

Meta-Chronicle: マルチデータベース環境における時空間的データ結合方式と人文・科学史への適用可能性

Meta-Chronicle: A Spatial and Temporal Data Integration Method and its Feasibility in Histories

石橋 直樹^{*} 清木 康^{*}

Naoki ISHIBASHI Yasushi KIYOKI

本論文では、ローカルデータベース群のメタレベルにおいて、異種データベース群を連結し、時間的、および、空間的関連性の動的な計量によるデータベース間結合を実現するマルチデータベースシステムとして、Meta-Chronicle システムアーキテクチャを提案する。Meta-Chronicle は、データベース間の時空間的関連性を計量することにより、異種データベース間の時空間的文脈に応じた結合を実現し、時間的、空間的関連性が重要となる複数の情報資源を結合、連携することによって、時間的、空間的共起性という視点からの新しい情報獲得、分析、発見を可能とする。本論文では、Meta-Chronicle を人文・科学史の分野へ適用し、適用可能性を検証する。

In this paper, we propose a multidatabase system *Meta-Chronicle* which interconnects heterogeneous databases by dynamically computing spatial and temporal relationships in a meta-level of local databases. The feature of *Meta-Chronicle* is to realize data integration among heterogeneous databases by computing spatial and temporal inter-relationships context-dependently, and such integration realizes to retrieve, analyze and extract new information generated with a viewpoint of spatial and temporal cooccurrence among legacy databases. An application to histories of distinct domains is also shown to evaluate feasibility of *Meta-Chronicle*.

1. はじめに

急速な広域ネットワークの普及・高速化、および、データベース技術の普及に伴って、多数のデータベース群が広域ネットワークへ接続されている。これら既存のデータベース群を相互に接続する機構は、既存のデータベース群の価値を高めることができる[1,2]。

既存のデータベース群の多くは、時間的、もしくは、空間的データを有する。たとえば、新聞データベースは、特定の空間的所在において、特定の時間に起こった事象に関する記事を有する。本論文では、明示的/暗黙的に時間的、または、空間的意味を有するデータを、時間的データ、または、空間的データ、これらの時空間的データを有するデータベース群を、時空間データベース群と記す。

異文化の衝突とその体系的な評価・分析[3,11]、社会・政治・経済における科学技術の波及効果分析[13]、文学作品を対象とした時空間分析[5]、特定地域における社会変動分析

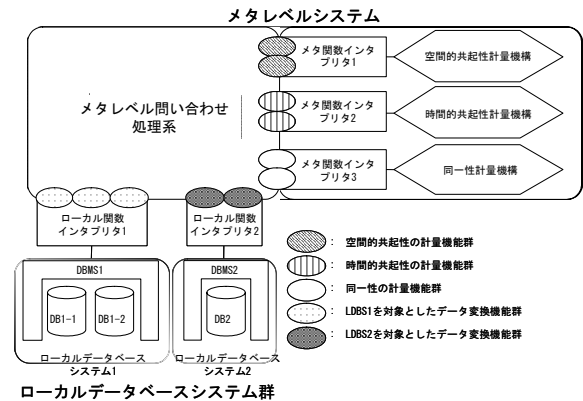


図1 Meta-Chronicle のシステムアーキテクチャ
Fig.1 A System Architecture of Meta-Chronicle

[12]等、多くの社会科学分野において、時空間分析が極めて本質的役割を果たしている。これらの研究は、文献、統計など様々な一時データを対象として、同時性、前後関係、距離性、包含性などの時空間的関連性に応じた時空間分析を適用することで、時空間的事象間に内在する関連性を抽出し、その時空間的関連性から、時空間的事象間の因果関係を示した。

本論文では、時空間的事象間に生じる共起性の計量機構を有するマルチデータベースシステムであるMeta-Chronicleの実現方式、および、人文・科学史を対象とした応用の実現を示す。Meta-Chronicleは次の特徴を有する。

第1に、時空間的事象間の共起性計量機構を用いて、人文・科学史などの時間的、空間的関連性が重要となる複数の情報資源を結合、連携することによって、時間的、空間的共起性という視点からの新しい情報獲得、分析、発見を実現可能とする。複数の専門領域においてそれぞれ構築されたデータベース間を連携することによって得られる情報は、時系列、および、空間分析において極めて重要であり、この情報獲得を実現する枠組みが重要となる。

第2に、人文・科学史などの専門領域において個別に構築されたデータベース群を対象としてMeta-Chronicleを適用することにより、時空間的事象の背景、および、人・物体など時空間的に推移する移動体(以下ムービング・オブジェクトと記す)が経験した環境・状況が動的に抽出可能となる。各専門領域の時空間データベースは、それぞれ、固有のドメインを有し、そのドメインに限定された検索・分析環境が提供されている。これらの時空間データベース群を対象としてMeta-Chronicleを適用し、異なるデータベースの有するデータと時空間的に結合することにより、異なるドメインに属し、かつ、時空間的に共起しているデータ間の連携が実現され、異なる視点、統合的な視点からの時空間分析が可能となる。

また、本論文では、人文・科学史を対象とした応用の実現を示し、Meta-Chronicleの適用可能性を検証する。この応用は、独立に構築された人文・科学史データベース群を対象として、時空間的共起性の計量を適用し、複数の歴史記述を結合することにより、歴史間に内在する知識の抽出を実現する。

2. Meta-Chronicle のシステム構成

異種データベース群の有するデータ間の時空間的共起性を動的に計算する枠組みを実現するために、我々は既存のデータ結合方式[7,8,9]を用いて、時間的共起性計量機構を統合的に有するマルチデータベースシステム Meta-Chronicle を、既存のデータベース群の抽象層(以下単にメタレベルと記す)

^{*} 正会員 慶應義塾大学大学院政策・メディア研究科
naoki@mdbl.sfc.keio.ac.jp
^{*} 正会員 慶應義塾大学環境情報学部
kivoki@mdbl.sfc.keio.ac.jp

に実現した(図 1).

異種データベース群の有する時空間的事象を対象とした共起性計量機構として、以下に示すデータ構造、および、共起性計量機能を設計・構築した。共起性の計量機能は、それぞれ、時間的データ表現間、および、空間的データ表現間の共起性を計量する。ここで、本稿における共起性とは、隣接、内包、包含関係などを含む任意の空間的、または、時間的領域の共有を表す関連性である。

Meta-Chronicle は次のデータ構造、共起性計量機能を用いて時空間的共起性に応じたデータベース間結合を実現する。

$(VSC(1,1))$ 2次元空間上の四辺形, $\{(x_1, y_1), (x_2, y_2)\}$

$(VSC(2,1))$ 1次元空間上の区間, $\{(t_1), (t_2)\}$

$VSC(1,1)$ は2次元空間上のベクター型のデータ構造を示し、 $VSC(2,1)$ は1次元空間上のベクター型のデータ構造を示す。

$(PF_{[1,1]})$ $gt_cooccur(s, gv_1, gv_2) \rightarrow boolean$

$PF_{[1,1]}$ は2次元ユークリッド空間において、その共起性を計量する機能である。 $PF_{[1,1]}$ は、特定の2次元ユークリッド空間 s 上において、 $VSC(1,1)$ によって記述されたベクター値 gv_1 とベクター値 gv_2 が、次式で示す条件を満たす場合、真値を返す計量機能である。式中 G_1° は、ベクター値 gv_1 の内部領域に対応する点集合、 $\mathcal{A}G_1$ は、ベクター値 gv_1 の境界線に対応する点集合を表す[4]。

$$R_g = \begin{cases} true, & (G_1^\circ \cap G_2^\circ = -\emptyset) \vee (G_1^\circ \cap \mathcal{A}G_2 = -\emptyset) \vee \\ & (\mathcal{A}G_1 \cap G_2^\circ = -\emptyset) \vee (\mathcal{A}G_1 \cap \mathcal{A}G_2 = -\emptyset) \\ false, & else \end{cases}$$

$(PF_{[2,1]})$ $tt_cooccur(s, tv_1, tv_2) \rightarrow boolean$

$PF_{[2,1]}$ は1次元ユークリッド空間において、その共起性を計量する機能である。 $PF_{[2,1]}$ は、特定の1次元ユークリッド空間 s 上において、 $VSC(2,1)$ によって記述されたベクター値 tv_1 とベクター値 tv_2 が、次式で示す条件を満たす場合、真値を返す計量機能である。式中 T_1° は、ベクター値 tv_1 の内部領域に対応する点集合、 $\mathcal{A}T_1$ は、ベクター値 tv_1 の境界(具体的には始点と終点)に対応する点集合を表す。

$$R_t = \begin{cases} true, & (T_1^\circ \cap T_2^\circ = -\emptyset) \vee (T_1^\circ \cap \mathcal{A}T_2 = -\emptyset) \vee \\ & (\mathcal{A}T_1 \cap T_2^\circ = -\emptyset) \vee (\mathcal{A}T_1 \cap \mathcal{A}T_2 = -\emptyset) \\ false, & else \end{cases}$$

3. 人文・科学史を対象とした応用の実現

本節では、*Meta-Chronicle* システムにおける共起性計量機能を、実際の多様な人文・科学史のデータベース群に適用し、それらの情報の結合、連携による新たな情報生成のプロセスを示す。人文・科学史のデータベース群を対象として、共起性計量を適用することで、異なるドメインに属する歴史情報間の共起性を発見することが可能となる。本論文で示す応用では、科学技術史 DB (LDB1, 53 タプル)、日本史 DB (LDB2, 1860 タプル)、アメリカ史 DB (LDB3, 854 タプル)、ヨーロッパ史 DB (LDB4, 1751 タプル)、経済史 DB (LDB5, 231 タプル)、映画史 DB (LDB6, 110 タプル)、アインシュタインの個人史 DB (LDB7, 26 タプル) [10] を、*Meta-Chronicle* へ接続することにより、歴史情報間の共起性の獲得、分析システムを構築した。

本応用におけるデータベース間結合の流れを図 2 に示す。人文・科学史データベース群が保有するデータは、対応する空間的領域、および、時区間を明示的に有する。本応用では、まず、これらの時空間的データ表現を、地理空間、および、

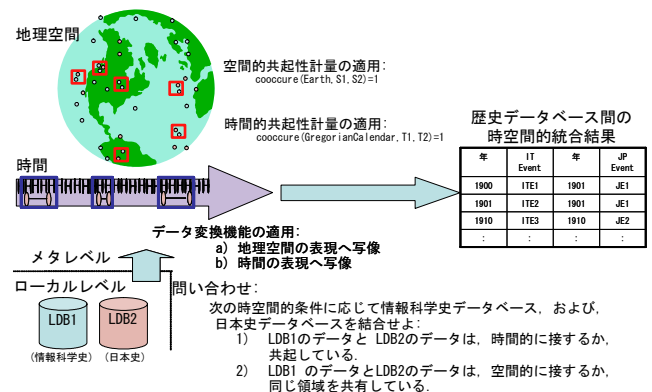


表1 時空間的メタデータ例
Table 1 An Example of the Spatial and Temporal Metadata

時間	事象	時間表現	空間表現
1492	コロンブスの第1回航海, <u>アメリカ</u> 発見	T1dVi {10, (149201010000), (149301010000)}	S2dVr {10, (-179, 51), (-129, 71)}

表2 Q1 結果
Table 2 Results of Q1

時間	事象(情報科学史)	時間	事象(日本史, ヨーロッパ史, アメリカ史を動的に統合)	LDB
1834	バベジが解析機関の構想を樹立 (イギリス)	1833	イギリス, 一般工場法成立	LDB4
1834	バベジが解析機関の構想を樹立 (イギリス)	1833	幕府, 江戸・大阪へ廻米を奨励	LDB2
:	:	:	:	:
1936	チューリングがチューリング機械の理論 (イギリス)	1936 7	スペイン内戦始まる	LDB4
1936	チューリングがチューリング機械の理論 (イギリス)	1936 2-26	二・二六事件起こる	LDB2
:	:	:	:	:
1940	J・アタナソフが世界初のコンピュータを開発 (アメリカ)	1940 5-10	ドイツ軍, マジノ線突破。オランダ・ベルギー、ドイツに降伏	LDB4
1940	J・アタナソフが世界初のコンピュータを開発 (アメリカ)	1940 9-27	日独伊三国同盟調印	LDB2
:	:	:	:	:
1944	暗号解読用の Colossus が完成 (イギリス)。リレー式電算機 <u>Har</u> <u>d Mar1</u> が完成 (イギリス)	1944 7-1~22	ブレトン・ウッズ会議	LDB3
1945	フォン・ノイマン、プログラム内蔵式コンピュータを提案 (アメリカ)	1945 5-8	ドイツ降伏	LDB4
1945	フォン・ノイマン、プログラム内蔵式コンピュータを提案 (アメリカ)	1945 7-15	原子爆弾実験成功	LDB3
1945	フォン・ノイマン、プログラム内蔵式コンピュータを提案 (アメリカ)	1945 8-6	アメリカ軍、広島に原子爆弾投下	LDB2
1945	フォン・ノイマン、プログラム内蔵式コンピュータを提案 (アメリカ)	1945 8-9	長崎に原子爆弾投下	LDB2
:	:	:	:	:
1998	マイクロソフト社が Windows98 を発売 (アメリカ)	1998 4-18	キューバを除く 34 か国による第2回南北アメリカ首脳会議 (米州サミット) がチリのサンティアゴで開幕	LDB3
1998	マイクロソフト社が Windows98 を発売 (アメリカ)	1998 7-13	ブラジルが包括的核実験禁止条約と核拡散防止条約に署名	LDB3

表3 Q2 結果
Table 3 Results of Q2

時間	事象(情報科学史)	時間	事象(日本史, ヨーロッパ史, アメリカ史を動的に統合)	LDB
1879 3-14	ドイツのウルムに生まれる	1878	ドイツのビスマルク, 社会主義者鎮圧法を制定	LDB4
1879 3-14	ドイツのウルムに生まれる	1879	ドイツ・オーストリア同盟成立	LDB4
:	:	:	:	:
1921 11-10	ノーベル物理学賞授与決定	1921 1-12	イタリア、ファシスト党成立	LDB4
1921 11-10	ノーベル物理学賞授与決定	1921 3-18	リガ条約でソ連・ポーランドの国境画定	LDB4
1922 10-8	訪日の途に。11-17日神戸着、日本各地を訪問。12月19日離日	1922 2-6	ワシントン会議終わる。四か国条約・五か国条約・九か国条約調印	LDB2
1922 10-8	訪日の途に。11月17日神戸着、日本各地を訪問。12月19日離日	1922 6-12	加藤友三郎内閣成立 (~23・9-2)	LDB2
1929	統一場理論を提唱	1929 2-11	イタリア、ラテラン協定でバチカン市国成立	LDB4
:	:	:	:	:
1936	妻エルザ没	1936 11-3	ルーズベルト大統領再選	LDB3
1936	妻エルザ没	1936 12	特別汎米平和会議 (ブエノス・アイレス)	LDB3
1939 8-2	アメリカ大統領に原子爆弾製造の研究を勧告する手紙を出す	1939 7-27	アメリカ、日米通商航海条約破棄通告	LDB3
1947	国連総会に世界政府樹立を提唱するメッセージ	1947 3-12	トルーマン・ドクトリン	LDB3
:	:	:	:	:
1955 4-11	ラッセル-アインシュタイン声明に署名 (→バグウォッシュ会議開催のきっかけ)。4月18日永眠	1955 12-6	アラバマ州モントゴメリーでキング牧師の指導下、バス・ボイコット運動始まる	LDB3

```

data_conversion("LDB1",
  {NameToName, "EventName", CS(3, 1), Vn},
  {IntervalToInterval, "Year", CS(2, 1), Vsp},
  {RegionToRegion, "Location", CS(1, 1), Vgp})->V1
data_conversion("LDB2",
  {NameToName, "EventName", CS(3, 1), Vn},
  {IntervalToInterval, "Date", CS(2, 1), Vsp},
  {RegionToRegion, "Location", CS(1, 1), Vgp})->V2
data_conversion("LDB3",
  {NameToName, "EventName", CS(3, 1), Vn},
  {IntervalToInterval, "Date", CS(2, 1), Vsp},
  {RegionToRegion, "Location", CS(1, 1), Vgp})->V3
data_conversion("LDB4",
  {NameToName, "EventName", CS(3, 1), Vn},
  {IntervalToInterval, "Date", CS(2, 1), Vsp},
  {RegionToRegion, "Location", CS(1, 1), Vgp})->V4
unionV(V2, V3)->MV1
unionV(MV1, V4)->MV2
mjoin(V1, MV2,
  {SCR(2), CS(2, 1), cooccure, NULL})->RESULT

```

図 3 Q1: LDB1, LDB2, LDB3, LDB4 を対象とした時間結合
Fig. 3 Q1: Temporal Join among LDB1, LDB2, LDB3 and LDB4

計量機構を用いることにより、動的にムービング・オブジェクトの経験を、他のデータベース群との結合から、抽出可能であることが観測された。表 3 より、アインシュタインが、時空間的推移に応じて経験した事象が抽出されている。`事象(日本史, ヨーロッパ史, アメリカ史を動的に統合)`と記述された属性は、LDB2, LDB3, LDB4 をメタレベルにおいて動的に統合することにより生成され、LDB7 との共起性に応じて結合されている。時空間的メタデータ表現へ写像されたアインシュタインの個人史は、時空間的に推移するムービング・オブジェクトの履歴として捉えられるため、*Meta-Chronicle* は、他のデータベース群との時空間的連結により、そのムービング・オブジェクトが経験した環境、状況を抽出可能であることが確認された。また、*Meta-Chronicle* の有する共起性計量機構を適用することにより、アインシュタインが、不安定な情勢のヨーロッパにおいて幼少期を過ごし、国際的紛争の中、研究を推進していたことがわかる。

5. おわりに

本論文では、異種データベース間を時空間的に連結する *Meta-Chronicle* を提案し、また、人文・科学史を対象とした応用の実現を示した。人文・科学史を対象とした応用において、*Meta-Chronicle* は、個々の専門家が記述した過去に関する知識(すなわち既存のデータベース群)を時空間的に結合、連携し、共起性という視点からの情報獲得、分析、発見を可能とした。

今後は、社会科学の分野において、*Meta-Chronicle* を総合的な時空間分析による情報獲得、分析、発見、総合的な政策形成を目的として適用し、社会科学の学際的インフラストラクチャ形成を目指す。具体的には、美術館の多様なマルチメディアコンテンツを対象とした応用および作品の分析環境を実現する予定である。

[謝辞]

本研究に当たって貴重な御助言を頂いた筑波大学教授の北川高嗣氏、名古屋工業大学講師の細川宜秀氏、慶應義塾大学講師の吉田尚史氏、また、実験システムの構築において御支援頂いた慶應義塾大学の倉林修一氏に感謝の意を表します。本研究の一部は、日本学術振興会学術創成研究プロジェクト「人文社会科学と自然科学を連携するメタレベル知識ベースシステムの開発」によるものである。ここに記して謝意を表します。

[文献]

[1] Bouguettaya, A., Benatallah, B. and Elgarmid, A.: "An Overview of Multidatabase Systems: Past and Present", In Elgarmid, A.,

```

data_conversion("LDB2",
  {NameToName, "EventName", CS(3, 1), Vn},
  {IntervalToInterval, "Date", CS(2, 1), Vsp},
  {RegionToRegion, "Location", CS(1, 1), Vgp})->V1
data_conversion("LDB3",
  {NameToName, "EventName", CS(3, 1), Vn},
  {IntervalToInterval, "Date", CS(2, 1), Vsp},
  {RegionToRegion, "Location", CS(1, 1), Vgp})->V2
data_conversion("LDB4",
  {NameToName, "EventName", CS(3, 1), Vn},
  {IntervalToInterval, "Date", CS(2, 1), Vsp},
  {RegionToRegion, "Location", CS(1, 1), Vgp})->V3
data_conversion("LDB7",
  {NameToName, "EventName", CS(3, 1), Vn},
  {IntervalToInterval, "Year", CS(2, 1), Vsp},
  {RegionToRegion, "Location", CS(1, 1), Vgp})->V4
unionV(V1, V2)->MV1
unionV(MV1, V3)->MV2
mjoin(V4, MV2,
  {SCR(1), CS(1, 1), cooccure, NULL})->R1
mjoin(V4, MV2,
  {SCR(2), CS(2, 1), cooccure, NULL})->R2
intersectionP(R1, R2)->RESULT

```

図 4 Q2: LDB2, LDB3, LDB4, LDB7 を対象とした時空間結合
Fig. 4 Q2: Spatial and Temporal Join among LDB2, LDB3, LDB4 and LDB7

- Rukinkiewicz, M. and Sheth, A. editors, *Management of Heterogeneous and Autonomous Database Systems*, Morgan Kaufmann, pp.1--27 (1998).
- [2] Bright, M.W., Hurson, A.R. and Pakzad, S.: "A Taxonomy and Current Issues in Multidatabase Systems", *Computer*, Vol.25, No.3, pp.50--60 (1992).
- [3] Chomsky, N.: *9-11*, Seven Stories Press (2001).
- [4] Egenhofer, M. and Franzosa, R.: "Point-Set Topological Spatial Relations," *International Journal of Geographical Information Systems*, Vol.5, No.2, pp.161--174. (1991)
- [5] 江藤 淳: *漱石とその時代*, 新潮社 (1970).
- [6] ESRI, *Digital Chart of the World*, 1991.
- [7] 細川 宜秀, 石橋 直樹, 八代 夕紀子, 清木 康: "マルチデータベース環境における時間的・空間的関連性評価によるデータ結合方式", *情報処理学会論文誌:データベース*, Vol.40, No.SIG 8(TOD4), pp.95--111 (1999).
- [8] 石橋 直樹, 細川 宜秀, 清木 康: "時空間的文脈に応じた動的関連性計量機構を有する異種データベース間結合方式," *情報処理学会論文誌:データベース*, Vol.43, No.SIG2(TOD13), pp.128--145. (2002)
- [9] Kiyoki, Y., Hosokawa, Y. and Ishibashi, N.: "A Metadatabase System Architecture for Integrating Heterogeneous Databases with Temporal and Spatial Operations," *Advanced Database Research and Development Series Vol. 10, Advances in Multimedia and Databases for the New Century, A Swiss/Japanese Perspective*, pp.158--165, World Scientific Publishing. (2000)
- [10] Net Advance Inc.: "年表," *Japan Knowledge*, available via WWW from <http://www.japanknowledge.com/>.
- [11] Said, E.W.: *Orientalism*, Vintage (1979).
- [12] 富永 健一: *日本の近代化と社会変動-チュービンゲン講義*, 講談社 (1990).
- [13] 薬師寺 泰蔵: *テクノヘゲモニー*, 中央公論社 (1989).

石橋 直樹 Naoki ISHIBASHI

慶應義塾大学大学院政策・メディア研究科助手。データベースシステム、マルチメディアシステムの研究開発に従事。日本データベース学会、ACM、情報処理学会各会員。

清木 康 Yasushi KIYOKI

慶應義塾大学環境情報学部教授。工学博士。データベースシステム、知識ベースシステム、マルチメディアシステムの研究に従事。日本データベース学会、ACM、IEEE、情報処理学会、電子情報通信学会各会員。