

Redagowanie informacji wyjściowej jest często procesem odwrotnym w stosunku do redagowania informacji wejściowej i może obejmować następujące czynności:

- 1) tłumaczenia informacji zakodowanej,
- 2) podawania okresu ważności dla zbioru danych,
- 3) upakowywania i przegrupowywania danych,
- 4) włączania lub wyłączania znaków, szczególnie mocno rozbudowane w razie wydruku wyników. Występuje wówczas konieczność wymazywania zer z lewej strony, spacjowania, wstawiania przecinków, symboli (np. takich jak „zł”),
- 5) tworzenia sumarycznych zapisów po analitycznym przetwarzaniu większych grup informacji, np. kartotek,
- 6) dokonania podziału wydruku zgodnie z logicznym podziałem zagadnienia,
- 7) zredagowania informacji na temat czasu i kosztu przetwarzania itp.

4.5. Sortowanie

Celem sortowania jest ułożenie transakcji w określonym porządku, np. umożliwiającym sekwencyjną aktualizację kartoteki. Przy sortowaniu posłużymy się np. taśmami magnetycznymi. Istnieje wiele metod sortowania. Przedstawimy metodę sortowania przez dobieranie (*by merging*), używając 4 taśm magnetycznych. Rozróżniamy trzy główne fazy sortowania:

A. Utworzenie tzw. szeregów monotonicznych, zwanych też łańcuchami (*strings*). Jeśli sortowanie odbywa się według indeksów w kolejności wzrastającej, to wyszukuje się sekwencyjnie ułożone indeksy i przepisuje je z jednej taśmy wejściowej na przemian na dwie taśmy wyjściowe.

7, 8, 13 6, 14, 30 11, 40 5, 9, 10, 12
stringi

B. Dobieranie łańcuchów z dwóch taśm wejściowych na dwie taśmy wyjściowe. Operacji dokonuje się wielokrotnie tak długo, aż na obu taśmach wyjściowych transakcje zostaną ułożone w żądanej kolejności.

Wielokrotność operacji wyniesie:

$$\log_K = \frac{N}{G}$$

gdzie:

K — liczba taśm wyjściowych (w naszym przykładzie dwie, może być więcej),

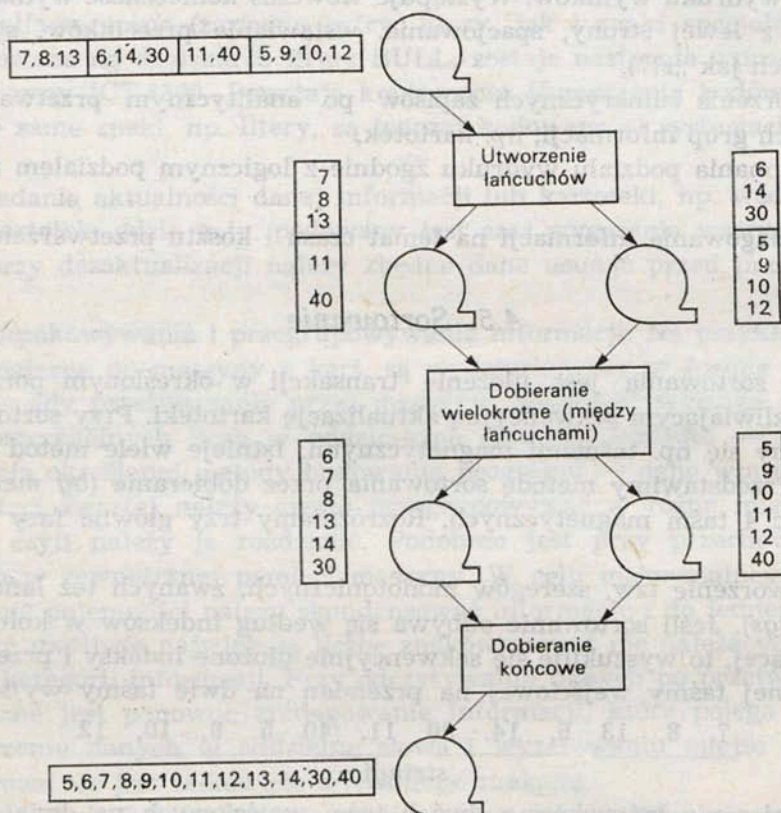
N — liczba transakcji.

G — liczba łańcuchów.

C. Dobieranie końcowe w celu utworzenia posortowanej taśmy. Przebieg sortowania ilustruje ryc. 4.5.

Z praktyki wynika, że średnio 30% czasu pracy maszyny w zakresie

przetwarzania danych przypada na sortowanie. Z tego względu poświęca się wiele uwagi technice sortowania. Każda dobrze oprogramowana maszyna jest zaopatrzona w tzw. tablice sortowania, gdzie w zależności od: prędkości taśm magnetycznych, liczby transakcji, długości transakcji, liczby kanałów sterowania, a dla niektórych maszyn od długości indeksu, można wyznaczać czasy sortowania z dokładnością do 1 sekundy. Poda-
jemy dla przykładu wycinek tablicy sortowania maszyny Gamma 30:

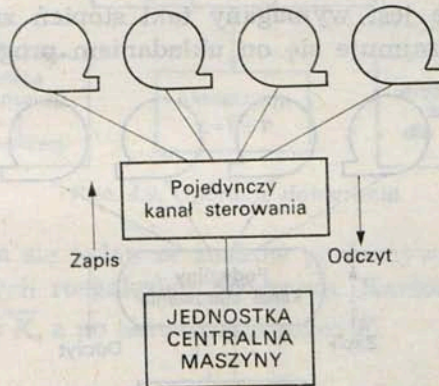


Ryc. 4.5. Przebiegi sortowania

Duży wpływ na prędkość sortowania ma liczba kanałów sterowania jednostkami taśm magnetycznych. Przy jednym kanale (por. ryc. 4.6) nie ma możliwości współbieżnego pisania i czytania taśmy.

Przy dwóch kanałach (por. ryc. 4.7) istnieje taka możliwość, ale w jednostkach taśm przyłączonych na stałe do tych kanałów. Najbardziej sprawny jest tzw. kanał podwójny (por. ryc. 4.8), który umożliwia równoczesne pisanie i czytanie przy dwóch dowolnie wybranych jednostkach taśmy magnetycznej. Takie rozwiązanie zostało zastosowane po raz pierwszy przez firmę GE dla maszyn serii 400 i 600. Czas sortowania na taśmach o prędkości 20 Kc ze sterowaniem podwójnym odpowiada taśmom średnio o 80% szybszym, tzn. w granicach 36 Kc.

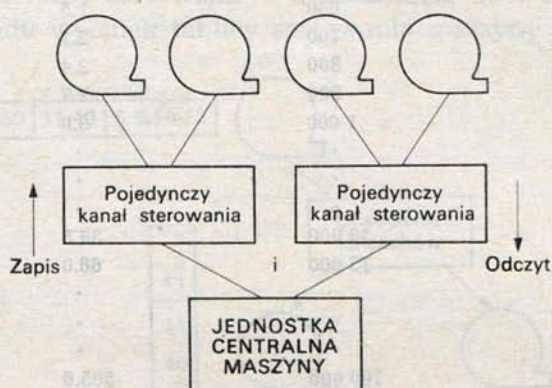
Pojemność PAO 20 000 znaków Długość pozycji	Liczba pozycji	Czas sortowania w min. przy prędkości	
		33 Kc	66 Kc
400	500	1.3	0.9
•	600	1.5	1.1
•	700	2.1	1.5
•	800	2.4	1.7
•	900	2.7	1.9
•	1 000	3.0	2.2
•	•	•	•
•	•	•	•
•	10 000	39.5	28.6
•	15 000	66.0	47.1
•	•	•	•
•	•	•	•
•	100 000	505.6	372.1
450	500	1.4	1.0
•	•	•	•
•	•	•	•
•	•	•	•
•	100 000	560.7	408.2



Ryc. 4.6. Synchronizacja taśm magnetycznych i jednostki centralnej maszyny pojedynczym kanałem sterowania

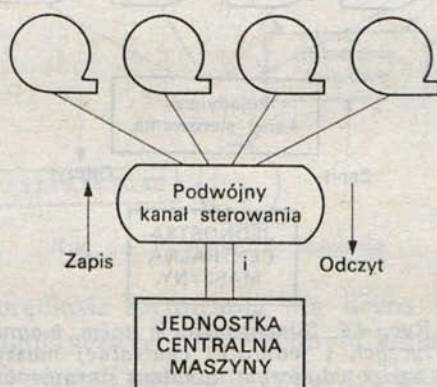
W wypadku transakcji umieszczonych na kartach powstaje często problem, czy lepiej je sortować pośrednio przy użyciu sortera, czy też wykonywać tę czynność na komputerze. Odpowiedź zależy przede wszystkim od liczby transakcji. Mała liczba transakcji będzie szybciej uporządkowana na sorterze, gdyż czas potrzebny na przygotowanie do przetwarzania na maszynie jest stosunkowo długi. Przy większej liczbie transakcji bardziej efektywne z punktu widzenia czasu jest użycie komputera.

Sortowanie za pomocą komputerów jest średnio 20—100-krotnie szybsze, jednakże koszt przetwarzania na sorterach jest mniej więcej tyle razy mniejszy. Przy stosowaniu komputerów do sortowania nie powstają błędy, a przy stosowaniu sorterów — liczba błędów może być duża ze względu na duży udział pracy ręcznej.



Ryc. 4.7. Synchronizacja taśm magnetycznych i jednostki centralnej maszyny dwoma pojedynczymi kanałami sterowania

Sortowanie jest czynnością bardzo często wykonywaną. Budowa programów sortowania wymaga pewnego zasobu wiedzy. Jednakże od projektanta systemu nie jest wymagany taki stopień znajomości tego zagadnienia, gdyż nie zajmuje się on układaniem programów sortowania,



Ryc. 4.8. Synchronizacja taśm magnetycznych i jednostki centralnej maszyny podwójnym kanałem sterowania

lecz tylko z nich korzysta, najczęściej z programów uniwersalnych. Projektant systemu musi tylko podać, na których indeksach (kluczach) pozycji odbędzie się sortowanie i w jakiej kolejności. Może on również — wykorzystując niektóre cechy programów sortowania — włączać w toku przebiegów A i C własne sekwencje rozkazów.

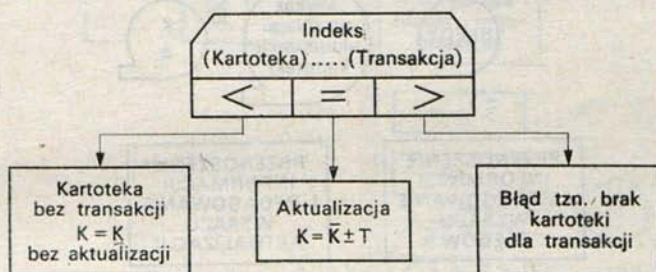
4.6. Dobieranie

Dobieranie jest jedną z najczęściej powtarzających się sekwencji operacji w przetwarzaniu maszynowym. Podczas dobierania dane muszą być uprzednio posortowane. Dobieranie obejmuje m.in. następujące czynności:

- a) dobieranie dwóch zbiorów danych według określonego klucza w celu utworzenia jednego zbioru,
- b) rozłączanie jednego zbioru danych na kilka podzbiorów według określonego klucza,
- c) wybieranie z danego zbioru danych tylko tych, które są zgodne z kluczem wzorcowym.
- d) dołączanie pewnych danych do zbioru, jeśli są zgodne z kluczem wzorcowym.

Dobieranie jest podstawową operacją w procesie aktualizacji kartoteki, którą ilustruje ryc. 4.9.

W symbolu porównywania, w nawiasach wstawia się indeksy aktualnie porównywanej pozycji kartotekowej i transakcji. W miejsce podwójnego



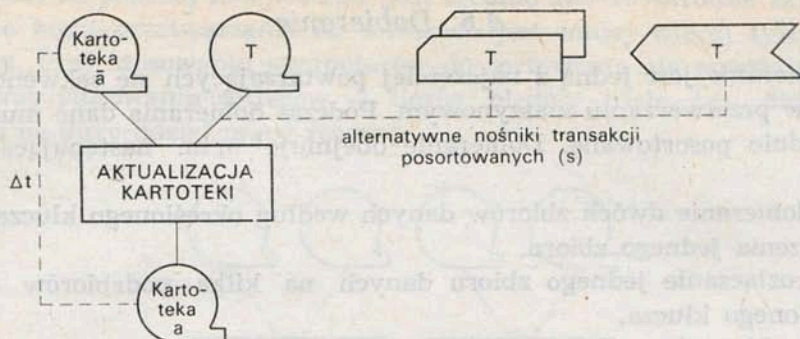
Ryc. 4.9. Operacje dobierania

dwukropka wstawia się jeden ze znaków porównywania podczas analizowania poszczególnych rozgałęzień programu. Kartotekę przed aktualizacją oznaczono przez \bar{K} , a po aktualizacji przez K .

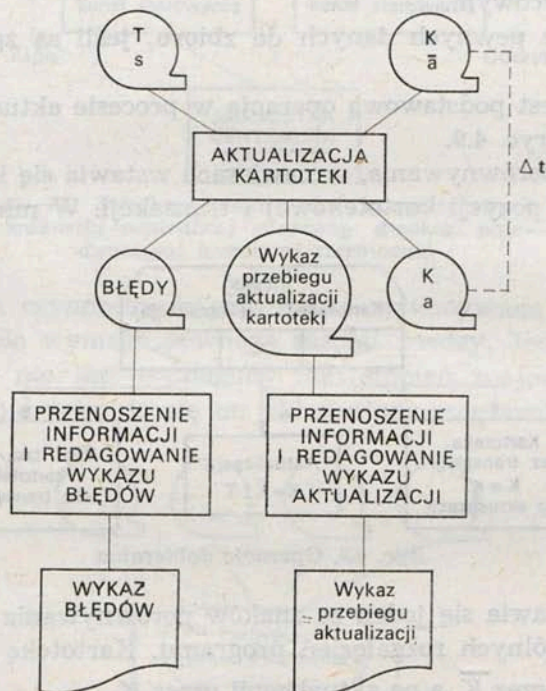
4.7. Aktualizacja kartotek

W procesie przetwarzania danych jednym z podstawowych i najczęściej stosowanych przebiegów jest aktualizacja kartoteki. Przedstawia to ryc. 4.10.

Przy zastosowaniu taśm magnetycznych uaktualniona na dany dzień kartoteka jest w całości przepisywana na nową taśmę magnetyczną. Nie-raz pożądane jest uzyskanie do celów analitycznych informacji na temat każdorazowego przebiegu aktualizacji, np. ujęcie procentowe przyrostów stanów, przychodów, rozchodów. Powstawanie takiego wykazu ilustruje ryc. 4.11.



Ryc. 4.10. Schemat aktualizacji kartoteki



Ryc. 4.11. Schemat przetwarzania dla otrzymania wykazu przebiegu aktualizacji kartoteki

W toku aktualizacji mogą powstać błędy niedobrania kartoteki z transakcjami, spowodowane np. brakiem kartoteki dla danej transakcji lub złą symbolizacją. Odpowiednio zastosowane metody kontroli (por. pkt. 4.3.) mogą przeciwdziałać powstawaniu błędów lub szybko je wykrywać.

4.7.1. Typowe metody aktualizacji kartotek

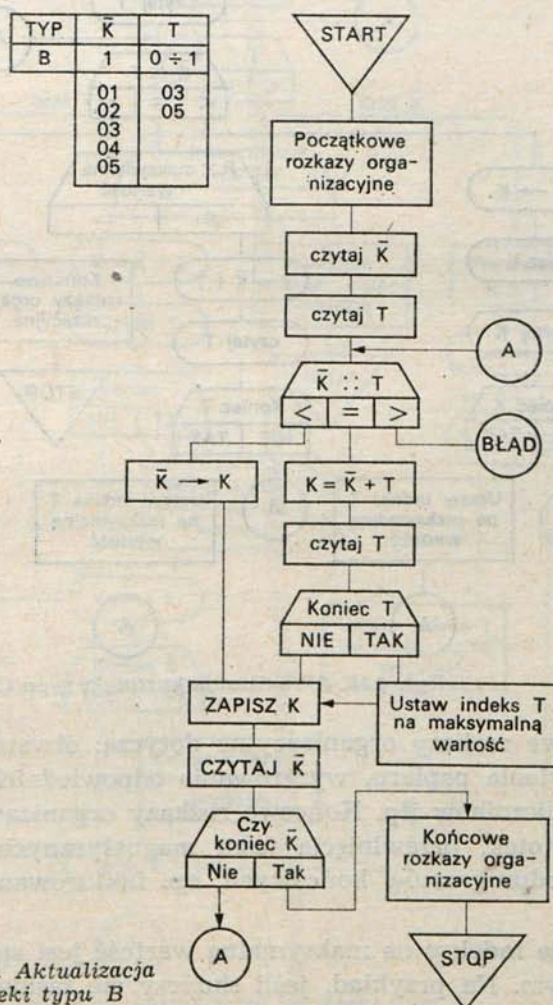
Aktualizację kartotek można sklasyfikować według kryterium liczby transakcji i kartotek, związanych ze sobą przetwarzaniem (por. tabela 4.1).

Tablica 4.1

Rodzaje aktualizacji kartotek

Typ	Kartoteka K	Transakcje T
A	1	1
B	1	0—1
C	1	0—n
D	0—1	0—1
E	0—1	0—n

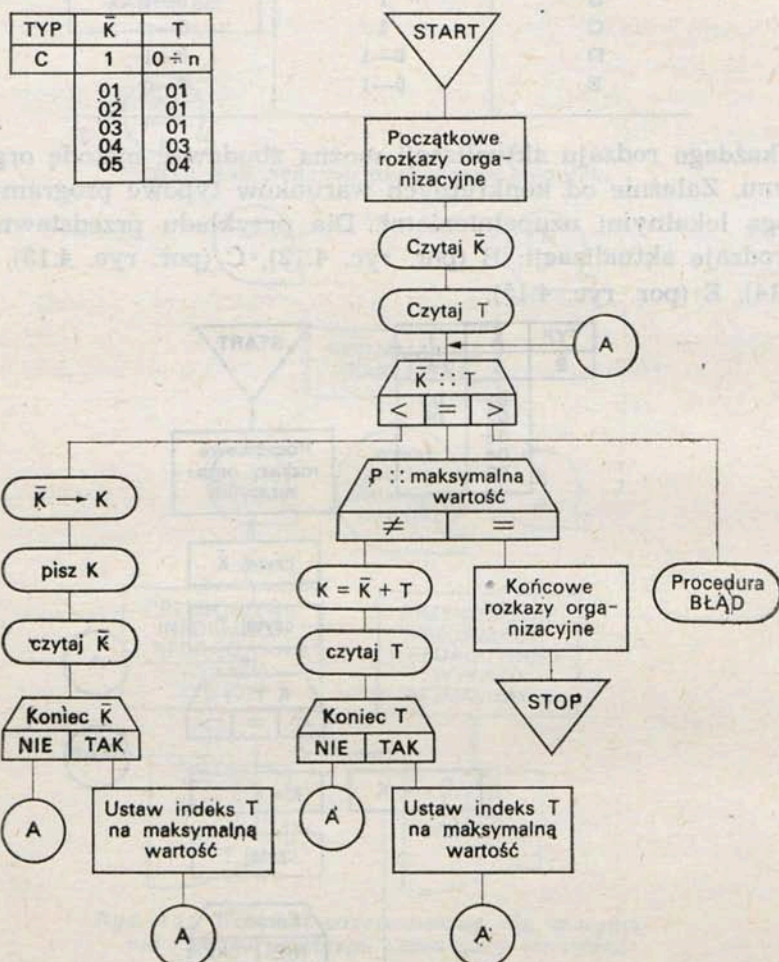
Dla każdego rodzaju aktualizacji można zbudować metodę organizacji programu. Zależnie od konkretnych warunków typowe programy różnić się mogą lokalnymi uzupełnieniami. Dla przykładu przedstawimy niektóre rodzaje aktualizacji: B (por. ryc. 4.12), C (por. ryc. 4.13), D (por. ryc. 4.14), E (por. ryc. 4.15).



Ryc. 4.12. Aktualizacja kartoteki typu B

W przedstawionych schematach blokowych programów wyjaśnienia wymagają:

- początkowe rozkazy organizacyjne,
- końcowe rozkazy organizacyjne,
- ustawienie indeksu na maksymalną wartość.

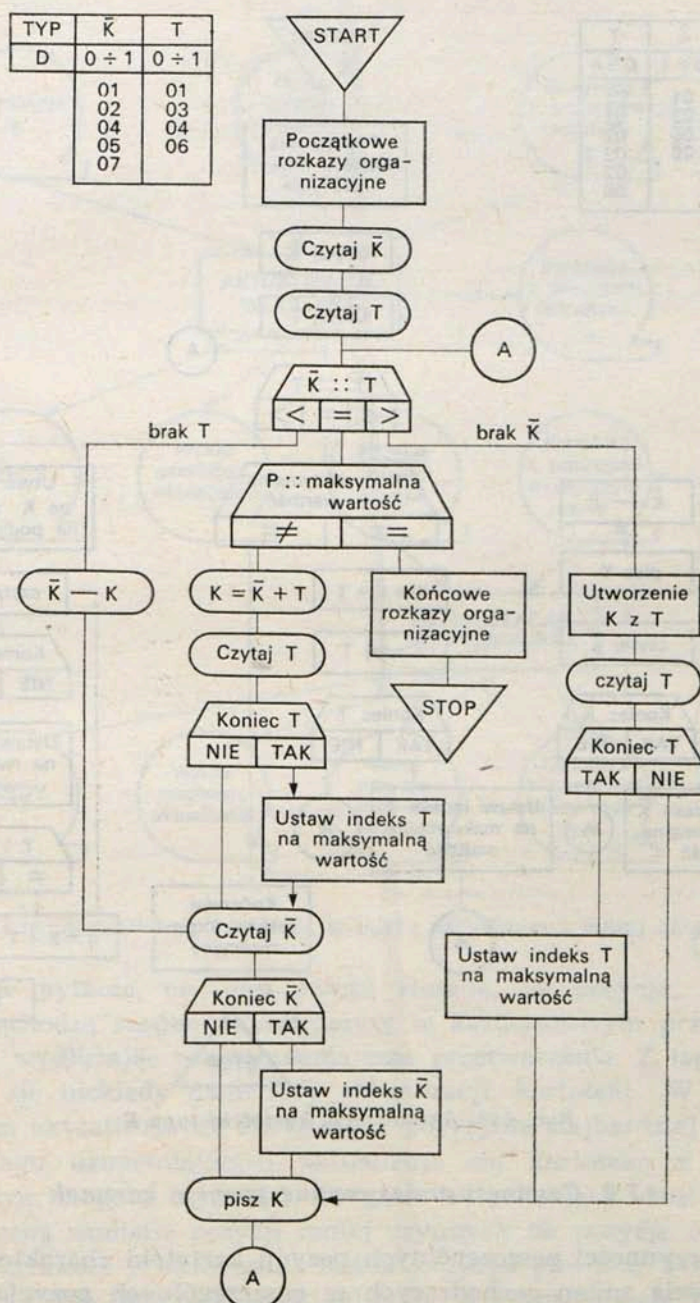


Ryc. 4.13. Aktualizacja kartoteki typu C

Początkowe rozkazy organizacyjne dotyczą: otwarcia kartotek, pierwszego ustawienia papieru, wyzerowania odpowiednich komórek pamięci, ustawienia liczników itp. Końcowe rozkazy organizacyjne dotyczą: zamknięcia kartotek, przewinięcia taśm magnetycznych, wywołania odpowiednich podprogramów końcowych, np. fakturowania za czas przetwarzania.

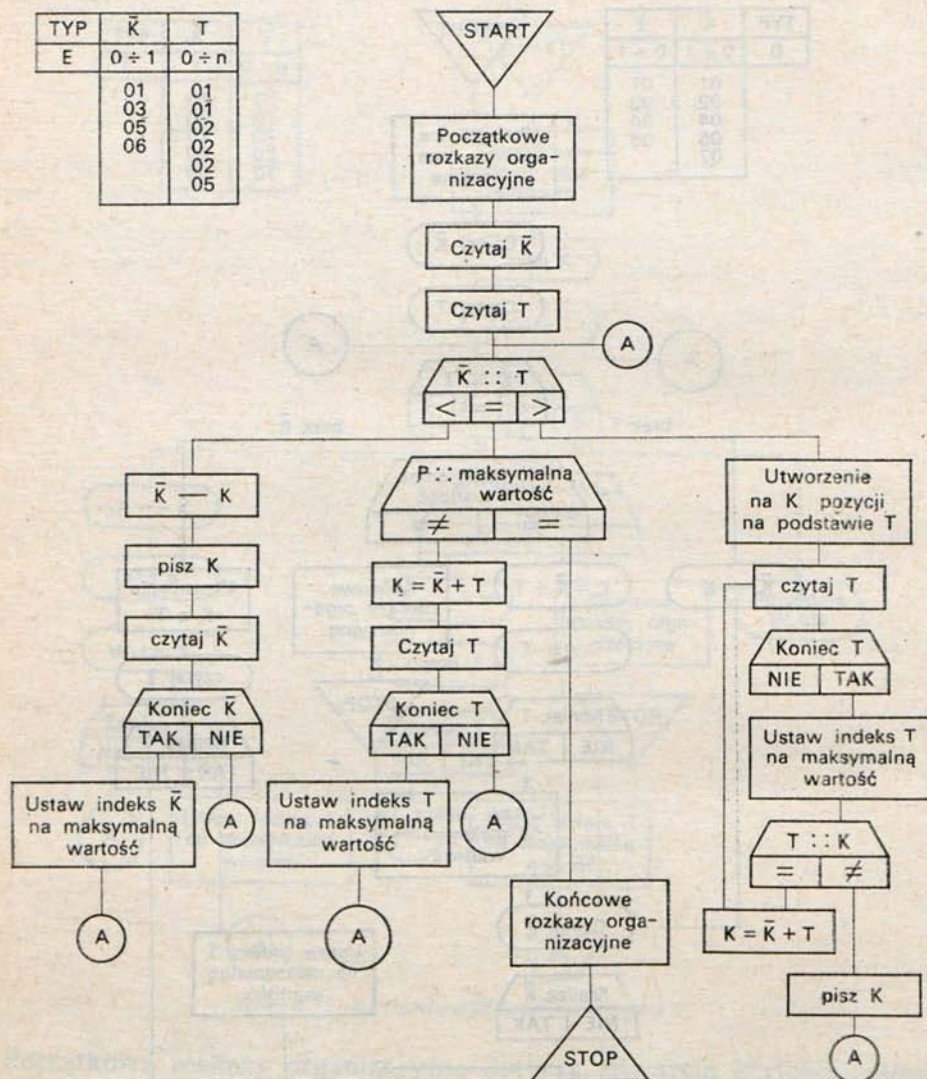
Ustawienie indeksu na maksymalną wartość jest specjalnym zabiegiem programowym. Na przykład, jeśli skończy się taśma z transakcjami, to

taśma z kartotekę \bar{a} może być jeszcze nie przeczytana do końca. Wszystkie pozycje tej kartoteki należy przepisać na kartotekę a . Przy skoku do porównania $K :: T$ należy przewidzieć stałe przechodzenie gałęzią $\bar{K} - \bar{K}$. Wówczas — po wykryciu końca T w komórce pamięci przechowującej



Ryc. 4.14. Aktualizacja kartoteki typu D

indeks T — programem generuje się maksymalną wartość indeksu, który będzie się składał z wielu tych samych znaków. Na przykład dla GE 400 będzie to „1” o aktualnej wartości $(77)_8$. Umożliwia to stałe przechodzenie gałęzią $K - K$; wynik porównania z góry będzie znany.

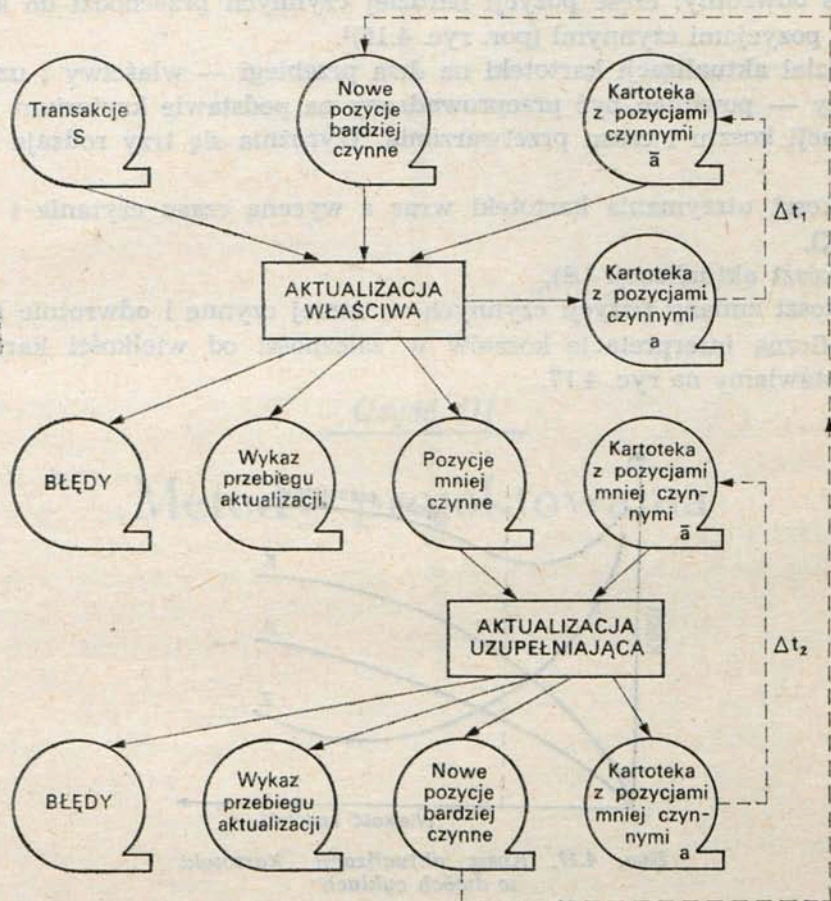


Ryc. 4.15. Aktualizacja kartoteki typu E

4.7.2. Czynne i mniej czynne pozycje kartotek

Stopień czynności poszczególnych pozycji kartoteki charakteryzuje się częstotliwością zmian zachodzących w poszczególnych pozycjach kartoteki podczas przebiegu aktualizacji. W wielu kartotekach pewne pozycje

kartotekowe podlegają znacznie większej liczbie zmian w odróżnieniu od innych. W kartotece magazynowej można wyróżnić takie rodzaje materiałów, które są zawsze w obrocie, i inne, które są rzadko w użyciu, np. ewidencja obrotu towarów sezonowych.



Ryc. 4.16. Aktualizacja kartoteki w cyklu aktywnym i mniej aktywnym

Powstaje pytanie, czy jest rzeczą słuszną, by pozycje, w których zmiany zachodzą rzadko, uczestniczyły w każdorazowym przebiegu aktualizacji, wydłużając niepotrzebnie czas przetwarzania. Z tego względu rozróżnia się niekiedy dwie fazy aktualizacji kartoteki. W przebiegu właściwym aktualizuje się kartotekę z pozycjami najbardziej czynnymi, w przebiegu uzupełniającym aktualizuje się kartotekę z pozycjami o mniejszym stopniu czynności. Program aktualizujący musi zapewniać automatyczną zamianę pozycji mniej czynnych na pozycje o większym stopniu czynności i odwrotnie. Każda pozycja powinna zawierać datę ważności i liczby przebiegów aktualizacji, w których może brać udział.

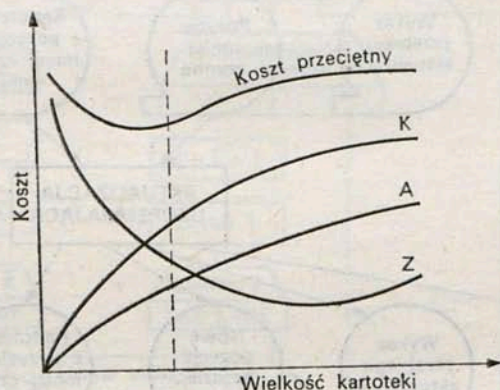
Równoległe z taśmą błędów powstaje nowy rodzaj zapisu wymagający

oddzielnej taśmy w celu zgrupowania tych wszystkich transakcji, które powinny aktualizować kartotekę z pozycjami mniej czynnymi, oraz tych pozycji kartoteki, które z czynnych stały się mniej czynne. Podczas aktualizacji pozycji kartoteki o mniejszym stopniu aktywności występuje proces odwrotny: część pozycji bardziej czynnych przechodzi do kartoteki z pozycjami czynnymi (por. ryc. 4.16)¹.

Podział aktualizacji kartoteki na dwa przebiegi — właściwy i uzupełniający — powinien być przeprowadzony na podstawie kryterium minimalizacji kosztu i czasu przetwarzania. Wyróżnia się trzy rodzaje kosztów:

- a) koszt utrzymania kartoteki wraz z wyceną czasu czytania i pisania (K),
- b) koszt aktualizacji (A),
- c) koszt zmiany pozycji czynnych na mniej czynne i odwrotnie (Z).

Graficzną interpretację kosztów w zależności od wielkości kartoteki przedstawiamy na ryc. 4.17.



Ryc. 4.17. Koszt aktualizacji kartoteki w dwóch cyklach

Minimum kosztu przeciętnego wskazuje na optymalną wielkość kartoteki z pozycjami czynnymi, przy której opłacalne jest rozróżnianie stopnia aktywności pozycji kartotekowych.

Należy podkreślić statyczny moment rozpatrywania opłacalności podziału przebiegu aktualizacyjnego. Zmiana stopnia czynności pozycji kartoteki zachodzi w sposób dynamiczny, przez co minimum funkcji kosztu jest ruchome i trudne do określenia.

¹ Rysunek pochodzi z pracy H.N. Laden, T.R. Gidersleeve, *System Design for Computer Applications*, John Wiley and Sons, wyd. II, New York 1964, s. 60.