

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B1)

(11) 特許番号

特許第6269884号  
(P6269884)

(45) 発行日 平成30年1月31日(2018.1.31)

(24) 登録日 平成30年1月12日(2018.1.12)

(51) Int.Cl. F I  
**G06F 17/30 (2006.01)**  
 G06F 17/30 180D  
 G06F 17/30 230Z  
 G06F 17/30 210D

請求項の数 11 (全 31 頁)

<p>(21) 出願番号 特願2017-99789 (P2017-99789)                  (22) 出願日 平成29年5月19日(2017.5.19)                  審査請求日 平成29年5月29日(2017.5.29)                   早期審査対象出願</p>	<p>(73) 特許権者 592218300                  学校法人神奈川大学                  神奈川県横浜市神奈川区六角橋3丁目27番1号                   (74) 代理人 100098626                  弁理士 黒田 壽                   (74) 代理人 100134728                  弁理士 奥川 勝利                   (72) 発明者 桑原 恒夫                  神奈川県横浜市神奈川区六角橋三丁目27番1号 学校法人 神奈川大学内                   審査官 齊藤 貴孝</p>
---	---

最終頁に続く

(54) 【発明の名称】 情報検索装置、検索用プログラム、データベースの更新装置、データベース更新用プログラム

(57) 【特許請求の範囲】

【請求項1】

複数の関係項目の中から指定された関係項目に関連付けられた検索対象項目を、複数の検索対象項目の中からコンピュータによって検索する情報検索装置であって、

前記複数の関係項目のうちの2つ以上の関係項目に同一の上位項目を関連づけた第一データと、

前記複数の検索対象項目のうち、前記同一の上位項目が関連づけられた関係項目に係る検索対象項目に、該上位項目を関連づけた第二データと、

前記第一データ及び前記第二データでは関連づけられない関係項目と検索対象項目とを関連づけた第三データと、を有するデータ構造をもつデータベースを用い、

前記複数の関係項目の中から1つ以上の関係項目が指定された所定の検索条件での検索を行う場合、前記コンピュータが、前記第一データを参照し、該1つ以上の関係項目のうちの一部又は全部が関連づけられた上位項目が存在すると判断したときは、少なくとも前記第二データを検索し、前記関連づけられた上位項目が存在しないと判断したときは、前記第三データのみを検索することを特徴とする情報検索装置。

【請求項2】

請求項1に記載の情報検索装置において、

前記第一データは、所定の上位項目作成条件を満たす2以上の関係項目に同一の上位項目を関連づけたものであることを特徴とする情報検索装置。

【請求項3】

請求項 1 又は 2 に記載の情報検索装置において、

前記データベースは、複数の前記上位項目のうちの 2 つ以上の上位項目に同一の第二上位項目を関連づけた第四データと、前記複数の検索対象項目のうち、前記同一の第二上位項目が前記上位項目を介して関連づけられた関係項目に関係する検索対象項目に、該第二上位項目を関連づけた第五データとを有し、

前記複数の関係項目の中から 2 つ以上の関係項目が指定された所定の検索条件での検索を行う場合、前記コンピュータが、前記第四データを参照して、該所定の検索条件に一致する 2 以上の関係項目が上位項目を介して関連づけられた第二上位項目が存在すると判断したときは、前記第五データを検索することを特徴とする情報検索装置。

【請求項 4】

請求項 1 乃至 3 のいずれか 1 項に記載の情報検索装置において、

前記複数の関係項目の中から 2 つ以上の関係項目が指定された所定の検索条件での検索を行う場合、前記コンピュータが、前記第一データを参照し、該所定の検索条件に一致する 2 以上の関係項目が関連づけられた上位項目が存在すると判断したときは、前記第二データを検索し、該所定の検索条件に一致する 2 以上の関係項目が関連づけられた上位項目が存在しないと判断したときは、前記第三データを検索することを特徴とする情報検索装置。

【請求項 5】

複数の関係項目の中から指定された関係項目に関連付けられた検索対象項目を、複数の検索対象項目の中からコンピュータによって検索する情報検索装置であって、

前記複数の検索対象項目のうちの 2 つ以上の検索対象項目に関連づけられた上位項目を、該 2 つ以上の検索対象項目に関係する関係項目に関連づけた第六データと、

前記第六データには含まれない関係項目と該関係項目に関係する検索対象項目とを関連づけた第七データと、

前記複数の検索対象項目のうちの 2 つ以上の検索対象項目に同一の上位項目を関連づけた第八データと、を有するデータ構造をもつデータベースを用い、

前記コンピュータが、前記第六データを参照して前記指定された関係項目に関連づけられた上位項目が存在すると判断したときは、前記第八データから該上位項目に関連づけられた検索対象項目を抽出して、該検索対象項目を検索結果として出力し、前記第七データを参照して前記指定された関係項目に関係する検索対象項目が存在すると判断したときは、該検索対象項目を検索結果として出力することを特徴とする情報検索装置。

【請求項 6】

複数の関係項目の中から指定された関係項目に関連付けられた検索対象項目を複数の検索対象項目の中から検索する工程をコンピュータに実行させるための検索用プログラムであって、

前記工程では、前記複数の関係項目のうちの 2 つ以上の関係項目に同一の上位項目を関連づけた第一データと、前記複数の検索対象項目のうち、前記同一の上位項目が関連づけられた関係項目に関係する検索対象項目に、該上位項目を関連づけた第二データと、前記第一データ及び前記第二データでは関連づけられない関係項目と検索対象項目とを関連づけた第三データとを有するデータ構造をもつデータベースを用い、前記複数の関係項目の中から 1 つ以上の関係項目が指定された所定の検索条件での検索を行う場合、前記コンピュータに、前記第一データを参照させ、該 1 つ以上の関係項目のうちの一部又は全部が関連づけられた上位項目が存在すると判断したときは、少なくとも前記第二データを検索させ、前記関連づけられた上位項目が存在しないと判断したときは、前記第三データのみを検索させることを特徴とする検索用プログラム。

【請求項 7】

複数の関係項目の中から指定された関係項目に関連付けられた検索対象項目を複数の検索対象項目の中から検索する工程をコンピュータに実行させるための検索用プログラムであって、

前記工程では、前記複数の検索対象項目のうちの 2 つ以上の検索対象項目に関連づけら

10

20

30

40

50

れた上位項目を、該2つ以上の検索対象項目に係る関係項目に関連づけた第六データと、

前記第六データには含まれない関係項目と該関係項目に係る検索対象項目とを関連づけた第七データと、

前記複数の検索対象項目のうち2つ以上の検索対象項目に同一の上位項目を関連づけた第八データと、を有するデータ構造をもつデータベースを用い、前記コンピュータに、前記第六データを参照させて前記指定された関係項目に関連づけられた上位項目が存在すると判断したときは、前記第八データから該上位項目に関連づけられた検索対象項目を抽出させて、該検索対象項目を検索結果として出力させ、前記第七データを参照させて前記指定された関係項目に係る検索対象項目が存在すると判断したときは、該検索対象項目を検索結果として出力させることを特徴とする検索用プログラム。

10

【請求項8】

複数の関係項目の中から指定された関係項目に関連付けられた検索対象項目を、複数の検索対象項目の中からコンピュータによって検索する情報検索装置に用いられるデータベースであって、前記複数の関係項目のうち2つ以上の関係項目に同一の上位項目を関連づけた第一データと、前記複数の検索対象項目のうち、前記同一の上位項目が関連づけられた関係項目に係る検索対象項目に、該上位項目を関連づけた第二データと、前記第一データ及び前記第二データでは関連づけられない関係項目と検索対象項目とを関連づけた第三データと、を有するデータ構造をもつデータベースの更新装置であって、

検索対象項目とこれに係る関係項目とを追加する追加手段と、

所定の上位項目作成条件を満たす2以上の関係項目に同一の上位項目を関連づけて前記データベースを更新する更新手段とを有することを特徴とするデータベースの更新装置。

20

【請求項9】

複数の関係項目の中から指定された関係項目に関連付けられた検索対象項目を、複数の検索対象項目の中からコンピュータによって検索する情報検索装置に用いられるデータベースであって、前記複数の検索対象項目のうち2つ以上の検索対象項目に関連づけられた上位項目を、該2つ以上の検索対象項目に係る関係項目に関連づけた第六データと、前記第六データには含まれない関係項目と該関係項目に係る検索対象項目とを関連づけた第七データと、前記複数の検索対象項目のうち2つ以上の検索対象項目に同一の上位項目を関連づけた第八データと、を有するデータ構造をもつデータベースの更新装置であって、

30

検索対象項目とこれに係る関係項目とを追加する追加手段と、

所定の上位項目作成条件を満たす2以上の検索対象項目に同一の上位項目を関連づけて前記データベースを更新する更新手段とを有することを特徴とするデータベースの更新装置。

【請求項10】

複数の関係項目の中から指定された関係項目に関連付けられた検索対象項目を、複数の検索対象項目の中からコンピュータによって検索する情報検索装置に用いられるデータベースであって、前記複数の関係項目のうち2つ以上の関係項目に同一の上位項目を関連づけた第一データと、前記複数の検索対象項目のうち、前記同一の上位項目が関連づけられた関係項目に係る検索対象項目に、該上位項目を関連づけた第二データと、前記第一データ及び前記第二データでは関連づけられない関係項目と検索対象項目とを関連づけた第三データと、を有するデータ構造をもつデータベースを更新する工程をコンピュータに実行させるためのデータベース更新用プログラムであって、

40

検索対象項目とこれに係る関係項目とを追加する追加工程と、

所定の上位項目作成条件を満たす2以上の関係項目に同一の上位項目を関連づけて前記データベースを更新する更新工程とを、前記コンピュータに実行させることを特徴とするデータベース更新用プログラム。

【請求項11】

複数の関係項目の中から指定された関係項目に関連付けられた検索対象項目を、複数の

50

検索対象項目の中からコンピュータによって検索する情報検索装置に用いられるデータベースであって、前記複数の検索対象項目のうち2つ以上の検索対象項目に関連づけられた上位項目を、該2つ以上の検索対象項目に係る関係項目に関連づけた第六データと、前記第六データには含まれない関係項目と該関係項目に係る検索対象項目とを関連づけた第七データと、前記複数の検索対象項目のうち2つ以上の検索対象項目に同一の上位項目を関連づけた第八データと、を有するデータ構造をもつデータベースを更新する工程をコンピュータに実行させるためのデータベース更新プログラムであって、

検索対象項目とこれに係る関係項目とを追加する追加工程と、

所定の上位項目作成条件を満たす2以上の検索対象項目に同一の上位項目を関連づけて前記データベースを更新する更新工程とを、前記コンピュータに実行させることを特徴とするデータベース更新プログラム。

10

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、複数の関係項目の中から指定された関係項目に関連付けられた検索対象項目を、複数の検索対象項目の中からコンピュータによって検索するためのデータベースを用いる情報検索装置及び検索プログラム、並びに、当該データベースの更新装置及び当該データベース更新用のプログラムに関する。

【背景技術】

【0002】

従来、例えばネットスーパーで商品を検索するにあたり、顧客はまず先の選択カテゴリー（肉の種類）からいずれかの先選択項目（関係項目：「牛肉」、「豚肉」、「鶏肉」）を選択（指定）した後に、後の選択カテゴリー（部位の種類）からいずれかの後選択項目（関係項目：「サーロイン」、「タン」、「ヒレ」、「モモ」、「肩」、「ササミ」、「レバー」）を選択（指定）することで、選択された選択項目に該当する商品（検索対象項目：例えば牛のモモ肉）が表示される絞り込み商品検索システムが知られている。

20

【0003】

特許文献1には、先選択項目に対して重複して関連づけられる後選択項目の重複を減らすために、いくつかの先選択項目（関係項目）を上位の分類（上位項目）でグループ化したデータテーブルを含むデータ構造が開示されている。このデータ構造には、後選択項目（関係項目）を先選択項目や上位分類と関連づけたデータテーブルも含まれている。

30

【0004】

例えば、先の選択カテゴリー（肉の種類）について、「牛肉」及び「豚肉」の先選択項目を、「哺乳類」という上位分類でグループ化する。更に、「哺乳類」の上位分類及び「鶏肉」の先選択項目を、「動物（肉）」という上位分類でグループ化する。この場合、下記の表1に示すようなデータテーブルが作成される。一方で、後の選択カテゴリー（部位の種類）については、上述した上位分類（「哺乳類」、「動物（肉）」）や先選択項目（「牛肉」、「豚肉」、「鶏肉」）と関連づけた下記の表2に示すようなデータテーブルを作成する。これらのデータテーブルを用いることで、例えば、先選択項目として「牛肉」が選択された場合、下記の表1のデータテーブルから、「牛肉」に関連づけられた「哺乳類」と、その「哺乳類」に関連付けられた「動物（肉）」とが選択される。その後、その選択結果である「哺乳類」、「動物（肉）」、更には選択された「牛肉」のいずれかに関連付けられた肉の部位を、下記の表2のデータテーブルから抽出し、後の選択カテゴリーとして顧客に提示される。これを受けて、顧客が後選択項目として「モモ」を選択した場合、下記の表3に示すようなデータテーブルから、「牛肉・モモ」に関連付けられた商品番号（検索対象項目）が検索され、商品番号4に対応する「牛のモモ肉」の商品が顧客に提示される。

40

【0005】

【表 1】

哺乳類	牛肉
哺乳類	豚肉
動物(肉)	哺乳類
動物(肉)	鶏肉

【0006】

【表 2】

牛肉	サーロイン
牛肉	タン
鶏肉	ササミ
哺乳類	ヒレ
哺乳類	肩
動物(肉)	モモ
動物(肉)	レバー

【0007】

【表 3】

牛肉	ヒレ	商品番号1
牛肉	肩	商品番号2
牛肉	サーロイン	商品番号3
牛肉	モモ	商品番号4
牛肉	タン	商品番号5
牛肉	タン	商品番号6
牛肉	レバー	商品番号7
豚肉	ヒレ	商品番号8
豚肉	肩	商品番号9
豚肉	肩	商品番号10
豚肉	モモ	商品番号11
豚肉	レバー	商品番号12
鶏肉	ササミ	商品番号13
鶏肉	ササミ	商品番号14
鶏肉	モモ	商品番号15
鶏肉	レバー	商品番号16

【先行技術文献】

【特許文献】

【0008】

【特許文献1】特許第4853797号公報

【発明の概要】

【発明が解決しようとする課題】

【0009】

近年、複数の関係項目の中から指定された関係項目に関連付けられた検索対象項目を、

10

20

30

40

50

複数の検索対象項目の中からコンピュータによって検索する検索処理において、処理の高速化が重要性を増している。特に、近年のコンピュータの処理能力の向上や記憶容量の向上に応じて、様々なデータベースは、蓄積される検索対象項目の数や関係項目の数が膨大になる一方、コンピュータの処理能力の向上スピードが従来と比べて鈍化していることから、より効率的な検索方法によって検索処理の高速化を実現することが望まれる。

【0010】

前記特許文献1で示された例における最終目的は、商品分類の区分けを記録するだけの表1や表2に示すようなデータテーブルよりもレコード数が圧倒的に多い前記表3に示すようなデータテーブルから、目的とする商品番号を検索することである。すなわち、前記特許文献1に開示のデータ構造において、指定された関係項目（選択された先選択項目（例えば「牛肉」）及び選択された後選択項目（例えば「モモ」））に関連づけられた検索対象項目（例えば「牛のモモ肉」の商品番号）を検索するために用いられるデータテーブルは、前記表3に示すように、各検索対象項目とその検索対象項目に關係する關係項目（先選択項目及び後選択項目）とが関連づけられた一般的なデータ構造と変わらない。したがって、前記特許文献1に開示のデータ構造を用いた検索処理では、検索処理の高速化を図ることはできない。

10

【0011】

本発明は、上述した課題に鑑みてなされたものであり、その目的とするところは、検索処理を高速化できる情報検索装置、検索用プログラム、データベースの更新装置、データベース更新用プログラムを提供することである。

20

【課題を解決するための手段】

【0012】

前記目的を達成するために、請求項1の発明は、複数の関係項目の中から指定された関係項目に関連付けられた検索対象項目を、複数の検索対象項目の中からコンピュータによって検索する情報検索装置であって、前記複数の関係項目のうち2つ以上の関係項目に同一の上位項目に関連づけた第一データと、前記複数の検索対象項目のうち、前記同一の上位項目が関連づけられた関係項目に關係する検索対象項目に、該上位項目に関連づけた第二データと、前記第一データ及び前記第二データでは関連づけられない関係項目と検索対象項目とを関連づけた第三データと、を有するデータ構造をもつデータベースを用い、前記複数の関係項目の中から1つ以上の関係項目が指定された所定の検索条件での検索を行う場合、前記コンピュータが、前記第一データを参照し、該1つ以上の関係項目のうちの一部又は全部が関連づけられた上位項目が存在すると判断したときは、少なくとも前記第二データを検索し、前記関連づけられた上位項目が存在しないと判断したときは、前記第三データのみを検索することを特徴とするものである。

30

本発明は、検索に用いられるデータベースが、上述した第一データ、第二データ及び第三データを含むデータ構造をもつことで、検索処理の高速化を実現するものである。

詳しくは、本発明に係るデータ構造をもつデータベースでは、第一データを参照して、2つ以上の関係項目が関連づけられた上位項目と検索対象項目とが関連づけられた第二データと、第一データ及び第二データでは関連づけられない関係項目と検索対象項目とが関連づけられた第三データとから、指定された関係項目（指定関係項目）に關係する検索対象項目を検索する。

40

第二データは、2つ以上の関係項目が上位項目にまとめられて検索対象項目と関連づけされたものであるため、第二データ及び第三データの両方を足し合わせたレコード数は、関係項目と検索対象項目との関連のすべてを両者の直接的な対応関係のみで記録した一般的なデータ構造をもつデータベースのレコード数よりも少ない。したがって、本発明に係るデータ構造をもつデータベースを用いた場合、第二データ及び第三データの両方から指定関係項目に關係する検索対象項目を検索するときの検索処理速度は、上述した一般的なデータ構造をもつデータベースを用いた場合よりも高速化できる。

ここで、本発明において特筆すべきは、2つ以上の関係項目が指定された所定の検索条件の検索処理を行う場合に飛躍的な高速化が望める点である。

50

例えば、第一データにおいて同一の上位項目に関連づけられた2つ以上の関係項目の全部に  
関係する検索対象項目を第二データに含めた場合、指定された2つ以上の関係項目を  
AND条件で検索する際には、その指定された2つ以上の関係項目が当該同一の上位項目  
に関連づけられた2つ以上の関係項目と一致するときは、第三データに対する検索を行わ  
ずに、その上位項目に関連づけられた検索対象項目を第二データから検索するだけで、検  
索処理を完了できる。この場合の第二データは、上位項目が関連づけられた2以上の関係  
項目の全部に  
関係する検索対象項目しか含まないので、そのレコード数が少なく、その検索処理速度は、  
一般的なデータ構造をもつデータベースを用いた場合よりも大幅に高速化できる。また、  
この場合、指定された2つ以上の関係項目が当該同一の上位項目に関連づけられた2つ  
以上の関係項目と一致しない場合には、第二データに対する検索を行わずに、その指定  
された2つ以上の関係項目に関連づけられた検索対象項目を第三データからAND条件で  
検索するだけで、検索処理を完了できる。第三データは、第一データ及び第二データ  
では関連づけられない関係項目と検索対象項目とを関連づけるレコードしか含まない  
ので、そのレコード数は少なく、その検索処理速度は、一般的なデータ構造をもつデ  
ータベースを用いた場合よりも、大幅に高速化できる。

10

また、例えば、第一データにおいて同一の上位項目に関連づけられた2つ以上の関係項  
目の少なくとも1つに  
関係する検索対象項目を第二データに含めた場合、指定された2つ以上の関係項目をOR  
条件で検索する際には、その指定された2つ以上の関係項目が当該同一の上位項目に  
関連づけられた2つ以上の関係項目に含まれているときは、第三データに対する検索を  
行わずに、その上位項目に関連づけられた検索対象項目を第二データから検索する  
だけで、検索処理を完了できる。この場合の第二データは、上位項目が関連づけら  
れた2以上の関係項目の少なくとも1つに  
関係する検索対象項目しか含まないので、そのレコード数が少なく、その検索処理速度  
は、一般的なデータ構造をもつデータベースを用いた場合よりも大幅に高速化できる。  
また、この場合、指定された2つ以上の関係項目のいずれも当該同一の上位項目に  
関連づけられた2つ以上の関係項目に含まれていない場合には、第二データに対する  
検索を行わずに、その指定された2つ以上の関係項目に関連づけられた検索対象項目  
を第三データからOR条件で検索するだけで、検索処理を完了できる。第三データは、  
第一データ及び第二データでは関連づけられない関係項目と検索対象項目とを  
関連づけるレコードを含まないので、そのレコード数は少なく、その検索処理速度は、  
一般的なデータ構造をもつデータベースを用いた場合よりも、大幅に高速化できる。  
ただし、第一データにおいて同一の上位項目に関連づけられた2つ以上の関係項目の  
少なくとも1つに  
関係する検索対象項目をその上位項目と関連づけて記録する第二データのテーブル（以下  
「テーブルB」という。）は、第一データにおいて同一の上位項目に関連づけられた2  
つ以上の関係項目の全部に  
関係する検索対象項目をその上位項目と関連づけて記録する第二データのテーブル（以下  
「テーブルA」という。）とは別のテーブルで記録することが望ましい。両者の関連は  
内容的に異なるにもかかわらず、両者を同一のテーブルに記録するとその差異が区  
別できなくなるからである。また、第一データにおいて同一の上位項目に関連づけ  
られた2つ以上の関係項目の少なくとも1つに  
関係する検索対象項目は、当然に第一データにおいて同一の上位項目に関連づけら  
れた2つ以上の関係項目の全部に  
関係する検索対象項目を包含する。したがって、テーブルAに記録したレコードは  
テーブルBには記録しないようにすることが、データの重複記録防止の観点から望ま  
しい。

20

30

40

このように、上述したような所定の検索条件での検索処理が高速化する結果、本発明  
に係るデータ構造をもつデータベースを用いた場合の検索処理速度の期待値は、一般  
的なデータ構造をもつデータベースを用いた場合よりも大幅に向上する。

そして、本発明によれば、このような所定の検索条件を見込んで、上位項目に  
関連づける2つ以上の関係項目を適切に設計すれば、単一条件での検索が可能な  
第二データに対する検索処理を行う頻度を高めることができる。単一条件での  
検索が可能な第二データに対する検索処理は、レコード数が同じ場合、AND  
条件やOR条件での検索を行う第三データに対する検索処理よりも高速な  
検索が可能である。したがって、検索処理速度の期待値

50

は、より一層向上する。

【0013】

また、請求項2の発明は、請求項1に記載の情報検索装置において、前記第一データは、所定の上位項目作成条件を満たす2以上の関係項目に同一の上位項目を関連づけたものであることを特徴とするものである。

本発明において、同一の上位項目が関連づけられる2以上の関係項目を選ぶための上位項目作成条件は適宜設定されるが、例えば、レコード数を効率よく削減して検索処理速度の高速化を図る場合には、2以上の関係項目の全部に検索対象項目の数が所定数以上であるという条件を満たしたときに、当該2以上の関係項目に同一の上位項目を関連づけるようにしてもよい。また、所定の検索条件での検索処理速度の高速化を図る場合には、例えば、過去の検索処理における指定された関係項目の履歴などから、今後

10

【0014】

また、請求項3の発明は、請求項1又は2に記載の情報検索装置において、前記データベースは、複数の前記上位項目のうち2つ以上の上位項目に同一の第二上位項目を関連づけた第四データと、前記複数の検索対象項目のうち、前記同一の第二上位項目が前記上位項目を介して関連づけられた関係項目に検索対象項目に、該第二上位項目を関連づけた第五データとを有し、前記複数の関係項目の中から2つ以上の関係項目が指定された所定の検索条件での検索を行う場合、前記コンピュータが、前記第四データを参照して、該所定の検索条件に一致する2以上の関係項目が上位項目を介して関連づけられた第二上位項目が存在すると判断したときは、前記第五データを検索することを特徴とするものである。

20

本発明によれば、検索に用いられるデータベースが、上述した第一データ、第二データ及び第三データのほかに、第四データ及び第五データも含むデータ構造をもつことで、更なる検索処理の高速化を実現することが可能となる。特に、指定された関係項目が第四データにおいて同一の第二上位項目に関連づけられた上位項目に関連づけられている2つ以上の関係項目と一致する場合、第二データ及び第三データに対する検索を行わずに、第四データを参照してその第二上位項目に関連づけられた検索対象項目を第五データから検索するだけで、検索処理を完了できる。第五データのレコード数は、第二データのレコード数よりも更に少ないものであり、検索処理速度の更なる高速化を図ることができる。そして、本発明によれば、指定される関係項目の所定の検索条件を見込んで、第二上位項目に

30

また、請求項4の発明は、請求項1乃至3のいずれか1項に記載の情報検索装置において、前記複数の関係項目の中から2つ以上の関係項目が指定された所定の検索条件での検索を行う場合、前記コンピュータが、前記第一データを参照し、該所定の検索条件に一致する2以上の関係項目が関連づけられた上位項目が存在すると判断したときは、前記第二データを検索し、該所定の検索条件に一致する2以上の関係項目が関連づけられた上位項目が存在しないと判断したときは、前記第三データを検索することを特徴とするものである。

40

本発明によれば、2つ以上の関係項目が指定された所定の検索条件での検索を行う場合の検索処理が高速化できる。

【0015】

また、請求項5の発明は、複数の関係項目の中から指定された関係項目に関連付けられた検索対象項目を、複数の検索対象項目の中からコンピュータによって検索する情報検索装置であって、前記複数の検索対象項目のうち2つ以上の検索対象項目に関連づけられた上位項目を、該2つ以上の検索対象項目に検索対象項目に検索対象項目に検索対象項目とを関連づけた第六データと、前記第六データには含まれない関係項目と該関係項目に検索対象項目とを関連づけた第七データと、前記複数の検索対象項目のうち2つ以上の検索対象項目に同一の

50



上位項目を関連づけた第八データと、を有するデータ構造をもつデータベースを用い、前記コンピュータが、前記第六データを参照して前記指定された関係項目に関連づけられた上位項目が存在すると判断したときは、前記第八データから該上位項目に関連づけられた検索対象項目を抽出して、該検索対象項目を検索結果として出力し、前記第七データを参照して前記指定された関係項目に検索対象項目が存在すると判断したときは、該検索対象項目を検索結果として出力することを特徴とする。

本発明は、検索に用いられるデータベースが、上述した第六データ、第七データ及び第八データを含むデータ構造をもつことで、検索処理の高速化を実現するものである。

詳しくは、本発明に係るデータ構造をもったデータベースでは、2つ以上の検索対象項目に関連づけられた上位項目と該2つ以上の検索対象項目に検索対象項目とを関連づけた第六データと、前記第六データには含まれない関係項目と検索対象項目とを関連づけた第七データとから、指定された関係項目（指定関係項目）に検索対象項目を検索する。

第六データは、2つ以上の検索対象項目が上位項目にまとめられて関係項目に関連づけられたものであるため、第六データ及び第七データの両方を足し合わせたレコード数は、関係項目と検索対象項目との関連のすべてを両者の直接的な対応関係のみで記録した一般的なデータ構造をもつデータベースのレコード数よりも少ない。したがって、本発明に係るデータ構造をもったデータベースを用いた場合、第六データ及び第七データの両方から指定関係項目に検索対象項目を検索するときの検索処理速度は、上述した一般的なデータ構造をもつデータベースを用いた場合よりも高速化できる。

更に、本発明においては、単一の関係項目を指定して検索対象項目を検索する場合等において、検索処理速度の高速化が望める。これは、第六データに含まれる関係項目と第七データに含まれる関係項目とが排他的な関係になっているため、第六データ及び第七データのうち、関係項目が指定された際、その関係項目を含むいずれか一方のデータだけを検索するだけで、検索処理を完了できるからである。特に、第六データは、上位項目が関連づけられた2以上の検索対象項目の全部に検索対象項目しか含まないので、そのレコード数が少なく、その検索処理速度は、一般的なデータ構造をもつデータベースを用いた場合よりも大幅に高速化できる。

このように、上述したような特定の検索条件（単一の関係項目を指定して検索対象項目を検索する場合など）での検索処理が高速化する結果、本発明に係るデータ構造をもったデータベースを用いた場合の検索処理速度の期待値は、一般的なデータ構造をもつデータベースを用いた場合よりも向上する。

【0019】

また、請求項6の発明は、複数の関係項目の中から指定された関係項目に関連付けられた検索対象項目を複数の検索対象項目の中から検索する工程をコンピュータに実行させるための検索プログラムであって、前記工程では、前記複数の関係項目のうちの2つ以上の関係項目に同一の上位項目を関連づけた第一データと、前記複数の検索対象項目のうち、前記同一の上位項目が関連づけられた関係項目に検索対象項目に、該上位項目を関連づけた第二データと、前記第一データ及び前記第二データでは関連づけられない関係項目と検索対象項目とを関連づけた第三データとを有するデータ構造をもつデータベースを用い、前記複数の関係項目の中から1つ以上の関係項目が指定された所定の検索条件での検索を行う場合、前記コンピュータに、前記第一データを参照させ、該1つ以上の関係項目のうちの一部又は全部が関連づけられた上位項目が存在すると判断したときは、少なくとも前記第二データを検索させ、前記関連づけられた上位項目が存在しないと判断したときは、前記第三データのみを検索させることを特徴とするものである。

また、請求項7の発明は、複数の関係項目の中から指定された関係項目に関連付けられた検索対象項目を複数の検索対象項目の中から検索する工程をコンピュータに実行させるための検索プログラムであって、前記工程では、前記複数の検索対象項目のうちの2つ以上の検索対象項目に関連づけられた上位項目を、該2つ以上の検索対象項目に検索対象項目に検索対象項目とを関連づけた第六データと、前記第六データには含まれない関係項目と該関係項

目に関係する検索対象項目とを関連づけた第七データと、前記複数の検索対象項目のうち  
の2つ以上の検索対象項目に同一の上位項目を関連づけた第八データと、を有するデータ  
構造をもつデータベースを用い、前記コンピュータに、前記第六データを参照させて前記  
指定された関係項目に関連づけられた上位項目が存在すると判断したときは、前記第八デ  
ータから該上位項目に関連づけられた検索対象項目を抽出させて、該検索対象項目を検索  
結果として出力させ、前記第七データを参照させて前記指定された関係項目に関係する検  
索対象項目が存在すると判断したときは、該検索対象項目を検索結果として出力させるこ  
とを特徴とするものである。

これらの発明によれば、高速な検索処理を実施することが可能となる。

【0020】

また、請求項8の発明は、複数の関係項目の中から指定された関係項目に関連付けられ  
た検索対象項目を、複数の検索対象項目の中からコンピュータによって検索する情報検索  
装置に用いられるデータベースであって、前記複数の関係項目のうち2つ以上の関係項  
目に同一の上位項目を関連づけた第一データと、前記複数の検索対象項目のうち、前記同  
一の上位項目が関連づけられた関係項目に関係する検索対象項目に、該上位項目を関連  
づけた第二データと、前記第一データ及び前記第二データでは関連づけられない関係項目と  
検索対象項目とを関連づけた第三データと、を有するデータ構造をもつデータベースの更  
新装置であって、検索対象項目とこれに関係する関係項目とを追加する追加手段と、所定  
の上位項目作成条件を満たす2以上の関係項目に同一の上位項目を関連づけて前記デー  
タベースを更新する更新手段とを有することを特徴とするものである。

また、請求項9の発明は、複数の関係項目の中から指定された関係項目に関連付けられ  
た検索対象項目を、複数の検索対象項目の中からコンピュータによって検索する情報検索  
装置に用いられるデータベースであって、前記複数の検索対象項目のうち2つ以上の検  
索対象項目に関連づけられた上位項目を、該2つ以上の検索対象項目に関係する関係項  
目に関連づけた第六データと、前記第六データには含まれない関係項目と該関係項目に関  
係する検索対象項目とを関連づけた第七データと、前記複数の検索対象項目のうち2つ以  
上の検索対象項目に同一の上位項目を関連づけた第八データと、を有するデータ構造をも  
つデータベースの更新装置であって、検索対象項目とこれに関係する関係項目とを追加す  
る追加手段と、所定の上位項目作成条件を満たす2以上の検索対象項目に同一の上位項  
目を関連づけて前記データベースを更新する更新手段とを有することを特徴とするもので  
ある。

また、請求項10の発明は、複数の関係項目の中から指定された関係項目に関連付けられ  
た検索対象項目を、複数の検索対象項目の中からコンピュータによって検索する情報検  
索装置に用いられるデータベースであって、前記複数の関係項目のうち2つ以上の  
関係項目に同一の上位項目を関連づけた第一データと、前記複数の検索対象項目のうち、  
前記同一の上位項目が関連づけられた関係項目に関係する検索対象項目に、該上位項目を  
関連づけた第二データと、前記第一データ及び前記第二データでは関連づけられない関係  
項目と検索対象項目とを関連づけた第三データと、を有するデータ構造をもつデータベー  
スを更新する工程をコンピュータに実行させるためのデータベース更新プログラムであ  
って、検索対象項目とこれに関係する関係項目とを追加する追加工程と、所定の上位項目  
作成条件を満たす2以上の関係項目に同一の上位項目を関連づけて前記データベースを更  
新する更新工程とを、前記コンピュータに実行させることを特徴とするものである。

また、請求項11の発明は、複数の関係項目の中から指定された関係項目に関連付けられ  
た検索対象項目を、複数の検索対象項目の中からコンピュータによって検索する情報検  
索装置に用いられるデータベースであって、前記複数の検索対象項目のうち2つ以上の  
検索対象項目に関連づけられた上位項目を、該2つ以上の検索対象項目に関係する関係項  
目に関連づけた第六データと、前記第六データには含まれない関係項目と該関係項目に関  
係する検索対象項目とを関連づけた第七データと、前記複数の検索対象項目のうち2つ  
以上の検索対象項目に同一の上位項目を関連づけた第八データと、を有するデータ構造を  
もつデータベースを更新する工程をコンピュータに実行させるためのデータベース更新用  
プログラムであって、検索対象項目とこれに関係する関係項目とを追加する追加工程と、

10

20

30

40

50

所定の上位項目作成条件を満たす2以上の検索対象項目に同一の上位項目を関連づけて前記データベースを更新する更新工程とを、前記コンピュータに実行させることを特徴とするものである。

これらの発明では、検索対象項目とこれに関係する関係項目の追加により所定の上位項目作成条件を満たすこととなった2以上の関係項目又は2以上の検索対象項目に対し、同一の上位項目を関連づけて、データベースの更新を行う。所定の上位項目作成条件は、適宜設定される。例えば、レコード数を効率よく削減して検索処理速度の高速化を図る場合には、2以上の関係項目又は2以上の検索対象項目の全部に関係する検索対象項目の数が所定数以上になったという条件を満たしたときに、当該2以上の関係項目に同一の上位項目を関連づけるようにしてもよい。また、請求項1乃至4のいずれか1項に記載のデータ構造をもつデータベースについて、所定の検索条件での検索処理速度の高速化を図る場合には、例えば、過去の検索処理における指定された関係項目の履歴などから、今後に所定の検索条件で指定される2つ以上の関係項目を見込んで、この条件を満たす2つ以上の関係項目に同一の上位項目を関連づけるようにしてもよい。

【0021】

なお、上述したプログラムは、CD-ROM等の記録媒体に記録された状態で配布したり、入手したりすることができる。また、上述したプログラムを乗せ、所定の送信装置により送信された信号を、公衆電話回線や専用線、その他の通信網等の伝送媒体を介して配信したり、受信したりすることでも、配布、入手が可能である。この配信の際、伝送媒体中には、コンピュータプログラムの少なくとも一部が伝送されていればよい。すなわち、コンピュータプログラムを構成するすべてのデータが、一時に伝送媒体上に存在している必要はない。上述したプログラムを乗せた信号とは、コンピュータプログラムを含む所定の搬送波に具現化されたコンピュータデータ信号である。また、所定の送信装置からコンピュータプログラムを送信する送信方法には、プログラムを構成するデータを連続的に送信する場合も、断続的に送信する場合も含まれる。

【発明の効果】

【0022】

本発明によれば、検索処理を高速化することができるという優れた効果が奏される。

【図面の簡単な説明】

【0023】

【図1】実施形態1における情報検索装置を模式的に示す外観図である。

【図2】同情報検索装置の構成を示すブロック図である。

【図3】同情報検索装置が用いるデータベースで管理する、顧客識別番号1, 2, 3, 4, 5, ...と、その顧客が過去に購入した商品の商品名a, b, c, dとの関係をまとめた表の一例である。

【図4】同情報検索装置が用いるデータベースで管理される商品名と上位項目との関係をまとめた表の一例である。

【図5】同情報検索装置が用いるデータベースで管理される、顧客識別番号1, 2, 3, 4, 5, ...と、商品名a, b, c, d又は上位項目W, X, Y, Zとの関係をまとめた表の一例である。

【図6】同情報検索装置が用いるデータベースの第二データテーブルのデータ構造を示す説明図である。

【図7】同情報検索装置が用いるデータベースの第三データテーブルのデータ構造を示す説明図である。

【図8】実施形態1におけるAND条件での検索処理の流れを示すフローチャートである。

【図9】実施形態1における単一条件での検索処理の流れを示すフローチャートである。

【図10】実施形態1におけるOR条件での検索処理の流れを示すフローチャートである。

【図11】実施形態1におけるデータベースの更新処理の流れを示すフローチャートであ

10

20

30

40

50

る。

【図12】変形例における情報検索装置が用いるデータベースの第四データテーブルのデータ構造を示す説明図である。

【図13】同情報検索装置が用いるデータベースの第五データテーブルのデータ構造を示す説明図である。

【図14】同情報検索装置が用いるデータベースの第一データテーブルのデータ構造を示す説明図である。

【図15】同情報検索装置が用いるデータベースの第二データテーブルのデータ構造を示す説明図である。

【図16】検索されるデータテーブルと階層とを関連づけたデータテーブルのデータ構造を示す説明図である。

10

【図17】効果確認試験における、顧客識別番号と、その顧客が過去に購入した商品の商品名との関係をまとめた表の一例である。

【図18】効果確認試験における商品名a, bと上位項目Wとを関連づけた第一データテーブルを示す説明図である。

【図19】効果確認試験における、顧客識別番号1~60000と、商品名a, b又は上位項目Wとの関係をまとめたデータテーブルを示す説明図である。

【図20】効果確認試験における、顧客識別番号1~60000のうち、上位項目Wが関連づけられた2つの商品名a, bに共通して関係する顧客識別番号20001~40000に、当該上位項目Wを関連づけたデータテーブルを示す説明図である。

20

【図21】効果確認試験における、顧客識別番号1~60000のうち、上位項目Wが関連づけられていない商品名a, bに關係する顧客識別番号1~20000, 40001~60000に、当該商品名a, bを関連づけたデータテーブルを示す説明図である。

【図22】実施形態2における単一条件での検索処理の流れを示すフローチャートである。

【発明を実施するための形態】

【0024】

〔実施形態1〕

以下、本発明の一実施形態（以下、本実施形態を「実施形態1」という。）を、添付図面を参照しながら説明する。

30

なお、本実施形態1は、スーパーや百貨店などの店舗において顧客の商品購入履歴を管理するためのデータベースを用い、複数の商品名（関係項目）の中から指定された商品名に関連付けられた顧客（当該商品名の商品を購入した顧客）の識別番号（検索対象項目）を、複数の顧客識別番号の中からコンピュータによって検索する検索処理を行う例について説明する。なお、複数の関係項目の中から指定された関係項目に関連付けられた検索対象項目を複数の検索対象項目の中からコンピュータによって検索するためのデータベースを用いるのであれば、どのようなものであってもよい。

【0025】

また、以下の説明では、リレーショナルデータベースの例で説明するが、NoSQLデータベースなどの他のデータベースにも適用可能である。なお、NoSQLデータベースでいえば、以下で述べるデータテーブルはバケットやコレクションなどに対応する。

40

【0026】

図1は、本実施形態1における情報検索装置100を模式的に示す外觀図である。

図2は、本実施形態1における情報検索装置100の構成を示すブロック図である。

本実施形態1における情報検索装置100は、汎用のパーソナルコンピュータあるいはサーバ装置（以下「コンピュータ装置」という。）に、所定のデータベースソフト（検索用プログラム、データベース製造用プログラム）をインストールしたものである。なお、情報検索装置100は、このような汎用のコンピュータ装置に限らず、情報検索専用の装置であってもよい。

【0027】

50

本実施形態 1 の情報検索装置 100 は、主に、装置本体 101 と、表示手段としてのモニター 102 と、キーボード 103 やポインティングデバイス 104 等の操作受付手段としての入力装置とから構成される。装置本体 101 は、主に、CPU (Central Processing Unit) 等からなる演算部 101a と、ROM (Read Only Memory) や RAM (Random Access Memory) あるいは HDD (Hard Disk Drive) 等の記憶装置からなる記憶部 101b とから構成されるが、そのほか、通信インターフェース等からなる通信部など、汎用のコンピュータ装置が備える一般的な機能も備えている。

#### 【0028】

装置本体 101 の記憶部 101b には、情報検索装置 100 の機能を発揮させるためのデータベースソフトを含む各種コンピュータプログラムが記憶されている。また、記憶部 101b は、データ記憶手段として機能し、後述するデータ構造をもつデータベースも記憶される。装置本体 101 の演算部 101a は、記憶部 101b に記憶されている各種コンピュータプログラムを CPU で実行することにより、本実施形態 1 の情報検索装置 100 としての動作に必要な各種機能を実現する。

#### 【0029】

特に、本実施形態 1 の演算部 101a は、データベースソフトに含まれる検索用プログラムを CPU で実行することにより、記憶部 101b に記憶されたデータベースを用いて、指示された商品名 (関係項目) に関連付けられた顧客識別番号 (検索対象項目) を検索する検索処理を行う。また、本実施形態 1 の演算部 101a は、データベースソフトに含まれるデータベース製造用プログラムを CPU で実行することにより、追加入力された顧客識別番号 (検索対象項目) とこれに関係する商品名 (関係項目) に応じて所定の上位項目作成条件を満たすこととなった 2 以上の商品名 (関係項目) に対し、同一の上位項目を関連づけてデータベースを更新し、新たなデータベースを作成 (製造) する処理を行う。

#### 【0030】

図 3 は、本実施形態 1 における情報検索装置 100 が用いるデータベースで管理する、顧客識別番号 1, 2, 3, 4, 5, … (検索対象項目) と、その顧客が過去に購入した商品の商品名 a, b, c, d (関係項目) との関係をもとめた表の一例である。この表によれば、顧客識別番号 1 の顧客は商品名 a, b, c の商品を購入し、顧客識別番号 2 の顧客は商品名 a, c の商品を購入し、顧客識別番号 3 の顧客は商品名 b の商品を購入し、顧客識別番号 4 の顧客は商品名 a, c の商品を購入し、顧客識別番号 5 の顧客は商品名 a の商品を購入し、顧客識別番号 10 の顧客は商品名 d の商品を購入したことが管理されている。

#### 【0031】

従来一般的なデータベースでは、図 3 に示す表のようなデータテーブルを作成し、そのデータテーブルに対して検索処理を行っていた。この場合、顧客識別番号 1 ~ 5 の顧客の購入商品分だけでも、レコード数は 9 個であり、例えば、商品名 a の商品を購入した顧客の識別番号を検索するときは、9 個のレコード数を検索する必要がある。

#### 【0032】

図 4 は、本実施形態 1 における情報検索装置 100 が用いるデータベースで管理される商品名 (関係項目) と上位項目との関係をまとめた表の一例である。

本実施形態 1 においては、所定の 2 つ以上の商品名 (関係項目) を上位項目にグループ化する。図 4 の表の例では、商品名 a と商品名 b を上位項目 W にグループ化し、商品名 a と商品名 c を上位項目 X にグループ化し、商品名 b と商品名 c を上位項目 Y にグループ化し、商品名 a と商品名 b と商品名 c を上位項目 Z にグループ化している。

#### 【0033】

図 5 は、本実施形態 1 における情報検索装置 100 が用いるデータベースで管理される、顧客識別番号 1, 2, 3, 4, 5, … (検索対象項目) と、商品名 a, b, c, d (関係項目) 又は上位項目 W, X, Y, Z との関係をもとめた表の一例である。

仮に図 5 に示す表のようなデータテーブルを作成し、そのデータテーブルに対して検索処理を行った場合、顧客識別番号 1 ~ 5 の顧客の購入商品分だけで見ると、レコード数は

10

20

30

40

50

5個であり、図3に示す表のようなデータテーブルを用いる従来の一般的なデータベースよりも、検索するレコード数が少なく済む。このように検索するレコード数を少なくできれば、1つの商品名(関係項目)を指定して顧客識別番号を検索する場合でも、2つ以上の商品名(関係項目)を指定してOR条件やAND条件の検索処理を行う場合でも、検索処理の高速化を図ることができる。

【0034】

ただし、図5に示す表のようなデータテーブルに対する検索処理の場合、1つの商品名(関係項目)を指定して顧客識別番号(検索対象項目)を検索するときでも、2つ以上の商品名をAND条件で顧客識別番号を検索するときでも、2つ以上の商品名をOR条件で顧客識別番号を検索するときでも、検索するレコード数は同じである。

10

【0035】

本実施形態1においては、以下のようなデータ構造をもつデータベースを用いることで、2つ以上の商品名(関係項目)をAND条件で顧客識別番号を検索する場合の検索処理速度の期待値を向上させる。

【0036】

図6は、本実施形態1における情報検索装置100が用いるデータベースの第二データとしての第二データテーブルのデータ構造を示す説明図である。

図6に示す第二データテーブルは、複数の顧客識別番号1, 2, 3, 4, 5, ... (検索対象項目)のうち、同一の上位項目W, X, Y, Zが関連づけられた商品名a, b, c(関係項目)の全部に關係する顧客識別番号1, 2, 4, 6, 7, ...に、当該上位項目W, X, Y, Zを関連づけたデータである。第二データテーブルは記憶部101bに記憶されており、演算部101aは、検索用プログラムを実行して第二データテーブルから検索を行うことができる。

20

【0037】

図7は、本実施形態1における情報検索装置100が用いるデータベースの第三データとしての第三データテーブルのデータ構造を示す説明図である。

図7に示す第三データテーブルは、複数の顧客識別番号1, 2, 3, 4, 5, ... (検索対象項目)のうち、同一の上位項目W, X, Y, Zが関連づけられた商品名a, b, c(関係項目)の一部のみに關係する顧客識別番号3, 5, 8, ...、及び、同一の上位項目W, X, Y, Zが関連づけられた商品名a, b, c(関係項目)のいずれにも關係しない顧客識別番号10に、当該商品名d(関係項目)を関連づけたデータである。第三データテーブルは記憶部101bに記憶されており、演算部101aは、検索用プログラムを実行して第三データテーブルから検索を行うことができる。

30

【0038】

図8は、本実施形態1におけるAND条件での検索処理の流れを示すフローチャートである。

例えば、商品名aと商品名bのAND条件で両商品を購入したことがある顧客識別番号を検索する場合を考える。入力装置103, 104をユーザーが操作するなどして、AND条件で検索する商品名aと商品名bが入力された場合(S1)、演算部101aは、検索用プログラムを実行し、まず、図4に示す表のような関連づけがなされた第一データとしての第一データテーブルを参照し(S2)、商品名aと商品名bの2つの商品名をグループ化した上位項目が存在するか否かを判断する(S3)。この第一データテーブルは、記憶部101bに記憶されている。

40

【0039】

ここでは、商品名aと商品名bの2つの商品名をグループ化した上位項目Wが存在するので(S3のYes)、演算部101aは、図6に示す第二データテーブルの全体から、当該上位項目Wに関連づけられた顧客識別番号を検索する(S4)。そして、その検索結果に係る顧客識別番号6を出力する(S6)。この出力は、検索結果に係る顧客識別番号そのものをモニター102に表示させるものでもよいが、例えば、顧客識別番号6に関連づけられた各種顧客情報(性別、年齢層など)を別のデータテーブルから読み出しあるいは

50

それらの読み出した結果に様々な分析や加工を加えてうえでモニター 102 に表示させるなど、検索結果に関連づけられた他の情報を出力するものであってもよい。また、出力方法は、モニター 102 に表示する方法に限らず、例えば、通信部からネットワークを通じて外部装置へ送信する方法や、コンピュータ内部の記憶装置に記録させる方法などの他の出力方法であってもよい。

【0040】

一方で、もし商品名 a と商品名 b の 2 つの商品名をグループ化した上位項目が存在しない場合には (S3 の No)、演算部 101 a は、図 7 に示す第三データテーブルの全体から、当該 2 つの商品名 a と商品名 b に関連づけられた顧客識別番号を AND 条件で検索し (S5)、その検索結果を出力することになる (S6)。

10

【0041】

本実施形態 1 によれば、指定された 2 つの商品名 a と商品名 b (関係項目) の全部が上位項目 W に関連づけられた 2 つの商品名 a, b と一致する場合には、図 6 に示す第二データテーブルに対し、その上位項目 W に関連づけられた顧客識別番号 (検索対象項目) を検索するだけで (第三データテーブルに対する検索処理を行わずに)、検索処理を完了する。第二データテーブルは、上位項目 W, X, Y, Z が関連づけられた 2 以上の商品名の全部に關係する顧客識別番号しか含まない。したがって、図 3 に示す表のようなデータテーブルに対して検索処理を行う従来の場合はもとより、図 5 に示す表のようなデータテーブルに対して検索処理を行う場合よりも、検索するレコード数が少なく済み、検索処理速度を高速化できる。

20

【0042】

また、指定された 2 つの商品名 (関係項目) の全部が上位項目 W に関連づけられた 2 つの商品名 a, b と一致しない場合には、図 7 に示す第三データテーブルに対し、その 2 つの商品名に関連づけられた顧客識別番号 (検索対象項目) を検索するだけで (第二データテーブルに対する検索処理を行わずに)、検索処理を完了する。第三データテーブルは、第二データテーブルに含まれている関連性をもった顧客識別番号 (検索対象項目) と商品名 (関係項目) とが関連づけられたレコードを含まないので、そのレコード数は少なく、検索処理速度を高速化できる。

【0043】

このように、いずれの場合であっても AND 条件の検索処理が高速化するので、本実施形態 1 に係るデータ構造をもったデータベースを用いた場合の検索処理速度の期待値を向上させることができる。

30

【0044】

特に、本実施形態 1 によれば、検索の際に AND 条件で指定される頻度が高いと見込まれる 2 つ以上の商品名 (関係項目) を上位項目に適切に関連づけることで、第二データテーブルに対する検索処理の実行頻度を高めることができる。第二データテーブルに対する検索処理 (S4) は、単一条件での検索処理であるため、レコード数が同じである場合でさえも、AND 条件での検索を行う第三データテーブルに対する検索処理 (S5) よりも高速な検索が可能である。したがって、検索処理速度の期待値は、より一層向上する。

【0045】

また、本実施形態 1 においては、上述したデータ構造をもつデータベースを用いることで、1 つの商品名 (関係項目) に關係する顧客識別番号を単一条件で検索する検索処理についても、同様に、検索処理速度が向上する。

40

【0046】

図 9 は、本実施形態 1 における単一条件での検索処理の流れを示すフローチャートである。

例えば、商品名 a の単一条件で、その商品を購入したことのある顧客識別番号を検索する場合を考える。入力装置 103, 104 をユーザーが操作するなどして商品名 a が入力された場合 (S11)、演算部 101 a は、検索用プログラムを実行し、まず、図 4 に示す記憶部 101 b 内の第一データテーブルを参照し (S12)、商品名 a に関連づけられ

50

た上位項目が存在するか否かを判断する（S13）。

【0047】

ここでは、商品名 a に関連づけられている上位項目 W, X, Z が存在するので（S13 の Yes）、演算部 101a は、図 6 に示す第二データテーブルの全体から、これらの上位項目 W, X, Z に関連づけられた顧客識別番号を OR 条件で検索する（S14）。ただし、商品名 a が関連づけられている上位項目が 1 つだけしか存在しない場合には、演算部 101a は、第二データテーブルの全体から、当該 1 つの上位項目に関連づけられた顧客識別番号を単一条件で検索することになる。更に、演算部 101a は、図 7 に示す第三データテーブルについても、商品名 a に関連づけられた顧客識別番号を単一条件で検索する（S15）。そして、これらの検索結果に係る顧客識別番号 1, 2, 4, 5, 6 を出力する（S16）。

10

【0048】

一方で、もし商品名 a に関連づけられた上位項目が存在しない場合には（S13 の No）、演算部 101a は、図 7 に示す第三データテーブルの全体から、商品名 a に関連づけられた顧客識別番号を単一条件で検索し（S15）、その検索結果に係る顧客識別番号を出力する（S16）。

【0049】

本実施形態 1 によれば、指定された 1 つの商品名 a（関係項目）に関連づけられた上位項目が存在しない場合には、図 7 に示す第三データテーブルに対し、その 1 つの商品名 a に関連づけられた顧客識別番号（検索対象項目）を単一条件で検索するだけで（第二データテーブルに対する検索処理を行わずに）、検索処理を完了する。第三データテーブルは、上述したように、第二データテーブルに含まれている関連性をもった顧客識別番号（検索対象項目）と商品名（関係項目）とが関連づけられたレコードを含まないので、図 3 に示す表のようなデータテーブルに対して検索処理を行う従来の場合よりも、図 5 に示す表のようなデータテーブルに対して検索処理を行う場合よりも、検索するレコード数が少なく済み、検索処理速度を高速化できる。

20

【0050】

また、指定された 1 つの商品名 a（関係項目）に関連づけられた上位項目が存在する場合には、図 6 に示す第二データテーブルと図 7 に示す第三データテーブルの両方に対して検索処理を行うことになる。このときのレコード数の合計は、図 3 に示す表のようなデータテーブルに対して検索処理を行う従来の場合と比較すると少ないので検索処理速度を高速化できるが、図 5 に示す表のようなデータテーブルに対して検索処理を行う場合と比較すると、レコード数は同じであり、レコード数の減少による検索処理速度の高速化を図ることはできない。

30

【0051】

ただし、図 5 に示す表のようなデータテーブルに対して検索処理を行う場合には、OR 条件の項目数が、上位項目 W, X, Z と商品名 a の合計 4 つであるのに対し、本実施形態 1 の場合の OR 条件の項目数は、図 6 に示す第二データテーブルに対する検索処理の際の上位項目 W, X, Z の 3 つである。したがって、本実施形態 1 の場合に必要となる OR 条件での検索処理（図 6 に示す第二データテーブルに対する検索処理）は、図 5 に示す表のようなデータテーブルに対する OR 条件での検索処理と比べて、OR 条件の項目数が少なく、かつ、OR 条件で検索するレコード数も少ない。OR 条件の項目数が少ないほど OR 条件での検索処理時間の増加（オーバーヘッド）が少なく済み、また、OR 条件で検索するレコード数が少ないほど OR 条件でのオーバーヘッドが少なく済み。よって、本実施形態 1 によれば、図 5 に示す表のようなデータテーブルに対して検索処理を行う場合と比較しても、検索処理速度を高速化できる。

40

【0052】

更に、本実施形態 1 においては、上述したデータ構造をもつデータベースを用いることで、2 つ以上の商品名（関係項目）に係る顧客識別番号を OR 条件で検索する検索処理についても、同様に、検索処理速度が向上する。

50



## 【 0 0 5 3 】

図 1 0 は、本実施形態 1 における O R 条件での検索処理の流れを示すフローチャートである。

例えば、商品名 a と商品名 b の O R 条件でいずれかの商品を購入したことのある顧客識別番号を検索する場合を考える。入力装置 1 0 3 , 1 0 4 をユーザーが操作するなどして、O R 条件で検索する商品名 a と商品名 b が入力された場合 ( S 2 1 )、演算部 1 0 1 a は、検사용プログラムを実行し、まず、図 4 に示す記憶部 1 0 1 b 内の第一データテーブルを参照して ( S 2 2 )、商品名 a 又は商品名 b に関連づけられた上位項目が存在するかどうかを判断する ( S 2 3 )。

## 【 0 0 5 4 】

ここでは、商品名 a に関連づけられている上位項目として W , X , Z が存在し、商品名 b に関連づけられている上位項目として W , Y , Z が存在するので ( S 2 3 の Y e s )、演算部 1 0 1 a は、図 6 に示す第二データテーブルの全体から、これらの上位項目 W , X , Y , Z に関連づけられた顧客識別番号を O R 条件で検索する ( S 2 4 )。ただし、商品名 a 又は商品名 b が関連づけられている上位項目が 1 つだけしか存在しない場合には、演算部 1 0 1 a は、第二データテーブルの全体から、当該 1 つの上位項目に関連づけられた顧客識別番号を単一条件で検索することになる。更に、演算部 1 0 1 a は、図 7 に示す第三データテーブルについても、商品名 a と商品名 b に関連づけられた顧客識別番号を O R 条件で検索する ( S 2 5 )。そして、これらの検索結果に係る顧客識別番号を出力する ( S 2 6 )。

## 【 0 0 5 5 】

一方で、もし商品名 a 又は商品名 b に関連づけられた上位項目が存在しない場合には ( S 2 3 の N o )、演算部 1 0 1 a は、図 7 に示す第三データテーブルの全体から、商品名 a と商品名 b に関連づけられた顧客識別番号を O R 条件で検索する ( S 2 5 )。そして、これらの検索結果に係る顧客識別番号を出力する ( S 2 6 )。

## 【 0 0 5 6 】

本実施形態 1 によれば、指定された 2 つの商品名 a 又は商品名 b ( 関係項目 ) に関連づけられた上位項目が存在しない場合には、図 7 に示す第三データテーブルに対し、その 2 つの商品名 a と商品名 b に関連づけられた顧客識別番号 ( 検索対象項目 ) を O R 条件で検索するだけで ( 第二データテーブルに対する検索処理を行わずに )、検索処理を完了する。第三データテーブルは、上述したように、第二データテーブルに含まれている関連性をもった顧客識別番号 ( 検索対象項目 ) と商品名 ( 関係項目 ) とが関連づけられたレコードを含まないので、図 3 に示す表のようなデータテーブルに対して検索処理を行う従来の場合よりも、図 5 に示す表のようなデータテーブルに対して検索処理を行う場合よりも、検索するレコード数が少なく済み、また、O R 条件の項目数も少なく済むので、検索処理速度を高速化できる。

## 【 0 0 5 7 】

また、指定された 2 つの商品名 a 又は商品名 b ( 関係項目 ) に関連づけられた上位項目が存在する場合には、図 6 に示す第二データテーブルと図 7 に示す第三データテーブルの両方に対して検索処理を行うことになる。このときのレコード数の合計は、図 3 に示す表のようなデータテーブルに対して検索処理を行う従来の場合と比較すると少ないので検索処理速度を高速化できるが、図 5 に示す表のようなデータテーブルに対して検索処理を行う場合と比較すると、レコード数は同じであり、レコード数の減少による検索処理速度の高速化を図ることはできない。

## 【 0 0 5 8 】

ただし、図 5 に示す表のようなデータテーブルに対して検索処理を行う場合には、O R 条件の項目数が、上位項目 W , X , Y , Z と商品名 a , b の合計 6 つであるのに対し、本実施形態 1 の場合の O R 条件の項目数は、図 6 に示す第二データテーブルに対する検索処理の際の上位項目 W , X , Y , Z の 4 つと、図 7 に示す第三データテーブルに対する検索処理の際の商品名 a , b の 2 つである。本実施形態 1 の場合に必要となる O R 条件での検

10

20

30

40

50

索処理は、図5に示す表のようなデータテーブルに対するOR条件での検索処理と比べて、個々の検索処理におけるOR条件の項目数が少なく、かつ、個々の検索処理でOR条件で検索するレコード数も少ない。OR条件の項目数が少ないほどOR条件での検索処理時間の増加（オーバーヘッド）が少なくて済み、また、OR条件で検索するレコード数が少ないほどOR条件でのオーバーヘッドが少なくて済むので、本実施形態1によれば、図5に示す表のようなデータテーブルに対して検索処理を行う場合と比較しても、検索処理速度を高速化できる。

#### 【0059】

以上のように、本実施形態1によれば、図3に示す表のようなデータテーブルに対して検索処理を行う従来の場合はもとより、図5に示す表のようなデータテーブルに対して検索処理を行う場合と比較しても、単一条件、AND条件、OR条件のいずれの条件での検索処理速度についても高速化することが可能である。

10

#### 【0060】

なお、本実施形態1においては、図4で示したように、所定の2つ以上の商品名（関係項目）を上位項目にグループ化するが、そのグループ化の条件（上位項目作成条件）は適宜設定することができる。例えば、2つ以上の商品名（関係項目）の全部に関係している顧客識別番号（検索対象項目）の数が所定数以上存在する場合に（あるいはデータ蓄積の進行に伴って一定数以上になった時点で）、当該2つ以上の商品名（関係項目）を上位項目にグループ化するという条件としてもよい。この場合、図6に示す第二データテーブルに登録されるレコード数が、図7に示す第三データテーブルに登録されるレコード数に対して相対的に増えることで、より高速な検索処理が実現可能となる。また、例えば、過去の検索処理における指定された商品名（関係項目）の履歴などから、今後指定される2つ以上の商品名（関係項目）のAND条件を見込み（推定し）、今後AND条件で検索される頻度が高いと推定される2つ以上の商品名（関係項目）を上位項目にグループ化するという条件としてもよい。この場合、図6に示す第二データテーブルのみを用いた検索処理（図7に示す第三データテーブルを用いない検索処理）の頻度が増える結果、AND条件での検索処理速度の高速化を図ることができる。

20

#### 【0061】

また、本実施形態1は、指定される2つ以上の商品名（関係項目）がAND条件で検索されるときの検索処理速度を高速化する例であったが、指定される2つ以上の商品名（関係項目）がOR条件で検索されるときの検索処理速度を高速化することも可能である。この場合、第二データテーブルは、例えば、第一データテーブルにおいて同一の上位項目に関連づけられた2つ以上の商品名（関係項目）の少なくとも1つに関係する検索対象項目を含むように構築すればよい。

30

#### 【0062】

次に、本実施形態1におけるデータ構造をもったデータベースの更新方法（製造方法）について説明する。

本実施形態1のようにスーパーや百貨店などの店舗において顧客の商品購入履歴を管理するためのデータベースは、新たなデータが日々追加されていき、また必要に応じて古いデータが日々削除されていく。本実施形態1では、ユーザーが入力装置103、104を用いた手入力により顧客識別番号とその顧客が購入した商品名とを入力することで、図6に示す第二データテーブルや図7に示す第三データテーブルが更新される。新たなデータの入力方法は、このような手入力に限らず、例えば、顧客識別番号とその顧客が購入した商品名とのデータを店舗のレジシステムから送信し、これを情報検索装置100の通信部で受信することで、入力する方法であってもよい。なお、入力タイミングは、例えば、1件の新規データが発生するたびに入力するようにしてもよいし、複数件の新規データを定期的に入力するようにしてもよい。

40

#### 【0063】

図11は、本実施形態1におけるデータベースの更新処理の流れを示すフローチャートである。

50

データベースの製造装置として機能する情報検索装置100に対し、新規のデータが入力されると(S31)、演算部101aは、データベース製造用プログラムを実行して、追加手段として機能し、新規のデータを記憶部101b内のデータベースに追加する追加工程を実施する。そして、所定の2つ以上の商品名(関係項目)を上位項目にグループ化するための上位項目作成条件を満たすことになったか否かを判断する(S32)。この上位項目作成条件は、上述したように、適宜設定することができる。このとき、新規のデータ入力によって上位項目作成条件を満たすことにはならなかった場合には(S32のNo)、演算部101aは、更新手段として機能し、新規のデータ入力によって新たに既存の上位項目に関連付けられることになった顧客識別番号のレコードを新たに図6に示す第二データテーブルに追加する更新処理(更新工程)を行うとともに(S35)、第二データ  
10  
テーブルに追加した分を図7に示す第三データテーブルから削除する更新処理(更新工程)を行う(S36)。また、新規のデータ入力によって新たに既存の上位項目に関連付けられない顧客識別番号のレコードについては、図7に示す第三データテーブルに追加する更新処理(更新工程)を行う(S36)。

#### 【0064】

一方、新規のデータ入力によって上位項目作成条件を満たすことになった場合(S32のYes)、当該上位項目作成条件に対応する新たな上位項目を作成する(S33)。例えば、上位項目作成条件が、2つ以上の商品名(関係項目)の全部に関係している顧客識別番号(検索対象項目)の数が所定数以上存在するという条件であった場合、新規のデータ入力によって当該顧客識別番号(検索対象項目)の数が所定数以上に達したら、新たな  
20  
上位項目を作成して、当該2つ以上の商品名(関係項目)を当該新たな上位項目でグループ化する。そして、当該2つ以上の商品名(関係項目)に対し、当該新たな上位項目を関連づけたレコードを、図4に示す第一データベースに追加する更新処理(更新工程)を実行する(S34)。

#### 【0065】

続いて、新規のデータ入力によって新たに既存の上位項目又は前記新たな上位項目に関連付けられることになった顧客識別番号のレコードを新たに図6に示す第二データテーブルに追加する更新処理を行うとともに(S35)、第二データテーブルに追加した分を図7に示す第三データテーブルから削除する更新処理を行う(S36)。また、新規のデータ入力によって新たに既存の上位項目又は前記新たな上位項目に関連付けられない顧客識別  
30  
番号のレコードについては、図7に示す第三データテーブルに追加する更新処理を行う(S36)。

#### 【0066】

以上のようなデータベースの更新処理を実行することにより、上位項目に適切をグループ化された最新のデータを上述したデータテーブル別で管理されたデータベースを用いて、検索処理を行うことができる。しかも、本実施形態1では、新規のデータ入力によって上位項目作成条件を満たすことになった場合には、新たな上位項目が作成され、当該新たな上位項目を含むデータベースに更新される。これにより、最新のデータを適切な上位項目でグループ化し、上述したデータテーブル別で管理されたデータベースを用いた検索処理が実現される。  
40

#### 【0067】

##### 〔変形例〕

次に、上述した実施形態1におけるデータベースのデータ構造の一変形例について説明する。

図12は、本変形例における情報検索装置100が用いるデータベースの第四データとしての第四データテーブルのデータ構造を示す説明図である。

本変形例では、2つ以上の上位項目W, X, Yを同一の第二上位項目でグループ化し、これらの上位項目W, X, Yと第二上位項目とを関連づけた第四データテーブルを作成して記憶部101bに記憶している。なお、第二上位項目が上位項目W, X, Yを介して関連づけられる商品名はa, b, cであり、上述した実施形態1における上位項目Z  
50

と同じである。本変形例では、上位項目Zは作成されていない。

【0068】

図13は、本変形例における情報検索装置100が用いるデータベースの第五データとしての第五データテーブルのデータ構造を示す説明図である。

図13に示す第五データテーブルは、第二上位項目が上位項目W, X, Yを介して関連づけられた商品名a, b, c(関係項目)に關係する顧客識別番号1, 9に、当該第二上位項目を関連づけたデータである。第五データテーブルは記憶部101bに記憶されており、演算部101aは、検索用プログラムを実行して第五データテーブルから検索を行うことができる。

【0069】

図14は、本変形例における情報検索装置100が用いるデータベースの第一データテーブルのデータ構造を示す説明図である。

図14に示す第一データテーブルは、上位項目W, X, Yが関連づけられた商品名a, b, c(関係項目)に、当該上位項目W, X, Yを関連づけたデータである。本変形例の第一データテーブルは、上述した実施形態1の図4に示した第一データテーブルから上位項目Zに対応するレコードが除外されたものに相当する。

【0070】

図15は、本変形例における情報検索装置100が用いるデータベースの第二データテーブルのデータ構造を示す説明図である。

図15に示す第二データテーブルは、上位項目W, X, Yが関連づけられた商品名a, b, c(関係項目)に關係する顧客識別番号2, 4, 6, 7, ...に、当該上位項目W, X, Yを関連づけたデータである。本変形例の第二データテーブルは、上述した実施形態1の図6に示した第二データテーブルから上位項目Zに対応するレコードが除外されたものに相当する。

【0071】

本変形例における検索処理の基本的な流れは上述した実施形態1と同様である。つまり、検索する商品名が入力された場合、演算部101aは、検索用プログラムを実行し、まず、図14に示す第一データテーブルを参照し、指定された商品名に対応する上位項目が存在するか否かを判断する。ただし、指定された商品名に対応する上位項目が存在する場合には、更に、図12に示す第四データテーブルを参照し、指定された商品名に対応する上位項目についての第二上位項目が存在するか否かを判断することになる。そして、この判断で第二上位項目が存在するとされた場合には、演算部101aは、図13に示す第五データテーブルの全体から、当該第二上位項目に関連づけられた顧客識別番号を検索することになる。

【0072】

本変形例のデータベースでは、上述した実施形態1よりも更にデータテーブルを細分化(階層化)して管理することにより、個々のデータテーブルのレコード数を削減して、検索処理を高速化できる。また、本変形例によれば、上述した実施形態1よりも検索時のOR条件の項目数を更に削減して、検索処理を高速化することが可能となる。更に、本変形例によれば、AND条件での検索処理において、指定されたすべての商品名が第二上位項目に関連づけられている場合、図13に示す第五データテーブルに対して、当該第二上位項目に関連づけられた顧客識別番号(検索対象項目)を単一条件で検索するだけで、検索処理を完了することができ、特に高速な検索処理が実現される。

【0073】

なお、本変形例のようにグループ化の階層が複数段にわたる場合には、検索処理においては、図16に示すようなデータテーブルを利用するようにしてもよい。図16に示すデータテーブルは、検索されるデータテーブルと階層とを関連づけたものであり、階層の数字が大きいほど、より上の上位項目と検索対象項目とを関連づけたデータテーブルであることを示している。

【0074】

図16のようなデータテーブルを利用することは、特に新規データの追加によるデータベースの更新時に特にメリットがある。具体的には、新規データの追加により上位項目作成条件に応じて新たな上位項目あるいは新たな第二上位項目が増えていくと、上位項目と検索対象項目とを関連づけた第二データテーブルや、第二上位項目と検索対象項目とを関連づけた第五データテーブルのレコード数が増大してく。その結果、これらのデータテーブルに対する検索処理時のレコード数が増えて検索処理の時間がかかるようになってくる。このような場合には、2つ以上の第二上位項目を更に第三の上位項目でグループ化し、これらの第二上位項目と第三上位項目とを関連づけた新たなデータテーブル（第六テーブル）を作成するとともに、第三上位項目が第二上位項目を介して関連づけられた関係項目に  
10

【0075】

このように新たなデータテーブルを増やす場合、通常は、これに対応して検索用プログラムの内容を変更することが必要になるが、図16に示すようなデータテーブルを利用することで、検索用プログラムの内容変更を伴わずに、新たなデータテーブルを増やすことが可能となる。すなわち、検索用プログラムが図16に示すようなデータテーブルを参照してデータベースのデータ構造を特定するプログラム内容とすることで、新たなデータテーブルが増えたときに図16に示すデータテーブルを更新するにすれば、検索用プログラムの内容を変えずに、新たに増えたデータテーブルも利用した検索処理が可能となる。  
20

【0076】

次に、本実施形態1における効果確認試験について説明する。

図17は、本効果確認試験における、顧客識別番号（検索対象項目）と、その顧客が過去に購入した商品の商品名（関係項目）との関係をまとめた表の一例である。

本効果確認試験は、商品名（関係項目）が商品名aと商品名bの2つであり、顧客識別番号（検索対象項目）が1～60000の6万個であり、商品名aは顧客識別番号1～40000の4万個に関連づけられており、商品名bは顧客識別番号20001～60000の4万個に関連づけられている例で行った。また、本効果確認試験では、2つの商品名a, b（関係項目）を上位項目Wにグループ化した。  
30

【0077】

図18は、本効果確認試験における商品名a, b（関係項目）と上位項目Wとを関連づけた第一データテーブルであり、上述した実施形態1における図4に示す第一データテーブルに相当する。

図19は、本効果確認試験における、顧客識別番号1～60000（検索対象項目）と、商品名a, b（関係項目）又は上位項目Wとの関係をまとめたデータテーブルであり、図5に示す表のデータテーブルに相当する。

図20は、顧客識別番号1～60000（検索対象項目）のうち、上位項目Wが関連づけられた2つの商品名a, b（関係項目）の全部に  
40

関係する顧客識別番号20001～40000に、当該上位項目Wを関連づけたデータであり、上述した実施形態1における図6に示す第二データテーブルに相当する。  
図21は、顧客識別番号1～60000（検索対象項目）のうち、同一の上位項目Wに関連づけられている商品名a, b（関係項目）の一部のみに関係する顧客識別番号1～20000, 40001～60000、及び、同一の上位項目Wが関連づけられた商品名a, b（関係項目）のいずれにも関係しない顧客識別番号に、a, b以外の商品名（関係項目）を関連づけたデータ（ただし、本例ではそのようなデータは存在しない。）であり、上述した実施形態1における図7に示す第三データテーブルに相当する。

【0078】

本効果確認試験では、図17に示す表のようなデータテーブルに対して検索処理を行うケース1と、図19に示すデータテーブルに対して検索処理を行うケース2と、図20及び図21に示す2つのデータテーブルに対して検索処理を行うケース3（実施形態1）と  
50

の比較試験を行った。

本効果確認試験では、商品名 a に関する顧客識別番号を検索し、その件数を出力する単一条件の検索条件 ( 1 ) と、商品名 a 又は商品名 b に関する顧客識別番号を検索し、その件数を出力する O R 条件の検索条件 ( 2 ) と、商品名 a 及び商品名 b に関する顧客識別番号を検索し、その件数を出力する A N D 条件の検索条件 ( 3 ) とを行った。そして、いずれの検索条件 ( 1 ) ~ ( 3 ) も、それぞれ 1 0 0 回連続で繰り返し行ったときの時間を検索時間として計測した。

【 0 0 7 9 】

なお、本効果確認試験のその他の主な条件は以下のとおりである。

情報検索装置：富士通 SELSIUS W510

O S : Windows10

D B M S : MySQL Server 5.1.40-community

【 0 0 8 0 】

本効果確認試験の結果は、下記の表 4 に示すとおりである。

【表 4】

	検索条件(1)	検索条件(2)	検索条件(3)
ケース1	3.57秒	4.06秒	7.74秒
ケース2	3.03秒	3.17秒	2.67秒
ケース3	2.79秒	2.99秒	0.92秒

【 0 0 8 1 】

いずれの検索条件 ( 1 ) ~ ( 3 ) であっても、ケース 2 及びケース 3 では、ケース 1 よりも検索時間が短縮されている。これは、ケース 1 では図 1 7 に示すデータテーブルのレコード数の 8 万個であるのに対し、ケース 2 では図 1 9 に示すデータテーブルのレコード数の 6 万個、ケース 3 では図 2 0 及び図 2 1 に示す 2 つのデータテーブルのレコード数を合わせても 6 万個であり、主に、検索するレコード数の削減による効果であると言える。

【 0 0 8 2 】

また、検索条件 ( 1 ) の場合、ケース 2 では、図 1 8 に示す第一データテーブルを参照して、図 1 9 に示すデータテーブルに対して商品名 a と上位項目 X の O R 条件での検索処理を行う。これに対し、ケース 3 では、図 1 8 に示す第一データテーブルを参照して、図 2 0 に示す第二データテーブルに対して上位項目 X の単一条件での検索処理を行うとともに、図 2 1 に示す第三データテーブルに対して商品名 a の単一条件での検索処理を行う。

【 0 0 8 3 】

このように、検索条件 ( 1 ) の場合、ケース 2 とケース 3 とでは、検索するレコード数はいずれも 6 万個であるが、ケース 2 では O R 条件での検索処理を行う結果、O R 条件のオーバーヘッドが生じるのに対し、ケース 3 では O R 条件のオーバーヘッドが生じない。これがケース 2 よりもケース 3 の方が検索時間を短縮できている理由の 1 つであると考えられる。

【 0 0 8 4 】

また、検索条件 ( 2 ) の場合、ケース 2 では、図 1 8 に示す第一データテーブルを参照して、図 1 9 に示すデータテーブルに対して商品名 a と商品名 b と上位項目 X の O R 条件での検索処理を行う。これに対し、ケース 3 では、図 1 8 に示す第一データテーブルを参照して、図 2 0 に示す第二データテーブルに対して上位項目 X の単一条件での検索処理を行うとともに、図 2 1 に示す第三データテーブルに対して商品名 a と商品名 b の O R 条件での検索処理を行う。

【 0 0 8 5 】

このように、検索条件 ( 2 ) の場合、ケース 2 とケース 3 とでは、検索するレコード数はいずれも 6 万個であり、ともに O R 条件での検索処理を行うため、O R 条件のオーバーヘッドが生じる。ただし、ケース 3 では、O R 条件の検索処理を行う第三データテーブルのレコード数が 4 万個であり、ケース 2 が O R 条件の検索処理を行うデータテーブルのレ

10

20

30

40

50

コード数（6万個）よりも少ない。更に、ケース3では、OR条件の項目数が商品名aと商品名bの2項目であり、OR条件の項目数が商品名aと商品名bと上位項目Xの3項目であるケース2よりも少ない。これらの理由から、ケース2よりもケース3の方が検索時間を短縮できていると考えられる。

【0086】

また、検索条件（3）の場合、ケース2では、図18に示す第一データテーブルを参照して、図19に示すデータテーブルに対して上位項目Xの単一条件での検索処理を行う。これに対し、ケース3では、図18に示す第一データテーブルを参照して、図20に示す第二データテーブルに対して上位項目Xの単一条件での検索処理を行い、図21に示す第三データテーブルに対する検索処理は行わない。

10

【0087】

このように、検索条件（3）の場合、ケース2では、単一条件で検索するレコード数が6万個であるが、ケース3では、単一条件で検索するレコード数が2万個で済む。これがケース2よりもケース3の方が検索時間を大幅に短縮できている理由の1つであると考えられる。

【0088】

上述したケース2によれば、ケース1と比べて、検索するレコード数が常に $3/4$ （ $= 60000 / 80000$ ）と少ないものとなる。したがって、ケース2は、ケース1と比べて、理論的には検索時間が約 $3/4$ になる。

【0089】

一方、ケース3の場合は、検索条件によって検索するレコード数が変わってくるので、検索時間の短縮効果も検索条件によって変わってくる。

20

【0090】

例えば、図18に示す第一データテーブルを参照して、図20に示す第二データテーブルを検索する必要がない検索条件である場合がある。この場合、図21に示す第三データテーブルだけを検索すればよく、検索するレコード数は、図21に示す第三データテーブルのレコード数である4万個となり、ケース2と比べても検索時間は理論上 $2/3$ （ $= 40000 / 60000$ ）となる。

【0091】

一方、例えば、図18に示す第一データテーブルを参照して、図21に示す第三データテーブルを検索する必要がない検索条件である場合がある。この場合、図20に示す第二データテーブルだけを検索すればよく、検索するレコード数は、図20に示す第二データテーブルのレコード数である2万個となり、ケース2と比べても検索時間は理論上 $1/3$ （ $= 20000 / 60000$ ）となる。

30

【0092】

他方、例えば、図18に示す第一データテーブルを参照して、図20に示す第二データテーブルも図21に示す第三データテーブルも検索する必要がある場合がある。この場合、検索するレコード数は、ケース2と同じ6万個となり、検索時間は理論上ケース2と同じとなる。しかしながら、検索時のOR条件によるオーバーヘッドはケース2よりケース3の方が少ないため、レコード数が同じ場合であっても、ケース2よりケース3の検索時間が短くなることが実験から確認された。

40

【0093】

〔実施形態2〕

次に、本発明の他の実施形態（以下、本実施形態を「実施形態2」という。）について、説明する。

本実施形態2も、前記実施形態1と同様、スーパーや百貨店などの店舗において顧客の商品購入履歴を管理するためのデータベースを用いるが、複数の顧客識別番号（関係項目）の中から指定された顧客識別番号に関連付けられた商品名（当該顧客が購入した商品）（検索対象項目）を、複数の商品名の中からコンピュータによって検索する検索処理を行う例である。したがって、前記実施形態1の例とは、関係項目と検索対象項目との関係が

50

逆の関係になっている。以下、本実施形態 2 の説明は、前記実施形態 1 と異なる点を中心に説明する。

【 0 0 9 4 】

本実施形態 2 の説明は、上述した実施形態 1 の効果確認試験で用いた図 1 7 ~ 図 2 1 を参照して行う。

すなわち、本実施形態 2 は、商品名（検索対象項目）が商品名 a と商品名 b の 2 つであり、顧客識別番号（関係項目）が 1 ~ 6 0 0 0 0 の 6 万個であり、商品名 a は顧客識別番号 1 ~ 4 0 0 0 0 の 4 万個に関連づけられており、商品名 b は顧客識別番号 2 0 0 0 1 ~ 6 0 0 0 0 の 4 万個に関連づけられている例である。また、本実施形態 2 では、2 つの商品名 a , b（検索対象項目）が上位項目 W にグループ化されている。本実施形態 2 で用いるデータテーブルは、図 2 0 に示すデータテーブル（以下「第六データテーブル」という。）と、図 2 1 に示すデータテーブル（以下「第七データテーブル」という。）である。

10

【 0 0 9 5 】

図 2 0 に示す第六データテーブルは、上述したとおり、顧客識別番号 1 ~ 6 0 0 0 0（関係項目）のうち、上位項目 W が関連づけられた 2 つの商品名 a , b（検索対象項目）の全部に関係する顧客識別番号 2 0 0 0 1 ~ 4 0 0 0 0 に、当該上位項目 W を関連づけたデータである。第六データテーブルは記憶部 1 0 1 b に記憶され、演算部 1 0 1 a は、検索用プログラムを実行して第六データテーブルから検索を行うことができる。

【 0 0 9 6 】

図 2 1 に示す第七データテーブルは、顧客識別番号 1 ~ 6 0 0 0 0（関係項目）のうち、第六データテーブルに含まれない顧客識別番号 1 ~ 2 0 0 0 0 , 4 0 0 0 1 ~ 6 0 0 0 0 に、商品名 a , b（検索対象項目）を関連づけたデータである。第七データテーブルは記憶部 1 0 1 b に記憶され、演算部 1 0 1 a は、検索用プログラムを実行して第七データテーブルから検索を行うことができる。

20

【 0 0 9 7 】

図 2 2 は、本実施形態 2 における単一条件での検索処理の流れを示すフローチャートである。

例えば、顧客識別番号 9 9 9 の単一条件で、その顧客識別番号 9 9 9 に対応する顧客が購入したことのある商品名を検索する場合を考える。入力装置 1 0 3 , 1 0 4 をユーザーが操作するなどして顧客識別番号 9 9 9 が入力された場合（S 4 1）、演算部 1 0 1 a は、検索用プログラムを実行し、まず、図 2 0 に示す第六データテーブルの全体から、顧客識別番号 9 9 9 に関連づけられた上位項目 W を単一条件で検索する（S 4 2）。

30

【 0 0 9 8 】

ここでは、顧客識別番号 9 9 9 が上位項目 W に関連づけられていないので、第六データテーブルに対する検索処理では顧客識別番号 9 9 9 に関連づけられた上位項目 W がヒットしない（S 4 3 の No）。そのため、演算部 1 0 1 a は、図 2 1 に示す第七データテーブルの全体から、顧客識別番号 9 9 9 に関連づけられた商品名を単一条件で検索する（S 4 4）。これにより、顧客識別番号 9 9 9 に関連づけられた商品名 a がヒットするので、その検索結果に係る商品名 a を出力する（S 4 6）。

【 0 0 9 9 】

一方で、もし顧客識別番号 9 9 9 が上位項目 W に関連づけられていた場合、第六データテーブルに対する検索処理において、顧客識別番号 9 9 9 に関連づけられた上位項目 W がヒットする（S 4 3 の Yes）。この場合、演算部 1 0 1 a は、図 1 8 に示す第一データテーブルを参照し（S 4 5）、上位項目 W に関連づけられている商品名 a , b を抽出して、これを検索結果として出力する（S 4 6）。

40

【 0 1 0 0 】

本実施形態 2 によれば、第六データテーブルに含まれる顧客識別番号（関係項目）と第七データテーブルに含まれる顧客識別番号（関係項目）とが排他的な関係になっている。これにより、指定された 1 つの顧客識別番号 9 9 9（関係項目）に関連づけられた上位項目が存在する場合には、図 2 0 に示す第六データテーブルに対し、その 1 つの顧客識別番

50



号 999 に関連づけられた上位項目を単一条件で検索するだけで（第七データテーブルに対する検索処理を行わずに）、検索処理を完了する。第六データテーブルは、第七データテーブルに含まれる顧客識別番号（関係項目）と商品名（検索対象項目）とが関連づけられたレコードを含まないので、検索するレコード数が少なく、検索処理速度を大幅に高速化できる。

**【 0 1 0 1 】**

また、指定された 1 つの顧客識別番号 999（関係項目）に関連づけられた上位項目が存在しない場合には、図 20 に示す第六データテーブルと図 21 に示す第七データテーブルの両方に対して検索処理を行うことになる。このときのレコード数の合計は、図 17 に示す表のようなデータテーブルに対して検索処理を行う従来の場合と比較すると少ないので検索処理速度を高速化できる。

10

**【 0 1 0 2 】**

なお、本実施形態 2 では、レコード数が相対的に少ない第六テーブルから先に検索処理を行う例となっているが、第七テーブルから先に検索処理を行うようにしてもよい。この場合には、指定された 1 つの顧客識別番号 999（関係項目）に関連づけられた上位項目が存在しないとき、図 21 に示す第七データテーブルに対し、その 1 つの顧客識別番号 999 に関連づけられた商品名を単一条件で検索するだけで（第六データテーブルに対する検索処理を行わずに）、検索処理を完了するので、検索処理速度を大幅に高速化できるようになる。

**【 0 1 0 3 】**

次に、本実施形態 2 における効果確認試験について説明する。

本効果確認試験でも、上述した実施形態 1 の効果確認試験と同じデータテーブルを用い、図 17 に示す表のようなデータテーブルに対して検索処理を行うケース 1 と、図 19 に示すデータテーブルに対して検索処理を行うケース 2 と、図 20 及び図 21 に示す 2 つのデータテーブルに対して検索処理を行うケース 3（実施形態 2）との比較試験を行った。ただし、本効果確認試験では、前記実施形態 1 の効果確認試験とは関係項目と検索対象項目との関係が逆の関係になっており、1 つの顧客識別番号から商品名を検索し、その件数を出力する単一条件の検索処理を行い、それぞれのケースで、検索処理を 100 回連続で繰り返し行ったときの時間を検索時間として計測した。

20

**【 0 1 0 4 】**

ケース 2 では、ケース 1 よりも検索時間が短縮された。これは、ケース 1 では図 17 に示すデータテーブルのレコード数の 8 万個であるのに対し、ケース 2 では図 19 に示すデータテーブルのレコード数の 6 万個であり、主に、検索するレコード数の削減によって検索処理の高速化が実現される。ケース 2 によれば、ケース 1 と比べて、検索するレコード数が  $3/4$ （ $= 60000 / 80000$ ）と少ないものとなる。なお、検索結果を出力する際には、上位項目から商品名を抽出するために図 18 に示す第一データテーブルを検索する処理が発生するが、この検索処理時間は相対的には非常に短いため、無視してよい。したがって、ケース 2 は、ケース 1 と比べて、理論的には検索時間が約  $3/4$  になる。

30

**【 0 1 0 5 】**

一方、ケース 3 の場合は、検索条件や検索処理手順によって検索するレコード数が変わってくるので、検索時間の短縮効果も検索条件や検索処理手順によって変わってくる。

40

**【 0 1 0 6 】**

まず、図 20 に示す第六データテーブルを先に検索する場合を考える。この場合、第六データテーブルを検索してヒットしたときには、図 21 に示す第七データテーブルを検索する必要がない。よって、検索するレコード数は、図 20 に示す第六データテーブルのレコード数である 2 万個となり、上位項目から商品名を抽出するために図 18 に示す第一データテーブルを検索する処理が必要になるとしても、ケース 2 と比べて検索時間は理論上  $1/3$ （ $= 20000 / 60000$ ）となる。

**【 0 1 0 7 】**

一方、第六データテーブルを検索してヒットしないときには、図 21 に示す第七データ

50

テーブルも検索する必要がある。よって、検索するレコード数は、図 20 に示す第六データテーブルと図 21 に示す第七データテーブルのレコード数の合計である 6 万個となり、ケース 2 と同じである。よって、検索時間は理論上ケース 2 と同じとなる。

【0108】

この場合のケース 2 に対するケース 3 の検索時間の短縮率の期待値は、 $1/3$  (第六データテーブルを検索してヒットする場合の確率)  $\times 1/3$  (その場合のケース 2 に対するケース 3 の検索時間の理論上の短縮率)  $+ 2/3$  (第六データテーブルを検索してヒットしない場合の確率)  $\times 1$  (その場合のケース 2 に対するケース 3 の検索時間の理論上の短縮率)  $= 7/9$  (約 78%) となる。

【0109】

次に、図 21 に示す第七データテーブルを先に検索する場合を考える。この場合、第七データテーブルを検索してヒットしたときには、図 20 に示す第六データテーブルを検索する必要がない。よって、検索するレコード数は、図 21 に示す第七データテーブルのレコード数である 4 万個となり、ケース 2 と比べて検索時間は理論上  $2/3$  ( $= 40000 / 60000$ ) となる。

【0110】

一方、第七データテーブルを検索してヒットしないときには、図 20 に示す第六データテーブルも検索する必要がある。よって、検索するレコード数は、図 21 に示す第七データテーブルと図 20 に示す第六データテーブルのレコード数の合計である 6 万個となり、ケース 2 と同じである。よって、検索時間は理論上ケース 2 と同じとなる。

【0111】

この場合のケース 2 に対するケース 3 の検索時間の短縮率の期待値は、 $2/3$  (第七データテーブルを検索してヒットする場合の確率)  $\times 2/3$  (その場合のケース 2 に対するケース 3 の検索時間の理論上の短縮率)  $+ 1/3$  (第七データテーブルを検索してヒットしない場合の確率)  $\times 1$  (その場合のケース 2 に対するケース 3 の検索時間の理論上の短縮率)  $= 7/9$  (約 78%) となる。

【0112】

本効果確認試験の測定結果では、ケース 1 が 3.14 秒、ケース 2 が 2.37 秒、ケース 3 が 1.82 秒であった。ケース 2 はケース 1 に対して検索時間が約 75% となり、ケース 1 に対して約 25% の高速化が実現された。ケース 3 はケース 2 に対して検索時間が約 77% となり、ケース 2 に対して約 23% の高速化が実現された。これらの実験結果は理論値とほぼ同様の結果である。

【符号の説明】

【0113】

- 100 情報検索装置
- 101 a 演算部
- 101 b 記憶部
- 101 装置本体
- 103, 104 入力装置

【要約】

【課題】検索処理を高速化することを課題とする。

【解決手段】複数の商品名 a, b, c, d (関係項目) の中から指定された商品名に関連付けられた顧客識別番号 1, 2, 3, ... (検索対象項目) を、複数の顧客識別番号の中からコンピュータによって検索するためのデータベースのデータ構造であって、前記複数の商品名のうちの 2 つ以上の商品名に同一の上位項目 W, X, Y, Z を関連づけた第一データと、前記複数の顧客識別番号のうち、前記同一の上位項目が関連づけられた商品名に係る顧客識別番号に、該上位項目を関連づけた第二データと、前記第一データ及び前記第二データでは関連づけられない商品名と顧客識別番号とを関連づけた第三データとを有する。

【選択図】図 8

10

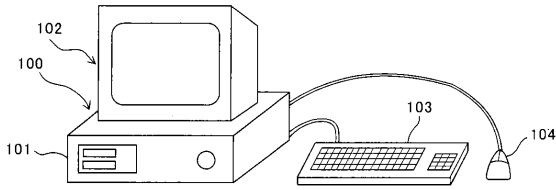
20

30

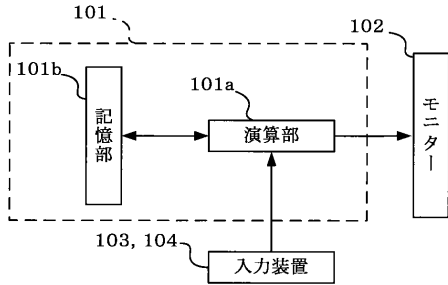
40

50

【図1】



【図2】



【図3】

顧客識別番号	商品名
1	a
1	b
1	c
2	a
2	c
3	b
4	a
4	c
5	a
.	.
.	.
.	.
10	d
.	.
.	.
.	.

【図4】

上位項目	商品名
W	a
W	b
X	a
X	c
Y	b
Y	c
Z	a
Z	b
Z	c

【図5】

顧客識別番号	上位項目 or 商品名
1	Z
2	X
3	b
4	X
5	a
.	.
.	.
.	.
10	d
.	.
.	.
.	.

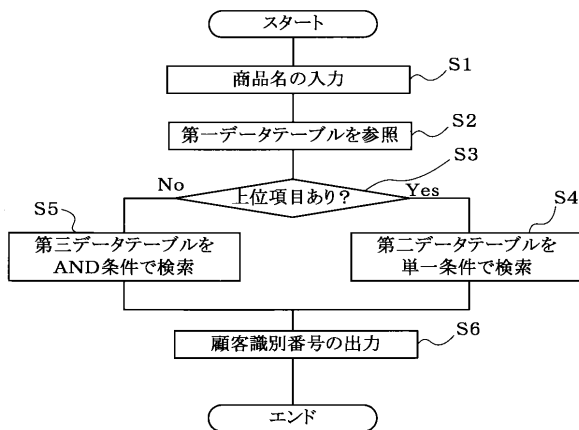
【図6】

顧客識別番号	上位項目
1	Z
2	X
4	X
6	W
7	Y
・	・
・	・
・	・

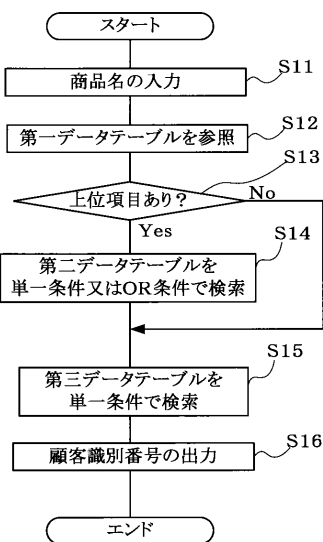
【図7】

顧客識別番号	商品名
3	b
5	a
8	c
・	・
・	・
・	・
10	d
・	・
・	・
・	・

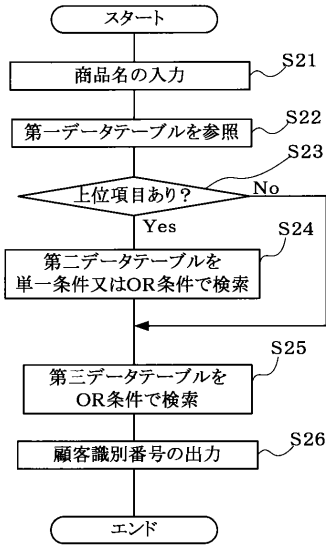
【図8】



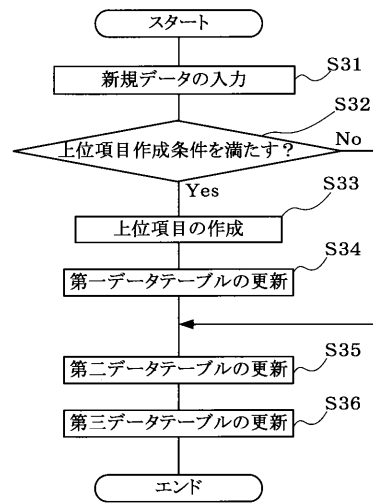
【図9】



【図10】



【図11】



【図12】

第二上位項目	上位項目
α	W
α	X
α	Y

【図13】

顧客識別番号	第二上位項目
1	α
9	α
.	.
.	.
.	.

【図15】

顧客識別番号	上位項目
2	X
4	X
6	W
7	Y
.	.
.	.
.	.

【図14】

上位項目	商品名
W	a
W	b
X	a
X	c
Y	b
Y	c

【図16】

テーブル名	階層	参照テーブル
第三データテーブル	1	
第二データテーブル	2	第一データテーブル
第五データテーブル	3	第四データテーブル

【図17】

顧客識別番号	商品名
1	a
.	.
.	.
40000	a
20001	b
.	.
.	.
60000	b

【図18】

上位項目	商品名
W	a
W	b

【図19】

顧客識別番号	商品名
1	a
.	.
.	.
20000	a
20001	W
.	.
.	.
40000	W
40001	b
.	.
.	.
60000	b

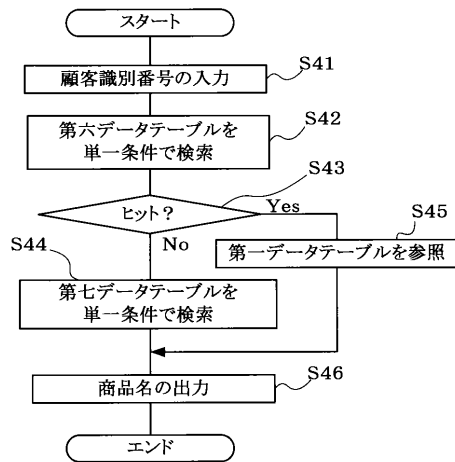
【図20】

顧客識別番号	商品名
20001	W
.	.
.	.
40000	W

【図21】

顧客識別番号	商品名
1	a
.	.
.	.
20000	a
40001	b
.	.
.	.
60000	b

【図22】



---

フロントページの続き

- (56)参考文献 特開2007-310842(JP,A)  
特開2002-157407(JP,A)  
米国特許第06622126(US,B1)  
特開平10-011469(JP,A)  
特開2001-216307(JP,A)  
国際公開第2014/167702(WO,A1)  
特開2005-099938(JP,A)  
特開2009-086815(JP,A)

- (58)調査した分野(Int.Cl., DB名)  
G06F 17/30