

**OŚRODEK POSTĘPU TECHNICZNEGO
W KATOWICACH**

Eugeniusz Kurzydem

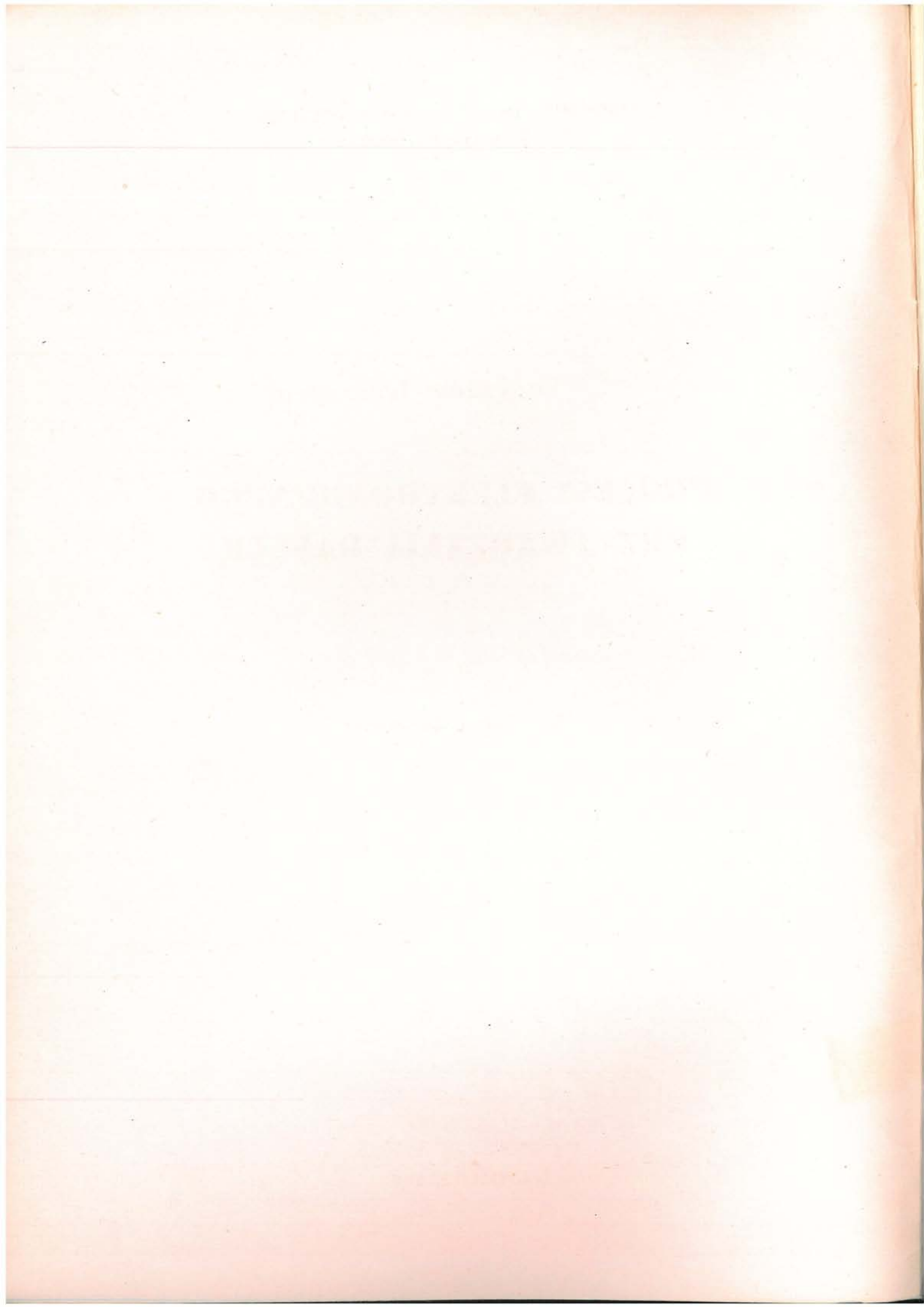
**PROCESY ELEKTRONICZNEGO
PRZETWARZANIA DANYCH**

Materiały szkoleniowe

do użytku wewnętrznego

KATOWICE 1973

136



OŚRODEK POSTĘPU TECHNICZNEGO

w Katowicach

Eugeniusz K U R Z Y D E M

PROCESY ELEKTRONICZNEGO PRZETWARZANIA

DANYCH

Materiały szkoleniowe

Katowice, 1973 r.

Spis treści

1. Opis danych.

- 1.1. Znak.
- 1.2. Pole informacyjne,
- 1.3. Grupa pól.
- 1.4. Rekord informacyjny.
- 1.5. Zbiór danych.
- 1.6. Bank danych systemu.

2. Podstawowe przebiegi przetwarzania.

- 2.1. Przebieg scalania /merging/.
- 2.2. Przebieg dobierania /collating/.
- 2.3. Przebieg sortowania /sorting/.
- 2.4. Przebieg rozdziału /separating/.
- 2.5. Przebieg konwersji /converting/.
- 2.6. Przebieg wybierania.
- 2.7. Przebieg aktualizacji prostej.

3. Spis rysunków

4. Wykaz literatury.

1. Opis danych

Przedmiotem działania systemów elektronicznego przetwarzania są dane.

W procesie przetwarzania danych występuje zazwyczaj kilka podstawowych typów danych. Rozważania w niniejszym rozdziale dotyczyć będą definicji i hierarchicznej struktury danych złożonych z wyrażeń składających się z danych elementarnych.

Na wstępie trzeba zaznaczyć, iż mamy do czynienia z dwoma głównymi klasyfikacjami struktur danych. Pierwsza z nich wynika z właściwości i cech fizycznych sprzętu dostępnego w procesach elektronicznego przetwarzania danych. Klasyfikacja ta zwana jest klasyfikacją Komputerowo zorientowana ^{1/}. Podstawowe jednostki danych tej klasyfikacji to:

- bit - czyli cyfra liczby przedstawiona w dwójkowym systemie liczenia,
- byte - czyli zespół ośmiu bitów, przy pomocy którego możemy przedstawić jeden dowolny znak,
- słowo - czyli zespół większej ilości bitów niż osiem, np. dwanaście, dwadzieścia cztery, czterdzieści osiem, /przy czym istnieją również komputery, w których słowo posiada zmienną długość/.
- blok fizyczny na nośniku wejściowym, wyjściowym lub pamięci zewnętrznej,
- szpula taśmy magnetycznej lub pakiet dysku magnetycznego

Druga klasyfikacja struktur danych wynika z potrzeb opisu danych dla całokształtu procesu przetwarzania i zwana jest klasyfikacją, procesowo zorientowana.

Podstawowe jednostki tej klasyfikacji to:

- znak,
- pole informacyjne,
- grupa pól,
- - - - -

1/ zgodnie z nomenklaturą podaną w /2/.

- rekord informacyjny,
- zbiór,
- bank danych systemu.

Poniżej postaramy się możliwie dokładnie omówić poszczególne jednostki struktury klasyfikacji procesowo zorientowanej. Klasyfikacja komputerowo zorientowana nie jest przedmiotem rozważań niniejszego opracowania.

1.1. Znak.

Znak jest najmniejszą jednostką danych przetwarzania. Różniamy cztery typy znaków: znak numeryczny dziesiętny, znak alfabetyczny, znak alfanumeryczny i znak binarny.

Znak numeryczny może przyjmować jedną z poniższych wartości: 0, 1, 2, 3, 4, 5, 6, 7, 8 i 9.

Znak alfabetyczny może przyjmować jedną z wartości alfabetu od A do Z.

Znak alfanumeryczny może przyjmować jedną z wartości spośród w/w dziesięciu cyfr dziesiętnych, liter alfabetu, lub innych symboli pisarskich jak np: kropka, przecinek, +, -, =, /, *, spacja /czyli blank - puste miejsce/.

Znak binarny /zwany również znakiem dwójkowym/ posiada wartość nadanego mu kodu dwójkowego, zgodnie z przyjętym kodem maszyny.

Należy w tym miejscu zaznaczyć, iż zarówno odpowiednim literom jak i symbolom pisarskim, a także cyfrom przyporządkowany zostaje wewnątrz maszyny jednoznaczny kod dwójkowy w odpowiedniej /zgodnie z przyjętą organizacją wewnętrzną komputera/ liczbie bitów /np. sześciu lub ośmiu/.

1.2. Pole informacyjne.

Pojedyńcze znaki najczęściej występują w większych jednostkach, które zwane są polami informacyjnymi. I tak np. liczba 528 - składająca się z trzech znaków numerycznych - będzie

połem informacyjnym. Również zapis KATOWICE składający się z ośmiu znaków alfabetycznych - będzie stanowił pole informacyjne. W zależności od charakteru znaków - numerycznych czy alfabetycznych reprezentowanych przez dane pole - poszczególne pola informacyjne nazywać będziemy odpowiednio polami numerycznymi bądź polami alfabetycznymi. W praktyce bardzo często występują również pola informacyjne, których zakodowana treść jest kombinacją pola numerycznego i alfabetycznego - łącznie z innymi symbolami pisarskimi /czyli znakami alfanumerycznymi/. Pola te noszą nazwę pól alfanumerycznych. Przykładem takiego pola może być pole zawierające adres jakiejś instytucji np. KATOWICE, UL. WIECZORKA 35/36.

Przez pole informacyjne należy zatem rozumieć skończony ciąg znaków przeznaczonych do zapisu informacji.

Z pojęciem pola informacyjnego nierozdzielnie związane jest pojęcie rozmiaru pola. Przez rozmiar pola rozumieć należy maksymalną ilość znaków, jaka może być zapisana w danym polu.

Charakter i rozmiar pola przyjęto w praktyce oznaczać następująco:

- symbolem 9 oznacza się pola informacyjne numeryczne, zaś jego rozmiar /czyli maksymalną ilość znaków, jaka może w polu wystąpić/ podaje się w nawiasie występującym bezpośrednio za symbolem. Np. 9(4), 9(7), 9(11).
- symbolem A oznaczone są pola alfabetyczne, a ich rozmiar określa również liczba w nawiasie, występującym bezpośrednio za tym symbolem. Np. A(4), A(25), A(17), A(76).
- symbolem X oznacza się pola informacyjne alfanumeryczne przy czym podobnie jak w przypadku pól numerycznych czy alfabetycznych, o rozmiarze pola informuje liczba ujęta w nawiasie. Np. X(12), X(32), X(120).

Przyjmuje się zazwyczaj konwencję, że pola informacyjne numeryczne są adiustowane do prawego brzegu pola. Oznacza to, że zapisu dokonuje się od najmniej znaczącej pozycji

cyfrowej liczby. Natomiast zarówno pola alfabetyczne jak i pola alfanumeryczne adiustowane są do lewego brzegu pola. Poniższa tablica przedstawia przykłady opisu różnych pól.

Tablica nr 1.

LP.	Charakter i rozmiar pola	Zawartość informacyjna pola	Określenie pola
1.	9(6)	<u>1712191312171</u>	pole numeryczne
2.	9(9)	<u>1 1 17121815131211</u>	pole numeryczne
3.	A(10)	<u>111N1F1O1R1M1A1C1J1A1</u>	pole alfabetyczne
4.	A(11)	<u>111K1S111N1S1K111 1 1 1</u>	pole alfabetyczne
5.	A(11)	<u>1F1E1N1D1E1R1O1W1S1K111</u>	pole alfabetyczne
6.	X(14)	<u>1W1-1W1A1 1U1L1.1M1111A1 1 71</u>	pole alfanumeryczne
7.	X(14)	<u>1K1A1T1O1W111C1E1-11121 1 1 1</u>	pole alfanumeryczne
8.	9(6)	<u>1 1 1 1912181</u>	pole numeryczne

W przypadku pól numerycznych określenie rozmiaru pola nie należy identyfikować z zawartością informacyjną pola. Symbol 9/5/ nie oznacza pola numerycznego, w którym zapisana jest wartość pięć. Wartość pola opisanego przez 9(5) może bowiem być zawarta w granicach od wartości 00000 do 99.999.

W tablicy nr 1 w pozycjach 1 i 8 rozmiar pola jest identyczny, - wartości natomiast różne.

To samo dotyczy oczywiście pól alfabetycznych i alfanumerycznych /patrz tabl. nr 1 poz. 4 i 5 oraz 6 i 7/. Proszę również zwrócić uwagę na adiustację pól informacyjnych numerycznych oraz pól alfabetycznych i alfanumerycznych.

Należy zwrócić uwagę na to, że każde pole alfabetyczne jest równocześnie polem alfanumerycznym. Twierdzenie odwrotne natomiast nie jest prawdziwe. Ponadto w każdym polu alfanumerycznym może występować zapis "czysto" numeryczny, np. w polu zadeklarowanym jako X(4) może wystąpić wartość 2734.

1.3. Grupa pól

Kolejną, większą jednostką klasyfikacji procesowo zorientowanej jest grupa pól. Grupa pól jest to zespół kilku pól tworzących całość z punktu widzenia danego procesu przetwarzania. Przykładem grupy pól mogą być:

- pola: dzień, miesiąc, rok-mogą być grupą pól, którą nazwać można wspólnym mianem "data".
- pola: długość, szerokość, wysokość-mogą być grupą pól, którą nazwać można wspólnym mianem "wymiary".
- pola: miejscowość, ulica, numer domu-mogą stanowić grupę pól zwaną "adresem".
- pola: nazwisko, imię, adres, może być grupą pól nazwaną "nadawca".

We wszystkich wyżej podanych przykładach pola informacyjne wchodzące w skład grupy pól mogą tworzyć z punktu widzenia przetwarzania - rodzaj identyfikatora w danym procesie przetwarzania.

1.4. Rekord informacyjny

Zespół pól informacyjnych związanych z sobą, tworzących logiczną całość w procesie przetwarzania danych nazywać będziemy rekordem informacyjnym.

Rekordem informacyjnym nazywać zatem możemy zespół pól informacyjnych stanowiących kompletną informację o pracowniku - w procesie przetwarzania danych związanych z gospodarką osobową w jakiejś instytucji lub przedsiębiorstwie. Poszczególne pola informacyjne będą zawierały:

- numer identyfikacyjny pracownika
- nazwisko,
- imię,
- data urodzenia,
- stanowisko służbowe,
- wykształcenie,
- płeć,
- stan cywilny,
- stawkę zaszeregowania,
- miejsce zamieszkania,
- ulica,
- nr domu,
- przynależność do organizacji politycznych, społecznych,
- znajomość języków obcych,
- informacje o deputacie węglowym,

Przykładem rekordu może również być zespół pól zawierający całość informacji związanych z zamówieniem. Pola informacyjne będą wtedy zawierały:

- numer ewidencyjny zamówienia,
- datę wystawienia
- datę wpływu zamówienia
- specyfikację zamawianego wyrobu
- wymiary wyrobu
- zamówiona ilość
- nazwa zamawiającego
- adres zamawiającego
- warunek płatności
- żądania specjalne w zakresie opakowania, transportu
- żądany termin dostawy.

Nazwa "rekord" wywodzi się od angielskiego słowa "record" /zapis/, a którego pochodzenie powstało na bazie powszechnego nawyku administracji do używania typowych dokumentów, jak np. zamówienia, dowody przyjęcia, dowody wydania, faktury itp. obejmujące przecież szereg szczegółowych pól informacyjnych. Rekord jest zatem odpowiednikiem jednostkowego dokumentu lub pozycji dokumentu wielopozycyjnego.

Rekord informacyjny może być przechowywany na różnych maszynowych nośnikach informacji; kartach perforowanych, magnetycznych nośnikach /taśmach lub dyskach magnetycznych/, w pamięci operacyjnej komputera, wreszcie jako linia druku na drukarce wierszowej czy też karta perforowana na dziurkarce kart.

Rekordy składają się zwykle z dwu podstawowych grup pól:

- 1/ grupy pól przeznaczonej do identyfikacji danego rekordu, tę grupę pól nazywać będziemy kluczem rekordu.
- 2/ grupy pól informacyjnych nie wchodzących do klucza rekordu, tę grupę pól nazywać będziemy ciałem rekordu.

W procesie przetwarzania mamy do czynienia najczęściej z trzema rodzajami rekordów informacyjnych.

- rekordy informacyjne wejściowe,
- rekordy informacyjne wyjściowe,
- rekordy informacyjne wejściowo-wyjściowe

Rekordy informacyjne wejściowe, są to rekordy wprowadzane do pamięci operacyjnej komputera z urządzeń wejścia, np. czytnika kart lub czytnika taśmy perforowanej. W pierwszym przypadku maszynowym nośnikiem informacji rekordu informacyjnego jest zatem karta perforowana. Wymiary rekordu ograniczone są pojemnością karty perforowanej /tj. przeważnie 80 znaków/.

Na ogół przyjmuje się bowiem, że na jednej karcie perforowanej przechowywać można jeden rekord. W drugim przypadku maszynowym nośnikiem informacji jest taśma perforowana. Na taśmie perforowanej wielkość rekordu może być dowolnie duża. Poszczególne rekordy informacyjne oddzielone są od siebie tzw. znacznikiem końca rekordu, tj. umownym symbolem /a w zasadzie umowną kombinacją perforacji, w zależności od przyjętego kodu/ oznaczającym koniec rekordu informacyjnego.

Rekordy informacyjne wyjściowe są to rekordy wyprowadzane z pamięci operacyjnej komputera do urządzeń wyjścia np. drukarki wierszowej, dziurkarki kart perforowanych lub dziurkarki taśmy papierowej.

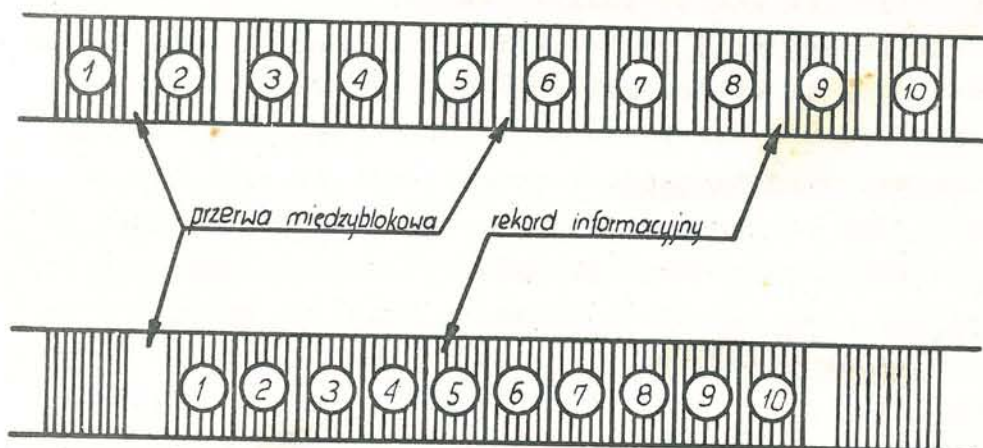
W pierwszym przypadku wymiary rekordu ograniczone są ilością znaków w linii urządzenia drukującego. Przyjmuje się bowiem, że jedna linia druku przeznaczona jest na jeden rekord informacyjny. Podobnie w drugim przypadku: wymiary rekordu ograniczone są pojemnością karty perforowanej. Analogicznie jak to miało miejsce z rekordami informacyjnymi wejścia - przyjmuje się, iż na jednej karcie perforowanej przechowywać można jeden rekord informacyjny. W przypadku gdy maszynowym nośnikiem informacji rekordu wyjściowego jest taśma perforowana organizacja sygnalizacji wielkości rekordów wyjściowych odbywa się przy pomocy znaczników końca rekordu.

Rekordy informacyjne wejściowo-wyjściowe są to rekordy wprowadzane lub wyprowadzane do i z pamięci operacyjnej komputera z i do jednostek pamięci zewnętrznej w trakcie procesu przetwarzania. Ich maszynowymi nośnikami informacji są zatem magnetyczne urządzenia pamięciowe komputera. Organizacja zapisu rekordów na magnetycznych nośnikach informacji wynika z technicznych własności tych urządzeń. Organizację zapisu na tego rodzaju nośnikach informacji zilustrujemy poniżej na przykładzie taśmy magnetycznej.

Dla dokonania zapisu lub odczytu z taśmy magnetycznej musi się ona poruszać z pewną określoną szybkością. W momencie kiedy nie dokonujemy ani zapisu, ani odczytu taśma musi się zatrzymać. Od chwili startu do chwili osiągnięcia wymaganej szybkości przesuwu taśmy - nie dokonuje się zapisu. Również od momentu dokonania ostatniego zapisu - do całkowitego zatrzymania się taśmy, pewna jej część nie zostanie zapisana. Między poszczególnymi zapisami na taśmie magnetycznej powstają zatem pewne przerwy, zwane przerwami międzyblokowymi. Mając to na uwadze zrozumiałą rzeczą jest, że osiągniemy tym pełniejsze wykorzystanie taśmy im mniej będzie przerw w zapisie. Jednostki zapisu powinny mieć zatem rozmiar rzędu

kilkuset lub nawet kilku tysięcy znaków uzyskując tym samym optymalne wykorzystanie nośnika.

Zapisując na taśmie pojedyncze rekordy informacyjne /których rozmiary oscylują w granicach stu znaków/ przerw międzyblokowych byłoby bardzo dużo. Rysunek nr 1 przedstawia w pozycji a/ zapis dziesięciu rekordów na taśmie magnetycznej, przy czym każdy z tych rekordów został zapisany osobno. W pozycji b/ - te same rekordy zapisane zostały w skumulowanej jednostce przesyłu.



Rys. nr 1. Organizacja zapisu rekordów w bloku fizycznym na taśmie magnetycznej.

Skumulowana jednostka przesyłu nosi nazwę bloku fizycznego. W zależności od potrzeb blok fizyczny składać się może z kilku lub kilkunastu rekordów informacyjnych. W krańcowym przypadku blok fizyczny może oczywiście składać się również tylko z jednego rekordu informacyjnego. Wielkość bloku fizycznego jest w większości współczesnych komputerów zmienna. Aby umożliwić technikę zapisu i odczytu rekordów w blokach fizycznych, w pamięci operacyjnej komputera alokowany jest

automatycznie obszar pamięci, zdolny do przechowania jednego bloku fizycznego. Realizowany program przetwarzania ma jednak w zasadzie dostęp tylko do jednego rekordu informacyjnego danego bloku. Jeśli rekord ostatnio przetwarzany był ostatnim rekordem informacyjnym bloku fizycznego następuje przesłanie nowego bloku do lub z pamięci operacyjnej.

Technika zapisu na innych magnetycznych nośnikach informacji jest zbliżona do wyżej opisanej.

W większości współczesnych komputerów możliwy jest również zapis w jednym bloku fizycznym rekordów informacyjnych o zmiennym rozmiarze /różnej długości/. Możliwość taka zwiększa wydatnie efektywność wykorzystania nośnika.

Należy dodatkowo wyjaśnić, w odniesieniu do magnetycznych nośników informacji, iż oprócz rekordów informacyjnych w trakcie procesu przetwarzania niezbędne są rekordy pomocniczo - identyfikacyjne, zwane etykietami. W odniesieniu do taśm magnetycznych z reguły przyjmuje się następującą konwencję identyfikacji: na początku taśmy magnetycznej znajduje się rekord identyfikacyjny /tzw. etykieta początku zbioru/, dalej występują kolejno rekordy informacyjne, zaś na końcu taśmy - drugi rekord identyfikacyjny /tzw. etykieta końca zbioru/. Rekordy identyfikacyjne umożliwiają jednoznaczne sprawdzenie /poprzez specjalny program organizacyjny komputera/ zawartości danego nośnika pamięci zewnętrznej, zabezpieczając je tym samym przed niewłaściwym użyciem w ramach danego procesu informacyjnego.

1.5. Zbiór danych

W trakcie procesu przetwarzania danych przez komputer przepływają strumienie danych. Strumienie danych wchodzące bądź wychodzące przez urządzenia peryferyjne komputera - w postaci ciągu/zespołu/rekordów informacyjnych-nazywać będziemy zbiorami danych.

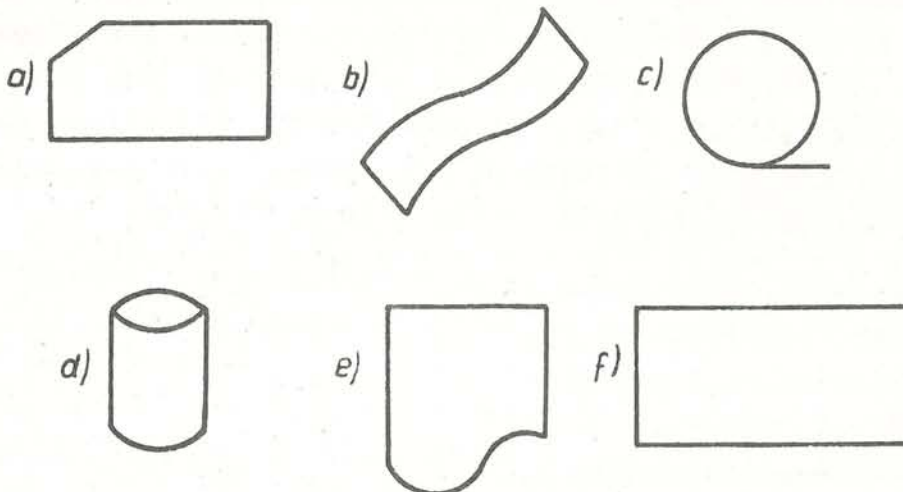
Takie określenie w odniesieniu do zbioru danych nie jest ani pełne ani jednoznaczne. W dalszym ciągu niniejszego punktu metodę kolejnych przybliżeń będziemy starali się wyjaśnić i uzupełnić to pojęcie.

W praktyce codziennej zbiory danych - jako ciągi /zespoły/ rekordów informacyjnych - przechowywane są również na maszynowych nośnikach informacji takich jak taśmy i dyski magnetyczne, karty perforowane, taśma dziurkowana itd. Mówimy zatem o zbiorze danych na kartach perforowanych, zbiorze drukarki /każda linia druku - jak pamiętamy stanowi jeden rekord informacyjny wyjściowy, całe wydawnictwo zatem tworzyć będzie zbiór/, zbiorze taśmy magnetycznej itd.

Można przyjąć następującą klasyfikację zbiorów danych:

- zbiory wejściowe zewnętrzne, dla których typowymi nośnikami będą karty perforowane, czy też taśma perforowana, związane odpowiednio z urządzeniem czytnika kart lub czytnika taśmy perforowanej ,
- zbiory wyjściowe zewnętrzne, dla których typowymi nośnikami informacji będą: papier na drukarce wierszowej /wydawnictwo, tabulogram/ karty perforowane na dziurkarce kart lub taśma perforowana na dziurkarce taśmy papierowej.
- zbiory podstawowe systemu, najczęściej związane z magnetycznymi nośnikami informacji.

Rysunek nr 2 przedstawia oznaczenia przyjęte w technice przetwarzania danych dla poszczególnych zbiorów.



Rys. nr 2. Oznaczenia zbiorów danych. a/ zbiór kart perforowanych, b/ zbiór taśmy perforowanej, c/ zbiór na taśmie magnetycznej, d/ zbiór na dysku magnetycznym, e/ zbiór drukarki wierszowej, f/ jednostka centralna /procesor/.

W praktyce codziennej mamy do czynienia nie ze zbiorami danych, których zawartość informacyjna jest stała i niezmienna, a kolejnymi generacjami zbiorów. Zbiór danych jest bowiem pojęciem dynamicznym. W trakcie przebiegu przetwarzania, zwanego przebiegiem aktualizacji, kolejną generacją zbioru zastąpimy przez następną, która w danym momencie będzie generacją najaktualniejszą z punktu widzenia zawartości informacyjnej. Pojęcie generacji zbiorów omówimy szczegółowo w pkt. 2.7. W tym miejscu należy jedynie dodać, że pojęcie generacji zbiorów związane jest najczęściej ze zbiorami podstawowymi.

W odniesieniu do zbiorów danych zmuszeni jesteśmy podać szereg parametrów charakteryzujących zbiory. Wartości tych parametrów decydują o doborze najbardziej racjonalnych metod przetwarzania oraz ściśle z metodą przetwarzania związane: organizację zbiorów, metody dostępu, sposób adresowania oraz wybór nośnika informacji.

Parametrami tymi są:

- Aktywność zbioru danych która określana jest stosunkiem ilości rekordów do których żądany jest dostęp w przeciętnym przebiegu - do ogólnej ilości rekordów w zbiorze. Aktywność zbioru podaje się w procentach. Jeżeli aktywność zbioru jest niska, zbiór danych winien być na ogół tak zorganizowany aby dostęp do rekordów był bezpośredni, bez konieczności przeglądania całego zbioru. Na ogół przyjmuje się, że jeżeli aktywność zbioru jest niższa niż 5 %, to zbiór winien być przetwarzany z dostępem bezpośrednim /dyski magnetyczne/.
- Stabilność objętości zbioru, związana jest z ilością rekordów wprowadzanych do zbioru i rekordów wymazywanych /anulowanych/ w zbiorze w określonej jednostce czasu. Zbiory w których ilość wymian rekordów jest niska nazywamy zbiorami statycznymi w odróżnieniu od zbiorów dynamicznych, w których ilość wymian rekordów jest wysoka. Te ostatnie skolei dzielimy na:
 - zbiory dynamiczne rosnące
 - zbiory dynamiczne malejące.
 - zbiory dynamiczne stabilne.

Stabilność objętości zbioru ma decydujący wpływ na organizację zbioru.

- Wielkość zbioru, czyli ilość rekordów w zbiorze, ma bardzo wielki wpływ na metodę przetwarzania. Jeśli zbiór jest mały, a zatem i czas jego przetwarzania jest krótki, problem doboru metod organizacji zbioru i dostępu nie ma istotnego znaczenia. Jeżeli natomiast zbiór jest duży, a zwłaszcza jeśli ilość informacji przekracza objętość wybranego nośnika informacji /tom dysku magnetycznego, szpula taśmy magnetycznej/ wybór metody dostępu i organizacji zbioru jest sprawą niezmiernie ważną.

W odniesieniu do wyżej podanych parametrów należy przypomnieć podstawowe cechy najczęściej stosowanych magnetycznych nośników informacji tj. taśm magnetycznych i dysków magnetycznych.

Taśmy magnetyczne, ze względu na ich fizyczne właściwości, wymagają sekwencyjnej organizacji przetwarzania. Na tym nośniku informacji niemożliwa jest jakakolwiek zmiana informacji, bez konieczności przepisania całej zawartości taśmy.

W odróżnieniu od taśm, dyski magnetyczne - poprzez możliwość ustawienia głowic nad dowolną /nie koniecznie kolejną/ ścieżką tarcz, pozostawienia głowic nad wybranymi ścieżkami oraz możliwość bezpośredniego adresowania każdego bloku fizycznego umożliwiają dokonanie zmian w zapisach w tym samym miejscu dysku, bez konieczności przepisania całego zbioru danych na nim zapisanych. Udogodnienia organizacji przetwarzania wynikające z zastosowania dysków magnetycznych jako nośnika informacji narzucają jednak określone konsekwencje /ograniczenia, odpowiedni reżym postępowania/ w odniesieniu do projektanta systemu.

Warto również nadmienić, iż dysk magnetyczny jako nośnik informacji, kosztuje dziesięć razy tyle ile taśma magnetyczna, przy czym jedna szpula taśmy magnetycznej może przechowywać trzy razy tyle informacji ile jeden tom dysku magnetycznego. Przez organizację zbiorów danych rozumie się sposób uporządkowania ich na nośniku informacji.

Rozróżnia się dwa poziomy organizacji:

- organizację rekordów w bloku,
- organizację bloków w zbiorze.

Identyfikatorem rekordu w zbiorze/bloku/jest grupa pól zwana kluczem rekordu /patrz pkt. 1.4/.

Rekordy w bloku mogą być umieszczone w sposób:

- uporządkowany, tj. według wzrastającej lub malejącej wartości klucza rekordu, taka organizacja nazywa się powszechnie organizacją sekwencyjną,
- nieuporządkowany; taka organizacja zwana jest organizacją przypadkową /random/.

Dostęp do żadanego rekordu następuje zawsze przez kolejne przeglądanie rekordów w bloku. Dla udostępnienia żadanego bloku do przetwarzania stosuje się najczęściej niżej opisane metody dostępu:

- dostęp bezpośredni, żądany blok jest wybierany do przetwarzania bez kolejnego przeglądania bloków zbioru /możliwy jedynie przy zastosowaniu dysków magnetycznych/
- dostęp przez kolejne przeglądanie, kolejne bloki zbioru są przeglądane, aż do napotkania bloku żadanego; kolejne przeglądanie może w tym przypadku oznaczać fizyczną kolejność /np. taśma magnetyczna/ lub logiczną kolejność; przeglądanie może odbywać się od początku zbioru, lub innego, wyznaczonego przez programistę miejsca zbioru.

Zbiór podstawowy uporządkowany wg malejącej lub rosnącej wartości klucza rekordów wchodzących w skład zbioru - nazywać będziemy kartoteką.

Mówić zatem będziemy o Kartotece Pracowników, Kartotece Materiałowej, Kartotec-e Zamówień itd.

1.6. Bank danych systemu.

Przez bank danych systemu rozumieć będziemy zespół wszystkich aktualnych generacji zbiorów podstawowych systemu.

2. Podstawowe przebiegi przetwarzania

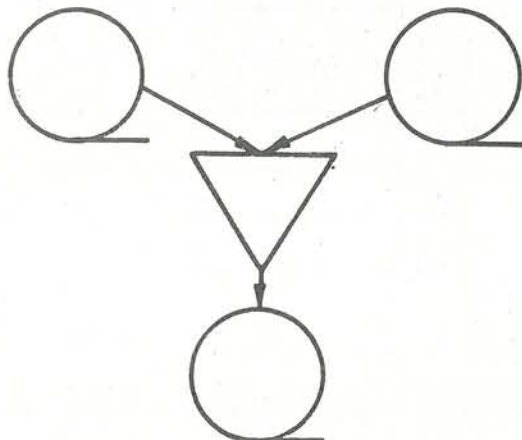
Ze względu na powszechność stosowania w Polsce taśm magnetycznych jako nośnika informacji zbiorów podstawowych systemu, a także ze względu na rozmiary niniejszego opracowania w dalszych punktach omówimy jedynie podstawowe przebiegi, związane z sekwencyjnym przetwarzaniem danych.

Przez przebieg przetwarzania należy rozumieć zespół operacji wykonywanych na jednym lub więcej zbiorach wejściowych, realizowanych w sposób automatyczny przez komputer na podstawie jednego programu. W wyniku działania przebiegu powstaje jeden lub więcej zbiorów wyjściowych. W ramach niniejszej pracy omówiono jedynie podstawowe, najbardziej charakterystyczne typy przebiegów. Sporadycznie tylko przedstawiono również pewne odmiany /wersje/ tych "klasycznych" przebiegów. Główny nacisk położono na tzw. przebiegi elementarne, występujące najczęściej w systemach elektronicznego przetwarzania danych.

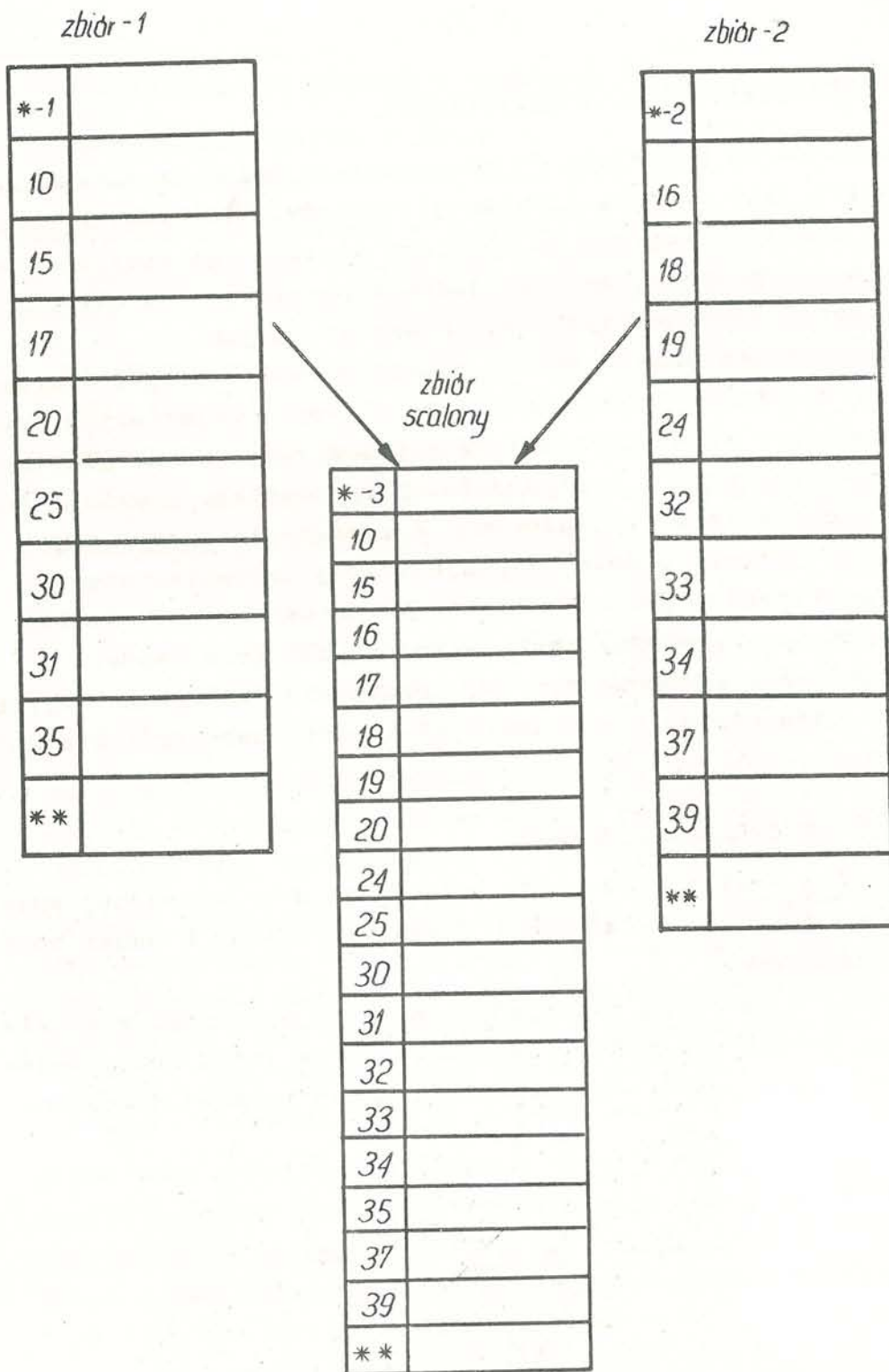
2.1. Przebieg scalania /merging/

Przebieg scalania służy do połączenia dwu lub więcej zbiorów zgodnie uporządkowanych /czyli Kartotek/ w jeden zbiór uporządkowany.

Zgodnie z powyższą definicją w prostym przebiegu scalania biorą udział dwa zbiory wejściowe, oraz jeden zbiór wyjściowy. Oznaczenie tego przebiegu przedstawiono na rysunku 3.



Rys. nr 3. Przebieg scalania prostego.



Rys. nr 4. Przykład scalania zbiorów o różnych wartościach kluczy.

Prześledźmy algorytm przebiegu na prostym przykładzie.

Rysunek 4 przedstawia w sposób skrajnie uproszczony zawartość zbiorów wejściowych /zbiór-1 i zbiór-2/ oraz zawartość zbioru wyjściowego /zbiór scalony/. W poszczególnych rekordach wydzielono wyraźnie pola stanowiące klucz rekordów. Przyjęto w nich symboliczne wartości. W praktyce oczywiście ani miejsca kluczy w rekordach ani też ich rzeczywista wartość nie musi być taka jaką podano w przykładzie. Z reguły mamy wtedy do czynienia ze zbiorami o większej ilości rekordów. Wartości kluczy nie muszą również być wartościami wyłącznie numerycznymi. Mogą to być zarówno wartości alfabetyczne jak i wartości alfanumeryczne. W przykładzie przyjęto uporządkowanie zbiorów wg wzrastającej wartości kluczy.

Pominięto również sposób organizacji rekordów w bloku fizycznym. Symbolicznie przyjęto także zawartość informacyjną rekordów pomocniczo-identyfikacyjnych /czyli tzw. etykiet/ /patrz pkt. 1.4./.

Procedura postępowania będzie następująca:

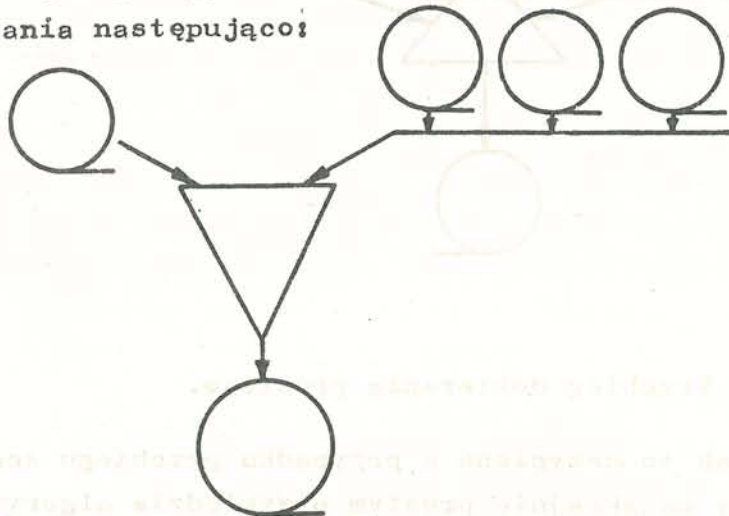
Do pamięci operacyjnej komputera wczytane zostaną rekordy pomocniczo-identyfikacyjne zbioru-1 /etykieta * -1/ i zbioru-2 /etykieta * -2/. Procedura ta zezwala na jednoznaczne określenie czy te właśnie zbiory mają wejść do przebiegu scalania. Po tej kontroli następuje zapis rekordu pomocniczo-informacyjnego /etykieta * -3/ w zbiorze wyjściowym. Następnie wczytane zostaną do pamięci operacyjnej pierwsze rekordy obu zbiorów wejściowych. /rekord o wartości klucza 10 ze zbioru-1 oraz rekord o wartości klucza 16 ze zbioru-2/. Następuje porównanie wartości kluczy. Rekord, którego wartość klucza jest mniejsza zostaje zapisany w zbiorze wyjściowym. W naszym przypadku jest to rekord ze zbioru-1 o wartości klucza 10. Rekord o wartości klucza 16 /ze zbioru-2/ przechowywany /zapamiętany/ jest w pamięci operacyjnej. Do pamięci wczytany zostaje rekord ze zbioru o wartości klucza 15. Następuje porównanie wartości/obu rekordów egzystujących w danym momencie w pamięci operacyjnej. W wyniku porównania wartości kluczy w zbiorze wyjściowym zapisany zostanie rekord ze zbioru-1 o wartości klucza 15. Rekord ze zbioru-2 /o wartości klucza 16/ nadal zapamiętany jest w pamięci operacyjnej. Kolejny trzeci rekord ze zbioru-1 wprowadzony zostanie do pamięci /rekord wartości klucza 17/. Następuje porównanie wartości kluczy. Tym razem wartość klucza rekordu zbioru-1 jest większa od wartości klucza rekordu zbioru-2. W zbiorze wyjściowym zapisany zostanie wobec tego rekord ze zbioru-2 o wartości klucza 16. Rekord zbioru-1 /o wartości klucza 17/ przechowany zostanie w pamięci operacyjnej. W kolejnym zatem kroku do pamięci wczytany zostanie rekord ze zbioru-2 /o wartości klucza 18/. Następuje porównanie wartości kluczy. Rekord o mniejszej wartości klucza zapisany zostanie w zbiorze wyjściowym. Dalsze kroki postępowania są identyczne z opisanymi powyżej aż do momentu, w którym w pamięci operacyjnej przechowywany jest rekord ze zbioru-1 o wartości klucza 35, /przy czym wcześniej ze zbioru-2 na zbiór wyjściowy zapisane zostały

rekordy o wartościach kluczy 32, 33 i 34/ a do pamięci wczytany został rekord ze zbioru-2 o wartości klucza 37.

W wyniku porównania wartości kluczy rekord ze zbioru-1 /wartość klucza 35/ zapisany zostaje w zbiorze wyjściowym. i ze zbioru-1 wczytany ma zostać do pamięci operacyjnej kolejny rekord. Rekord o wartości klucza 35 był jednak już ostatnim rekordem zbioru-1. Przeczytana zostanie etykieta końca zbioru-1. Stanowi to sygnał dla algorytmu przebiegu scalania, iż w dalszym toku postępowania należy jedynie przepisać rekordy ze zbioru-2 na zbiór wynikowy, aż do momentu wystąpienia w zbiorze-2 etykiety końca zbioru.

Zapisany wobec tego zostaje rekord o wartości klucza 37 /egzystował w danym momencie przetwarzania już w pamięci operacyjnej/ a następnie przeczytany i przepisany na zbiór wyjściowy zostanie ostatni rekord /wartość klucza 39/ zbioru-2. W kolejnym kroku odczytana zostanie etykieta końca zbioru-2. Stanowi to sygnał dla algorytmu przetwarzania, że zbiór scalony jest kompletny. Następuje zapis etykiety końca zbioru w zbiorze scalonym.

Przedstawiony powyżej algorytm postępowania odnosi się - jak to już zaznaczono - do prostego scalania dwu zbiorów. Zgodnie z definicją przebiegu scalania algorytm może przewidywać scalenie trzech bądź większej ilości zbiorów wejścia. Przebieg tego typu scalania oznacza się na schematach przetwarzania następująco:



Rys. nr 5. Przebieg scalania złożonego.

Mamy wtedy do czynienia z tzw. przebiegiem scalania złożonego.

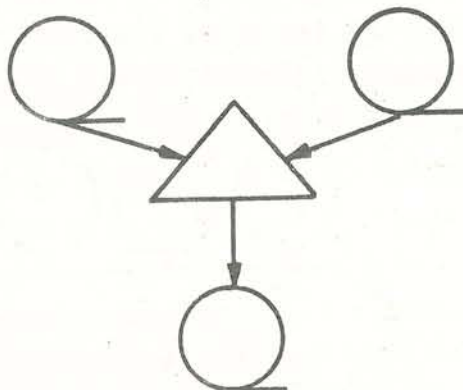
Należy również zaznaczyć, iż w przykładzie przyjęto fakt, iż rekordy obu zbiorów wejściowych są jednakowej i stałej długości. W praktyce mamy jednak do czynienia najczęściej ze zbiorami, w których rekordy są zmiennej długości. Najczęściej również przebieg scalania zbiorów daje możliwość otrzymania informacji /np. na konsoli operatorskiej/ o łącznej ilości rekordów występujących w nowoutworzonym zbiorze scalonym /dla celów kontrolnych/. Warto w tym miejscu zaznaczyć, iż przebieg scalania zbiorów wchodzi z reguły w zestaw standardowego oprogramowania komputera.

2.2. Przebieg dobierania /collating/.

Przebieg dobierania służy do wybierania z dwu lub więcej zbiorów zgodnie uporządkowanych /kartotek/ wspólnego ze względu na wartość kluczy, zbioru uporządkowanego.

Przebieg dobierania prosty, to taki przebieg w którym biorą udział dwa zbiory wejściowe oraz jeden zbiór wyjściowy.

Symboliczne oznaczenie prostego przebiegu dobierania przedstawia rysunek 6.



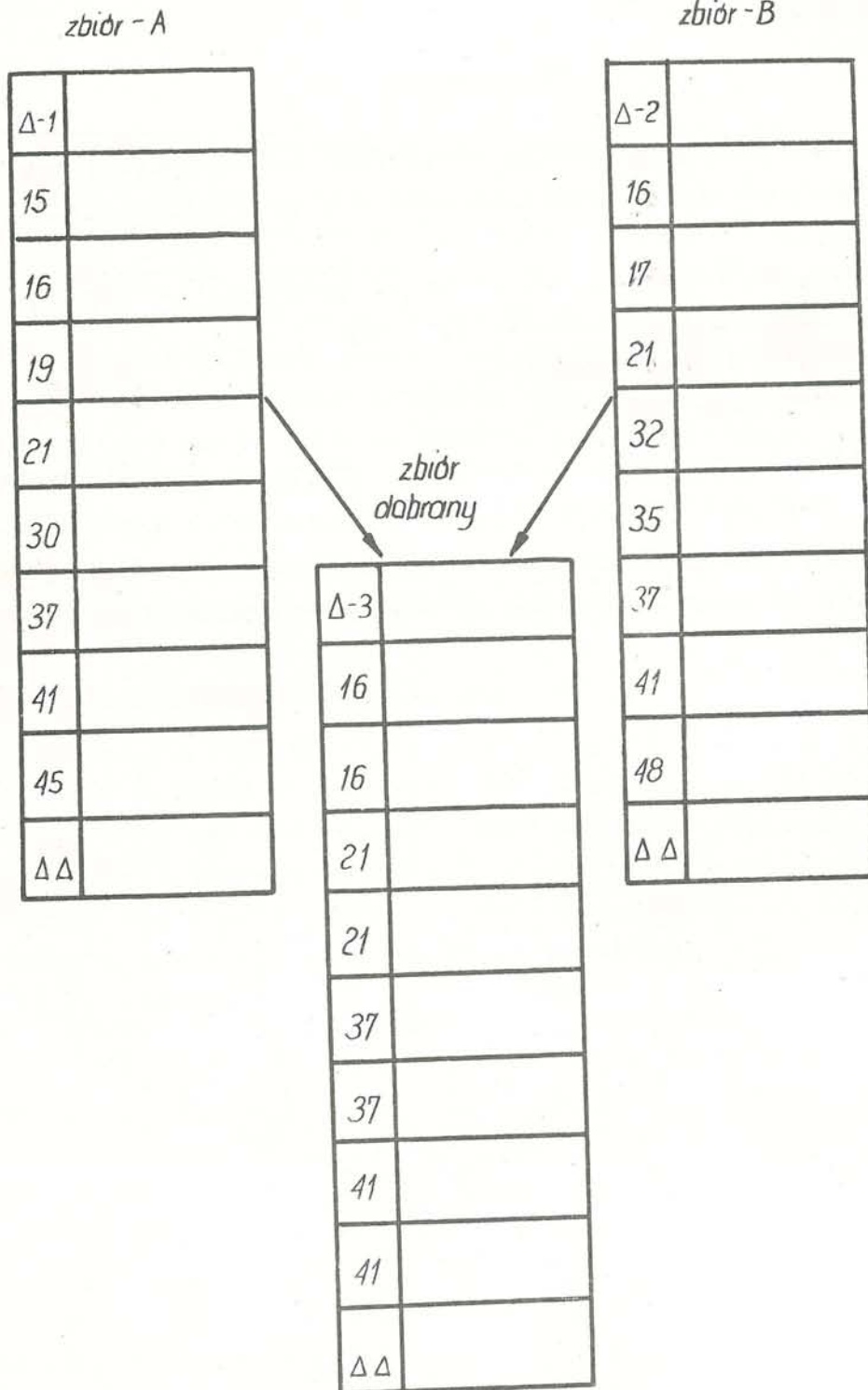
Rys. nr 6. Przebieg dobierania prostego.

Podobnie jak to uczyniono w przypadku przebiegu scalania - prześledźmy na skrajnie prostym przykładzie algorytm tego przebiegu. Uproszczenia w zakresie:

- nie definiowania pełnej zawartości informacyjnej rekordów,
- symbolicznego oznakowania rekordów pomocniczo-identyfikacyjnych,
- przyjęcia numerycznych wartości kluczy,
- rosnącego uporządkowania zbiorów,
- rekordów stałej długości,
- małej ilości rekordów,

przyjęte w przypadku wyjaśniania procedury przebiegu scalania mają również i tutaj miejsce.

Rysunek 7 przedstawia zawartość zbiorów wejściowych /zbiór-A i zbiór-B/ oraz zawartość zbioru wyjściowego /zbiór dobrany/.



Rys. nr 7. Przykład dobierania prostego na podstawie wartości kluczy.

Po kontroli etykiet początków zbiorów i zapisaniu etykiety początku zbioru wyjściowego do pamięci operacyjnej wczytane zostają: pierwszy rekord zbioru-A /o wartości klucza 15/ oraz pierwszy rekord zbioru-B /o wartości klucza - 18/. Następuje porównanie wartości kluczy obu wczytanych rekordów czy są równe. W przypadku nierówności kluczy do zbioru wyjściowego nie zostaje wprowadzona żadna informacja. Rekord o wyższej wartości klucza przechowywany jest w pamięci operacyjnej/w naszym przykładzie rekord zbioru-B o wartości klucza 16/. Na miejsce rekordu o mniejszej wartości klucza, rekord ze zbioru-A o wartości klucza 15/ wczytany zostaje kolejny rekord tego zbioru. W naszym przykładzie wczytany zostanie rekord ze zbioru-A o wartości klucza 16. Następuje sprawdzenie czy wartości kluczy rekordów są równe. Jeśli tak - oba rekordy /ze zbioru-A i ze zbioru-B/ zapisane są na zbiorze wyjściowym. Ten przypadek ma właśnie miejsce w naszym przykładzie. Stąd obydwa rekordy o wartości kluczy 16 zapisane zostały w zbiorze dobranym. Przeczytane zostają kolejne dwa rekordy: jeden ze zbioru-A /wartość klucza 19/ oraz jeden ze zbioru-B /o wartości klucza 17/.

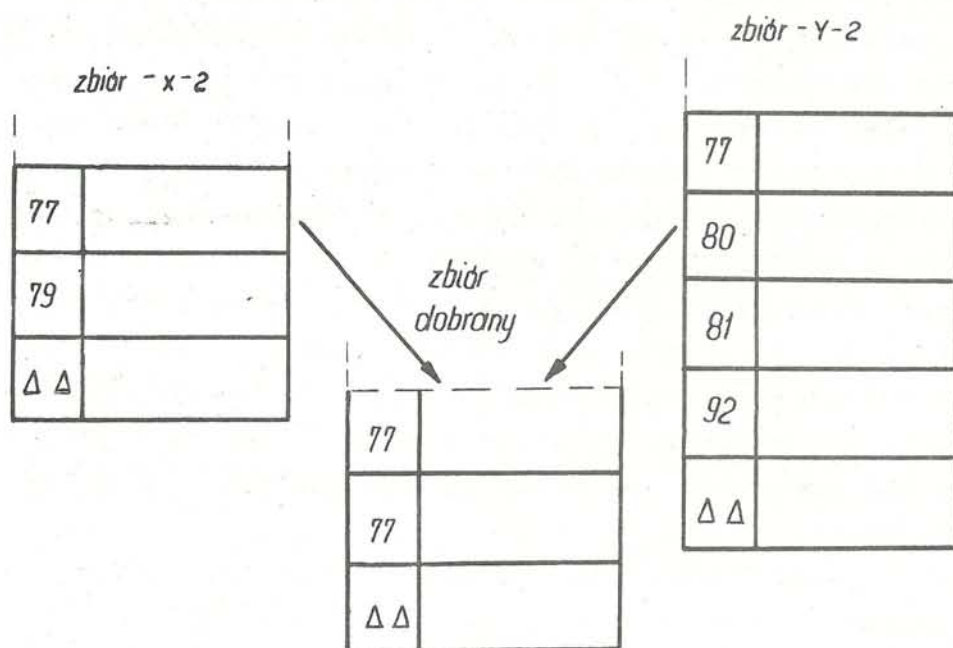
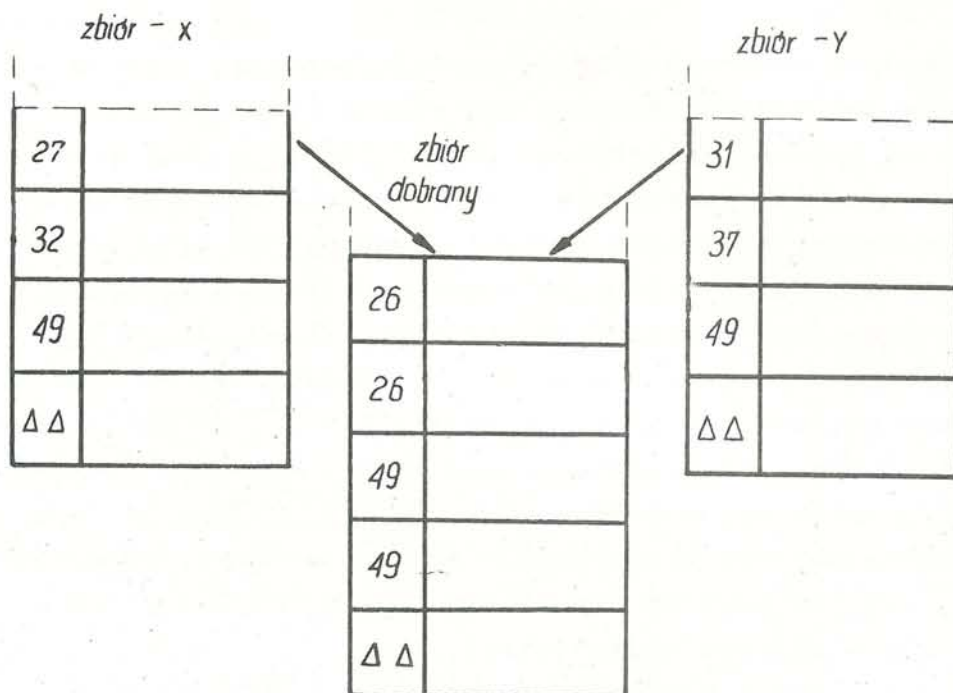
Z porównania wartości kluczy wynika, że nie są one sobie równe, w związku z czym ze zbioru-B wczytany jest następny rekord informacyjny /wartość klucza rekordu zbioru-B była niższa/. Jest nim rekord o wartości klucza 21. W dalszym ciągu wartości klucza nie są równe. Przeczytany zostaje wobec tego rekord o wartości klucza 21 ze zbioru-A. Następuje porównanie wartości kluczy. Są one teraz równe, dlatego też obydwa te rekordy informacyjne zapisane zostają w zbiorze dobranym. W dalszym ciągu procedura postępowania jest taka sama. Kolejne dwa rekordy /jeden ze zbioru-A i jeden ze zbioru-B/ wczytywane są do pamięci, następuje porównanie wartości kluczy itd.

W przykładzie naszym jeszcze rekordy o wartości kluczy 37 i 41 zostaną dobrane i zapisane w zbiorze wyjściowym. Po zapisaniu rekordów o wartości klucza 41 zgodnie z algorytmem - przeczytane zostaną ze zbioru-A rekord o wartości klucza 45 oraz ze zbioru B rekord o wartości klucza 48.

Wartości kluczy nie są równe.

Gdyby w zbiorze-A występowały jeszcze dalsze rekordy o odpowiednich wartościach kluczy procedura dobierania trwałaby dalej, zgodnie z wyżej opisanymi krokami postępowania. W przykładzie naszym rekord o wartości klucza 45 był ostatnim rekordem zbioru-A. Przeczytana zostanie etykieta końca zbioru. Pojawienie się etykiety końca zbioru w zbiorze-A kończy proces dobierania w naszym przykładzie. Zostaje zapisana etykieta końca zbioru w zbiorze dobranym /wyjściowym/.

Oczywiście układ dobranych rekordów w zbiorach wejściowych może być taki jak to przedstawiono niżej na rysunku 8 i 9.




Rys. 9. Przykład dobierania w którym jeden ze zbiorów wejściowych jest większy od drugiego zbioru wejściowego.

W przypadku gdy ostatnie rekordy zostają dobrane, a co za tym idzie - zapisane na zbiorze wyjściowym /rys. 8/ w kolejnym kroku zostaną przeczytane obie etykiety końca zbiorów. Przebieg dobierania oczywiście zostaje zakończony. W przykładzie z rysunku 9, natomiast, zbiór Y-2 zawiera o wiele więcej rekordów/po ostatnich rekordach o równej wartości klucza z rekordem zbioru X-2/. Proces dobierania zakończy się z chwilą natrafienia w przebiegu etykiety ze zbioru X-2.

Można zatem powiedzieć, iż pojawienie się etykiety końca zbioru w którymkolwiek ze zbiorów wejściowych lub w obu równocześnie - kończy przebieg dobierania.

W "czystym" przebiegu dobierania - a taki przebieg omawialiśmy wyżej - obydwa rekordy /ze zbioru-A i ze zbioru-B/ zapisane zostają w zbiorze wyjściowym.

W przykładzie, który poniżej przeanalizujemy /rysunek 10/ jednym ze zbiorów wejściowych jest kartoteka-A, drugim zaś Zbiór Wybiorczy-B. W Kartotece-A wyodrębniono wyraźnie pole klucza rekordu, nadając mu odpowiednie dla różnych rekordów wartości numeryczne. Uporządkowanie Kartoteki-A według rosnących wartości klucza. W "ciele" rekordów Kartoteka-A poszczególne pola informacyjne oznaczono różnym zakreskowaniem. Podobnie uczyniono w odniesieniu do rekordów Zbioru Wybiorczego. Uporządkowanie tej kartoteki jest również rosnące. Rekord Zbioru Dobranego jest zupełnie różny od rekordów^w obu kartotekach. Zawiera on pola informacyjne zarówno rekordu Kartoteki-A, jak i rekordu Zbioru Wybiorczego a ponadto wygenerowane zostało w nim zupełnie nowe pole /oznaczone /. To nowe pole informacyjne może zawierać np. wynik jakiejś procedury obliczeniowej, stanowiącej integralną część przebiegu dobierania.

Jak wynika z powyższego przykładu, w przebiegu dobierania rekordy zbioru wyjściowego mogą być jakościowo różne od rekordów kartotek wejściowych. Zwróćmy również uwagę, iż klucze rekordów w kartotekach nie muszą być wcale pierwszymi polami rekordów.

Omówimy jeszcze krótko ten sam problem na przykładzie na danych bardziej szczegółowych.

zbiór wybiórczy

kartoteka - A

0-1	
0325	+++++
0332	+++++
0333	+++++
0339	+++++
0341	+++++
0345	+++++
0347	+++++
0350	+++++
0362	+++++
00	

0-2	
0324	
0332	
0334	
0339	
0340	
0345	
0347	
0360	
0362	
00	

Zbiór
dobrany

0-3	
0332	+++++
0339	+++++
0345	+++++
0347	+++++
0362	+++++
00	

Bys. nr 10. Przykład dobierania prostego z procedurą obliczeniową.

Jednym ze zbiorów wejściowych przebiegu jest Kartoteka Pracowników stworzona i wykorzystywana dla potrzeb Systemu Gospodarki Osobowej w jakiejś instytucji czy przedsiębiorstwie. Pola informacyjne poszczególnych rekordów kartoteki zawierają te same elementy informacyjne jak to przedstawiono w pkt. 1.4. Rysunek 11. przedstawia fragment tej kartoteki.

Drugim zbiorem wejściowym do przebiegu dobierania jest /bardzo uproszczona/ Kartoteka Przepracowanych Roboczogodzin. Zawiera ona następujące elementy informacyjne:

- numer pracownika
- ilość godzin przepracowanych przez pracownika
- ilość godzin urlopowych,
- miesiąc i rok, którego zarobek dotyczy

W przebiegu dobierania dodatkowo wprowadzono procedurę obliczeniową, której zadaniem jest obliczenie wysokości zarobku wynikającego z faktycznie przepracowanych godzin na podstawie elementów informacyjnych zawartych w obu kartotekach /ilość godzin w Kartotece Przepracowanych Roboczogodzin oraz stawek zaszeregowania w Kartotece Pracowników/.

Kartoteka przepracowanych
roboczogodzin

NR. PRAC	JLÓŚĆ GODZ	URLOP	M	R
325	192	0	07	72
NR PRAC	JLÓŚĆ GODZ	URLOP	M	R
326	492	0	07	72
NR PRAC	JLÓŚĆ GODZ	URLOP	M	R
327	192	0	07	72
NR PRAC	JLÓŚĆ GODZ	URLOP	M	R
340	168	24	07	72

Kartoteka pracowników

NR PRAC	NAZWISKO			IMIE		
325	IKSIŃSKI			ADAM		
DATA UR	STANOWISKO			W	P	STAN
02 03 36	ŚLUSARZ			5	2	5
STAWKA	MIEJSCE ZAMIESZKANIA					
11 20	KATOWICE					
ULICA						NR
KRAKOWSKA						17
PRZYN	JEZYKI			DEP		
1	1	0	0	0	0	1
NR PRAC	NAZWISKO			IMIE		
327	YGREKOWSKA			EWA		
DATA UR	STANOWISKO			W	P	STAN
12 08 42	TOKARZ			3	1	3
STAWKA	MIEJSCE ZAMIESZKANIA					
9 80	KATOWICE - 17					
ULICA						NR
3-60 MAJA						12
PRZYN	JEZYKI			DEP		
2	0	0	1	0	0	2
NR PRAC	NAZWISKO			IMIE		
340	ZETOWICZ			HENRYK		
DATA UR	STANOWISKO			W	P	STAN
12 12 23	FREZER			3	2	5
STAWKA	MIEJSCE ZAMIESZKANIA					
8 40	SIEMIANOWICE ŚL					
ULICA						NR
BRATKÓW						122
PRZYN	JEZYKI			DEP		
2	0	0	0	0	0	1

kartoteka wynagrodzeń

NR PRAC	ŁĄCZNIE WYNAGR		SKŁADNIKI			M	R
325	2.400	40	25.000	2150,40	0,00	07	72
NR PRAC	ŁĄCZNIE WYNAGR		SKŁADNIKI			M	R
327	2.001	60	120,00	1881,60	0,00	07	72
NR PRAC	ŁĄCZNIE WYNAGR		SKŁADNIKI			M	R
340	1.862	80	250,00	1411,20	201,60	07	72

Rys.nr 11. Przykład przebiegu dobierania prostego w którym rekord wynikowy jest jakościowo inny od rekordów kartotek wejściowych.

Ponadto w rekordzie zbioru wynikowego winny znaleźć się elementy informacyjne o poszczególnych składnikach wynagrodzenia. Na łączne wynagrodzenie składają się takie elementy jak: zarobek wynikający /osobne pole informacyjne/ z przepracowanych godzin, /czyli procedury obliczeniowej/, deputat węglowy /informacja na temat otrzymywania deputatu węglowego w całości lub części znajduje się Kartotece Pracowników w polu informacyjnym DEP, przy czym cyfra 1 oznacza pełny deputat 250 zł., zaś cyfra 2 oznacza deputat niepełny/a także wynagrodzenie za dni urlopowe /ilość dni urlopowych podano w godzinach, w rekordzie Kartoteki Przepracowanych Roboczo-godzin/.

Rekord w zbiorze dobranym /tzw. Kartoteka Wynagrodzeń/ zawiera zatem następujące pola informacyjne:

- numer pracownika
- łączne wynagrodzenie pracownika
- składniki wynagrodzenia takie jak:

deputat węglowy

zarobek wynikający z przepracowanych godzin

zarobek za dni urlopowe

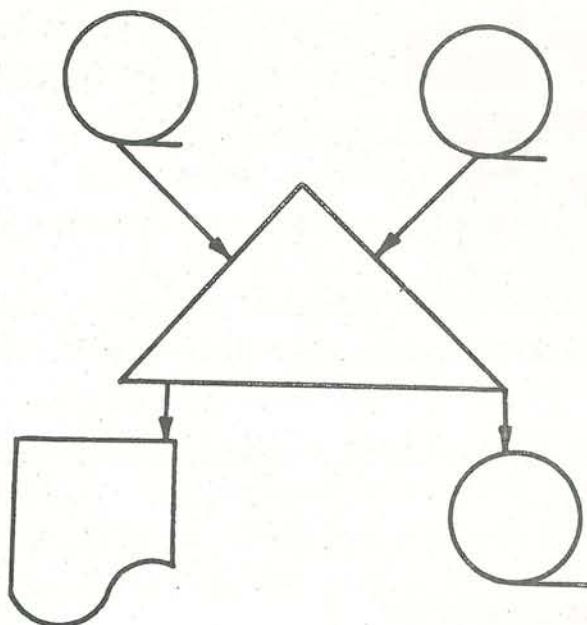
- miesiąc i rok obrachunkowy

Poszczególne pola informacyjne w rekordzie Kartoteki Wynagrodzeń niezbędne są dla emisji listy płac w dalszych przebiegach systemu.

Kluczem uporządkowania kartotek jest rosnąca wartość pola informacyjnego "numer pracownika". Poza spostrzeżeniami o których już mówiliśmy wyżej omawiając przykład poprzedni, warto zwrócić uwagę, iż w ostatnim przykładzie w Kartotece Przepracowanych Roboczo-godzin występuje rekord o numerze pracownika 326. Nie występuje natomiast rekord o takiej wartości w Kartotece Pracowników. Nasuwa się wniosek, iż albo wystąpił błąd w Kartotece Przepracowanych Roboczo-godzin /np. miał to być nr. pracownika 362 - a tzw. "czeski błąd" popełniła dziurkarka podczas perforowania maszynowego nośnika i nie został on wykryty przez program wczytujący/ albo też Kartoteka

Pracowników nie zawiera aktualnego rejestru wszystkich pracowników. W obu tych przypadkach powinna nastąpić sygnalizacja tego błędu. Dlatego też w niektórych przypadkach w przebiegu dobierania występują również zbiór drukarki wierszowej, na której sygnalizowane są błędy.

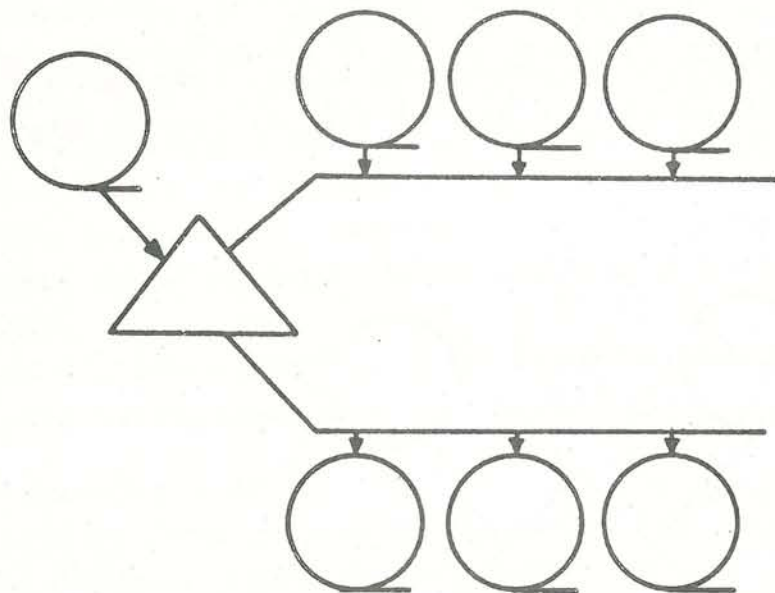
Symboliczne oznaczenia przebiegu dobierania prostego z równoczesnym wydrukiem tzw. raportu dobierania przedstawia się następująco:



Rys.nr 12. Przebieg dobierania z równoczesnym wydrukiem raportu.

Raport z przebiegu dobierania może zawierać również inne elementy informacyjne, np. ilość rekordów dobranych, ilość rekordów niedobraniych w Kartotekach Wejściowych itp. W przypadku gdy w przebiegu dobierania występuje większa ilość kartotek wejściowych i wyjściowych mamy do czynienia z przebiegiem dobierania wielofazowego. Symboliczne

oznaczenie przebiegu dobierania wielofazowego przedstawiono niżej.



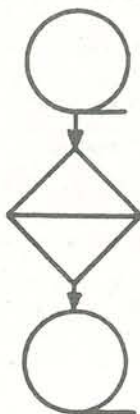
Rys.nr 13. Przebieg dobierania wielofazowego.

Przebieg dobierania nie wchodzi w zestaw standardowego oprogramowania komputera.

2.3. Przebieg sortowania /sorting/

Przebieg sortowania służy do uporządkowania zbioru według wzrastających lub malejących wartości kluczy.

Symboliczne oznaczenie przebiegu sortowania jest następujące:

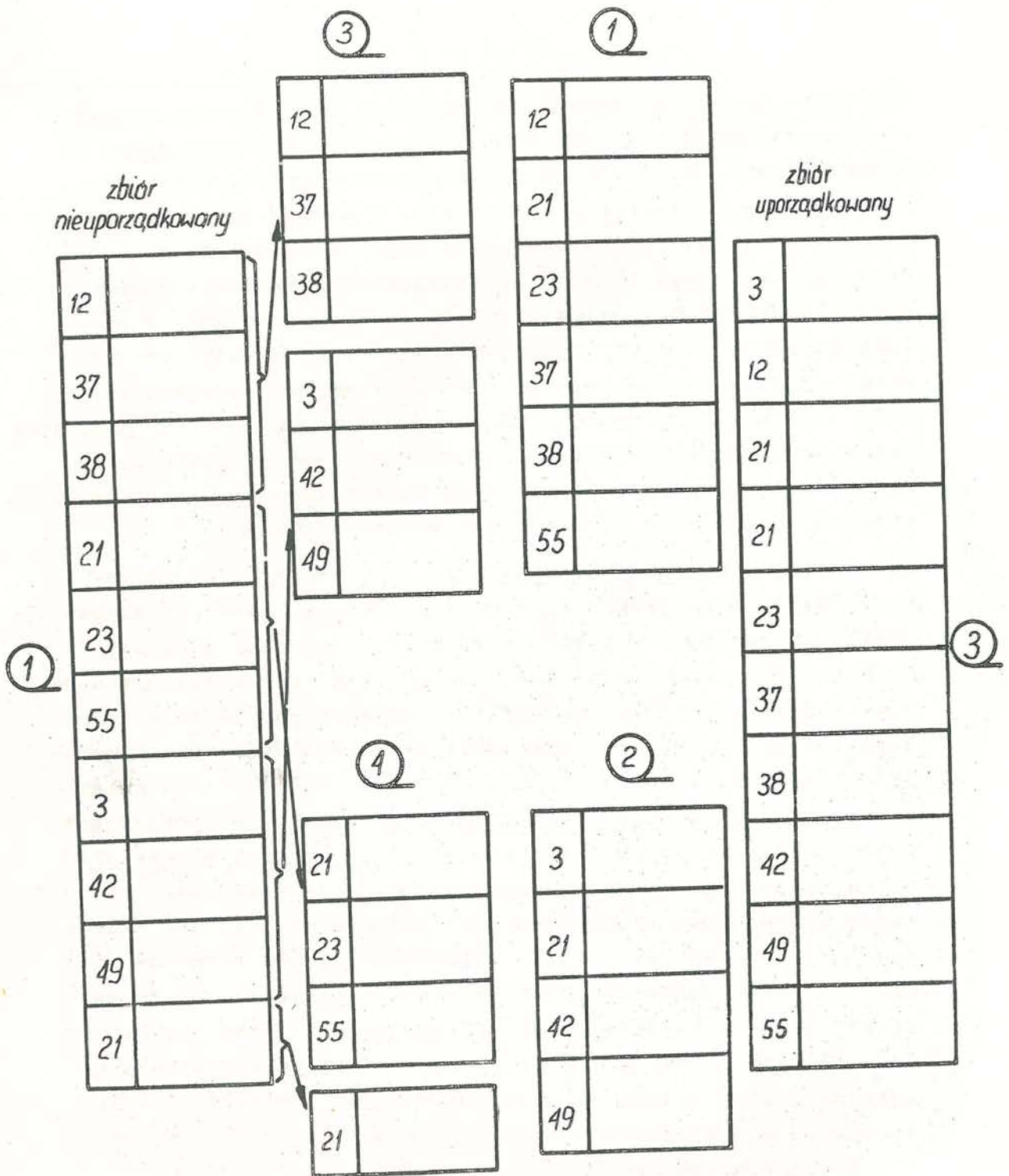


Rys.nr 14. Przebieg sortowania.

Efektywne algorytmy sortowania wymagają /w przypadku przetwarzania sekwencyjnego/ conajmniej czterech jednostek taśmy magnetycznej. Na rysunku 15 przedstawiono przykładowy tok postępowania w przebiegu, sortowania, w którym wykorzystuje się cztery taśmy magnetyczne. Podobnie jak to miało miejsce przy omawianiu poprzednich przebiegów w przykładzie podano skrajnie uproszczone elementy informacyjne rekordów. Z pół informacyjnych rekordu wyodrębniono jedynie pole klucza rekordu. Przyjęto rosnącą kolejność uporządkowywania zbioru. Zbiór nie uporządkowany/ znajdujący się fizycznie na jednostce taśmy nr 1/w pierwszym kroku przebiegu jest rozpisywany na dwa podzbiory /fizycznie na taśmę magnetyczną nr 3 i 4/ częściowo uporządkowane wg rosnących wartości klucza.

Do pamięci komputera wczytany zostaje pierwszy i drugi rekord zbioru nieuporządkowanego. Z porównania wartości kluczy wynika iż rekordy występują w rosnącym porządku. Pierwszy z nich zostaje zapisany na taśmę nr 3, drugi natomiast przechowywany jest w pamięci operacyjnej. Przeczytany zostaje następny rekord ze zbioru nieuporządkowanego /o wartości klucza 38/. Z porównania wartości klucza wynika, iż w dalszym ciągu zachowana jest rosnąca kolejność wartości. Rekord dotychczas przechowywany w pamięci operacyjnej / o wartości klucza 37/ zapisany zostaje na taśmę magnetyczną /w dalszym ciągu na taśmę nr 3/. Przeczytany jest kolejny rekord. Wartość klucza czwartego rekordu wynosi 21. Przerwana została rosnąca kolejność uporządkowana. Rekord o wartości klucza 38 zostaje zapisany jeszcze na taśmie nr 3. Fakt przerwania rosnącej kolejności zostaje zapamiętany, a do pamięci zostaje wczytany kolejny rekord zbioru nieuporządkowanego. Pamiętać należy, że rekord o wartości klucza 21 przechowywany jest również w pamięci operacyjnej.

Następuje porównanie wartości kluczy. Wynika z niego że od rekordu o wartości klucza 21 utrzymuje się rosnące uporządkowanie rekordów. Rekord o wartości klucza 21 zostaje zapisany na taśmę nr 4. Dalsze kroki postępowania są identyczne. Rekordy o wartościach kluczy 23 i 55 /tzn. aż do chwili kolejnego załamania się rosnącej sekwencji wartości kluczy rekordów/ zostają również zapisane na taśmie nr 4.

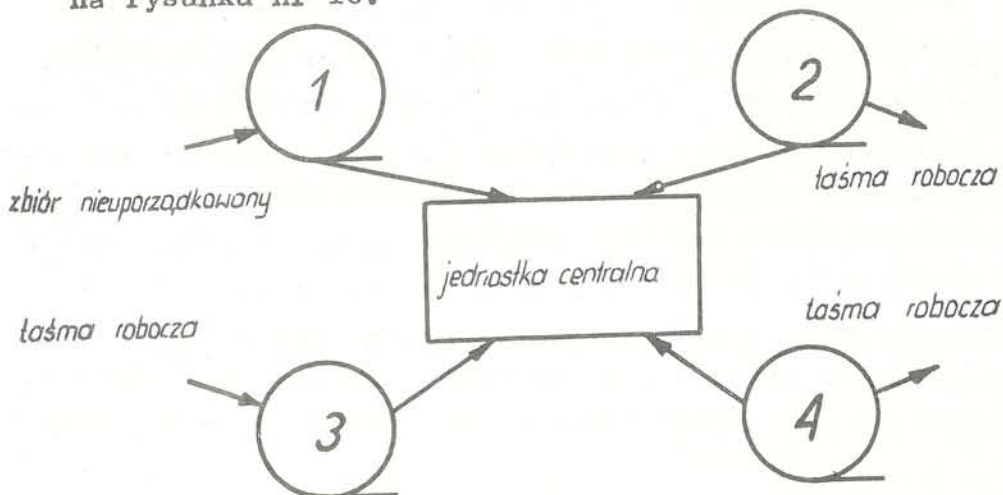


Rys. nr 15. Przykład przebiegu sortowania.

W momencie przeczytania rekordu o wartości klucza równym 3, zapis rekordów na taśmie nr 4 zostaje zaprzestany. Kolejna sekwencja rosnąca uporządkowanych rekordów zbioru nieuporządkowanego zapisana zostanie na taśmie nr 3 itd. Zapis kolejnych sekwencji rosnących wg kolejności zapisywania na taśmach 3 lub 4 przedstawiono na rysunku nr 15. odpowiednim zakreśleniem i oznakowaniem miejsc zapisu strzałkami.

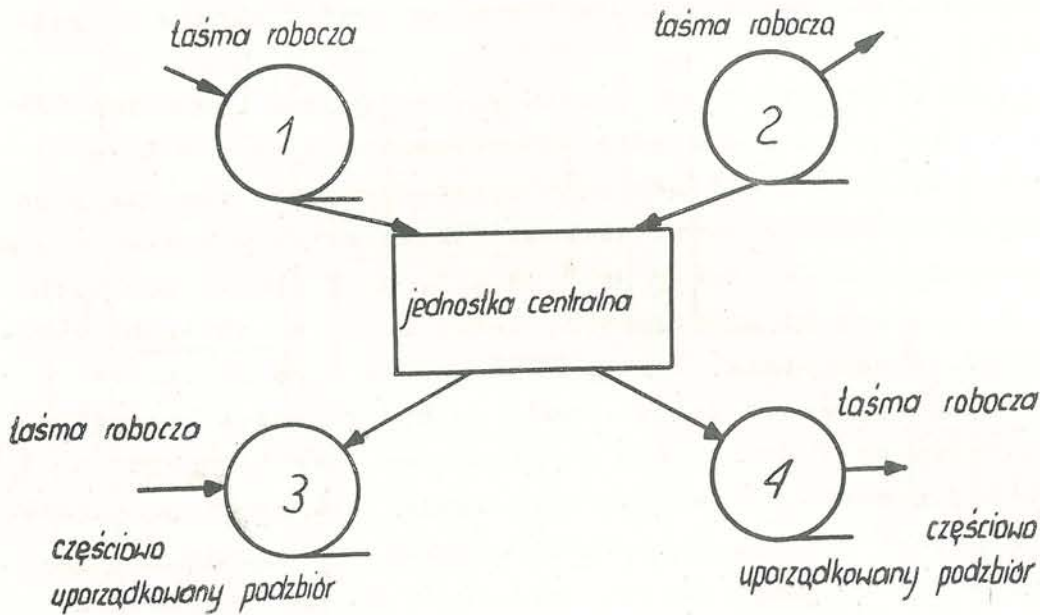
Następne kroki przebiegu sortowania są niczym innym jak podprzebiegami scalania częściowo uporządkowanych podzbiorów w coraz dłuższe ciągi rekordów uporządkowanych zgodnie z rosnącą sekwencją wartości klucza. Wykorzystywane w tym celu są na przemian parami taśmy nr 3, 4 i 1, 2. W chwili gdy taśmy 3 i 4 są zbiorami wejściowymi, taśmy 1 i 2 są zbiorami wyjściowymi. I odwrotnie. Postępowanie takie trwa do chwili uzyskania na jednej taśmie /może to być w naszym przypadku albo taśma nr 3 albo taśma nr 4/ zbioru posortowanego.

Powyższy przykład przebiegu sortowania, wykorzystujący cztery jednostki taśm magnetycznych, oparto o algorytm von Neumana. Warto w tym miejscu wyjaśnić, że na instalacjach komputerowych wyposażonych jedynie w cztery fizyczne jednostki taśm magnetycznych tok postępowania operatorskiego jest następujący. Zbiór nieuporządkowany zakładany jest na fizycznie pierwszą jednostkę taśmy, tak jak to przedstawiono na rysunku nr 16.



Rys. nr 16. Założenie zbioru nieuporządkowanego.

Po pierwszym kroku przebiegu, /tzn. po rozpisaniu zbioru nieuporządkowanego na częściowo uporządkowane zbiory na taśmach 3 i 4/ następuje wymiana szpul taśmy magnetycznej na jednostce nr 1. Zbiór nieuporządkowany zostaje zastąpiony tzw. taśmą roboczą, tak jak to przedstawiono na rysunku nr 17.

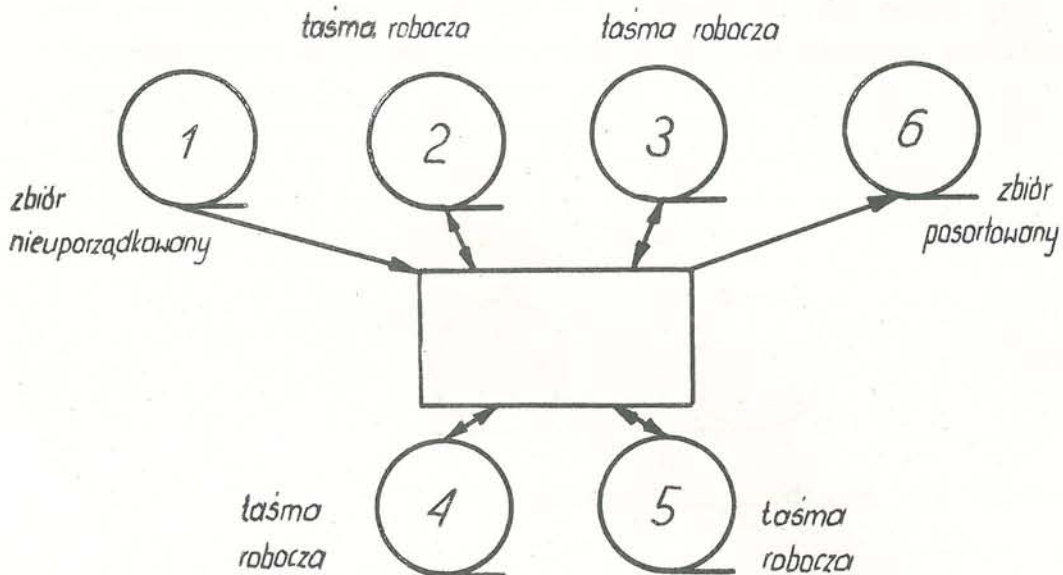


Rys.nr 17. Wymiana szpuli taśmy na fizycznej jednostce nr 1.

Technika postępowania w tym przypadku jest taka, iż zbiór nieuporządkowany jest w dalszym ciągu dostępny. Należy bowiem pamiętać że zbiór nieuporządkowany z punktu widzenia jednego klucza uporządkowania, może być równocześnie zbiorem uporządkowanym z punktu widzenia innego klucza, i jako taki potrzebny w dalszym procesie przetwarzania.

W przypadku gdy instalacja komputera wyposażona jest w fizycznie większą ilość jednostek taśm magnetycznych technika postępowania jest nieco inna. Przedstawiono to schematycz-

nie na rysunku nr 18.



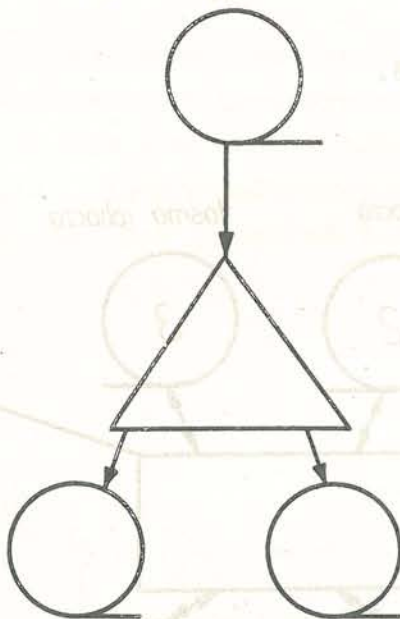
Rys.nr 18. Wykorzystanie fizycznych jednostek taśm magnetycznych w przebiegu sortowania.

W uzupełnieniu przebiegu sortowania należy dodać że współczesne komputery wyposażone są w standardowe programy sortowania. Dla standardowego programu wystarczy jedynie podać takie informacje jak: wielkość bloku, miejsce położenia i wielkość klucza w rekordach oraz wymagany porządek sortowania /rosnący lub malejący/.

2.4. Przebieg rozdziału /separating/

Przebieg rozdziału służy do podziału jednego zbioru głównego na dwa lub więcej podzbiorów.

Schematycznie przebieg rozdziału przedstawiamy następująco:

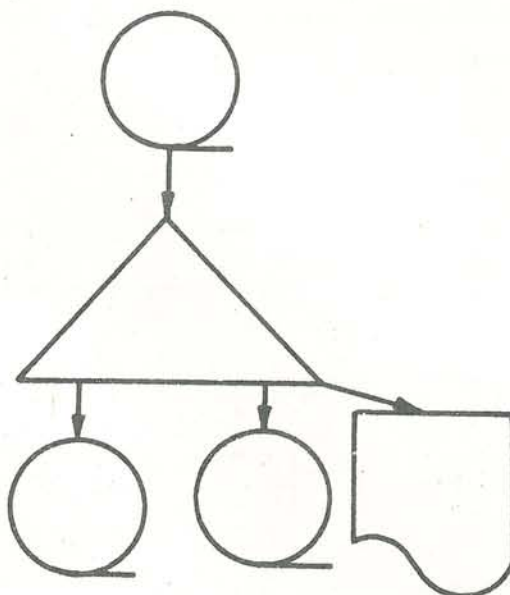


Rys.nr 19. Przebieg rozdziału.

W praktyce projektowej niejednokrotnie napotyka się na konieczność podzielenia jakiegoś zbioru - nazwijmy go umownie "głównym" na dwa lub więcej podzbiorów. Kryteria podziału zbioru głównego mogą być bardzo różne. Przykładowo może to być rozdział Kartoteki Pracowników, /która dotyczy ogółu pracowników przedsiębiorstwa/ na dwa podzbiory kartotekowe zawierające: Kartotekę Pracowników Fizycznych i Kartotekę Pracowników Umysłowych. Rozdział zbioru dokonywany jest najczęściej z tytułu szybszej /czasowo/ współpracy z wygenerowanymi podzbiórami /w przypadku przetwarzania sekwencyjnego/ w trakcie obliczeń odcinkowych zadań danego systemu przetwarzania danych. W skrajnie prostym przebiegu rozdziału każdy rekord zbiorów wyjściowych jest identyczny z jednym z rekordów zbioru wejściowego /nierozdzielonego/. Ma tu zatem miejsce jedynie przenoszenie rekordów według zadanych kryteriów do podzbiorów wyjściowych. Reprezentantem tego typu przebiegu jest pierwszy krok /pierwsza faza/ przebiegu sortowania /patrz pkt. 2.3./.

Podobnie jak to miało miejsce w przebiegu dobierania /pkt. 2.2./, niektóre typy /wersje/ przebiegu rozdziału mogą w zbiorach wyjściowych generować rekordy jakościowo zupełnie różne od rekordów zbioru wejściowego.

W trakcie przebiegu rozdziału może być również generowany raport przebiegu, zawierający szczególnie informacje na temat zawartości ilości rekordów w zbiorze wejściowym, ilości rekordów w poszczególnych zbiorach wyjściowych itp. Przebieg ten oznaczamy na schematach tak jak to przedstawiono na rys. 20.



Rys. nr 20. Przebieg rozdziału z równoczesnym wydrukiem raportu.

2.5. Przebieg konwersji /converting/

Przebieg konwersji umożliwia zmianę nośnika informacji zbioru. Zadaniem przebiegu jest zatem przeniesienie zawartości informacyjnej zbioru z jednego nośnika informacji na drugi. Przebieg konwersji może mieć różny charakter. Ogólnie różniamy trzy typy przebiegów konwersji:

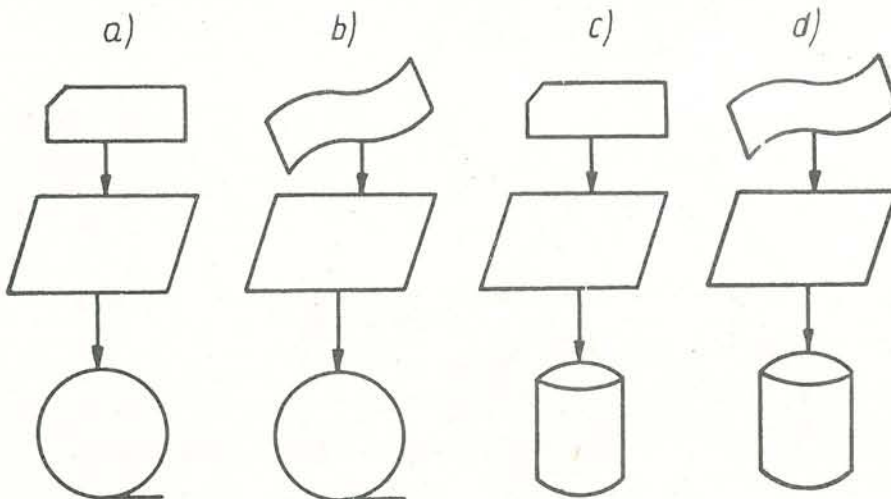
- przebieg konwersji wejścia,
- przebieg konwersji wyjścia-
- przebieg konwersji wewnętrznej.

Do pierwszej grupy zaliczymy przebiegi związane z wprowadzeniem zbiorów kartowych lub zbiorów taśmy perforowanej do pamięci zewnętrznych komputera. Przedstawicielami tej grupy

przebiegów konwersji są:

przebieg konwersji karta - taśma magnetyczna,
przebieg konwersji taśma perforowana - taśma magnetyczna,
przebieg konwersji karta perforowana - dysk magnetyczny,
przebieg konwersji taśma perforowana - dysk magnetyczny,

Rysunek 21 przedstawia oznaczenia tych przebiegów na schematach przetwarzania.

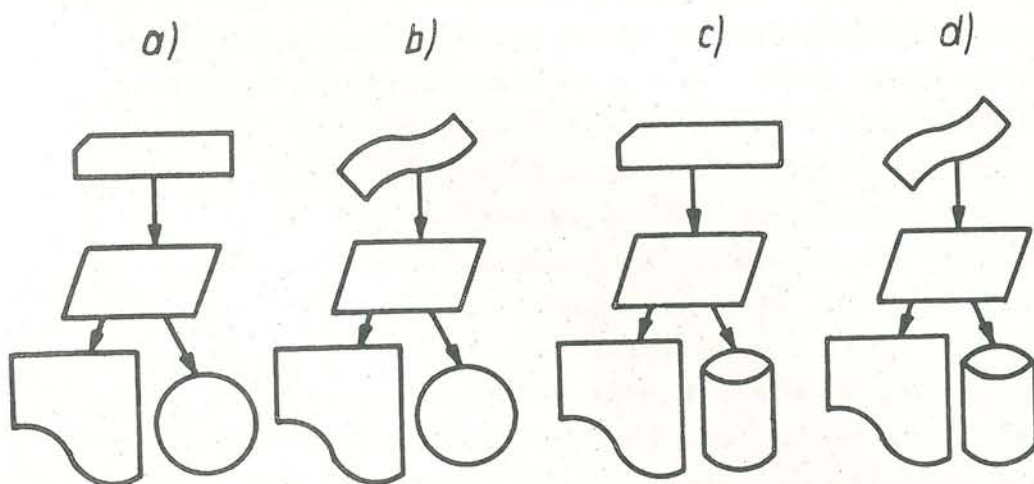


Rys. nr 31. Przebiegi konwersji wejścia: a/ konwersja karta-taśma magnetyczna, b/ konwersja taśma-perforowana-taśma magnetyczna, c/ konwersja karta-dysk magnetyczny, d/ konwersja taśma perforowana-dysk magnetyczny.

Wprowadzenie informacji źródłowych do systemu najczęściej jednak odbywa się przy równoczesnym kontrolowaniu poprawności danych wprowadzanych.

W przebiegach konwersji o których mówiliśmy przed chwilą, mówiąc założyliśmy, że przebieg konwersji wejścia nie zmienia ani postaci ani rozmiaru informacji przenoszanej z jednego maszynowego nośnika na drugi. W praktyce przebieg konwersji wejścia zawiera procedury kontrolne lub procedury obliczeniowe, których celem jest wyselekcjonowanie błędnych informacji wejściowych. Wyselekcjonowane formalne lub merytoryczne błędne rekordy wejściowe generowane są najczęściej w tzw. raporcie błędów. Raport błędów drukowany jest na drukarce wierszowej.

Na jego podstawie można ze zbioru np. kart perforowanych wybrać karty błędne i poprawić je. W przypadku gdy przebieg konwersji wejścia zawiera procedury kontrolne i drukowany jest raport błędów, oznaczenia na schematach przetwarzania przedstawiają się tak jak to pokazano na rysunku 22 / w kolejności identycznej jak na rysunku poprzednim/.



Rys.nr 22. Przebiegi konwersji wejścia z równoczesną kontrolą poprawności informacji wprowadzanych.

Rodzaje funkcji kontrolnych w przebiegach konwersji wejścia są bardzo zróżnicowane. Do najczęściej stosowanych należą:

- badanie rekordów wejściowych czy należą one do zbioru /każdy rekord zawiera wtedy identyfikator przynależności rekordu do danego zbioru/,
- sprawdzenie czy zawartości poszczególnych pól informacyjnych są zgodne z zadeklarowanym charakterem tych pól /np. czy w polach informacyjnych numerycznych nie występują wartości alfabetyczne lub alfanumeryczne/,
- badanie zawartości niektórych pól informacyjnych /np. sprawdzenie, czy pole informacyjne "stan cywilny" jest

poprawny tzn. czy symbol stanu cywilnego mieści się w zakresie przyjętych w systemie symboli dla stanów cywilnych/,

- badanie relacji pomiędzy poszczególnymi polami informacyjnymi /np. gdy pole informacyjne "jednostka miary" zawiera symbol określający sztuki - to pole informacyjne "ilość" nie może zawierać wielkości ułamkowej/.

Jeśli procedura kontrolna znajdzie przynajmniej jeden błąd, rekord zawierający błąd - nie jest przenoszony na magnetyczny nośnik informacji. Rodzaj błędu oraz zawartość informacyjna rekordu błędnego zostaje wydrukowana w raporcie.

Wczytany zostaje następny rekord wejściowy.

Dla niektórych pól informacyjnych również dla celów kontrolnych stosuje się tzw. sumy kontrolne. Wartości wybranych pól informacyjnych - dla z góry określonej ilości rekordów zbioru - są sumowane /manualnie lub przy wykorzystaniu innych technik/. Powstają tą drogą tzw. sumy kontrolne. Sumy kontrolne wprowadzane są także do komputera przebiegiem konwersji. Procedura obliczeniowa przebiegu wylicza również sumy kontrolne i sprawdza je. W przypadku zgodności - przyjmuje się, iż z góry określona ilość rekordów /tzw. paczka/ nie zawiera w wybranych polach informacyjnych błędu. W przypadku niezgodności sum kontrolnych - cała paczka traktowana jest jako paczka błędna.

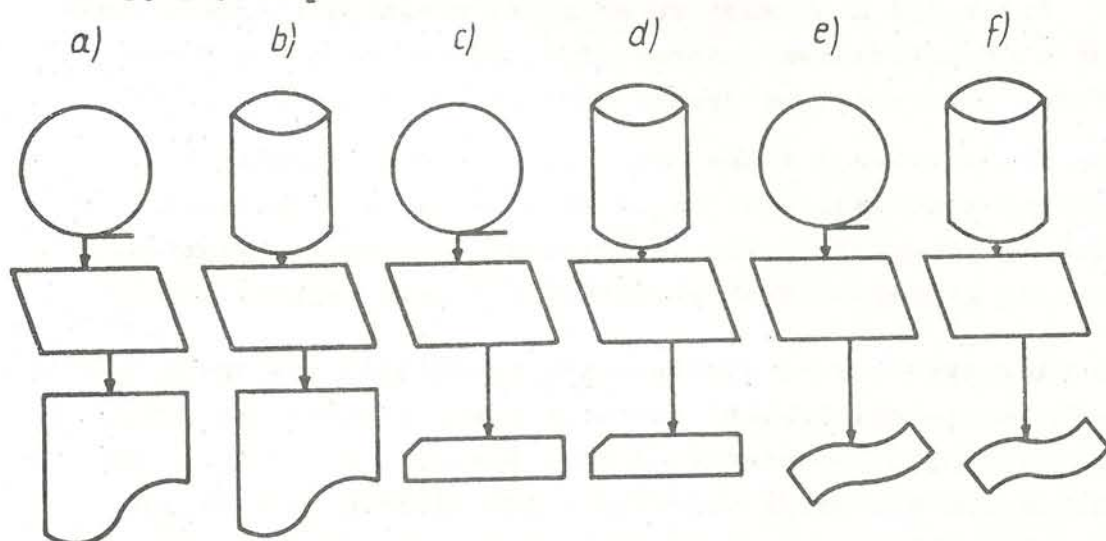
Należy dodać, że w bardziej skomplikowanych przebiegach konwersji wejścia jeden rekord zbioru wejściowego nie koniecznie generuje jeden tylko rekord na magnetycznym nośniku informacji. W praktyce niejednokrotnie jeden rekord wejściowy generuje dwa lub więcej rekordów na taśmie lub dysku magnetycznym.

Przechodząc do omówienia drugiej grupy przebiegów konwersji, a mianowicie przebiegów konwersji wyjścia należy stwierdzić, iż do grupy tej zaliczyć należy wszystkie te programy, których celem jest przekazywanie wyników pracy systemu. Prze-

biegi konwersji wyjścia wyprowadzają z systemu odpowiednio wyselekcjonowane i zagregowane informacje **bezpośrednie** lub **pośrednie**. Do grupy tej zaliczamy:

przebieg konwersji taśma magnetyczna - drukarka wierszowa,
przebieg konwersji dysk magnetyczny - drukarka wierszowa,
przebieg konwersji taśma magnetyczna - dziurkarka kart,
przebieg konwersji dysk magnetyczny - dziurkarka kart
przebieg konwersji taśma magnetyczna - perforator taśmy
przebieg konwersji dysk magnetyczny - perforator taśmy.

Rysunek nr 23 przedstawia oznaczenia tych przebiegów na schemacie przetwarzania.



Rys.nr 23. Przebieg konwersji wyjścia: a/ konwersja taśma magnetyczna-drukarka wierszowa, b/ konwersja dysk magnetyczny-drukarka wierszowa, c/ konwersja taśma magnetyczna - perforator kart, d/ konwersja dysk magnetyczny - perforator kart, e/ konwersja taśma magnetyczna - perforator taśmy, f/ konwersja dysk magnetyczny-perforator taśmy.

Zadaniem dwu pierwszych z wyżej wymienionych przebiegów konwersji jest wyprowadzenie przede wszystkim wyników bezpośrednich tzn. wyników końcowych, użytkowych. Są to wszelkiego rodzaju wydawnictwa wykorzystywane bezpośrednio przez użytkownika systemu.

Pozostałe przebiegi realizują nieco inne zadania. Najczęściej służą one do wyprowadzania wyników pośrednich obliczeń, stanowiąc równocześnie maszynowy nośnik informacji dla następnych przebiegów wejścia, na tym samym lub innym komputerze.

W praktyce dość często spotykamy się z połączeniem dwu lub więcej typów przebiegów konwersji wyjścia. Jako przykład można przedstawić przebieg, w którym emitowane jest wydawnictwo, syntetyczne a równocześnie perforowana jest taśma dziurkowana, której zawartość informacyjną stanowią dane analityczne emitowanego wydawnictwa.

Przebiegi konwersji wewnętrznej obejmują programy związane z przenoszeniem informacji z jednego magnetycznego nośnika informacji na drugi. Będą to przebiegi:

przebieg konwersji taśma magnetyczna - dysk magnetyczny,
przebieg konwersji dysk magnetyczny - taśma magnetyczna,
przebieg konwersji taśma magnetyczna - taśma magnetyczna,
przebieg konwersji dysk magnetyczny - dysk magnetyczny,

Dwa pierwsze z wyżej wymienionych przebiegów występują dość rzadko w projektowanych systemach przetwarzania. Najczęściej bowiem - /należy pamiętać iż mogą one wystąpić dopiero po przebiegach konwersji wejścia/ w ich miejsce wchodzi już przebiegi realizujące właściwą pracę systemu przetwarzania tj. przebiegi sortowania, dobierania itd.

Do celów zabezpieczenia zbiorów systemu stosunkowo często stosowane są natomiast pozostałe dwa rodzaje przebiegów konwersji wewnętrznej. Noszą one wtedy nazwę programów kopiujących.

2.6. Przebieg wybierania

Przebieg wybierania służy do wyselekcjonowania z jednego lub więcej zbiorów - jednego lub więcej podzbiorów wg zadanego kryterium.

W prostym przebiegu wybierania bierze udział jeden zbiór wejściowy oraz jeden zbiór wyjściowy.

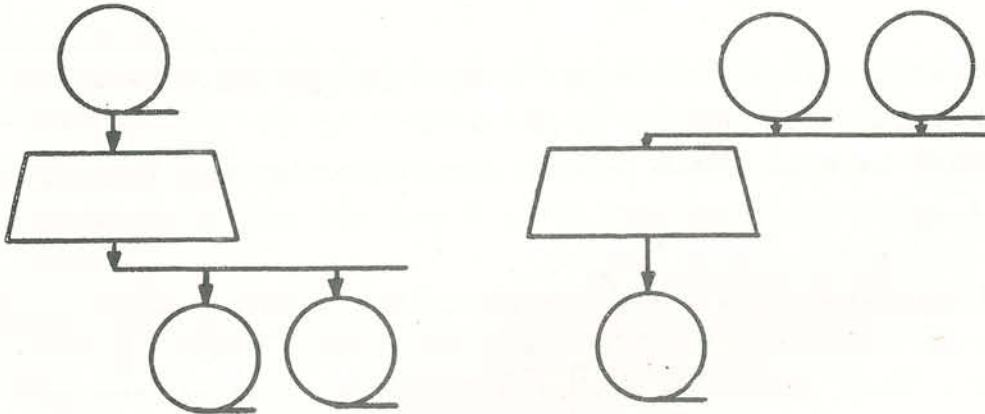
Oznaczenie na schematach przetwarzania przedstawia się następująco:



Rys.nr 24. Przebieg wybierania prostego.

Przykładem przebiegu wybierania może być program, który wybiera z Kartoteki Zamówień /zawierającej ogół zamówień złożonych przez wszystkich klientów/ zamówienia jednego /lub kilku/ klienta. Innym przykładem przebiegu wybierania może być program, który z Kartoteki Zamówień /zawierającej całoroczny portfel zamówień/ wybiera zamówienia jednego kwartału. Jeśli zbiorów wejściowych lub wyjściowych jest więcej niż jeden mamy do czynienia ze złożonym przebiegiem wybierania. Przykładem tego typu przebiegu może być program, który z Kartoteki Pracowników wybiera na jeden podzbiór wyjściowy - wszystkich mężczyzn poniżej 25 lat, zaś na drugi podzbiór - wszystkich pracowników danego zawodu /np. ślusarzy, tokarzy itp./.

Oznaczenie na schematach przetwarzania złożonego przebiegu wybierania przedstawiono na rys. 25.



Rys.nr 25. Złożone przebiegi wybierania.

Przebiegi wybierania są przebiegami występującymi bardzo często w systemach elektronicznego przetwarzania danych.

2.7. Przebiegi aktualizacji prostej /updating/

W przebiegu aktualizacji prostej biorą udział dwa zbiory wejściowe zgodnie uporządkowane /czyli kartoteki/ i jeden zbiór wyjściowy. Jednym ze zbiorów wejściowych jest tzw. stara generacja kartoteki /old master file/. Drugim zbiorem wejściowym jest tzw. zbiór transakcji /transactions/, który zawiera informacje zezwalające na dokonanie koniecznych zmian w starej generacji kartoteki. W wyniku przebiegu powstaje nowy zbiór tzw. nowa generacja kartoteki /new master file/. Jak to już powiedziano w pkt. 1.5. pojęcie zbioru /kartoteki/ jest pojęciem dynamicznym. W kolejnych okresach pracy systemu przetwarzania danych, odpowiednią kartotekę zastępować musi następna, która w danym momencie zawierać będzie najaktualniejsze informacje niezbędne w systemie. Posłużmy się przykładem. Kartoteka Pracowników w Systemie Gospodarki Osobowej musi być stale uaktualniana. Fluktuacja kadry wymaga aby z Kartoteki Pracowników anulować wszystkich tych pracowników, którzy odeszli z przedsiębiorstwa. Na ich miejsce przyjęci zostali nowi pracownicy. O dane personalne tych pracowników musi być wzbogacona Kartoteka Pracowników, aby spełniała dostatecznie swoje zadanie w trakcie obliczania np. listy płac. Zmiana miejsc zamieszkania pracowników, sta-

nowisk, grup zaszeregowania, zmiana nazwiska /u kobiet/ itp. powoduje również konieczność uaktualnienia kartoteki.

Jak widać z powyższego w trakcie przebiegu aktualizacji występują trzy typy czynności, które sterują dokonywaniem zmian w kartotece.

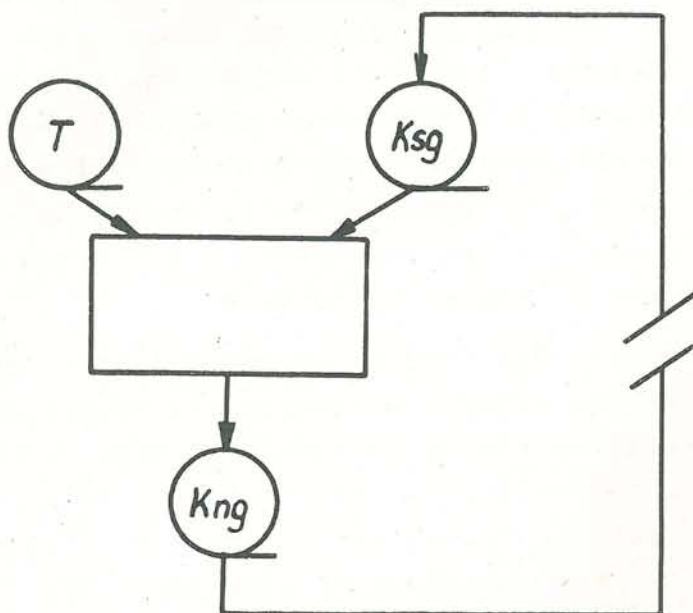
Typami czynności tymi są:

- a/ Anulacja rekordu /zmazanie/ - ze zbioru starej generacji ma zostać usunięty rekord o tej samej wartości klucza co wartość klucza w rekordzie transakcji. W praktyce najczęściej temu typowi czynności przypisana jest wartość = 1. /wracając do podanego wyżej przykładu: typ czynności = 1 będzie zawierał rekord zbioru transakcji dla pracowników, którzy odeszli z przedsiębiorstwa i których dane mają być z Kartoteki Pracowników usunięte/.
- b/ Zmiana zawartości niektórych pól informacyjnych rekordu - wartość pól rekordu w zbiorze starej generacji, o tej samej wartości klucza co wartość klucza w rekordzie transakcji ma być zmieniona zgodnie z wartością rekordu transakcji. W praktyce temu typowi czynności przypisana jest wartość = 2. /W odniesieniu do przykładu: typ czynności = 2 zastosujemy w rekordach transakcji dla tych rekordów Kartoteki Pracowników, w których ma nastąpić: zmiana nazwiska, zmianę stanowiska, zmiana adresu zamieszkania itd./.
- c/ Wprowadzenie nowego rekordu - w zbiorze kartotekowym, ma zostać wygenerowany rekord o różnej wartości klucza od wartości kluczy rekordów znajdujących się już w Kartotece. Wartość klucza określa rekord transakcji.

Miejsce wprowadzenia określa sekwencja uporządkowania kartoteki. Jeżeli kartotekę uporządkowano według rosnącej wartości klucza, wprowadzenie następuje pomiędzy rekord kartoteki o najbliższej niższej wartości klucza a rekord o najbliższej wyższej wartości klucza od wartości klucza rekordu wprowadzanego. W praktyce ten typ czynności ma wartość = 1. /w przykładzie: typ czynności = 1 otrzymują

rekordy transakcji dla nowoprzyjętych pracowników, których dane personalne mają od tego momentu figurewać w Kartotece Pracowników/.

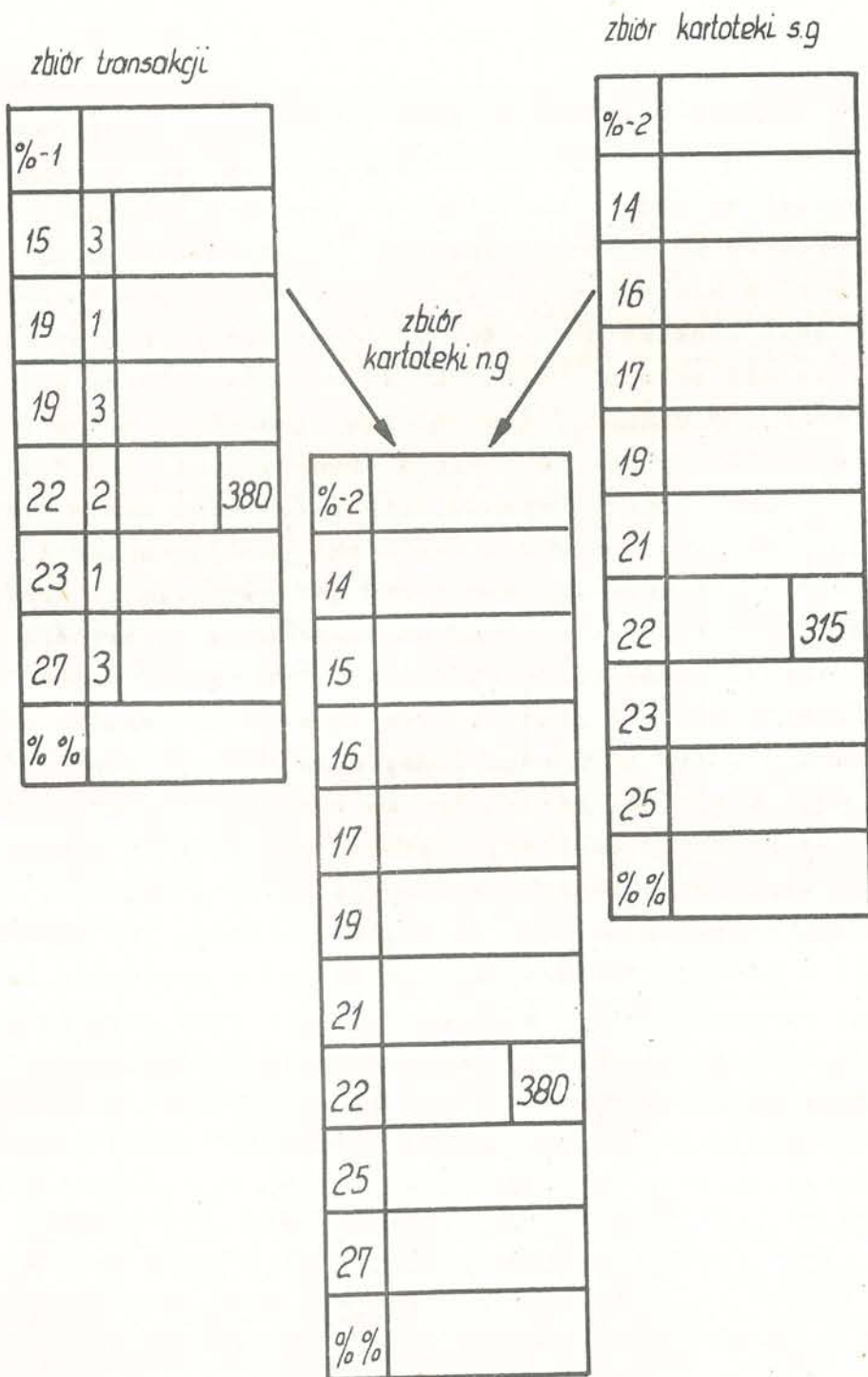
Możnaby zatem powiedzieć, że w czasie przebiegu aktualizacji następuje - sterowana transakcjami - zmiana zawartości zbioru kartotekowego. Oznaczenie przebiegu aktualizacji na schematach przetwarzania przedstawiono na rys.26.



Rys.nr 26. Przebieg aktualizacji prostej: T - zbiór transakcji, Ksg - stara generacja kartoteki, Kng - nowa generacja kartoteki.

Prześledźmy algorytm przebiegu aktualizacji na przykładzie: Rys.27 przedstawia zbiór transakcji, zbiór starej generacji kartoteki, oraz zbiór nowej generacji kartoteki. W rekordach zbioru transakcji - poza polem klucza rekordu - wyodrębniono również pole zawierające typ czynności, a poza tym dla rekordu o wartości klucza 22, wyodrębniono pole informacyjne, którego wartość ma być wprowadzona do pola rekordu /o tej samej wartości klucza/ w kartotece /rekord o wartości klucza 22 w starej generacji kartoteki zawiera w polu tym wartość 315/.

Po kontroli etykiet początku zbiorów wejściowych i zapisaniu etykiety początku w zbiorze wyjściowym przeczytane zostaną: pierwszy rekord ze zbioru kartoteki /o wartości klucza 14/ i pierwszy rekord ze zbioru transakcji /o wartości klucza 15/. Następuje porównanie wartości kluczy. Wartość klucza zbioru transakcji jest większa od wartości klucza zbioru kartoteki. Znaczy to, że dla rekordu kartoteki /o wartości klucza mniejszej/ nie może już ukazać się w zbiorze transakcji rekord, który zasterowałby wykonanie jakiejś czynności na rekordzie kartoteki /zgodne uporządkowanie zbiorów/, a tym samym rekord ten zostaje przepisany na zbiór kartoteki nowej generacji. Rekord zbioru transakcji przechowywany jest w pamięci. Następuje kolejny odczyt ze zbioru kartoteki /rekord o wartości klucza 16/ i porównanie wartości kluczy. Tym razem wartość klucza rekordu kartoteki jest większa od wartości klucza rekordu transakcji. Typ czynności rekordu transakcji /wartość = 1/ informuje o tym, że rekord ten ma zostać wprowadzony do zbioru kartoteki nowej generacji. Rekord /o wartości klucza 15/ zostaje zapisany w zbiorze kartoteki nowej generacji. Rekord zbioru starej generacji /o wartości klucza 16/ przechowywany jest w pamięci operacyjnej, do której w następnym kroku wczytany również zostaje kolejny rekord zbioru transakcji /o wartości klucza 19/. Następuje porównanie wartości kluczy. Wartość klucza rekordu zbioru starej generacji jest mniejszy. Rekord ten zostaje zapisany w zbiorze kartoteki nowej generacji. Rekord transakcji /wartość klucza = 19, typ czynności = 1/ przechowywany jest w pamięci operacyjnej. Wczytany zostaje kolejny rekord zbioru kartoteki /o wartości klucza = 17/. Następuje porównanie wartości kluczy, które przesądza o tym że rekord ten zostaje zapisany w zbiorze kartoteki nowej generacji, z której to w następnym kroku przeczytany jest kolejny rekord /o wartości klucza 19/.



Rys. nr 27. Przykład przebiegu aktualizacji prostej.

Następuje porównanie wartości kluczy. Wartości są równe. Typ czynności zbioru transakcji wskazuje na to, aby rekord zbioru kartoteki zanulować /tj. nie zapisywać go już więcej w zbiorze kartoteki nowej generacji/. Zostaje przeczytany kolejny rekord zbioru transakcji i zbioru kartoteki. Ze zbioru transakcji wczytany został rekord o wartości klucza 19 /ale już z typem czynności = 3/ natomiast ze zbioru kartoteki rekord o wartości klucza = 21. Z porównania wartości kluczy wynika, że rekord zbioru transakcji ma wartość mniejszą, a typ czynności tego rekordu /wartość 3/ każe rekord ten wprowadzić do zbioru kartoteki nowej generacji. Przeczytany zostaje kolejny rekord zbioru transakcji, którego wartość /22/ jest większa od wartości klucza zbioru kartoteki /przechowywanego aktualnie w pamięci operacyjnej/. Rekord z kartoteki zapisany zostaje zatem w zbiorze kartoteki nowej generacji. Z kartoteki starej generacji wczytany zostaje kolejny rekord /o wartości klucza 22/. Klucze są równe w obu rekordach przechowywanych aktualnie w pamięci operacyjnej.

Typ czynności 2 wskazuje na dokonanie zmiany w określonym polu informacyjnym. Do rekordu kartoteki wprowadzona zostaje nowa wartość z pola informacyjnego rekordu transakcji /w miejsce wartości 315 wprowadzona zostaje wartość 380/ i rekord ten zapisany zostaje w zbiorze kartoteki nowej generacji. Po wykonaniu tej operacji do pamięci wczytane zostają kolejne dwa rekordy, jeden ze zbioru transakcji /wartość klucza 23/ oraz następny rekord zbioru kartoteki /o wartości klucza 23/. Wartości kluczy są równe. Typ czynności rekordu transakcji wskazuje na konieczność zanulowania rekordu o wartości klucza 23. Rekord ten nie zostaje zapisany w kartotece nowej generacji, wczytane zostają do pamięci kolejne dwa rekordy: ze zbioru transakcji /o wartości klucza 27/ i ze zbioru kartoteki /o wartości klucza 25/. Następuje porównanie wartości kluczy. Rekord kartoteki starej generacji zapisany zostaje w zbiorze kartoteki nowej generacji. Zgodnie z procedurą, z kartoteki starej generacji wczytany zostaje kolejny rekord.

Jest nim etykieta końca zbioru starej kartoteki. Typ czynności /o wartości 3/ dla ostatniego rekordu transakcji wskazuje na to, aby wprowadzić ten rekord do zbioru kartoteki nowej generacji. Rekord zostaje zapisany. Ze zbioru transakcji wczytana jest etykieta końca zbioru. Stanowi to sygnał, że przebieg aktualizacji kartoteki został zakończony. Wygenerowana zostaje etykieta końca zbioru kartoteki nowej generacji.

Gdyby w zbiorze transakcji występowały kolejne rekordy o typie czynności = 3, byłyby one również zapisane w zbiorze kartoteki nowej generacji. I odwrotnie, gdyby w zbiorze starej generacji występowały jeszcze dalsze rekordy /tj. jeśli zbiór transakcji skończyłby się wcześniej niż zbiór kartoteki/ rekordy te przepisane zostałyby do zbioru wyjściowego.

Rekordy zbioru kartoteki starej generacji oraz rekordy zbioru transakcji wczytywane są do tzw. obszarów wejścia zbiorów w pamięci operacyjnej. W zależności od aktualnej zawartości tych obszarów definiujemy tzw. stan zbiorów w przebiegu aktualizacji.

Rozróżnia się następujące stany zbiorów w czasie omawianego przebiegu:

- stan zero; w obszarze wejścia transakcji i w obszarze wejścia zbioru kartoteki znajdują się rekordy,
- stan jeden; w obszarze wejścia zbioru kartoteki znajduje się etykieta końca zbioru, w obszarze wejścia transakcji znajduje się rekord.
- stan dwa; w obszarze wejścia transakcji znajduje się etykieta końca zbioru, a w obszarze wejścia zbioru kartoteki znajduje się rekord.
- stan trzy; w obszarze wejścia obu zbiorów znajdują się etykiety końca zbiorów.

W zależności od stanu zbiorów istnieją dla przebiegu aktualizacji następujące zastrzeżenia:

- a/ dla stanu zero lub trzy nie wolno czytać zbioru kartoteki starej generacji

b/ dla stanu dwa lub trzy nie wolno czytać zbioru transakcji ani badać zakresu wartości typu czynności.

Zwróćmy jeszcze uwagę na następujące reguły:

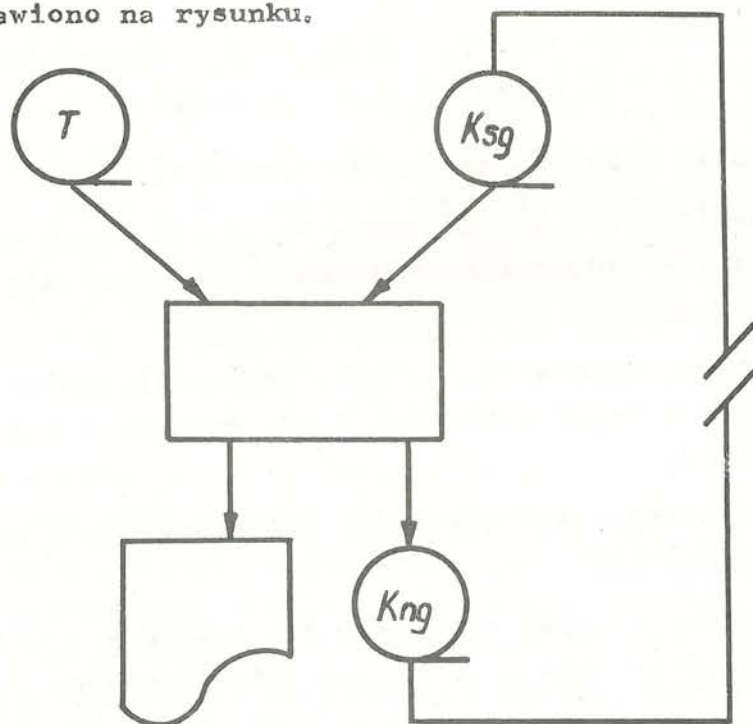
- a/ typ czynności /1,2 lub 3/ muszą znajdować się w rekordach transakcji. W jednym rekordzie może znajdować się tylko jeden typ czynności.
- b/ rekord zbioru transakcji o typie czynności 1 może poprzedzać rekord o typie czynności 3 dla rekordu o tej samej wartości klucza.
- c/ pole informacyjne typ czynności wchodzi w klucz sortowania zbioru transakcji.
- d/ inna kombinacja typów czynności dla rekordów o tej samej wartości kluczy niż w p. b/ jest niedozwolona.

W przebiegu aktualizacji występuje najczęściej również zbiór drukarki wierszowej. Drukowany jest na niej tzw. raport aktualizacji. Raport aktualizacji obejmuje kompleksową informację na temat zmian, anulacji i wprowadzeń dokonanych w kartotece. Informuje również o zauważonych błędach występujących w trakcie przebiegu. Do najczęściej występujących błędów w przebiegu aktualizacji zalicza się:

- wystąpienie w zbiorze t-ransakcji rekordów o typie czynności 1 /anulacja/ o wartości kluczy nie mających swych odpowiedników w rekordach kartoteki,
- wystąpienie w zbiorze transakcji rekordów o typie czynności 3 /wprowadzenie/ o wartości kluczy już istniejących w kartotece,
- wystąpienie w zbiorze transakcji rekordów o typie czynności 2 /zmiana/ o wartości kluczy nie mających swych odpowiedników w rekordach kartoteki.

Oznaczenie przebiegu aktualizacji prostej z równoczesnym wydrukiem raportu aktualizacji na schematach przetwarzania

przedstawiono na rysunku.



Rys. nr 28. Przebieg aktualizacji prostej z równoczesnym wydrukiem raportu aktualizacji.

W kontekście tego co wyżej powiedziano należy zdać sobie sprawę, iż aktualność odpowiednich kartotek systemu przetwarzania danych uzależniona jest od obszaru i zakresu działania danego systemu elektronicznego przetwarzania danych. Tym samym, od tych czynników zależy częstotliwość aktualizacji zbiorów kartotekowych. Odstępy czasu pomiędzy aktualizacjami kartoteki mogą wahać się od kilku miesięcy do kilku zaledwie dni. W systemach operatywnego planowania produkcji, w których planuje się produkcję dobowo-zamianową- kartoteka zawierająca planowane obciążenie musi być aktualizowana codziennie. Aktualna kartoteka dnia dzisiejszego staje się kartoteką starej generacji już w dniu jutrzejszym. Stwarza to wymóg jaknajprecyzyjniejszego i najekonomiczniejszego opracowywania przebiegów aktualizujących kartoteki. Z tej racji też należy pamiętać w tym, że przebieg aktualizacji jest przebiegiem trudnym i skomplikowanym. ^{1/}

1/ Rozszerzenie tego tematu znajdzie Czytelnik w pracy /3/.

3. Spis rysunków

- Rys. nr 1. Organizacja zapisu rekordów w bloku fizycznym na taśmie magnetycznej.
- Rys. nr 2. Oznaczania zbiorów danych.
- Rys. nr 3. Przebieg scalania prostego.
- Rys. nr 4. Przykład scalania zbiorów o różnych wartościach kluczy.
- Rys. nr 5. Przebieg scalania złożonego.
- Rys. nr 6. Przebieg dobierania prostego.
- Rys. nr 7. Przykład dobierania prostego na podstawie wartości kluczy.
- Rys. nr 8. Przykład przebiegu dobierania, w którym ostatnie rekordy posiadają równe wartości kluczy.
- Rys. nr 9. Przykład przebiegu dobierania, w którym jeden ze zbiorów wejściowych jest większy od drugiego zbioru wejściowego.
- Rys.nr 10. Przykład dobierania prostego z procedurą obliczeniową.
- Rys.nr 11. Przykład przebiegu dobierania prostego, w którym rekord wynikowy jest jakościowo inny od rekordów kartotek wejściowych.
- Rys.nr 12. Przebieg dobierania z równoczesnym wydrukiem raportu.
- Rys.nr 13. Przebieg dobierania wielofazowego.
- Rys. nr 14. Przebieg sortowania.
- Rys.nr 15. Przykład przebiegu sortowania.
- Rys.nr 16. Założenie zbioru nieuporządkowanego.
- Rys.nr 17. Wymiana szpuli taśmy na fizycznej jednostce nr 1.
- Rys.nr 18. Wykorzystanie fizycznych jednostek taśm magnetycznych w przebiegu sortowania.
- Rys.nr 19. Przebieg rozdziału.
- Rys.nr 20. Przebieg rozdziału z równoczesnym wydrukiem raportu.
- Rys.nr 21. Przebiegi konwersji wejścia.
- Rys.nr 22. Przebiegi konwersji wejścia z równoczesną kontrolą poprawności informacji wprowadzanych.

- Rys.nr 23. Przebiegi konwersji wyjścia.
Rys.nr 24. Przebieg wybierania prostego.
Rys.nr 25. Złożone przebiegi wybierania.
Rys.nr 26. Przebieg aktualizacji prostej.
Rys. nr 27. Przykład przebiegu aktualizacji prostej.
Rys.nr 28. Przebieg aktualizacji prostej z równoczesnym wydrukiem raportu aktualizacji.

4. Wykaz literatury.

1. M. Greniewski: "Robot kierownictwa - automatyczne przetwarzanie danych" PWN Warszawa 1967 r.
2. Praca zbiorowa pod redakcją
M. Greniewskiego: "Technologie procesów przetwarzania danych dla zarządzania. PWE Warszawa 1972 r.
3. U.Szmidt, E. Kurzydem, "Zasady projektowania zbiorów przechowywanych na TM lub DM, stanowiących masyw danych systemu" Warszawa 1970 r.







