

(19) 世界知的所有権機関
国際事務局



(43) 国際公開日
2009年5月28日 (28.05.2009)

PCT

(10) 国際公開番号
WO 2009/066442 A1

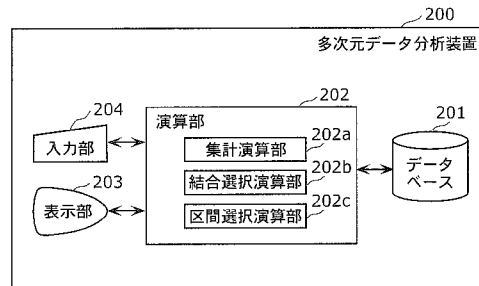
- (51) 国際特許分類:
G06F 17/30 (2006.01) G06F 19/00 (2006.01)
- (21) 国際出願番号: PCT/JP2008/003366
- (22) 国際出願日: 2008年11月18日 (18.11.2008)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2007-301025
2007年11月20日 (20.11.2007) JP
- (71) 出願人 (米国を除く全ての指定国について): 国立大学法人大阪大学 (OSAKA UNIVERSITY) [JP/JP]; 〒5650871 大阪府吹田市山田丘1番1号 Osaka (JP).
- (72) 発明者; および
(75) 発明者/出願人 (米国についてののみ): 猪口 明博 (INOKUCHI, Akihiro). 高林 健登 (TAKABAYASHI, Kiyoto). 鷺尾 隆 (WASHIO, Takashi).
- (74) 代理人: 新居 広守 (NII, Hiromori); 〒5320011 大阪府大阪市淀川区西中島5丁目3番10号タナカ・イトーピア新大阪ビル6階新居国際特許事務所内 Osaka (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, 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, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG,

[続葉有]

(54) Title: MULTI-DIMENSIONAL DATA ANALYZING METHOD, MULTI-DIMENSIONAL DATA ANALYZING DEVICE, AND PROGRAM

(54) 発明の名称: 多次元データ分析方法、多次元データ分析装置、及びプログラム

[図2]



200 MULTI-DIMENSIONAL DATA ANALYZING DEVICE
204 INPUT UNIT
203 DISPLAY UNIT
202 OPERATION UNIT
202a AGGREGATION OPERATION BLOCK
202b JOIN SELECTION OPERATION BLOCK
202c SECTION SELECTION OPERATION BLOCK
201 DATABASE

(57) Abstract: A multi-dimensional data analyzing method having high usability for interactively analyzing, e.g., medical data and management data stored in a hospital information system so as to assist discovery of knowledge about clinical decision. A multi-dimensional data analyzing device (200) comprises a database (201) holding separately a section table (I) showing sections and a hierarchy table (T) showing the hierarchical structure of the dimensions of multi-dimensional data, a section selection operation block (202c) for selecting a section (I') having a property (c) requested by a user from the section table (I) by using a section selection operation g, a join selection operation block (202b) for joining a set of sections with a join selection operation β by using a join selection operation β in the selected section (I'), and aggregation operation block (202a) for creating a multi-dimensional cube from the result obtained by the join selection operation block (202b) by using an aggregation operation α .

(57) 要約: 例えば病院情報システムに蓄積された診療データ、および管理データに対して対話的に分析を行うことで臨床判断に関する知識発見を支援する、ユーザビリティに富んだ多次元データ分析方法を提案する。多次元データ分析装置(200)は、区間を表す区間テーブルI、及び多次元データの次元が有する階層構造を表す階層テーブルTとに分離して保持するデ

[続葉有]



WO 2009/066442 A1



SK, SL, SM, ST, SV, SY, TJ, TM, TN, TR, TT, TZ, UA,
UG, US, UZ, VC, VN, ZA, ZM, ZW.

CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,
IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE,
SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,
GW, ML, MR, NE, SN, TD, TG).

(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, 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 (AT, BE, BG,

添付公開書類:
— 国際調査報告書

ータベース (201)、区間選択演算 g を用いて、区間テーブル I よりユーザの要求する性質 c を有する区間 I' を選択する区間選択演算部 (202c)、区間選択演算部 (202c) において選択された区間 I' において、結合選択演算 β を用いて、区間の集合を結合選択演算 β と結合する結合選択演算部 (202b)、結合選択演算部 (202b) における結果を用いて、集計演算 α を用いて多次元キューブを生成する集計演算部 (202a) 等を備える。

明 細 書

多次元データ分析方法、多次元データ分析装置、及びプログラム 技術分野

[0001] 本発明は、次元と事象が多対多の関係にある時系列データを多次元分析するための多次元データ分析方法、多次元データ分析装置、及びプログラムに関するものである。

背景技術

[0002] 近年、計算機環境やそれを取り巻くネットワーク技術の著しい向上とデータベースを代表とするミドルウェア等の基盤技術の発展によって、膨大な量の情報を蓄積・管理する技術が整備されてきている。それに加えて、厚生労働省によって「医療・健康・介護・福祉分野の情報化グランドデザイン」（非特許文献１５参照）が策定され、徐々に電子カルテシステムが導入されてきたことによって、診療・管理データを蓄積するシステムの普及が進み、診療業務の効率化が図られてきている。

[0003] 一方、日々蓄積される膨大な情報を活用し、知的生産性を高める情報マネジメント技術や新たな知識発見を可能にする解析技術に対する期待は高まりつつある。医療を取り巻く最近の情勢として、国民医療費の増加や少子・高齢化による健康保険制度の財政の逼迫と *e-japan* 構想に代表される公的サービスのＩＴ化によって、情報システムを利用した病院経営の変革が求められている（非特許文献９参照）。

[0004] 現在、医療分野の情報化グランドデザインに沿って徐々にではあるが医療情報システムの導入が進み、診療業務・病院運営管理業務の効率化の兆しが見られ、診療の透明性を高めたことによって患者に安心感を与えることに成功している。

[0005] しかしながら、せっかく蓄積された膨大な量の医療情報を経営の効率化や根拠に基づいた医療（ＥＢＭ）の確立に活用する技術には改善の余地が残されている。

- [0006] というのも、医療情報データは患者の診療や検査データ、投薬、手術など時系列データからなり、またそれぞれの項目が非常に複雑な階層構造を有し、マスターデータとして管理されている。各患者はそれぞれ異なる診療科の異なる診療、手術、投薬、検査を複数回受けている。これらのデータの分析は、診療プロセスのより詳細な解析やクリティカルパス（クリニカルパス）の評価等に貢献できるが（非特許文献 9 参照）、複雑で大規模なデータベース全体から問題発見のため、可能な仮説空間を全探索するようなデータマイニングの技術で分析を行うことは容易ではない。したがって、ユーザが興味のある項目を対話的に、あるいは試行錯誤しながら絞り込む分析を行うことが計算機の処理能力の点からも現実的である。
- [0007] また、対話的に分析を行うことは複雑な構造を有するデータからの問題発見プロセスとしても有効である。データベースの分野では時系列データを対話的に分析する技術として、多次元データベースがある（非特許文献 1, 2, 4, 6, 11 参照）。
- [0008] この多次元データベースは計数と次元値を持った事象の集合としてデータを扱う。例えば小売データでは、各購入履歴がファクトであり、量や価格が計数であり、商品種別や購入時間、購入場所等が次元値である。膨大な量の元データに対して検索・抽出・加工し、多次元データベースに格納し、結果を出力する処理をオンライン分析処理（OLAP：OnLine Analytical Processing）という。多次元データベースの次元は階層構造を持ち、処理要求に応じたデータの粒度でデータを選択・集計することが可能である。
- [0009] そして、例えば、従来の多次元DBでの分析での典型例として購買履歴の例がある。各店舗では、どの商品が、いつ、どこで、いくら売れたかをデータベースに蓄積し、図 15 に示すような 3 次元のデータベースで売り上げ合計などを集計する。
- [0010] 図 15 は、多次元キューブ 1500 の例であり、図 15 の例の購買位置の軸（次元）は、市レベルでの集計であるが、次元には階層があり、県レベル、地域（関東、関西など）レベルなど、ユーザの分析意図に応じた極度で対

話的に分析することができる。

非特許文献1 : S. Agarwal, R. Agrawal, P. Deshpande, A. Gupta, J. F. Naughton, R. Ramakrishnan, and S. Sarawagi. On the Computation of Multidimensional Aggregates. Proc. of International Conference on Very Large Data Bases, pp. 506–521, 1996.

非特許文献2 : P. Baumann, A. Dehmel, P. Furtado, R. Ritsch, and N. Widmann. Spatio-Temporal Retrieval with RasDaMan. Proc. of International Conference on Very Large Data Bases, pp. 746–749, 1999.

非特許文献3 : P. F. Dietz. Maintaining order in a linked list. Proc. of Annual ACM Symposium on Theory of Computing, pp. 122–127, 1982.

非特許文献4 : S. Goil and A. N. Choudhary. High Performance Multi-dimensional Analysis of Large Datasets. Proc. of International Workshop on Data Warehousing and OLAP, pp. 34–39, 1998.

非特許文献5 : H. Gupta, V. Harinarayan, A. Rajaraman, and J. D. Ullman. Index Selection for OLAP. Proc. of International Conference on Data Engineering, pp. 208–219, 1997.

非特許文献6 : M. Gyssens and L. Lakshmanan. A Foundation for Multi-dimensional Databases. Proc. Of International Conference on Very Large Data Bases, pp. 106–115, 1997.

非特許文献7 : A. Inokuchi, K. Takeda, N. Inaoka, and F. Wakao. MedTAKMI-CDI: Interactive knowledge discovery for clinical decision intelligence. IBM Systems Journal, Volume 46, Number 1, pp. 115–134, 2007.

非特許文献8 : A. Inokuchi and K. Takeda. A Method for Online Analytical Processing of Text Data. Proceedings of ACM Conference on Information and Knowledge Management, (CIKM 2007), 2007. (to appear)

非特許文献9 : 紀ノ定保臣, 梅本敬夫, 猪口明博, 武田浩一, 稲岡則子. マイニング技術を活用した診療プロセス分析への挑戦. 医療情報学, 第26巻, 第3号, pp. 191–199, 2006.

非特許文献10 : T. Pedersen and C. Jensen. Multidimensional Data Modeling for Complex Data. Proceedings of the 15th International Conference on Data Engineering, pp. 336-345, 1999.

非特許文献11 : T. B. Pedersen and C. S. Jensen. Multidimensional Database Technology. IEEE Computer, Vol. 34, No. 12: pp. 40-46, 2001.

非特許文献12 : 若尾文彦, 石川ベンジャミン光一, 稲岡則子, 猪口明博, 鈴木進. がん診療プロセス解析システムの検討. 第25回医療情報学連合大会, 2-F-6-6, 2005.

非特許文献13 : L. Wang, A. Zhang, and M. Ramanathan. BioStar Models of Clinical and Genomic Data for Biomedical Data Warehouse Design. International Journal of Bioinformatics Research and Applications, Vol. 1, No. 1, pp. 63-80, 2005.

非特許文献14 : 五十嵐健夫, 芦原貴司, 永田啓, 高田雅弘, 中沢一雄. 最新パソコンコンピューティング技術に基づく電子カルティンターフェース. 医療情報学, 第20巻, 第2号, pp. 482-483, 2000.

非特許文献15 : 厚生労働省. 医療・健康・介護・福祉分野の情報化グランドデザイン. <http://www.mhlw.go.jp/houdou/2007/03/h0327-3.html>.

非特許文献16 : 西堀真弘, 椎名晋一. 医療情報システムのユーザーインターフェース. 医療情報学, 第10巻, 第1号, pp. 3-14, 1990.

非特許文献17 : 山野辺裕二, 相澤志優, 本多正幸. 電子カルテシステムGUIの問題点. ITヘルスケア, Vol. 2, No. 1, pp. 28-31, 2007. 8

発明の開示

発明が解決しようとする課題

- [0011] しかしながら、前記のような既存多次元DBで例えば電子カルテ内の医療データを分析する場合、医療情報データの特性として、従来の多次元データベースで用いられているスキーマによってデータを格納することが困難である点や分析の際にデータの時間的順序を考慮しなければならない点などがあり、これまで多く研究されてきた購入履歴等のデータよりも複雑なデータの

モデル化および解析のための新たな手法が必要であるという課題がある。

- [0012] すなわち、医療情報データに対する従来のOLAPの課題として、医療情報データに対して従来のOLAPを用いて解析を行う場合、以下の4つの問題に直面する。
- [0013] 1つ目に、スタースキーマによる多次元データベースでは、ファクトと次元が1対nの関係であるのに対して、診療履歴は必ずしも1対nでなく、n対mの関係のことが多い。具体的には、小売データの分析ではファクトである1つの購買履歴に対して商品種別や購入時間、購入場所等の各次元値が1つだけ関連付けされている。これに対して、診療履歴において一患者の履歴をファクトとし、診療や手術、投薬、検査データを次元値とした場合、1ファクトに対して次元値が複数存在し、次元となりうる項目に複数のファクトが対応することになり、従来のスタースキーマでは対応できない。一方、一入院をファクト、“主”病名、“主要な”手術などを次元として扱えばスタースキーマにデータを格納できるが、外来と入院をまたがった分析、複数入院にまたがった分析を困難にする。
- [0014] 2つ目に、医療情報データでは事象の時間的順序が重要な意味を持ち、事象の前後関係を意識した分析の問い合わせをしなければならない。具体的には、喉頭がんが発症している患者に対して、外科手術を行う前に化学療法や放射線治療で腫瘍を小さくする場合と外科手術を行った後にがんの再発予防のために化学療法や放射線治療をした場合とでは異なる診療プロセスとして捉えることが要求される。
- [0015] 3つ目に、これまで記述してきた課題を考慮すると複雑な条件を組み合わせた問い合わせであるため、対話的に分析を行うための効率的な処理が必要である。ただし、次元となりうる項目の種類が多い診療データに対して、事前集計を必要とするMOLAPのような形態は困難である。
- [0016] 4つ目に、このような複雑な処理を実行するにはSQL等の問い合わせ言語を用いて複雑なクエリを与えなければならない。SQLに不慣れな医療従事者をユーザと想定した場合、対話的に分析を行うためには、直感的に操作

可能なユーザインターフェースを装備していなければならない。

[0017] このように、個々の購買は別々のレコードとして使われるが、電子カルテの各検査履歴、手術歴、入退院歴、病歴などは、同一の患者であれば一連のデータであるので、データの特性が異なり、十分な分析ができないという問題がある。購買履歴であれば、1つの購買レコードに対して、次元別である購買位置、購買時間、商品種が1つずつ関連付けられているが、医療データでは患者に対して、各項目に対して複数の検査履歴、手術歴、入退院歴、病歴は複数個関連付けられている。主病名、主要な手術、検査の有無などの主要な1つのデータを関連付けて商用システムで分析した実施例もあるが、十分な分析は不可能である。

[0018] また、後述するPedersenの手法では、診療プロセスの順番を考慮した分析が困難であり、Biostar（非特許文献7参照）と呼ばれる手法では、データの格納法が中心の提案で、ユーザが分析したい結果を得るための手順（演算）はユーザにゆだねられている。またさらに、MedTAKMI-CDI（非特許文献13参照）の手法では、イベント中心にデータを保持しているが、効率面で悪い。また、個々の機能が別々に実装されており拡張性・柔軟性がないという問題がある。

[0019] 本発明は、以上の課題に鑑みてなされたものであり、従来のOLAPでは柔軟な分析が容易ではない医療情報データ等に対して、データを事象の開始時間と終了時間の情報を持つ区間データとして取り扱うことで、時間的順序の扱いを容易にするデータモデルとテーブルスキーマを有する多次元データ分析方法を提供する。

[0020] また、ユーザの様々な問い合わせを統一的に扱うことができる多次元データ分析方法を提供する。

[0021] またさらに、ユーザの分析意図を直感的に表現して、分析を容易に実行することができるユーザインターフェースを有する多次元データ分析方法を提供することを目的とする。

課題を解決するための手段

[0022] 以上の課題を解決するために、本発明に係る多次元データ分析方法は、次元と事象が多対多の関係にある時系列データを多次元分析するための多次元データ分析方法であって、前記事象の開始時間と終了時間との情報を持つ区間を表す区間テーブル I 、及び多次元データの次元が有する階層構造を表す階層テーブル T とに分離してデータベースに保持する保持ステップと、区間をあらわすテーブルを返す演算である区間選択演算 g を用いて、前記区間テーブル I よりユーザの要求する性質 c を有する区間 I' を選択する区間選択演算ステップと、前記区間選択演算ステップにおいて選択された区間 I' において、前記区間 I' と所定の結合の条件とを結合する演算である結合選択演算 β を用いて、区間の集合を前記結合選択演算 β と結合する結合選択演算ステップと、前記結合選択演算ステップにおける結果を用いて、データテーブルから n 次元の多次元キューブを生成する演算である集計演算 α を用いて多次元キューブを生成する集計演算ステップとを含むことを特徴とする。

[0023] この構成により、前記区間テーブル I を用いてデータを事象の開始時間と終了時間の情報を持つ区間データとして取り扱うことで、時間的順序の扱いを容易にするデータモデルとテーブルスキーマを有し、前記区間選択演算 g 、前記結合選択演算 β 、前記集計演算 α を用いることによりユーザの様々な問い合わせを統一的に扱うことができる多次元データ分析方法とすることができる。

[0024] また、本発明に係る多次元データ分析方法は、さらに、ユーザからの入力指示を受け付ける入力ステップと、前記集計演算ステップにより生成された前記多次元キューブ、及び前記入力ステップにおけるユーザ操作時に使用されるユーザインターフェースを画面に表示する表示ステップとを含み、前記表示ステップにおいて表示されるユーザインターフェースは、長方形オブジェクトの左辺、右辺を区間の開始、終了時間とし、異なる長方形オブジェクトの2区間を線で結ぶことにより区間に時間的な前後関係を指定すると共に、前記長方形オブジェクトと集計の演算用長方形オブジェクトとに線を結ぶことで、集計演算の入力とされることを特徴とする。

[0025] この構成により、入力ステップにおいてユーザが前記ユーザインターフェースを用いて対話的に分析するので、操作演算子やプログラミングに不慣れた医療従事者等のユーザでも操作でき、ユーザの分析意図を直感的に表現して、分析を容易に実行することができる多次元データ分析方法とすることができる。

[0026] 尚、前記目的を達成するために、本発明は、多次元データ分析方法の特徴的なステップを手段とする多次元データ分析装置として実現したり、コンピュータに各ステップを実行させるためのプログラムとして実現することもできる。そのようなプログラムは、CD-ROM等の記録媒体やインターネット等の伝送媒体を通じて配信することができるのは言うまでもない。

発明の効果

[0027] 本発明に係る多次元データ分析方法では、データを事象の開始時間と終了時間の情報を持つ区間データとして取り扱うことで、時間的順序の扱いを容易にするデータモデルとテーブルスキーマを実現できる。また、ユーザの様々な問い合わせを統一的に扱うことでできるデータ操作演算を提供できる。またさらに、ユーザの分析意図を直感的に表現でき、分析を容易に実行することを可能にするユーザインターフェースをも提供できる。

図面の簡単な説明

[0028] [図1] 図1は、本発明に係る多次元データ分析装置のデータ分析の説明図である。

[図2] 図2は、本発明の多次元データ分析装置の機能ブロックの一例を示す図である。

[図3] 図3は、本発明の多次元データ分析装置の演算部の動作手順を示すフローチャートである。

[図4] 図4は、国際疾病分類の一部を示す参考図である。

[図5] 図5は、関数 g を示す参考図である。

[図6] 図6は、クエリ例1の出力イメージを示す参考図である。

[図7] 図7は、クエリ例2の出力イメージを示す参考図である。

[図8] 図8は、クエリ例3の出力イメージを示す参考図である。

[図9] 図9は、クエリ例4の出力イメージを示す参考図である。

[図10] 図10は、クエリ例5の出力イメージを示す参考図である。

[図11] 図11(a)は、GUI上で区間を表すオブジェクトを示す参考図、
図11(b)は2区間の関係を表している参考図である。

[図12] 図12(a)(b)は、クエリ例1, 2のクエリ描写を示す参考図である。

[図13] 図13は、擬似データを用いて本発明に適用した例を示す参考図である。

[図14] 図14は、BioStarのテーブルスキーマを表している参考図である。

[図15] 図15は、従来のOLAPの説明図である。

符号の説明

- [0029] 200 多次元データ分析装置
 - 201 データベース
 - 202 演算部
 - 202a 集計演算部
 - 202b 結合選択演算部
 - 202c 区間選択演算部
 - 203 表示部
 - 204 入力部
- 400 国際疾病分類
- 500 ユーザ定義関数

発明を実施するための最良の形態

- [0030] 以下、本発明に係る多次元データ分析方法の実施の形態を図面を参照しながら説明する。

- [0031] (実施の形態)

図1は、本発明に係る多次元データ分析装置のデータ分析の説明図である。

。

[0032] 本発明の多次元データ分析方法においては、例えば電子カルテ等の医療データを、事象の開始時間と終了時間との情報を持つ区間を表すテーブルであるテーブル I と多次元データの次元が有する階層構造を表す階層テーブル T とに分離してデータベースに格納している。ここで、I は例えば患者の入退院期間、病気の期間、手術時間帯などを保持し、T は例えば術式の階層、病気の階層（ICD）などを保持している。そして、後述する各演算の区間選択演算 g 、結合選択演算 β 及び集計演算 α によりユーザの要求する検索結果を多次元キューブとして表示できる。

[0033] また、後述の図 11 から図 13 に示すように、ユーザは区間を表示する長方形オブジェクトで所望の検索条件でデータ検索できる。本発明のユーザインターフェースでは、長方形オブジェクトの左辺、右辺を区間の開始、終了時間とし、2 区間を線で結ぶことにより区間に時間的な前後関係を指定する。また集計の演算用長方形オブジェクトに線を結ぶことで、集計演算の入力ができる。

[0034] 図 2 は、本発明の多次元データ分析装置の機能ブロックの一例を示す図である。

[0035] 多次元データ分析装置 200 は、電子カルテの情報を用いて分離された区間テーブル I 及び階層テーブル T が保持されるデータベース 201 と、集計演算部 202 a、結合選択演算部 202 b、及び区間選択演算部 202 c を含む演算部 202 と、演算部 202 の演算結果である多次元キューブや入力部 204 を介して操作されるユーザインターフェースを表示する表示部 203 と、操作入力部であるキーボード等の入力部 204 とを備える。

[0036] 図 3 は、本発明の多次元データ分析装置の演算部の動作手順を示すフローチャートである。

[0037] 最初に、区間選択演算部 202 c は、区間選択演算 g により、I よりユーザの要求する性質 c を有する区間 $l' = g(l, T, c)$ を選択する（S301）。続いて、結合選択演算部 202 b は、結合選択演算 β により、区間の集合を $\beta(\{l' 1, \dots, l' n\}, 0, W) = \pi o(\sigma p(l' 1 \times \dots \times l' n))$ と結合する（S302）。ここ

で、WとOは選択条件と出力のカラムである。最後に、集計演算部202aは集計演算 α により多次元キューブを生成し(S303)、表示部203に表示する。

[0038] 以下、本発明の多次元データ分析方法の詳細について説明する。

[0039] 最初に、本発明で提案する手法を文献（非特許文献8及び10参照）に従い定義すると、分析の対象データDを $D = \{(f_i, \{p_{i1}; p_{i2}, \dots, p_{im}\})\}$ ($i=1; 2, \dots, n$)と定義する。

[0040] ここで、 $\{f_i \mid i=1; 2, \dots, n\}$ は患者IDの集合であり、 p_{ij} は区間情報であるとする。また、 $(f_i; \{p_{i1}; p_{i2}, \dots, p_{im}\})$ は各患者 f_i が区間に関する情報 p_{ij} の集合をもつことを意味する。

[0041] 区間を $p_{ij} = (t_s, t_e, \{c:v\})$ と定義する。ここで、 t_s 、及び t_e は区間の開始時刻、及び終了時刻とする。特に、 $t_s = t_e$ であるとき、区間 p_{ij} をイベントと呼ぶ。vは区間を説明する値であり、cは値vの属するカテゴリとする。またcは階層を持つデータのある節点である。

[0042] 具体的には、 $\{p_{ij}\}$ が入退院に関する区間(期間)であれば、 $c:v$ は、入院中病名、主治医などからなる。病名のカテゴリは図4に示されるような階層構造をもつ国際疾病分類(ICD: International Classification of Diseases)400を用いる。また入院期間中の病名は1つとは限らないので、同一cであるが異なるvをもつ $c:v$ が存在する可能性もある。また p_{ij} が手術であれば、 $c:v$ は術式、手術部位、執刀医などであり、執刀医のカテゴリは所属する診療科などで階層化されている。また従来のOLAPシステムであれば、cとvを区別することなく用いるが、本発明ではcを検体検査における白血球数の検査項目、vをその検査値と扱うので、区別して扱う。またcはカテゴリ階層の最下位の節点である必要はなく、内部節点でもよいものとする。

[0043] 階層の集合 $D = \{T_k\}$ が与えられたとき、スキーマは $S = (F; D)$ と定義される。ここで、Fはファクトタイプ、 T_k は階層タイプ $T_k = (C_l; <_{T_k})$ である。タイプ T_k の階層インスタンス T_k は、 $T_k = (C_k; <_{T_k})$ である。ここで、 C_k はカテゴリ c_j の集合、 $<_{T_k}$ は C_k 間の半順序関係を示している。

- [0044] 本発明で用いる階層は従来のOLAPのシステムの多くが採用するバランス木である必要はなく、有向非循環グラフ(DAG: Directed Acyclic Graph)を想定している(非特許文献8参照)。各カテゴリ $c \in C$ は定義域 $\text{dom}(c)$ を持ち、前述のように $\text{dom}(c)$ の各要素は、 $\{c:v\}$ と表される。
- [0045] 集計演算の計算速度をあげるために、階層を以下のように索引付けする。人工のルートノード c_{root} を考え、 C のうち上位の概念を持たない c_j の親節点とする。 c_{root} からはじめて、各節点に前順、後順、深さを割り当てながら深さ優先に探索していく。ただし、内部節点ではバックトラックせずに、葉節点でのみバックトラックする。入力であるカテゴリ c とデータのカテゴリが子孫関係にあるかを判定するために、以下の条件で容易に判定することが可能である。もし節点Aが節点Bの祖先であるなら、以下の数(1)が成り立つ(非特許文献3参照)。

[0046] [数1]

$$A's \text{ preorder1} (=A \text{ の前順}) < B \text{ の前順} \\ \leq A's \text{ preorder2} = A \text{ の後順} + A \text{ の深さ}$$

- [0047] 階層関係と区間情報をストアするために、テーブルCATEGORY TとINTERVAL Iを以下のように定義する。

CATEGORY	(CATENAME	CHARACTER,
	PATH	CHARACTER,
	PREORDER1	INTEGER,
	PREORDER2	INTEGER,
	PARENT	INTEGER)
INTERVAL	(ID	INTEGER,
	START	TIMESTAMP,
	END	TIMESTAMP,
	PREORDER	INTEGER,
	VALUE	CHARACTER,
	INTERVALID	INTEGER)

- [0048] Tの各レコードは階層の各節点に相当し、CATENAME、PATH、PREORDER1、PREORDER2、PARENTは、それぞれ節点のカテゴリ名、ルート節点からのパス、前順、後順と深さの和、親節点の前順である。
- [0049] Iの各レコードは、 $(ts; te; \{c:v\})$ を $|\{c:v\}|$ 個に分割した情報に相当し、ID、START、END、PREORDER、VALUE、INTERVALIDは、それぞれ患者ID、区間の開始時間、区間の終了時間、カテゴリcの前順、 $\text{dom}(c)$ 内の値v、区間の識別子である。区間識別子INTERVALIDを用いる理由は、 $(ts; te; \{c:v\})$ を $|\{c:v\}|$ 個に分割したためである。
- [0050] 以上の2つのテーブルを用いて、集計演算を以下のように定義する。以下の定義において、Tcは入力カテゴリに対して、テーブルTの1タプルを返すSQL文である“ $\sigma_p(T)$ FETCH FIRST 1 ROWS ONLY”を意味する。
- [0051] (1) 集計演算 α : テーブルA(v1, v2, vn, id) に対して、 $\alpha(A) =_{v1:v2, \dots, vn} X_{v1:v2, \dots, vn; \text{count}(\text{distinct id})}$ を返す集計演算を $\alpha(A)$ と定義する。演算 α はテーブルAから、n次元の多次元キューブを生成する関数であると理解できる。
- [0052] (2) 結合演算 β : 結合演算 β を $\beta(\{I'_1, I'_2; \dots; I'_n\}; 0; W) = \pi_0(I'_1 \times I'_2 \times \dots \times I'_n)$ と定義する。ここで、各テーブル I'_i は区間 $I'(id; \text{start}; \text{end}; \text{value}; \text{interval_id})$ とする。またWは結合の条件式の集合であり、 $I'_i \times \dots \times I'_j$ は、Wの条件式、及び $I'_i.\text{id} = I'_j.\text{id}$ に従って結合される。0は出力されるカラムの集合であるとする。
- [0053] (3) 区間選択演算 g : 区間選択演算 $g(T; I; c)$ を区間を表すテーブル $I'(id; \text{start}; \text{end}; \text{value}; \text{interval_id})$ を返す演算と定義する。関数 g は分析の意図に応じて定義されるユーザ定義関数（非特許文献8参照）500であり、例えば図5に示される。 $g^{(1)}$ は指定したカテゴリcとその子孫カテゴリに属するvを有する区間を選択する演算である。 $g^{(2)}$ は指定したカテゴリcに属するvを有する区間を選択する演算である。
- [0054] $g^{(3)}$ は $g^{(1)}$ と同様の区間を選択する演算であるが、vをテーブルTのCATEGORYNAMEに置き換えて区間を選択する演算である。 $g^{(4)}$ もまた $g^{(1)}$ と同様の区間を選択する演算であるが、指定されたカテゴリcの子カテゴリのCATEGORYNAMEに置き

換えて区間を選択する演算である。 $g^{(5)}$ は $g^{(1)}$ と同様の区間を選択する演算であるが、 v を区間の開始時間に置き換えて区間を選択する演算である。

[0055] 上記で定義した演算によってどのような集計が可能かを具体例を用いて示す。

[0056] (1) クエリ例1を数(2)を用いて表す。

[0057] [数2]

$$\alpha(\beta(\{g^{(1)}(T, I, c_1), g^{(1)}(T, I, c_2)\}, O_1, W_1))$$

[0058] c_1 と c_2 をそれぞれ、手術と入退院のカテゴリとし、数(3)とすると、

[0059] [数3]

$$\begin{aligned} O_1 &= \{I'_1.value, id\}, \\ W_1 &= \{I'_2.start \leq I'_1.start, I'_1.end \leq I'_2.end\}, \end{aligned}$$

[0060] 上記のクエリは、入院期間中に手術を施した患者数を術式ごとに集計した結果を返す。出力のイメージは図6のとおりである。この図6は、クエリ例1の出力イメージ600を示す参考図である。

[0061] (2) クエリ例2を数(4)を用いて表す。

[0062] [数4]

$$\alpha(\beta(\{g^{(4)}(T, I, c_1), g^{(1)}(T, I, c_2), g^{(1)}(T, I, c_3)\}, O_2, P_2))$$

[0063] c_1, c_2, c_3 をそれぞれ、手術、入退院、放射線検査(X線, CT, MRI)のカテゴリとし、数(5)とすると、

[0064] [数5]

$$\begin{aligned} O_2 &= \{date(I'_1.start), I'_1.value, id\}, \\ W_2 &= \{I'_2.start \leq I'_3.start, I'_3.end \leq I'_1.start, I'_1.end \leq I'_2.end\} \end{aligned}$$

[0065] 上記のクエリは、入院期間中に放射線検査、手術を“順に”実施した患者について、手術の診療科、及び手術日毎に患者数を集計した結果を返す。出力のイメージ700は図7のとおりである。手術に関するデータはテーブルIでは術式で保持されていて、術式の上位の階層として術式に適した診療科があるものとする。図7は術式毎の集計から診療科毎の集計にロールアップしたことに相当する。

[0066] (3) クエリ例3を数(6)を用いて表す。

[0067] [数6]

$$\alpha(\beta(\{g^{(4)}(T, I, c_1), g^{(1)}(T, I, c_2), g^{(1)}(T, I, c_4)\}, O_3, W_3))$$

[0068] c_1, c_2, c_4 をそれぞれ、手術、入退院、性別のカテゴリとし、数(7)とすると、

[0069] [数7]

$$\begin{aligned} O_3 &= \{I'_1.value, date(I'_1.start) - date(I'_2.start), \\ &I'_3.value, interval_id\}, \\ W_3 &= \{year(I'_2.start) = 2007\} \end{aligned}$$

[0070] 上記のクエリは、入院が2007年である患者について、入院日から手術日までの経過日数を性別毎、診療科毎に手術件数を集計した結果を返す。出力のイメージ800は図8のとおりである。

[0071] 図8に示す、縦線は入院日を示し、横軸は入院日を基準とした左から右へ流れる時間であり、縦軸は各診療科において横軸で示される時間に手術を受けた男性患者（実線）および女性患者（点線）の人数である。条件式である $year(I'_2.start)=2007$ は、入院日が2007年である入退院期間に限定する操作であり、従来のOLAPのスライスに相当する操作である。また前述の2つのクエリが患者数を集計していたのに対し、クエリ例3は手術件数を集計

しており、本発明のテーブルスキーマでは計数の属性を別に扱わない。

[0072] (4) クエリ例 4 を数 (8) を用いて表す。

[0073] [数8]

$$\alpha(\beta(\{g^{(1)}(T, I, c_3), g^{(1)}(T, I, c_3), g^{(1)}(T, I, c_3), \\ g^{(1)}(T, I, c_2)\}, O_4, W_4))$$

[0074] c_3 を放射線検査のカテゴリとし、数 (9) とすると、

[0075] [数9]

$$O_3 = \{ I'_1.value, I'_2.value, I'_3.value, id \}, \\ W_4 = \{ I'_4.start \leq I'_1.start < I'_2.start < I'_3.start \leq \\ I'_4.end \}$$

[0076] 上記のクエリは、入院期間中に放射線検査を 3 回以上受けた患者で、その種類の順番を種類毎に集計したクエリである。出力のイメージ 900 は図 9 に示す。

[0077] 図 9 が示すように、生成されるキューブの各次元は同じ放射線検査の種類となっている。スタースキーマによって実装される従来の OLAP がテーブルを定義する際にキューブの各次元が定義されるのに対し、本発明に係る手法はクエリが生成される際に、キューブの次元が定義される。図 9 は、クエリ例 4 の出力イメージ 900 を示す参考図である。

[0078] (5) クエリ例 5 を数 (10) を用いて表す。

[0079] [数10]

$$\alpha(\beta(\{g^{(7)}(T, I, c_5)\}, O_5, \emptyset))$$

[0080] c_5 を白血球数のカテゴリとし、 $O_5 = \{ I'_1.value, id \}$, $g^{(7)}$ を白血球数を離散化する関数とすると、上記のクエリは、図 10 のような結果を返すクエリとなる。なお、図 10 はクエリ例 5 の出力イメージ 1000 を示す参考図であ

る。

[0081] 次に、本発明に係る多次元データ分析装置において用いるユーザインターフェースの説明を行う。

[0082] 電子カルテの情報が関係データベースに蓄積されている環境で、SQLを用いたことがある人であれば、上述したようなテーブルを用いなくても、業務システム（あるいは、そのレプリカ）に直接問い合わせれば所望の分析結果を得ることができる。

[0083] しかしながら、本発明を利用するユーザはSQLを用いたことのない医療従事者を対象としている。例としてG大学病院に導入されている電子カルテシステムは、100を超えるマスタ情報と数十の実施テーブルからなるため、SQLに不慣れなユーザが所望な分析結果を得るための問い合わせを表現することは容易ではない。

[0084] また上記で示したような関数 α 、 β 、 g の組み合わせを表現することも困難である。そこで本発明では、ユーザの分析意図であるクエリを容易に表現できるユーザインターフェースを提案する。

[0085] 図11(a)は、GUI上で区間を表すオブジェクトである。長方形の左辺は区間の開始時刻を、右辺は区間の終了時刻に対応するものとする。図11(b)は2区間の関係を表しており、入退院区間の開始以降に手術区間の開始があり、手術区間の終了以降に入退院区間の終了点があるので、手術が入院期間中に行われたことを表している。

[0086] このようなユーザインターフェースを用いることで、上述したクエリ例1、2は図12(a)、及び図12(b)によって表される。図12に示されるように、網がけの長方形が演算 g 、及びその入力を表している。網がけの長方形間の辺によって区間の間の関係 W が指定される。演算を記された長方形に繋がる辺が0であり、演算 β の出力、演算 α の入力となる。

[0087] 以上で述べた本発明をJava（登録商標）で実装し、JDBC(Java Database Connectivity)経由で関係データベースにあるデータを集計するシステムHealth Cubeを実装した。

- [0088] またG大学病院のマスタ情報を用いて、擬似的に患者の診療履歴情報1300を作成した。擬似データを用いて本発明手法に適用した例を図13に示す。左フレームがカテゴリの階層である。右上のフレームが問い合わせクエリを作成するインターフェースであり、右下のフレームに集計結果を表示する。図13では入院後に検体検査を行った後に呼吸器外科の手術を受けた患者の人数を表している。
- [0089] 具体的には、表中の数字は縦軸の検査を受けた後に横軸の手術を受けた人数である。擬似データの患者数は50,400人、総区間数は4,187,845であり、クエリに含まれる条件となる区間数や集計結果の次元数にもよるが、ほとんどのクエリを数秒で返すことができる。
- [0090] 次に、考察および関連研究を述べる。
- [0091] 1995年に厚生省の電子カルテ開発プロジェクトが開始される以前から継続して医療情報システムについて議論されてきているが、今もなおその操作性やインターフェースに関する議論は絶えない（非特許文献14, 16, 及び17参照）。
- [0092] 問題点としては利用環境の把握が不十分な点やユーザのシステムの利用に割くことのできる時間が少ない点、操作手順が複雑な点、柔軟な思考を反映できない点など多く挙げられる。医療情報分析ツールに関しても同様の声が聞かれる。OLAPのような対話的な分析手法において、利便性と効率を高めるにはツールの操作性はもちろんのこと、ユーザが分析したい事柄を直感的に表現でき、これを出力結果に反映させることが重要である。以上のことを加味して本発明と関連する研究について考察する。
- [0093] 上述したように、本発明は演算機能 α , β , g とテーブル T, I を組み合わせることで様々な種類の問い合わせ文を同じ形式で作成することが可能である。スペースの都合上、上述した例は比較的単純な例であるが、より複雑な問い合わせを生成することが可能である。問い合わせで生成する区間やイベントの順序関係は全順序である必要はなく、半順序でよい。例えば、区間AとBはC以降であるが、AとBの区間の順序を指定しない問い合わせを生成すること

が可能である。

[0094] 本発明に関連する研究として、非特許文献 10 がある。非特許文献 10 は、医療データを O L A P で分析する上で、必要とされる 9 つ課題を挙げ、それを解決するデータモデルとそれに対する演算を提案している。しかし、多次元キューブを生成する定義されている演算では、同じ次元を選択することはできず、図 9 のような結果を得ることはできない。

[0095] 図 14 は BioStar（非特許文献 13）のテーブルスキーマを表している参考図である。患者と投薬、患者と手術間の n 対 m の関係を表すためにファクト表と次元表の間に M-テーブルを配置し、n 対 m の関係を保持することを可能にしている。ただし、手術の場合、執刀医が複数人いる可能性があり、かつその人数が患者によって異なる場合の情報を保持するには適切とはいえない。また、本実施の形態で示したように、診療履歴の分析の場合、区間やイベントの時間的前後関係を考慮した分析が重要であるが、非特許文献 13 で述べられているのは、データの格納方法が中心であり、それに対する処理についてはあまり述べられていない。

[0096] 上述したように、入退院期間中に施された術式を得る演算は、数（11）で表される。

[0097] [数11]

$$\begin{aligned} & \beta(\{g^{(1)}(T, I, c_1), g^{(1)}(T, I, c_2)\}, O_1 W_1) \\ & = \pi o(g^{(1)}(c_1) \bowtie w g^{(1)}(c_2)) \end{aligned}$$

[0098] ここで、 c_1 と c_2 は手術と入退院のカテゴリであり、数（12）である。

[0099] [数12]

$$\begin{aligned} O_1 &= \{I'_1.value, id\}, \\ W_1 &= \{I'_2.start \leq I'_1.start, I'_1.end \leq I'_2.end\} \end{aligned}$$

[0100] また、便宜上一部の T と I を省略している。一方、MedTAKMI-CDI（非特許文

献 7 参照) もまた、上記で挙げた問題の一部を解決するために提案された手法である。MedTAKMI-CDI は、区間単位ではなくイベント単位でデータを保持する。従って、入退院区間はイベント時間を持った入院イベントと退院イベントでデータを保持する。MedTAKMI-CDI で、入退院期間中に施された術式を得る演算を行うと数 (13) となる。

[0101] [数13]

$$\pi_{O_3}(\pi_{O_2}(GX_{O_1}(g^{(1)}(c_2) \bowtie_{P_1} g^{(1)}(c_3))) \bowtie_{P_2} g^{(1)}(c_1))$$

[0102] ここで、 c_1 , c_2 , c_3 は手術イベント、入院イベント、退院イベントのカテゴリであり、下記の数 (14) である。

[0103] [数14]

$$\begin{aligned} P_1 &= \{2.id = 3.id \text{ and } 2.start < 3.start\}, \\ P_2 &= \{2.id = 1.id \text{ and } 2.start \leq 1.start \leq end\}, \\ O_1 &= \{2.id, 2.start, \min(3.start - 2.start) \text{ as } min\}, \\ O_2 &= \{2.id, 2.start, 2.start + min \text{ as } end\}, \\ O_3 &= \{1.value, 3.id\}, G = \{2.id, 2.start\} \end{aligned}$$

[0104] $i.start$ は $g^{(1)}(c_i)$ から返されるカラム名とする。クエリ (2) と (3) を比較すると、クエリ (2) が結合が 1 回であるのに対し、クエリ (3) は結合が 2 回、集計が 1 回必要となる。 $g(c_i) \times \cdots \times p g(c_j)$ は、スタースキーマのファクト表が保持するタプル数と同じくらいのタプル数をもつテーブルの結合であるので、後者の方が計算時間を要するのは明らかである。

[0105] 更に、本発明が様々な分析要求を演算 α , β , g によって生成し、同じ形式の問い合わせで図 6、図 7、図 8 などを集計できるのに対し、MedTAKMI-CDI はそれぞれの機能毎に実装されており、同じ形式でのクエリを生成できない。

[0106] 以上の説明のように、本発明に係る多次元データ分析方法においては、従来では十分に分析できなかった区間やイベントの前後関係を考慮した多次元

キューブを生成することができ、また、演算 α 、 β 、 g の組み合わせにより、様々なクエリを表現することができる。また、対話的分析を支援するクエリを生成できる直感的なインターフェースを提供することができる。

[0107] 従って、例えば病院情報システムに蓄積された診療データ、および管理データに対して、区間データの概念を取り入れた演算機能とテーブルを組み合わせることで様々な種類の問い合わせ文を作成し、対話的に柔軟な分析を行うことが可能となる。

[0108] また、本発明に係る多次元データ分析方法を用いて過去の診療履歴のデータを分析することで、診療の質の向上、評価が可能となる。また、診療報酬の改定などで病院の経営の見直しが必要となるが、診療科同士の比較や在院日数の長期化の原因を調べることで、経営改善の効果を期待することができる。

[0109] なお、本発明は医療情報データに特化して記述しているが、本発明は汎用であり、異なるデータにも適用可能である。

産業上の利用可能性

[0110] 本発明に係る多次元データ分析方法は、電子カルテの医療データに適用することで、診療プロセス分析やクリニカルパスの定量的評価に用いることができる。ただし、本発明に係る多次元データ分析方法は、非常に汎用であり、医療データに限らず、例えば品質管理やマーケット分析にも利用可能である。

請求の範囲

- [1] 次元と事象が多対多の関係にある時系列データを多次元分析するための多次元データ分析方法であって、
- 前記事象の開始時間と終了時間との情報を持つ区間を表す区間テーブル I、及び多次元データの次元が有する階層構造を表す階層テーブル T とに分離してデータベースに保持する保持ステップと、
- 区間をあらわすテーブルを返す演算である区間選択演算 g を用いて、前記区間テーブル I よりユーザの要求する性質 c を有する区間 l' を選択する区間選択演算ステップと、
- 前記区間選択演算ステップにおいて選択された区間 l' において、前記区間 l' と所定の結合の条件とを結合する演算である結合選択演算 β を用いて、区間の集合を前記結合選択演算 β と結合する結合選択演算ステップと、
- 前記結合選択演算ステップにおける結果を用いて、データテーブルから n 次元の多次元キューブを生成する演算である集計演算 α を用いて多次元キューブを生成する集計演算ステップとを含む
- ことを特徴とする多次元データ分析方法。
- [2] 前記集計演算ステップにおける前記集合演算 α は、テーブル $A (v_1, v_2, \dots, v_n, id)$ に対して、 $\alpha(A) = v_1; v_2, \dots, v_n \wedge v_1; v_2, \dots, v_n; count(distinct id)$ を返す演算と定義される
- ことを特徴とする請求項 1 記載の多次元データ分析方法。
- [3] 前記結合演算ステップにおける前記結合演算 β は、 $\beta(\{l'_1, \dots, l'_n\}, 0, W) = \pi_O(\sigma_P(l'_1 \times \dots \times l'_n))$ と定義され、各テーブル l'_i は区間 $l' (id; start; end; value; interval_id)$ とされ、前記 W は結合の条件式の集合であり、 $(l'_i \times \dots \times l'_j)$ は、前記 W の条件式及び $l'_i.id = l'_j.id$ に従って結合され、前記 O は出力されるカラムの集合である
- ことを特徴とする請求項 1 記載の多次元データ分析方法。
- [4] 前記区間選択ステップにおける前記区間選択演算 g は、分析の意図に応じて定義されるユーザ定義関数であり、前記区間テーブル I よりユーザの要求する性質 c を有する区間を表すテーブル $I' (id; start; end; value; interv$

al_id)を返す演算と定義される

ことを特徴とする請求項1記載の多次元データ分析方法。

[5] 前記多次元データ分析方法は、さらに、

ユーザからの入力指示を受け付ける入力ステップと、

前記集計演算ステップにより生成された前記多次元キューブ、及び前記入力ステップにおけるユーザ操作時に使用されるユーザインターフェースを画面に表示する表示ステップとを含み、

前記表示ステップにおいて表示されるユーザインターフェースは、長方形オブジェクトの左辺、右辺を区間の開始、終了時間とし、異なる長方形オブジェクトの2区間を線で結ぶことにより区間に時間的な前後関係を指定すると共に、前記長方形オブジェクトと集計の演算用長方形オブジェクトとに線を結ぶことで、集計演算の入力とされる

ことを特徴とする請求項1から4のいずれか1項に記載の多次元データ分析方法。

[6] 前記表示ステップにおいて表示されるユーザインターフェースは、前記長方形オブジェクトが区間を、前記長方形オブジェクト間の線が前記条件Wを、前記演算用長方形オブジェクトへの線が前記Oとなり集計問い合わせクエリを生成する

ことを特徴とする請求項5記載の多次元データ分析方法。

[7] 次元と事象が多対多の関係にある時系列データを多次元分析するための多次元データ分析装置であって、

前記事象の開始時間と終了時間との情報を持つ区間を表す区間テーブルI、及び多次元データの次元が有する階層構造を表す階層テーブルTとに分離して保持するデータベースと、

区間をあらわすテーブルを返す演算である区間選択演算gを用いて、前記区間テーブルIよりユーザの要求する性質cを有する区間I'を選択する区間選択演算手段と、

前記区間選択演算手段において選択された区間I'において、前記区間I'と

所定の結合の条件とを結合する演算である結合選択演算 β を用いて、区間の集合を前記結合選択演算 β と結合する結合選択演算手段と、

前記結合選択演算手段における結果を用いて、データテーブルから n 次元の多次元キューブを生成する演算である集計演算 α を用いて多次元キューブを生成する集計演算手段とを備える

ことを特徴とする多次元データ分析装置。

[8] 前記多次元データ分析装置は、さらに、

ユーザからの入力指示を受け付ける入力手段と、

前記集計演算手段において生成された前記多次元キューブ、及び前記入力ステップにおけるユーザ操作時に使用されるユーザインターフェースを画面に表示する表示手段とを備え、

前記表示手段において表示されるユーザインターフェースは、長方形オブジェクトの左辺、右辺を区間の開始、終了時間とし、異なる長方形オブジェクトの2区間を線で結ぶことにより区間に時間的な前後関係を指定すると共に、前記長方形オブジェクトと集計の演算用長方形オブジェクトとに線を結ぶことで、集計演算の入力とされる

ことを特徴とする請求項7記載の多次元データ分析装置。

[9] 次元と事象が多対多の関係にある時系列データを多次元分析するための多次元データ分析装置に用いられるプログラムであって、

前記事象の開始時間と終了時間との情報を持つ区間を表す区間テーブル I 、及び多次元データの次元が有する階層構造を表す階層テーブル T とに分離してデータベースに保持する保持ステップと、

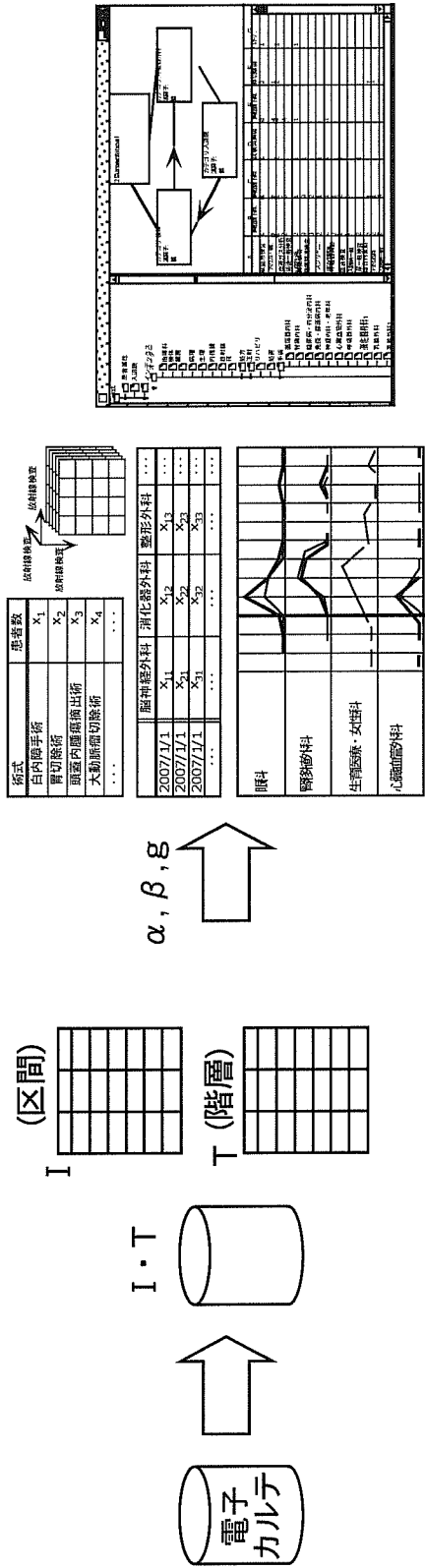
区間をあらわすテーブルを返す演算である区間選択演算 g を用いて、前記区間テーブル I よりユーザの要求する性質 c を有する区間 I' を選択する区間選択演算ステップと、

前記区間選択演算ステップにおいて選択された区間 I' において、前記区間 I' と所定の結合の条件とを結合する演算である結合選択演算 β を用いて、区間の集合を前記結合選択演算 β と結合する結合選択演算ステップと、

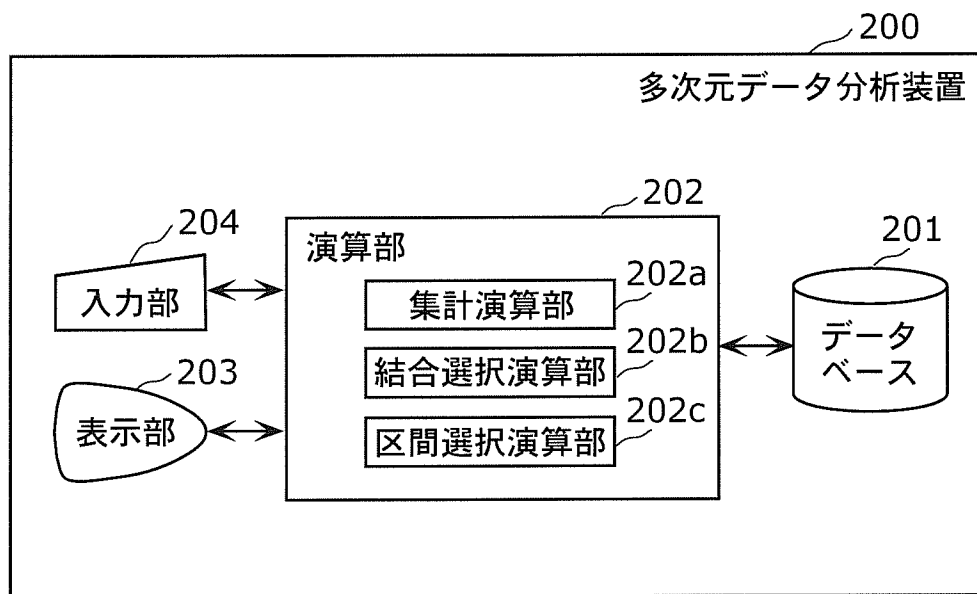
前記結合選択演算ステップにおける結果を用いて、データテーブルから n 次元の多次元キューブを生成する演算である集計演算 α を用いて多次元キューブを生成する集計演算ステップと

をコンピュータに実行させることを特徴とするプログラム。

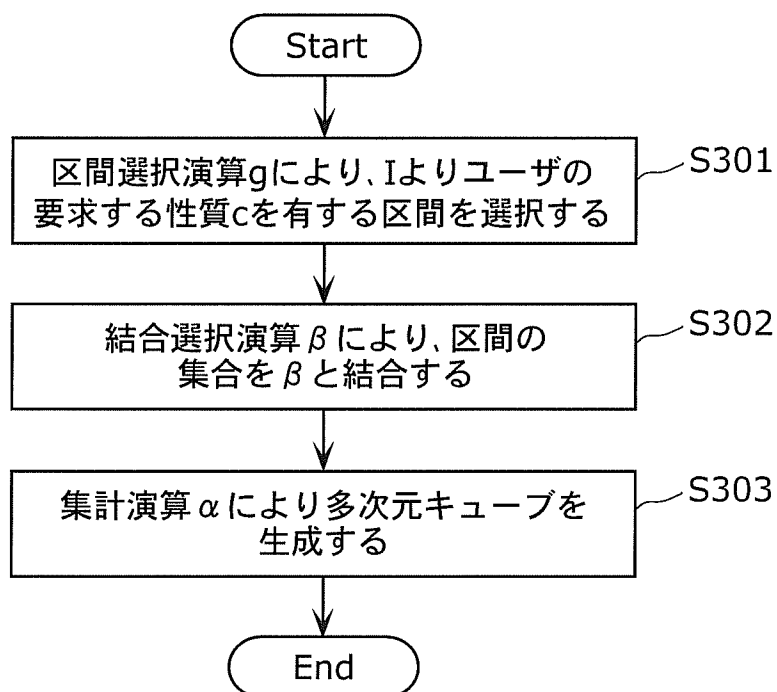
[図1]



[図2]



[図3]



[図4]

国際疾病分類400

- ・A00 - B99:感染症・寄生虫症
 - ・...
- ・C00 - D48:新生物
 - ・C00 口唇の悪性新生物
 - ・C00.0 口唇の悪性新生物, 外側上唇
 - ・C00.1 口唇の悪性新生物, 外側下唇
 - ・...
 - ・C01 舌根<基底>部の悪性新生物
 - ・C02 舌のその他及び部位不明の悪性新生物
 - ・...
- ・D50 - D89:血液・造血器疾患および免疫機能障害
 - ・...
- ・E00 - E90:内分泌・栄養・代謝疾患
 - ・...

[図5]

500

$g^{(1)}(T, I, c) = \mathcal{T}_{id, start, end, value, interval_id}(I \bowtie preorder1 \leq preorder \leq preorder2 \ T_c)$
 $g^{(2)}(T, I, c) = \mathcal{T}_{id, start, end, value, interval_id}(I \bowtie preorder1 = preorder \ T_c)$
 $g^{(3)}(T, I, c) = \mathcal{T}_{id, start, end, T.catename \text{ as value}, interval_id}(I \bowtie preorder1 \leq preorder \leq preorder2 \ T_c)$
 $g^{(4)}(T, I, c) = \mathcal{T}_{id, start, end, T.catename \text{ as value}, interval_id}((T_c \bowtie T.parent = T_c.preorder1 \ T) \bowtie T.preorder1 \leq I.preorder \leq T.preorder2 \ I)$
 $g^{(5)}(T, I, c) = \mathcal{T}_{id, start, end, start \text{ as value}, interval_id}(I \bowtie preorder1 \leq preorder \leq preorder2 \ T_c)$
 $g^{(6)}(T, I, c) = \mathcal{T}_{id, start, end, end \text{ as value}, interval_id}(I \bowtie preorder1 \leq preorder \leq preorder2 \ T_c)$
 ...

[図6]

600

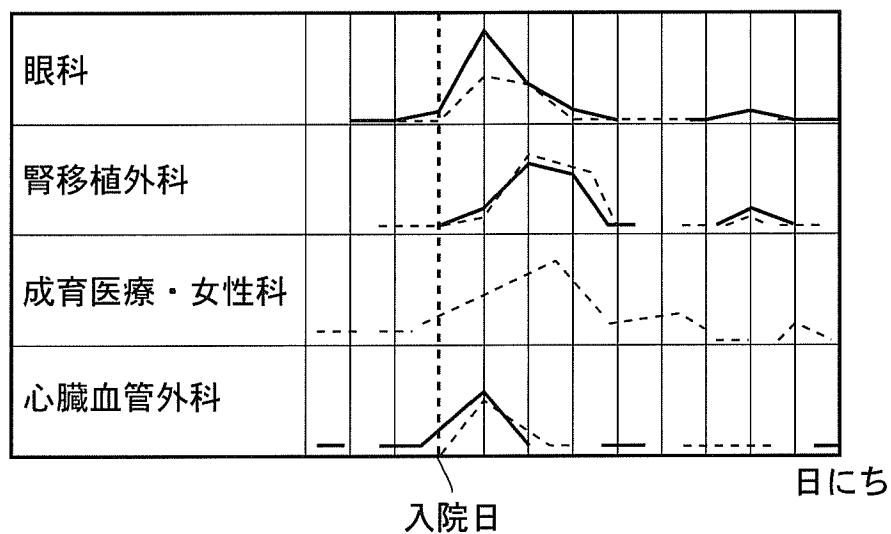
術式	患者数
白内障手術	x_1
胃切除術	x_2
頭蓋内腫瘍摘出術	x_3
大動脈瘤切除術	x_4
...	...

[図7]

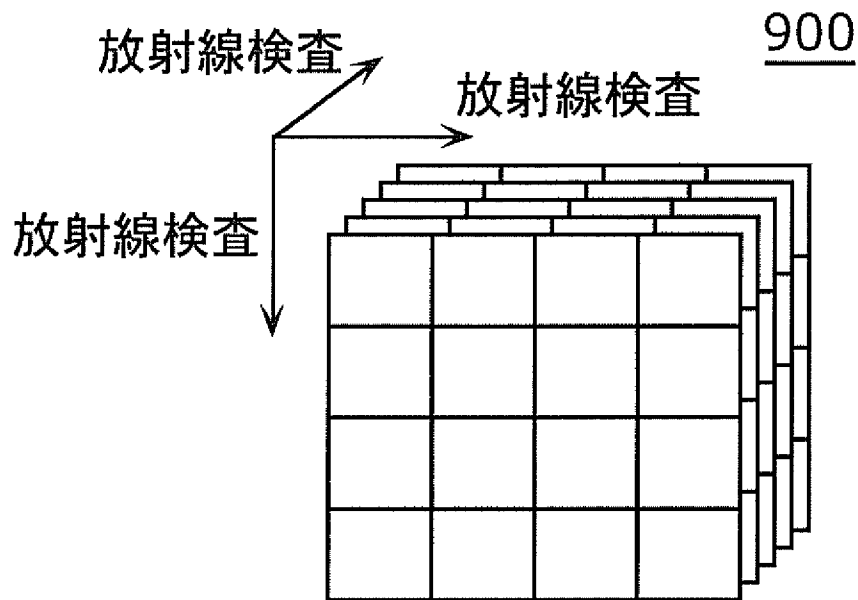
700

	脳神経外科	消化器外科	整形外科	...
2007/1/1	x_{11}	x_{12}	x_{13}	...
2007/1/2	x_{21}	x_{22}	x_{23}	...
2007/1/3	x_{31}	x_{32}	x_{33}	...
...	...	...	...	...

[図8]

800

[図9]

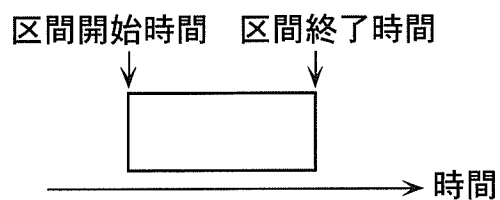


[図10]

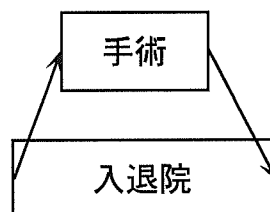
1000

白血球数	患者数
0 -1000	x_1
1000 -2000	x_2
2000 -3000	x_3
...	...

[図11]

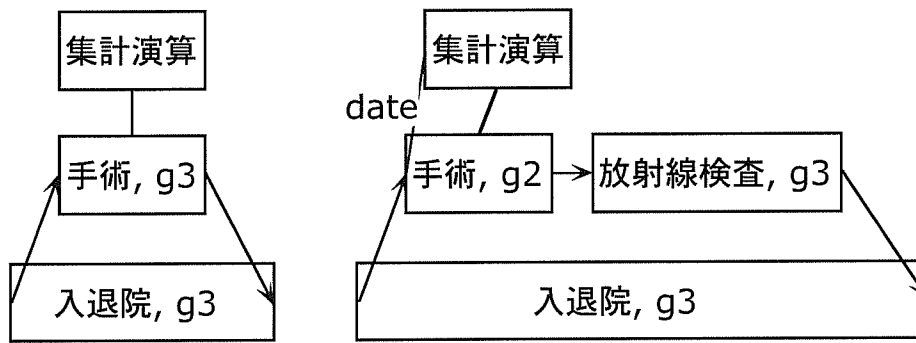


(a) 区間の表現



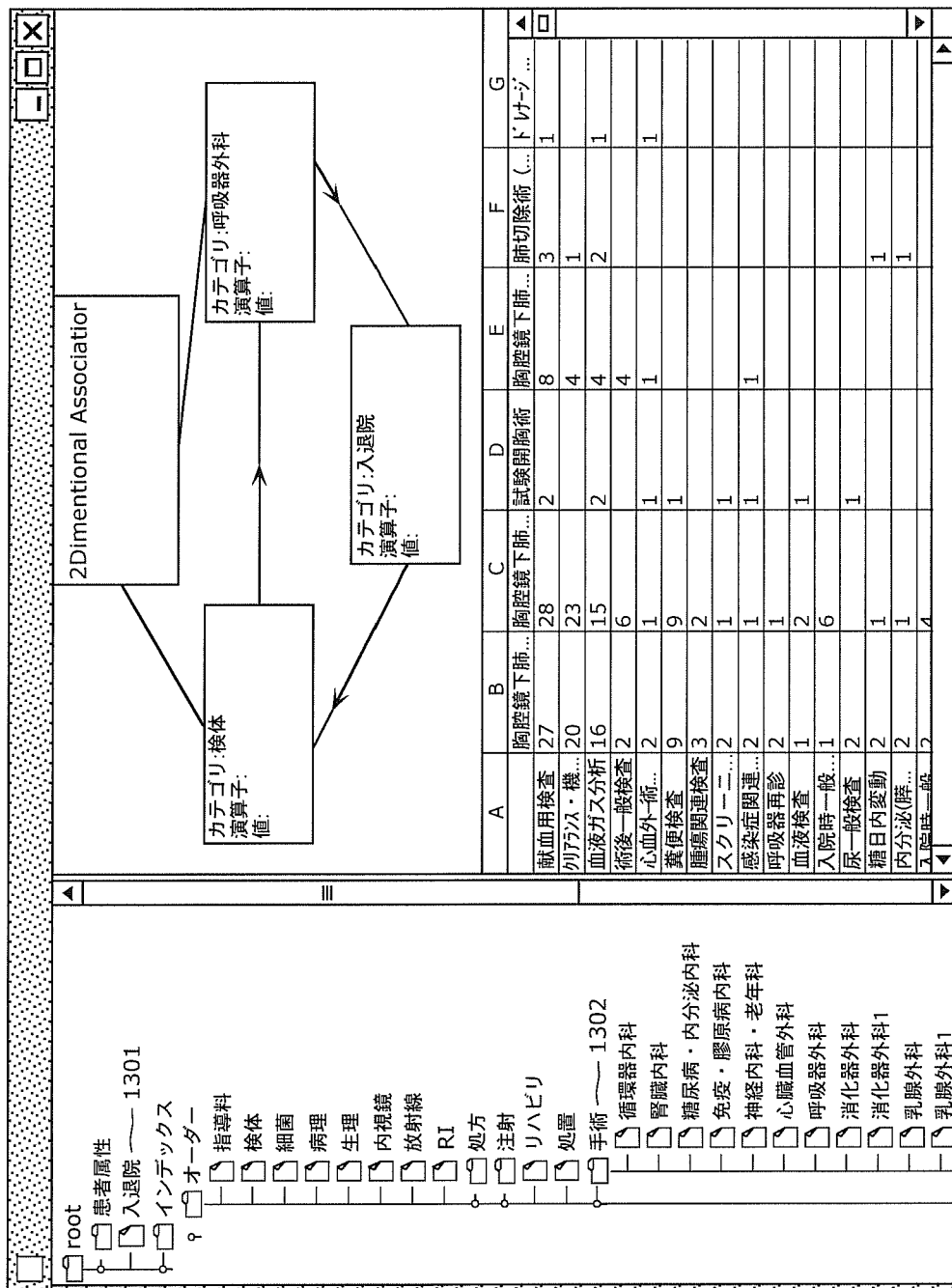
(b) 2区間の関係の表現

[図12]

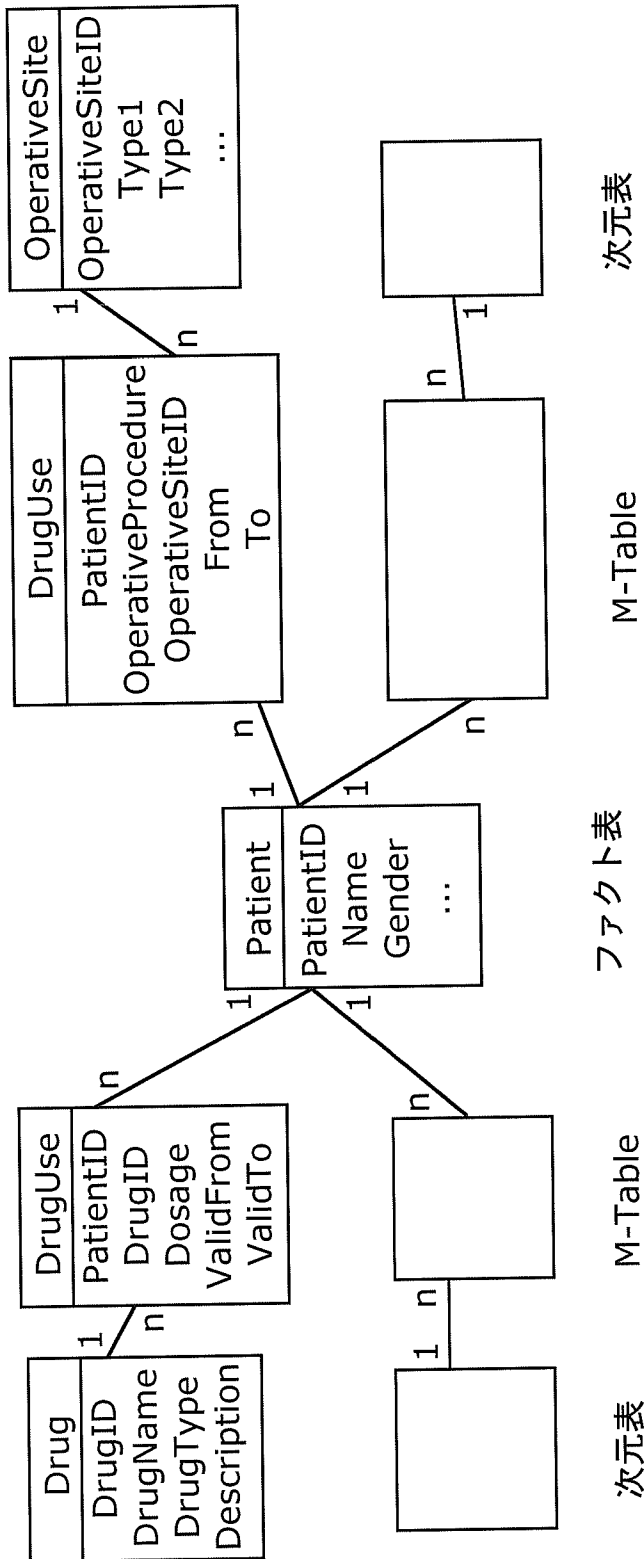


[図13]

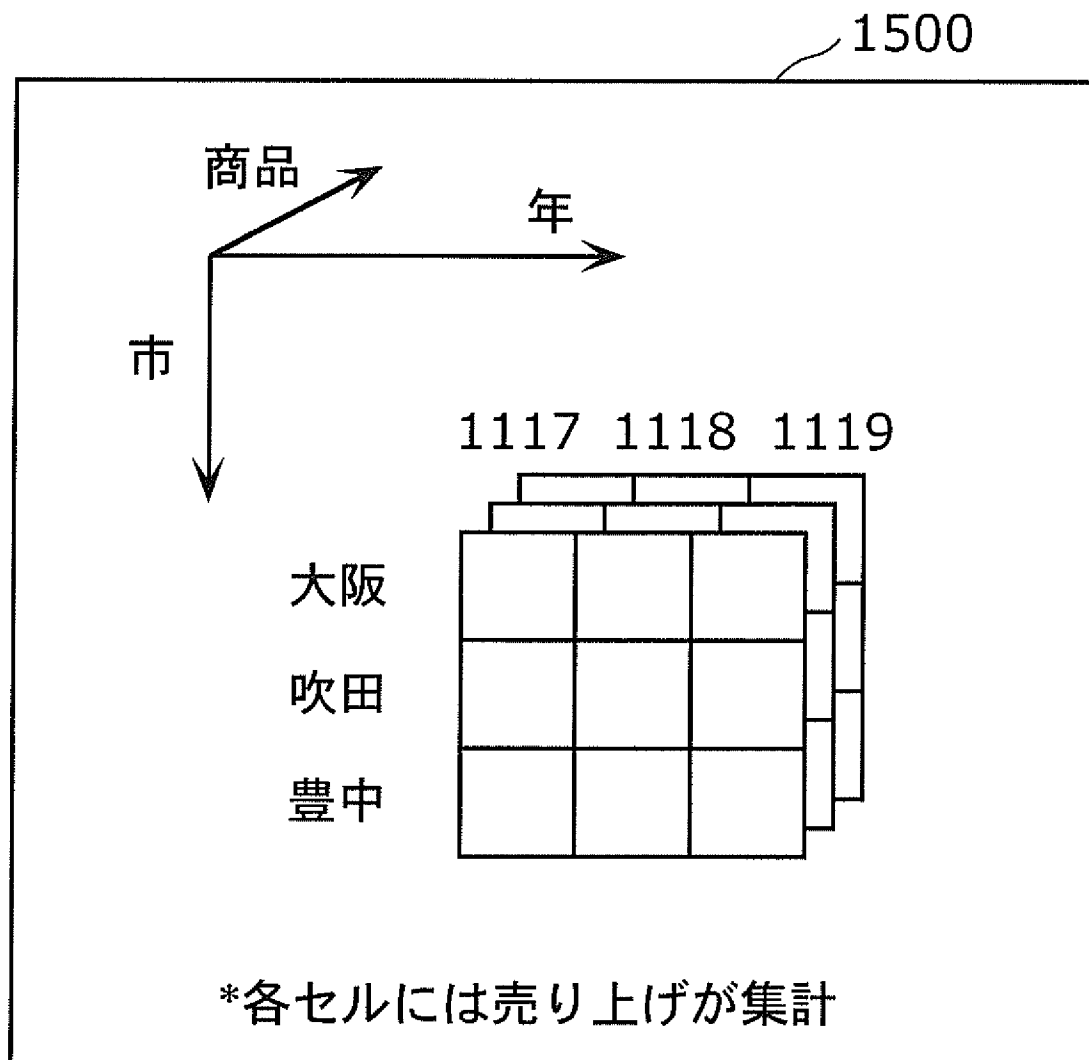
1300



[図14]



[図15]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2008/003366

A. CLASSIFICATION OF SUBJECT MATTER

G06F17/30(2006.01) i, G06F19/00(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)

G06F17/30, G06F19/00

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	1996-2009
Kokai Jitsuyo Shinan Koho	1971-2009	Toroku Jitsuyo Shinan Koho	1994-2009

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.
Y	A. Inokuchi et al., MedTAKMI-CDI: Interactive knowledge discovery for clinical decision intelligence, IBM SYSTEMS JOURNAL[online], Vol.46, No.1, 09 January, 2007 (09.01.07), pages 115 to 133, [retrieval date 09 January, 2009 (09.01.09)], Internet<URL:http://www.research.ibm.com/journal/sj/461/inokuchi.pdf>	1-9
Y	JP 2005-352934 A (Fuji Electric Holdings Co., Ltd.), 22 December, 2005 (22.12.05), Full text; Figs. 1 to 27 (Family: none)	1-9

☒ 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
09 January, 2009 (09.01.09)

Date of mailing of the international search report
20 January, 2009 (20.01.09)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2008/003366

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	Akihiro INOKUCHI et al., "Text Bunseki no Tameno OLAP System", Transactions of Information Processing Society of Japan, 15 June, 2007 (15.06.07), Vol.48, No.SIG11(TOD34), pages 58 to 68	1-9
P,X	Kiyoto TAKABAYASHI et al., "Rinsho Process Kaiseki o Shien suru OLAP System no Kosatsu", Database to Web Joho System ni Kansuru Symposium (DBWeb2007) Ronbunshu [CD-ROM], Vol.2007, No.3, Information Processing Society of Japan, 27 November, 2007 (27.11.07), pages 1 to 8	1-9

A. 発明の属する分野の分類（国際特許分類（I P C））
Int.Cl. G06F17/30(2006.01)i, G06F19/00(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. G06F17/30, G06F19/00

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	1922-1996年
日本国公開実用新案公報	1971-2009年
日本国実用新案登録公報	1996-2009年
日本国登録実用新案公報	1994-2009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 請求の範囲の番号
Y	A. Inokuchi、外3名, MedTAKMI-CDI: Interactive knowledge discovery for clinical decision intelligence, IBM SYSTEMS JOURNAL[online], 第46巻, 第1号, 2007.01.09, p. 115-133, [検索日: 2009/1/9], インターネット< URL:http://www.research.ibm.com/journal/sj/461/inokuchi.pdf>	1-9
Y	JP 2005-352934 A (富士電機ホールディングス株式会社) 2005.12.22, 全文, 第1-27図 (ファミリーなし)	1-9

☒ 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

09.01.2009

国際調査報告の発送日

20.01.2009

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100-8915

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）

鈴木 和樹

電話番号 03-3581-1101 内線 3599

5M

3252

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 請求の範囲の番号
A	猪口明博、外 1 名，テキスト分析のためのOLAPシステム，情報処理学会論文誌，2007.06.15，第 48 巻，第SIG11 (TOD34) 号，p. 58－68	1－9
P, X	高林健登、外 3 名，臨床プロセス解析を支援するOLAPシステムの考察，データベースと Web 情報システムに関するシンポジウム (DBWeb2007) 論文集 [CD-ROM]，第 2007 巻，第 3 号，社団法人情報処理学会，2007.11.27，p. 1－8	1－9