

Ein Modell zur Abbildung topologischer Beziehungen in GIS

Norbert Rösch

Universität Karlsruhe, Geodätisches Institut

Englerstraße 7

76128 Karlsruhe

E-Mail: roesch@gik.uni-karlsruhe.de

Zusammenfassung

Es wird ein Konzept zur Abbildung und Speicherung von ein-, zwei- und dreidimensionalen topologischen Beziehungen in Geo-Informationssystemen (GIS) vorgestellt, das auf dem Modell der Quasi-Zellenzerlegung (q -Zellenzerlegung) basiert. Diese Zerlegung ist das Ergebnis einer Anpassung der aus der algebraischen Topologie bekannten Zellenzerlegung an die Anforderungen in GIS. Das Modell berücksichtigt die wesentlichen topologischen Invarianten, so daß der Raum mit Hilfe der Beziehungen zwischen den Elementen der q -Zellenzerlegung, den q -Zellen, hinreichend beschrieben ist. Die Zerlegung des Raumes in Verbindung mit der Abbildung der topologischen Relationen seiner Elemente ermöglicht die *Analyse* ebenso wie die Prüfung auf *Konsistenz* und *Plausibilität* dieser Beziehungen.

Dabei ist zu berücksichtigen, daß die q -Zellenzerlegung im Unterraum \mathbb{R}^p ($p = 2, 3$) durchgeführt wird und lediglich strukturelle Beziehungen der q -Zellen untereinander abgebildet werden. Damit kann beispielsweise eine zielgerichtete Analyse der topologischen Relationen nur unter Hinzuziehung der Semantik erfolgen. Die Gesamtzahl der Beziehungen bleibt dadurch verhältnismäßig gering.

In diesem Modell können alle topologischen Relationen als Mengenbeziehungen betrachtet werden. Damit lassen sich sämtliche Operationen auf Mengenoperationen zurückführen. Berücksichtigt man diese Eigenschaft bei der Implementierung und verwendet zur Abbildung der Beziehungen geeignete Mengendatentypen, dann kann die Bearbeitung des q -Zellenkomplexes sehr einfach über Mengenoperatoren erfolgen.

1 Einleitung

In Ermangelung geeigneter Werkzeuge werden vor allem zwei- und dreidimensionale topologische Beziehungen oft als beschreibende Attribute geführt oder bei Bedarf aus geometrischen Beziehungen abgeleitet. So wird u. U. die Information über die Zugehörigkeit einer Stadt zu einem bestimmten Land als beschreibendes Attribut abgelegt oder im Bereich der Analyse wird die Frage nach einem Biotop innerhalb einer Gemeinde über den Point-in-Polygon-Test geklärt. Diese Lösungen sind teilweise nicht nur umständlich und zeitaufwendig, sondern sie unterlaufen auch das Konzept der klaren Trennung zwischen semantischer, geometrischer und topologischer Datenmodellierung.

Während topologische Modelle, die auf graphentheoretischen Ansätzen beruhen, schon sehr weit entwickelt sind, besteht bei weitergehenden Modellen noch Forschungsbedarf. Dies betrifft in erster Linie den \mathbb{R}^2 , da die meisten GIS derzeit noch in diesem Raum arbeiten. Denkt man an zukünftige Entwicklungen, dann wären Implementierungen im \mathbb{R}^3 durchaus wünschenswert. Geht man noch weiter, kommt zwangsläufig die Zeit als weitere Dimension ins Spiel. In den folgenden Abschnitten wird ein Modell vorgestellt, mit Hilfe dessen viele der oben genannten Nachteile vermieden und Anforderungen bzgl. topologischer Beziehungen demgegenüber erst ermöglicht werden.

2 Die q-Zellenzerlegung

Das Modell der q-Zellenzerlegung basiert auf einer Verallgemeinerung der aus der algebraischen Topologie bekannten Zellenzerlegung, die auf J.H. Whitehead zurückgeht ([WHITEHEAD 1949]). Die Gesamtheit aller Zellen wird dabei als CW-Zellenkomplex bezeichnet. Während dieses Modell allerdings von einer tatsächlichen Zerlegung des Raumes in einzelne Punkte ausgeht, wobei man beachten muß, daß diese Punkte auch n-dimensional sein können, ist dieser Ansatz für GIS nicht geeignet. Statt dessen wird bei der q-Zellenzerlegung der Raum nicht in einzelne Punkte aufgeteilt, sondern es werden auch Eigenschaften der mengentheoretischen Topologie eingeführt. Dies führt dazu, daß das so entstandene Gebilde nicht mehr Zellenkomplex, sondern lediglich Quasi-Zellenkomplex oder q-Zellenkomplex genannt werden kann.

Voraussetzung für eine q-Zellenzerlegung ist ein separierter Raum (sogenannter Hausdorffraum) der Form $\mathbb{R}^p \times \mathbb{D}^q$, wobei \mathbb{R}^p einen p-dimensionalen Raum bezeichnet und \mathbb{D}^q das kartesische Produkt der paarweise zueinander orthogonalen Mengen \mathbb{D}^i ist, die die Semantik darstellen (analog zum Relationenmodell). Die q-Zellenzerlegung, deren Elemente q-Zellen genannt werden, findet dabei im Unterraum \mathbb{R}^p statt und muß den folgenden Bedingungen genügen (ein hochgestellter Index steht dabei für die Dimension einer q-Zelle, ein Punkt über der Zelle bezeichnet ihren Rand).

- Q_1 : Eine Quasi- oder q-Zelle z^p hat einen äußeren Rand \dot{z}_a^p und für $p > 1$ n ($n = 1, 2, 3, \dots, m$) innere Ränder \dot{z}_i^p . Es gilt: $\dot{z}_j^p \cap \dot{z}_k^p = \emptyset$ für $j \neq k$, wobei k, j die Menge der Ränder durchlaufen. Die q-1-Zelle ist homöomorph zu $]0, 1[$. Die Menge aller q-Zellen wird als Quasi-Zellenkomplex oder als q-Zellenkomplex Q bezeichnet.
- Q_2 : Für die randbildenden z^{p-1} -Zellen muß gelten, daß jeweils die endliche Vereinigungsmenge eines Teils der z^{p-1} homöomorph zur $(p-1)$ -dimensionalen offenen Einheitskugel ist.
- Q_3 : Es gilt $z_i^p \cap z_j^q = \emptyset$ für $p = q$ und $z_i^p \cap z_j^q = \emptyset$ oder $z_i^p \cap z_j^q = z_i^p \vee z_j^q$ für $p \neq q$.
- Q_4 : Es sind nur endliche q-Zellenkomplexe zugelassen.

Aus obigen Forderungen kann man entnehmen, daß die q-Zellenzerlegung sowohl Eigenschaften der mengentheoretischen als auch solche der algebraischen Topologie in sich vereinigt. Das Hausdorffkriterium muß im Unterraum allerdings nicht erfüllt sein. Dies ergibt sich über den Umkehrschluß aus der Voraussetzung, die diese Bedingung lediglich an den gesamten Raum $\mathbb{R}^p \times \mathbb{D}^q$ stellt. \mathbb{R}^p ist dabei derjenige Raum, in dem ausschließlich die topologischen Beziehungen abgebildet werden. Der Anschauung aus der algebraischen Topologie folgend wird er aus Punkten in allgemeiner Lage aufgespannt.

Im weiteren werden die verschiedenen Forderungen näher erläutert. Am einfachsten ist dabei die Bedingung Q_4 zu interpretieren, die lediglich dazu dient q-Zellenkomplexe mit unendlicher Ausdehnung zu unterbinden. Sie wurde einzig wegen der sinngemäß gleichen Forderung bei den Zellenkomplexen übernommen. In Q_2 wird für jede Zelle eine endliche Hülle gefordert. Dies entspricht ebenfalls den Forderungen der CW-Zellenkomplexe, obwohl sie in wesentlichen Teilen abgeschwächt wurden. Als Konsequenz daraus ergibt sich die Tatsache, daß der Umring eines Gebietes nicht etwa in Form einer Funktion wie z. B. $x^2 + y^2 = r^2$ definiert sein darf, sondern vielmehr muß der Rand eines Gebietes im \mathbb{R}^2 durch eine endliche Menge von Kanten gegeben sein. Darüber hinaus darf ein Gebiet G_k weitere Gebiete (siehe Q_1) enthalten, die nicht zu G_k selbst gehören. Diese Möglichkeit besteht bei den CW-Zellenkomplexen nicht. Abbildung 1 gibt einige Beispiele für das zuvor Ausgeführte. Weitere Details zur q-Zellenzerlegung können [RÖSCH 1998] entnommen werden.

Aus den Bedingungen $Q_1 - Q_4$ ergeben sich für die am Zellenkomplex beteiligten Elemente klar definierte topologische Relationen. Die Anzahl dieser Beziehungen, ebenso wie deren Art können aufgrund des Bildungsgesetzes der q-Zellen a priori für beliebige Räume angegeben werden. Dies bedeutet allerdings nicht, daß sie damit auch zu interpretieren und zu analysieren sind. Derzeit ist dies im Raum \mathbb{R}^p nur für $p \leq 3$ möglich.

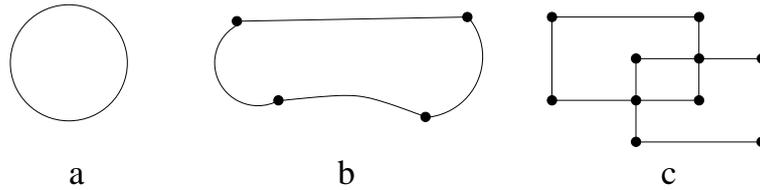


Abbildung 1: Beispiele für eine unzulässige (a) und zwei zulässige (b,c) q-Zellenzerlegungen

Während bisher lediglich topologische Aspekte der Datenmodellierung behandelt wurden, folgen im nächsten Kapitel die semantischen, die erst die Eindeutigkeit der topologischen Beziehungen von fachlichen Objekten sicherstellen.

3 Der Einsatz der q-Zellenzerlegung in GIS

Betrachtet man das Objekt als das Aggregat seiner topologischen, semantischen und geometrischen Eigenschaften (siehe dazu Abb. 2), dann stellt sich die Frage, wie das vorgestellte Modell in GIS implementiert werden kann. Hierzu wurde eine Klassenbibliothek entwickelt, die die strukturellen Beziehungen der am q-Zellenkomplex beteiligten Elemente, abbildet. Für die 1-Zelle und die 2-Zelle im \mathbb{R}^2 sind diese Beziehungen mit Hilfe der Object Modeling Technique in Abbildung 3 dargestellt.

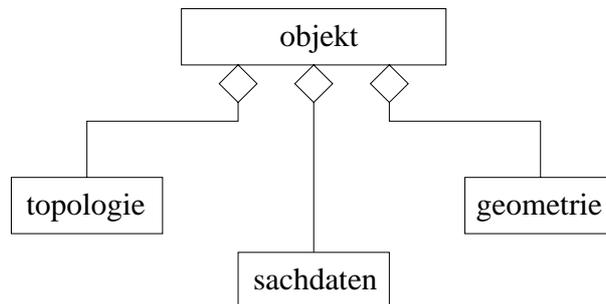


Abbildung 2: Das Objekt als Aggregat seiner verschiedenen Eigenschaften

Wie aus der Übersicht ersichtlich ist, gestaltet sich die Beziehung der einzelnen Zelle zu den übrigen als reine Mengenbeziehung und kann somit auch entsprechend implementiert werden. Moderne Programmiersprachen unterstützen dies durch entsprechende Datentypen. In *C++* sind sie beispielsweise in der *Standard Template Library* realisiert. Damit können diese Mengenbeziehungen sehr einfach in Programmcode umgesetzt werden.

Besonderes Augenmerk muß in diesem Zusammenhang auf die Beziehung q-Identität gelegt werden. Diese Relation bringt die Mehrdeutigkeit der Zugehörigkeit des Elements zu einem bzw. mehreren Objekten zum Ausdruck. Der Grund dafür liegt in der Tatsache, daß die q-Zellenzerlegung in einem Unterraum durchgeführt wird und von daher die Häufigkeit der Zuordnung des Elements zu Objekten nicht bekannt ist. Beispielsweise kann eine q-Zelle einmal die Bedeutung *Grenzpunkt* und gleichzeitig auch *Gebäudeeckpunkt* haben. Dies bereitet allerdings keine größeren Probleme, da über die Semantik jederzeit Eindeutigkeit hergestellt werden kann. Weiterhin kann man davon ausgehen, daß jede topologische Analyse vom fachlichen Objekt ausgeht, wodurch die Semantik bereits vorgegeben ist. So stützt sich beispielsweise die Frage nach Flurstücken innerhalb eines Wasserschutzgebiets ausschließlich auf die Semantik, obwohl die Analyse selbst topologischer Natur ist. Die Beziehung q-Identität stellt somit

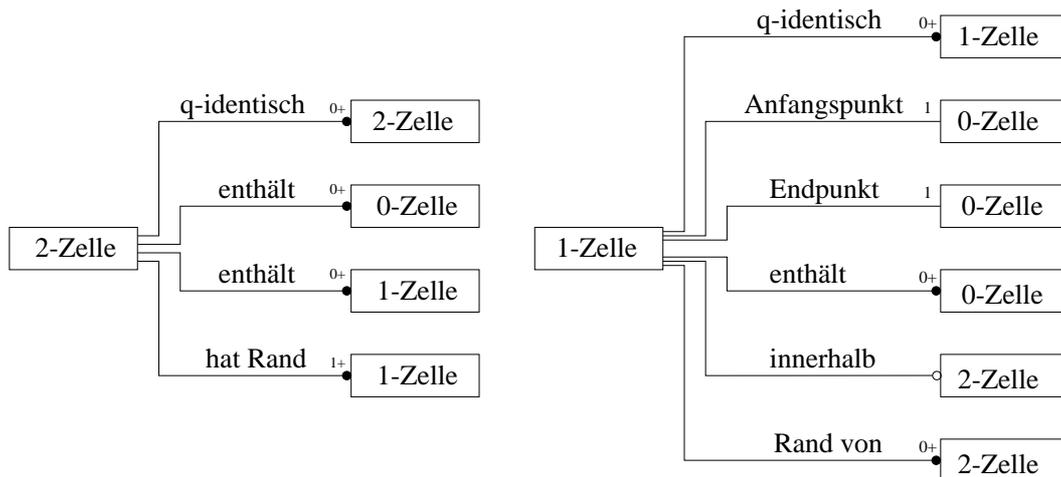


Abbildung 3: Die 2- und die 1-Zelle im \mathbb{R}^2

die Verbindung zwischen Topologie und Semantik her. Genau genommen ist diese Relation somit nicht rein topologischer Art.

Im Rahmen der topologischen Modellbildung erhält das Objekt die Eigenschaften einer bestimmten q -Zelle. Die Dimension ergibt sich dabei aus dem Kontext. So kann beispielsweise eine Stadt in einer bestimmten Anwendung als Verkehrsknoten und somit als 0-Zelle modelliert sein, während die gleiche Stadt bei der Darstellung seiner Stadtteile als 2-Zelle repräsentiert wird. Es ist auch denkbar, daß sich ein Objekt in verschiedenen Kontexten topologisch unterschiedlich verhält, d. h. es vereinigt das topologische Verhalten von unterschiedlichen q -Zellen in sich.

4 Die Vorteile einer objektbezogenen Topologie

Aus der strengen, objektbezogenen Modellierung ergeben sich im wesentlichen drei Vorteile. Es besteht zum einen die Möglichkeit den Datenbestand auf *Vollständigkeit* zu prüfen zum anderen können die Daten auf *Plausibilität* bzw. *Konsistenz* getestet werden und schließlich kann der Datenbestand auch noch topologisch *analysiert* werden.



Abbildung 4: Ein unvollständiger Datenbestand

Einige Beispiele sollen das oben Gesagte kurz erläutern. Zunächst wird auf die Prüfung der Vollständigkeit näher eingegangen. Dies läßt sich am einfachsten an einem Beispiel verdeutlichen. Abbildung 4 zeigt zwei nicht näher spezifizierete Stromkabel, die zunächst nicht miteinander verbunden sind (dies wird durch die Lücke zum Ausdruck gebracht). Als galvanisch verknüpft können sie erst gelten, wenn die beiden Kabelenden beispielsweise durch eine Verbindungsmuffe verbunden sind. Sollte also in einem Stromnetzwerk überprüft werden, ob alle Kabel miteinander verknüpft sind, dann kann dies

neben anderen Kriterien beispielsweise auch bedeuten, daß je zwei Kabelenden mit einer Muffe versehen sein müssen. Dies wäre in Abb. 4 nicht gegeben. Ganz abgesehen davon müßten hinsichtlich der Spannungsebene usw. noch weitere Kriterien erfüllt sein, die allerdings mehr im semantischen Bereich angesiedelt sind und somit in diesem Zusammenhang keine Rolle spielen.

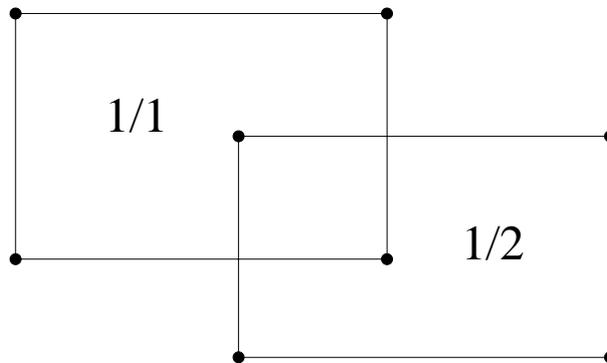


Abbildung 5: Zwei sich überlappende Flurstücke

Als nächstes gilt es die Prüfung auf Konsistenz bzw. Plausibilität näher zu beleuchten. In diesem Zusammenhang geht es vor allem darum zu untersuchen, ob der Datenbestand, so wie man ihn vorfindet überhaupt denkbar ist, d. h. bestimmten vordefinierten Kriterien nicht widerspricht. Abbildung 5 gibt dazu ein anschauliches Beispiel. Flurstücke bilden per Definition eine sogenannte ebene Pflasterung und können sich folglich aufgrund dessen nicht überlappen. Die dargestellte Situation entspricht somit nicht den Voraussetzungen und müßte von daher bereinigt werden. Der Widerspruch liegt dabei nicht etwa in der Topologie an sich, sondern er entsteht erst durch die Hinzuziehung der Semantik.

Zuletzt bleibt noch die Analyse zu erläutern. Sie kann als die Ergänzung der beiden zuvor genannten Vorteile betrachtet werden. Nachdem der Datenbestand vollständig ist (soweit dies mit den aufgeführten Mitteln festzustellen ist) und auch auf Plausibilität getestet wurde, kann er analysiert werden. Abbildung 6 gibt dazu ein Beispiel aus dem Bereich der Stromversorgung, wo es oftmals um Spannungsverfolgung geht. Auf diese Art kann festgestellt werden, welche Betriebsmittel galvanisch miteinander verbunden sind.



Abbildung 6: Drei Kabel und zwei Schalter als Teil eines Stromnetzes

5 Anwendungsbeispiel

In diesem Abschnitt soll das Modell an einem kleinen Beispiel näher erläutert werden. Ausgangspunkt sei dabei die q-Zellenzerlegung, wie sie in Abbildung 7 dargestellt ist. Man beachte, daß damit noch keine Semantik vorgegeben ist.

Im nächsten Schritt wird zur reinen q-Zellenzerlegung noch die Semantik hinzugezogen. Dabei sollen die 0-Zellen S_i Städte darstellen, bei den 1-Zellen G_j handle es sich um Landesgrenzen, die mit F_k bezeichneten Zellen stellen Fahrbahnen dar und die 2-Zelle L sei damit ein von G_j berandetes Land. Die übrigen 0-Zellen P_l sind damit folgerichtig als Grenzpunkte zu interpretieren. Auf die Verfeinerung der Semantik kann im weiteren verzichtet werden, da sie für das Beispiel ohne Konsequenz bleibt. Ebenso

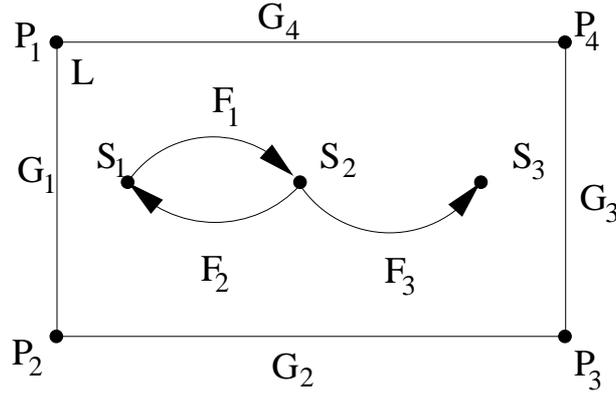


Abbildung 7: Beispiel für eine q-Zellenzerlegung

wird auf die exakte Darstellung der Geometrie verzichtet, da diese gleichfalls von untergeordneter Bedeutung ist. Denn es ist für das topologische Modell völlig uninteressant, wie der tatsächliche geometrische Verlauf der Straßen zwischen den Städten ist.

Nimmt man das in Abschnitt 3 vorgestellte Modell zur Hand, dann können die topologischen Relationen wie folgt abgebildet werden. Zunächst die Städte S_i :

$$S_1 = \begin{cases} \text{q-identisch} & \{\emptyset\} \\ \text{anfang von} & \{F_1\} \\ \text{ende von} & \{F_2\} \\ \text{innerhalb} & \{L\} \\ \text{rand von} & \{\emptyset\} \end{cases} \quad S_2 = \begin{cases} \text{q-identisch} & \{\emptyset\} \\ \text{anfang von} & \{F_2, F_3\} \\ \text{ende von} & \{F_1\} \\ \text{innerhalb} & \{L\} \\ \text{rand von} & \{\emptyset\} \end{cases} \quad S_3 = \begin{cases} \text{q-identisch} & \{\emptyset\} \\ \text{anfang von} & \{\emptyset\} \\ \text{ende von} & \{F_3\} \\ \text{innerhalb} & \{L\} \\ \text{rand von} & \{\emptyset\} \end{cases}$$

Es folgen die Fahrbahnen (Anzahl und Art der Beziehungen wie in Abb. 3):

$$F_1 = \begin{cases} \text{q-identisch} & \{\emptyset\} \\ \text{beginnt} & \{S_1\} \\ \text{endet} & \{S_2\} \\ \text{enthält} & \{\emptyset\} \\ \text{innerhalb} & \{L\} \\ \text{rand von} & \{\emptyset\} \end{cases} \quad F_2 = \begin{cases} \text{q-identisch} & \{\emptyset\} \\ \text{beginnt} & \{S_2\} \\ \text{endet} & \{S_1\} \\ \text{enthält} & \{\emptyset\} \\ \text{innerhalb} & \{L\} \\ \text{rand von} & \{\emptyset\} \end{cases} \quad F_3 = \begin{cases} \text{q-identisch} & \{\emptyset\} \\ \text{beginnt} & \{S_2\} \\ \text{endet} & \{S_3\} \\ \text{enthält} & \{\emptyset\} \\ \text{innerhalb} & \{L\} \\ \text{rand von} & \{\emptyset\} \end{cases}$$

Zuletzt noch die 2-Zelle, die als Land interpretiert wird (Anzahl und Art der Beziehungen ebenfalls wie in Abb.3):

$$L = \begin{cases} \text{q-identisch} & \{\emptyset\} \\ \text{enthält} & \{S_1, S_2, S_3\} \\ \text{enthält} & \{F_1, F_2, F_3\} \\ \text{hat rand} & \{G_1, G_2, G_3, G_4\} \end{cases}$$

	S_1	S_2	S_3
F_1	1	-1	0
F_2	-1	1	0
F_3	0	1	-1

Tabelle 1: Die Inzidenzmatrix der F_i und S_j zu Abbildung 7

Die Aufstellung macht deutlich, daß beispielsweise die Herleitung der Inzidenzmatrix für die Städte S_j und die Fahrbahnen F_i ohne Probleme möglich ist, da sie sofort aus der Datenstruktur abgelesen werden kann. Damit können auch alle gängigen Algorithmen, wie z. B. kürzeste Wege zwischen zwei Städten, angewendet werden. Sollten dabei die Kanten auch gewichtet werden, so ist dies nach der Logik dieses Modells eine semantische Information, die auch auf dieser Ebene geführt werden muß. Der Vollständigkeit halber sollte noch ergänzt werden, daß auf die Darstellung der topologischen Beziehungen der übrigen Objekte (z. B. der Grenzen G_i) verzichtet wird, da sie in diesem Zusammenhang keine weiteren Einsichten bringen.

Darüber hinaus können aber auch weitere Analysen durchgeführt werden. So kann beispielsweise die Frage, welche Städte liegen in L ebenso direkt aus der Struktur abgelesen werden wie die dazu inverse Beziehung 'in welchem Land liegt die Stadt S_j '. Es kann natürlich prinzipiell auch geklärt werden, welche weitere Länder an L grenzen, obwohl diese Analyse in diesem einfachen Beispiel nicht vorgesehen ist.

6 Abschließende Bemerkungen

Dieser Beitrag befaßt sich mit einem Modell zur Abbildung topologischer Beziehungen in GIS, das auf dem Ansatz der q -Zellenzerlegung beruht. Die Anwendung in der Praxis wurde am Beispiel eines Verkehrsnetzes aufgezeigt. Dabei wurde nicht auf die Erzeugung des q -Zellenkomplexes eingegangen, die sich algorithmisch durchaus anspruchsvoll gestalten kann, sondern es wurden lediglich die strukturellen Beziehungen der beteiligten q -Zellen untereinander berücksichtigt.

Eine erste prototypische Implementierungen hat die Leistungsfähigkeit dieses Modells bereits unter Beweis gestellt. Ein Programm für den graphentheoretischen Ansatz konnte erfolgreich getestet werden. Die Erweiterung auf den \mathbb{R}^2 wurde ebenfalls realisiert, hier wurde allerdings bisher lediglich die ebene Pflasterung, wie sie beispielsweise bei Flurstücken gegeben ist, berücksichtigt. Demgegenüber sind die Analysemöglichkeiten im \mathbb{R}^2 wesentlich weitreichender umgesetzt.

Die ersten Erfahrungen, die mit dem Prototypen bisher gewonnen werden konnten waren sehr zufriedenstellend. Die Vermutung, daß die Hauptschwierigkeit in der Zerlegung selbst liegt hat sich im wesentlichen bestätigt. Hier gilt es noch leistungsfähige Algorithmen zu entwickeln, damit die Zerlegung selbst möglichst schnell durchgeführt werden kann. Liegt ein Datenbestand bereits als q -Zellenzerlegung vor, dann ist der Aufwand bei der Fortführung ungleich geringer.

Literatur

Whitehead 1949 Whitehead J. H.: Combinatorial homotopy. I. Bull. AMS.

Rösch 1998 Rösch N.: Topologische Beziehungen in Geo-Informationssystemen. Dissertation. Im Internet veröffentlicht. <http://www.ubka.uni-karlsruhe.de/cgi-bin/psview?document=1998/bauverm/2>