

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局



(43) 国際公開日
2011年2月24日(24.02.2011)

PCT

(10) 国際公開番号
WO 2011/021347 A1

- (51) 国際特許分類:
G06F 7/24 (2006.01)
- (21) 国際出願番号: PCT/JP2010/004595
- (22) 国際出願日: 2010年7月15日(15.07.2010)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2009-189602 2009年8月18日(18.08.2009) JP
- (71) 出願人 (米国を除く全ての指定国について): 株式会社エスグランツ (S.GRANTS CO., LTD.) [JP/JP]; 〒2610004 千葉県千葉市美浜区高洲三丁目5番3棟1210号 Chiba (JP).
- (72) 発明者: および
- (75) 発明者/出願人 (米国についてのみ): 新庄敏男 (SHINJO, Toshio) [JP/JP]; 〒2610004 千葉県千葉市美浜区高洲三丁目5番3棟1210号株式会社エスグランツ内 Chiba (JP). 國分光裕 (KOKUBUN, Mitsuhiro) [JP/JP]; 〒2610004 千葉県千葉市美浜区高洲三丁目5番3棟1210号株式会社エスグランツ内 Chiba (JP).

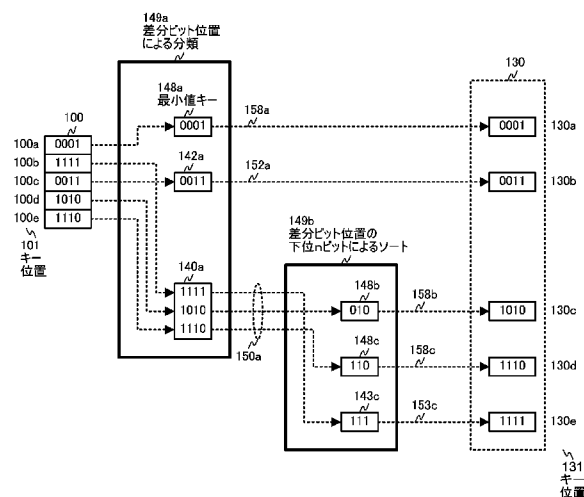
- (74) 代理人: ▲徳▼永民雄 (TOKUNAGA, Tamio); 〒2020012 東京都西東京市東町3丁目2番19号 Tokyo (JP).
- (81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.
- (84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[続葉有]

(54) Title: BIT-STRING DATA SORTING DEVICE, SORTING METHOD AND PROGRAM

(54) 発明の名称: ビット列データソート装置、ソート方法及びプログラム

[図2A]



101 KEY POSITION
131 KEY POSITION
148a MINIMUM VALUE KEY
149a CLASSIFICATION BY DIFFERENCE BIT POSITION
149b SORTING BY LOWER-ORDER n BITS OF DIFFERENCE BIT POSITION

(57) Abstract: Disclosed is an efficient sorting method in which wasted processing does not occur in sorting of bit-string data. First, bit-string comparison of a key to be a reference and a key comprising a bit-string to be sorted is sequentially performed starting from a leading bit, and a difference bit position, which is a first bit position where bit values differ between the key to be the reference and the key comprising the bit-string to be sorted, is obtained, and then the key to be sorted is classified according to the difference bit position. Next, a sorted key string is obtained by sorting, in difference bit position order, the keys that have been classified as having the same difference bit position.

(57) 要約: ビット列データのソート処理において、無効となる処理が発生しない、効率的なソート手法を提供する。初めに基準となるキーとソート対象であるビット列からなるキーのビット列比較を先頭位置のビットから順次行い、基準となるキーとソート対象であるビット列からなるキーにおいて最初に異なるビット値となるビット位置である差分ビット位置を求め、ソート対象のキーの差分ビット位置による分類を行う。次に差分ビット位置順に、同一の差分ビット位置に分類されたキーをソートすることにより、ソート済みキー列を取得する。



添付公開書類:

— 国際調査報告 (条約第 21 条(3))

明 細 書

発明の名称：

ビット列データソート装置、ソート方法及びプログラム

技術分野

[0001] 本発明は、ビット列データのソート装置、ソート方法及びプログラムに関する。

背景技術

[0002] 近年、社会の情報化が進展し、大規模なデータベースが各所で利用されるようになってきている。このような大規模なデータベースからレコードを検索するには、各レコードの記憶されたアドレスと対応づけられたレコード内の項目をインデックスキーとして検索をし、所望のレコードを探し出すことが通例である。また、全文検索における文字列も、文書のインデックスキーと見なすことができる。

[0003] そして、それらのインデックスキーはビット列で表現されることから、データベースの検索はビット列データの検索に帰着されるということができる。

一方、データベースに関連した処理として、データベース中のレコードのインデックスキーによるソート処理が行われている。このソート処理もビット列データからなるキーのソート処理に帰着される。なお、以下においては、ビット列データを、単にキーということもある。

[0004] ソートの手法は各種のものが開発されており、下記特許文献 1 には、クイックソート、ラディックソート（基数ソート）等が紹介されている。また、特許文献 2 にも基数ソートが記載されている。

[0005] 図 1 に示すのは、従来の基数ソートの概念を説明する図である。基数ソートによれば、図 1 に例示するソート対象である 4 ビットのビット列であるキーは、0 ビット目から 3 ビット目に至る各ビット位置におけるビット値による分類を繰り返すことにより、ソートが実行される。以下、図 1 の例示によ

り、基数ソートの概念を説明する。

[0006] 図1には、ソート対象であるキーからなるキー列100が示されている。図1の例示では、キー列100に含まれるキーの存在するキーの位置であるキー位置101が100a（以下、キー位置100aのように表記する。）である記憶領域にはキー“0001”が存在する。また、キー位置100b、100c、100d、100eには、それぞれキー“1111”、“0011”、“1010”、“1110”が存在する。

[0007] 図1に示すように、まずビット位置毎の分類（0ビット目）110aによりキー列100に含まれるキーの0ビット目による分類が行われる。その結果、キー“0001”とキー“0011”からなる0ビット目の値0の組120aと、キー“1111”、キー“1010”、キー“1110”からなる値1の組121aが得られる。次に値0の組120aはビット位置毎の分類（1ビット目）110bにより、次に値1の組121aはビット位置毎の分類（1ビット目）110eによりそれぞれ1ビット目の値による分類が行われる。

[0008] ビット位置毎の分類（1ビット目）110bでは、キー“0001”とキー“0011”の1ビット目が共に0であることから、1ビット目の値0の組120bしか得られず、0ビット目の値0の組120aと同じキーが2ビット目による分類の対象となり、ビット位置毎の分類（2ビット目）110cにより2ビット目による分類が行われる。ビット位置毎の分類（2ビット目）110cでは、2ビット目の値0の組120dとして1つのキー“0001”が得られるので、最小値であるキーを格納する、ソート済みキー列130のキーを格納する位置であるキー位置131が130a（以下、キー位置130aのように表記する。）である記憶領域に格納される。同様に、2ビット目の値1の組121dとして1つのキー“0011”が得られるので、ソート済みキー列130の最小値の次の値のキーを格納するキー位置130bに格納される。なお、ソート済みキー列には、キー位置の符号順に小さ

いほうのキーから格納されるものとする。

[0009] 一方、ビット位置毎の分類（１ビット目）１１０_eの分類では、１ビット目の値０の組１２０_eとして１つのキー“１０１０”が得られるので、ソート済みキー列１３０の次の格納位置であるキー位置１３０_cに格納される。また、１ビット目の値１の組１２１_eとして、キー“１１１１”とキー“１１１０”の組が得られ、ビット位置毎の分類（２ビット目）１１０_fの分類で、２ビット目の値に基づく分類が行われる。

[0010] ビット位置毎の分類（２ビット目）１１０_fでは、キー“１１１１”とキー“１１１０”の２ビット目が共に１であることから、１ビット目の値１の組１２１_fしか得られず、１ビット目の値１の組１２１_eと同じキーが３ビット目による分類の対象となり、ビット位置毎の分類（３ビット目）１１０_gにより３ビット目の値による分類が行われる。ビット位置毎の分類（３ビット目）１１０_gでは、３ビット目の値０の組１２０_gとして１つのキー“１１１０”が得られるので、ソート済みキー列１３０の次の格納位置であるキー位置１３０_dに格納される。同様に、３ビット目の値１の組１２１_gとして１つのキー“１１１１”が得られるので、ソート済みキー列１３０の次の格納位置であるキー位置１３０_eに格納される。

[0011] 以上の処理により、キー列１００のキーは、ソート済みキー列１３０のキー位置１３０_a～１３０_eにソートされて格納される。しかし、上述の基数ソート方法によりソートを実行する場合には、図１Ａに示すビット位置毎の分類（１ビット目）１１０_bやビット位置毎の分類（２ビット目）１１０_fにみられるように、分類が行われない無効な処理が発生する。

先行技術文献

特許文献

[0012] 特許文献１：特開２００２－１１６９０７号公報

特許文献２：特開２００５－３１６６６３号公報

発明の概要

発明が解決しようとする課題

- [0013] そこで本発明の解決しようとする課題は、ビット列データのソート処理において、無効となる処理の発生を少なくした、効率的なソート手法を提供することである。

課題を解決するための手段

- [0014] 本発明のソート処理によれば、基準となるビット列データである基準キーとソート対象であるビット列データであるソートキーのビット列比較を行い、最初に異なるビット値となるビット位置である差分ビット位置を求め、差分ビット位置によるソート対象のキーの分類を行い、ソートキーの差分ビット位置の下位 n ビット値によるソートを行うことにより、ソート済みキー列を取得する。ここで、ソートキーの差分ビット位置の下位 n ビット値とは、ソートキーの差分ビット位置が b とすると、差分ビット位置は 0 から始まることとしているので、ソートキーから上位 $b + 1$ ビットを除く残りのビット列のビット値である。

発明の効果

- [0015] 本発明によれば、最初に差分ビット位置によりキーの分類を行うため、分類が行われない無効な処理の発生をなくすることができる。したがって、効率的なソート処理を実現することができる。

図面の簡単な説明

- [0016] [図1]従来の基数ソートの概念を説明する図である。
- [図2A]本発明の一実施の形態における差分ビット位置と下位ビット値によるソート処理の概念を説明する図である。
- [図2B]本発明の一実施の形態におけるキーを差分ビット位置順に分類するためのデータ構造例を説明する図である。
- [図2C]本発明の一実施の形態におけるビット列データソート装置の機能ブロック構成例を説明する図である。
- [図3]本発明を実施するためのハードウェア構成例を説明する図である。

[図4A] 本発明の一実施の形態におけるソート処理の前段の処理フロー例を説明する図である。

[図4B] 本発明の一実施の形態におけるソート処理の後段の処理フロー例を説明する図である。

[図5A] 本発明の一実施の形態における各キーを差分ビット位置毎に分類する処理フロー例を説明する図である。

[図5B] 本発明の一実施の形態における各差分ビット位置に分類されたキーをソートし、ソート済みキーとして出力する処理フロー例を説明する図である。

発明を実施するための形態

[0017] 以下、本発明を実施するための形態を、図1に示したキー列100に含まれるキーと同一のキーをソート対象のキーとして例示して説明する。

[0018] 図2Aに示すのは、本発明の一実施の形態における差分ビット位置と下位ビットのビット値によるソート処理の概念を説明する図である。ソート対象のキー列100及びソート済みキー列130は、図1に示すものと同じである。

[0019] キー列100を構成する各キーは、差分ビット位置による分類149aにおいて、最小値キー148a、最小値キーとの差分ビット位置が2であるキー群142a及び最小値キー148aとの差分ビット位置が0であるキー群140aに分類される。図の例では、最小値キー148aは“0001”である。キー群142aには、キー“0011”のみが含まれており、キー群140aにはキー“1111”、“1010”及び“1110”が含まれている。最小値キー148aは、キー列100からキーを差分ビット位置によるソート処理のための入力バッファに読み込むときに求めることができる。

[0020] 最小値キー148aは、図の点線の矢印158aに示すように、ソート済みキー列130のキー位置130aに格納される。また、キー群142aのキーは1つのみであるので、図の点線の矢印152aに示すように、ソート済みキー列130のキー位置130bに格納される。

[0021] 最小値キー 1 4 8 a との差分ビット位置が 0 であるキー群 1 4 0 a にはキーが複数含まれているので、図の点線の矢印の組 1 5 0 a に示すように、差分ビット位置の下位 n ビットによるソート 1 4 9 b におけるソートの対象となる。下差分ビット位置の下位 n ビットによるソート 1 4 9 b におけるソートでは、差分ビット位置による分類 1 4 9 a で分類に用いられた差分ビット位置 0 の下位ビット、図 2 A の例では 1 ビット目から 3 ビット目までの 3 ビットを切り出してソートする。

[0022] 上述のソートにより、切り出された差分ビット位置の下位 n ビットはそれぞれ別々の組に完全に分類される。図 2 A の例では、ビット値が “0 1 0” の組 1 4 8 b、ビット値が “1 1 0” の組 1 4 8 c、ビット値が “1 1 1” の組 1 4 3 c に分類される。そして、点線の矢印 1 5 8 b、1 5 8 c、1 5 3 c にそれぞれ示すように、差分ビット位置までのビット値を付加して復元したキー “1 0 1 0”、“1 1 1 0”、“1 1 1 1” がキー位置 1 3 0 c、1 3 0 d、1 3 0 e に格納され、全体のソート処理が完了する。

[0023] なお、差分ビット位置の下位 n ビットによるソート方法、図 2 A の例示ではソート 1 4 9 b におけるソート方法には限定はなく、任意のソート方法を用いることができる。

また、上述の例では、最小値キーを、差分ビット位置を計算する基準のキーとしているが、最大値キーを基準のキーとすることもできる。この場合には、図 2 A のソート対象キーの例では、最大値キーはキー “1 1 1 1” であり、差分ビット位置による分類の結果として、差分ビット位置が 3 のキー “1 1 1 0”、差分ビット位置が 1 のキー “1 0 1 0”、差分ビット位置が 0 のキー “0 0 1 1” と “0 0 0 1” の 3 つの組が得られる。

[0024] 図 2 B は、本発明の一実施の形態における差分ビット位置によるソート処理で用いるデータ構造を説明するものである。図 2 B に示すように、キー表 3 0 9、差分ビット位置表 3 1 0、リンク表 3 1 1、ソート済みキー表 3 1 3 が用いられる。

[0025] キー表 3 0 9 はソート要求があったときにソート対象のキー 6 0 2 が読み

込まれて設定されるものである。図2Bの例示では、ソート対象の5つのキー602は、図1及び図2A示すキー列100に含まれるものが読み込まれている。キー表309の読出位置601の先頭はP0であり、以下、P1、P2、P3、P4と続いている。読出位置P0には、最小値キー“0001”が設定されているが、別領域に設定しておくことも可能である。点線の矢印689cで示すように、最小値キーはソート済みキー表313の所定の位置に格納される。

[0026] 差分ビット位置表310は、最小値キーに対する同一の差分ビット位置を有するキーのうち、キー表309から最初に読み出されたキーの読出位置601が親リンク612として差分ビット位置611毎に格納されている。図2Aの例示では、点線の矢印660aで示すように、最小値キー“0001”との差分ビット位置が0であるキー“1111”、“1010”、“1110”のうち最初に読み出されるキー“1111”の読出位置P1が差分ビット位置表310の差分ビット位置611が0である位置に格納されている。また、点線の矢印662aで示すように、最小値キー“0001”との差分ビット位置が2であるキーはキー“0011”だけであるので、キー“0011”

の読出位置P2が差分ビット位置表310の差分ビット位置611が2である位置に格納されている。なお、親リンクに設定される読出位置は、上述のように最初に読み出されるキーの読出位置に限ることはない。後述のリンク表311に格納されたリンク関係と併せて、最小値キーに対して同一の差分ビット位置を有するキーにアクセスすることができればよい。

[0027] リンク表311は、最小値キーに対する同一の差分ビット位置を有するキーにアクセスするためのものである。図に示すように、キーの読出位置621に対して、同一の差分ビット位置を有するキーの読出位置を示すリンク622が格納されている。

[0028] 差分ビット位置表310からリンク表311への点線の矢印671bで示すように、同一の差分ビット位置0を有する親リンクのキーの読出位置P1に

対して、その読出位置 P 1 を読出位置 6 2 1 とするリンク表 3 1 1 の読出位置（以下、リンク表 3 1 1 の読出位置 P 1 のようにいう。）に、最小値キー“0 0 0 1”に対して差分ビット位置 0 を有する親リンクのキー“1 1 1 1”と同一の差分ビット位置 0 を有するキー“1 1 1 0”のキー表 3 0 9 の読出位置 P 4 が格納されている。

[0029] そして、図の点線の矢印 6 7 4 c で示すように、リンク表 3 1 1 の読出位置 P 4 に、同一の差分ビット位置 0 を有するキー“1 0 1 0”のキー表 3 0 9 の読出位置 P 3 が格納されている。また、図の点線の矢印 6 7 3 c で示すように、リンク表 3 1 1 の読出位置 P 3 には、同一の値のキー表 3 0 9 の読出位置 P 3 が格納されている。これにより、それ以上同一の差分ビット位置 0 を有するキーは存在しないことを表している。

[0030] また、差分ビット位置表 3 1 0 からリンク表 3 1 1 への点線の矢印 6 7 2 b で示すように、同一の差分ビット位置 2 を有する親リンクのキーの読出位置 P 2 に対して、リンク表 3 1 1 の読出位置 P 2 には、リンク表 3 1 1 の読出位置 P 2 と同一の値のキー表 3 0 9 の読出位置 P 2 が格納されている。これは、キー表 3 0 9 の読出位置 P 2 のキー“0 0 1 1”が最小値キー“0 0 0 1”との差分ビット位置が 2 であるキーの唯一のものであることを示している。

[0031] 上述の図 2 B に示す差分ビット位置表 3 1 0 とリンク表 3 1 1 の状態は、図 2 A に示す差分ビット位置による分類 1 4 9 a が実行されて得られるものである。

以上の説明から明らかなとおり、上述のキー表 3 0 9 は請求項 1 に係る発明のソート対象キー記憶手段の一実施例である。また、差分ビット位置表 3 1 0 とリンク表 3 1 1 で分類対象キー差分ビット位置記憶手段の一実施例を構成しており、キー表 3 0 9 の読出位置 6 0 1 はキーの識別情報に相当する。

[0032] 図 2 C は、本発明の一実施の形態におけるビット列データソート装置の機能ブロック構成を説明する図である。

図に示すように、ビット列データソート装置 2 0 0 は、差分ビット位置計

算手段 210、差分ビット位置分類手段 220、下位 n ビットソート手段 225、及びソート済みキー出力手段 230 を含む。以下、各手段の動作の概要について、図 2B の例示を参照して説明する。

[0033] 差分ビット位置計算手段 210 は、キー表 309 に記憶された全てのキー 602 のうち、基準キー以外のキーについて、基準キーに対する差分ビット位置を計算する。この場合、差分ビット位置を計算する基準となるキーは、読出位置 P0 に格納された最小値キー “0001” である。ソート済みキー出力手段 230 は、最小値キー “0001” をソート済みキー表 313 に出力する。

[0034] 差分ビット位置分類手段 220 は、差分ビット位置計算手段 210 の計算結果に基づいて、差分ビット位置表 310 とリンク表 311 の書き込みを行う。その結果、キーの識別情報をキー表 309 の読出位置 601 として、差分ビット位置が 0 の組 (P3、P4、P1) と差分ビット位置が 2 の組 (P2) に分類される。

[0035] 下位 n ビットソート手段 225 は、差分ビット位置分類手段 220 により分類された各差分ビット位置の組について、差分ビット位置の下位の n ビットによりソートを実行する。

ソート済みキー出力手段 230 は、下位 n ビットソート手段 225 によるソート結果に基づいて、ソート済みのキーをソート済みキー表 313 に出力する。

[0036] 図 3 は、本発明を実施するためのハードウェア構成例を説明する図である。

本発明によるソート処理は中央処理装置 302 及びキャッシュメモリ 303 を少なくとも備えたデータ処理装置 301 によりデータ格納装置 308 を用いて実施される。キー表 309、差分ビット位置表 310、リンク表 311、ソート済みキー表 313 を有するデータ格納装置 308 は、主記憶装置 305 または外部記憶装置 306 で実現することができ、あるいは通信装置 307 を介して接続された遠方に配置された装置を用いることも可能である。

。

[0037] 図3の例示では、主記憶装置305、外部記憶装置306及び通信装置307が一本のバス304によりデータ処理装置301に接続されているが、接続方法はこれに限るものではない。また、主記憶装置305をデータ処理装置301内のものとすることもできるし、キー表309は外部記憶装置306に、他の表は主記憶装置305に持つなど、使用可能なハードウェア環境、ソート対象キーの集合の大きさ等に応じて適宜ハードウェア構成を選択できることは明らかである。

[0038] また、特に図示されてはいないが、処理の途中で得られた各種の値を後の処理で用いるためにそれぞれの処理に応じた主記憶装置305の一時記憶領域が用いられることは当然である。そして、以下の説明においては、一次記憶領域に格納されるあるいは設定される値を一時記憶領域の名前で呼ぶことがある。

[0039] 次に、本発明の一実施の形態におけるソート処理について、図4A及び図4Bを参照して詳細に説明する。また、その説明において、図2Aあるいは図2Bの例示を適宜用いて説明する。

[0040] 図4Aは、本発明の一実施の形態におけるソート処理の前段の処理フロー例を説明する図である。この前段の処理は、ソートキーを差分ビット位置毎に分類するとともに、基準キーである最小値キーをソート済みキー表に設定するものである。キー表にはソート対象のキーが記憶されており、キー表の先頭の読出位置には最小値キーが記憶されていることを前提とする。すなわち、図2Bに示すようにキー表309の読出位置P0には最小値キーが記憶されているものとする。

[0041] まず、ステップS401において、キー表の先頭位置をキー表の読出位置（以下、単に読出位置ということがある。）に設定する。ここでいう「キー表の読出位置」は、先に述べた「処理の途中で得られた各種の値を後の処理で用いるためにそれぞれの処理に応じた一時記憶領域」の1つである。以下の説明では、「キー表の先頭位置をキー表の読出位置に設定する。」のよう

に、「キー表の読出位置」が図示しない一次記憶領域であることを省略して述べることもある。

上記の設定により、図2Bの例示では、キー表の読出位置にP0が設定される。

[0042] 次にステップS402において、キー表より、先頭位置の指すキーを最小値キーとして取り出す。ここでいう「最小値キーとして取り出す」は、図示しない一次記憶領域である最小値キーに設定する、の意味である。以下においても、同様な表記を用いる場合がある。

[0043] ステップS403に進み、読出位置に次の読出位置を設定する。次にステップS404において、全てのキーは処理済みか、すなわち、全てのキーについての差分ビット位置による分類処理が終了したか判定する。

[0044] 全てのキーが処理済みでなければ、ステップS405に進み、キー表より、読出位置の指すキーをソートキーとして取り出すとともに、読出位置をソートキーの読出位置に設定する。そしてステップS406において、ソートキーと最小値キーの差分ビット位置を得て、差分ビット位置によりソートキーを分類し、ステップS403に戻る。

ステップS406の処理においては、差分ビット位置表とリンク表への設定が行われる。その処理の詳細については、後に図5Aを参照して説明する。

[0045] 一方、ステップS404で全てのキーが処理済みであると判定されると、ステップS407に分岐し、ステップS402で得た最小値キーを、ソート済みキー表に設定し、図4Bに示すステップS408以下の処理に進む。以上の処理により、図2Bの例示のように、差分ビット位置表310とリンク表311の各エントリの値が設定される。また、最小値キー“0001”がソート済みキー表313に格納される。

[0046] 図4Bは、本発明の一実施の形態におけるソート処理の後段の処理フロー例を説明する図である。この後段の処理は、順次、差分ビット位置毎に分類されたソートキーを差分ビット位置の下位nビットによりソートし、ソート

済みキーとしてソート済みキー表に出力するものである。

- [0047] ステップS 4 0 8においては、図示しない一次記憶領域である差分ビット位置に、初期値として差分ビット位置表の差分ビット位置、すなわちソート対象のキーが有する可能性のある差分ビット位置の最大値を設定する。図2 Bの例示では、差分ビット位置の最大値として3が設定される。
- [0048] ステップS 4 0 9において、差分ビット位置表の全ての差分ビット位置のエントリについて処理済みであるか判定し、処理済みであればソート処理を終了し、処理済みでなければステップS 4 1 0に進む。
- [0049] ステップS 4 1 0では、差分ビット位置表より、差分ビット位置の指すエントリから親リンクを読み出す。差分ビット位置の指すエントリに親リンクが格納されていれば、この読み出された親リンクは一次記憶領域である親リンクに設定される。すなわち、本実施の態様の説明における「親リンクを読み出す」等の表記は、親リンクを読み出して図示しない一次記憶領域である親リンクに設定することを意味することがある。親リンク以外のものに対しても同様である。
- [0050] ステップS 4 1 1では、親リンクは設定済みであるか判定する。設定済みでなければステップS 4 1 6で差分ビット位置から値を1減らしてステップS 4 0 9に戻り、一方、設定済みであればステップS 4 1 3 aに進む。
- [0051] ステップS 4 1 3 aでは、リンク表より、ステップS 4 1 0で読み出した親リンク（キー表の読出位置）の指すリンク列に含まれるキーをソートし、ソート済みキー表に設定してステップS 4 1 6に進む。ステップS 4 1 3 aの処理の詳細は、後に図5 Bを参照して説明する。
- [0052] 上述の処理を、ステップS 4 0 9において、差分ビット位置表の全ての差分ビット位置のエントリについて処理済みである、と判定されるまで繰り返し、その判定が得られるとソート対象のキーのソートが完了しているので処理を終了する。
- [0053] 図4 Bに例示する処理は、差分ビット位置を計算する基準となるキーを最小のキーとし、差分ビット位置の降順で処理を行ってソート済みキーを昇順

で取り出すものである。一方、先に述べたように、差分ビット位置を計算する基準となるキーを最大のキーとする場合でも、差分ビット位置の大きいキーほど基準キーの値に近いのであるから、差分ビット位置の降順で処理を行ってソート済みキーを降順で取り出すことができることは当業者に明らかである。

[0054] 図5Aは、図4Aに示すステップS406の処理の詳細を示すものであり、本発明の一実施の形態における各キーを差分ビット位置毎に分類する処理フロー例を説明する図である。本実施の形態においては、差分ビット位置表とリンク表を生成することにより、分類対象の各キーは差分ビット位置毎に分類される。

[0055] 図5Aに示すように、ステップS501で、ソートキーに設定されているキーと最小値キーに設定されているキーをビット列として比較し、上位0ビット目からみた最初の不一致ビットのビット位置を、差分ビット位置に設定する。ソートキーは、図4Aに示すステップS405で設定されるものであり、図5Aに示す分類処理の対象のキーである。

[0056] 次にステップS502で差分ビット位置表より、ステップS501で設定した差分ビット位置の指す親リンクを読み出し、ステップS503で親リンクは設定済みであるか判定する。

[0057] 親リンクが設定済みでなければ、ステップS504に進み、差分ビット位置表の、差分ビット位置の指す親リンクにソートキーの読出位置を設定し、ステップS505において、リンク表の、ソートキーの読出位置の指すリンクに、ソートキーの読出位置を設定して処理を終了する。

なお、ステップS504の処理におけるソートキーの読出位置は、図4Aに示すステップS405で設定されるものである。

[0058] 差分ビット位置表には、先に述べたように同一の差分ビット位置を有するキーのうち、最初に読み出されたキーの読出位置が親リンクとして格納されるが、上記ステップS504の処理は、まさに最初に読み出されたソートキーの読出位置を差分ビット位置表の差分ビット位置の指す親リンクに設定す

るものである。

- [0059] 一旦親リンクが設定済みとなった後、すなわちステップS 5 0 3での判定が親リンクは設定済みになると、ステップS 5 1 0に進み、リンク表より、親リンクの指すリンクを次リンクとして読み出し、ステップS 5 1 1で、次リンクと親リンクは同一か判定する。
- [0060] ステップS 5 1 1で次リンクと親リンクは同一であると判定されると、ステップS 5 1 2において、リンク表の親リンクの指すリンクに、ソートキーの読出位置を設定する。そして、ステップS 5 1 3において、リンク表のソートキーの読出位置の指すリンクに、ソートキーの読出位置を設定して処理を終了する。
- [0061] 一方、ステップS 5 1 1で次リンクと親リンクは同一でないと判定されると、ステップS 5 1 4に進み、リンク表の親リンクの指すリンクに、ソートキーの読出位置を設定する。そして、ステップS 5 1 5において、リンク表のソートキーの読出位置の指すリンクに、ステップS 5 1 0で得た次リンクを設定して処理を終了する。
- [0062] 上述のステップS 5 1 1の判定は、同一の差分ビット位置を有するキーのうちで分類済みのキーが1つのみであったかの判定と同じである。分類済みのキーが1つのみであればこのキーの読出位置の指すリンク表のリンクにはその読出位置が格納されている。そこで、それをソートキーの読出位置で更新し、ソートキーの読出位置の指すリンク表のリンクにソートキーの読出位置を設定してソートキーが、その時点での最後の分類済みのキーであることを示す。
- [0063] 分類済みのキーが2つ以上のときは、ステップS 5 1 4とステップS 5 1 5の処理により、リンク表で表現される親リンクと次リンクのリンク関係の間に、ソートキーの読出位置を挿入する。
- [0064] 上述の説明から明らかなとおり、同一の差分ビット位置を有するソートキーの読出位置は、差分ビット位置表に格納された親リンクからリンク表の読出位置とリンクをたどることにより順次全てアクセスすることができる。そこ

で、以下の説明において、この親リンクからたどることのできる親リンクも含めた読出位置の列を親リンクの指すリンク列といい、親リンクの指すリンク列に含まれる読出位置のキー表に記憶されたキーを、親リンクの指すリンク列に含まれるキーということがある。

[0065] 図5Bは、図4Bに示すステップS414aの処理である、本発明の一実施の形態における各差分ビット位置に分類されたキーをソートし、ソート済みキーとして出力する処理フロー例を説明する図である。本実施の形態においては、リンク表に基づいて親リンクの指すリンク列に含まれるキーを取り出してソートし、ソート済みキー表に設定する。

[0066] 図に示すように、ステップS541において、リンク表より、親リンクの指すリンク列のリンクを取り出す。親リンクは、図4Bに示すステップS410で読み出したものである。

[0067] ステップS542に進み、キー表よりステップS541で取り出したリンクの指すキーを読み出し、リンクとともにキーの差分ビット位置の下位nビットをソートキーとしてソート対象データに設定する。

[0068] 図2Bに示す例で、差分ビット位置が0の場合には、親リンクはP1、親リンクの指すリンク列のリンクは、P1、P4、P3であるから、ソート対象データには、(P1、“111”)、(P4、“110”)、(P3、“010”)の3つのデータが設定される。ソートキーは、それぞれ差分ビット位置の下位3ビットである、“111”、“110”、“010”である。

[0069] 次に、ステップS543においてソート対象データをソートする。上述の例で昇順にソートすると、ソート結果のデータの並びは、(P3、“010”)、(P4、“110”)、(P1、“111”)となる。

[0070] 最後にステップS544で、ソート順に、ソート対象データのリンクを取り出し、キー表より、リンクの指すキーを読み出し、ソート済みキー表に順次設定し、処理を終了する。

[0071] 上述の例では、まず、リンクP3が取り出され、キー表より読出位置P3の

キー“1010”が読み出されてソート済みキー表に設定される。次に、リンクP4が取り出され、キー表より読出位置P4のキー“1010”が読み出されてソート済みキー表に設定される。最後に、リンクP1が取り出され、キー表より読出位置P1のキー“1111”が読み出されてソート済みキー表に設定される。

[0072] 以上本発明を実施するための形態について詳細に説明したが、本発明の実施の形態はそれに限ることなく種々の変形が可能であることは当業者に明らかである。

例えば、図5Bに示すソート処理では、キー表よりステップS541で取り出したリンクの指すキーを読み出し、リンクとともにキーの差分ビット位置の下位nビットをソートキーとしてソート対象データに設定することにより、ソートキーのサイズを小さくしているが、キー自体をソート対象データに設定し、ソートキーをキーの差分ビット位置の下位nビットとしてソートしても同一のソート結果が得られることは明らかである。

[0073] また、リンクをソート対象データに含めず、差分ビット位置までの上位のビット値を記憶しておき、ソート後にキーの差分ビット位置の下位nビットと合成してキーを復元することが可能であることも明らかである。処理は複雑になるが、リンクに要するビット数と差分ビット位置の大きさとの関係により、ソート対象データを、リンクを含めたものとするかしないかを切り替えることも可能である。

差分ビット位置が大きい場合には、先頭から差分ビット位置までのビットをソート対象から除くことにより、ソート対象データのデータサイズを縮小することができる。

[0074] さらに、本発明のビット列データソート装置が、図2Bに例示するデータ構造を有するデータを記憶する記憶手段と図4A及び図4Bに示す処理をコンピュータに実行させるプログラムによりコンピュータ上に構築可能なことは明らかである。

したがって、上記プログラム、及びプログラムを記録したコンピュータ読

み取り可能な記録媒体は、本発明の実施の形態に含まれる。

以上詳細に説明した本発明が提供する新しいビット列データソート手法を用いることにより、より高速なビット列データのソートを行うことが可能となる。

請求の範囲

[請求項1]

ビット列で表されるソート対象のキーのソート処理を行うビット列データソート装置において、

前記ソート対象のキーを記憶するソート対象キー記憶手段と、

前記ソート対象キー記憶手段に記憶された前記ソート処理を構成する分類処理の分類対象である前記キーと、該分類対象であるキーのうち基準となる基準キーとのビット列比較により、最初に異なるビット値となる差分ビット位置を求める差分ビット位置計算手段と、

前記差分ビット位置計算手段で求めた差分ビット位置毎に、該差分ビット位置を有するキーの識別情報を記憶する分類対象キー差分ビット位置記憶手段と、

前記差分ビット位置計算手段で求めた前記差分ビット位置毎に該差分ビット位置を有するキーの識別情報を前記分類対象キー差分ビット位置記憶手段に書き込むことにより、前記分類対象であるキーを同一の差分ビット位置を有する組に分類する差分ビット位置分類手段と、

前記差分ビット位置分類手段で分類された同一の差分ビット位置を有するキーの組毎に差分ビット位置順で各キーの当該差分ビット位置の下位 n ビットによるソートを行う下位 n ビットソート手段と、

前記基準キーと、前記下位 n ビットソート手段により差分ビット位置順で差分ビット位置毎にソートされたキーとを、ソート済みキーとして出力するソート済みキー出力手段と、

を備えることを特徴とするビット列データソート装置。

[請求項2]

請求項 1 に記載のビット列データソート装置において、

前記分類対象であるキーのうち基準となるキーを、分類対象であるキーのうち最小の値をとる最小値キーあるいは最大の値をとる最大値キーとすることを特徴とするビット列データソート装置。

[請求項3]

請求項 2 に記載のビット列データソート装置において、

前記キーを識別する情報は、該キーが記憶された前記ソート対象キ

一記憶手段における該キーのアドレスである読出位置であり、
前記分類対象キー差分ビット位置記憶手段は、前記分類対象であるキーと該分類対象であるキーのうち基準となるキーである前記最小値キーあるいは最大値キーとの差分ビット位置毎に、該最小値キーあるいは最大値キーに対する同一の差分ビット位置を有するキーのうちの任意のキーの前記読出位置である親リンクを格納する差分ビット位置表と、前記読出位置毎に該読出位置のキーと前記同一の差分ビット位置を有するキーの読出位置であるリンクを格納するリンク表から構成されることを特徴とするビット列データソート装置。

[請求項4]

請求項3に記載のビット列データソート装置において、

前記同一の差分ビット位置を有するキーの前記ソート対象キー記憶手段の読出位置をリンク表の読出位置とするリンクは、該リンクをリンク表の読出位置とし、さらに該読出位置に設定されたリンクを順次たどることにより、前記同一の差分ビット位置を有する全てのキーの前記ソート対象キー記憶手段の読出位置を参照するために設定されており、該同一の差分ビット位置を有するキーのうちの最後のキーの読出位置のリンクには、該最後のキーの読出位置が格納されていることを特徴とするビット列データソート装置。

[請求項5]

請求項4に記載のビット列データソート装置において、

前記下位 n ビットソート手段は、前記差分ビット位置順に、前記差分ビット位置表の当該差分ビット位置に格納された親リンクと該親リンクの指す前記ソート対象キー記憶手段に記憶されたキーの当該差分ビット位置の下位 n ビットをソートキーとするソート対象データとし、該親リンクと同一の差分ビット位置を有する全てのキーのリンクと該リンクの指す前記ソート対象キー記憶手段に記憶されたキーの当該差分ビット位置の下位 n ビットをソートキーとするソート対象データとして、前記ソート対象データをソートすることを前記差分ビット位置順に実行するものであり、

前記ソート済みキー出力手段は、差分ビット位置順に、前記下位 n ビットソート手段によりソートされたソート対象データの前記リンクの指す前記ソート対象キー記憶手段からキーを読み出してソート済みキーとして出力する、
ことを特徴とするビット列データソート装置。

[請求項6]

ビット列で表されるソート対象のキーのソート処理を行うビット列データソート方法において、

前記ソート対象のキーを記憶するソート対象キー記憶手段に前記ソート対象のキーを記憶するソート対象キー記憶ステップと、

前記ソート対象キー記憶手段に記憶された前記ソート処理を構成する分類処理の分類対象である前記キーと、該分類対象であるキーのうち基準となる基準キーとのビット列比較により、最初に異なるビット値となる差分ビット位置を求める差分ビット位置計算ステップと、

前記差分ビット位置計算ステップで求めた前記差分ビット位置毎に、前記分類対象キー差分ビット位置記憶手段に該差分ビット位置を有するキーの識別情報を書き込むことにより、前記分類対象であるキーを同一の差分ビット位置を有する組に分類する差分ビット位置分類ステップと、

前記差分ビット位置分類ステップで分類された同一の差分ビット位置を有するキーの組毎に差分ビット位置順で各キーの当該差分ビット位置の下位 n ビットによるソートを行う下位 n ビットソートステップと、

前記基準キーと、前記下位 n ビットソートステップにおいて差分ビット位置順で差分ビット位置毎にソートされたキーとを、ソート済みキーとして出力するソート済みキー出力ステップと、
を備えることを特徴とするビット列データソート方法。

[請求項7]

請求項6に記載のビット列データソート方法において、

前記分類対象であるキーのうち基準となるキーを、分類対象である

キーのうち最小の値をとる最小値キーあるいは最大の値をとる最大値キーとすることを特徴とするビット列データソート方法。

[請求項8]

請求項7に記載のビット列データソート方法において、

前記キーを識別する情報は、該キーが記憶された前記ソート対象キー記憶手段における該キーのアドレスである読出位置であり、
前記分類対象キー差分ビット位置記憶手段は、前記分類対象であるキーと該分類対象であるキーのうち基準となるキーである前記最小値キーあるいは最大値キーとの差分ビット位置毎に、該最小値キーあるいは最大値キーに対する同一の差分ビット位置を有するキーのうちの任意のキーの前記読出位置である親リンクを格納する差分ビット位置表と、前記読出位置毎に該読出位置のキーと前記同一の差分ビット位置を有するキーの読出位置であるリンクを格納するリンク表から構成されることを特徴とするビット列データソート方法。

[請求項9]

請求項8に記載のビット列データソート方法において、

前記同一の差分ビット位置を有するキーの前記ソート対象キー記憶手段の読出位置をリンク表の読出位置とするリンクは、該リンクをリンク表の読出位置とし、さらに該読出位置に設定されたリンクを順次たどることにより、前記同一の差分ビット位置を有する全てのキーの前記ソート対象キー記憶手段の読出位置を参照するために設定されており、該同一の差分ビット位置を有するキーのうちの最後のキーの読出位置のリンクには、該最後のキーの読出位置が格納されていることを特徴とするビット列データソート方法。

[請求項10]

請求項9に記載のビット列データソート方法において、

前記下位 n ビットソートステップは、前記差分ビット位置順に、前記差分ビット位置表の当該差分ビット位置に格納された親リンクと該親リンクの指す前記ソート対象キー記憶手段に記憶されたキーの当該差分ビット位置の下位 n ビットをソートキーとするソート対象データとし、該親リンクと同一の差分ビット位置を有する全てのキーのリン

クと該リンクの指す前記ソート対象キー記憶手段に記憶されたキーの当該差分ビット位置の下位 n ビットをソートキーとするソート対象データとして、前記ソート対象データをソートすることを前記差分ビット位置順に実行し、

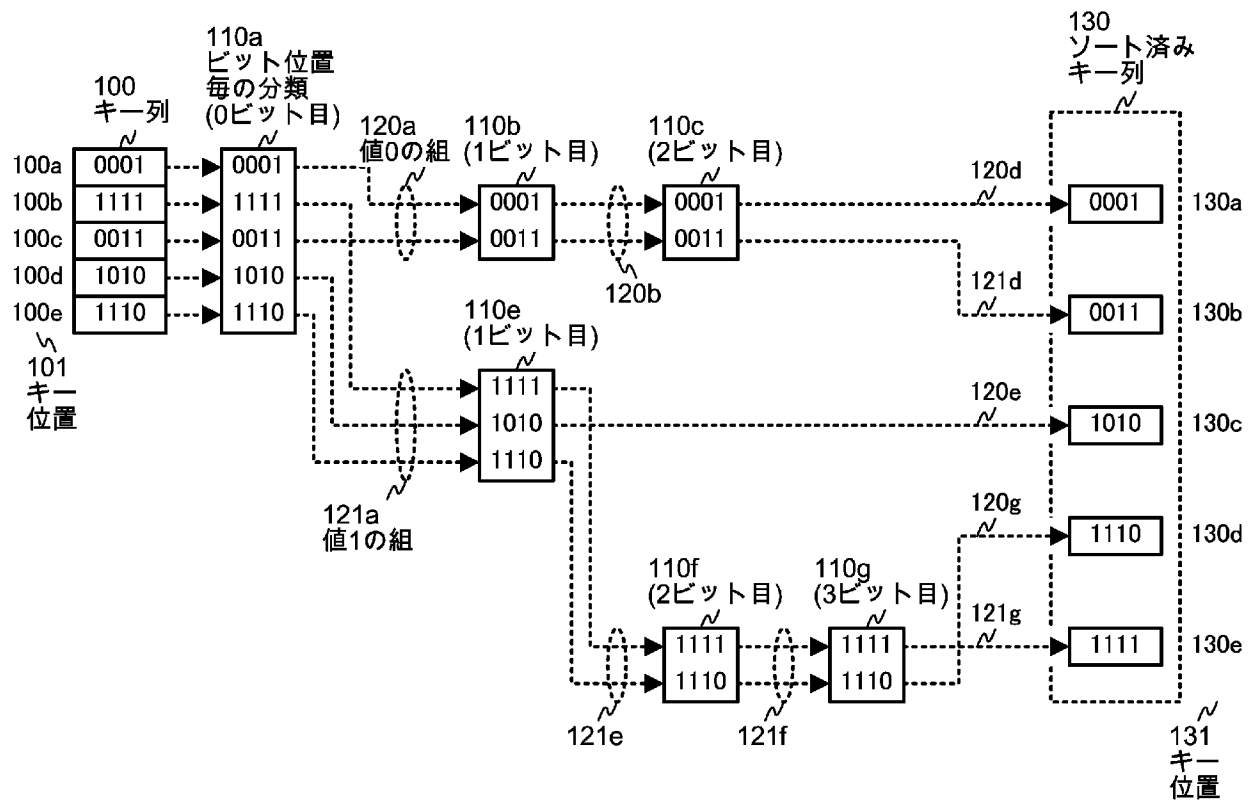
前記ソート済みキー出力ステップは、差分ビット位置順に、前記下位 n ビットソートステップにおいてソートされたソート対象データの前記リンクの指す前記ソート対象キー記憶手段からキーを読み出してソート済みキーとして出力する、

ことを特徴とするビット列データソート方法。

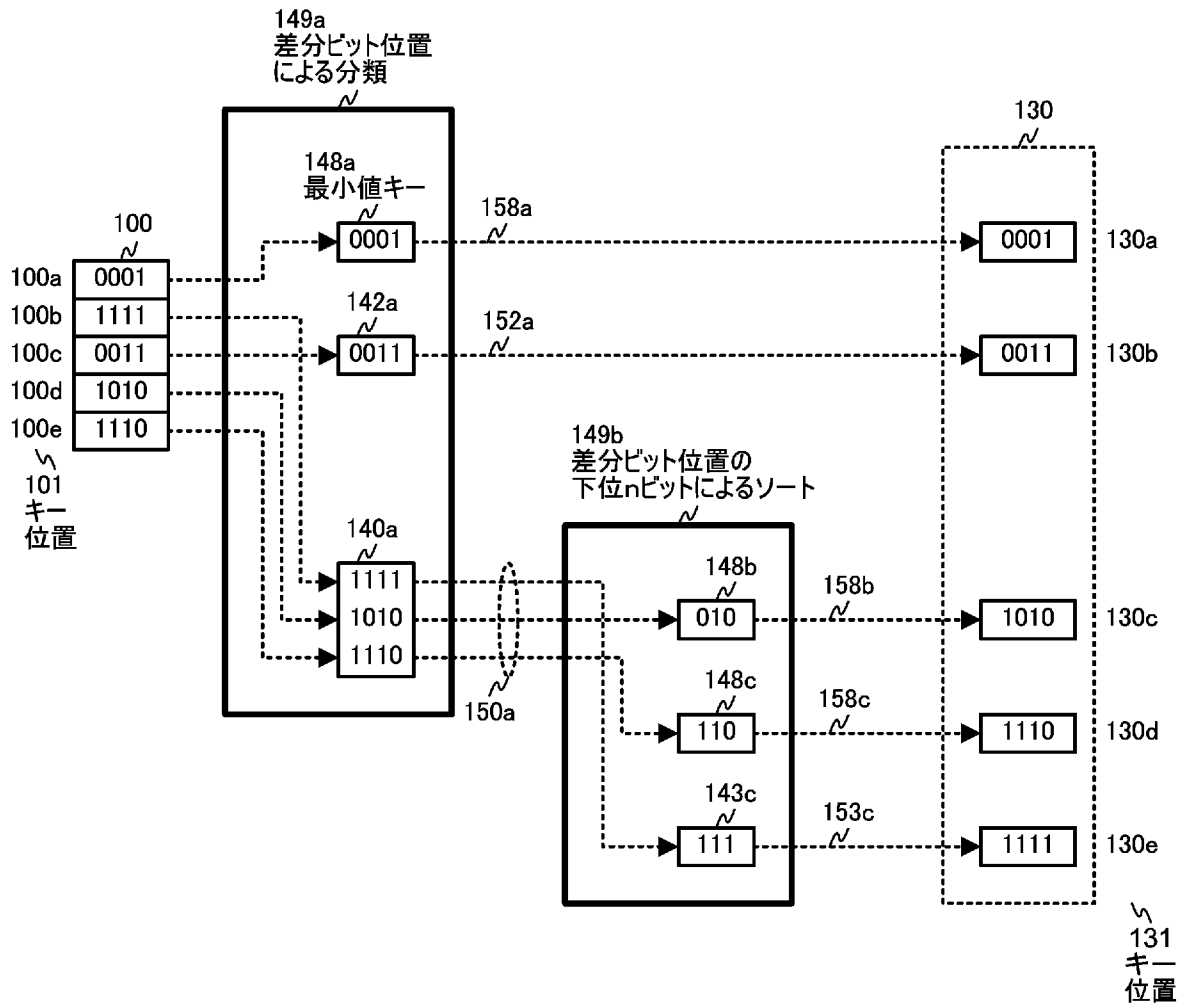
[請求項11] 請求項6～請求項10のいずれか1項に記載のビット列データソート方法をコンピュータに実行させるプログラム。

[請求項12] 請求項6～請求項10のいずれか1項に記載のビット列データソート方法をコンピュータに実行させるプログラムを記憶したコンピュータ読み取り可能な記憶媒体。

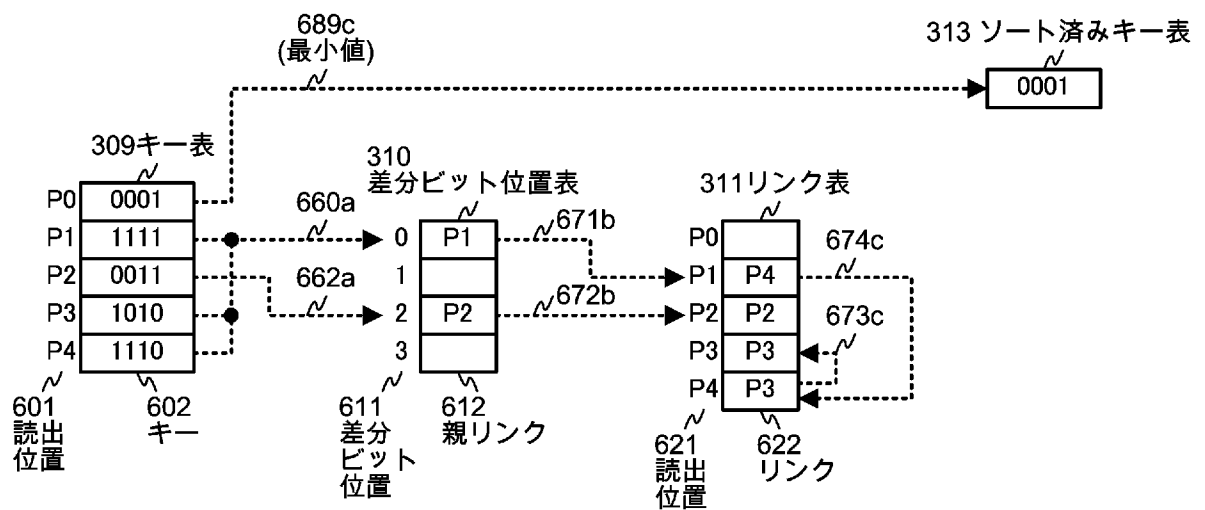
[図1]



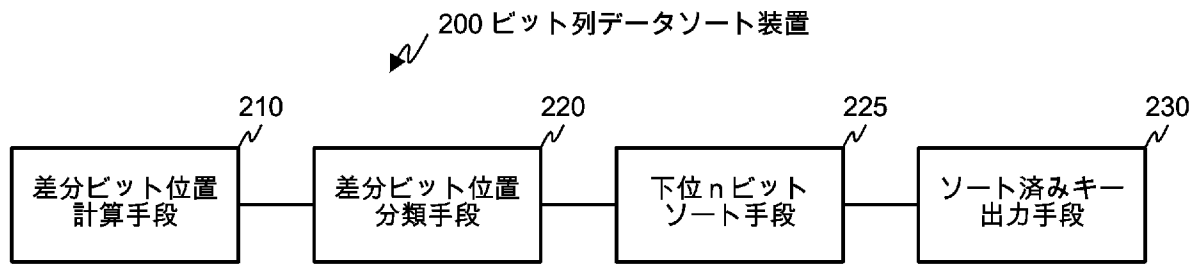
[図2A]



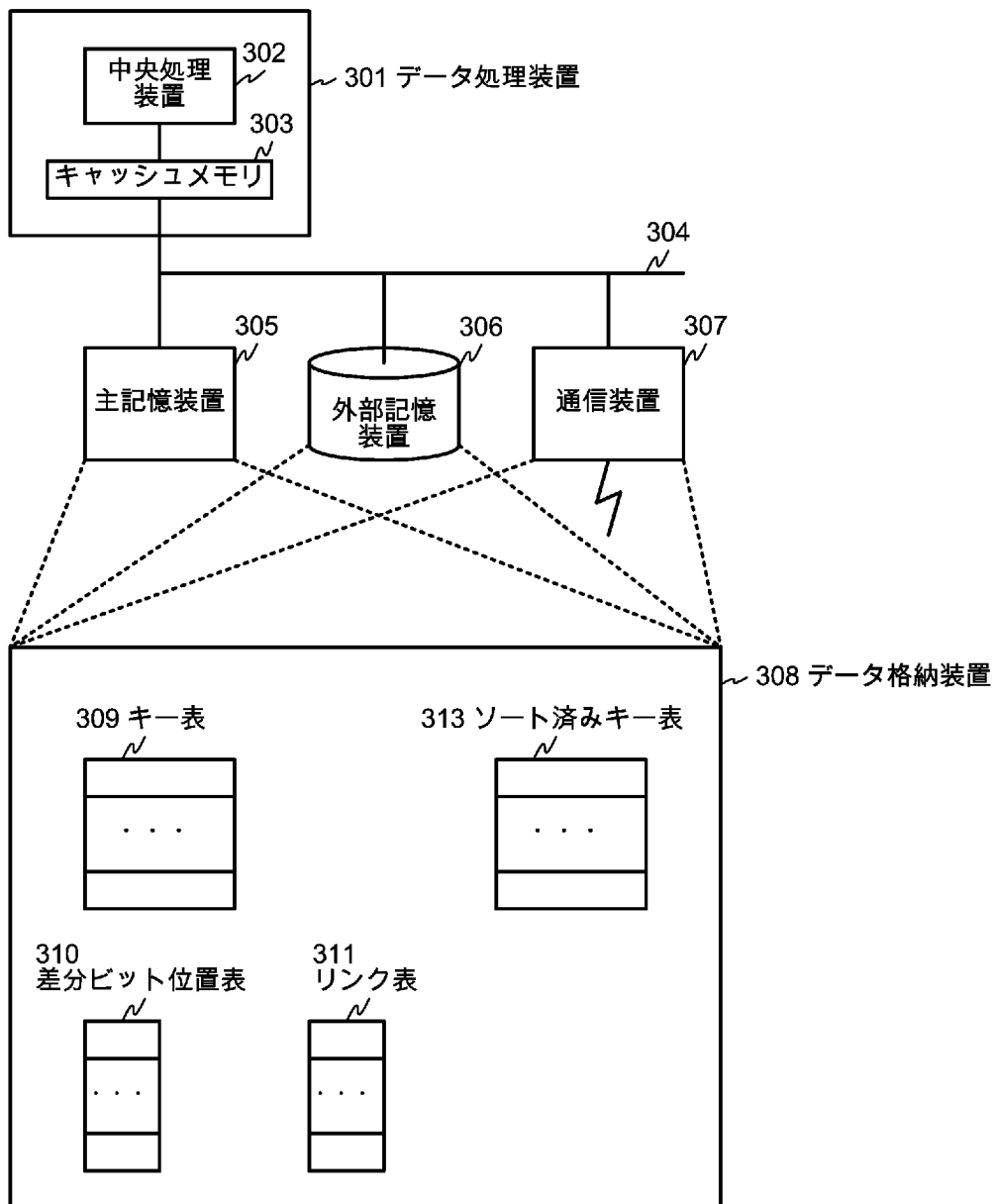
[図2B]



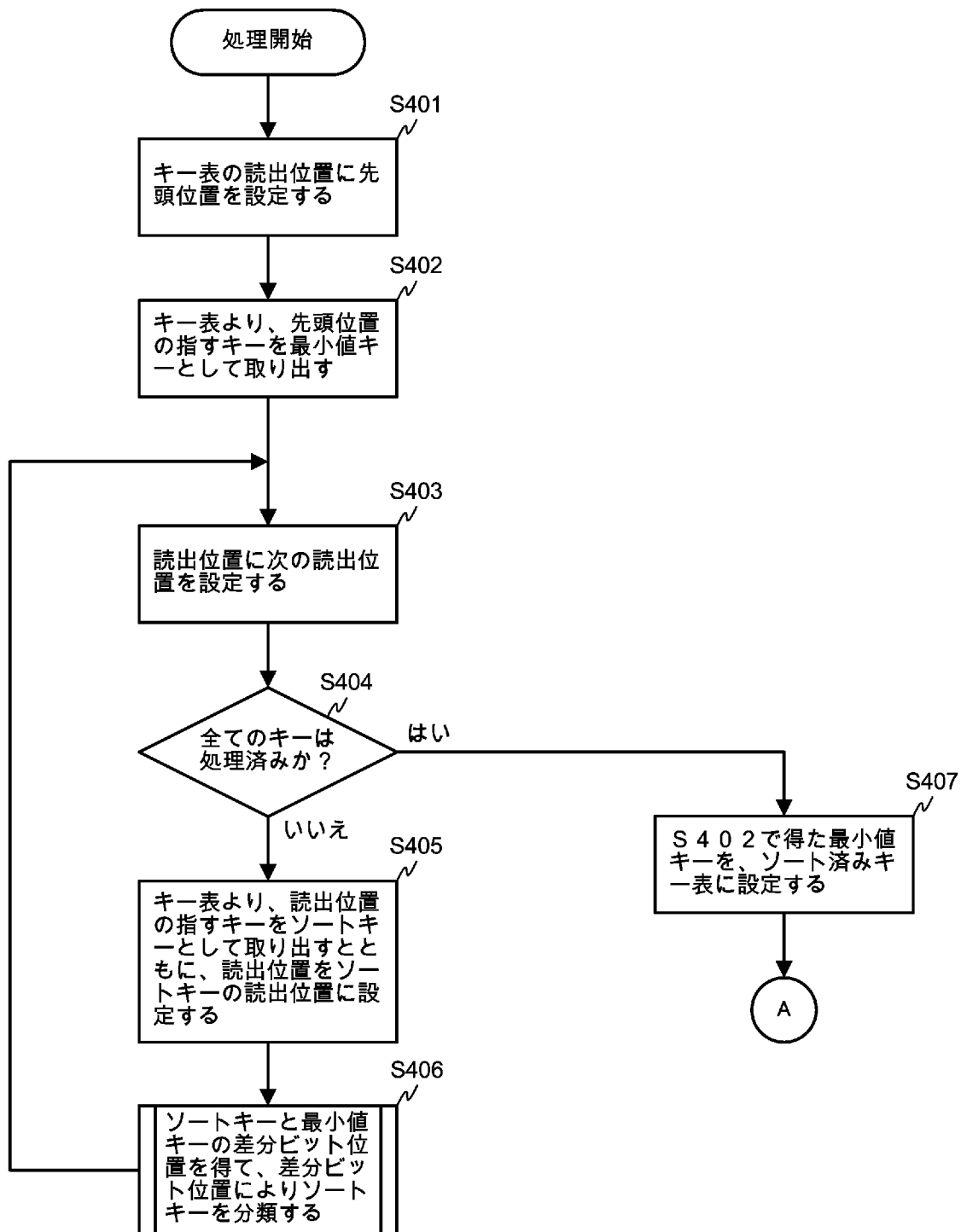
[図2C]



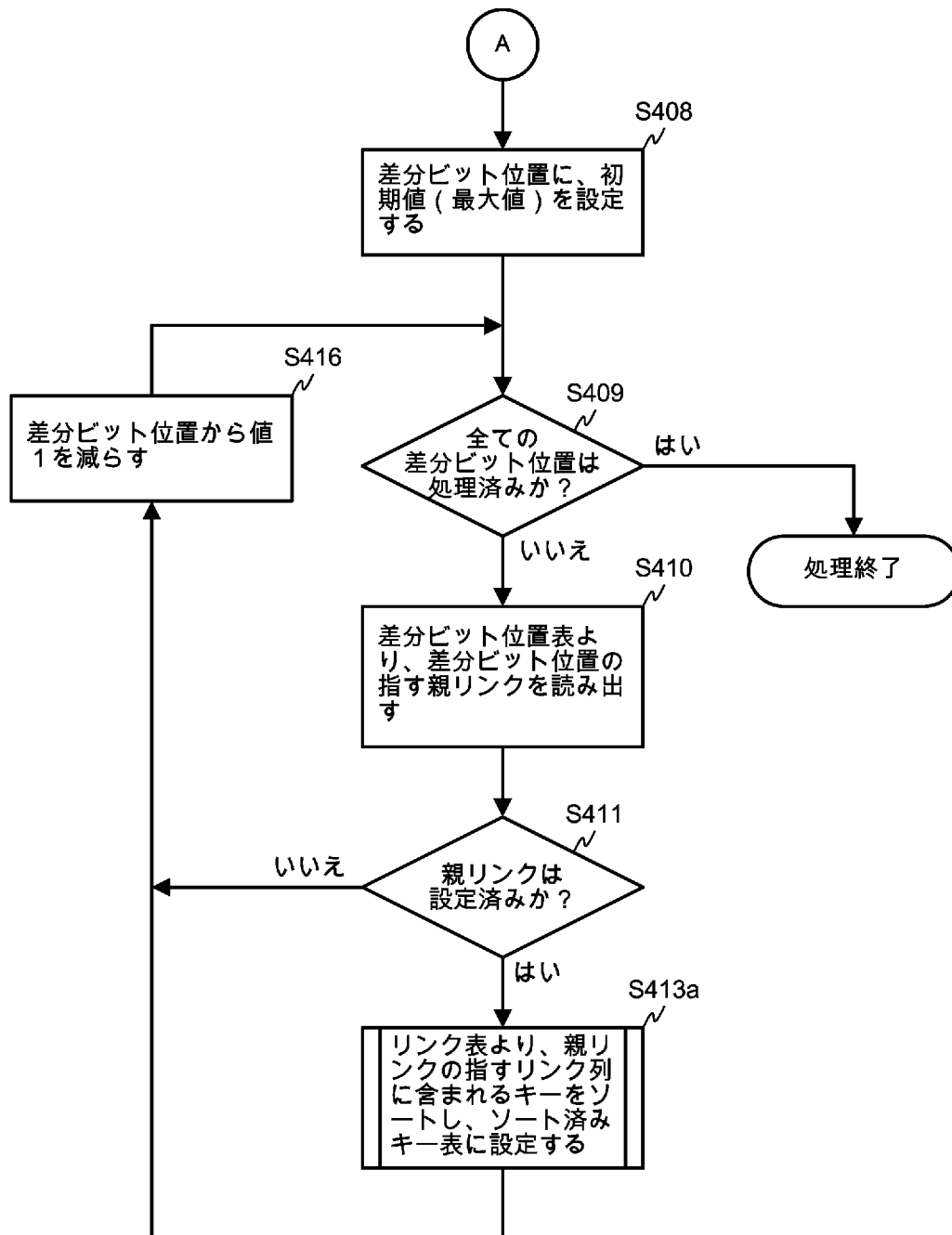
[図3]



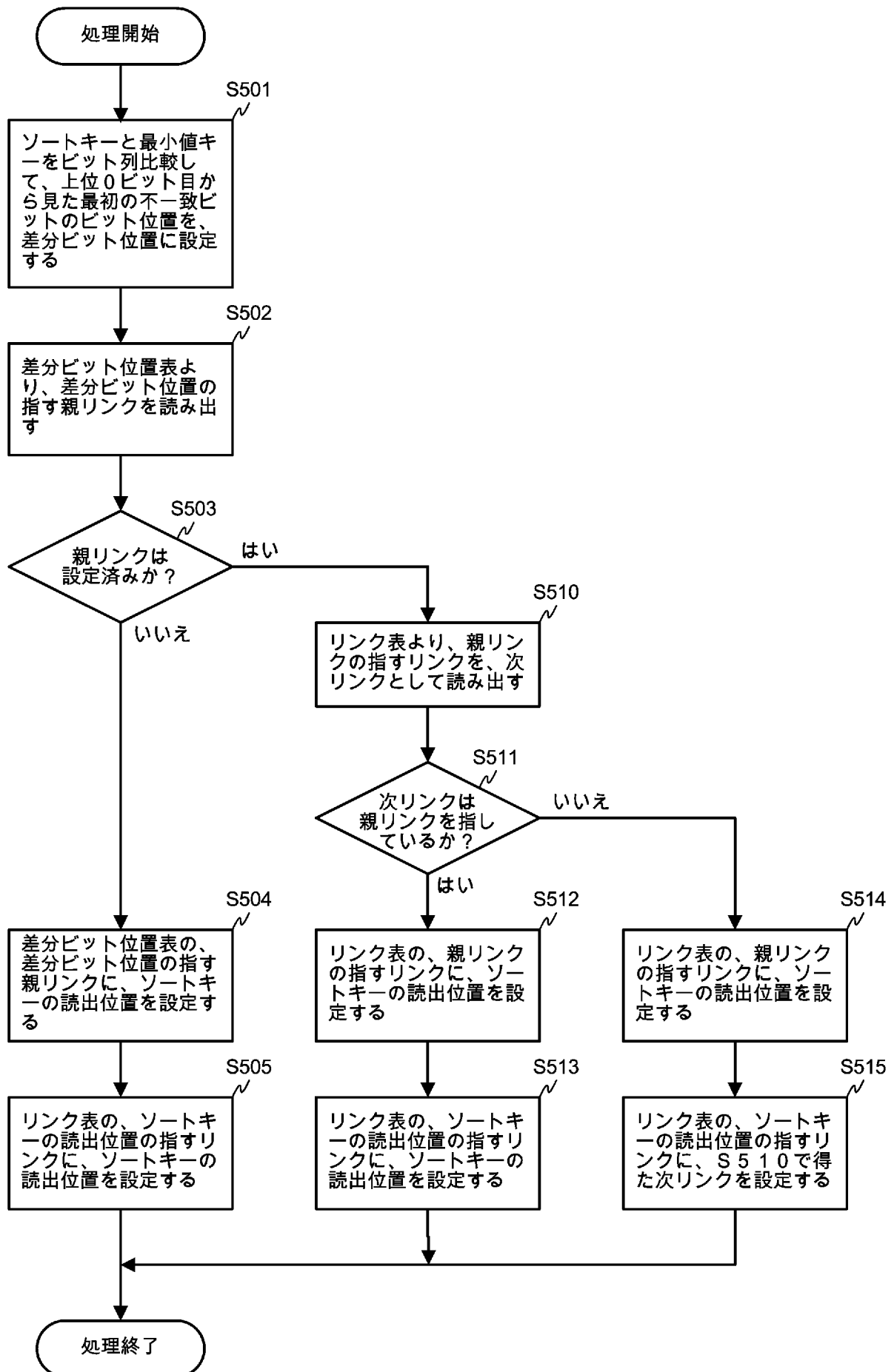
[図4A]



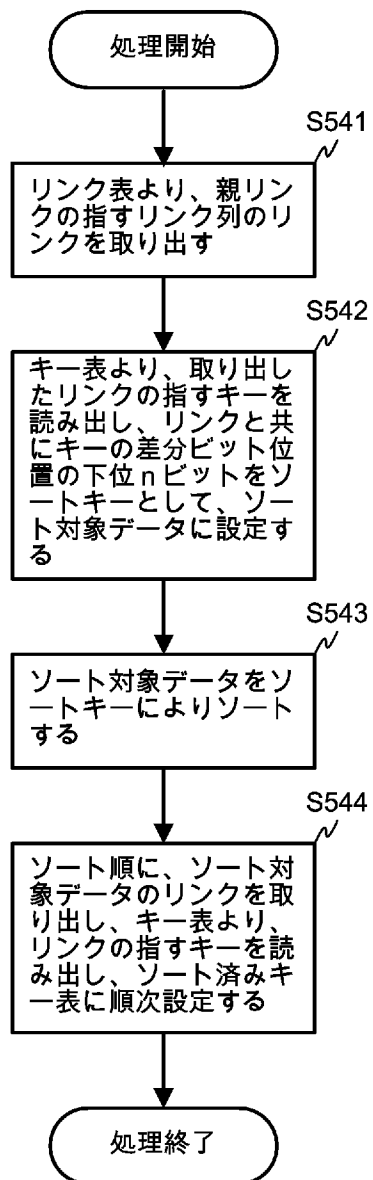
[図4B]



[図5A]



[図5B]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2010/004595

A. CLASSIFICATION OF SUBJECT MATTER

G06F7/24 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F7/24

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Jitsuyo Shinan Koho 1922-1996 Jitsuyo Shinan Toroku Koho 1996-2010

Kokai Jitsuyo Shinan Koho 1971-2010 Toroku Jitsuyo Shinan Koho 1994-2010

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P, A	WO 2010/073471 A1 (S. Grants Co., Ltd.), 01 July 2010 (01.07.2010), entire text; all drawings (Family: none)	1-12
A	US 4595995 A (AT&T Bell Laboratories), 17 June 1986 (17.06.1986), entire text; all drawings (Family: none)	1-12
A	JP 7-302187 A (Nippon Steel Corp.), 14 November 1995 (14.11.1995), entire text; all drawings (Family: none)	1-12

☒ Further documents are listed in the continuation of Box C.☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search
22 September, 2010 (22.09.10)Date of mailing of the international search report
05 October, 2010 (05.10.10)Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2010/004595

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2010-92088 A (NEC Corp.), 22 April 2010 (22.04.2010), entire text; all drawings (Family: none)	1-12
A	JP 4-127321 A (Fujitsu Ltd.), 28 April 1992 (28.04.1992), entire text; all drawings (Family: none)	1-12

A. 発明の属する分野の分類 (国際特許分類 (I P C))

Int.Cl. G06F7/24(2006.01)i

B. 調査を行った分野

調査を行った最小限資料 (国際特許分類 (I P C))

Int.Cl. G06F7/24

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1 9 2 2 - 1 9 9 6 年
日本国公開実用新案公報	1 9 7 1 - 2 0 1 0 年
日本国実用新案登録公報	1 9 9 6 - 2 0 1 0 年
日本国登録実用新案公報	1 9 9 4 - 2 0 1 0 年

国際調査で使用した電子データベース (データベースの名称、調査に使用した用語)

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
P, A	WO 2010/073471 A1 (株式会社エスグランツ) 2010.07.01, 全文、全図 (ファミリーなし)	1-12
A	US 4595995 A (AT&T Bell Laboratories) 1986.06.17, 全文、全図 (ファミリーなし)	1-12
A	JP 7-302187 A (新日本製鐵株式会社) 1995.11.14, 全文、全図 (ファミリーなし)	1-12

☒ C 欄の続きにも文献が列挙されている。☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの

「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの

「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す)

「O」口頭による開示、使用、展示等に言及する文献

「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&」同一パテントファミリー文献

国際調査を完了した日

2 2 . 0 9 . 2 0 1 0

国際調査報告の発送日

0 5 . 1 0 . 2 0 1 0

国際調査機関の名称及びあて先

日本国特許庁 (I S A / J P)

郵便番号 1 0 0 - 8 9 1 5

東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

田中 友章

5 E

9 3 7 6

電話番号 0 3 - 3 5 8 1 - 1 1 0 1 内線 3 5 2 1

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
A	JP 2010-92088 A (日本電気株式会社) 2010.04.22, 全文、全図 (ファミリーなし)	1-12
A	JP 4-127321 A (富士通株式会社) 1992.04.28, 全文、全図 (ファミリーなし)	1-12