

異種地理データベース間での地図合成のためのオブジェクト併合

6U-3

高倉 弘喜[†] 濱野 寿彦[‡] 有川 正俊^{*} 上林 彌彦[†]

京都大学工学研究科[†] 京都大学工学部[‡] 広島市立大学情報科学部^{*}

1 はじめに

現在、インターネット上や駅のタウンガイドシステム等、様々な場面で地図を用いた情報サービスが行われている。このように計算機上で地理情報を取り扱い、その応用範囲を広げていく上では、いかにして既存のデータ資産を活用していくかが重要となる。そのための有効な手段として、複数の異種地理データベースを統合することが挙げられる。これにより、現在使用している地理データベースに外から地理データを取り入れ、地理データベースの拡充を図ることができる。

本稿では、異種地理データベースを効果的に統合するための手法を検討する。まず地理オブジェクトに求められる形式について考察する。地理オブジェクト (Geographic Object: GO) とは、実世界の個々の地理実体に対応する地理データベースでのデータ要素である。次に、それらの地理オブジェクトを併合する手法について検討する。

2 統合対象としての地理データベース

2.1 地理データベース

本稿で言う地理データベース (地理DB) とは、地理情報が数値化されたベクトルデータの集合体とする。ベクトルデータの中にも大別して、(A) 実世界を写像する目的で作られた歪みの小さな存在、(B) 利用者の目的に合わせて意図的に大きく実世界を歪めた略地図的存在、とがある。このうち、本稿では統合の対象として次のものを想定する。基盤となる地理DB上に、利用者が任意にGOを追加することにより構成された、基盤拡張形の地理DBである。その基盤部分は、(A) の様に歪みの小さな特定のデータセットを元に構成された地理DBである。即ち基盤部分は統合対象間で共通である。地理DB統合の問題をより高い一般性をもって扱うためには、(A) と (B) の様に全体に渡って異なる地理DBを取り扱うことは重要である。しかし、統合対象を上記のように基盤を揃えたものに限定することは無意味ではない。アメリカではTIGER/Line Files、日本では国土地理院やゼンリンの数値地図等、実際に基盤となりうるデータセットが存在するからである。また、地理データの標準化の動きも世界的に進みつつある。このような状況を踏まえると、本稿での地理DBの取り扱い決りは決して特異なことではない。同じデータセットの中に縮尺や作成時期が異なるサブセットが存在することはあるが、これは基本的なデータ形式が同等であるため対応が可能である。また、ここでの技術がより一般性の高い場合にも還元され、基盤が異なる場合であっても基盤自身を統合する方向で拡張することが考えられる。

本稿では、統合対象となる複数の地理DBを単に統合対象と呼ぶ。統合により生成された新たな地理DBを統合DBと呼ぶ。また、地理DBの“無矛盾な統合”とは、以下の条件を満たすこととする。

- 統合DB中に同じ地理実体を表すGOが複数存在しないこと。

- 利用者が併合不要と指定していないGOが、統合後に消失しないこと。
- 統合の前後でGO間の位置関係に矛盾が生じないこと。矛盾の例として、併合前に道路Sの南側にあった駅Kが併合後では北側にある、ということがある。但し、あらかじめ統合対象間で矛盾が存在した場合は、この限りではない。

2.2 地理オブジェクト (GO)

統合対象中の個々の地理データ要素を地理オブジェクト (GO) と呼ぶ。一般にGOは複数の属性項目を持つ (2.3節にて詳述)。また、GOが地図上に表示される際の幾何的な形態には、点、線分、ポリライン、ポリゴン (領域) がある。これらはいずれも点の集合として定義される [1]。なお、幾何データが点であるGOを点オブジェクトと呼ぶ。線分、ポリライン、ポリゴンを構成する点は、単独ではオブジェクトとして扱われない。

2.3 地理オブジェクトの形式

GOにも、それを構成する属性項目によって様々な形式のものが考えられる。属性項目には以下のようなものが考えられる。以下では [] は属性項目を表す。

- [ID]: 地理DB内でGO個々を一意に指す識別子。
- [座標]: GOの幾何データ (2.2節参照)。
- [名前]: GOを表現する名前。

単純なGOはこれらだけで構成されるが、これらの属性項目に入る値は、数値や文字列等、同一性判定 (3.1節参照) に際して、曖昧さを生じるものである。数値の座標値が完全に一致することは稀である。また、同一物を表す名前にも様々な表現がある。例えば、“京都大学”、“京大”、“Kyoto University” は同じ地理実体を表す。逆に、同じ名前のものが異なる地理実体を指すこともある。

属性項目の数が少ないものや、各項目値の詳細度、精度が低いものはGOとして低品質である。低品質なGOほどこの曖昧さが大きくなり、無矛盾な統合に支障をきたす。そこで以下の属性項目が要求される。

- [クラス]: GOの種類。例えば“鴨川”、“宇治川”は共に“川”というクラスに属する。これを照合することにより同一性判定における曖昧さをある程度解消できる。
- [位置関係]: GO間の位置の相対的な関係 (図1参照)。
[座標] は個々の地理DB内に固有の値であるため、無矛盾な統合のためには、あらかじめGO個々が周辺のGOとの[位置関係]を記憶しておくことが望まれる。記憶対象として、点オブジェクトの場合、それを含む最小ポリゴン、線に接する場合にはその情報等がある。

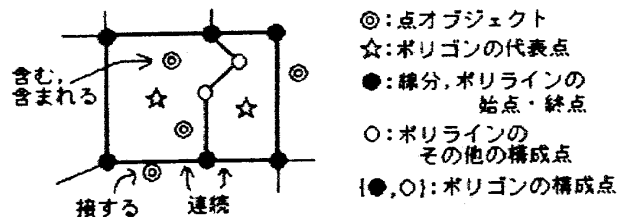


図1: GOの位置関係, 点の性質

- GOとは比較階層が異なるが、次の属性も必要となる。
- [性質]: GOの幾何データを構成する個々の点の役割 (図1参照)。これは、点オブジェクト以外の点を併

Merging Objects for Generating Maps from Heterogeneous Geographic Databases

Hiroki TAKAKURA[†], Toshihiko HAMANO[‡], Masatoshi ARIKAWA^{*} and Yahiko KAMBAYASHI[†]

Department of Information Science, Kyoto University[†],

Faculty of Engineering, Kyoto University[‡],

Faculty of Computer Science, Hiroshima City University^{*}

合する際に照合される。

また、利用者要求を的確に取り入れ、利用者にとって信頼性の高い統合処理とするため、以下のようなメタデータの存在も望まれる。これらは主として、統合対象間で共通するGOを併合する際に参照される。

- [品質] : GOの信頼性の尺度となる。上記の地理的属性各項目についての詳細度、精度から設定される。
- [作成者] : GOの作成者で個人または組織。信頼性の尺度となるほか、利用者の特定にも利用可能である。これが組織であれば、基盤部分の作成元である場合もある。
- [作成時] : GOが作成された時期。信頼性の尺度となる。[作成者]が同じであれば、これが新しい方を信頼性が高いものとして取り扱う。利用者の追加したGOだけでなく、基盤部分についてもこれにより作成時期を保持する。
- [目的] : GOが作成された理由、GOの利用目的。
- [重要度] : GOの利用者における必要性の尺度。

最後に、備考と今後の拡張のため[その他]を用意する。なお、関連情報を参照可能とするために、地図上からインターネット等にアクセスするための、リンク先を保持させるのも実用上有効である。

3 異種地理データベースの統合手法

3.1 点オブジェクト同士の併合

- i) 統合対象間で、[名前]、[クラス]、[座標]が全て一致するGOの存在を確認する。[名前]については、一致判定に際して、正式名称、略称、別称等様々な表現に対応できる柔軟性が必要となる。但し、これは地理DBに限った問題ではなく、データベース統一般に生じることである。[座標]については、一定の範囲を定め、その許容内であれば一致とする。また、[クラス]、[座標]が一致して[名前]が違う場合等、判定が曖昧な場合には[位置関係]も参照する。
- ii) 同一判定されたGOは、次のように再構成して統合DBに格納する。
 - [座標]、[名前]については、利用者の選択要求が無い場合、[重要度]、[目的]及び信頼性から総合的に判断して優先するものを決定する。
 - [ID]を統合DB中で唯一のものとする。
 - [クラス]、[性質]は共通事項で不変である。
 - [品質]を再設定する。
 - 他の属性項目については次の方針を挙げる。
 - [座標]決定の判断で優先されたものを採用
 - 属性値の精度、詳細度が高いものを採用
 - 利用者の選択する属性値を採用

なお、[目的]、[重要度]は使用状況に合わせて利用者が設定可能とする。

- iii) 同一判定されなかったGOは、統合対象の中で唯一のものである。これらのGOは、統合DBの拡充を図るために取り入れる。
- iv) 以上のステップで、統合DBに格納したGOが不適切であった場合、現状では、その対処を利用者に委ねることとする(4章の(8)参照)。

3.2 点オブジェクトと非点オブジェクトとの併合

同じ地理実体を表すGOであっても、地図上の表現方法によってGOの幾何データは異なることがある。例えば、ある利用者は学校を点で表現し、別の利用者はそれを領域で表現するかもしれない。

このように表現が異なるGOを併合する場合、

- 1) GOの同一性判定を行う。ここではまず[座標]が評価対象から外れ、[位置関係]が重視される。他は点オブジェクト同士の場合と同等である。ここで同一

の候補が存在した場合に限り、[座標]に関して、点と線分の距離、点と領域の包含関係等を評価する。

- 2) 同一地理実体を表す異形態のGOが存在した場合、利用者の[重要度]、[目的]に従って幾何データを選択する。それが判断出来ないときには、そのGOには複数の幾何データを保持させ、信頼性を加えた判断により優先順位を付ける。表示の際にどの幾何データを使用するかは、優先順位及びその時々地図への要求に応じて表示機構が選択する。

4 プロトタイプシステム

以下のシステムを実装し、WWWブラウザから参照可能とする予定である。

- (1) 国土地理院数値地図2500を用いて、基本図(道路や川、学校etc)を表示。
- (2) 利用者は基本図上に好きなGO(店、家etc)を点、線、領域の任意の図形で描き込む。その際、GOの属性として[名前]の記入、[クラス]の選択、[作成者]、[目的]、[重要度]、[その他]の記入、表示色の決定等をして、GOを確定する。
- (3) システムでは、確定されたGOについて、描き込まれた図形の[座標]、周辺のGOとの[位置関係]、図形中の個々の点の[性質]を割り出し、これに[ID]、[作成時]、[品質]、利用者が記入した各種属性を付加してシステムの内部オブジェクトを生成する。
- (4) (2)、(3)の繰り返しにより、利用者は複数のGOを地理DBに登録する。
- (5) (4)を複数の人が行うことで、複数の異種地理DBが生成される。
- (6) (5)で生成された複数の異種地理DBについて、各内部オブジェクトを併合し、1つの統合DBを生成する。この際、利用者は統合対象の個々について、併合が不要であるGOを指定できる。また、特定のGO間の対応関係をあらかじめ定義することも可能とする。
- (7) 統合DBを用いて地図を表示する。
- (8) GOの併合結果が利用者の意図に合わない場合に備え、システムには誤り訂正機能が必要となる。具体的には、次のような事を考える。

- 利用者が結果を判断して、手動でGOを修正する。または、GO間の対応関係を定義して、統合処理を再実行する。
- システムがGOの併合時に、疑わしいものをチェックし、データベースの統合中、または統合後に利用者に問い合わせる。

5 おわりに

本稿では、異種地理データベースを矛盾無く統合するために求められる地理オブジェクトの形式及び主に点オブジェクトに関する併合手法について検討した。さらに線オブジェクト同士や領域オブジェクト同士の併合手法が今後の課題である。

また、本稿では取り扱いの対象としなかったが、縮尺や作成元が異なるデータセットを元とした地理データベースを統合する際には、統合対象を同じ観点から比較するため、当該地理データベース内での相対座標系を絶対座標(緯度、経度)系に変換する必要が生じる。このような場合の手法も今後の発展的課題である。

謝辞

種々の御助言を頂きました上林研究室の皆様へ感謝致します。なお、本研究は文部省科学研究費補助金基盤研究(A)(2)展開により援助を受けている。

参考文献

- [1] 有川正俊, “地理情報システムにおける空間表現・ユーザインタフェース”, 日本ソフトウェア科学会チュートリアル 地理情報システム, 1997.10.27.