

地理情報標準応用スキーマ設計手法を用いた文化情報学的研究

藤本 悠

Study for Culture and Information Science with Methods of Designing Application Schema (ISO19109)

FUJIMOTO Yu

Abstract: Study for Cultural Phenomena depends on ideal value of each scholar. Therefore, specifying his ideal value is critically important. Information modeling is the effective method for this problem. With using modeling language such as UML, ideal value is strictly designed as Ideal Type Model. Additionally, if the model can be designed by according to Application Schema of ISO 19109, it can be encoded to specific dataset. Because the encoded dataset reflects the ideal aspects, we can understand cultural phenomena with analyzing structure of dataset.

In this paper, I propose the new method of study, ITMA. This method provides the series of procedure from ideal value modeling to analyzing structure of dataset. Actual State, There are two analytical operations for ITMA.

Keywords:

地理情報標準(ISO 191xx Series), 応用スキーマ(Application Schema), 理念型モデル化分析法(Ideal Type Modeling and Analysis), 文化情報学(Culture And Information Science)

1. はじめに

自然現象と異なり, 文化現象と呼ばれる現象は非常に漠然としているため, 文化現象を研究するためには核となる主題を取り出す必要がある. 何が重要であるかという判断は, 個々の研究者の主観的な価値観や理念像に強く依存することになる. そのため, 個々の研究

者がどのような理念像から文化現象を扱っているかを明示することが肝要となる. この問題に対し, 従来は論文などの文章表現を通して明示する試みが行われてきたが (Weber, 1904), 文章による表現では曖昧さが残り, 誤解を生む危険性がある.

文化現象の理念像をより厳密な方法で明示するには, モデリング言語によるスキーマ設計が有効な手段となり得る. UML(Unified Modeling Language)などのモデリング言語は, シンプルかつ厳密な文法によって記述されるため, 文章と比較して誤解を生む危険性は低減される. 文化現象をモデル化するためには, メタモデ

ルの定義やオブジェクト群の準備が必要となるが、その現象が時空間的側面から描きだし得るのであれば、地理情報標準における地物として扱うことができ、モデル化作業は大幅に簡略化できる。また、同標準に準じてモデル化を行うことでモデルに忠実な符号化ができ、文化現象に対する理念像を反映したデータの作成も可能となる。

文化現象に対する理念像を理念型モデルとして定義し、符号化することができれば、文化現象の構造をスキーマ構造から分析することができる。このモデル化から分析までの一連のプロセスを「理念型モデル化分析法(ITMA: Ideal Type Modeling and Analysis)」と呼ぶことにし、本稿ではその概要および開発中の二つの分析法について述べる。なお、本稿では原規格の ISO 191xx シリーズ、JIS x7100 シリーズ、JSGI2.0、JPGIS をそれぞれ区別せず、「地理情報標準」という言葉を用いている。

2. 地理情報標準と文化現象の情報

2.1. 応用スキーマの有効性および設計方法

文化現象の多くは時空間的な現象で、時間と空間によって整理される。したがって、文化現象は地理情報標準が定める地物として定義することが可能で、同標準の応用スキーマ設計手法を用いたモデリングを適用することができる。同標準の応用スキーマ設計手法に従う意義は二つある。一つ目は、メタモデルの定義が不要であることで、二つ目は時空間オブジェクトの設計が不要であることである。この二つを完全に標準に委ねることで、設計者は対象とする現象の設計のみに専念することができ、応用システムと切り離してモデル化作業を進めることができる。

同標準における応用スキーマの設計は非常にシンプルである。最初にモデル化する領域(論議領域)を設定し、その領域内で使われている用語を整理し、UMLを用いてモデル化を行う。モデル化では、抽象化した事象示す用語(クラス名)とその事象を説明する属性を示す用語(属性名)を整理してクラスを生成し、さらにクラス間の関係を UML の文法に従って定義する。このと

き、同標準が準備している時空間オブジェクトは、クラスの属性の型として用いることができる。このような手順で設計したモデルが応用スキーマとなる。設計した応用スキーマは、同標準における符号化規則(ISO 19118)や GML(ISO 19136)といったインスタンス化規則を通して符号化できるほか(地理情報標準推進委員会ほか, 2002), 設計した応用スキーマをさらに正規化し、RDBMS(Relational DataBase System)に実装することもできる(藤本, 2008)。

地理情報標準はあくまで地理空間情報を標準化するためのものであるが、時空間的な視点で整理し得る情報を同標準に従って整理することは、標準化の問題と切り離して考えても有効な手段である。

2.2. 文化現象における情報の特徴とモデル化

地理情報標準の応用スキーマ設計手法に従ってモデル化した文化現象をインスタンス化する場合、対象とした資料によっては実装できないクラスが存在したり、関連先が変わるということがあり得る。この状況は、論議領域を広く設定した場合や、非常にミクロなレベルまでオブジェクトを整理した場合に顕著に現れる。例えば、古代都市を対象とした場合、城壁が存在する都市もあれば、存在しない都市もある。城を中心とした都市が存在するかと思えば、存在しない港湾都市もある。したがって、スキーマ上においてクラスとして定義しても、資料の性質上、城壁や城といったクラスはインスタンス化できないことがある。また、関連に注目して考えると、都市と港とが関連を持つことがあれば、都市が付随しない代わりに砦が関連付けられた軍港も存在し得る。このように、スキーマ上では定義されている関連が選択的になる可能性も存在し得る。

文化現象に関わる情報の特徴は時間にも存在する。文化現象を考える上で「歴史」という言葉は時間を表すキーワードであるが、歴史は時間位相の考え方に極めて近く、歴史上の様々なイベントをノードとした位相構造の中で考えることが多い。先史時代のように時間的な遷移が相対的にしか分からないような場合、土器型式の変遷を時間軸に当てはめる。しかし、土器型式は

地域によって複数型式が並行して存在したり、消失した型式が別地域では存続し続けたりといった状況がある。このような場合、イベント間の関係は相対的かつ位相的な時間で整理することができる。絶対的な時間が分かる歴史時代においても時間位相の考え方は重要で、別々の原因で発生した異なる諸事象が一つの事象に結びついたり、ある一つの事象から複数の諸事象が派生していく過程を観察することができる。つまり、歴史的なイベントの重要性を他のイベントとの関係から評価することが可能となる。このように、時間位相的な関連を重視して文化現象をモデル化できる。

現状の文化現象研究では、スキーマのインスタンス化時に生じる欠失や関連先クラスの違いといった形質的構造や、時間位相によって関連付けられたイベントといったものは、型式や形式、様式といった言葉で表現された類型基準によって整理されているが、類型基準を設けることは過度な細分化を招くだけでなく、属性値として整理することも困難にする。しかし、スキーマ構造を通して見てみると、類型基準を設ける必要はなく、常に全体の中で個々の事象の位置づけを析出することができる。

3. 理念型モデル化分析法(ITMA)

3.1. ITMA の概要

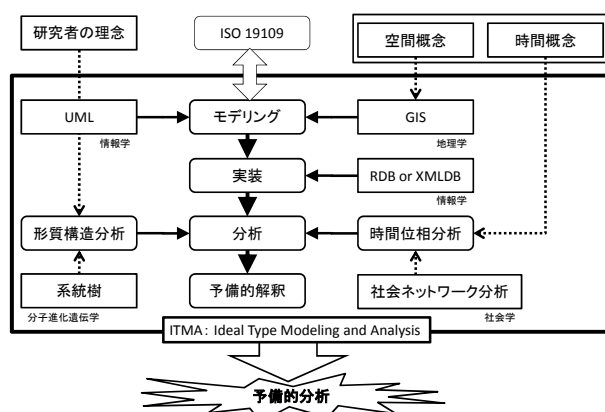


図 1 ITMA における分析

ITMA は、個々の研究者の文化現象に対する理念像をモデル化し、モデル構造を通して分析するための一連の方法を指し、情報の標準化はその目的ではない。したがって、地理情報標準の応用スキーマ設計手

法を用いていても ITMA における理念型モデルは同標準が定める応用スキーマではない。ITMA の特徴は、理念像の写像と明示が目的である点と、設計するモデルを使用して直接的にデータベースに実装できる点、そして実装したデータベースのインスタンスを理念像に従って分析することである。ここで重要なことは、「モデル設計の意義」と「実装化の意義」と「分析の意義」はそれぞれが独立している。これは、非常に重要で、研究の過程で収集した情報とそれを格納したデータベースの運用性や有効性は分析結果の良し悪しに依存しないことを意味する。

ITMA は、モデル化、実装化、分析という三段階の手順から成り、モデル化の段階では、地理情報標準に準じたモデル設計を前提としている。しかし、これは現状において最も理想的であるためで、メタモデルの定義や時空間オブジェクトの定義が可能であれば、同標準に従う必要はない。したがって、より理想的に理念像を写像できるモデル化言語が存在する場合、UML 以外のモデリング言語の使用も可能である。実装化の段階では、設計したモデルの構造を維持したまま符号化できるのであれば符号化の方式を限定するものではが、現状では安定性の高い RDBMS や地理情報標準に準じた XML による符号化の方式が現実的である。ITMA における分析では、モデル構造を利用した分析は全てを認めるし、現在開発を進めている形質構造分析と時間位相分析以外の分析手法も否定しない。しかし、現状において、ITMA が地理情報標準の応用スキーマ設計手法を用いているのは、文化現象が地理空間的な視点で考察されることが多く、同標準に準じることで、より理想的なモデル化を実現できるためである。

3.2. ITMA における形質構造分析

文化現象研究では、インスタンス化時に生じるクラスの欠損や相異なる関連先クラスを通して、それぞれのインスタンスの親近性を評価する。型式や形式、様式といった類型基準は、親近性の高いものを格納するためのコンテナのような役割を果たす。したがって、ITMA における分析では、インスタンス間の親近性を評価す

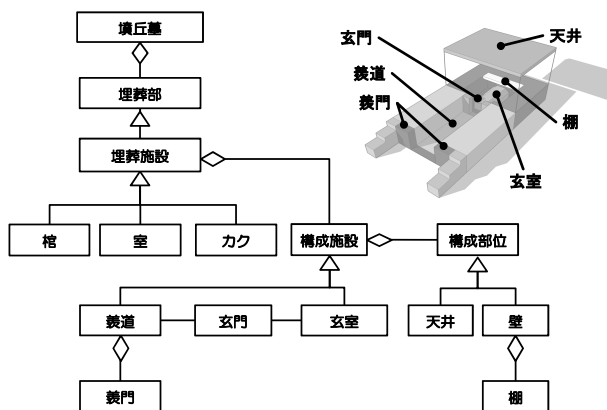


図 5 横穴石室の理念型モデル

るための分析手法として「形質構造分析」を準備する。インスタンス間の親近性の評価する分析方法には、対応分析などの統計学的方法があるが、ここでは分子進化遺伝学における系統樹の作成手法を利用することにした。

系統樹の作成では、異なる遺伝子間の中に現れる置換(Substitution)、欠失(Deletion)、挿入(Insertion)、転換(Inversion)という四つの状態を検出して系統樹を作成する。系統樹の作成手法は大きく二つに大別でき、進化距離ないし遺伝距離による計算から系統樹を得る方法(「距離行列法」と、現存種の塩基配列アミノ酸配列から先祖種を推定して系統樹全体における進化的変化の数を最小化して系統樹を得る方法(「最大節約法」とがある(根井, 2002)。本稿では、確立済みの分析手法として、「欠失」の状態から文化現象の形質構造を分析する方法について述べる(藤本, 2008)。

インスタンス化における「欠失」の状態は、XML によって符号化を行うことで簡単に検出できる。古墳時代における横穴石室を例に、UPGMA による系統樹作成の具体的方法を示しておこう。まず、横穴石室という文化事象を図 2 のようにモデル化する。このスキーマはあらゆる状況に対応しうる論理的に矛盾の無いモデルである。ここでは対象資料 A と B の存在を前提に、それぞれをインスタンス化する(図 3)。インスタンス A は、スキーマ上で定義されたすべてのクラスを実装できているが、インスタンス B は羨門、玄門、棚といった部位が先天的に存在していないから実装できない。この両者の違いは、地域差であるかもしれないし、時期差かもしれ

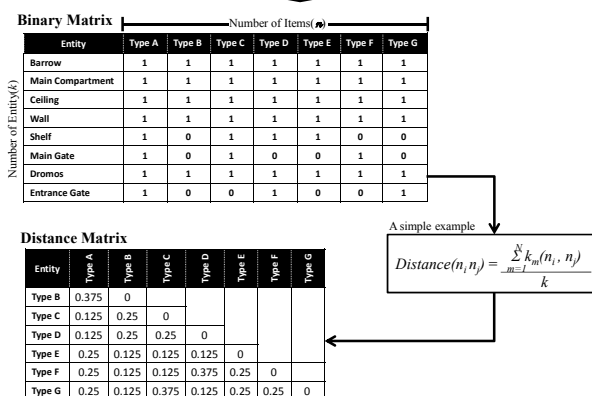
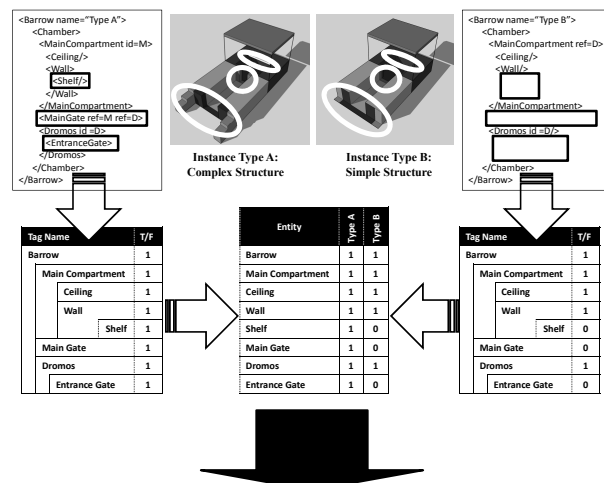


図 2 UPGMA による系統樹作成の手順

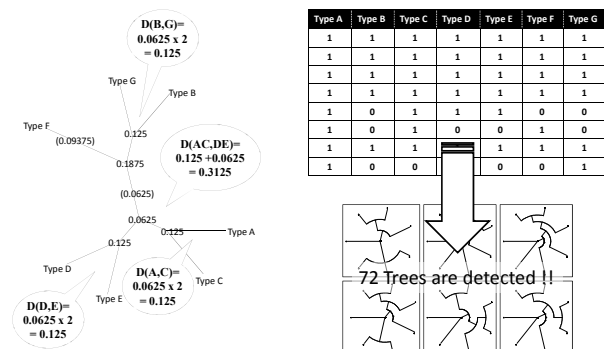


図 3 UPGMA から得られた系統樹の例 図 4 最大節約法による系統樹の例

ない。実装したタグの実装状況を検出した結果は全て「1」と「0」に置き換えることができる。同様に他の石室に関してもタグの有無を検出していけば、距離行列を得ることができる。系統樹はこの行列から導くことができる(図 4)。この例では全サイトに対する置換数を用いて距離行列を作成したが、タグ階層性や実装されたデータの属性値から重みづけを与えたり、適切な進化距離

系統樹の作成方法は、UPGMA 以外にも多く存在し、系統樹作成のためのパッケージも数多く存在する。距離行列法を用いたパッケージとしては本研究で用いた SplitsTree4 があり、このパッケージでは、異なる系統樹を組み合わせる SuperNetwork や ConsecusTree を作成することもできる(Huson ほか, 2004)。また、「最大節約法」による分析では、PAUP*というパッケージがあり、「0 | 1」のデータやカテゴリーな配列データから系統樹を作成することができる。なお、この例で用いたデータを PAUP*で分析した結果、72 の系統樹が作成された(図 5)。

地理情報標準では、時間位相に関する情報もオブジェクト(TM_TopologicalPrimitive)として定義されている(村尾ほか, 2008). これを用いることで、異なる文化現象間の時間位相から時間位相ネットワークの構築が可能となり、これによって歴史的イベントの重要性を評価できる。ネットワーク分析によって歴史的なイベントを評価する方法としては中心性分析がある。中心性の評価にはノードに結び付けられたエッジ数から算出する次数中心性や媒介中心性といった分析がある(Wasserman, S et al, 1994). 図6は中心性評価の例である。図中、四角で表したものが時間ノードで、ノード間を結ぶ線分が時間エッジである。この例では架空のデータを用いたが、実際のデータを扱う場合、このノー

ITMAは文化情報学における研究手法として重要な手法であるが、根本的に大きな問題を抱えている。文化現象の電子化の問題である。ITMAはその性質上かなり広範囲かつ詳細な情報を必要とするが、現状ではこの要件を満たすような総合的な文化情報データベースは存在しない。たとえば、考古学では未だに紙媒体の報告書の情報が主流で、電子化されている情報は報告書のメタデータ程度でしかない。また、報告書自身の情報はページ数という物理的制限があり、記載されていない情報を入手することは極めて困難である。このような状況では、ITMAを用いても効果的な分析は

できない。この問題を解決するためには、調査の過程で蓄積した全データを体系的に蓄積する仕組みと、ITMA で設計した理念型モデルと調査データのモデルを相互に変換するような仕組みも必要となる。

今後は、ITMA における分析パッケージの充実化を図ると同時に、文化現象に関わる調査業務からシームレスに情報構築する仕組みやモデル変換の仕組みについても検討したい。

参考文献

- [1] Huson, D. H., DeZulian, T., Klopper, T., & Steel, M. A. (2004). *Phylogenetic Super-Networks from Partial Trees*. IEEE/ACM Transactions on Computational Biology and Bioinformatics, Vol.1, No.4, pp.151-158, IEEE CS, CI and EMB Societies & the ACM.
- [2] Wasserman, S., Faust, K. (1994). *Social Network Analysis: Methods and Applications*, Cambridge University Press.
- [3] Weber, M. (1904). Die "Objektivitat" sozialwissenschaftlicher und sozialpolitischer Erkenntnis. (祇園寺信彦, 祇園寺則夫(訳) (1994). 『社会科学の方法』, 講談社学術文庫.)
- [4] 根井政利 (2002) 『分子進化遺伝学 Molecular Evolutionary Genetics』, p.23, 培風館.
- [5] 地理情報標準推進委員会・国土交通省・国土地理院 (2002) 『地理情報標準第二版』, 国土地理院 技術資料 A・1・No.257. 国土地理院.
- [6] 藤本悠 (2008) 学際研究としての文化情報学研究 -考古学における遺跡情報モデリングとその実装-, 「文化情報学」, 第3巻, 第1号, pp.15-24, 同志社大学文化情報学会.
- [7] 藤本悠. (2008) 文化現象研究における OOGIS 理論の有効性-考古学における古墳時代研究での実践的検討-, 「情報処理学会研究報告 2008-CH-78 人文科学とコンピュータ」, Vol.2008, No.47, pp.1-8, 情報処理学会.
- [8] 村尾吉章・碓井照子・森本晋・清水啓治・藤本悠・山本由佳 (2008) 遺構情報モデルへの時間スキーマ適用法の検討, 「地理情報システム学会講演論文集」, 地理情報システム学会.