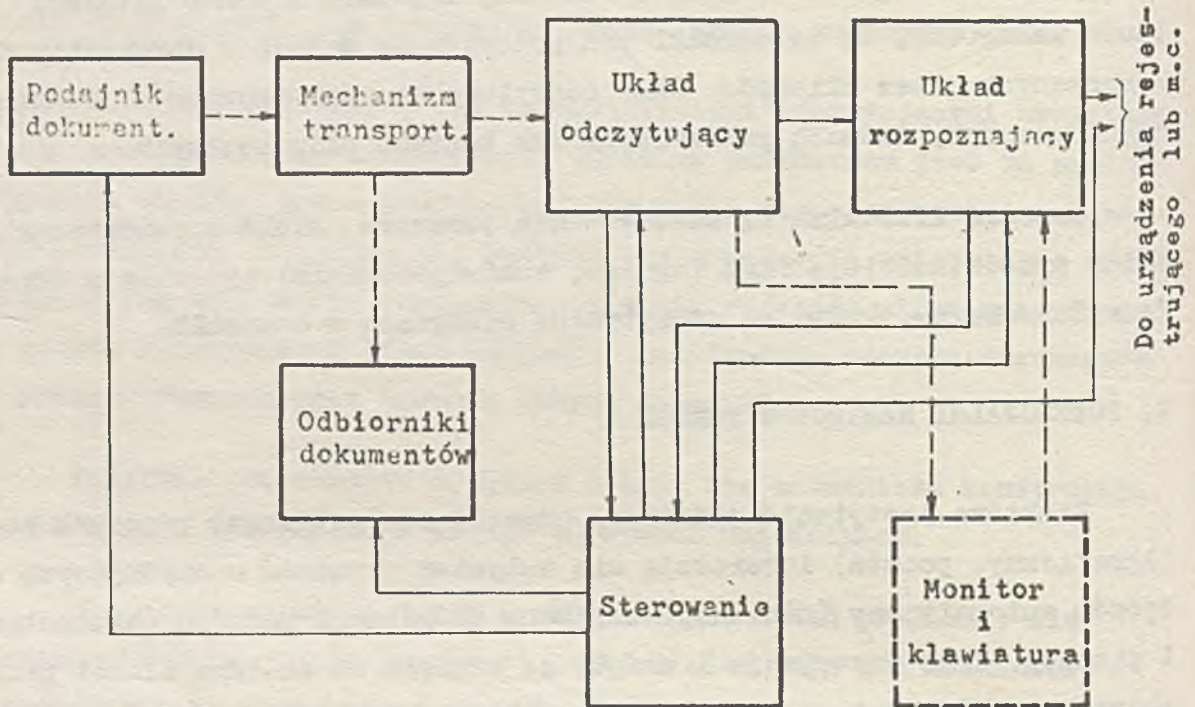
Zdaniem autora instytucje takie jak Urząd Patentowy PRL, poligrafia i duże biblioteki powinny dysponować urządzeniami odczytującymi znaki zapisane na mikrofilmie. Poczta, duże ośrodki obliczeniowe, domy towarowe i biura turystyczne powinny stosować bezpośredni odczyt z dokumentów. Natomiast urzędy statystyczne powinny wykorzystywać zarówno urządzenia odczytujące dokumenty, jak również informację naniesioną na nośnik filmowy.

Należałoby wymienić jeszcze jedną dziedzinę zastosowań ORZ, a mianowicie wykorzystanie układów odczytujących (patrz p. 5.1.) lub kompletnych czytników znaków w czytelniach dla niewidomych.

5. ZASADY BUDOWY URZĄDZEŃ ORZ

5.1. Układy elektroniczne

Na rys. 1 przedstawiono uproszczony schemat blokowy czytnika. Podstawowymi zespołami elektronicznymi optycznego czytnika znaków są: receptor czyli układ odczytujący oraz klasyfikator czyli układ rozpoznający. W dalszym ciągu zostaną omówione rodzaje spotykanych rozwiązań ww układów.



Rys. 1. Schemat blokowy optycznego czytnika znaków alfanumerycznych

U k ł a d y o d c z y t u j ą c e

Receptory współcześnie budowanych optycznych czytników znaków wykorzystują następujące techniki odczytu:

- względny ruch dokumentów pod kolumną lub siatkówką fotoelementów,
- oświetlanie punktów rastru znaku promieniem laserowym lub wąską wiązką światła niekoherentnego prowadzoną systemem ruchomych lusterek, względnie oświetlanie za pomocą wirującej tarczy z otworami,
- oświetlanie punktów rastru wąską wiązką światła lampy promieniowej (metoda ruchomej plamki),

- d) odczyt za pomocą lampy akumulującej,
- e) śledzenie konturu znaku.

Technika odczytu wymieniona w punkcie a) daje dużą prędkość odczytu, ale jest kosztowna. Rozwiązanie typu b) należy do najtańszych, natomiast prędkość odczytu jest nieduża.

Techniki a) i b) należą do najpopularniejszych. Metoda ruchomej plamki pozwala na uproszczenie mechanizmu, natomiast jest metodą bardzo kosztowną, podczas gdy prędkość odczytu jest średnia (np. w czytniku M6000 Philco-Ford wynosi 1250 zn/s przy odczycie znaków alfanumerycznych).

W metodzie d) prędkość odczytu zależy od rodzaju zastosowanej lampy, którą może być albo widikon albo lampa typu dysektor obrazowy. W pierwszym przypadku prędkość jest mała (250-500 zn/s), a w drugim średnia (2000 zn/s). Metoda ta nie jest szeroko rozpowszechniona. W czytnikach będących obecnie w sprzedaży nie spotyka się rozwiązań z lampami typu widikon. Natomiast tylko w jednym czytniku (typu 20/20 firmy Seean Optics) wykorzystuje się dysektor obrazowy (lampę zbudowaną przede wszystkim do optycznego rozpoznawania).

Metoda e) daje średnią prędkość odczytu. Przede wszystkim znajduje ona zastosowanie w rozpoznawaniu znaków pisma ręcznego.

U k ł a d y r o z p o z n a j ą c e

Metody rozpoznawania stosowane w czytnikach dostępnych na rynkach krajów zachodnich można podzielić na dwie grupy:

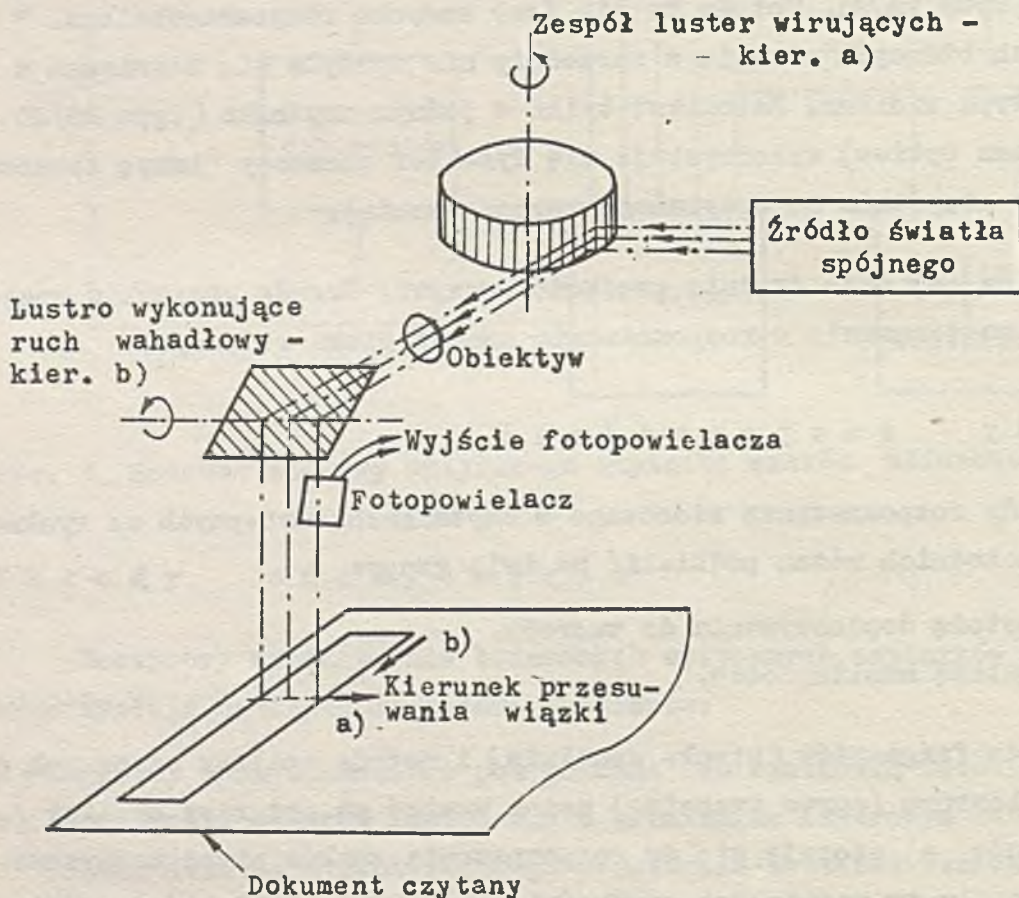
- a) metodę dopasowywania do wzorców,
- b) metodę analizy cech.

Metodę fragmentów (stroke analysis) i metodę analizy stycznych do odcinków konturu (curve tracing) można uważać za podgrupy analizy cech. Metodę a) stosuje się do rozpoznawania znaków pisma maszynowego. Osiąga się tu największe prędkości odczytu (do 3600 zn/s). Jest bardzo popularna. Wadą jej jest brak możliwości rozpoznawania znaków ręcznych. Większość metod z grupy analizy cech jest stosowana do roz-

poznawania znaków ręcznych lub w urządzeniach odczytujących kilka rodzajów pism. Wyjątek stanowi metoda fragmentów, która nie nadaje się do analizy znaków ręcznych. W niektórych czytnikach stosuje się jednocześnie dwie metody: jedną do rozpoznawania znaków maszynowych (a), a drugą do ręcznych (b).

Przykład rozwiązania

Poniżej przedstawiamy skrócony opis czytnika typu H8959 firmy Hitachi (Japonia) [2], którego model został wyprodukowany w 1973 r. Jest to czytnik stron rozpoznający znaki numeryczne pisma ręcznego lub znaki alfanumeryczne maszynowe. Stanowi on przykład prostego i jednocześnie nowoczesnego rozwiązania. Receptor czytnika jest typu elektromechanicznego, w którym wykorzystuje się źródło światła spójnego (laser).



Rys. 2. Schemat kinetyczny elektromechanicznego układu odczytującego występującego w czytniku H8959 firmy Hitachi

Na rys. 2 przedstawiono schemat kinetyczny układu odczytującego. Odpowiednie ustawienie zwierciadeł w zespole lusterek wirujących oraz dobór prędkości tego zespołu do okresu wahan zwierciadła wykonującego ruch wahadłowy powodują kolejne (w czasie) przesuwanie się bardzo wąskiej wiązki światła emitowane przez laser, w taki sam sposób, jak to ma miejsce w systemie telewizyjnym. Strzałki na rysunku pozwalają prześledzić kierunek przesuwania się promienia laserowego po dokumencie. Światło odbite od papieru jest odbierane przez fotopowielacz, a następnie wzmacniane i przetwarzane na sygnały cyfrowe.

Zastosowanie źródła koherentnego w technice odczytu daje wiele korzyści, wśród których najważniejszą jest poprawienie stosunku sygnał/szum. Pozwala to lepiej rozdzielać miejsca zaczerpnięte od niezaczerpniętych, jak również nie wymaga stosowania złożonych wielostopniowych wzmacniaczy. Powyższe rozwiązanie receptora oznacza się niskimi kosztami. Prędkość odczytu jest nieduża ze względu na wykorzystanie mechanicznego układu odchylającego.

W układzie rozpoznającym wyróżnia się dwa zasadnicze zespoły, a mianowicie:

- zespół przetwarzania wstępnego
- właściwy układ rozpoznający

Przy rozpoznawaniu znaków pisma ręcznego w zespole przetwarzania wstępnego realizowane są między innymi następujące funkcje:

- normalizacja wymiarów
- uzupełnienie przerw w linii znaku
- usuwanie miejsc zaczerpniętych nie związanych ze znakami
- pocienianie

Ww funkcje wykonują układy logiczne. Wyniki działania tych układów można zilustrować na przykładzie obróbki znaku "2" (rys. 3 a i b).

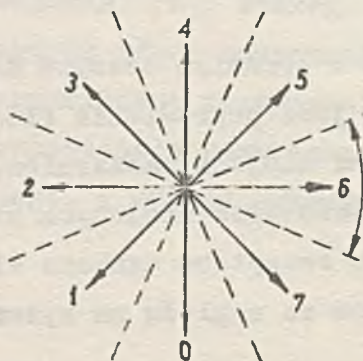
Sygnały obrazu pocienionego wprowadza się do zespołu rozpoznającego, w którym dokonuje się podziału linii znaku na odcinki o zbliżonym współczynniku kierunkowym stycznej do krzywej oraz przydziela tzw. łańcuch kodu i ostatecznie rozpoznaje. Wszystkie możliwe kierunki sty-



a) Macierz sygn. wejściowych



b) Macierz sygn. pocienionych



c) Sposób oznaczania kierunków



Łańcuch: = [6,7,1,6]

d) Sposób kodowania rozpoznawanych znaków

Rys. 3. Graficzna ilustracja funkcji realizowanych w układzie rozpoznającym znaki pisma ręcznego, w czytniku H8959

cznej do linii znaku redukuje się do ośmiu kierunków róży wiatrów, którym przyporządkowuje się cyfry od 0 do 7. W rezultacie działania specjalnej sieci logicznej każdemu rozpoznawanemu znakowi zostaje przydzielony skończony ciąg (czyli łańcuch) cyfr, na podstawie którego zostaje dokonane rozpoznanie znaku. Sposób budowy łańcucha kodowego obrazuje rys. 3 (c i d).

Powyższa metoda rozpoznawania znaków ręcznych obejmująca m.in. funkcje takie jak eliminowanie zakłóceń (uzupełnianie przerw i odrzucanie plam) oraz pocienianie należy do nowoczesnych. Znaki ręczne nie mogą być pisane dowolnie. Oznacza to, że osoba wypełniająca dokument musi znać reguły pisania (rys 4).

Do rozpoznawania znaków pisma maszynowego w czytniku H8959 zastosowano konwencjonalną metodę dopasowywania masek (wzorców).

	Poprawnie	Źle										
Nie łącz ze sobą znaków	<table border="1"><tr><td></td><td></td><td>1</td><td>0</td><td>0</td></tr></table>			1	0	0	<table border="1"><tr><td></td><td></td><td>1</td><td>0</td><td>0</td></tr></table>			1	0	0
		1	0	0								
		1	0	0								
Linie znaków nie mogą mieć przerw	<table border="1"><tr><td>8</td><td>6</td><td>9</td><td>7</td><td>5</td></tr></table>	8	6	9	7	5	<table border="1"><tr><td>8</td><td>6</td><td>9</td><td>7</td><td>5</td></tr></table>	8	6	9	7	5
8	6	9	7	5								
8	6	9	7	5								
Kształt znaków powinien być prosty i bez "ogonków"	<table border="1"><tr><td>7</td><td>2</td><td>4</td><td>1</td><td>0</td></tr></table>	7	2	4	1	0	<table border="1"><tr><td>7</td><td>2</td><td>4</td><td>1</td><td>0</td></tr></table>	7	2	4	1	0
7	2	4	1	0								
7	2	4	1	0								

Rys. 4. Pouczenie dla osób wypełniających dokumenty pismem ręcznym

Przyjęty tu układ klasyfikujący rozpoznaje do 50 różnych typów znaków jednocześnie. Zbiór wzorców dla jednego typu pisma maszynowego jest zmontowany w formie jednego zespołu. W przypadku zmiany rodzaju pisma maszynowego należy zamienić ten zespół na inny.

Podstawowe parametry czytnika H8959

1) Rodzaje pism:

- znaki numeryczne ręcznie kaligrafowane oraz znaki C, S, T, X i Z
- znaki numeryczne maszynowe stylizowane typu OCR-A, OCR-B, 407 (IBM) i 12F (Farrington)
- Opcje
 - znaki alfanumeryczne stylizowane OCR-A
 - znaki numeryczne różnych typów maszyn do pisania

2) Urządzenie odczytuje jednocześnie (tzn. z jednego dokumentu) znaki numeryczne maszynowe i ręczne

3) Wymiary dokumentów od 145 x 95 do 300 x 220 mm

4) Maksymalna liczba w wierszu

- znaków ręcznych - 36
- znaków maszynowych - 72

5) Maksymalna liczba wierszy na dokumencie

- przy znakach ręcznych - 25
- przy znakach maszynowych - 29

- 6) Maksymalna prędkość odczytu
 - znaków ręcznych - 50 zn/s
 - znaków maszynowych - 100 zn/s
 - formularzy - 36 dok./min
- 7) Wyprowadzanie danych: taśma papierowa
- 8) Wymiary urządzenia w mm (bez dziurkarki taśmy papierowej):
1160 x 630 x 1360
 - waga: około 360 kg
 - wymiary dziurkarki taśmy papierowej - 490 x 425 x 340 mm
- 9) Dopuszczalna temperatura i wilgotność przestrzeni otaczającej urządzenie: 5 - 35°C i 30 - 85 % wilgotności względnej.

Z powyższych danych wynika, że parametry urządzenia nie dorównują parametrom wielu czytników firm takich, jak REJ, CDC, Scan Data Corporation lub IBM. Urządzenie H8959 ma ograniczone możliwości w zakresie liczby znaków przypadających na jeden dokument, w zakresie liczby rodzajów pism, oznacza się także małą prędkością odczytu. Z drugiej natomiast strony jest urządzeniem o małych wymiarach oraz wg opinii zalicza się do stosunkowo tanich.

5.2. Mechanizmy przenoszące dokumenty

Najbardziej skomplikowane są mechanizmy czytników dokumentów i czytników stron. W nowoczesnych urządzeniach stosuje się podajniki cierne, próżniowe oraz z kołami stożkowymi. W zakresie mechanizmów transportujących stosuje się bębny próżniowe. Nowoczesne odbiorniki wykorzystują mechanizmy bębnowe. Mechanizmy występujące we współczesnie produkowanych czytnikach są stosunkowo kosztowne. Istnieje opinia [1], że koszt ich produkcji stanowi 50 % kosztów całego czytnika.

6. PRACE PROWADZONE W POLSCE

Od kilku lat w Instytucie Organizacji i Kierowania PAN są prowadzone zarówno badania podstawowe, jak również pewne prace aplikacyjne z zakresu rozpoznawania. Rozwija się tu kierunek biocybernetyczny

zmierzający do skonstruowania modelu sieci sztucznych neuronów, naśladującej system nerwowy człowieka, jak również tzw. kierunek programowy zajmujący się poszukiwaniem ogólnych zasad rozpoznawania za pomocą prostych operacji realizowanych przez elektroniczną maszynę cyfrową lub wyspecjalizowane urządzenie. Prace te dotyczą wielu dyscyplin rozpoznawania obiektów, a w tej liczbie rozpoznawania znaków alfanumerycznych oraz informacji graficznej. Między innymi opracowano tu i uruchomiono model urządzenia odczytującego znaki alfanumeryczne pisma maszynowego z prędkością odczytu do 40 znaków na sekundę.

7. TENDENCJE W BUDOWIE URZĄDZEŃ ORZ

Do odczytu znaków zaczyna się stosować źródło światła koherentnego. Rozwiązanie to zapewnia przede wszystkim duży stosunek sygnał szum sygnałów analogowych odbieranych przez receptory, przez co zwiększa się dokładność odczytu, a także można stosować gorsze gatunki papieru jak również eliminować ciemnię w otoczeniu głowicy odczytowej. Rozwiązanie to stosuje się w czytnikach komercyjnych takich jak: System 70 firmy Cognitronics (1970 r.), 921 DR firmy CDC (1971 r) oraz opisany w pkt 5.1. czytnik stron typu H8959 firmy Hitachi (1973 r.).

Na konferencji na temat ORZ, która odbyła się w 1967 r. w Delft (Holandia) omawiano możliwość wykorzystania holografii do rozpoznawania znaków [3]. Badania w tym kierunku są prowadzone w wielu krajach. Jednak, jak do tej pory brak informacji o sprzedaży czytników wykorzystujących zjawisko holografii.

Ostatnio coraz częściej stosuje się system zwany rozpoznawaniem kontekstowym, w którym analizuje się rodzaj znaków sąsiadujących z rozpoznawanymi.

W nowoczesnych rozwiązaniach stosuje się dwustopniową analizę cech. Najpierw rozpoznaje się tzw. cechy małe (lokalne), a następnie przeprowadza się końcową identyfikację w ramach grupy znaków o zbliżonych kształtach. Stosowane jest również "douczenie" w przypadkach pojawiania się znaków o odmiennych kształtach lub też pełne uczenie przeprowadzane przez użytkownika.

Prowadzone są prace nad rozpoznawaniem znaków ręcznych alfabetu, czego wyrazem jest propozycja stworzenia normy dotyczącej znaków ręcznych [5] .

Wobec małego popytu, wskutek wysokich cen, producenci zaczęli budować znacznie tańsze urządzenia specjalizowane do konkretnych zastosowań.

8. PROPOZYCJE

Wobec wzrostu zastosowań komputerów w ostatnich latach w Polsce zwiększa się liczba potencjalnych nabywców czytników znaków alfanumerycznych. Ponieważ urządzenia tego rodzaju są bardzo kosztowne należałoby obecnie rozpocząć odpowiednie prace przygotowawcze, w wyniku których można by uzyskać żądane efekty techniczne przy możliwie najmniejszych kosztach. W tym celu należałoby określić jednolity krój czcionek dla wszystkich nowonabywanych maszyn do pisania z importu i produkcji krajowej, a zwłaszcza tych, które będą służyć do wypełniania dokumentów źródłowych do przetwarzania danych.

Należałoby ustalić listę typów czytników znaków, które mogą być produkowane w kraju lub będą sprowadzane z zagranicy.

Powinna powstać instytucja, która zajęłaby się dziedziną zastosowań urządzeń ORZ. Komórka ta między innymi powinna zająć się sprawą przyszłego zaopatrzenia użytkowników urządzeń w materiały piśmienne produkcji krajowej zapewniające wysoką jakość zapisu (formularze do odczytu maszynowego, gatunki papieru i taśmy maszyn do pisania).

Literatura

- [1] Auerbach Computer Technology Reports 1971-73 tom E2 raporty nr 280.0000.050 _ 280.7940.020.
- [2] KĘDZIOR Z.: Zastosowanie czytników znaków alfanumerycznych. Maszyny Matematyczne 1968, nr 12.
- [3] TAICHIRO SANO i TOSHIHIRO HANANOI: Model H-8959 Optical Character Reader. Hitachi Review, 1973, nr 8.
- [4] SEGER G.: Holographische Methoden zur Zeicherkennung. Informatik 1967, tom 5, nr 5.
- [5] Computer Weekly 1973, nr 330.

WYKAZ NIEKTÓRYCH TYPÓW OPTYCZNYCH CZYTNIKÓW ZNAKÓW ALFANUMERYCZNYCH PRODUKCJI FIRM AMERYKAŃSKICH

Tabela 1

Rozdaj	Typ	Producent	Podstawowe parametry						Zasada pracy	Rok wyprodukowania pierwszego egzemp.	Cena w dol.USA	
			Prędkość podawania dok. szt/min	Prędkość czytania znaki/s	Liczba znaków na dokumencie	Rodzaje czytanego pisma i znaków	Wyprowadzanie danych	Wymiary dokumentów (mm)			urządzenia	urządzenia
1	2	3	4	5	6	7	8	9	10	11	12	13
Czytnik dokumentów	2703 Document Reading Systems	UNIVAC	300;600	1500	1 wiersz do 80 znaków	OCRA i UNIVAC H 14 numeryczne + 4 znaki specjalne	bespoór.do komputera	od 67,7x76 do 108x222	Kolumna fotoelem.i dopasow.masek	1970	42 000	64 560
	921 Document Reader	CDC	1200	2210	1 wiersz; do 80 znaków	OCRA, 7B, OCRB - numeryczne	taśma magnetyczna	od 66x114 do 114x229	promień laserowy i dopasow.masek	1971	brak danych	61 000
Czytnik stron	1288 Optical Page Reader	JEM	14 328	1000	84 wiersze do 96 znaków/wiersz	OCRA alfanumeryczne; ręczne drukow.-numeryczne	bespoór.do komputera	od 76x166 do 227x355	metoda ruchomej plamki	brak danych	223 390	296 480
	Model 6 000 Print Reader	Philco-Ford	180-360	1250	78 wierszy; do 90 znaków/wiersz	kilka pism maszynowych, znaki alfanumeryczne litery duże i małe	taśma magnet. karty dziurkow. taśma papierowa besp.do komputera	od 127x178 do 216x229	metoda ruchomej plamki i dopasowanie masek	1965	400 000	brak danych
Mikrofilm czytnik stron	370 OCR System	Compucan	brak danych	400 4000	brak dokładnych danych	dowolne pismo maszynowe; znaki alfanumeryczne	bespoór.do komputera	taśma mikrofilmowa 16 lub 35 mm	metoda ruchomej plamki+dopasow.masek + analiza cech.	1970	900 000	brak danych
Czytnik stron i dokumentów	System/70	Cognitronics	do 60	250	dowolna liczba wierszy, do 76 znaków/wiersz	numeryczne maszynowe-Kilka repertuarów + numeryczne-ręcznie drukowane; opcje: OCRA lub OCRB - alfanumeryczne	taśma magnetyczna, taśma papierowa, karty dziurkowane	od 102x82,5 do 227x355	promień laserowy i analiza cech	1970	33 600	41 250
	Model 250	Scan Data	50; 180	800	81 wierszy- do 100 znaków/wiersz	1-5 rodzajów pisma maszynowego znaki alfanumeryczne małe i duże; więcej repertuarów -opcja	taśma papierowa, taśma magnetyczna, karty dziurkowane	od 127x76 do 280x356	metoda ruchomej plamki i analiza cech	1971	215 000	315 000
Czytnik taśmy rejestracyjnej	Model 4040	Lundy-Farrington	2000 wierszy/min	2000	do 32 znaków/wiersz	numeryczne 12 F, 7 B, OCRA, MCR NCF, 1428 /+ do 8 znaków specjalnych/	taśma magnetyczna	-	metoda ruchomej plamki	1969	brak danych	brak danych

mgr inż. Andrzej KOJEMSKI

68/322-181.4.004.14:68/327.8

mgr inż. Jerzy KASPRZYK

Instytut Maszyn Matematycznych

WYKORZYSTANIE MAŁYCH MASZYN CYFROWYCH W SIECIACH TELEINFORMATYCZNYCH

1. Wstęp

Obecny stan rozwoju wykorzystania komputerów uwidacznia coraz większą rolę środków służących przekazywaniu danych do i z komputera na większe odległości. Dziedzina telekomunikacji zajmująca się tymi zagadnieniami nazywa się transmisją danych (data transmission). Według ocen podawanych w literaturze (np [1]) ponad 60% komputerów w 1975 r. będzie pracowało w konfiguracjach związanych z transmisją danych. Wzrastająca liczba informacji w postaci publikacji, sprawozdań z konferencji itp. imprez świadczy o lawinowym narastaniu zainteresowania tą tematyką. Transmisja danych rozwija się, gdyż pozwala:

- udostępnić wielu użytkownikom dużą moc obliczeniową w wybranym przez nich miejscu i czasie (np. rozdział mocy obliczeniowej w wielodostępnych systemach abonenckich oraz możliwości jej skupienia w przypadku sieci komputerowych);
- przesyłać dane do tych miejsc, gdzie są one potrzebne (np. systemy informacyjne z bankami danych, systemy zbierania danych);
- wymieniać programy i specjalizowane usługi obliczeniowe.

Wszystkie te czynniki sprawiają, że oprócz tak podstawowych korzyści jak rozpowszechnienie dostępu do środków obliczeniowych i zwiększenie ich wydajności, uzyskuje się zupełnie nowe możliwości wykorzystywania sprzętu komputerowego. Przykładem mogą tu być np. systemy rezerwacji biletów lotniczych o zasięgu międzynarodowym oraz przystaw-

ki ze sprzężeniem akustycznym [2] pozwalające na współpracę z komputerem z budki telefonicznej. Najogólniej, przestrzennie rozmieszczony sprzęt w postaci urządzeń informatycznych, połączony za pomocą linii transmisji danych tworzy sieć, którą będziemy nazywali siecią teleinformatyczną. Jeśli w węzłach takiej sieci występują komputery, wówczas mówimy o sieci komputerowej.

Z ogólnego problemu powiązania komputerów z sieciami teleinformatycznymi przedmiotem naszego zainteresowania w tym opracowaniu są przede wszystkim małe komputery, zwane minikomputerami. Podstawowym kryterium zaliczenia procesora do klasy minikomputerów jest jego niska cena (poniżej 25 tys. dolarów), związana z długością słowa zawartą w zakresie od 8 do 16 bitów, uproszczoną listą rozkazów oraz rozwiązaniem technologicznym opartym na układach scalonych.

W sieciach teleinformatycznych typowe wykorzystanie minikomputerów polega na stosowaniu ich w charakterze procesorów satelitarnych, spełniających podstawową rolę w realizacji procesu przesyłania danych. Ze względu na taką rolę, minikomputer jest wówczas najczęściej nazywany procesorem komunikacyjnym. Większe maszyny cyfrowe stosowane w sieciach teleinformatycznych spełniają właściwą rolę urządzeń przetwarzania danych i będą przez nas nazywane głównymi komputerami.

Należy zaznaczyć, że stosowanie minikomputerów do budowy sieci teleinformatycznych nie jest konieczne. Istnieją przykłady rozwiązań takich sieci bez wykorzystywania minikomputerów, jak np. sieć typu TSS (Time Sharing System) wykorzystująca sprzęt firmy IBM [3]. Niemniej występują tendencje do wykorzystywania minikomputerów coraz szerzej w sieciach teleinformatycznych, związane z dążnością do oddzielenia w sieciach zadań komunikacyjnych od zadań przetworzeniowych. Tendencje takie są zgodne z kierunkami rozwoju architektury komputerów, zmierzającymi do wydzielenia modułów ukierunkowanych zadaniowo i odpowiadają np. stosowaniu procesorów specjalizowanych do obsługi komunikacji z urządzeniami zewnętrznymi (w roli kanałów maszyny cyfrowej).

2. Linie transmisji danych

Do przekazywania danych w sieciach teleinformatycznych można wykorzystywać różne powszechnie stosowane środki łączności przewodowej, radiowej, a także świetlnej. Nawet łączność satelitarna jest wykorzystywana w komputerowych sieciach transmisji danych, jak np. w sieci GE Information Services Network [11].

Największe znaczenie odgrywają obecnie środki łączności telegraficznej i telefonicznej, które są ogólnie dostępne i powstały dużo wcześniej niż rozpowszechniły się komputery. Zmusza to projektantów sieci teleinformatycznych do dopasowywania się do istniejących rozwiązań technicznych. Najbardziej rozpowszechniona sieć telefoniczna jest przystosowana do transmisji sygnałów analogowych (akustycznych), podczas gdy sygnały wejściowe i wyjściowe w komputerach mają charakter cyfrowy. Ograniczenie pasma częstotliwości przesyłanych przez linie telefoniczne od strony najniższych i najwyższych częstotliwości do typowego zakresu $300 \div 3400$ Hz oraz występowanie w tym zakresie częstotliwości sygnałowych, związanych z pracą central telefonicznych, powoduje konieczność stosowania specjalnych urządzeń pośredniczących, tj. modulujących przy wprowadzaniu sygnałów cyfrowych na linię telefoniczną i demodulujących przy odbiorze sygnałów z linii, nazywanych modemami. Podobnie przy stosowaniu linii telegraficznych stosuje się specjalne urządzenia do przekształcania sygnałów.

Szybkości transmisji w liniach telegraficznych nie przewyższają 200 bitów/s, a w liniach telefonicznych typowo wynoszą $600 \div 1200$ bitów/s. W wydzielonych (niekomutowanych) liniach telefonicznych osiąga się szybkości transmisji 9600 bitów/s.

Charakterystyczną cechą telefonicznych linii transmisyjnych jest ich stosunkowo mała pewność działania, wynikająca z tzw. szumów komutacyjnych, chwilowych krótkotrwałych przerw i zakłóceń, nieistotnych zwykle w przypadku przesyłania sygnałów analogowych ze względu na charakter przesyłanych informacji (małe wykorzystanie możliwości przesyłowego kanału i duży poziom redundancji charakteryzującej sygnał), a mających zasadnicze znaczenie w przypadku transmisji cyfrowych.

Przy stosowaniu linii telefonicznych do transmisji danych należy uwzględnić nie tylko wymienione dotąd cechy, ale również występowanie niejednorodności takich parametrów, jak zniekształcenia amplitudowe i fazowe. Wszystko to powoduje, że większość istniejących sieci telefonicznych nie odpowiada potrzebom rozwijających się sieci teleinformatycznych. Najbardziej atrakcyjne przy wykorzystywaniu linii telefonicznych jest ich powszechna dostępność i działanie na zasadzie sieci otwartej, o nieograniczonych możliwościach rozbudowy. Budowa oddzielnych sieci teleinformatycznych w początkowym okresie rozwoju teleinformatyki jest ekonomicznie nieuzasadniona. Dla transmisji większych ilości danych gdy potrzeba zapewnić większą przepustowość, stosuje się wydzielone linie nielokutowane. Linie takie ze względu na trwały charakter połączenia mogą być lepiej fizycznie dopasowane do charakterystyk nadajników i odbiorników, odporniejsze na zakłócenia, a nawet tańsze od połączeń komutowanych przy stosowanej polityce taryfowej. Niemniej postępujący rozwój systemów informatycznych wywołuje potrzebę rozwoju sieci transmisyjnych lepiej dopasowanych do cyfrowej postaci sygnałów. Stąd przewiduje się, że powstaną specjalne sieci teleinformatyczne w poszczególnych krajach, a potem sieci o zasięgu międzynarodowym.

Należy zwrócić uwagę, że sama telefonia coraz bardziej wykorzystuje technikę cyfrową. Elektroniczne centrale telefoniczne oraz wprowadzanie systemu impulsowej modulacji kodowej (PCM) powodują, że klasyczna telekomunikacja ciąży w kierunku transmisji danych cyfrowych. Stąd przewiduje się, że w dalszej przyszłości powstaną zintegrowane sieci cyfrowe powszechnego użytku, o uniwersalnym przeznaczeniu. Pozwoli to dołączać się do złożonych sieci komputerowych praktycznie w dowolnym, nie ograniczonym położeniem geograficznym - miejscu. Sieci takie mogą być podstawą do tworzenia publicznych sieci usług komputerowych.

Osobno należy rozpatrywać kanały transmisji o bardzo dużych przepustowościach. Są to tory kabli współosiowych, tory falowodowe i linie radiowe, a nawet tory światłowodowe. Zwykle stosuje się wówczas połączenia typu "punkt - punkt". Pozwalają one na uzyskiwanie w pojedynczych doświadczalnych instalacjach [10] przepustowości nawet ok. 10000 bitów/s.

3. Sieci teleinformatyczne

Struktura sieci teleinformatycznej jest określone przez rozkład topograficzny węzłów tej sieci, wyposażenie tych węzłów i własności linii transmisyjnych łączących te węzły.

Najogólniej możemy wyróżnić dwa rodzaje sieci teleinformatycznych:

- scentralizowane, w których występuje jeden główny komputer ¹⁾, wykorzystywany przez wielu użytkowników lokalnych i zdalnych,
- zdecentralizowane, w których występuje kilka odległych od siebie głównych komputerów.

Sieci scentralizowane są prostą kontynuacją dążenia do pełnego wykorzystania komputera przez dołączenie większej liczby urządzeń umożliwiających dostęp do maszyny większej liczbie użytkowników.

Sieci zdecentralizowane są zwykle niejednorodne pod względem typów głównych procesorów oraz ich wyposażenia. Podstawą do ich tworzenia jest wymiana programów, przekazywanie danych i udostępnianie mocy obliczeniowej, co także stanowi powód do różnicowania poszczególnych węzłów sieci. Klasycznym przykładem zdecentralizowanej sieci komputerowej jest znana sieć ARPA, która ulegała stopniowej rozbudowie i obecnie liczy ponad 30 węzłów.

Ważnymi czynnikami określającymi właściwości sieci teleinformatycznej, obok jej rozwiązania topologicznego, są:

- wymagane przepustowości międzywęzłowe,
- dostępność i koszty użytkowania lub dzierżawy (ewentualnie budowy) linii transmisyjnych,
- stopa błędów linii transmisyjnych,
- możliwe do osiągnięcia szybkości transmisji,
- charakter wykorzystywanych urządzeń końcowych i komputerów,
- wymagana niezawodność i typowy bądź maksymalny czas realizacji przesłania.

1) W literaturze komputer taki często po angielsku jest nazywany "host computer" lub "main computer", dla odróżnienia od komputerów pomocniczych, spełniających funkcje komunikacyjne w sieci teleinformatycznej.

W sieciach teleinformatycznych przesyłanie odbywa się zwykle na zasadzie przekazywania, np. na podstawie adresu, określonej porcji informacji z jednego węzła do drugiego gdzie jest przechowywana i skąd ewentualnie jest dalej wysyłana. Procedurę tę nazywa się komutacją informacji (message switching) w odróżnieniu od klasycznej w telekomunikacji komutacji połączeń kanałów łączności (line switching). Metoda szybkiego przesyłania stosunkowo niedużych porcji informacji (do ok. 1000 bitów), zwanych pakietami, przez kolejne węzły sieci, aż do miejsca przeznaczenia, znana jest powszechnie jako komutacja pakietów informacji (packet switching).

Trudności przy synchronizacji występujące w szybkiej transmisji sygnałów w sieciach o strukturze drzewiastej, spowodowały zainteresowanie się sieciami mającymi strukturę pętlową. W sieciach takich synchronizacja jest wymuszana przez impulsy przesyłane w pętli. Przykładem takiej sieci może być sieć DCS (Distributed Computer System) [3] opracowana na Uniwersytecie Kalifornijskim w Irvine. Przesyłanie w pętli odbywa się z szybkością powyżej 2 milionów bitów/s, a występujące w pętli specjalne urządzenia umożliwiają dołączenie do pętli komputerów, urządzeń końcowych tworzących stacje abonenckie oraz dodatkowych podpętli. Informacja przesyłana w pętli jest adresowana według nazwy odbiorcy, a nie jego rzeczywistego położenia. Stwarza to możliwość pracy tego samego użytkownika w dowolnych punktach sieci pętlowej.

Przy projektowaniu sieci teleinformatycznej występuje wiele złożonych problemów. Nie istnieją uniwersalne i kompleksowe metody projektowania. Powszechnie natomiast stosowane są metody analizy oraz symulacji w celu weryfikowania proponowanych rozwiązań.

Praktycznie każda sieć jest inna i inne są wymagania jej dotyczące. Inaczej wygląda np. problem optymalizacji wykorzystania istniejącej sieci, a inaczej budowy nowej sieci. Najogólniej można wydzielić zagadnienia związane z kanałami transmisji danych oraz węzłami sieci. Przykładem są takie zagadnienia rozpatrywane przy projektowaniu sieci jak dobór trasy przesyłania danych, rozmieszczenie koncentratorów i długości przesyłanych meldunków. Szersze rozwinięcie tych spraw można znaleźć w pracach [6], [12].

Jedną z najbardziej pracochłonnych czynności przy projektowaniu sieci jest opracowanie procedur "protocols" sterowania transmisją danych, określających szczegółowo reakcje na każde zdarzenie występujące w sieci. Zbiory tych procedur wyznaczają wymagania dotyczące sprzętu i oprogramowania wykorzystywanego w sieci. Pojawia się w związku z tym zagadnienie standaryzacji "przepisów ruchu" w sieciach transmisji danych. Istotne to jest też przy rozbudowie istniejących sieci, albo ich łączeniu w większe systemy. Przykładem wytyępowania takiego zagadnienia może tu być łączenie różnych krajowych sieci informatycznych między sobą oraz dowiązywanie ich do sieci zagranicznych, celem tworzenia sieci międzynarodowych.

4. Rola minikomputera w sieci teleinformatycznej

4.1. Dlaczego i do czego

Podstawowe czynniki powodujące wzrost zainteresowań oraz coraz szersze wykorzystanie minikomputera w sieciach teleinformatycznych (zwykle zamiast specjalizowanych urządzeń) wynikają przede wszystkim z tego, że:

- jest to urządzenie umożliwiające łatwą i elastyczną zmianę swych funkcji przez wymianę programów sterujących (programowalność) oraz zmianę zestawu dołączonych urządzeń,
- szybkość działania tego urządzenia jest duża (wydajność),
- urządzenie ma pamięć, która może łatwo służyć do chwilowego przechowywania danych (zdolność buforowania),
- jest to urządzenie stosunkowo tanie, a ceny minikomputerów maleją w wyniku postępu technologii, np. przez stosowanie układów scalonych o coraz wyższym stopniu scalenia (tanieść); wg pracy [5] koszt jednostkowy operacji w przypadku minikomputerów maleje ok. 60 % rocznie, co odpowiada przeszko stokrotnemu zmniejszeniu tego kosztu w okresie 10 lat.

W obecnej chwili wykorzystanie minikomputerów przy przesyłaniu danych cyfrowych można najogólniej sklasyfikować wg następujących punktów:

- jako zdalne koncentratory danych i jednostki sterujące w urządzeniach końcowych (terminalach),
- jako wspomagające urządzenia dla większego (głównego) procesora, wypełniające funkcje związane z transmisją danych oraz ewentualnie innymi pomocniczymi zadaniami przetwarzania danych (procesor buforowy).

Szczegółowe omówienie specyfiki powyższych zastosowań przedstawiono dalej. Tutaj omówimy najogólniej funkcje jakie powinien spełniać minikomputer w ramach transmisji danych.

4.2. Funkcje minikomputera w sieci

Poniżej omówimy najogólniej funkcje, jakie spełnia procesor minikomputerowy przy transmisji danych w sieciach teleinformatycznych.

- Organizacja procesu przesyłania danych - funkcje sterowania
Procesor minikomputerowy gromadzi nadsyłane dane i przygotowuje je do transmisji przez formowanie odpowiednich bloków danych, uzupełnianie ich w odpowiednie znaki sterujące i kontrolne (np. bity parzystości) i adresy. W niektórych przypadkach istotną rolą procesora jest nadążanie łączności, a nawet optymalizacja kierunku przekazywania informacji, w przypadku złożonych sieci zapewniających swobodę w wyborze trasy przy przesyłaniu danych.
- Weryfikacja i korekta danych
Ze względu na duże odległości i rodzaje stosowanych środków łączności wykorzystywanych w sieciach teleinformatycznych dużą rolę przydaje się do metod weryfikacji i korekty przesyłanych danych. Jak skuteczne mogą być te metody świadczy fakt, że np. w znanej powszechnie wielkiej sieci ARPA w czasie pierwszego roku eksploatacji nie wykryto ani jednego błędnego działania wynikającego z transmisji danych [6] .

Typowe ilości błędów przy transmisji w sieciach teleinformatycznych, bez uwzględnienia korekty, zawierają się w zakresie od jednego niepoprawnego bitu na 50 tys. przesłanych bitów, do jednego niepoprawnego bitu na 500 tys. przesłanych bitów [7] .

Weryfikacja poprawności danych polega zwykle na wykorzystaniu redundancji przesyłanej informacji. Do najprostszyc sposobów wprowadzania redundancji należą dodawanie bitów parzystości oraz znaków kontrolnych.

Korekca błędów odbywa się automatycznie różnymi sposobami i najczęściej polega na zażądaniu powtórnego przesłania wadliwych danych lub odtworzeniu właściwych znaczeń przy wykorzystaniu struktury użytego kodu.

- **Dopasowanie do punktów nadawczo-odbiorczych**

Punktami nadawczo-odbiorczymi mogą być tu urządzenia dołączone do procesora jak również węzły sieci teleinformatycznej.

Procesor może spełniać rolę urządzenia pośredniczącego, które dopasowuje się do postaci danych, szybkości oraz reguł sterowania poszczególnych urządzeń współpracujących - innych procesorów lub urządzeń zewnętrznych.

Do działań tych należy między innymi zmiana przedstawienia kodowego danych, istotna ze względu na to, że umożliwia łatwe dołączanie do sieci teleinformatycznej urządzeń dostępnych na rynku, pomimo różnorodności stosowanych kodów (np. ASCII, CCITT No 2, itp.).

Takie oraz inne funkcje dopasowujące realizowane przez procesor pozwalają na elastyczne rozszerzanie sieci teleinformatycznych przez kolejne dołączanie nowych węzłów.

- **Dopasowanie do linii transmisyjnych**

Dopasowanie do linii transmisyjnej może polegać również na sposobie przedstawienia danych, np. zamianie postaci równoległej na szeregową. Jednak główną rolę odgrywa tu dopasowanie do optymalnej szybkości transmisji w linii. Zagadnienie to jest szczególnie istotne przy transmisji danych na większe odległości, gdy koszty linii transmisyjnych są znaczne (np. połączenia międzynarodowe, międzykontynentalne) i opłaca się rozbudowywać urządzenia współpracujące z liniami, aby lepiej wykorzystać przepustowość tych linii.

Funkcje takie spełniają przede wszystkim urządzenia zwane koncentratorami (bardzo często z udziałem minikomputerów), a częściowo multipleksory.

Wykorzystanie minikomputerów pozwala na elastyczne dopasowanie się do zmiennych warunków transmisji między węzłami, niezbędne np. przy uszkodzeniach linii transmisyjnych i stosowaniu połączeń zapasowych (np. gdy zamiast wydzielonych linii telefonicznych trzeba wykorzystać połączenie komutowane zwykłej sieci telefonicznej).

4) Buforowanie w operacjach przesyłania danych

Buforowanie może służyć również jako środek do dopasowywania do szybkości transmisji w linii, ale zostało ono wyróżnione w osobnym punkcie, ze względu na szersze znaczenie tej funkcji. Wykorzystuje się przy tym szybkie pamięci operacyjne minikomputera, jak również pamięci zewnętrzne.

Proste gromadzenie w pamięci minikomputera danych wolno napływających z klawiatur obsługiwanych przez operatorów, aby potem po odpowiednim uformowaniu ich przekazać je dalej, jest najprostszą ilustracją funkcji buforowania.

Przekazywanie określonych porcji danych, zwanych pakietami, między pamięciami operacyjnymi kolejnych procesorów komunikacyjnych w sieci komputerowej ARPA, pracującym systemem "store-and-forward", jest innym przykładem wykorzystywania funkcji buforowania.

W złożonych systemach z przypadkowym dostępem użytkowników buforowanie pozwala złagodzić szczyty obciążeń transmisyjnych, a tym samym uniknąć konieczności rezerwacji dodatkowych zdolności przepustowych linii.

Buforowanie z wykorzystaniem pamięci zewnętrznych minikomputera komunikacyjnego, takich jak np. pamięci dyskowe, pozwala na "złagodne" przerwanie pracy sieci w przypadku awarii procesora głównego.

5) Adaptacyjność funkcji i parametrów do potrzeb systemu

Adaptacyjność minikomputera w sieci teleinformatycznej wynika z jego programowalności. Należy przy tym rozróżnić dwa stopnie tej adaptacyjności.

Pierwszy stopień wynika z możliwości zmiany sposobu działania minikomputera, przez prostą wymianę programu przez operatora.

Taki przypadek może wystąpić w sytuacji, gdy następuje zmiana konfiguracji urządzeń dołączonych do procesora komunikacyjnego, lub zmienia się linie transmisyjne współpracujące z procesorem komunikacyjnym.

Drugi stopień adaptacyjności, bardziej atrakcyjny, wynika z możliwości przesyłania do pamięci komputera komunikacyjnego całych programów, albo też parametrów dotyczących tych programów, z głównych procesorów występujących w sieci teleinformatycznej.

Przykładem takiego wykorzystania minikomputera może być wymiana programu wyznaczającego sekwencję obsługi ("przepytowania") urządzeń współpracujących z minikomputerem jako jednostką sterującą w urządzeniu końcowym, w ramach tzw. "pollingu". Zmienia to priorytety obsługi poszczególnych urządzeń. Na podobnej zasadzie główny procesor może przesyłać do minikomputera pracującego w urządzeniu końcowym program diagnostyczny, aby ocenić stan techniczny (zdalna diagnoza) poszczególnych urządzeń.

Innym przykładem adaptacyjnej współpracy procesora komunikacyjnego z głównym procesorem, mogłaby być poprzedzająca właściwą współpracę wstępną, konwersacyjną wymiana informacji sterujących, doprowadzająca do ustalenia odpowiedniej szybkości transmisji danych między urządzeniem końcowym a głównym procesorem. Ustalenia takie mogą być dokonywane automatycznie, po dołączeniu urządzenia końcowego do sieci teleinformatycznej.

5. Minikomputer w roli koncentratora

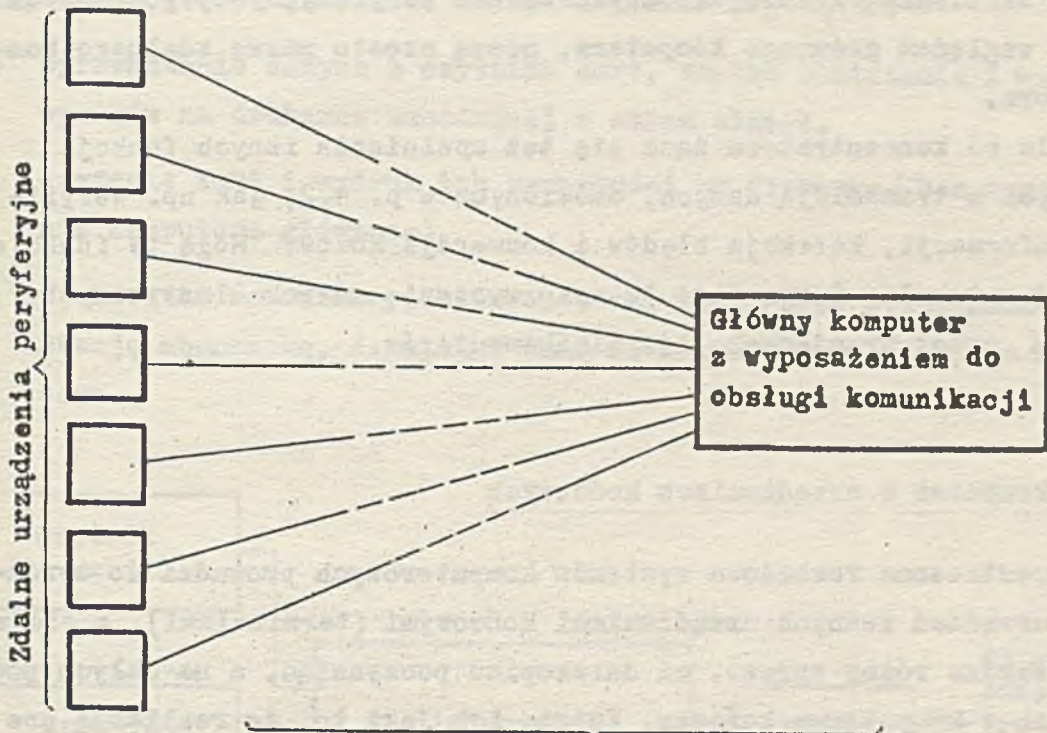
Konwencjonalny sposób wykrczystywania przepustowości kanałów telekomunikacyjnych przez wielokrotnienie (multipleksowanie) oparte na podziale częstotliwości (FDM - Frequency Division Multiplexing) lub podziale czasu (TDM - Time Division Multiplexing), jest prostą metodą przesyłania danych z wielu źródeł jednym torem. Charakterystyczną cechą urządzenia spełniającego tę rolę, a zwanego multipleksozem jest to, że przepustowość na wyjścia jest równa sumie przepustowości linii wejściowych. Efektywność takiej metody jest duża, gdy dane wejściowe nadchodzą w sposób rytmiczny, tj. są równomiernie rozłożone w czasie.

Już w prostych kanałach transmisji danych, np. łączących końcowe urządzenie typu dalekopisowego z głównym komputerem, rozkład czasowy przesyłanych danych podlega dużym wahanom - operator naciskający kła-

wieze zastanawia się, robi przerwy dla pobrania nowych danych itp. Tym bardziej w dużych sieciach komputerowych należy brać pod uwagę statystyczne wahania obciążeń przepustowości kanałów transmisji danych. Zapewnienie takiej przepustowości kanału transmisji danych, aby można było zapewnić prawidłową pracę w przypadku losowo występujących największych obciążeń, może być związane z koniecznością zapewnienia dużych zapasów przepustowości kanałów, wykorzystywanych w efekcie tylko częściowo. Jaskrawym przykładem niegospodarności takiego podejścia może być praca pojedynczego dalekopisu przez wydzielony kabel podmorski z komputerem zainstalowanym na innym kontynencie.

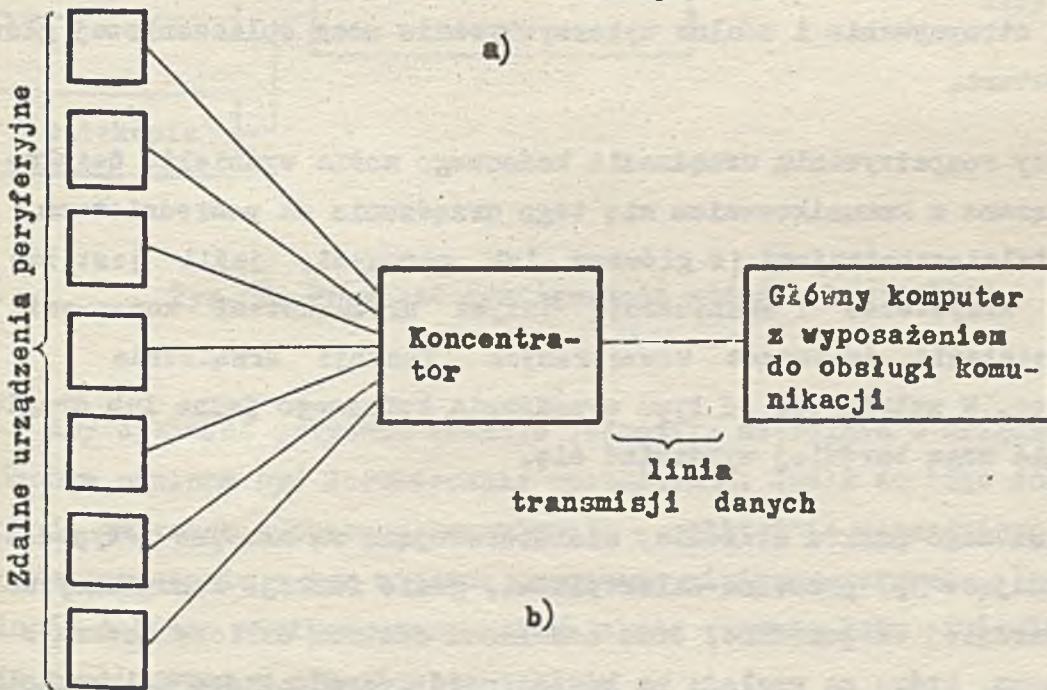
Rozwiązanie takich problemów może polegać na zbieraniu danych z różnych źródeł w specjalnym urządzeniu buforującym i formowaniu z nich bardziej zagęszczonych strumieni danych. Zanika wówczas synchroniczność między otrzymaniem danych na wejściu, a wysyłaniem ich na wyjściu. Urządzenia transmisji danych spełniające takie funkcje noszą nazwę koncentratorów i stanowią wyższy stopień rozwoju urządzeń niż multipleksory, które ze względu na swą niższą cenę i prostotę działania powinny być i są stosowane obok koncentratorów. Cechą charakterystyczną koncentratora jest więc to, że ma on zdolność buforowania, umożliwiającą zmniejszenie przepustowości linii wyjściowej w stosunku do sumy przepustowości linii wejściowych. Pozwala to zmniejszyć koszty użytkowania kanału transmisji danych lub lepiej ten kanał wykorzystywać przez przesyłanie większej ilości danych. Pojemność pamięci buforowej koncentratora, stanowiąca jego podstawowy parametr, jest wybierana na podstawie rozkładów obciążeń transmisyjnych linii wejściowych.

Omawiana rola koncentratora związana jest głównie z transmisją w kierunku od wielu źródeł informacji do jednego ujścia, na którym dane są przedstawione w zmodyfikowanej postaci. Cozywiste jest, że koncentrator również zwykle spełnia funkcje związane z transmisją danych w drugim kierunku, które odpowiadają "dekonzcentracji" danych, przez rozsyłanie danych do wielu odbiorców. Przykład wykorzystania koncentratora do zmniejszenia liczb linii łączących urządzenia wejścia-wyjścia z głównym komputerem pokazano na rys. 1a i b.



7 linii transmisji danych

a)



linia transmisji danych

b)

Rys. 1. Przykład konfiguracji połączeń urządzeń peryferyjnych z głównym komputerem bez użycia koncentratora a) oraz z użyciem koncentratora b). Dla uproszczenia na rysunkach nie przedstawiono modemów

Koncentratory wykorzystywane, jako ośrodki zbierania danych w miejscu skupienia większej liczby urządzeń peryferyjnych, położonych zdalnie względem głównego komputera, noszą często nazwę zdalnego koncentratora.

Zwykle od koncentratora żąda się też spełnienia innych funkcji związanych z transmisją danych, omówionych w p. 4.2, jak np. weryfikacja informacji, korekcja błędów i konwersja kodów. Mogą te funkcje bardziej naturalną formą jest tu wykorzystanie małych elastycznych urządzeń przetwarzających, tj. minikomputerów.

6. Minikomputer w urządzeniach końcowych

Przestrzenna rozbudowa systemów komputerowych prowadzi do wyróżnienia urządzeń zwanych urządzeniami końcowymi (terminalami), a obejmującymi bardzo różny sprzęt, od dalekopisu poczynając, a na całych podsystemach z komputerem kończąc. Istotą ich jest to, że realizują one takie funkcje, jak zdalne wprowadzanie danych do głównego komputera lub ich otrzymywanie i zdalne wykorzystywanie mocy obliczeniowej głównego komputera.

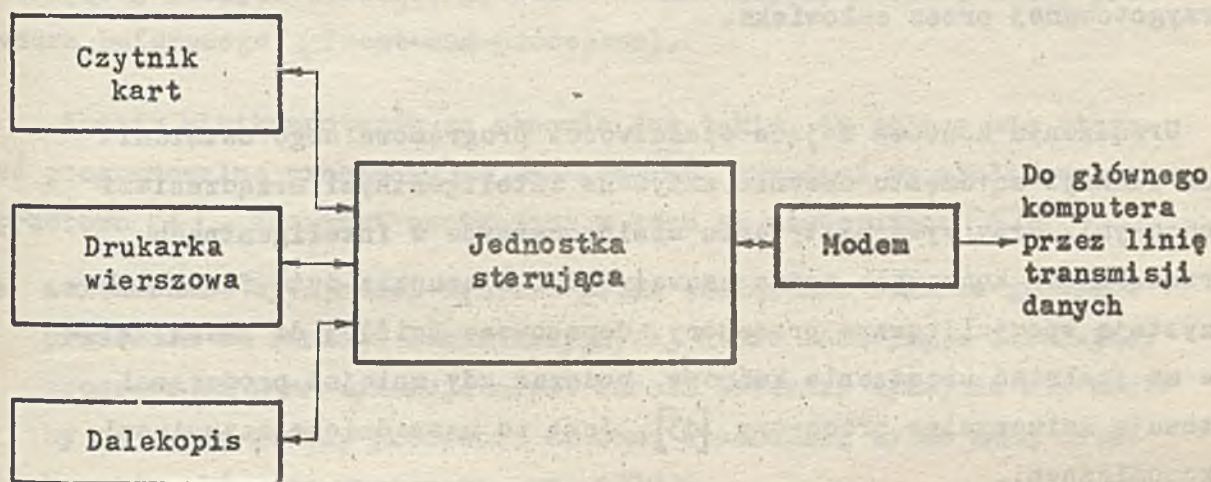
Przy rozpatrywaniu urządzenia końcowego można wydzielić działania związane z komunikowaniem się tego urządzenia za pośrednictwem sieci teletransmisyjnej (z głównym lub głównymi, jeśli jest ich więcej) komputerem i ewentualnie innymi urządzeniami końcowymi, oraz działania dotyczące wewnętrznych funkcji urządzenia końcowego. W zależności od typu urządzenia końcowego jedno lub drugie działania mogą bardziej wyróżniać się.

Z naszego punktu widzenia, nieinteresujące są skrajne przypadki, odpowiadające np. prostemu dalekopisowi, gdzie funkcje komunikacyjne są najbardziej eksponowane, oraz lokalnemu centrum obliczeniowemu z komputerem, które ze względu na konieczność sporadycznego zwiększania mocy obliczeniowej aby można było rozwiązać szczególnie trudne zadania, jest okresowo łączone z innym, oddalonym komputerem.

Interesuje nas przypadek mniej lub więcej rozbudowanej stacji abonenckiej, która w klasycznym rozwiązaniu (np. wg rys. 2) stanowi

jednostkę sterującą z zestawem urządzeń do wprowadzania i wyprowadzania danych. Ta stacja abonencka ma np. takie funkcje [13]:

- wprowadzanie danych z czytnika kart, zdalne obliczanie i wydruk wyników na drukarce wchodzącej w skład stacji,
- czytanie kart i wydruk ich zawartości na drukarce (bez angażowania komputera głównego),
- drukowanie formularzy, których stała treść jest wyznaczona przez stację abonencką, a zmienne dane są uzyskiwane z głównego komputera.



Rys. 2. Przykład konfiguracji stacji abonenckiej

Aby spełniać powyższe funkcje jednostka sterująca w urządzeniu końcowym powinna być dostatecznie rozbudowana. Jeśli do tego dołożymy funkcje związane z obsługą komunikacji - całkowicie uzasadnione staje się wykorzystanie w tym miejscu procesora minikomputerowego. Programowalność takiego minikomputera potencjalnie pozwala łatwo zmieniać funkcje realizowane w stacji abonenckiej.

Przykładem rozbudowanych stacji abonenckich wykorzystujących minikomputery mogą być wielostanowiskowe urządzenia do przygotowania danych na dysku magnetycznym.

Minikomputer w roli jednostki sterującej urządzenia końcowego może spełniać również istotną rolę jako urządzenie przetwarzające dane i w ten sposób odciążające główny komputer. Ważne może być wówczas również dążenie do zmniejszenia wymagań dotyczących przepustowości linii transmisji danych łączącej urządzenie końcowe z głównym komputerem. Przykładem jest tu zdalnie pracujący graficzny monitor ekranowy. Minikomputer może w tym przypadku obsługiwać wszystkie działania związane z wprowadzaniem na ekran monitora obrazów i ich modyfikacją przez operatora. Transmisja danych między jednostką sterującą, a głównym procesorem dotyczy wówczas tylko wyprowadzania wyników przetwarzania w głównym procesorze oraz odbierania przez ten procesor informacji przygotowanej przez człowieka.

Urządzenia końcowe mające właściwości programowalnego ustalania ich funkcji są często obecnie nazywane inteligentnymi urządzeniami końcowymi. Przy wykorzystywaniu miniprocessorów w inteligentnych urządzeniach końcowych można zauważyć dwa kierunki: duże firmy wykorzystują specjalizowane procesory, dopasowane ściślej do zadań, które ma spełniać urządzenie końcowe, podczas gdy mniejsi producenci stosują uniwersalne procesory [13]. Jest to uzasadnione czynnikami ekonomicznymi.

7. Minikomputer jako procesor buforowy

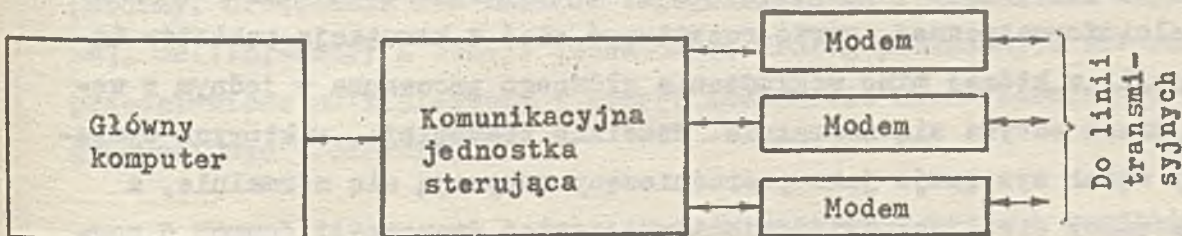
Dotychczas nie zajmowaliśmy się bliżej obsługą linii transmisyjnych doprowadzonych do głównego komputera. Zwykle linie te są doprowadzone do procesora tego komputera nie bezpośrednio lecz przez urządzenie (rys. 3) zwane komunikacyjną jednostką sterującą (transmission control unit, line control unit). Jednostka ta może w najprostszym przypadku zbierać docierające do niej bity w znaki i przekazywać do głównego procesora oraz w drugim kierunku - otrzymywać znaki i wysyłać pojedyncze bity. Bardziej złożone urządzenia umożliwiają wymianę danych z głównym komputerem na poziomie bloków lub słów złożonych ze znaków. Oprócz tych funkcji komunikacyjna jednostka sterująca może, a czasem powinna spełniać podstawowe działania związane z obsługą transmisji danych omówione w p. 4.2, jak np. korekta błędów.

Przykładem urządzenia zbudowanego specjalnie do obsługi linii transmisji danych może tu być urządzenie IBM 2701 o nazwie Data Adapter Unit, dołączane do procesorów IBM 360 tak, jak inne jednostki sterujące. Wadą tego rozwiązania jest to, że obciąża główny procesor takimi funkcjami komunikacyjnymi jak buforowanie, sterowanie wydawaniem i przyjmowaniem danych, zmiany formatu danych itp.

Wzrost liczby urządzeń końcowych, wzrastające wymagania dotyczące jakości i elastyczności zmian konfiguracji systemów teleinformatycznych spowodowały dążenie do uwolnienia głównego procesora od zadań komunikacyjnych. W miejsce specjalizowanej komunikacyjnej jednostki sterującej zaczęto stosować procesor komunikacyjny noszący nazwę procesora buforowego¹ (front-end-processor).

Koszty minikomputerów są obecnie już takie, że opłaca się stosować programowalne procesory buforowe zamiast urządzeń specjalizowanych sprzętowo [4]. Korzyści wynikające z tego są następujące [5]:

- zwiększenie wydajności systemu przez odciążenie czasowe głównego procesora od funkcji komunikacyjnych, które równoległe działający procesor buforowy wykonuje nawet czasem bardziej wydajnie niż mógłby to robić główny procesor (za cenę stanowiącą tylko małą część kosztów głównego procesora, np. 1/20),
- odciążenie pamięci operacyjnej w głównym procesorze przez przerzucenie funkcji buforowania do procesora buforowego,



Rys. 3. Komunikacyjna jednostka sterująca we współpracy z głównym komputerem

¹ W literaturze polskiej brak jest przyjętego powszechnie tłumaczenia angielskiej nazwy "front-end-processor". W pracy [14] np. zaproponowano nazwę "procesor czołowy". Nazwę "procesor buforowy" należy traktować jako propozycję autorów.

- odciążenie pamięci masowej głównego komputera przez wykorzystywanie do przechowywania komutowanej informacji pamięci masowej związanej z procesorem buforowym, co również zwiększa efektywność dostępu do niej przy obsłudze transmisji danych,
- elastyczność wynikająca z programowalności procesora buforowego i pozwalająca łatwo dostosowywać się do zmian konfiguracji sprzętu, zmian wymagań funkcjonalnych (np. adaptacyjność omówiona w p.4.2),
- modularność oprogramowania wynikająca z oddzielenia oprogramowania głównego procesora i procesora komunikacyjnego, co powoduje łatwiejsze dostosowanie tego oprogramowania do zmian w systemie,
- zwiększenie odporności systemu na przekłamania przez rozszerzenie funkcji weryfikacji i korekcji danych realizowanych w procesorze buforowym oraz lepsze odizolowanie procesora głównego od błędów wynikających z transmisji danych,
- rozdzielenie funkcjonowania obsługi części komunikacyjnej od działania głównego procesora, co pozwala na ciągle podtrzymywanie obsługi transmisji danych, nawet w przypadku chwilowych niesprawności głównego procesora oraz umożliwia testowanie na bieżąco w czasie pracy dołączonych urządzeń komunikacyjnych.

Dla zilustrowania korzyści wynikających z zastosowania procesora buforowego jako samodzielnego gospodarza funkcji komunikacyjnych w sieci teleinformatycznej, warto rozpatrzeć sieć z komutacją pakietów informacji, w której mimo uszkodzenia głównego procesora w jednym z węzłów, praca odbywa się normalnie. Wszelkie transmisje, w których wymieniony węzeł występuje jako pośredniczący odbywają się normalnie, a użytkownicy sieci są zawiadomieni przez sieć transmisji danych o powstałym uszkodzeniu i "łagodnym" odłączeniu (dzięki buforowaniu w procesorze komunikacyjnym) tego głównego procesora.

Szczegółowsze omówienie funkcji różnych typów procesorów buforowych można znaleźć w literaturze [5].

8. Stan rozwoju sieci teleinformatycznych w poszczególnych krajach

Najbardziej rozwinięto prace nad sieciami teleinformatycznymi w USA. Przejawem tego jest duży rozwój sieci zdecentralizowanych. Powstano tu kilkadziesiąt sieci komputerowych, z których najbardziej znana jest wspomniana już sieć ARPA (Advanced Research Projects Agency) [15]. Do innych znanych sieci należy TYMNET, INFONET, CYBERNET, MERIT, OCTOPUS, DCS, TUCC [3], [11]. Z prac prowadzonych w Europie Zachodniej najbardziej znana jest eksperymentalna sieć komputerowa opracowana w Anglii przez National Physical Laboratory (NPL), podobna funkcjonalnie do sieci ARPA.

Rozwój przestrzennych systemów informatycznych w poszczególnych krajach związany jest ściśle z rozwojem sieci służących do transmisji danych. Sprawy te są przedmiotem międzynarodowych ustaleń, a zajmuje się nimi głównie od 1970 r. [8], [9], tzw. grupa NRD (Nouveaux Réseaux du Données - Nowe Sieci Danych) powołana przez CCITT (Comité Consultatif International Télégraphique et Téléphonique). Jako przykład można podać, że Wielka Brytania przewiduje uruchomienie krajowej sieci transmisji danych w 1977 r. Plany rozwojowe sieci transmisji danych zostały również sprecyzowane we Francji [16] i NRF [17].

Stan opracowań urządzeń Jednolitego Systemu EMC związanych z transmisją danych został przedstawiony w "Informatyce" [18], [19]. Opisano tu urządzenia transmisji danych służące do przekształcania sygnałów (modemy, urządzenia dla kanałów telegraficznych i urządzenia zwykowe), multipleksory i stacje abonenckie. Według posiadanych informacji przedstawione multipleksory i stacje abonenckie nie wykorzystują programowalnego procesora.

Z innych informacji o pracach prowadzonych w krajach socjalistycznych nad transmisją danych w powiązaniu z minikomputerami należy wymienić prace realizowane poza JS EMC w NRD i na Węgrzech. W NRD opracowano system przesyłania informacji [20] oparty na 16-to bitowym minikomputerze Robotron 4000 z pamięcią operacyjną o pojemności 32 kilobajtów. Węgry w październiku 1972 r. zaprezentowali na pokazie w Moskwie współpracę minikomputera typu 1010B z EMC typu Mińsk-32. Według przedstawionych propozycji [21] minikomputer 1010B pełni tu rolę procesora

buforowego umożliwiającego współpracę EMC typu Mińsk-32 z różnymi urządzeniami końcowymi, poczynając od programowanych (z tego samego typu minikomputerem 1010B) a kończąc na elektrycznych maszynach do pisania.

Prace prowadzone w Polsce dotyczące sieci teleinformatycznych są stosunkowo słabo rozwinięte. Budowanie systemów wielodostępnych, stanowiących pierwszy krok do budowy sieci teleinformatycznych umożliwiającą niedawno wprowadzone do produkcji maszyny ODRA 1305 i ODRA 1325. Pierwsze systemy teleinformatyczne są dopiero wprowadzone do użytku. Z ciekawych prac warto wymienić budowany na Politechnice Wrocławskiej Wielodostępny Abonencki System Cyfrowy (WASC) oparty na EMC typu ODRA 1305 i 1325 i umożliwiający przetwarzanie wadowe oraz pracę konwersacyjną [22], oraz oparty na importowanej EMC system CYFRONET [23]. Inne systemy informatyczne znajdują się w opracowaniu. Rozwój sieci komputerowych znajduje się dopiero na etapie dyskusji.

Należy sądzić, że duży wpływ na rozwój teleinformatyki w Polsce będzie miał zakup licencji na elektroniczny system komputacji CITEDIS, umożliwiający budowę nowoczesnych central telefonicznych i powstawanie zintegrowanej sieci łączności z wykorzystaniem systemów teletransmisyjnych PCM. Przejawem rozwijających się coraz silniej zainteresowań teleinformatyką jest konferencja "Sieci komputerowe" [24], która odbyła się we wrześniu 1973 r., oraz projekt zorganizowania konferencji "Sieci teleinformatyczne" w najbliższym okresie. Szersze wykorzystanie minikomputerów w sieciach transmisji danych cyfrowych należy przewidywać wraz z ogólnym rozwojem teleinformatyki w Polsce.

9. Zakończenie

Powszechnie akceptowany jest pogląd, że maszyny cyfrowe będą rozwijały się w dwóch kierunkach: przez stosowanie w wielu miejscach takich, małych mini, albo nawet mikroprocesorów oraz bardzo dużych procesorów, o wielkich możliwościach przetwarzania danych.

Ze względu na postęp w technologii elektronicznej ceny procesorów maleją kilka razy szybciej, niż ceny urządzeń zewnętrznych, głównie związanych z konstrukcjami mechanicznymi [25]. Stąd coraz mniej opłaca

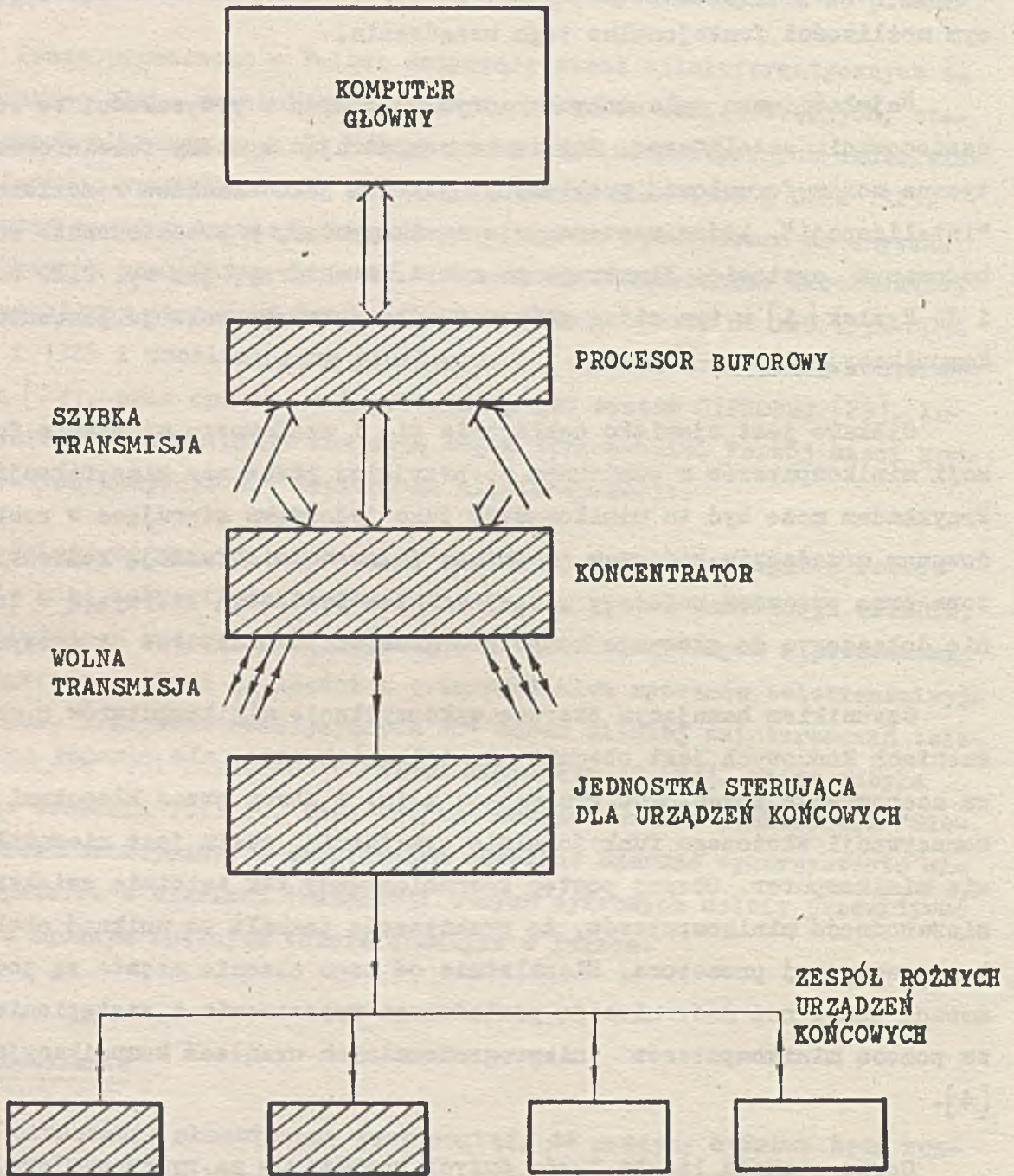
się obudowywać jeden, względnie tani procesor drogimi urządzeniami zewnętrznymi. Można przewidywać, że w przyszłości każde urządzenie zewnętrzne będzie połączone z mikroprocesorem znakomicie zwiększającym możliwości funkcjonalne tego urządzenia.

Najważniejsza rola małych maszyn cyfrowych w przyszłości to różne zastosowania satelitarne. Już teraz rozpatrując systemy teleinformatyczne można formułować problemy projektowe jako właściwe rozdzielenie "inteligencji", którą reprezentuje minikomputer, w przestrzennie rozbudowanym systemie. Ilustruje to rys. 1. Niektórzy, jak np. C.B. Newport i J. Ryzlak [5] w tym widzą główne źródło dalszego rozwoju procesorów komunikacyjnych.

Ciekawe jest zjawisko nakładania się i wzajemnego mieszania funkcji minikomputerów w odmiennym do przyjętej przez nas klasyfikacji. Przykładem może być tu minikomputer jako jednostka sterująca w rozbudowanym urządzeniu końcowym, pełniący jednocześnie funkcję koncentratora oraz procesor buforowy będący zarazem jednostką sterującą w lokalnie dołączonym do głównego komputera graficznym monitorze ekranowym.

Czynnikiem hamującym szersze wykorzystanie minikomputerów w urządzeniach końcowych jest obecnie nie tylko ich cena, ale również postawa zachowawcza inwestorów, wynikająca m.in. z obawy przed kłopotami przy konserwacji złożonego funkcjonalnie urządzenia, jakim jest niewątpliwie minikomputer. Obecny postęp technologiczny tak istotnie zwiększa niezawodność minikomputerów, że praktycznie pozwala to uniknąć obsługi konserwacyjnej procesora. Niezależnie od tego obecnie często są podejmowane prace nad modernizacją posiadanego wyposażenia i zastąpieniem za pomocą minikomputerów nieprogramowalnych urządzeń komunikacyjnych [4].

Coraz większa liczba typów dużych komputerów na rynku zawiera oferowane przez producentów wyposażenie komunikacyjne. Od stanu, w którym minikomputerowy procesor komunikacyjny był głównie przedmiotem troski projektantów sieci teleinformatycznej, przechodzi się do stanu, w którym producent dużej maszyny sam dostarcza minikomputerowy procesor komunikacyjny.



Rys. 4. Minikomputer jako "urządzenie inteligentne" w sieci teletransmisyjnej zakreskowano prostokąty reprezentujące minikomputer lub urządzenie końcowe zawierające minikomputer

Początkowy rozwój komputerów nie wiązał się z telekomunikacją, reprezentującą bogatą i długoletnią tradycję techniczną. Początkowe próby powiązania tych dwóch dziedzin dostarczyły sporo kontrowersji, np. co do przyjęcia linii rozdzielającej zakresy działania tych dwóch dziedzin. Wprowadzenie techniki cyfrowej do telekomunikacji (np. systemy PCM) z jednej strony, a zwiększenie zapotrzebowania na usługi poczty przez informatykę - z drugiej strony, znacznie zaawansowało godzenie interesów tych dziedzin. Chyba tylko sprawą czasu jest zastąpienie telefonu w domu abonentów przez podręczne cyfrowe urządzenia końcowe o różnorodnym przeznaczeniu. Zintegrowana sieć powszechnego użytku udostępni moc komputerową każdemu abonentowi. Prekursorem takich powszechnych sieci będą chyba sieci o zasięgu lokalnym, instalowane w różnych instytucjach, fabrykach itp.

Zakres tematyki związanej z teleinformatyką jest bardzo rozległy i w tym opracowaniu zostało przedstawione tylko szersze tło, na którym pokazano bliżej sprawy związane z minikomputerami.

Literatura

- [1] Data Communication Concepts, Univac 1969.
- [2] IBM System 370, System Summary, GA22 - 7001.
- [3] FABER D.J.: Networks: An Introduction. Datamation. 1972, t. 18 nr 4, s. 36-39.
- [4] BURNER M.B., MILLION R. i inni: The Use of a Small Computer as a Terminal Controller for a Large Computing System. AFIPS Conf. Proc. 1969, t. 34. Spring Joint Comput. Conf., s. 775-776; również w: Minicomputers: Hardware, Software and Applications. IEEE Press 1972, s. 317-318.
- [5] NEWPORT C.B., RYZLAK J.: Communication Processor, Proceedings of the IREE, 1972 t. 60, nr 11, s. 1321-1332.
- [6] MARTIN J.: System Analysis for Data Transmission, Prentice - Hall, Inc. 1972.

- [7] MARTIN J.: Teleprocessing Network Organization, Prentice -Hall, Inc., 1970.
- [8] FIJAŁKOWSKI W.J.: Nowe koncepcje sieci zdalnego przetwarzania danych. Informatyka, 1971, nr 11, s. 14.
- [9] ANDREWS M.C.: New Data Networks in Europe. Proceedings of the IEEE, 1972, t. 60, nr 11, s.1369.
- [10] SZPIGLER Z.: Linie telekomunikacyjne; w: Dziś i jutro telekomunikacji. Wydawnictwa Komunikacji i Łączności, Warszawa 1973, s.147.
- [11] SCHWARTZ M, BOORSTYN R.R., PICKHOLTZ R.L.: Terminal -Oriented Computer - Communication Networks, Proceeding of the IEEE, 1972, t. 60, nr 11, s.1408.
- [12] SEIDLER J.: Teoretyczne problemy cyfrowych systemów telekomunikacyjnych. W: Dziś i jutro telekomunikacji. Warszawa 1973 Wydawnictwa Komunikacji i Łączności, s. 33.
- [13] BALL C.J.: Communications and the Minicomputer. Computer 1971, t. 4, wrzesień/październik s. 13-21; również w: Minicomputers: Hardware, Software and Applications, IEEE Press 1972, s.307-314 .
- [14] LEPSZONEK A.: Kierunki rozwoju komunikacyjnych urządzeń sterujących. Informatyka 1973, nr 5, s. 8.
- [15] KAHN R.E.: Resource - Sharing Computer Communications Networks. Proceedings of the IEEE. 1972, t. 60, nr 11, s. 1397-1407.
- [16] JOUSSET A.M.: Plans and Principles for Public Data Switched Network in France. Proceedings of the IEEE. 1972, t. 60, nr 11, s. 1370 - 1374.
- [17] EABLER H.G.: Plans for Data Communications in Germany. Proceedings of the IEEE. 1972, t. 60, nr 11, s. 1374 - 1377.
- [18] CHEŁSTOWSKI T., SOBANIEC J.: Środki techniczne JSEMC. Urządzenia teleprzetwarzania i transmisji danych oraz przygotowania danych JS EMC. Informatyka, nr 10, s. 11.
- [19] RAWSKI R.: Środki i systemy teleprzetwarzania danych JSEMC. Informatyka, 1973, nr 11.
- [20] GRÖBER A., MONJAU J.: Nachrichtenvermittlungssystem mit dem Rechner ROBOTRON 4000. Rechentchnik Datenverarbeitung, Beiheft 1, 1973, s. 33 - 36.
- [21] Sovmestnaja rabota EVM 1010b - Mińsk-32. Prospekt firmy VIDEOTON.
- [22] BAZEWICZ M.: Wielotematyczny abonencki system cyfrowy na Politechnice Wrocławskiej. Informatyka, 1973, nr 9, s. 12.

- [23] ŻELAZNY R.: Cyfronet w Instytucie Badań Jądrowych w Świerku. Informatyka, 1973, nr 10, s. 6.
- [24] Streszczenie materiałów na konferencję naukową "Sieci komputerowe, 1973. PAN. Komitet Informatyki.
- [25] WITHINGTON F.G.: The Next (and last ?) Generation. Datamation, 1972, maj, s.71.

