

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6744179号
(P6744179)

(45) 発行日 令和2年8月19日 (2020.8.19)

(24) 登録日 令和2年8月3日 (2020.8.3)

(51) Int.Cl. F I
G O 6 F 16/22 (2019.01) G O 6 F 16/22
G O 6 F 16/25 (2019.01) G O 6 F 16/25

請求項の数 7 (全 17 頁)

| | | | |
|-----------|------------------------------|-----------|------------------------|
| (21) 出願番号 | 特願2016-179579 (P2016-179579) | (73) 特許権者 | 513269000 |
| (22) 出願日 | 平成28年9月14日 (2016.9.14) | | 株式会社エスペラントシステム |
| (65) 公開番号 | 特開2018-45441 (P2018-45441A) | | 千葉県流山市前平井61番地 E S S ビル |
| (43) 公開日 | 平成30年3月22日 (2018.3.22) | (74) 代理人 | 110000888 |
| 審査請求日 | 令和1年9月11日 (2019.9.11) | | 特許業務法人 山王坂特許事務所 |
| | | (72) 発明者 | 古庄 晋二 |
| | | | 神奈川県横浜市中区元浜町3丁目21番地 |
| | | | 2 株式会社ターボデータラボラトリー内 |
| | | 審査官 | 西村 直史 |

最終頁に続く

(54) 【発明の名称】 データ統合方法、データ統合装置、データ処理システム及びコンピュータプログラム

(57) 【特許請求の範囲】

【請求項1】

1 または複数のフィールドよりなる、各々レコード番号が付与されたレコードを配列したテーブルを表すインデックスデータを、統合データに統合するデータ統合方法であって、

前記インデックスデータは、前記各フィールドの各々に対応して設けられたインデックスを含み、

前記各インデックスは、

当該インデックスが含まれるインデックスデータが表す前記テーブルのレコードの当該インデックスに対応するフィールドの値として用いられている値を各エントリに登録した値リストと、

各エントリに、当該インデックスが含まれるインデックスデータが表す前記テーブルの、当該エントリの順位と同じ値のレコード番号のレコードの、当該インデックスに対応するフィールドの値が登録されている、当該インデックスの前記値リストのエントリの順位を登録した対応値リストとを含み、

前記統合データは、前記各フィールドの各々に対応して設けられた統合値リストと、前記各フィールドの各々に対応して設けられた対応統合値リストとを含み、

当該データ統合方法は、

各フィールドについて、統合データに統合するインデックスデータの当該フィールドに対応するインデックスの値リストに含まれる値のうちの、当該フィールドに対応する前記

10

20

統合値リストに登録されていない値を当該統合値リストに追加登録するステップと、

各フィールドについて、統合データに統合するインデックスデータの当該フィールドに対応するインデックスの対応値リストの各エントリについて求めた対応統合値を、当該エントリの当該対応値リスト中の順位に従った順番で並べた配列を、前記統合データの当該フィールドに対応する対応統合値リストに追加登録、もしくは、当該配列を当該フィールドに対応する新たな対応統合値リストとして前記統合データに追加するステップとを有し、

前記対応値リストのエントリについて求めた前記対応統合値は、当該対応値リストのエントリに登録されている順位と同じ順位の、当該対応値リストが属するインデックスの前記値リストに登録されている値と同じ値が登録される、当該対応値リストが属する前記インデックスに対応するフィールドに対応する前記統合値リストのエントリの順位に等しいことを特徴とするデータ統合方法。

10

【請求項2】

1または複数のフィールドよりなる、各々レコード番号が付与されたレコードを配列したテーブルを、統合データに統合するデータ統合方法であって、

前記統合データは、前記各フィールドの各々に対応して設けられた対応統合値リストと、前記各フィールドの各々に対応して設けられた統合値リストとを有し、

当該データ統合方法は、

統合データに統合するテーブルの各フィールドについて、当該テーブルの各レコードの当該フィールドの値として用いられている値のうちの、当該フィールドに対応する前記統合値リストに登録されていない値を当該統合値リストに追加登録するステップと、

20

統合データに統合するテーブルの各フィールドについて、当該テーブルの各レコードの当該フィールドの値と同じ値が登録される、当該フィールドに対応する前記統合値リストのエントリの順位を、前記テーブルにおけるレコードの順番に従った順番で並べた配列を、前記統合データの当該フィールドに対応する対応統合値リストに追加登録、もしくは、当該配列を当該フィールドに対応する新たな対応統合値リストとして前記統合データに追加するステップとを有することを特徴とするデータ統合方法。

【請求項3】

1または複数のフィールドよりなる、各々レコード番号が付与されたレコードを配列したテーブルを表すインデックスデータを、統合データに統合するデータ統合装置であって、

30

前記インデックスデータは、前記各フィールドの各々に対応して設けられたインデックスを含み、

前記各インデックスは、

当該インデックスが含まれるインデックスデータが表す前記テーブルのレコードの当該インデックスに対応するフィールドの値として用いられている値を各エントリに登録した値リストと、

各エントリに、当該インデックスが含まれるインデックスデータが表す前記テーブルの、当該エントリの順位と同じ値のレコード番号のレコードの、当該インデックスに対応するフィールドの値が登録されている、当該インデックスの前記値リストのエントリの順位を登録した対応値リストとを含み、

40

前記統合データは、前記各フィールドの各々に対応して設けられた統合値リストと、前記各フィールドの各々に対応して設けられた対応統合値リストとを含み、

当該データ統合装置は、

各フィールドについて、統合データに統合するインデックスデータの当該フィールドに対応するインデックスの値リストに含まれる値のうちの、当該フィールドに対応する前記統合値リストに登録されていない値を当該統合値リストに追加登録する値統合手段と、

各フィールドについて、統合データに統合するインデックスデータの当該フィールドに対応するインデックスの対応値リストの各エントリについて求めた対応統合値を、当該エントリの当該対応値リスト中の順位に従った順番で並べた配列を、前記統合データの当該

50

フィールドに対応する対応統合値リストに追加登録、もしくは、当該配列を当該フィールドに対応する新たな対応統合値リストとして前記統合データに追加する対応値登録手段とを有し、

前記対応値リストのエントリについて求めた前記対応統合値は、当該対応値リストのエントリに登録されている順位と同じ順位の、当該対応値リストが属するインデックスの前記値リストに登録されている値と同じ値が登録される、当該対応値リストが属する前記インデックスに対応するフィールドに対応する前記統合値リストのエントリの順位に等しいことを特徴とするデータ統合装置。

【請求項 4】

1 または複数のフィールドよりなる、各々レコード番号が付与されたレコードを配列したテーブルを、統合データに統合するデータ統合装置であって、

前記統合データは、前記各フィールドの各々に対応して設けられた対応統合値リストと、前記各フィールドの各々に対応して設けられた統合値リストとを有し、

当該データ統合装置は、

統合データに統合するテーブルの各フィールドについて、当該テーブルの各レコードの当該フィールドの値として用いられている値のうちの、当該フィールドに対応する前記統合値リストに登録されていない値を当該統合値リストに追加登録する値統合手段と、

統合データに統合するテーブルの各フィールドについて、当該テーブルの各レコードの当該フィールドの値と同じ値が登録される、当該フィールドに対応する前記統合値リストのエントリの順位を、前記テーブルにおけるレコードの順番に従った順番で並べた配列を、前記統合データの当該フィールドに対応する対応統合値リストに追加登録、もしくは、当該配列を当該フィールドに対応する新たな対応統合値リストとして前記統合データに追加する対応値登録手段とを有することを特徴とするデータ統合装置。

【請求項 5】

請求項 3 または 4 記載のデータ統合装置と、前記統合データを保存したストレージとを備えたデータ処理システムであって、

前記統合データが、各レコード番号の各フィールドに、当該フィールドに対応する前記対応統合値リストの当該レコード番号と同じ値の順位のエントリに登録されている順位と同じ順位の、当該フィールドに対応する前記統合値リストのエントリに登録されている値を有するレコードを含むテーブルを表しているものとして、前記ストレージに保存された前記統合データが表すテーブルを操作するテーブル操作手段を有することを特徴とするデータ処理システム。

【請求項 6】

コンピュータによって読み取られ実行されるコンピュータプログラムであって、

前記コンピュータを請求項 3 または 4 記載のデータ統合装置として機能させることを特徴とするコンピュータプログラム。

【請求項 7】

コンピュータによって読み取られ実行されるコンピュータプログラムであって、

前記コンピュータを、請求項 5 記載のデータ処理システムとして機能させることを特徴とするコンピュータプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、主としてデータベースのデータを統合する技術に関するものである。

【背景技術】

【0002】

データベースのデータ構造としては、図 8 a 1 に示すような RDB (Relational Database) のテーブルを表す、図 8 a 2 に示すようなインデックスセットのデータ構造が知られている (たとえば、特許文献 1)。

ここで、図 8 a 1 のテーブル (テーブル A) は、複数のフィールド (図では、「性別」

10

20

30

40

50

と「タイプ」)を持つレコードを行とする表であり、各レコードにはテーブル中の当該レコードの順番を示すレコード番号が付与されている。なお、レコード番号は0から始まる。

また、図8a2に示すインデックスセット(テーブルAインデックスセット)は、図8a1のテーブルのレコードのフィールド毎に設けられたインデックスより構成される。図8a1のテーブルでは、テーブルのレコードのフィールドは、「性別」と「タイプ」であるので、「性別」のインデックスと、「タイプ」のインデックスが図8a2のインデックスセットに含まれる。

【0003】

そして、図8a2に示すように各インデックスは、VNo、VLを含む。

10

VLは、対応するテーブルの対応するフィールドの値として用いられている値を所定の基準(たとえば、値の昇順)でソートして各エントリに登録したリストである。

すなわち、たとえば、図8a1に示すテーブルのテーブルデータセットのインデックスセットの「性別」のインデックスであれば、テーブルの「性別」のフィールドに登録されているのはFとMのみであるので、VLは、Fに登録したエントリとMに登録したエントリよりなるリストとなる。

【0004】

次に、VNoは、対応するテーブルのレコード数と同数のエントリよりなるリストであり、VNoの順位nのエントリには、対応するテーブルのレコード番号nのレコードの対応するフィールドの値が登録されているVLのエントリのVL中の順位を示す値が登録される。なお

20

【0005】

すなわち、たとえば、図8a1に示すテーブルのテーブルデータセットのインデックスセットの「性別」のインデックスであれば、テーブルのレコード番号2のレコードの「性別」のフィールドの値はFであり、Fが登録されているVLのエントリの順位は0であるので、VNoの順位2のエントリには1が登録される。

【0006】

ここで、このようなインデックスによれば、VNoのエントリ数より、対応するテーブルのレコード数を速やかに求めることができ、VNoとVLより、各レコード番号のレコードの対応するフィールドの値を速やかに求めることができる。

30

すなわち、たとえば、「性別」のインデックスの、レコード番号2に対応するVNoの順位0のエントリからVLの順位0が求まり、このVLの順位0のエントリにはFが登録されているので、レコード番号2の「性別」のフィールドの値はFとして求まる。

よって、このようなレコードのフィールド毎に設けられたインデックスより構成されるインデックスセットによってテーブルを完全に表現できると共に、当該インデックスセットを用いてテーブルを速やかに利用できるようになる。

【先行技術文献】

【特許文献】

【0007】

【特許文献1】特開2000-339390号公報

40

【発明の概要】

【発明が解決しようとする課題】

【0008】

さて、テーブルを各々が表す複数のインデックスセットを保存する場合がある。たとえば、期間毎にテーブルを作成し、作成した各期間のテーブルをログとして保存センターに蓄積していく場合や、地域毎にテーブルを作成し、作成された各地域のテーブルを収集して保存する場合などがある。

【0009】

しかし、このときにテーブルのインデックスセットを、そのまま保存すると以下のような問題が生じる。

50

すなわち、保存するインデックスセットの増加に伴い、保存するインデックスセットのデータ量の和に比例して保存に必要となるストレージのデータ容量も増加してしまう。

また、保存するインデックスセットの増加に伴い、保存した複数のインデックスセットが表す複数のテーブルを統合的に利用することが容易でなくなる。

すなわち、図 8 a 1 に示すテーブル A のインデックスセットである図 8 a 2 に示すテーブル A インデックスセットと、図 8 b 1 に示すテーブル B のインデックスセットである図 8 b 2 に示すテーブル B インデックスセットと、図 8 c 1 に示すテーブル C のインデックスセットである図 8 c 2 に示すテーブル C インデックスセットととの 3 つのインデックスセットをそのまま保存した場合、テーブル A、B、C を統合的に利用することが容易でなくなる。

10

【 0 0 1 0 】

たとえば、テーブル A、B、C のフィールドタイプ「性別」が F であるレコード数の総数を求めるためには、テーブル A インデックスセットを用いてテーブル A 中のフィールドタイプ「性別」が F であるレコード数を求め、テーブル B インデックスセットを用いてテーブル B 中のフィールドタイプ「性別」が F であるレコード数を求め、テーブル C インデックスセットを用いてテーブル C 中のフィールドタイプ「性別」が F であるレコード数を求め、各テーブル A、B、C について求めたレコード数の総和を算出するという比較的煩雑で負荷の大きい処理を行う必要が生じる。なお、このような処理の負荷は、保存されているインデックスセットの数の増加に伴い、逐次的に増大していくこととなる。

【 0 0 1 1 】

20

一方で、複数のインデックスセットを保存する代わりに、各インデックスセットが表すテーブルをまとめた一つのテーブルを表す一つのインデックスセットである統合インデックスセットを保存するようにすれば上述のような問題は生じないが、このようにする場合には、統合インデックスセットにまとめるべきインデックスセットが新たに発生する度に、統合インデックスセットの全体を新たに発生したインデックスセットに応じて編集し書き換える処理を行う必要が生じる。そして、この処理の負荷が、保存している統合インデックスセットの規模が大きくなると過剰な負荷となってしまいうという問題が生じることとなる。

【 0 0 1 2 】

そこで、本発明は、比較的簡易な処理によって、複数のテーブルのデータを、データ量を削減しつつ、各テーブルの統合的な利用が容易な形態で保存することを課題とする。

30

【課題を解決するための手段】

【 0 0 1 3 】

前記課題達成のために、本発明は、1 または複数のフィールドよりなる、各々レコード番号が付与されたレコードを配列したテーブルを表すインデックスデータを、統合データに統合するデータ統合方法を提供する。ここで、前記インデックスデータは、前記各フィールドの各々に対応して設けられたインデックスを含んでいる。また、前記各インデックスは、当該インデックスが含まれるインデックスデータが表す前記テーブルのレコードの当該インデックスに対応するフィールドの値として用いられている値を各エントリに登録した値リストと、各エントリに、当該インデックスが含まれるインデックスデータが表す前記テーブルの、当該エントリの順位と同じ値のレコード番号のレコードの、当該インデックスに対応するフィールドの値が登録されている、当該インデックスの前記値リストのエントリの順位を登録した対応値リストとを含んでいる。また、前記統合データは、前記各フィールドの各々に対応して設けられた統合値リストと、前記各フィールドの各々に対応して設けられた対応統合値リストとを含んでいる。

40

【 0 0 1 4 】

そして、当該データ統合方法は、各フィールドについて、統合データに統合するインデックスデータの当該フィールドに対応するインデックスの値リストに含まれる値のうちの、当該フィールドに対応する前記統合値リストに登録されていない値を当該統合値リストに追加登録するステップと、各フィールドについて、統合データに統合するインデックス

50

データの当該フィールドに対応するインデックスの対応値リストの各エントリについて求めた対応統合値を、当該エントリの当該対応値リスト中の順位に従った順番で並べた配列を、前記統合データの当該フィールドに対応する対応統合値リストに追加登録、もしくは、当該配列を当該フィールドに対応する新たな対応統合値リストとして前記統合データに追加するステップとを有する。ただし、前記対応値リストのエントリについて求めた前記対応統合値は、当該対応値リストのエントリに登録されている順位と同じ順位の、当該対応値リストが属するインデックスの前記値リストに登録されている値と同じ値が登録される、当該対応値リストが属する前記インデックスに対応するフィールドに対応する前記統合値リストのエントリの順位に等しいものである。

【0015】

10

また、本発明は、前記課題達成のために、このようなデータ統合方法を実行するデータ統合装置も提供する。

このようなデータ統合方法やデータ統合装置によれば、統合データに統合する各インデックスデータ間で、同じフィールドのインデックスの値リストのエントリに登録される値に重複が存在する場合でも、統合データの各統合インデックスの統合値リストの各エントリに登録される各値は、当該統合値リスト内においてユニークな値となる。したがって、各インデックスデータをそのまま保存する場合に比べ、保存に必要となるストレージのデータ容量を削減することができる。

【0016】

また、このような各データ統合方法やデータ統合装置によれば、統合データの既存の部分には一切変更を加えない比較的簡易な処理によって、新たなインデックスデータを統合データに統合することができるようになる。

20

また、統合データは、統合データに統合された各インデックスセットに対応するテーブルをつなげたテーブルを表すこととなるので、統合データを用いることにより、容易に当該各インデックスデータに対応するテーブルの統合的な操作や利用が行えるようになる。

また、本発明は、前記課題達成のために、1または複数のフィールドよりなる、各々レコード番号が付与されたレコードを配列したテーブルを、統合データに統合するデータ統合方法を提供する。ここで、前記統合データは、前記各フィールドの各々に対応して設けられた対応統合値リストと、前記各フィールドの各々に対応して設けられた統合値リストとを備えている。そして、当該データ統合方法は、統合データに統合するテーブルの各フィールドについて、当該テーブルの各レコードの当該フィールドの値として用いられている値のうちの、当該フィールドに対応する前記統合値リストに登録されていない値を当該統合値リストに追加登録するステップと、統合データに統合するテーブルの各フィールドについて、当該テーブルの各レコードの当該フィールドの値と同じ値が登録される、当該フィールドに対応する前記統合値リストのエントリの順位を、前記テーブルにおけるレコードの順番に従った順番で並べた配列を、前記統合データの当該フィールドに対応する対応統合値リストに追加登録、もしくは、当該配列を当該フィールドに対応する新たな対応統合値リストとして前記統合データに追加するステップとを有している。

30

【0017】

また、本発明は、前記課題達成のために、このようなデータ統合方法を実行するデータ統合装置も提供する。

40

このようなデータ統合方法やデータ統合装置によれば、統合データに統合する各テーブル間やレコード間で、同じフィールドに登録される値に重複が存在する場合でも、統合データの各統合インデックスの統合値リストの各エントリに登録される各値は、当該統合値リスト内においてユニークな値となる。したがって、各テーブルをそのまま保存する場合に比べ、保存に必要となるストレージのデータ容量を削減することができる。

【0018】

また、このようなデータ統合方法やデータ統合装置によれば、統合データの既存の部分には一切変更を加えない比較的簡易な処理によって、新たなテーブルを統合データに統合することができるようになる。

50

また、統合データは、統合データに統合された各テーブルをつなげたテーブルを表すこととなるので、統合データを用いることにより、容易に当該各テーブルの統合的な操作や利用が行えるようになる。

さて、本発明は、併せて、以上のデータ統合装置と、前記統合データを保存したストレージとを備えたデータ処理システムであって、前記統合データが、各レコード番号の各フィールドに、当該フィールドに対応する前記対応統合値リストの当該レコード番号と同じ値の順位のエントリに登録されている順位と同じ順位の、当該フィールドに対応する前記統合値リストのエントリに登録されている値を有するレコードを含むテーブルを表しているものとして、前記ストレージに保存された前記統合データが表すテーブルを操作するテーブル操作手段を備えたデータ処理システムも提供する。

10

【0019】

また、本発明は、コンピュータを以上のデータ統合装置やデータ処理システムとして機能させるコンピュータプログラムも提供する。

【発明の効果】

【0020】

以上のように、本発明によれば、比較的簡易な処理によって、複数のテーブルのデータを、データ量を削減しつつ、各テーブルの統合的な利用が容易な形態で保存することができる。

【図面の簡単な説明】

【0021】

20

【図1】本発明の実施形態に係るデータ処理システムの構成を示すブロック図である。

【図2】本発明の実施形態に係るデータ統合処理の概要を示す図である。

【図3】本発明の実施形態に係るデータ統合処理の概要を示す図である。

【図4】本発明の実施形態に係る統合インデックスの生成例を示す図である。

【図5】本発明の実施形態に係る統合インデックスの生成例を示す図である。

【図6】本発明の実施形態に係る統合インデックスの保存例を示す図である。

【図7】本発明の実施形態に係る統合インデックスの他の生成例を示す図である。

【図8】従来のテーブルのデータ構造を示す図である。

【発明を実施するための形態】

【0022】

30

以下、本発明の実施形態について説明する。

図1に、本実施形態に係るデータ処理装置の構成を示す。

図示するように、データ処理装置は、ストレージ1とプロセッサ2と入力装置3と表示装置4等を備えている。また、プロセッサ2はデータ統合部11と、RDBMS12（リレーショナルデータベースマネジメントシステム12）とを備えている。なお、データ統合部11、RDBMS12は、プロセッサ2が所定のコンピュータプログラムを実行することにより実現される機能部である。

【0023】

このような構成において、新たに保存すべきインデックスセットがストレージ1に格納されると、プロセッサ2のデータ統合部11は、格納されたインデックスセットを、ストレージ1に格納している統合インデックスセットに統合する。

40

なお、ストレージ1に格納される新たに保存すべきインデックスセットは、先に図8を用いて説明した先にVNo、VLとを含んだインデックスよりなるインデックスセットであり、当該インデックスセットは、図8a2のインデックスセットが図8a1のテーブルを表しているのと同様に、一つのテーブルを表している。

【0024】

一方、RDBMS12は、統合インデックスセットを用いて、それまで保存されたインデックスセットが表す各テーブルに対する統合的な操作を行う。

次に、データ統合部11が行う、インデックスセットの統合インデックスセットへの統合動作の概要について説明する。

50

いま、保存すべき最初のインデックスセットとして、図 2 a 1 に示すテーブル A を表すインデックスセットである図 2 a 2 のテーブル A インデックスセットが格納された場合、データ統合部 1 1 は、図 2 a 2 のテーブル A インデックスセットを、図 2 a 3 に示すように、そのまま統合インデックスセットとする。

【 0 0 2 5 】

したがって、統合インデックスセットは、インデックスセットと同様に、フィールド毎のインデックスを備え、各インデックスはVNo、VLを含むものとなり、統合インデックスセットは、テーブルを表すものとなる。

さて、ここで、以下では、インデックスセットのインデックスを単に「インデックス」と表記し、統合インデックスのインデックスを「統合インデックス」と表記することにより両者を区別する。また、以下では、インデックスセットによって表されるテーブルを単に「テーブル」と表記し、統合インデックスセットによって表されるテーブルを「統合テーブル」と表記することにより、両者を区別する。

【 0 0 2 6 】

さて、以上のようにして最初に生成された、図 2 a 3 の統合インデックスセットは、図 2 a 4 のような、テーブル A と同じ内容を持つ統合テーブルを表すものとなる。

次に、図 2 a 3 の統合インデックスセットが保存されている状態で、次に保存すべきインデックスセットとして、図 2 b 1 に示すテーブル B を表すインデックスセットである図 2 b 2 のテーブル B インデックスセットが格納された場合、データ統合部 1 1 は、テーブル B インデックスセットの各インデックスについて以下の処理を行う。

【 0 0 2 7 】

すなわち、テーブル B インデックスセットの各フィールドに対応するインデックスについて、インデックスのVLに含まれる値のうちの、当該フィールドに対応する統合インデックスのVLに含まれない値を、図 2 b 3 に示すように、当該フィールドに対応する統合インデックスのVLに追加登録する。

【 0 0 2 8 】

そして、インデックスのVNoのエントリ数が等しいリストであるT_VNoを作成し、T_VNoの各エントリに、当該T_VNoのエントリT_VNoE(i)と順位が等しい、当該インデックスのVNoのエントリVNoE(i)に格納されている値kを順位とする当該インデックスのVLのエントリVLE(k)に登録されている値jが登録されている、当該フィールドの統合インデックスのVLのエントリの順位を示す値を登録する。そしてT_VNoの各エントリを、図 2 b 3 に示すように、当該フィールドの統合インデックスのVNoに追加登録する。

【 0 0 2 9 】

ただし、T_VNoE(i)はT_VNoのi番目のエントリを、VNoE(i)はVNoのi番目のエントリを、VLE(k)は、VLのk番目のエントリを表している。

具体的には、たとえば、フィールド「タイプ」のインデックスであれば、インデックスのVNoのエントリ数は4であるので、4つのエントリを備えたT_VNoを作成する。

そして、たとえば、T_VNoの順位0のエントリについては、当該インデックスのVNoの順位0のエントリに登録されている値は1であり、当該インデックスのVLの順位1のエントリに登録されている値はBであり、Bはフィールド「タイプ」の統合インデックスのVLの順位3のエントリに登録されているので、T_VNoの順位0のエントリには3を登録する。

【 0 0 3 0 】

同様に、たとえば、T_VNoの順位2のエントリについては、当該インデックスのVNoの順位2のエントリに登録されている値は0であり、当該インデックスのVLの順位0のエントリに登録されている値はAであり、Aはフィールド「タイプ」の統合インデックスのVLの順位0のエントリに登録されているので、T_VNoの順位2のエントリには0を登録する。

【 0 0 3 1 】

そして、同様にしてT_VNoの各エントリに値を登録したならば、T_VNoの各エントリを、フィールド「タイプ」の統合インデックスのVNoに追加する。

さて、このようにして作成された図 2 b 3 に示す統合インデックスセットが表す統合テ

10

20

30

40

50

ーブルは、図 2 b 4 に示すように、テーブル A とテーブル B を当該順序で連結したものとなる。

次に、以上のようにして、テーブル B インデックスセットを統合インデックスセットに統合し、図 3 a 2 に示す統合テーブルを表す、図 3 a 1 の統合インデックスセットが保存されている状態で、次に保存すべきインデックスセットとして、図 3 b 1 に示すテーブル C を表すインデックスセットである図 3 b 2 のテーブル C インデックスセットが格納された場合、データ統合部 11 は、以上の同様に、テーブル C インデックスセットの各インデックスについて以下の処理を行う。

【 0 0 3 2 】

すなわち、テーブル C インデックスセットの各フィールドに対応するインデックスについて、インデックスの VL に含まれる値のうちの、当該フィールドの統合インデックスの VL に含まれない値を、図 3 b 3 に示すように、当該フィールドの統合インデックスの VL に追加登録する。

【 0 0 3 3 】

そして、インデックスの VNo のエン트리数が等しいリストである T_VNo を作成し、T_VNo の各エントリに、当該 T_VNo のエン트리 T_VNoE(i) と順位が等しい、当該インデックスの VNo のエン트리 VNoE(i) に格納されている値 k を順位とする当該インデックスの VL のエン트리 VLE(k) に登録されている値 j が登録されている、当該フィールドの統合インデックスの VL のエントリの順位を示す値を登録する。そして T_VNo の各エントリを、図 3 b 3 に示すように、当該フィールドの統合インデックスの VNo に追加登録する。

【 0 0 3 4 】

結果、このようにして作成された図 3 b 3 に示す統合インデックスセットが表す統合テーブルは、図 3 b 4 に示すように、テーブル A とテーブル B とテーブル C を当該順序で連結したものとなる。

以下、次に保存すべきインデックスセットが発生するたびに、同様の処理を行って、当該インデックスセットを統合インデックスセットに統合していく。

【 0 0 3 5 】

なお、統合インデックスセットには、統合インデックスの VNo のエントリのどの範囲が、どのインデックス、または、どのテーブルを表すインデックスから作成されたものであるのかを示すデータを対応データとして含めておくようにしてもよくこのようにすることにより、テーブルやインデックス単位での統合インデックスセットの多様な利用が容易に行えるようになる。

【 0 0 3 6 】

ここで、以上のように、本実施形態では、新たなインデックスセットを統合インデックスセットに統合する際に、統合インデックスセットの既に登録されている部分には一切の変更が加えられず、新たなインデックスセットの内容に応じた内容が統合インデックスセットに追加的に登録されていくことにより、インデックスセットの統合インデックスセットへの統合が実現されている。

【 0 0 3 7 】

また、統合インデックスセットに統合する各インデックスセット間で、同じフィールドのインデックスの VL のエントリに登録される値に重複が存在する場合でも、統合インデックスセットの各統合インデックスの VL の各エントリに登録される各値は、統合インデックスの VL 内においてユニークな値となる。したがって、各インデックスセットをそのまま保存する場合に比べ、保存に必要となるストレージ 1 のデータ容量を削減することができる。

【 0 0 3 8 】

さて、次に、以上のようにインデックスセットを統合インデックスセットに統合していく際の詳細な手順について説明する。

上述のようにインデックスセットの統合インデックスセットへの統合は、各フィールドのインデックスについて、当該インデックスの内容に応じて統合インデックスを編集する

10

20

30

40

50

ことにより実現される。

そして、このインデックスの内容に応じて統合インデックスの編集は、たとえば、以下に示す手順によって実現することができる。

ここでは、図 4 a に示すインデックスの内容に応じて、同図に示す統合インデックスを編集する場合を例にとり説明する。

この場合、まず、登録値調査処理を行う。

この登録値調査処理では、まず、ポインタ P を 0 に、カレントナンバ C N を統合インデックスの VL の末尾のエントリの次の順位である 3 に設定する。そして、インデックスの VL とエントリ数が等しい T_VL と、追加値リスト adVL を作成する。

そして、インデックスの VL のポインタ P が示す順位 0 のエントリに登録されている値 A を取得し、取得した値 A が統合インデックスの VL に格納されているかどうかを調べる。ここでは、統合インデックスの VL に格納されているので、値 A が格納されている統合インデックスの VL のエントリの順位 0 を取得し、取得した順位を T_VL の、ポインタ P が示す順位 0 のエントリに登録する。

【 0 0 3 9 】

そして、ポインタ P の値が、インデックスの VL の末尾のエントリの順位と一致しているかどうかを調べ、ここでは一致していないのでポインタ P の値を一つ進め 1 とする。

次に、ポインタ P の値を一つ進めたならば、図 4 b に示すように、インデックスの VL のポインタ P が示す順位 1 のエントリに登録されている値 B を取得し、取得した値 B が統合インデックスの VL に格納されているかどうかを調べる。ここでは、値 B は、統合インデックスの VL に格納されていないので、取得した値 B を追加値リスト adVL に追加登録する。また、カレントナンバ C N の値 3 を、T_VL の、ポインタ P が示す順位 1 のエントリに登録すると共に、カレントナンバ C N の値を一つ進め 4 とする。なお、カレントナンバ C N の値は、カレントナンバ C N の値を T_VL のエントリに登録したときに、一つ進める。

【 0 0 4 0 】

そして、そして、ポインタ P の値が、インデックスの VL の末尾のエントリの順位と一致しているかどうかを調べ、ここでは一致していないのでポインタ P を一つ進め 2 とする。

次に、ポインタ P の値を一つ進めたならば、図 4 c に示すように、インデックスの VL のポインタ P が示す順位 2 のエントリに登録されている値 F を取得し、取得した値 F が統合インデックスの VL に格納されているかどうかを調べる。ここでは、値 F は、統合インデックスの VL に格納されていないので、取得した値 F を追加値リスト adVL に追加登録する。また、カレントナンバ C N の値 4 を、T_VL の、ポインタ P が示す順位 2 のエントリに登録する。

【 0 0 4 1 】

そして、ポインタ P の値が、インデックスの VL の末尾のエントリと一致しているかどうかを調べ、ここでは一致しているので、登録値調査処理を終了する。

そして、登録値調査処理を終了したならば、次に、T_VNo 作成処理を行う。

T_VNo 作成処理では、図 5 a に示すように、インデックスの VNo とエントリ数が同じ T_VNo を作成する。

そして、インデックスの VNo の各エントリについて、当該順位 i のエントリに格納されている値 X を取得する。そして、取得した値 X と同じ順位 X の T_VL のエントリに登録されている値 Y を、順位 i の T_VNo のエントリに登録する。すなわち、たとえば、インデックスの VNo の順位 1 のエントリについては、インデックスの VNo の順位 1 のエントリに登録されている値は 1 であるので、順位 1 の T_VL のエントリに登録されている値 3 を取得する。そして、取得した値 3 を T_VNo の順位 1 のエントリに登録する。

【 0 0 4 2 】

そして、このようにして、T_VNo の各エントリに値を登録したならば、T_VNo 作成処理を終了する。

そして、次に、図 5 b に示すように、登録値調査処理で作成した追加値リスト adVL の各エントリを、統合インデックスの VL に追加し、T_VNo 作成処理で作成した T_VNo の各エン

10

20

30

40

50

トリを統合インデックスのVNoに追加する。

以上、データ統合部 1 1 が行う、インデックスセットの統合インデックスセットへの統合動作について説明した。

ところで、以上の図 5のように作成した統合インデックスセットは、図 6に示すように、異なるインデックスセットから作成した統合インデックスセットの各統合インデックスのVNoの部分毎に、別個のデータ構造体として作成、保存するようにしてもよい。すなわち、たとえば、テーブルAインデックスセットから最初に作成した統合インデックスセットの各統合インデックスのVNoの部分 6 0 1 と、テーブルBインデックスセットを統合インデックスセットに統合する際に作成し、統合インデックスセットに追加した各統合インデックスのVNoの部分 6 0 2 と、テーブルCインデックスセットを統合インデックスセットに統合する際に作成し、統合インデックスセットに追加した各統合インデックスのVNoの部分 6 0 3 とを、それぞれ別個のファイル、または、別個のファイル群などの、別個のデータ構造体としてストレージ 1 に保存するようにしてもよい。また、この場合には、各データ構造体の作成元となったインデックス、または、各データ構造体の作成元となったインデックスが表すテーブルと、データ構造体の対応を示すデータを対応データとして統合インデックスセットに含めておくようにしてもよく、このようにすることにより、テーブルやインデックス単位での統合インデックスセットの多様な利用が容易に行えるようになる。

【 0 0 4 3 】

このようにしても、各VNo部分 6 0 1、6 0 2、6 0 3の各フィールドに対応する部分を連結したものが、当該フィールドに対応する統合インデックスのVNoを形成することとなるので、論理的に統合インデックステーブルを表していることに変わりはない。

以下、プロセッサ 2 の R D B M S 1 2 が行う、統合インデックスセットを用いた、それまで保存されたインデックスセットが表す各テーブルに対する統合的な操作について説明する。

いま、図 3 b 3に示す統合インデックスセットがストレージ 1 に格納されているものとして、R D B M S 1 2 は、特定のフィールドに特定の値を持つレコードを求める操作を行う場合には次の処理を行う。

すなわち、まず、統合インデックスセットの当該特定のフィールドの統合インデックスのVLの、当該特定の値が登録されているエントリの順位を求める。そして、求めた順位が登録されている当該統合インデックスのVNoの各エントリの順位を求める。そして、求めた各順位をレコード番号とするレコードを求めるレコードのレコード番号として抽出する。

【 0 0 4 4 】

そして、次に、抽出した各レコード番号について、以下の処理を行って当該レコード番号のレコードの各フィールドの値を求める。

すなわち、統合インデックスの各フィールドの統合インデックスについて、当該統合インデックスのVNoの、当該レコード番号と等しい順位のエントリに登録されている値を求め、求めた値と等しい順位の当該統合インデックスのVLに登録されている値を、当該レコード番号のレコードの当該統合インデックスに対応するフィールドの値として算定する。

【 0 0 4 5 】

また、R D B M S 1 2 は、特定のフィールドに特定の範囲内の値を持つレコードを求める操作を行う場合には、当該特定の範囲内の各値について、当該値の各々を特定の値として、以上の特定のフィールドに特定の値を持つレコードを求める操作を行う処理を行う。

また、R D B M S 1 2 は、特定のフィールドに特定の値を持つレコードの数を求める操作を行う場合には次の処理を行う。

すなわち、この場合には、統合インデックスセットの当該特定のフィールドの統合インデックスのVLの、当該特定の値が登録されているエントリの順位を求める。そして、求めた順位が登録されている当該統合インデックスのVNoのエントリを数え、当該特定のフィールドに特定の値を持つレコードの数として算定する。

【 0 0 4 6 】

また、R D B M S 1 2 は、統合インデックスセットが表す統合テーブルのレコードを、特定のフィールドの値でソートする操作を行う場合には、以下の処理を行う。

すなわち、統合インデックスセットの当該特定のフィールドの統合インデックスのVLのエントリの順位を、当該エントリに登録されている値の順序（たとえば、値が小さい順）に従って配列する。そして、配列した各値の配列順に、当該値が登録されているVNoのエントリの順位を求め、求めたVNoのエントリの順位を当該順位を求めた順番に従って配列していく。そして、この結果得られた順位の配列を、ソートしたレコード番号の配列として算定する。そして、そして各レコード番号の各レコードの各フィールドの値を上述のように求めて、ソートしたレコード番号の順に配列してソート後のテーブルを作成する。

【 0 0 4 7 】

10

なお、以上のような操作は、上述した対応データを統合インデックスセットに含めた場合には、特定のテーブル/インデックスに対応するVNoの範囲のみを対象として行うこともできる。

以上のように、本実施形態に係る統合インデックスセットによれば、R D B M S 1 2 によって、統合インデックスセットに統合した各インデックスセットが表す各テーブルに対する統合的な操作を、統合インデックスセットに対する直接の操作によって速やかに行うことができる。

【 0 0 4 8 】

以上、本発明の実施形態について説明した。

なお、以上の実施形態において、フィールド構成が異なるインデックスセットがストレージ 1 に格納される場合には、値がブランクのインデックスを追加したり、インデックスの順序を入れ替えることにより、格納されたインデックスセットを、共通のフィールド構成のインデックスセットに変換した上で統合インデックスセットに統合するようにすればよい。

20

【 0 0 4 9 】

なお、以上の実施形態では、プロセッサ 2 のデータ統合部 1 1 は、新たに保存すべきインデックスセットがストレージ 1 に格納されたときに、格納されたインデックスセットを、ストレージ 1 に格納している統合インデックスセットに統合するものとして説明したが、本実施形態は、プロセッサ 2 のデータ統合部 1 1 において、新たに保存すべきテーブルがストレージ 1 に格納されたときに、格納されたテーブルから先に図 8 を用いて説明したようにインデックスセットを作成し、作成したインデックスセットを、ストレージ 1 に格納している統合インデックスセットに統合するものとしてもよい。

30

【 0 0 5 0 】

なお、この場合において、フィールド構成が異なるテーブルがストレージ 1 に格納される場合には、テーブルに値がブランクのフィールドを追加したり、テーブルのフィールドの順序を入れ替えることにより、格納されたテーブルを、共通のフィールド構成のテーブルに変換した上でインデックスセットを作成するようにすればよい。

【 0 0 5 1 】

また、このように、テーブルから統合インデックスセットを作成する場合には、プロセッサ 2 のデータ統合部 1 1 において、以下のように統合テーブルを作成するようにしてもよい。

40

すなわち、図 7 a 1 に示す様に、統合インデックスセットに統合する最初のテーブル A がストレージ 1 に格納されたならば、各フィールド（フィールド「性別」とフィールド「タイプ」）についてのVNoとVLを先に図 8 を用いて説明したように作成し、図 7 a 2 に示すように各フィールドについて作成したVNoを当該フィールドの部分インデックスとし、各フィールドについて作成した部分インデックスを含めたテーブル A 部分インデックスセットを統合インデックスセットに格納する。また、各フィールドについて作成したVLを含めた統合VLを統合インデックスセットに格納する。

【 0 0 5 2 】

そして、次に、図 7 a 2 に示す統合インデックスセットがストレージ 1 に格納されてい

50

る状態で図7b1に示すテーブルBがストレージ1に格納されたならば、テーブルBのフィールドについて、各レコードの当該フィールドに用いられている値（ユニーク値）のうちの、統合インデックスセットの統合VLの当該フィールドのVLに登録されていない値を、図7b2に示すように統合VLの当該フィールドのVLに追加登録する。

【0053】

すなわち、たとえば、フィールド「タイプ」については、テーブルBの各レコードのフィールド「タイプ」の値のうち、BとFが図7a2に示す統合インデックスセットの統合VLのフィールド「タイプ」のVLに登録されていないので、図7b2に示すようにBとFを統合VLのフィールド「タイプ」のVLに追加する。

【0054】

次に、テーブルBの各フィールドについて、VNoを以下のように作成する。

すなわち、テーブルBの各フィールドについて、テーブルBのレコード数と同数のエントリよりなるVNoを作成し、VNoの順位nのエントリに、対応するテーブルのレコード番号nのレコードの当該フィールドの値が登録されている、統合インデックスセットの統合VLの当該フィールドのVLのエントリのVL中の順位を示す値を登録する。

【0055】

すなわち、たとえば、フィールド「タイプ」については、テーブルBのレコード数が4であるので、4つのエントリを持つフィールド「タイプ」のVNoを作成し、作成したフィールド「タイプ」のVNoの順位iのエントリに、テーブルBのレコード番号iのレコードのフィールド「タイプ」の値が登録されている、統合VLのフィールド「タイプ」のVL（VL（タイプ））のエントリのVL中の順位を示す値を登録する。たとえば、フィールド「タイプ」のVNoの順位2のエントリには、テーブルBのレコード番号2のレコードのフィールド「タイプ」の値がAであり、統合VLのフィールド「タイプ」のVL（VL（タイプ））の順位0のエントリにAが登録されているので0が登録される。またフィールド「タイプ」のVNoの順位3のエントリには、テーブルBのレコード番号3のレコードのフィールド「タイプ」の値がFであり、統合VLのフィールド「タイプ」のVL（VL（タイプ））の順位4のエントリにFが登録されているので4が登録される。

【0056】

そして、以上のようにして、各フィールドについて作成したVNoを当該フィールドの部分インデックスとし、各フィールドについて作成した部分インデックスを含めたテーブルB部分インデックスセットを統合インデックスセットに格納する。

以降、新たなテーブルがストレージ1に格納される度に、テーブルBの場合と同様の処理を行い、格納されたテーブルを統合インデックスセットに統合していく。

なお、この場合には統合インデックスセットには、各部分インデックスセットが、どのテーブルから作成されたものであるのかを示すデータを、上述した対応データとして含めておくようにしてもよく、このようにすることにより、テーブル単位での統合インデックスセットの多様な利用が容易に行えるようになる。

【0057】

さて、図7b1の統合インデックスセットの各テーブルの各インデックスセットの各フィールドの部分インデックスのVNoは、これを連結して統合部分インデックスとして統合セットに保存するようにしてもよい。なお、この場合には統合インデックスセットには、統合部分インデックスのVNoのエントリのどの範囲が、どのテーブルから作成されたものであるのかを示すデータを上述した対応データとして含めておくようにしてもよく、このようにすることにより、テーブル単位での統合インデックスセットの多様な利用が容易に行えるようになる。

【0058】

ここで、このように部分インデックスのVNoを連結して統合部分インデックスとして保存した場合の各フィールドのVNoは、図7a1のテーブルAと図7b1のテーブルBと同じ2つのテーブル（図2a1のテーブルAと図2b2のテーブルB）のインデックスセットを図2に示したように統合した図2b3に示す統合インデックスセットの各フィールド

10

20

30

40

50

のインデックスのVNoと等しくなる。また、統合VLの各フィールドのVLも、図2b3に示す統合インデックスセットの各フィールドのVLと等しくなる。

【0059】

すなわち、部分インデックスのVNoを連結して統合部分インデックスとして保存した場合の統合インデックスセットのデータ構造体としての構成は、図2b3に示す統合インデックスセットのデータ構造体としての構成と等しくなる。

また、逆に、上述のように異なるインデックスセットから作成した統合インデックスセットの各統合インデックスのVNoの部分毎に、別個のデータ構造体として図6に示すように作成、保存した場合の図2b3に示す統合インデックスセットのデータ構造体としての構成は、図7b2の統合インデックスセットのデータ構造体としての構成と等しくなる。

10

【0060】

よって、図7に示したテーブルを統合インデックスセットに統合する処理によれば、インデックスセットを介さずに、テーブルを統合インデックスセットに統合することができる。また、上述のようにインデックスセットを統合して作成した統合インデックスセットと、同様に利用することができる。

【0061】

また、図7に示したテーブルを統合インデックスセットに統合する処理においても、新たなテーブルを統合インデックスセットに統合する際に、統合インデックスセットの既に登録されている部分には一切の変更が加えられず、新たなテーブルの内容に応じた内容が統合インデックスセットに追加的に登録されていくことにより、テーブルの統合インデックスセットへの統合が実現される。

20

【0062】

なお、図7に示した場合においても、フィールド構成が異なるテーブルがストレージ1に格納される場合には、テーブルに値がブランクのフィールドを追加したり、テーブルのフィールドの順序を入れ替えることにより、格納されたテーブルを、共通のフィールド構成のテーブルに変換した上で処理するようにする。

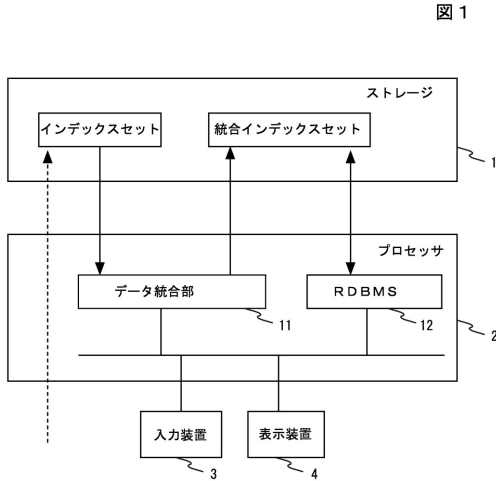
【符号の説明】

【0063】

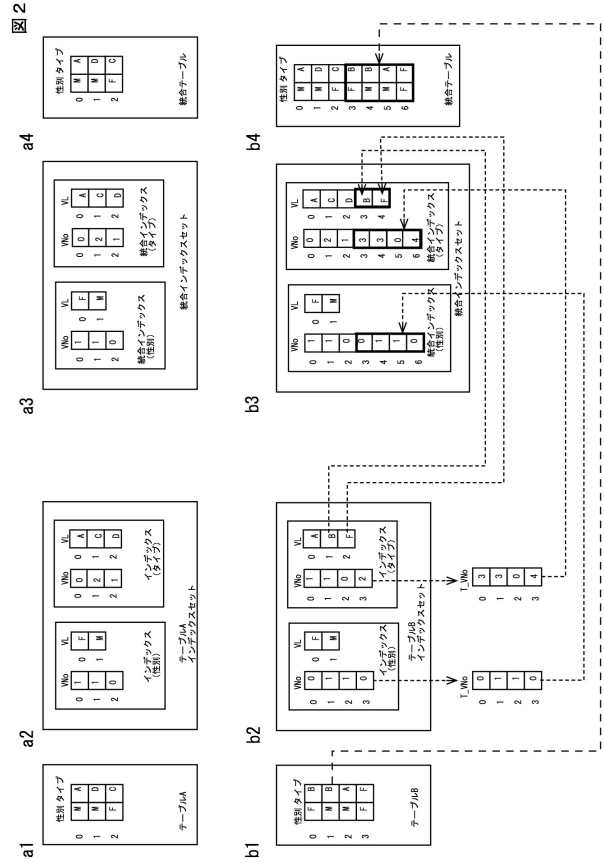
1 ... ストレージ、 2 ... プロセッサ、 3 ... 入力装置、 4 ... 表示装置、 11 ... データ統合部、 12 ... RDBMS。

30

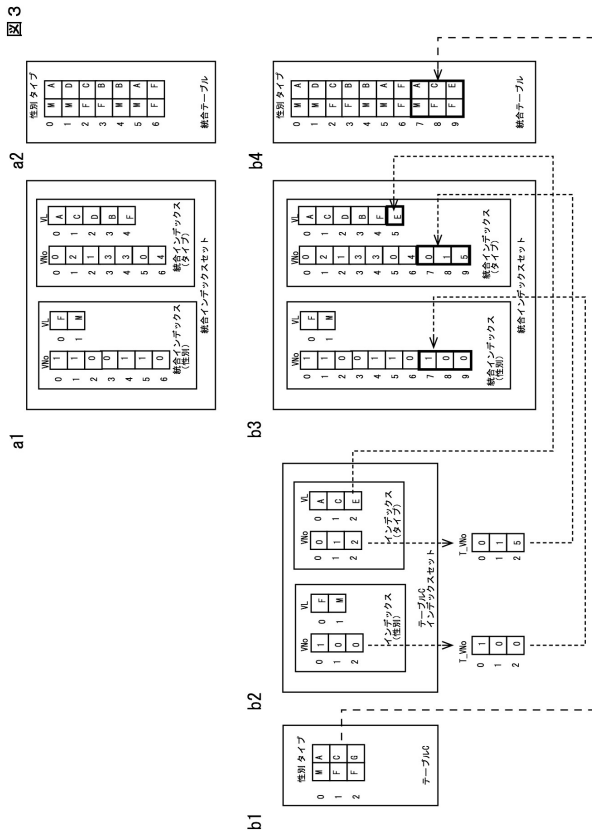
【図 1】



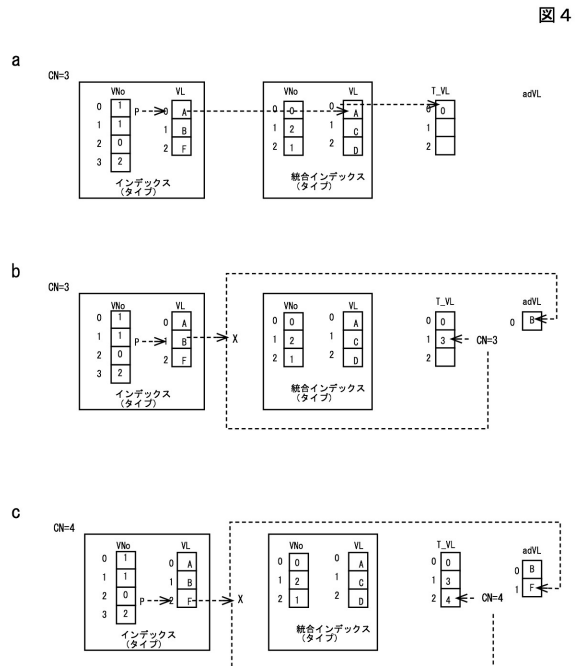
【図 2】



【図 3】



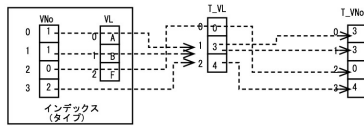
【図 4】



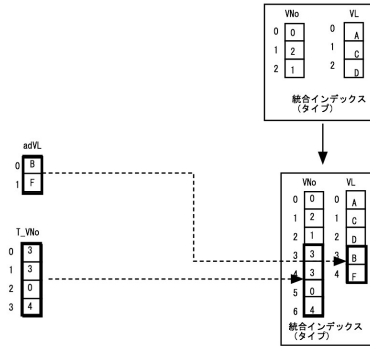
【図 5】

図 5

a

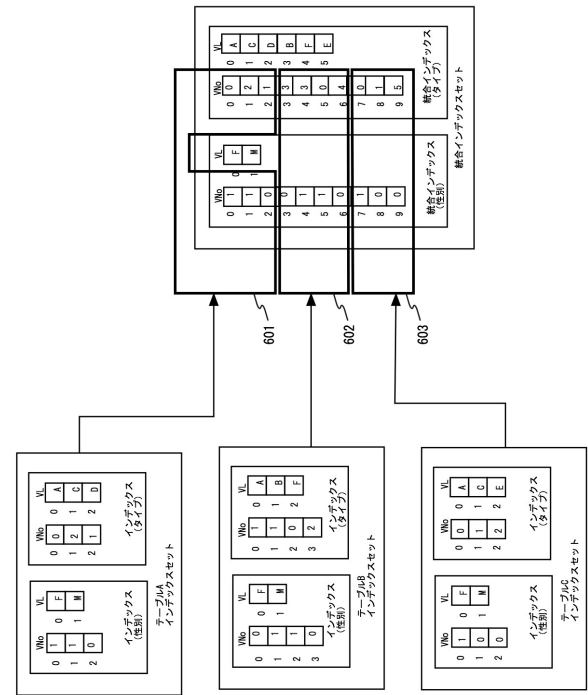


b



【図 6】

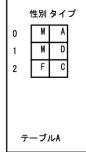
図 6



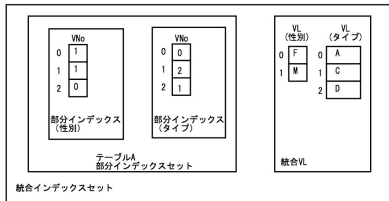
【図 7】

図 7

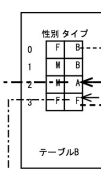
a1



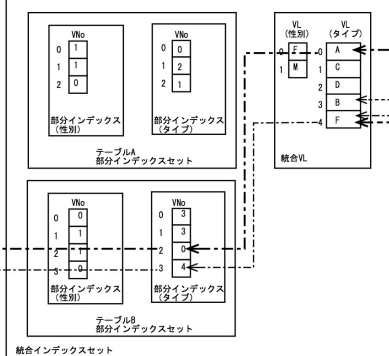
a2



b1



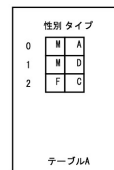
b2



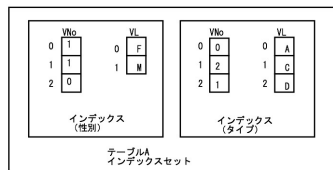
【図 8】

図 8

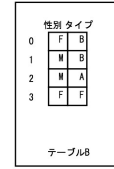
a1



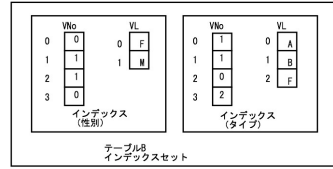
a2



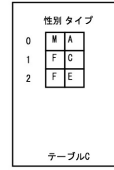
b1



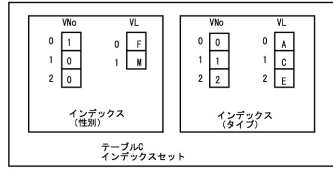
b2



c1



c2



フロントページの続き

(56)参考文献 特開 2 0 1 5 - 2 0 7 0 2 6 (J P , A)
特開平 1 - 1 6 3 8 2 6 (J P , A)
特開 2 0 1 5 - 1 0 6 3 4 7 (J P , A)

(58)調査した分野(Int.Cl. , D B 名)
G 0 6 F 1 6 / 0 0 - 1 6 / 9 5 8