

txt020262 ファイル編成法

① ファイル編成法の種類

① 順編成ファイル

一連の論理レコードを物理的に連続した順序で記憶したファイルである。レコードはキー順に並んでいるか、入力した順番に並んでいる。

② 直接編成ファイル

特定のレコードを見つけるのにそのレコードが記録されているアドレスを用いて、特定のレコードのみを処理できる。

③ 区分編成ファイル

メンバと呼ばれる区分された順次編成ファイルの集合で、直接アクセスによってメンバを検索できるディレクトリがある。

④ 索引編成ファイル

基本的には順編成ファイルであるが、特定のレコードを直接アクセスするために索引を利用する。順次アクセス、直接アクセスが可能である。

⑤ VSAMファイル

仮想記憶オペレーティングシステムのもとで利用できるファイル編成方式で、入力順データセット、キー順データセット、相対レコードデータセットの3種類の編成法がある。データセットの構造はシリンダに相当する制御域とトラックに相当する制御インタバルで構成される。

② 順編成

① 構造と用途

順編成は一連の論理レコードを物理的に連続した順序で記録したファイルである。通常、レコードはキー順に物理的に連続して並んでいる。キー順に入力するのは、キー順に処理する仕事が多いためである。順編成ファイルはマスターファイルやトランザクションファイル(変動ファイル)に広く利用され、ファイルの突き合わせ作業に用いられる。

② 領域確保とレコードの入力

領域の確保はファイルを記憶させるために必要な大きさの場所をディスク上に確保することであり、この場所をエクステントでVTOCに登録する。入力レコードをキー順に並べ、ユー

ユーティリティプログラムを利用して磁気ディスクに書き込む。入力レコードは入力した順番にエクステントの最初のトラックから連続して記録されていく。記録されるのは物理レコードである。固定長、可変長、不定長のすべての形式の書き込みが可能である。レコードのブロック化はユーティリティプログラムが行う。

㉓ 更新

順編成ファイルの更新とは状況の変化に対応して、ファイル内の特定のレコードのデータ項目を変えていくことである。新しいデータを含んだ変動ファイルとマスタファイルを突き合わせて、マスタファイルの内容を更新する場合に利用する。

磁気ディスクの更新は、更新したレコードをディスク上のもとの場所に戻す。磁気テープの更新は、別の磁気テープに新規の順編成ファイルとして作成する。更新時にレコードの長さの変更を伴うものは、磁気ディスクの場合でも新マスタファイルは新規のファイルとして別に用意したエクステントに作成する。

㉔ 追加・削除

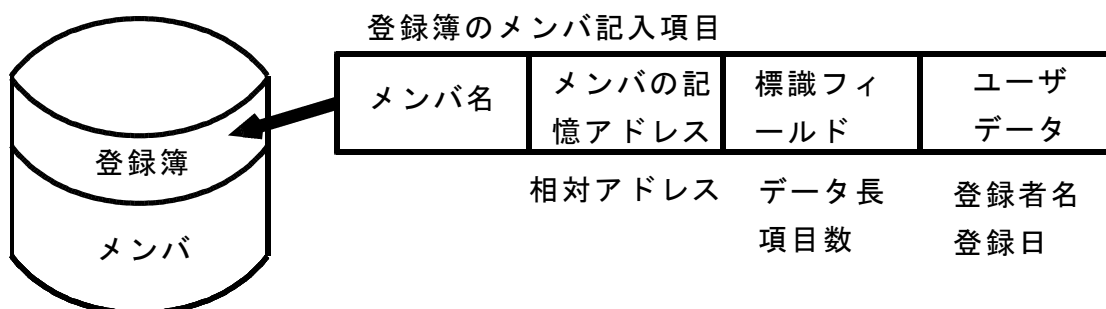
順編成の削除は、変動ファイル、マスタファイルからそれぞれのレコードを読み、キー項目の値が一致すればそのレコードを削除する。通常は、レコードの削除は論理的な削除であり、物理的な削除はしない。ファイルの再編成時に物理的な削除を行う。

磁気テープの場合、変動レコードとキー項目の一致しないレコードは新マスタファイルに書き写し、一致するレコードは書き写さない。書き写さない時は物理的に削除されてしまう。順編成の追加は、新しいレコードを既存のファイルの所定の箇所に挿入することである。

マスタファイルと変動ファイルを突き合わせ、順編成ファイルの連続したレコードの間に追加レコードを挿入する場合は、新マスタファイルを旧マスタファイルと別個に作り直す必要がある。磁気ディスク、磁気テープとも同様の処理になる。

③ 区分編成

㉑ 区分編成とは



区分編成は磁気ディスク等の直接アクセスが可能な媒体が対象となり、メンバとディレクト

リ(登録簿)で構成される。

⑥ メンバ

ファイル単位でメンバは構成される。メンバ内では、ファイルは順次編成で記録され、入力順に記憶される。メンバを処理するときは、名前を指定し、名前をディレクトリで探索し、該当のメンバを取り出す。メンバ毎に迅速なアクセスができる。

⑦ ディレクトリ(登録簿)

メンバ索引用登録簿で、メンバの名称、メンバの記憶アドレスをメンバ名の昇順に記録する。メンバ名は、メンバを作成するときに命名し記入項目に登録される。メンバの記憶アドレスは、そのメンバが記憶されている場所の先頭アドレスが相対トラック番号と相対レコード番号で記憶される。標識フィールドは、後続のユーザデータの長さ、その中のデータ項目の個数などが記憶される。ユーザデータはそのメンバの登録者名、登録日、バージョンなどが記憶される。

⑧ 区分編成ファイルの用途

各種プログラムファイル、システムの各種パラメータを記録するシステムファイル等に利用される。代表的なものが、プログラムライブラリで、一つのプログラムがメンバとなる。ジョブの実行順序を記述したジョブ制御言語、コマンドのプロシジャライブラリの格納に利用する。プロシジャ名をメンバ名とする。

⑨ 区分編成ファイルの作成

ディレクトリおよびメンバ記憶域の容量を決定し、記憶媒体上にファイルに必要な記憶領域を設定する。区分編成ファイル用処理プログラムを使用して、ディレクトリ領域を設定し、メンバを順次編成で読み込む。メンバ作成後にディレクトリにメンバ名とアドレスが記録される。

⑩ 変更・追加・削除

メンバ内のデータ項目を更新することは可能であるが、データ項目の新規追加、削除は不可能である。メンバ単位の追加・削除は、メンバおよびディレクトリ領域内の該当メンバの記入項目を追加削除して行う。削除の場合、メンバ領域のファイルはそのまま保存される。空きファイルの整理や追加メンバ記憶領域の確保を目的に、適宜、区分編成ファイルの再構成を行う。

⑪ 区分編成法の特徴

- ㊦ メンバと呼ばれる区分された順次編成ファイルと登録簿で構成される。
- ㊧ メンバ内のファイルは順次アクセスによる。
- ㊨ メンバはディレクトリを用いて直接アクセスにより検索が可能である。
- ㊩ ファイル単位での管理に適する。
- ㊪ 記憶媒体は、磁気ディスクなどのDASDを用いる。

- ㊦ 各種プログラムファイル、システムの各種パラメータなどを記録するシステムファイルなどに用いる。
- ㊧ ファイルの編成はメンバ領域にファイルを順次編成し、ディレクトリを作成する
- ㊨ 更新は、メンバ内のデータ項目でも可能であるが、追加や削除はメンバ単位でのみ可能である。
- ㊩ 追加削除は、メンバおよびディレクトリの該当領域の修正により行う。
- ㊪ 削除ファイルは、そのまま保存されるため、記憶領域の効率的な利用を考慮してファイルの再編成が必要である。

④ 索引編成

㊰ 索引編成ファイルの構造

基本データ域は、ファイルレコードそのものを記録する場所であり、レコードは、キー順に記録される。レコード形式は、キー付きになる。索引域は、トラック索引、シリンダ索引、マスタ索引の3レベルから構成される。

㊱ トラック索引

基本データ域のそれぞれのシリンダ毎に、各シリンダのトラック0に作られる。役割はそのシリンダ内のレコードを迅速に探索することにある。そのシリンダの各トラックに收容されているレコードの最大キーの値とそのトラック番号を記入する。トラック索引を調べると、特定のレコードがどのトラックにあるかが分かる。トラックが分かると、そのトラックの先頭から順に探して、該当レコードを見つければよい。

㊲ シリンダ索引



シリンダ索引は、ファイルに対して一つだけ作る。特定レコードが收容されているシリンダを迅速に確定するためである。各シリンダに收容されているレコードの最大キーとシリンダ番号を記入する。特定レコードを検索する場合、シリンダ索引からそのレコードが收容されてい

るシリンダを確定する。そのシリンダのトラック索引を調べ、該当トラックを見つけ、そのトラック内を順次探索して該当レコードを見つける。

④ マスタ索引

シリンダ索引自体が大きくなって、その探索に時間がかかるときに作成する。マスタ索引を作成することによってシリンダ索引の探索が効率的になる。マスタ索引の記入項目はシリンダ索引の各トラックに対応して1個ずつ作られる。シリンダ索引のトラックの最大キーとトラック番号が記入される。特定のレコードを検索する場合、最初にマスタ索引を調べシリンダ索引の該当トラックを確定し、そのシリンダ索引のトラックを調べ該当シリンダを確定する。そのシリンダのトラック索引を調べ該当レコードのトラックを確定し、そのトラック内を順次探索して該当レコードを見つける。

⑤ あふれ域

レコードは基本データ域にキー順に連続して並んでいる。レコードの追加を行う場合、キーの値に応じて途中に挿入する。追加レコードの挿入されたトラックは結果としてレコードがあふれる。あふれたレコードはあふれ域に収容される。あふれ域はシリンダあふれ域と独立あふれ域がある。シリンダあふれ域は、そのシリンダ内のトラックからあふれたレコードを収容する。独立あふれ域は、シリンダあふれ域からさらにあふれたレコードを収容する。

⑥ 索引編成ファイルの作成

順編成ファイルの作成と同じである。キーの昇順に並べた入力レコードを用意し、ユーティリティプログラムでファイルに読み込んでいく。レコードは、基本データ域にキー順に並べられる。索引は、ユーティリティプログラムが自動的に作成する。基本データ域、索引域、あふれ域は、ファイル設計時にあらかじめ用意しておく。

⑦ 順次アクセスによる処理

基本的には順編成ファイルでのやり方と同じである。レコードの更新や追加、削除は、変動ファイルと基本ファイルの突合せで行う。レコードの追加はファイルの途中に追加することができない。追加できるのはファイルの最後のレコード以降である。

⑧ 直接アクセスによる処理

レコードの更新と削除、追加は変動ファイルと基本ファイルの突合せの形で行う。読み取った変動レコードのキー値をセットすれば索引を調べて該当レコードを見つける。ファイル上にはレコードはキー順に連続して記録されている。追加レコードもこのキー順に該当の場所に挿入される。

挿入時にあふれたレコードはシリンダあふれ域に収容され、トラック索引の内容の正規記入項目、あふれ記入項目が更新される。基本データ域のレコードとあふれ域のレコードはチェーン法によって関係がとられている。あふれ記入項目にはあふれレコードを含めたそのトラックの最高キーを記憶し、あふれレコードが記憶されているアドレスが記憶される。

あふれレコードをあふれ域に記憶する場合、レコードの先頭に関係フィールドを設定し、次のレコードのアドレスを記憶する。関係フィールドを利用して基本データ域のレコードとあふれ域のレコードとの関係が図られて、レコードがキー順にアクセス可能になる。

トラック索引以降の検索の場合、トラック索引のトラック 1 に対する正規記入項目を調べる。正規記入項目になればトラック 1 のあふれ記入項目を調べる。あふれ領域にあれば、あふれ記入項目から最新のあふれレコードのアドレスを調べて、そのレコードの関係フィールドを順次辿って目的のレコードにたどり着く。

シリンダあふれ域はファイル作成時に各シリンダに必要なトラック数を指定する。追加レコードが各シリンダに均等に発生しないときは独立あふれ域を利用する。独立あふれ域はシリンダあふれ域からあふれたレコードを収容するために設ける。シリンダあふれ域を設けずに独立あふれ域だけを作り、基本データ域からあふれたレコードをすべて独立あふれ域に収容することもできる。

① ファイルの再編成

あふれレコードが多くなると処理時間が長くなる。レコードの削除が多くなると、場所の無駄遣いになる。この問題を解決するためにファイルの再編成を行う。ファイル再編成によってあふれ域のレコードはすべて基本データ域に移される。また、削除レコードはファイル上から削除される。再編成は基本データ域の余裕がないときは別の大きな区域に作り直す。再編成はユーティリティプログラムを使用して行う。

① 索引順次編成ファイルの特徴

- ㊦ レコードはキー値順に記録される。
- ① 索引はトラック毎のレコードに対して作成される。
- ㊵ キー値順に順次アクセスや、索引を使ってランダムにアクセスもできる。
- ㊶ ランダムにアクセス後、順次アクセスすることもできる。
- ㊦ ファイルは、索引域、基本域、あふれ域の 3 区域で構成される。
- ㊦ 基本域とあふれ域にレコードは記録される。
- ㊦ 索引域には、レコードの索引情報が記録される。

⑤ 直接編成

㊶ 直接編成とは

直接編成は特定のレコードを見つけるのに、そのレコードが記憶されているアドレスを用いて特定のレコードのみを処理する方式である。レコードアクセスを用いて特定のレコードだけを直接記録したり、取り出したりできるように編成する。レコードは指定したレコードアドレスに記録される。リアルタイム処理を必要とする業務に利用される。レコードの位置決めの方

法には、実アドレス指定法と相対アドレス指定法がある。

⑥ 実アドレス指定法

磁気ディスク上でレコードが実際に記録されているアドレスをそのまま指定するやり方である。アドレスとして、CCHHRを用いる。ファイル再編成時にレコードの記憶場所が移動したとき、それらの情報の変更が必要である。

⑦ 相対アドレス指定法

相対アドレスは、ファイル先頭の記憶場所から数えた相対的なトラック番号、レコード番号を用いて表したアドレスである。相対ブロック番号方式と相対トラック番号・相対レコード番号方式がある。

相対ブロック番号方式は、相対アドレスを相対ブロック番号で表す。相対レコード番号でアドレスを指定するやり方が該当する。

相対トラック番号と相対レコード番号方式は、相対トラック番号はファイルのために割り振られたエクステントの最初のトラック番号を00とし、続くトラック番号を01、02、…、と数えるやり方である。相対レコード番号は、それぞれのトラックの始めを起点として数えたレコード番号である。相対トラック番号と相対レコード番号でレコードアドレスを表す。

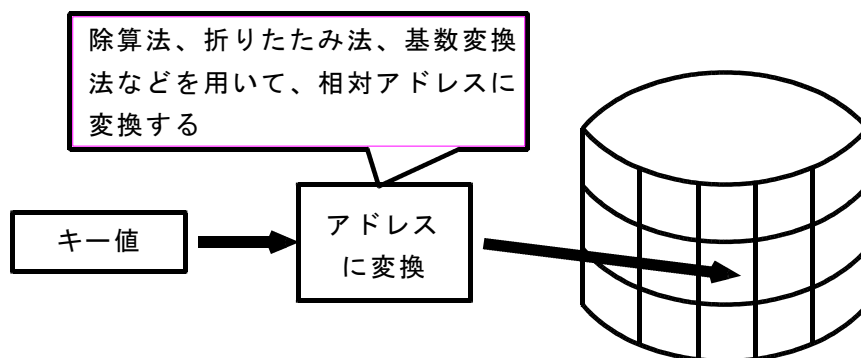
相対アドレス指定方式を用いた場合、ディスク上の実際のレコードをアクセスするときは、実アドレスに変換する必要がある。相対アドレスの実アドレスへの変換は、VTOCのそのファイルに対するエクステント情報の開始アドレスCCHHを用いて行う。

相対アドレスに変換する方法に除算法、折りたたみ法、基数変換法がある。

㊦ 除算法 : キー値をある除数で割り、その結果を相対アドレスとする。

㊧ 折りたたみ法 : キー値をいくつかの部分に分割し、それらを加え合わせ、その合計を相対アドレスにする。

㊨ 基数変換法 : キー値を10以外の基数で変換し、その結果を相対アドレスにする。



④ 直接編成の作成

直接アドレス方式で行う場合、入力レコードのキー値を昇順に並べてアドレスカウンタと比較し、一致すればキー値をアドレス変換する。不一致ならば、ダミーレコードを書き出す。間接アドレス方式で行う場合、最初にファイル全体にダミーレコードを書いて入力レコードをアドレス変換してファイルに書き込んでいく。レコードがすでに存在するときはシノニムレコードになる。シノニムレコードの処理は、シーケンシャル法またはチェーン法で行う。シーケンシャル法はホームアドレスの近くの空いているアドレスに書き出す方法である。チェーン法はシノニムレコードをポインタで繋いでいく方法である。

⑤ 更新・追加・削除

更新は順編成ファイルと同様に変動ファイルとマスタファイルの突合せによって行う。追加はファイル作成と同じ要領で行う。ダミーレコードを利用して追加レコードを入力レコードと同じように処理する。削除は順編成ファイルと同様に、削除文字をレコードの一部に書き込む。削除したレコードは、ダミーレコードと同様に取り扱う。

⑥ 直接編成法の特徴

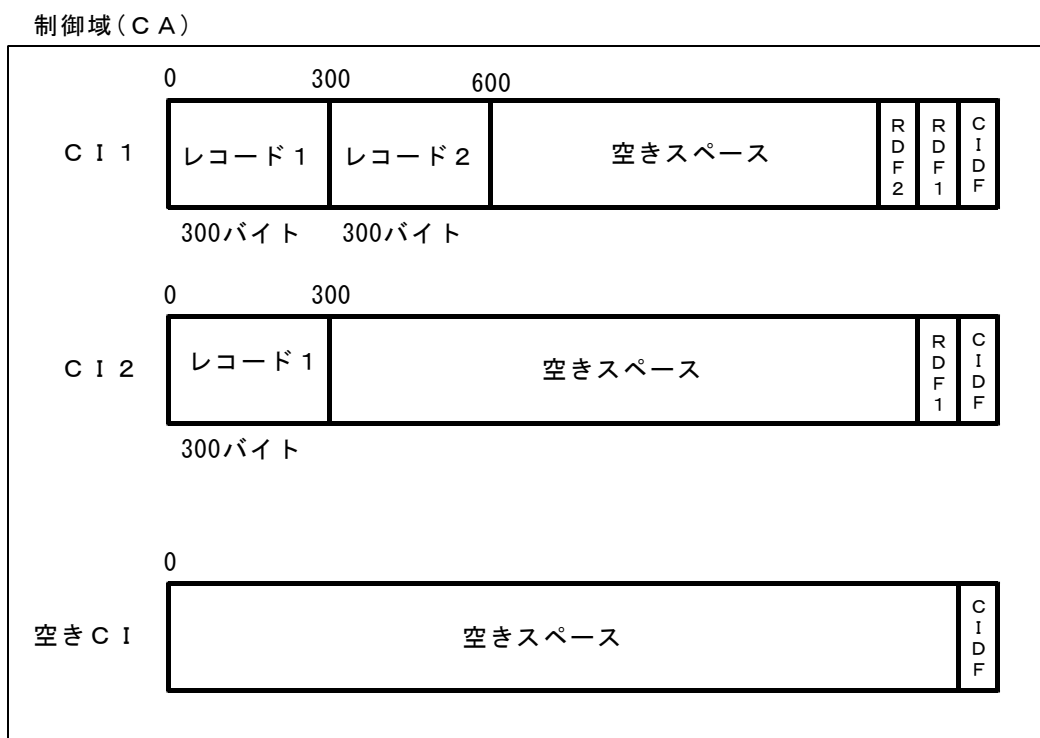
- ㊦ ランダムにレコードを検索する場合、使用効率が高い。
- ㊧ 順番にレコードを検索する場合、順編成に比較して使用効率が低い。
- ㊨ キーに基づくアドレスであるため、一般的に記憶領域の利用効率は高くない。
- ㊩ レコードの更新、追加、削除が比較的容易である。
- ㊪ 記憶媒体は、磁気ディスクやフロッピーディスクのDASDが利用される。
- ㊫ アクセス方式には、直接アドレス方式、間接アドレス方式がある。
- ㊬ 間接アドレス方式では、除算法、重ね合わせ法、基数変換法が使用される。
- ㊭ 異なるキーが同一のアドレスを指定する複数のレコードが発生するシノニム現象が生じる。
- ㊮ 編成は最初に領域全体にダミーレコードを記録した後、初期レコードを記録する。
- ㊯ 更新、追加、削除とも直接アクセスが可能である。削除はダミーレコードに置換する。

⑥ VSAMファイル

① VSAMファイルとは

仮想記憶オペレーティングシステムのもので利用できるファイル編成法とアクセス方式である。入力順データセット(ESDS)、キー順データセット(KSDS)、相対レコードデータセット(R

RDS)の3種類の編成法がある。ESDSは順編成ファイル、KSDSは索引編成ファイル、RRDSは直接編成ファイルに相当する。ファイルをデータセット、エクステントをデータスペースという。



⑥ データスペース

データスペースはVSAMデータセットのためにDASD上に割り当てられた記憶域であり、一つ又は複数のデータセットを收容する。ボリューム通し番号やスペースの大きさなどのデータスペースの情報はVSAMカタログに記憶される。データスペースの確保はアクセス方式サービス(AMS)というユーティリティプログラムを用いて行う。AMSは、クラスタの定義、データセットの複写、移動、削除といった機能を果たす。

⑦ VSAMデータセット

VSAMデータセットは、制御域(CA)や制御インタバル(CI)の論理的な構成要素の集合であり、この集合をクラスタという。データセットは、シリンダに相当するいくつかの制御域で構成される。また、一つの制御域はトラックに相当するいくつかの制御インタバルから構成され、論理レコードは制御インタバルのなかに記憶される。

制御域は、データセットを編成する上での論理的な区域で、レコードを記憶するために使用する。データセットの終わりの部分に空きCAを設ける。一つのデータセットは複数個のCAから構成され、空きCAは追加レコードが発生したとき、それを收容するのに使用する。空きCAがなくなると、新しいCAを設定する。

制御インタバルは、制御域を細分化したもので論理的領域である。制御インタバルは入出力操作の単位で、制御域の残りの部分に空き制御インタバルを設定し、レコードの追加の

ために使用する。レコードの追加が発生すると、制御域の中の空き制御インタバルに收容する。そこが使い尽くされると、空き制御域の制御インタバルを使用する。

㉔ 論理レコードの記憶形式

論理レコードは制御インタバルに記録される。論理レコードの形式は、固定長でも可変長でもよい。各レコードは制御インタバルの先頭から順番に記録される。同時に、レコード長などの情報を含んだレコード定義フィールド(RDF)が制御インタバルの右側に作成される。レコードとRDFの間があきスペースになる。空きスペースに関する情報が制御インタバルの最右端の制御インタバル定義フィールド(CIDF)に記憶される。CIDFには、空きスペースの長さや相対的位置などが記憶される。空きスペースは、追加や削除により、長さや相対位置が変わる。それに応じて、CIDFは、VSAMによって自動的に更新される。

㉕ 相対バイトアドレス

VSAMデータセットでは、各レコードの記憶アドレスを相対バイトアドレス(RBA)で表現する。RBAは、データセットの最初の制御域の最初の制御インタバルの先頭バイトを0として、そこから数えたバイト数である。レコードの記憶アドレスにRBAを用いることで、ハードウェアからの独立性が実現できる。DASDの種類に無関係に、どんなDASD上にあっても、また、DASDのどこにあってもRBAそのものは変わらない。これによってデータセットの保守作業や移動が容易になる。

例題演習

区分編成ファイルに関して、正しい記述はどれか。

- ア シノニムレコードの対策を考慮しなければならない。
- イ 直接アドレス指定方式と間接アドレス指定方式がある。
- ウ 登録簿とメンバに分かれ、登録簿によってメンバを直接アクセスすることができる。
- エ ボリューム管理情報、データセットに対するアクセスの履歴情報、データセットの属性、機密保護情報などを管理するための登録簿がある。

解答解説

区分編成ファイルの特徴に関する問題である。

アのシノニム現象が発生するのは直接編成ファイルである。シノニムとは、異なる2つのレコードが計算結果により同一のアドレスになることである。

イの直接アドレス指定方式、間接アドレス指定方式は直接編成ファイルの特徴である。直接アドレス指定方式はレコードのキー項目を直接記憶媒体のアドレスにすることであり、間接アドレス指定方式は主キーに、ある計算を実行し、その計算結果をアドレスとする方式である。

ウの登録簿でメンバに直接アクセスするのは区分編成ファイルで、求める答えはウとなる。

エのボリューム管理情報、アクセス履歴情報、データセットの属性、機密保護情報の管理の

ための登録簿はV S A Mファイルである。

例題演習

シノニムレコードの発生する可能性があるファイルアクセスはどれか。

- ア 区分編成ファイルへのレコードの追加
- イ 索引順編成ファイルのレコードの更新
- ウ 順編成ファイルのレコードの更新
- エ 直接編成ファイルへのレコードの追加

解答解説

シノニムレコードに関する問題である。

ハッシュ法は、データの値から直接、位置を計算する方法である。データの参照を高速に行う。キー値 X を関数 $H(X)$ を用いて、配列の添字やアドレスに変換する。関数 $H(X)$ をハッシュ関数といい、ハッシュ関数が返す値をハッシュ値という。ハッシュ関数を用いて、データを格納する添字やアドレスを求めることをハッシングという。

異なったキー値から同じハッシュ値が求まることを衝突という。先に格納されていたデータをホームデータ、後のデータをシノニムデータという。衝突が発生した場合、この衝突を回避する方法に、チェーン法とオープンアドレス法がある。

直接編成は、特定のレコードを見つけるのに、そのレコードが記憶されているアドレスを用いて特定のレコードのみを処理する方式である。レコードアクセスを用いて特定のレコードだけを直接記録したり、取り出したりできるように編成する。レコードの格納アドレスを求める際に衝突が発生する。シノニム現象が発生する。

シノニム現象が発生するのは、直接編成ファイルにレコードを追加する場合である。求める答えはエとなる。

例題演習

ファイル編成法の一つである索引編成ファイルの説明として、正しいものはどれか。

- ア 各レコードのもっているアドレスによって、レコードに直接アクセスできる。媒体の利用効率が悪くなることもある。
- イ データを記録する領域と、特定のレコードを直接アクセスするための情報を記録する領域で構成する。媒体の利用効率がよく、大規模なファイルを作ることができる。
- ウ 物理的に書き込んだ順番にレコードを記録する。順次アクセスだけができる。
- エ メンバと呼ぶデータ領域と、メンバの情報を管理する登録簿域で構成する。プログラムの格納に適している。

解答解説

索引編成ファイルは、索引キーによる直接アクセスができるファイル編成法である。

索引順次編成ファイルの特徴

- ① レコードはキー値順に記録される。
- ② 索引はトラック毎のレコードに対して作成される。
- ③ キー値順に順次アクセスすることもできるし、索引を使ってランダムにアクセスすることもできる。
- ④ ランダムにアクセス後、順次アクセスすることもできる。
- ⑤ ファイルは、索引域、基本域、あふれ域の3区域で構成される。
- ⑥ 基本域とあふれ域にレコードは記録される。
- ⑦ 索引域には、レコードの索引情報が記録される。

イのデータ域と索引域で構成されるのは索引編成ファイルである。求める答えはイである。
アは直接編成法、ウは順編成ファイル、エは区分編成ファイルである。

例題演習

ファイルをメンバと呼ぶ複数の単位に分割する区分編成ファイルに関する記述のうち、正しいものはどれか。

- ア 記憶媒体として磁気ディスクおよび磁気テープを使用できる。
- イ 区分編成ファイルをプログラム格納用ファイルに使うのは、メンバの容量を超えることが多いので適していない。
- ウ メンバ内のレコードはランダムにアクセスされる。
- エ メンバの格納位置を登録したディレクトリをもつ。

解答解説

区分編成ファイルの特徴に関する問題である。

区分編成ファイルの特徴は、ディレクトリはランダムアクセスが可能、メンバは順アクセスであり、ライブラリ管理に有効である。

アの磁気テープは不可である。

イの区分編成ファイルはプログラム格納ファイルとして利用する。

ウは、メンバは順アクセスのみで、ランダムアクセスはできない。

エのディレクトリの説明は区分編成ファイルに関するものである。求める答えはエである。