

Methodische Aspekte der Generalisierung von Geodaten

Inaugural-Dissertation
zur
Erlangung des Grades

Doktor-Ingenieur
(Dr.-Ing.)

der
Hohen Landwirtschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität
zu Bonn

vorgelegt am
28. November 2005
von
Matthias Bobzien
aus
Bonn

Angefertigt mit Genehmigung der Landwirtschaftlichen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

Referent: Professor Dr.-Ing. Dieter Morgenstern
Korreferent: Professor Dr. Lutz Plümer
Tag der mündlichen Prüfung: 10. Februar 2006

Diese Dissertation ist auf dem Hochschulschriftenserver der ULB Bonn
http://hss.ulb.uni-bonn.de/diss_online elektronisch publiziert.

Erscheinungsjahr: 2006

Zusammenfassung

Methodische Aspekte der Generalisierung von Geodaten

Diese Arbeit stellt die Generalisierung in der Kartographie auf eine theoretische Grundlage im Zusammenhang mit mathematischen Strukturen. Ausgegangen wird von einem erfolgreich entwickelten Verfahren zur Modellgeneralisierung von Digitalen Landschaftsmodellen. Die Verallgemeinerung dieses Verfahrens ist wünschenswert, um auch andere Anwendungen von den Ergebnissen profitieren zu lassen. Gezeigt wird, dass in bestimmten Bereichen eine Verallgemeinerung des gesamten Verfahrens prinzipiell nicht zu erreichen ist. Um der Verallgemeinerung der Modellgeneralisierung dennoch näher zu kommen, wird das Problem der Generalisierung auf einer abstrakteren Ebene definiert.

Das erste Ziel ist die Formalisierung der Generalisierungsziele. Hierbei wird auch die Formalisierung der Geodaten und des Generalisierungsvorganges angestrebt. Das zweite Ziel ist die Abstraktion und Verallgemeinerung der entwickelten Formalisierungen, um Gemeinsamkeiten zwischen verschiedenen Generalisierungsanwendungen zu finden. Das dritte Ziel ist die Erstellung von Bedingungen, die sich auf einzelne Komponenten der Geodaten beziehen. Die Erfüllung dieser Bedingungen ist Voraussetzung für eine Generalisierung mit korrekten und qualitativ hochwertigen Ergebnissen.

Die Methodik stützt sich auf folgende Mittel: (1) Es wird ein Datenmodell entwickelt, das auf eine große Bandbreite unterschiedlicher Arten von Geodaten anwendbar ist. Zum einen ermöglicht es durch seine klare Strukturierung in Thematik, Topologie und Geometrie eine differenzierte Gruppierung der Formalisierungen. Zum anderen ermöglicht es eine Multirepräsentation von Geodaten in unterschiedlichen Auflösungen. (2) Die Generalisierung wird durch eine mathematische Funktion ausgedrückt, die aus der ungeneralisierten Menge von Geodaten auf die Menge der generalisierten Geodaten abbildet. Die Formulierung der Generalisierung als mathematische Funktion bildet die Grundlage für die Abstraktion der Generalisierung. (3) Die Generalisierungsfunktion wird als Morphismus formuliert. Dazu werden Eigenschaften auf den ungeneralisierten Daten identifiziert, die durch eine Generalisierung nicht verändert werden dürfen oder sollen. Diese Eigenschaften bilden Invarianzen der Generalisierung. Beispiele für invariante Eigenschaften sind die Zugehörigkeit zu Objektklassen, die topologische Adjazenz und Inzidenz, der Netzzusammenhang sowie die Flächendeckung. Die gefundenen invarianten Eigenschaften werden als Bedingungen formuliert, und zwar so, dass sie durch einen Rechner automatisch überprüfbar sind. Sie sind unabhängig von einer konkreten Generalisierungsanwendung, da sie abstrahiert und allgemein formuliert sind. Mit der abstrakten und formalisierten Erstellung von Bedingungen zu invarianten Eigenschaften der Generalisierung leistet diese Arbeit einen Beitrag zur bedingungs-basierten Modellierung der Generalisierung.

Mit der Anwendung auf reale Daten wird die Praktikabilität der entwickelten Formalisierungen, Abstraktionen und Verallgemeinerungen gezeigt. Die Anwendung geschieht auf Basis der Implementation des eingangs erwähnten Verfahrens zur Modellgeneralisierung. An Hand von zwei weiteren Anwendungsbeispielen – der Generalisierung von geologischen Karten und der Generalisierung von Straßenkarten – wird gezeigt, wie die abstrahiert formulierten Bedingungen der Invarianzen auf andere Generalisierungsanwendungen übertragen werden können.

Abstract

Methodological Aspects of the Generalisation of Spatial Data

This thesis establishes a theoretical framework for the generalisation in cartography in the context of mathematical structures. The starting point is a procedure of model generalisation developed for digital landscape models that has been successfully implemented. A more general procedure that lets other generalisation applications benefit from the results would be desirable. It is shown that for some aspects it's impossible to generate a most general level of this procedure. To bypass this problem, an attempt is made to approximate a general level by defining the task of model generalisation in a more abstract way.

The first aim of this thesis is the formalisation of the goals of generalisation. Subordinate objectives are the formalisation of the spatial data and the formalisation of the generalisation process. The second aim is the abstraction of the formalisations developed and the identification of analogies that hold between different generalisation applications. The third aim is to establish constraints. These constraints concern different components of spatial data. The satisfaction of these constraints is a prerequisite for a generalisation that leads to correct and high quality results.

The following methods are used to reach these aims: (1) A data model is developed which is applicable to a large variety of different types of spatial data. It makes possible a differentiation of the formalisations by a clear classification of the spatial data into thematic, topological and geometrical elements. It also supports multiple representations of spatial data in different resolutions. (2) The generalisation is expressed by a mathematical function that maps an ungeneralised set of spatial data onto a generalised set of spatial data. The formulation as a mathematical function provides the basis for the abstraction of the generalisation. (3) The generalisation function is formulated as a morphism. For this purpose invariant properties of the sets of spatial data are identified. They must—or at least should—be preserved within the generalisation process. These properties function as the invariant parts of the generalisation. Examples of such invariants are: Membership of object classes, topological adjacency and incidence, net connectivity and area coverage. The identified invariants are formulated as constraints in such a way that they can be verified automatically by a computer. They are independent of the specification of a generalisation application because they are abstracted and formalised in a general way. With the abstracted and formalised constraints of invariant properties this thesis makes a contribution to the constrained-based modelling of generalisation.

The practicability of the developed formalisations and abstractions is shown by applying them to real data. Here the above mentioned procedure of model generalisation of digital landscape models is used. Two further examples—the generalisation of geological maps and the generalisation of road maps—show how the abstractly formulated constraints can be adapted to other generalisation applications.

Computers are good at following instructions, but not at reading your mind.

Donald E. Knuth
The T_EX-Book

Inhaltsverzeichnis

1	Einleitung	11
2	Generalisierung und ihre Automation	17
2.1	Generalisierung in der Kartographie	18
2.1.1	Definition der Generalisierung	18
2.1.2	Notwendigkeit der Generalisierung	20
2.1.3	Bestandteile der Generalisierung	21
2.1.4	Elementare Vorgänge der Generalisierung	23
2.2	Kartographische Modellbildung	25
2.2.1	Digitale Modelle	26
2.2.2	Arten der Generalisierung	27
2.2.3	Methoden der Datenhaltung für die Generalisierung	28
2.3	Automation der Generalisierung	31
2.3.1	Überblick über die Entwicklung der Automation der Generalisierung	31
2.3.2	Automatische Modellgeneralisierung	32
2.3.3	Klassifizierung von Generalisierungsmodellen	34
2.4	Methodik dieser Arbeit	36
3	Raumbezogene Datenmodelle	39
3.1	Modellierung raumbezogener Daten eines Strukturierungsgrades	40
3.1.1	Datenbankbezug räumlicher Datenmodelle	40
3.1.2	Grundtypen raumbezogener Datenmodelle	41
3.1.3	Nicht-räumliche Informationen: Thematik	47
3.1.4	Komplexe Objekte	49
3.1.5	Modellierungsbeispiele	50
3.1.6	Raumbezogene Datenmodelle in der amtlichen topographischen Kartographie	53
3.2	Modellierung raumbezogener Daten in mehreren Strukturierungsgraden .	56
3.2.1	Überblick	57
3.2.2	Erweitertes hierarchisches Datenmodell für mehrere Strukturierungsgrade	60
3.3	Fazit	64

4	Generalisierung als Funktion	65
4.1	Motivation	66
4.2	Generalisierungsrelation	68
4.2.1	Geodaten, Mengen von Geodaten und Zuordnungen zwischen Mengen von Geodaten	68
4.2.2	Mathematische Formulierung der Generalisierungsrelation	71
4.2.3	Generalisierung versus relationale Zuordnung (Matching)	72
4.2.4	Multiplizitäten der Generalisierungsrelation	73
4.3	Generalisierungsfunktion	77
4.3.1	Voraussetzungen	77
4.3.2	Spezielle Form der Generalisierungsfunktion	78
4.3.3	Allgemeine Formen der Generalisierungsfunktion	79
4.3.4	Beispiel: Zahlenmäßige Verringerung dreier Häuser	81
4.3.5	Vergleich mit der Arbeit von Ai und van Oosterom	85
4.3.6	Vergleich von Relation und Funktion	86
4.4	Mathematische Charakterisierung der Funktion	87
4.4.1	Nicht-Totalität	87
4.4.2	Nicht-Injektivität	88
4.4.3	Surjektivität	89
4.5	Nichtdeterminiertheit der Generalisierung	91
4.6	Fazit	94
5	Invariante Eigenschaften der Generalisierungsfunktion	97
5.1	Mathematische Grundlagen: Strukturen und Morphismen	99
5.1.1	Erweiterte Generalisierungsfunktion	100
5.1.2	Alternative zur erweiterten Generalisierungsfunktion	102
5.1.3	Morphismen	102
5.1.4	Mathematische Formulierung der Strukturhaltung	103
5.2	Invariante thematische Eigenschaften	105
5.2.1	Objektarten	108
5.2.2	Attribute	113
5.2.3	Auswirkungen von Aggregationen	117
5.3	Invariante topologische Eigenschaften	121
5.3.1	Inzidenz und Adjazenz	123
5.3.2	Inklusion	132
5.4	Invariante geometrische Eigenschaften	133
5.5	Invariante kombinierte Eigenschaften	141
5.5.1	Erreichbarkeit und Netzzusammenhang	141
5.5.2	Flächendeckung	144
5.5.3	Topologische Konsistenz und geometrische Generalisierung	145
5.6	Fazit	149

6 Anwendungen	151
6.1 Modellgeneralisierung	151
6.1.1 Invariante thematische Eigenschaften	154
6.1.2 Invariante topologische Eigenschaften	156
6.1.3 Invariante geometrische Eigenschaften	156
6.1.4 Invariante kombinierte Eigenschaften	157
6.2 Generalisierung geologischer Karten	159
6.3 Generalisierung von Straßenkarten	165
6.4 Fazit	169
7 Zusammenfassung und Ausblick	171
Literaturverzeichnis	183

Inhaltsverzeichnis

Kapitel 1

Einleitung

In den letzten Jahren ist ein enormer Aufwand in die Erfassung von Geodaten zur elektronischen Verarbeitung geflossen. In Deutschland zum Beispiel hat der Zusammenschluss der Landesvermessungsämter das Amtliche Topographisch-Kartographische Informationssystem (ATKIS) geschaffen und dieses mit Inhalten gefüllt (Jäger 2003). Die erfassten Daten bilden die *Geodatenbasis* in einem großmaßstäbigen Bereich¹. Diese Vorgehensweise ist ein international verwendeter Ansatz, der unter anderem in Frankreich (BD Topo, IGN 2004) und in der Schweiz (TLM, SwissTopo 2004) Anwendung findet.

Da die Erfassung der Daten mit hohen Kosten verbunden ist, ist ihre mehrfache Nutzung ein erstrebenswertes Ziel. Dazu müssen die Daten für unterschiedliche Zwecke aufgearbeitet und angepasst werden. Durch entsprechende Aufarbeitung kann ein breites Spektrum der Nutzbarkeit erzielt werden. Die Möglichkeit der Wiederverwendbarkeit bildet eines der größten Potenziale dieser Geodatenbasen und macht damit einen Großteil ihres Kapitals aus. Bereits bei der Modellierung der Daten wurde deshalb angestrebt, sie für unterschiedliche Zwecke und Nutzungen zu erfassen (Harbeck 1987). Diese Überlegungen gelten im übertragenen Sinne für fast alle Sammlungen von Geodaten.

Im Allgemeinen liegen die erfassten Daten in einem bestimmten Strukturierungsgrad vor, der einheitlich vorgegeben ist. Dieser ist für manche Anwendungen jedoch ungeeignet, weil er zu fein ist und zu viele Details enthält. Für viele Analysen ist die Datenmenge zu umfangreich. Ebenso kann eine adäquate Visualisierung ab einem bestimmten Maßstab mit einer festgelegten Signaturierung aus Platzgründen nicht mehr durchgeführt werden. Aus diesen Gründen ist eine Anpassung des Strukturierungsgrades für die gewählten Anwendungen nötig. Dieser Vorgang wird gemeinhin mit *Generalisierung* bezeichnet (siehe Kapitel 2).

Die Generalisierung ist ein komplexer und aufwändiger Prozess, der auch heute in vielen Bereichen noch nahezu vollständig manuell ausgeführt wird (SGK 2002, Schuurman 1999). Automatisierung kann diesen Prozess entscheidend beschleunigen und zur Kosteneffizienz beitragen. Es ist jedoch keinesfalls selbstverständlich, dass die Generalisierung automatisierbar ist. Die Grundvoraussetzungen für eine Automatisierbarkeit sind dennoch gegeben: Zum einen ist das gewünschte Ziel der Generalisierung, bestimmt

¹In Bayern im Bereich mittleren Maßstabes.

durch Zweck und Anwendung, bekannt. Zum anderen enthalten die Ausgangsdaten im Allgemeinen bereits alle für das Generalisierungsziel benötigten Informationen.

Die Automatisierung der Generalisierung wird seit vielen Jahren sowohl von Datenanbietern als auch von Datennutzern gefordert und gefördert, sowie wissenschaftlich erforscht (SGK 2002). Die Ergebnisse sind jedoch noch weit von einer vollständigen Automatisierung entfernt (Schuurman 1999). Die Relevanz der Generalisierung, sowie der weitere Forschungsbedarf ist allgemein anerkannt, wenn auch des Öfteren die Meinung vertreten wird, dieser Prozess lasse sich niemals vollständig automatisieren (Harbeck 1987, SGK 2002).

Die bisher erfolgten Ansätze zur Automatisierung der Generalisierung lassen sich grob in die folgenden drei Bereiche einteilen:

1. **Automatisierung von Teilen des Generalisierungsprozesses:** Einzelne *Teilaufgaben* der Generalisierung werden automatisiert, zum Beispiel die Ausdünnung der Geometrie, die Glättung von Linien, das Zusammenfassen von Objekten, deren Auswahl oder Verdrängung (siehe Abschnitt 2.1.4). Einige dieser Teilaufgaben sind in hohem Maße erforscht und haben beachtliche Erfolge in der Durchführung aufzuweisen, während andere noch vor großen Problemen stehen (siehe Abschnitt 2.3). Einige wenige Teillösungen wurden in kommerziellen GIS-Produkten implementiert und sind damit für die Allgemeinheit verfügbar (Lee 2003, Smith 2003).
2. **Konzeptionelle Modelle zur Generalisierung:** In diesen eher theoretisch orientierten Ansätzen wird der *Gesamtprozess* der Generalisierung modelliert. Lösungen zur Automation der Generalisierung werden durch ein Gesamtkonzept gesucht und entwickelt. McMaster und Shea (1992) stellen vier solcher Modelle unterschiedlicher Autoren vor und ergänzen dieses durch ein eigenes, fünftes Modell.
3. **Spezialisierte Gesamtlösungen:** In spezialisierten Gesamtlösungen wird die Generalisierung nicht wie in Punkt 1 auf Teilaufgaben beschränkt betrachtet, sondern in ihrer Gesamtheit implementiert. Diese Lösungen sind jedoch immer für spezielle, eingeschränkte Anwendungen entwickelt worden, und sind nicht ohne weiteres auf andere übertragbar. Beispiele für solche Gesamtlösungen sind die Modellgeneralisierung (Morgenstern und Schürer 1999, Schürer 2002) und das Projekt GiMoDig (Sester u. a. 2004).

Die Automatisierung der Generalisierung wurde in ihrer Gesamtheit noch nicht erschöpfend gelöst. Schwierigkeiten ergeben sich sowohl beim Zusammenfügen der Teillösungen, wie auch bei der Realisierung der konzeptionellen Modelle. Die spezialisierten Gesamtlösungen wiederum lassen sich nur mit größeren Anstrengungen und gravierenden Änderungen auf andere Anwendungen übertragen.

Im Wesentlichen lassen sich die Schwierigkeiten, eine allgemeingültige Gesamtlösung für die Automation der Generalisierung zu finden, auf drei Ursachen zurückführen:

- Es gibt keine natürliche Ordnung auf Geodaten.
- Es gibt keine natürliche Reihenfolge für die Ausführung der einzelnen Generalisierungsvorgänge.
- Das Generalisierungsziel ist meist gar nicht oder nur wenig formal ausgedrückt.

Diese Ursachen werden in den folgenden drei Abschnitten näher erläutert. Sie verdeutlichen die Komplexität des Generalisierungsprozesses, erklären den wissenschaftlichen Kontext dieser Arbeit und führen auf die spezifischen Ziele dieser Arbeit hin.

Die in dieser Arbeit betrachteten Geodaten sind im zweidimensionalen Raum der euklidischen Ebene modelliert.² Dort gibt es keine *natürliche Ordnung* im mathematischen Sinne. Als Hilfsmittel existieren Modelle zur Strukturierung im zweidimensionalen Raum. Diese stellen entweder eine rein *geometrische* Ordnung oder eine rein *thematische* Ordnung dar. Eine rein geometrische Ordnung, wie zum Beispiel *Quadtrees* (Worboys 2004), vernachlässigt die Gegebenheiten der Natur, zum Beispiel die der Topographie. Eine rein thematische Ordnung – zum Beispiel nach *Objektklassen* – wiederum vernachlässigt geometrische Zusammenhänge. Für die Kombination von Thematik und Geometrie kann allenfalls eine partielle Ordnung hergestellt werden.

Die fehlende totale Ordnung auf den Geodaten ist ein Grund dafür, dass die einzelnen Vorgänge der Generalisierung, wie zum Beispiel Vergrößern, Verdrängen und Fortlassen, nicht in eine sequentielle Reihenfolge gebracht werden können. Es ist aus der manuellen Generalisierung bekannt, dass Abhängigkeiten zwischen den Vorgängen bestehen (siehe Abschnitt 2.1.4). Es ist aber ebenso bekannt, dass diese Abhängigkeiten nicht linear sind, sondern Zyklen aufweisen können. Ein Kartograph ist in der Lage, eine Karte in ihrer Gesamtheit zusammen mit ihren räumlichen und semantischen Zusammenhängen zu erfassen (Harrie und Weibel 2006). Damit kann der Mensch von der Ausführung einzelner Generalisierungsschritte abstrahieren und die wechselseitigen Abhängigkeiten unterschiedlicher Vorgänge gleichzeitig beurteilen. Für eine automatische Durchführung der Generalisierung bedeutet dies, dass ein rein sequentieller Programmablauf für die Automatisierung der Generalisierung nicht ausreicht. Es gibt in der Forschung bereits Ansätze, die entsprechende Programmier- und Optimierungsmethoden nutzen, um mehrere oder gar alle Vorgänge parallel zu berücksichtigen (zum Beispiel Barrault u. a. 2001).

Das Generalisierungsziel – der angestrebte Zweck oder die angestrebte Anwendung der Karte oder der Geodaten – ist meist nur umgangssprachlich und damit oft unpräzise und nicht eindeutig ausgedrückt. In erster Linie sind diese Beschreibungen für Menschen erstellt. Dementsprechend schwierig ist die Umsetzung in eine für Rechner verarbeitbare

²Generalisierung im dreidimensionalen Raum ist ebenfalls ein Thema der Forschung, zum Beispiel im Kontext von Stadtmodellen. Die dortigen Fragestellungen weichen jedoch in erheblichem Maße von denen im zweidimensionalen Raum ab.

Kapitel 1 Einleitung

Form. Teilweise ist die Formalisierung angestrebt worden durch konkrete quantitative oder qualitative Kriterien für die Erfassung der Daten; ein Beispiel hierfür bilden die Objektartenkataloge von ATKIS. Jedoch finden sich auch dort im Detail immer wieder Formulierungen, die das menschliche Denken ansprechen und schwer formalisierbar sind. Die Anleitungen zur manuellen Generalisierung sprechen beim Kartographen oftmals die Intuition an. Hier spielt auch die Ästhetik der Darstellung eine wichtige Rolle. Sowohl Intuition als auch Ästhetik sind naturgemäß der automatischen Verarbeitung durch Rechner entgegen gesetzt.

Die ersten beiden Ursachen für die Komplexität des Generalisierungsprozesses – die fehlende Ordnung auf den Geodaten und die fehlende Reihenfolge der Generalisierungsvorgänge – sind kontinuierlich Thema in der Forschung (siehe Kapitel 2). Die dritte Ursache, die fehlende Formalisierung des Generalisierungszieles, ist hingegen in der Forschung bislang kaum untersucht worden. Hier setzt diese Arbeit an und versucht, einen Beitrag für die Formalisierung und damit für die Automatisierung der Generalisierung zu leisten.

Das erste Ziel dieser Arbeit ist demnach die *Formalisierung der Generalisierungsziele*. Um dies zu erreichen, müssen jedoch erst bestimmte Vorbedingungen erfüllt sein. Zum einen muss auch die Darstellung der Geodaten formellen Ansprüchen genügen, weil sich die Generalisierungsziele auf die Geodaten beziehen. Zum anderen muss auch der Generalisierungsvorgang formalisierbar sein, weil die Generalisierungsziele die Eigenschaften der ungeneralisierten Daten und die Zusammenhänge zwischen ungeneralisierten und generalisierten Daten berücksichtigen müssen. Die Arbeit befasst sich deshalb auch mit der Formalisierung der Geodaten und des Generalisierungsvorganges. Die *Formalisierung der Geodaten* geschieht durch eine objektorientierte Modellierung, in der die Geodaten in ihre thematischen und raumbezogenen Bestandteile aufgeteilt werden. Diese Aufteilung begünstigt die Differenzierung unterschiedlicher Generalisierungsziele. Die *Formalisierung des Generalisierungsvorganges* geschieht in Form einer mathematischen Funktion. Die Formulierung durch eine mathematische Funktion bringt die Generalisierung in eine für die Ausführung durch einen Rechner geeignete Form und erleichtert damit die Automatisierung.

Das zweite Ziel ist die *Abstraktion und Verallgemeinerung der oben beschriebenen Formalisierungen*. Das Ziel ergibt sich aus dem Kontext der Projektgeschichte. Diese Arbeit hat sich aus Projekten des Instituts für Kartographie und Geoinformation der Universität Bonn entwickelt, die vom Amt für Geoinformationswesen der Bundeswehr (vormals Amt für militärisches Geowesen der Bundeswehr) finanziert und gefördert wurden. Die Projekte hatten die Entwicklung und Implementation einer weitestgehend automatisierten Modellgeneralisierung von topographischen Daten aus dem Topographischen Informationssystem der Bundeswehr (TOPIS) zur Aufgabe. Aus den Projekten entstand eine Dissertation (Schürer 2002), in der für dieses spezielle Problem der Modellgeneralisierung ein Verfahren zu seiner Lösung entwickelt wurde. Bobzien (2000, 2001) beschreibt Teile der Implementationsarbeit. Angespornt durch die durchweg positiven Ergebnisse der Projekte entstand die Fragestellung nach einer Verallgemeinerung des entwickelten

Verfahrens. Die Verallgemeinerung sollte ermöglichen, auch Anforderungen anderer Generalisierungsanwendungen zu erfüllen. Diese Fragestellung hat die vorliegende Arbeit motiviert. Im Vorfeld dieser Arbeit wurden die Formalisierungen zunächst für die Modellgeneralisierung erstellt und damit ein erster Schritt zur Abstraktion und Verallgemeinerung vollzogen (Bobzien 2003, Bobzien und Morgenstern 2003). Die Arbeitshypothese lautet nun, dass allgemeine Probleme der Generalisierung existieren, die unabhängig sind von der Art der Generalisierungsanwendung und von der Art der Geodaten. Die Abstraktion und Verallgemeinerung der Formalisierungen soll ermöglichen, diese auch *für andere Generalisierungsanwendungen* verwendbar zu machen, zusätzlich zur Anwendung der Modellgeneralisierung. Die Hypothese wird verifiziert durch zwei Beispiele, die die Verallgemeinerbarkeit der Probleme demonstrieren.

Das dritte Ziel dieser Arbeit ist die *Erstellung von Bedingungen*, die sich auf einzelne Komponenten der Geodaten³ beziehen. Eine Generalisierung muss diese Bedingungen erfüllen, damit sie korrekte und qualitativ hochwertige Ergebnisse liefert. Beispiele für übertragbare Bedingungen sind die Erhaltung der topologischen Beziehungen, des Netzzusammenhanges und der Flächendeckung (siehe Kapitel 5). Je nach Generalisierungsanwendung können dann diejenigen Bedingungen ausgewählt werden, die für die Realisierung einer speziellen Anwendung notwendig sind und die für die Erstellung eines Verfahrens als Grundlage dienen können. Diese Arbeit beschränkt sich auf die Erstellung der Bausteine; es ist *kein Ziel* der Arbeit, ein allgemeines Verfahren zur Generalisierung anzugeben.

Darüber hinaus strebt diese Arbeit an, mit der Formulierung der Anforderungen und Bedingungen eine Grundlage für zukünftige Anwendungen zu schaffen. Für diese Anwendungen könnte dann im Vorhinein – vor der Generalisierung – entschieden werden, welche Daten gegeben und wie die Anforderungen an die generalisierten Daten formuliert sein müssen, um eine Generalisierung korrekt und qualitativ hochwertig auszuführen.

Die entwickelte Methodik dieser Arbeit verfolgt die drei genannten Ziele parallel. Deshalb ist die Arbeit nicht getrennt nach den drei Zielen aufgebaut. Im Einzelnen gliedert sich die Arbeit wie folgt: In Kapitel 2 werden die kartographischen Grundlagen gegeben, die für diese Arbeit von Bedeutung sind. In Kapitel 3 wird ein Datenmodell entwickelt, das für unterschiedliche Generalisierungsanwendungen nutzbar ist. Kapitel 4 beschreibt die Generalisierung durch eine mathematische Funktion. Kapitel 5 befasst sich mit denjenigen Strukturen der Mengen von Geodaten, die bei der Anwendung der Generalisierungsfunktion invariant sind. In Kapitel 6 werden Beispiele für die in dieser Arbeit entwickelten Konzepte gegeben. Die Arbeit endet mit einer Zusammenfassung und einem Ausblick in Kapitel 7.

³Komponenten sind hier die Thematik, die Topologie und die Geometrie, siehe Kapitel 3.

Kapitel 2

Generalisierung und ihre Automation

Dieses Kapitel beschreibt die Generalisierung in der Kartographie und die Aufgaben, die sich bei ihrer Automation stellen. Es liefert die Grundlagen für den Aufbau und den weiteren Verlauf dieser Arbeit.

Dieses Kapitel gliedert sich in vier Abschnitte. Der erste Abschnitt definiert die Generalisierung in der Kartographie, motiviert die Generalisierung durch das Aufzeigen ihrer Notwendigkeit und gibt eine erste Charakterisierung der Generalisierung ab, die in einer Aufteilung der Generalisierung in zwei Teile besteht. Der erste Bestandteil sind die elementaren Vorgänge der Generalisierung, die in diesem Abschnitt abgehandelt werden, während der zweite Bestandteil die Einschränkungen und Nebenbedingungen darstellt, die in den weiteren Abschnitten erläutert werden.

Der zweite Abschnitt beschäftigt sich mit der Modellbildung in der Kartographie. Besonderes Augenmerk wird hier auf die Bildung von digitalen Modellen gelegt, da nur diese automatisch generalisiert werden können. Aus den verschiedenen Arten von digitalen Modellen leiten sich verschiedene Arten von Generalisierungen ab. Betrachtet man dies im Zusammenhang mit der Datenhaltung von digitalen Modellen, so ergeben sich auch hier Unterschiede für die Generalisierung.

Der dritte Abschnitt befasst sich mit der Automation der Generalisierung. Zunächst erfolgt ein kurzer Überblick über die geschichtliche Entwicklung der Automation, bevor sich der Abschnitt auf die im Rahmen dieser Arbeit besonders wichtige Modellgeneralisierung konzentriert und ein Verfahren hierfür genauer beschreibt. Den Abschluss dieses Abschnittes bildet eine Klassifizierung von Generalisierungsmodellen, die es ermöglicht, unterschiedliche Ansätze zur automatischen Generalisierung besser zu verstehen und voneinander zu unterscheiden.

Für die Methodik dieser Arbeit, welche das Thema des vierten und letzten Abschnittes dieses Kapitels bildet, ist die am Ende des dritten Abschnittes getroffene Klassifizierung von besonderer Bedeutung, da an Hand dieser die Zielsetzung der Arbeit deutlich gemacht werden kann. Des Weiteren fasst dieser Abschnitt die Essenz der vorangegangenen Abschnitte zusammen und kann daraus das weitere Vorgehen dieser Arbeit auf beste motivieren.

2.1 Generalisierung in der Kartographie

Generalisierung ist einer der zentralen Begriffe in der Kartographie (Brassel 1990). Um sich dem Begriff zu nähern, wird im Folgenden zunächst die allgemeine, nicht auf die Kartographie beschränkte Bedeutung dieses Begriffes betrachtet. Brassel (1990) zitiert dazu aus der Brockhaus Enzyklopädie: „Generalisierung bedeutet Verallgemeinerung, den Übergang von Einzelfällen zum Allgemeinbegriff oder zum allgemeinen Gesetz“. Die Bedeutung der Generalisierung im Zusammenhang mit menschlichem Handeln betonen Brassel und Weibel (1988): „[Generalization] represents a basic human activity [...]“ Ebenda beschreiben sie den Gehalt der Generalisierung folgendermaßen: „The term generalization focuses on the extraction of the general, crucial elements of reality.“ Weiter führen sie aus, dass die Extraktion des Wesentlichen unlösbar verbunden ist mit dem Fortlassen des Spezifischen, des Zufälligen und des Unwichtigen.

In der Kartographie befasst sich die Generalisierung mit der Verallgemeinerung von Landkarten im Speziellen und Geodaten im Allgemeinen. Die Vorgänge umfassen die Extraktion des Wesentlichen und das Fortlassen von Unwesentlichem. Zudem wirken eine ganze Reihe von Randbedingungen auf den Vorgang der Generalisierung ein. Der erste Unterabschnitt versucht eine Definition des Begriffes der Generalisierung in der Kartographie, sowie eine erste Charakterisierung.

Der zweite Unterabschnitt begründet die *Notwendigkeit* der Generalisierung. Insbesondere werden dort die unterschiedlichen Motive für die Generalisierung von Karten gegenüber der Generalisierung von nicht visualisierten Geodaten hervorgehoben.

Im dritten Unterabschnitt erfolgt eine Gliederung der Generalisierung in zwei Bestandteile. Diese Gliederung ist für die Automatisierung der Generalisierung von zentraler Bedeutung, da aus ihr eine Begründung für prinzipiell unterschiedliche Ansätze zur Realisierung der Automation folgt. Zum einen ist dies eine Generalisierung „von unten“, aus den einzelnen Generalisierungsschritten heraus, zum anderen eine Generalisierung „von oben“, von der Gesamterscheinung der Karte oder den Charakteristika der Geodaten her.

Der erste Bestandteil dieser vorgenommenen Gliederung sind die *elementaren Vorgänge der Generalisierung* (auch: *Generalisierungsmaßnahmen*). Sie bilden das Thema des vierten Unterabschnittes. Verschiedene Ansätze zur Definition von elementaren Vorgängen werden vorgestellt, und es werden die Zusammenhänge dieser elementaren Vorgänge untersucht. Der zweite Bestandteil, die Randbedingungen und Einschränkungen, werden im weiteren Verlauf dieses Kapitels eingehender behandelt.

2.1.1 Definition der Generalisierung

Generalisierung in der Kartographie war vor dem Aufkommen der elektronischen Datenverarbeitung annähernd gleichbedeutend mit dem Begriff der *kartographischen Generalisierung*. Durch den Übergang von der analogen zur digitalen Kartographie hat sich der

2.1 Generalisierung in der Kartographie

Begriff der Generalisierung erweitert auf die Generalisierung *kartenunabhängiger* Geodaten, wie sie zum Beispiel in einer Datenbank gespeichert sind. Heute bildet die kartographische Generalisierung einen Spezialfall der Generalisierung in der Kartographie¹, wenn sie auch immer noch den weitaus größten Anteil davon ausmacht. Dennoch oder gerade deswegen werden Definitionen und Beschreibungen der kartographischen Generalisierung auch heute noch auf die Generalisierung in der Kartographie übertragen.

Eine aktuelle Definition der Generalisierung in der Kartographie findet sich in einer Publikation der Schweizer Gesellschaft für Kartographie (SGK 2002). Sie ist hervorgegangen aus der Definition der kartographischen Generalisierung und berücksichtigt die aktuellen Entwicklungen der elektronischen Datenverarbeitung in der Kartographie. Die Definition lautet wie folgt:

„Die Generalisierung ist die massstabgerechte inhaltliche und grafische Vereinfachung der komplexen Wirklichkeit auf der Grundlage digitaler Landschaftsmodelle, von Plänen oder Karten größeren Massstabes. Die Generalisierung besteht hauptsächlich aus der zweckentsprechenden Auswahl und Zusammenfassung der Objekte sowie der möglichst lagegenauen, charakteristischen, richtigen und eindeutigen grafischen Darstellung.“ (SGK 2002)

Diese Definition stellt zwei Punkte in den Vordergrund: Zum einen die *Vereinfachung* von Daten unterschiedlicher Herkunft. Diese Vereinfachung ist inhaltlich und grafisch zu sehen. Zum anderen betont sie die Hauptbestandteile der Generalisierung: *Auswahl* und *Zusammenfassung*, sowie die *Darstellung* der Objekte.

Die obige Definition bezieht sich auf einen bestimmten Maßstab („massstabgerechte [...] Vereinfachung“) und auf die Darstellung der Daten. Ein fester Maßstab muss jedoch nicht zwingend zum Ziel einer Generalisierung gehören. Schürer (2002) spricht im Zusammenhang der Modellgeneralisierung (siehe Abschnitt 2.2.2) von einem Übergang zwischen *Strukturierungsgraden*, um vom Begriff eines festen Maßstabes zu abstrahieren. Auch die graphische Darstellung steht heute nicht mehr ausschließlich im Zentrum der Generalisierung. Im Zuge des vermehrten Einsatzes von Geoinformationssystemen zur Verarbeitung von Geodaten verliert die graphische Darstellung an Bedeutung und wird zum Kommunikationshilfsmittel für den Menschen bei der Betrachtung von Geodaten herabgestuft. Denn Datenverarbeitung und Analyse kann mit Hilfe der EDV zum Teil auch ohne graphische Darstellung möglich gemacht werden. Schuurman (1999) verwendet in diesem Zusammenhang das Bild von der graphischen Darstellung als „tip of the iceberg“ (Spitze des Eisbergs), während die Datenbasis hinter der graphischen Darstellung weitaus mehr Informationen enthalten kann und in diesem Bild dem nicht sichtbaren Teil des Eisbergs unter der Wasseroberfläche entspricht. Dennoch ist der Mensch im Zentrum des Handelns und der Entscheidungen in vielen Fällen auf die graphische Darstellung der Geodaten angewiesen. Die graphische Darstellung wird somit auch in der Zukunft ihre Bedeutung und ihre Berechtigung haben.

¹Hake u. a. (2002) bezeichnen die kartographische Generalisierung als *Generalisierung im engeren Sinne*.

2.1.2 Notwendigkeit der Generalisierung

Dieser Abschnitt motiviert die Generalisierung, indem ihre *Notwendigkeit* aufgezeigt wird. Dabei wird eine Unterscheidung getroffen zwischen der Begründung für die Generalisierung von Karten und für die Generalisierung von nicht visualisierten Geodaten.

Notwendigkeit der Generalisierung von Karten

Laubert u. a. (1982) begründen die Notwendigkeit der Generalisierung folgendermaßen: Bezogen auf Karten kann eine „verkleinerte Darstellung [...] die Wirklichkeit auf Grund des reduzierten Platzangebotes nicht mehr vollständig wiedergeben. [...] Je kleiner der Maßstab wird, umso weniger Platz steht in der Karte für die Darstellung desselben Realweltausschnitts zur Verfügung. Auf einer begrenzten Kartenfläche wird mit der Verkleinerung des Kartenmaßstabes ein immer größeres Gebiet abgebildet. Würde man die Karten nur photographisch von Maßstab zu Maßstab verkleinern, wäre die Karte von einem bestimmten Grenzbereich an nicht mehr lesbar.“ Abbildung 2.1 verdeutlicht diese Argumentation am Beispiel von Ausschnitten topographischer Karten, in denen eine (photographische) Verkleinerung einer Generalisierung gegenübergestellt wird. Links in der Abbildung sieht man den Ausschnitt einer Karte im Maßstab 1 : 25 000, rechts oben denselben Ausschnitt vierfach verkleinert. Diesem verkleinerten Ausschnitt ist rechts unten der entsprechende Ausschnitt der topographischen Karte im Maßstab 1 : 100 000 gegenübergestellt (Kartenausschnitte entnommen aus SGK 2002). Es wird deutlich, dass die photographische Verkleinerung zwar alle Informationen der Ausgangskarte enthält; diese für den Menschen aber nicht mehr wahrnehmbar sind.² Eine Veränderung der Darstellung, mithin eine Generalisierung, ist also nötig.

Hake u. a. (2002) betonen zwei Prinzipien der Kartenerstellung, die bei der Generalisierung miteinander konkurrieren. Das Prinzip der *geometrischen Richtigkeit* wird dem Prinzip *der Vollständigkeit* gegenübergestellt. Beide Prinzipien können bei einer Generalisierung im Allgemeinen nicht gleich berücksichtigt werden. Verletzt man das erste Prinzip, so schränkt man die Lesbarkeit ein, verletzt man das zweite Prinzip, so verzichtet man auf Wiedergabe.

Notwendigkeit der Generalisierung von kartenunabhängigen Daten

Während die Generalisierung von Karten offensichtlich notwendig ist, erscheint die Generalisierung von Geodaten in einer Datenbank zunächst weniger zwingend. Im Zusammenhang mit dem Amtlichen Topographisch-Kartographischen Informationssystem ATKIS beschreibt Harbeck (1987) den *Idealzustand* eines einzigen Modelles von erfassten Daten – möglichst im Maßstab 1 : 1 –, aus dem Karten in verschiedenen Maßstäben abzuleiten

²Aus den Grenzen der Wahrnehmung folgt die Definition von Minimaldimensionen und daraus folgend die Festlegung von graphischen Mindestgrößen (Hake u. a. 2002). Aus diesen ergeben sich wiederum Vorgaben für die Kartengestaltung.



Topographische Karte 1 : 25 000



Photographisch
verkleinerte Karte
1 : 25 000 in 1 : 100 000



Generalisierte Karte
1 : 100 000

Abbildung 2.1: Notwendigkeit der Generalisierung (aus SGK 2002).

sind. Er sieht jedoch die ungelöste Problematik der automatischen Generalisierung, so dass Modelle mit geringerer Auflösung, entsprechend den Kartenauflösungen als Zwischenspeicher nötig sind. Die Probleme liegen hier insbesondere bei den (zur Zeit) begrenzten technischen und algorithmischen Möglichkeiten für die Realisierung eines solchen Idealzustands mit automatischer Ableitung. Trotz der unbestrittenen Erfolge, die auf dem Gebiet der Automation inzwischen gemacht wurden, ist dieser Idealzustand immer noch in weiter Ferne (Müller u. a. 1995, Weibel und Dutton 1999, Spaccapetra u. a. 2000). Deshalb ist eine Generalisierung auch für digitale Modelle nötig. Hake u. a. (2002) formulieren diese Problematik abstrahiert, indem sie darauf hinweisen, dass die Erzeugung von digitalen Modellen mit einer semantischen und geometrischen Auflösung, die für einen *bestimmten Zweck* ausgelegt sind, nötig ist.

2.1.3 Bestandteile der Generalisierung

Um das Wesen der Generalisierung besser zu erfassen, gliedert Brassel (1990) die Generalisierung in *zwei Bereiche* auf; er stützt sich dabei im Wesentlichen auf die Ergebnisse von Hake (1975) und Robinson u. a. (1984).

Generalisierungsprozesse

Als den ersten Bereich bezeichnet Brassel die *Generalisierungsprozesse* oder *Elemente der Generalisierung*. Unter diesen Begriffen werden Einzelschritte zusammengefasst, in welche die oben genannten Autoren den Gesamtprozess der Generalisierung aufteilen. Im angelsächsischen Sprachraum spricht man von *Generalisierungsoperatoren* (generalization operators, McMaster und Shea 1992). In deutschen Publikationen werden diese Einzelschritte unter den Begriffen der *elementaren Vorgänge der Generalisierung* (Hake u. a. 2002) oder der *Generalisierungsmaßnahmen* (Bollmann und Koch 2001) zusammengefasst. Diese elementaren Vorgänge wurden zum einen entwickelt aus der Nachbildung der Schritte der manuellen Generalisierung und zum anderen aus Techniken, die allein auf theoretischen und analytischen Überlegungen beruhen (McMaster und Shea 1992). Abschnitt 2.1.4 befasst sich mit der genaueren Beschreibung der elementaren Vorgänge der Generalisierung.

Randbedingungen und Einschränkungen

Der Bereich, den Brassel als den zweiten Bestandteil der Generalisierung bezeichnet, besteht aus *Randbedingungen* und *Einschränkungen*, die die elementaren Vorgänge der Generalisierung beeinflussen. Robinson u. a. (1984) koppeln diese Einschränkungen an die graphische Darstellung und halten sie allgemein. Sie umfassen zum Beispiel den Kartenzweck und die Kartenfunktion, den Maßstab und die graphischen Limite.

Nach Hake (1975) sind die Randbedingungen zum einen durch die Kategorien der kartographischen Aussage wie räumlicher Bezug, substantielle Merkmale und zeitliches Verhalten gegeben, zum anderen sind sie nach den Anwendungsbereichen der Generalisierung (Objektgeneralisierung und kartographische Generalisierung) unterteilt.

Aus der Gliederung der Generalisierung in diese zwei Bestandteile folgen unterschiedliche Methoden der Automation der Generalisierung. Stützt man sich auf die elementaren Vorgänge, so entsteht eine Generalisierung „von unten“, in der vom Kleineren (den einzelnen Generalisierungsschritten) ausgehend das Große (die Karte) generalisiert wird. Stützt man sich auf die Randbedingungen und Einschränkungen, so entsteht eine Generalisierung „von oben“, in der zunächst das Erscheinungsbild der Karte beziehungsweise die Charakteristika der Geodaten beschrieben werden (das Große) und von diesem Startpunkt ausgehend die einzelnen Elemente (das Kleine) generalisiert werden.

Die Generalisierung „von unten“, also die elementaren Vorgänge der Generalisierung, werden im nächsten Unterabschnitt weiter behandelt. Die Generalisierung „von oben“ ist jedoch der Ansatz, dem im weiteren Verlauf dieser Arbeit größere Bedeutung zukommen wird.

2.1.4 Elementare Vorgänge der Generalisierung

Entsprechend der vorgestellten Gliederung beschreibt dieser Abschnitt die elementaren Vorgänge der Generalisierung. In der Literatur gibt es verschiedene Unterteilungen der Generalisierungsvorgänge, die sich hauptsächlich durch unterschiedliche Granularität auszeichnen, aber auch zum Teil verschiedene Begrifflichkeiten verwenden. Besonders deutlich tritt der Unterschied zwischen den Unterteilungen in der deutschsprachigen Literatur (zum Beispiel Hake u. a. 2002) im Gegensatz zu den Unterteilungen im angelsächsischen Sprachraum (zum Beispiel McMaster und Shea 1992) zu Tage. Eine dritte Strömung entstammt aus dem Zusammenhang mit dem AGENT-Projekt³ (zum Beispiel Barrault u. a. 2001). In den letzten Jahren ist eine Annäherung der deutschsprachigen Sichtweise, vertreten durch Hake u. a. (2002) an die angelsächsische Sichtweise zu erkennen, belegbar durch Vergleich der verschiedenen Auflagen des Buches KARTOGRAPHIE (Hake u. a. 2002, Hake und Grünreich 1994, Hake 1982). Ein Vergleich der unterschiedlichen Unterteilungen findet sich bei Galanda (2003). Über die Uneinigkeit von Fachleuten über eine einheitliche Definition von Generalisierungsvorgängen haben Rieger und Coulson (1993) eine empirische Untersuchung durchgeführt.

An dieser Stelle wird die Unterteilung nach Hake u. a. (2002) vorgestellt (siehe Abbildung 2.2). Eine erste, grobe Unterteilung besteht in der Unterscheidung von geometrischen Vorgängen und sachlichen Vorgängen mit geometrischer Wirkung. Die rein geometrischen Vorgängen sind das *Vereinfachen*, mit dem Teilbereich *Glätten*, das *Vergrößern* und das *Verdrängen*. Die sachlichen Vorgänge mit geometrischer Wirkung sind das *Zusammenfassen*, das *Auswählen*, das *Klassifizieren* und das *Bewerten*. Bei McMaster und Shea (1992) wird der Vorgang des Zusammenfassens weiter unterteilt, abhängig von der geometrischen Ausprägung (Geometriety). Außerdem findet sich dort zusätzlich der Vorgang des *Collapse* (Qualitätsumschlag, Geometrietywechsel), der nach der Unterscheidung nach Hake u. a. (2002) am ehesten in den Bereich des Typisierens (Teilbereich vom Klassifizieren) fällt.

Zusammenhänge zwischen den elementaren Vorgängen

Der Ablauf einer Generalisierung kann im Allgemeinen nicht durch eine lineare Ordnung der elementaren Vorgänge wiedergegeben werden. Vielmehr lösen einige elementare Vorgänge andere aus (zum Beispiel Verdrängen als Folge von Vergrößern). Die Wirkung und Reihenfolge der elementaren Vorgänge ist nicht völlig unabhängig voneinander (Hake u. a. 2002).

Laubert u. a. (1982) schreiben, dass die einzelnen Generalisierungsvorgänge als Einheit zu sehen seien, da sie sich gegenseitig beeinflussen. Bei der manuellen Generalisierung

³Zum AGENT-Projekt siehe auch Abschnitt 2.3.1.













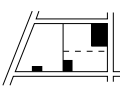


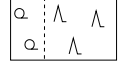





Elementarer Vorgang		Darstellung		
Merkmal	Bezeichnung <i>Teilbereiche</i>	Ausgangskarte im Ausgangsmaßstab 1 : m	Folgekarte	
			im Ausgangsmaßstab 1 : m	im Folgemaßstab 1 : 4 m
<i>Geometrisch</i>				
1	Vereinfachen <i>Teilbereich: Glätten</i>			
2	Vergrößern (Hauptfall: Verbreitern)			
3	Verdrängen (Folge von Nr. 2)			
<i>Sachlich mit geometrischer Wirkung</i>				
4	Zusammenfassen			
5	Auswählen (Erhalten oder Fortlassen)			
6	Klassifizieren <i>Teilbereich: Heraufstufen</i> <i>Teilbereich: Typisieren</i>			
7	Bewerten (Betonen oder Mindern)			

Abbildung 2.2: Die sieben elementaren Vorgänge der Generalisierung nach Hake u. a. (2002).

müssen alle notwendigen Maßnahmen nacheinander geprüft und im Zusammenhang entschieden und realisiert werden. Die *Auswahl* bezeichnen sie als einen Vorgang mit Vorrangstellung.

Im Zusammenhang mit dem kartographischen Generalisieren heißt es in (SGK 1980): „[Das] Wechselspiel zwischen Weglassung und Unterdrückung einerseits, Übertreibung und Betonung andererseits, begleitet die gesamte Entstehung der Karte von der Konzeption bis zum endgültigen Original für die Vervielfältigung. Wesentlich ist dabei das gute Zusammenspiel aller Massnahmen im Hinblick auf ein ausgewogenes, möglichst einheitlich bearbeitetes Gesamtbild.“

Die gegebenen Zitate belegen die Komplexität des Zusammenwirkens der elementaren Generalisierungsvorgänge. Eine eindeutige Abfolge kann bei der Generalisierung im Allgemeinen also nicht gegeben werden. Sie ist vielmehr situationsabhängig und kann während

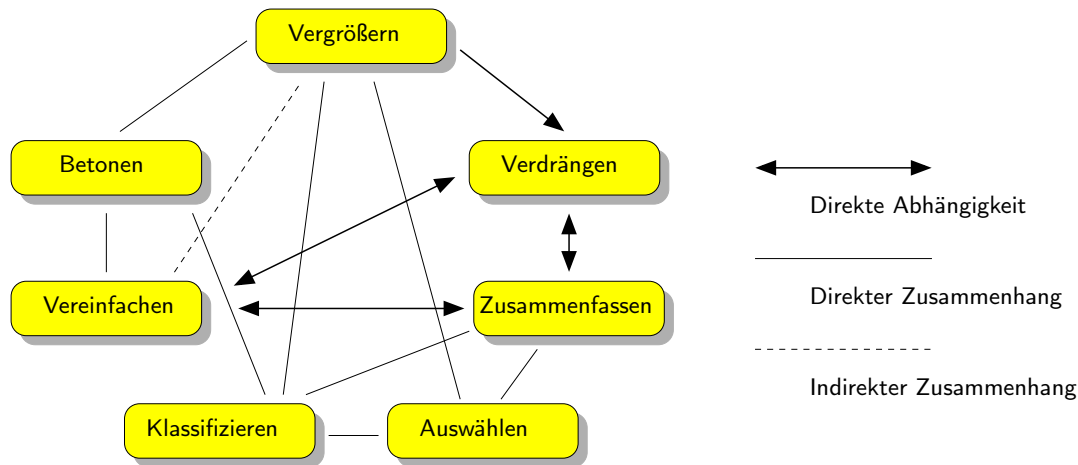


Abbildung 2.3: Zusammenhänge der sieben elementaren Generalisierungsvorgänge nach Heisser (1996).

der Generalisierung variieren. Abbildung 2.3 zeigt in einem Schema nach Heisser (1996) Zusammenhänge zwischen sieben elementaren Generalisierungsvorgängen. Er stützt sich dabei auf die Unterteilung von Hake und Grünreich (1994). Bei genauerer Betrachtung mag dieses Schema Lücken aufweisen. Dennoch zeigt es deutlich, dass sich Zyklen von Abhängigkeiten in den Beziehungen zwischen elementaren Generalisierungsvorgängen befinden.

Eine Automation der Generalisierung hat die Zusammenhänge zwischen den Generalisierungsvorgängen zu berücksichtigen. Insbesondere kann festgehalten werden, dass eine sequentielle Abarbeitung der Generalisierungsvorgänge im Allgemeinen nicht zu einem befriedigenden Erfolg führen kann, da ringförmige Abhängigkeiten zwischen den Generalisierungsvorgängen bestehen können. Für die Realisierung der Generalisierung in einem Programm ist deshalb auf geeignete Methoden zur Prozesssteuerung zurückzugreifen. Abschnitt 2.3.3 geht im Zusammenhang mit einer Klassifizierung von Methoden zur automatischen Generalisierung weiter auf diese Problematik ein.

2.2 Kartographische Modellbildung

Der Begriff der Modellbildung hat in der Kartographie eine lange Tradition. Ein Beispiel sind die kartographischen Modelle im Kontext von Kommunikationsprozessen. Man unterscheidet Primär-, Sekundär- und Tertiärmodelle (Hake u. a. 2002). Aus der Erfassung der Landschaft entsteht ein erstes Modell (Primärmodell), das kartographisch aufbereitet in die Karte (Sekundärmodell) überführt wird. Durch Betrachten der Karte entsteht beim Nutzer in der Vorstellung ein drittes Modell der Landschaft (Tertiärmodell).

Im Zusammenhang mit der elektronischen Datenverarbeitung erweitert sich diese Modellkette um digitale Modelle der Landschaft, wobei die technischen Mittel umfangreichere Abbildungen der Landschaft ermöglichen. Abschnitt 2.2.1 beschreibt heute übliche Unterscheidungen von digitalen Modellen in der Kartographie. Aus dieser Unterscheidung folgen verschiedene Arten von Generalisierungen. Diese sind Thema von Abschnitt 2.2.2.

Betrachtet man die Datenhaltung, so ergeben sich für die Generalisierung von digitalen Modellen unterschiedliche Ansätze, die sich vom Vorgehen bei der manuellen Generalisierung unterscheiden. Abschnitt 2.2.3 beschreibt zwei grundsätzlich unterschiedliche Ansätze der Generalisierung im Hinblick auf die Datenhaltung.

2.2.1 Digitale Modelle

Hake u. a. (2002) bezeichnen den allgemeinsten Fall eines Ausschnittes der Wirklichkeit in einer Datenbank mittels eines an Objekten orientierten Datenmodells als *Digitales Objektmodell* (DOM). Wenn das Thema der Darstellung die Topographie der Landschaft ist, so bezeichnet man das DOM als *Digitales Landschaftsmodell* (DLM).⁴

Für die Darstellung in Karten, beziehungsweise bei der Aufbereitung der DLM für Karten hat sich der Begriff *Digitales Kartographisches Modell* (DKM) herausgebildet (Hake u. a. 2002). In einem DKM werden Objekte kartennah gespeichert, das heißt mit der Geometrie, mit der sie in einer Karte erscheinen sollen, und mit einer Signaturierung oder einer Signaturierungsvorschrift.

Insofern ist ein DKM immer mit einer festen Auflösung, mit einem Maßstab, verbunden, während ein DLM nicht zwingend einer Auflösung zugeordnet sein muss (siehe Abschnitt 2.1.2). Aus den dort dargelegten Gründen, hauptsächlich technischer Art, werden aber auch DLM in verschiedenen Strukturierungsgraden, einem Begriff, der von Schürer (2002) mit geprägt wurde, vorgehalten.

Die amtliche Kartographie in Deutschland ist ursprünglich von einer klaren Trennung zwischen DLM und DKM ausgegangen (AdV 1989). Inzwischen jedoch ist das DKM als eigenständiges Modell in den Hintergrund getreten und in das DLM integriert worden (Birth 2003). Eine größere Bedeutung kommt inzwischen der Digitalen Topographischen Karte (DTK) zu, die als Endprodukt rasterorientiert ist und aus dem DLM abgeleitet wird.

Mit der Modellierung von digitalen Modellen einer Auflösung oder eines Strukturierungsgrades befasst sich Kapitel 3.1. Die Modellierung der Daten innerhalb *eines* Datenmodells ist das dortige Thema.

⁴Hake u. a. (2002, S. 163) bezeichnen hierbei die Landschaft als das Objekt, das im DOM abgebildet wird. In der Praxis besteht ein DLM jedoch aus der Darstellung einer *Menge* von topographischen Objekten, so dass dieser Argumentation von Hake u. a. (2002) hier nicht gefolgt werden kann.

Objektgeneralisierung		Kartographische Generalisierung			
Erfassungsgeneralisierung		Modellgeneralisierung			
Umwelt → Karte	Umwelt → DLM	DLM → DLM	DLM → DKM	DKM → Karte	Karte → Karte

Tabelle 2.1: Arten der Generalisierung nach Hake u. a. (2002).

Objektgeneralisierung		Kartographische Generalisierung				
Erfassungsgen.		Modellgen.				
Umwelt → Karte	Umwelt → DLM	DLM → DLM	DLM → DKM	DKM → DKM	DKM → Karte	Karte → Karte

Tabelle 2.2: Arten der Generalisierung nach Schürer (2002).

2.2.2 Arten der Generalisierung

Aus der Differenzierung von verschiedenen digitalen Modellen folgt dementsprechend eine Differenzierung in verschiedene Arten von Generalisierungen. Die Unterscheidung erfolgt im Allgemeinen danach, ob von einem DLM oder von einem DKM zu einem DLM oder zu einem DKM generalisiert wird.

Die Begrifflichkeiten sind in der Literatur jedoch nicht eindeutig. Da es sich bei den Ausgangs- und Zieldaten jeweils um digitale Modelle handelt, nennen Hake u. a. (2002) diese Art der Generalisierung folgerichtig *Modellgeneralisierung* und stellen sie der *Erfassungsgeneralisierung* gegenüber. Bei dieser findet eine Generalisierung bei der Aufnahme der Daten aus der Umwelt in ein digitales Objektmodell oder in eine Grundkarte statt. Des Weiteren unterscheiden sie zwischen *Objektgeneralisierung*, bei der die Generalisierung der *Landschaftsobjekte* im Vordergrund steht, und *kartographischer Generalisierung*, bei der die Generalisierung der *Kartenobjekte* im Vordergrund steht. Die Arten der Generalisierung nach Hake u. a. (2002) sind in Tabelle 2.1 zusammengefasst.

Für Schürer (2002) ist obige Definition der Modellgeneralisierung zu allgemein gehalten. Er nennt die von Hake u. a. (2002) geprägte Definition die *allgemeine Sichtweise* der Modellgeneralisierung. Demgegenüber stellt er eine Sichtweise, bei der die Modellgeneralisierung ganz aus dem Kontext der kartographischen Generalisierung gelöst wird und komplett dem Bereich der Objektgeneralisierung zugeordnet wird. Er beschreibt damit eine *engere Sichtweise* der Modellgeneralisierung, die sich exakt auf die Generalisierung von DLM zu DLM bezieht. Die Arten der Generalisierung nach Schürer (2002) sind in Tabelle 2.2 zusammengefasst.

Der Ansatz von Schürer (2002) schließt sich damit älteren Ausführungen von zum Beispiel Schuurman (1999) und Brassel und Weibel (1988) an, in der die Trennung zwischen Modellgeneralisierung (bei Brassel und Weibel auch *statistische Generalisierung* genannt) und kartographischer Generalisierung wie folgt charakterisiert ist: Die Modellgeneralisierung ist unabhängig von den Eigenschaften der Visualisierung, und unterliegt damit

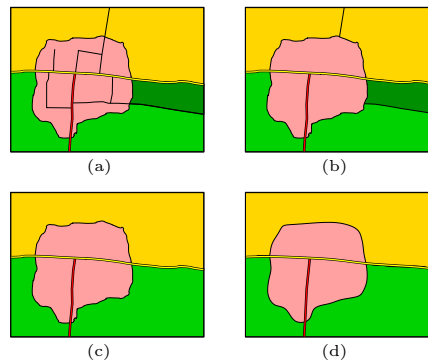


Abbildung 2.4: *Generalisierung als Reduktion (nach Bobzien 2003).*

hauptsächlich statistischen Betrachtungen, während die Visualisierung als das eigentliche Ziel der kartographischen Generalisierung bezeichnet wird.

Diese Arbeit ist im Kontext der Modellgeneralisierung entstanden, und zwar nach der Definition der engeren Sichtweise der Modellgeneralisierung, der Generalisierung von DLM zu DLM. Insgesamt wurde eine Ableitungskette von drei voneinander abgeleiteten DLM erzeugt. Da einige hartnäckige Probleme der kartographischen Generalisierung dabei nicht berücksichtigt werden mussten, konnte eine Automation leichter erfolgen. Die gewonnenen Erkenntnisse liefern Material für die Formalisierung der Generalisierung (siehe Kapitel 4, Generalisierung als Funktion) und die Benennung von Bedingungen (siehe Kapitel 5, Invariante Eigenschaften der Generalisierungsfunktion) auf Basis der Modellgeneralisierung im engeren Sinne. Durch die Ausführungen kann wiederum auf die Modellgeneralisierung im allgemeineren Sinne geschlossen werden.

2.2.3 Methoden der Datenhaltung für die Generalisierung

Bislang wurde Generalisierung *prozessorientiert* betrachtet: Die Generalisierung als ein Prozess, der aus einem Datenbestand einen anderen erzeugt. Weibel und Dutton (1999) setzen dieser Betrachtungsweise eine repräsentationsorientierte Sichtweise entgegen, bei der die Generalisierung aus der Sicht der *Datenhaltung* betrachtet wird. Die Datenhaltung auf einer abstrakten Ebene ist Thema dieses Unterabschnittes.

Bobzien (2003) unterscheidet im Zusammenhang mit der Datenhaltung bei der Generalisierung zwischen zwei Methoden: der Generalisierung als *Reduktion* und der Generalisierung als *Konstruktion*. Er führt dazu als Beispiel zwei unterschiedliche Abläufe einer Generalisierung an (siehe Abbildungen 2.4 und 2.5).

Bei der Anwendung der ersten Methode, der Generalisierung als Reduktion, *verändert* man die Ursprungsdaten so lange, bis sie dem gewünschten Ergebnis entsprechen. In Abbildung 2.4 sind beispielhaft die Schritte der Auswahl und des Fortlassens (von a

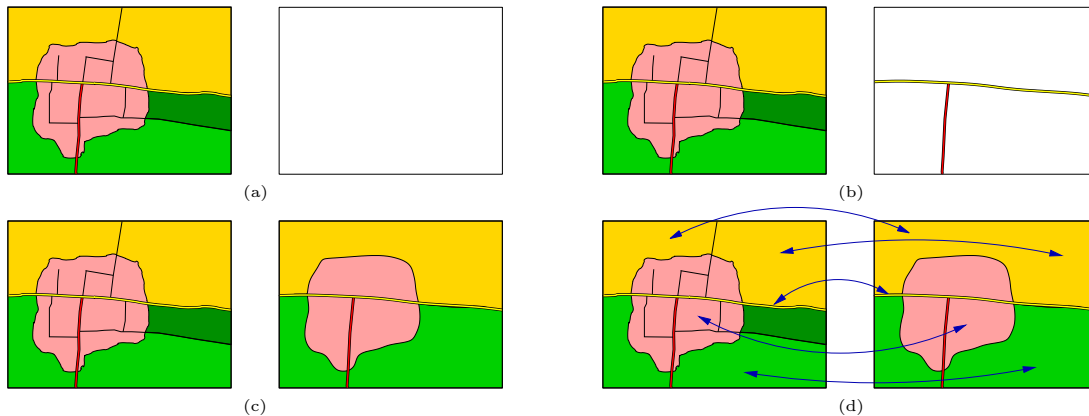


Abbildung 2.5: *Generalisierung als Konstruktion (nach Bobzien 2003).*

nach b), des Zusammenfassens (von b nach c) und der Vereinfachung (von c nach d) dargestellt.

Bei der Anwendung der zweiten Methode *behält* man die Ursprungsdaten *bei* und *erzeugt* einen neuen Datensatz neben dem ersten Datensatz.⁵ In Abbildung 2.5 ist der neue Datensatz zunächst leer (a), dann werden linienförmige Elemente erzeugt (b), danach die flächenförmigen Elemente (c). Die letzte Abbildung (d) zeigt die Verknüpfung zwischen den ungeneralisierten und generalisierten Objekten.

Mit dieser Unterscheidung motiviert Bobzien (2003) die Verknüpfung von ungeneralisierten und generalisierten Daten in zwei unterschiedlichen Datensätzen. Im Endeffekt ist es dafür allerdings unerheblich, ob die zu generalisierenden Daten zunächst kopiert und dann reduziert werden (Generalisierung als Reduktion), oder ob der neue Datensatz aus den zu generalisierenden Daten konstruiert wird (Generalisierung als Konstruktion). Entscheidend ist, dass Verknüpfungen zwischen den beiden Datensätzen bestehen. Im Falle der Reduktion werden sie beim Kopieren der ungeneralisierten Daten angelegt, im Falle der Konstruktion beim Konstruieren der generalisierten Daten.

Bobzien (2003) arbeitet zwei Vorteile der Verknüpfung von ungeneralisierten und generalisierten Daten heraus:

Vorteil 1: Situationsanalyse

Nach dem Ausführen von ersten Generalisierungsschritten kann bei Vorhandensein von Verknüpfungen zwischen ungeneralisierten und generalisierten Daten immer noch auf

⁵Selbstverständlich werden auch bei der Generalisierung als Reduktion die Ursprungsdaten in dem Sinne beibehalten, dass die Generalisierung im Allgemeinen auf einer *Kopie* der Daten stattfindet. Als entscheidenden Unterschied zwischen beiden Methoden sieht Bobzien (2003), dass während und nach der Generalisierung als Konstruktion jederzeit *der Bezug der generalisierten Daten zu den Ursprungsdaten* hergestellt werden kann.



Abbildung 2.6: In den generalisierten Daten sind Straßenzüge auch ohne Signaturierung erkennbar (Beispiele aus SGK 1980).

den ursprünglichen Datenbestand zurückgegriffen werden. Informationen gehen so nicht verloren und finden bei Bedarf Berücksichtigung im generalisierten Datenbestand. Es ist somit *während* der Generalisierung eine *Situationsanalyse* möglich. Es kann im Laufe der Generalisierung nötig sein, den Ursprungszustand der Daten zu analysieren, um weitere Generalisierungsschritte davon abhängig zu machen. Sowohl bei der Modellgeneralisierung im engeren als auch im weiteren Sinne finden sich Beispiele, bei denen diese Informationen benötigt werden und in dem generalisierten Modell noch sichtbar sind: Abbildung 2.6 zeigt als Beispiel die Generalisierung eines Wohnviertels, bei dem im generalisierten Datenbestand Straßen zwar nicht mehr signaturiert, aber durch die Anordnung der Häuser doch in ihrer Führung sichtbar sind (durch rote Ovale gekennzeichnet). Hätte man bei der Generalisierung zunächst die Straße fortgelassen und würde die Informationen darüber löschen, so wären die Häuser nicht notwendigerweise so generalisiert worden, dass man die Straßenführung noch erkennen kann. Bei SGK (2002) heißt es dazu, dass man „nicht dargestellte Verkehrswege mit der Orientierung der Häuser andeuten“ soll, und dass man bei einem „Wohnquartier [...] statt Quartierstrassensignatur [einen] durchgehenden Zwischenraum darstellen“ soll.

Ähnliches gilt für die Flächenzusammenfassung in der Modellgeneralisierung (siehe Bobzien 2001): Ein Zusammenfassen von Flächen kann die ursprünglichen Straßenführungen erkennbar machen, obwohl die Straßen entfallen sind. Ein von den Erfassungskriterien nicht erzwungenes Zusammenfassen über weggefallene Straßenzüge hinweg würde unnötig noch vorhandene Informationen zerstören. Im Gegenteil kann es auch für DLM wünschenswert sein, Landschaftsstrukturen zu erhalten, obwohl dies nach den Erfassungskriterien nicht erforderlich ist. Bobzien (2001) befasst sich mit Kriterien für die Flächenzusammenfassung bei der Modellgeneralisierung. Das hier skizzierte Kriterium der Landschaftsstrukturhaltung ist nach der dortigen Einteilung ein geometrisches Bewertungskriterium, das einer lokalen Bewertungsstrategie gehorcht. Beispiele für erhaltene Landschaftsstrukturen nach einer Modellgeneralisierung finden sich bei Schürer (2002) und in Kapitel 6 in dieser Arbeit.

Vorteil 2: Fortführung

Der zweite Vorteil von Verknüpfungen zwischen ungeneralisierten und generalisierten Daten ist die Ermöglichung der *automatischen Fortführung*. Wird im Ausgangsdatenbestand eine Änderung vollzogen, so muss nicht der gesamte Datenbestand neu generalisiert werden. Durch die Verknüpfung zwischen den ungeneralisierten und den generalisierten Objekten ist es möglich, genau die Objekte, die neu generalisiert werden müssen, zu identifizieren. Die Generalisierung kann sich somit auf die Generalisierung derjenigen Objekte und deren Umgebungen konzentrieren, die von der Fortführung betroffen sind.

Die beiden angeführten Vorteile von Verknüpfungen zwischen ungeneralisierten und generalisierten Daten motivieren die Einführung von Datenstrukturen, welche die Speicherung der Beziehungen zwischen den Objekten der verschiedenen Modelle erlauben. Diese Datenstrukturen sind Thema von Abschnitt 3.2.

2.3 Automation der Generalisierung

In diesem Abschnitt wird ein kurzer Überblick über die Entwicklung der automatischen Generalisierung gegeben (Unterabschnitt 2.3.1). Weil diese Arbeit im Kontext der Modellgeneralisierung entstanden ist, wird dabei speziell auf die automatisierte Modellgeneralisierung eingegangen (Unterabschnitt 2.3.2). Der dritte Unterabschnitt gibt eine aktuelle Klassifizierung von Ansätzen zur automatischen Generalisierung wieder. Die automatisierte Modellgeneralisierung wird dort in diese Klassifizierung eingeordnet. Dadurch wird das weitere Vorgehen in dieser Arbeit motiviert.

2.3.1 Überblick über die Entwicklung der Automation der Generalisierung

Die Entwicklung der Automatisierung der Generalisierung reicht zurück bis in die 60er und 70er Jahre des 20. Jahrhunderts. Bereits mit dem ersten Aufkommen der elektronischen Datenverarbeitung wurde die Automatisierung der Kartenerstellung sowie der kartographischen Generalisierung angedacht. Die ersten Algorithmen zur geometrischen Vereinfachung (Linienvereinfachung und -glättung) stammen aus dieser Zeit. Das vermutlich bekannteste Beispiel ist der Linienvereinfachungsalgorithmus nach Douglas und Peucker (1973). Der große Nachteil dieser in der damaligen Zeit entwickelten Ansätze ist die fehlende Berücksichtigung des Kontextes der generalisierten Objekte.

In den 80er und 90er Jahren entwickelte sich die automatische Generalisierung auf der Grundlage von *regel- und wissensbasierten Modellen*, und daraus resultierend zur Implementation von Expertensystemen, die die Regeln und das Wissen anwenden (Müller u. a. 1995). Diese Ansätze ermöglichten noch keine vollautomatische Generalisierung, so dass in kommerziellen Systemen zusätzlich interaktive, automationsunterstützende Werkzeuge angeboten wurden.

In den 90er Jahren wurden – als Folge der Misserfolge vorangegangener Ansätze – übergreifende Ansätze entwickelt, die das Fehlen des Kontextbezuges ausgleichen sollten. Prominentes Beispiel ist das AGENT-Projekt (Barrault u. a. 2001), das Geo-Objekten als „Agenten“ eine gewisse Autonomie verschafft, die darin besteht, mit benachbarten Objekten kommunizieren zu können. Durch eine Hierarchisierung der Agenten in Micro-, Meso- und Macro-Agenten wird der räumliche Kontext von Geo-Objekten in die Generalisierung mit einbezogen.

Die Menge der in der Vergangenheit entwickelten Ansätze zur automationsunterstützten und automatischen Generalisierung macht es schwierig, eine Auswahl zu treffen. Stattdessen sind an dieser Stelle Verweise auf Überblicksdarstellungen zur Geschichte der automatisierten Generalisierung angegeben, in denen sich eine Fülle weiterer Verweise finden lässt. Solche Überblicke finden sich bei Brassel und Weibel (1988), Nyerges (1991), Schoppmeyer und Beines (1991), Müller u. a. (1995), Schuurman (1999), Ruas (2002), Harrie und Weibel (2006).

Die Breite der Forschung brachte es mit sich, dass eine Fülle von Einzel- und Detaillösungen geschaffen wurden. Dies erkennend, entstand bei einigen Forschungsgruppen die Idee, die Ansätze durch eine gemeinsame Plattform kompatibel und austauschbar zu gestalten (Edwardes u. a. 2003). Eine erste Entwicklung in dieser Richtung stellen Burghardt u. a. (2005) mit der Konzeption eines Generalisierungs-Services vor.

Die Einschätzungen zur Zukunft der automatischen Generalisierung gehen heute davon aus, dass eine vollautomatische Generalisierung in naher Zukunft nicht zu erreichen ist. Eine pessimistische Einschätzung findet sich bei SGK (2002): „Eine vollständig automatisierte Generalisierung ist zur Zeit nicht absehbar. Eine interaktive Kontrolle und Nachbearbeitung ist notwendig, insbesondere im Hinblick auf das Zusammenspiel aller Kartenelemente nach dem Durchlaufen von Teilprozessen. Eine Weiterentwicklung der bestehenden Ansätze ist zweckmäßig und im Gange.“ Optimistischer sieht es Schuurman (1999): „... fully automated generalization is still a dream, albeit one within reach.“

2.3.2 Automatische Modellgeneralisierung

Im Jahr 2002 hat die Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (AdV) eine Entwicklung zur automatischen Modellgeneralisierung vom Basis-DLM zum DLM 50 angestoßen (Birth 2003). Das Projekt ist im Jahr 2004 abgeschlossen worden und das Produkt befindet sich in der produktiven Phase.

Bereits seit 1998 wurde im Auftrag des Amtes für Geowesen der Bundeswehr (AGeoBw, damals Amt für militärisches Geowesen, AMilGeo) an der Universität Bonn ein Prototyp zur automatisierten Modellgeneralisierung von TOPIS-Daten vom DLM 25 zum DLM 250 und vom DLM 250 zum DLM 1000 entwickelt und implementiert. Das Konzept für dieses Verfahren wurde von Schürer (2002) veröffentlicht. Weil die vorliegende Arbeit sich in

erheblichem Maße auf dieses Konzept stützt, wird an dieser Stelle detaillierter darauf eingegangen.

Das von Schürer entwickelte Verfahren zur automatisierten Modellgeneralisierung ist ein *umfassendes, automatisches, sequentielles* und *regelbasiertes* Verfahren zur Modellgeneralisierung:

1. **Umfassend:** Das Verfahren ist umfassend in dem Sinne, dass eine vollständige Generalisierung angestrebt und verwirklicht wurde. Der Generalisierungsprozess erhält als Eingabe ein DLM eines höheren Strukturierungsgrades und gibt als Ergebnis ein DLM eines niedrigeren Strukturierungsgrades aus. Damit ist dieses Verfahren eines der ersten, das eine Generalisierung von Anfang bis Ende umfasst.
2. **Automatisch:** Für den Fall einer korrekten und geeigneten Konfiguration vor Beginn der Generalisierung sind keine oder nur wenige manuellen Nacharbeiten zu leisten. Die Steuerung des Generalisierungsprozesses wird vor dem Beginn konfiguriert und geschieht ab dem Start des Prozesses automatisch.
3. **Sequentiell:** Schürer modelliert die Modellgeneralisierung als einen sequentiellen Prozess, der sukzessive auf den Ausgangsdaten abgearbeitet wird. Er unterteilt die Modellgeneralisierung dafür in die zwei Bereiche der semantischen und geometrischen Generalisierung, die in dieser Reihenfolge abgearbeitet werden. Die semantische Generalisierung umfasst dabei die in folgender Reihenfolge abzuarbeitenden Schritte der
 - *Klassifizierung und Typisierung:* Ganze Objektklassen werden ausgewählt beziehungsweise fortgelassen.
 - *Auswahl und Zusammenfassung:* DLM-Objekte werden an Hand von Selektionskriterien⁶ ausgewählt, fortgelassen oder mit anderen Objekten zusammengefasst.
 - *Klassifizierung und Typisierung sowie Auswahl und Zusammenfassung auf Attributebene:* Attributive Informationen von DLM-Objekten werden entsprechen der Anforderungen aus den Objektartenkatalogen generalisiert.

Die geometrische Generalisierung besteht aus zwei in folgender Reihenfolge abzuarbeitenden Schritten:

- *Geometriotypwechsel:* Die geometrischen und topologischen Veränderungen von Objekten, die ihre geometrische Ausprägung von einem höherwertigen Geometriotyp in einen niederwertigen Geometriotyp wechseln, werden berechnet. Dies sind die Wechsel von flächenförmig zu linienförmig, von flächenförmig zu punktförmig und von linienförmig zu punktförmig.
- *Anpassung an die Modellauflösung:* Geometrische Strukturen werden vereinfacht. Dies umfasst die Vereinfachung von Linien und von Flächengrenzen.

⁶Selektionskriterien sind innerhalb von ATKIS als Erfassungskriterien bezeichnet.

4. **Regelbasiert:** Die Generalisierung erfolgt in einzelnen Generalisierungsschritten nach vorher festgelegten Regeln. Die Regeln werden zum einen aus den Vorgaben der Objektartenkataloge gebildet, zum anderen sind sie abgeleitet aus kartographischem Wissen, das zusätzlich zu den Vorgaben in Regeln umformuliert wird.

Die Stärken des Verfahrens zur Modellgeneralisierung von Schürer liegen darin, dass es *vollständig* alle Schritte der Modellgeneralisierung umfasst, sowie *deterministisch* und damit *nachvollziehbar* die Generalisierung durchführt. Durch die Formulierung der Regeln ist die Generalisierung vorgegeben und wird im erwarteten Rahmen ablaufen.

Die Schwächen des Verfahrens liegen in seiner *Starrheit* und *Inflexibilität*. Generalisierungsschritte, die bereits ausgeführt wurden, können in einem späteren Schritt nicht mehr korrigiert oder verändert werden. Die wechselseitige Abhängigkeit von Generalisierungsvorgängen (siehe Abbildung 2.3 in Abschnitt 2.1.4) sind somit in dem Verfahren nicht modellierbar. Ein einfaches Beispiel kann dieses Problem verdeutlichen: Durch eine Vereinfachung der Geometrie ändert sich im Allgemeinen der Flächeninhalt flächenförmiger Objekte. Ist das Objekt vor der Geometrievereinfachung noch größer als eine Mindestgröße, so kann es nach der Vereinfachung kleiner sein. Dies würde eine erneute Ausführung des Generalisierungsvorganges der Auswahl oder der Zusammenfassung erfordern, dies kann aber wegen des sequentiellen Vorgehens nicht automatisch korrigiert werden.

Wegen seiner Vollständigkeit bietet sich das Verfahren von Schürer an, auch in Zukunft verwendet und auf andere Generalisierungsanwendungen übertragen zu werden. Die Schwächen des Verfahrens, bedingt durch Sequentialität und starre Regelbasierung, können dabei allerdings bei einigen Anwendungen stärker ins Gewicht fallen als bei der Modellgeneralisierung. Deshalb müssen hierfür Verbesserungen gesucht werden. Der nächste Unterabschnitt bietet eine Klassifizierung von Generalisierungsverfahren, in die die Modellgeneralisierung eingeordnet wird. Durch diese Klassifizierung werden Ansätze für Verbesserungsmöglichkeiten aufgezeigt.

2.3.3 Klassifizierung von Generalisierungsmodellen

Dieser Unterabschnitt bezieht sich auf eine Einteilung von Verfahren zur automatischen Generalisierung in drei *Modellierungsklassen* nach Harrie und Weibel (2006). Kriterium der Einteilung ist dabei die Art und Weise, wie die Steuerung des Generalisierungsprozesses in den Verfahren vonstatten geht. Für die Bezeichnungen der drei Modellierungsklassen haben sich noch keine einheitlichen Übersetzungen ins Deutsche durchgesetzt. Deshalb werden Übersetzungsvorschläge für die englischen Fachbegriffe eingeführt.

Harrie und Weibel (2006) unterscheiden zwischen den Modellierungsklassen der

1. Condition-action modelling (etwa: Vorbedingungs-Aktions-Modellierung)
2. Human interaction modelling (etwa: Mensch-Interaktions-Modellierung)
3. Constraint based modelling (etwa: Bedingungs-basierte Modellierung)

Diese drei Klassen werden im Folgenden kurz erläutert.

Vorbedingungs-Aktions-Modellierung

Diese Art der Generalisierungsmodellierung basiert auf Regeln, die eine von den Daten unabhängige Regelbasis bilden. Abhängig von diesen Regeln wird die automatische Generalisierung in zwei Phasen ausgeführt. In der ersten Phase werden die zu generalisierenden Daten analysiert und mit festgelegten Vorbedingungen verglichen. Sind diese Vorbedingungen erfüllt, so erfolgt die Durchführung einer bestimmten Aktion.

Beispielsweise bestimmt die Semantik eines Objektes zusammen mit seiner Größe die Art der Generalisierung: Abhängig von der Größe und seiner Bedeutung kann das Objekt fortgelassen oder mit benachbarten Objekten zusammengefasst werden, oder es wird zum Beispiel vergrößert.

Regelbasierte Systeme und Expertensysteme gehören zu dieser Klasse. Ebenso fällt das oben beschriebene Verfahren zur automatischen Modellgeneralisierung in diese Kategorie. Weitere Ansätze und Verfahren dieser Kategorie finden sich zum Beispiel bei Buttenfield und McMaster (1991).

Mensch-Interaktions-Modellierung

Hauptsächlich in den kommerziellen Programmen zur automatischen Generalisierung wird diese Modellierung verwendet. Weil vollautomatische Lösungen (noch) nicht existieren, und weil die Forschungsansätze bisher nicht zu den gewünschten Erfolgen in der Praxis führten, haben eine Reihe von kommerziellen GIS-Anbietern Teilschritte der Generalisierung automatisiert, die mit Hilfe von menschlicher Interaktion ausgeführt werden können.

Dies bedeutet, dass manuell der Workflow der Generalisierung vorgegeben wird, indem aus einer Menge von automatisierten Generalisierungsvorgängen ausgewählt werden kann. Die Generalisierung wird dann im Allgemeinen stückweise ausgeführt, die manuelle Auswahl der Generalisierungsvorgänge und ihrer Parameter und gegebenenfalls der Objekte wird für jeden Schritt wiederholt. In manchen Fällen können auch mehrere Schritte zu einer Stapelverarbeitung zusammengefasst und gemeinsam gestartet werden.

Diese Art der Generalisierung entspricht einer semi-automatischen oder automationsunterstützten Generalisierung, ist von einer automatischen Generalisierung aber noch weit entfernt. Die Ansätze sind jedoch praxisorientiert und präsentieren den derzeitigen Stand von kommerziell verfügbaren Programmen.

Bedingungsbasierte Modellierung

Die Klasse der bedingungsbasierten Generalisierungsmodellierungen gründet auf der Idee, dass die generalisierten Daten bestimmte Bedingungen erfüllen müssen, um eine dem Zweck der Generalisierung entsprechende Karte oder ein entsprechendes Datenmodell darstellen zu können. Im Prinzip ist der Ansatz ähnlich der regelbasierten Generalisierung (siehe oben), wobei dort der Fokus auf den *Prozessen* steht, während bei der bedingungsbasierten Generalisierung der Fokus auf den *Anforderungen an den Zustand des Ergebnisses* steht. Die regelbasierte Generalisierung gibt also ein *Verfahren* zur Generalisierung vor, während die bedingungsbasierte Modellierung das *Ziel* charakterisiert, jedoch den Weg dahin freistellt.

Harrie und Weibel (2006) beschreiben drei bereits realisierte Methoden einer bedingungs-basierten Modellierung, mittels *Agenten* (Barrault u. a. 2001), mittels *kombinierter Optimierung* (zum Beispiel Ware und Jones 1998) und mittels *kontinuierlicher Optimierung* (zum Beispiel Burghardt und Meier 1997).

Durch die Modellierung der Bedingungen wird keine Vorgabe für das Verfahren zum Erreichen der Bedingungserfüllung gemacht. Insofern können auch zukünftig entwickelte, neue Verfahren mit den aufgestellten Bedingungen arbeiten.

2.4 Methodik dieser Arbeit

„Computers are good at following instructions, but not at reading your mind.“
(Knuth 1991)

Das Leitmotiv dieser Arbeit beschreibt die Problematik der Automation der Generalisierung treffend: In der Kartographie existiert ein großes Wissen um die Generalisierung. Viele Kartographen sind tagtäglich mit der Generalisierung von Karten beschäftigt. Sie alle können im Allgemeinen gute von schlechten, richtige von falschen Generalisierungen sehr gut unterscheiden. Das Wissen um die Generalisierung wird in vielen Lehrbüchern ausführlich an Hand von Grundsätzen und Beispielen beschrieben (zum Beispiel Laubert u. a. 1982, SGK 2002). Doch sind diese Beschreibungen immer für die Durchführung der Generalisierung durch Menschen gemacht. Sie sind in der gegebenen Form für Rechner nicht anwendbar.

In den vorangegangenen Abschnitten dieses Kapitels wurde die Automation der Generalisierung und der aktuelle Wissensstand beschrieben. Sich darauf stützend geht dieser Abschnitt detaillierter als in der Einleitung auf die Methodik dieser Arbeit ein.

Ausgangspunkt der Arbeit ist das Verfahren zur Modellgeneralisierung nach Schürer (siehe Unterabschnitt 2.3.2), sowie eine im Rahmen dieser Arbeit entstandene Implementation dieses Verfahrens (Bobzien 2000, 2001). Der Erfolg des Projektes gab Anlass zu

der Hoffnung, dass sich dieser Erfolg auf andere kartographische Generalisierungsanwendungen außerhalb der Modellgeneralisierung übertragen ließe, mithin sich das Verfahren *verallgemeinern* ließe. Das Projekt war jedoch sehr speziell auf die Generalisierung von bestimmten DLM ausgerichtet, nämlich auf Daten des TOPIS-DLM⁷.

Es zeigte sich, dass zur Realisierung dieses Zieles der Verallgemeinerung mehrere Zwischenschritte nötig waren. Diese führten zu weiteren Ergebnissen, die nicht direkt mit dem oben gesetzten Ziel in Zusammenhang stehen. Dennoch konnte die Verallgemeinerung der Modellgeneralisierung als Ziel im Auge behalten werden, und die Zwischenergebnisse konnten für die Erlangung des Zieles Verwendung finden.

Insbesondere die *Regeln* sowie das *Lineare Verfahren* waren nicht zu verallgemeinern, da andere kartographische Generalisierungsanwendungen meistens differierende Anforderungen an die Generalisierung stellen. Deshalb wurde im Laufe der Verallgemeinerung von dem sequentiellen Verfahren Abstand genommen. Damit wurde ein Weg beschritten weg von dem mit der Modellgeneralisierung verknüpften Regelwerk, hin zu einer bedingungs-basierten Modellierung der Generalisierung, so wie sie in Abschnitt 2.3.3 beschrieben ist.

Als Basis wird eine Datenmodellierung benötigt, die für viele Arten von Daten zu verwenden ist, beziehungsweise eine Formalisierung der Anforderungen an die Daten, damit sie mit der entwickelten Generalisierungsmodellierung im Einklang stehen. Das Ergebnis ist ein Datenmodell, welches Elemente von bestehenden Datenmodellen enthält, aber mehrere Weiterentwicklungen beinhaltet. Kapitel 3 führt zu dieser Modellierung und erläutert sie detailliert.

Für die Umsetzung eines Generalisierungskonzeptes in ein Programm ist eine Formalisierung die Voraussetzung. Ein Programm P , das von einem Rechner ausgeführt werden soll, kann als eine Funktion f_P von der Menge der Eingabedaten E in die Menge der Ausgabedaten A , das heißt $f_P : E \rightarrow A$, verstanden werden (Claus und Schwill 2001). In Kapitel 4 erfolgt die Definition der Generalisierung als eine mathematische Funktion, konform mit und basierend auf dem Datenmodell aus Kapitel 3. Diese Funktion beschreibt die Generalisierung im Verständnis von Kartographen durch eine mathematische Formulierung. Dies ist der erste Schritt zu einer Formalisierung der Generalisierung.

Der umfangreichste Teil der Arbeit beschäftigt sich mit der Formulierung und Formalisierung von *Bedingungen* für die Generalisierungsfunktion und deren Ergebnis. Mit Kapitel 5 (Invariante Eigenschaften der Generalisierungsfunktion) wird der Schritt vollzogen von der regelbasierten Generalisierungsmodellierung hin zur bedingungsorientierten Generalisierungsmodellierung. Die formalisierten Bedingungen ergeben für ein vom Rechner ausgeführtes Programm (das ein Verfahren realisiert) die *Randbedingungen*, die erfüllt werden müssen. Diese Bedingungen können entweder *hart* (*kategorisch*) formuliert werden, so dass das Programm diese erfüllen muss, damit ein *korrektes* Ergebnis erzielt wird. Sie können jedoch auch *weich* formuliert werden, so dass das Programm auch ohne ihre Erfüllung zu einem korrekten Ergebnis kommt. Diese weichen Bedingungen bieten

⁷Zur Erläuterung des TOPIS-DLM siehe Abschnitt 3.1.6.

Kapitel 2 Generalisierung und ihre Automation

Qualitätsmaße, die die Güte einer Generalisierung (=Output des Programms) anzeigen. Da diese Arbeit sich auf der automatischen, regelbasierten Modellgeneralisierung gründet, liegt es nahe, dass aus den Regeln eher harte, kategorische Bedingungen entstehen. Deshalb sind die weichen Bedingungen unterrepräsentiert.

In Kapitel 6 werden die Bedingungen an der auf Basis des Verfahrens von Schürer (2002) erstellten Implementierung getestet und überprüft, inwieweit die Bedingungen mit der Modellgeneralisierung harmonieren. Danach werden diese Bedingungen auf andere Generalisierungsanwendungen übertragen (Abschnitte 6.2 und 6.3).

Es bleibt abschließend noch einmal darauf hinzuweisen, dass diese Arbeit *kein Verfahren* für eine automatisierte Generalisierung zur Verfügung stellt. Stattdessen wird – im Sinne der Einteilung von Harrie und Weibel (2006) – ein Rahmen für eine bedingungsorientierte Generalisierung entwickelt. Dies ist eine erste Stufe für eine allgemeine Realisierung der Generalisierung. Die in dieser Arbeit formalisierten Bedingungen können als Basis für zukünftige Forschungen, Verfahren und Implementierungen betrachtet werden.

Kapitel 3

Raumbezogene Datenmodelle

Für die Abbildung von Geodaten zur automatisierten Bearbeitung benötigt man ein *Datenmodell*, in dem festgelegt wird, wie Geodaten im Rechner modelliert werden. Ein Datenmodell ist eine formalisierte Beschreibung einer fachspezifischen Wirklichkeitsvorstellung, beziehungsweise eine Abstraktion ausgewählter Objekte der Realität (Hake und Grünreich 1994, S. 122). Es wird auf der konzeptionellen Ebene verwendet zur formalen Beschreibung von in einer Datenbank enthaltenen Daten und ihren Beziehungen untereinander (Claus und Schwill 2001). Im Kontext von Geodaten spricht man von *raumbezogenen Datenmodellen* (Hake u. a. 2002) oder *geographischen Datenmodellen* (Weber 1991). In der Literatur werden eine Vielzahl verschiedener Datenmodelle beschrieben und miteinander verglichen (Findeisen 1990, Weber 1991, Hake u. a. 2002).

Ziel dieses Kapitels ist die Entwicklung eines speziellen raumbezogenen Datenmodells, das für die Generalisierung geeignet ist. Die Modellierung erfolgt hier nur für Vektordaten (im Gegensatz zu Rasterdaten).¹ Im ersten Teil dieses Kapitels (Abschnitt 3.1) wird die Modellierung von Daten eines bestimmten Strukturierungsgrades behandelt, und als Spezialfall hiervon die Modellierung eines DLM eines Strukturierungsgrades. Verschiedene raumbezogene Datenmodelle werden vorgestellt, die aufeinander aufbauen. Dabei wird sukzessive das raumbezogene Datenmodell entwickelt, das für den restlichen Teil dieser Arbeit Verwendung findet. Im zweiten Teil dieses Kapitels erfolgt einer Erweiterung dieses Datenmodells auf die Modellierung mehrerer, miteinander verknüpfter DLM in verschiedenen Strukturierungsgraden. Man spricht hierbei von *Multirepräsentations-Datenmodellen*, welchen im Kontext von Generalisierung eine besondere Bedeutung zukommt. Diese werden in Abschnitt 3.2 diskutiert.

¹Die Verwendung von Vektordaten eignet sich sowohl für die räumliche als auch für die nicht-räumliche Analyse besser als Rasterdaten und erleichtert die weitere Nutzung und Analyse in großem Maße (siehe zum Beispiel Worboys 2004). Sie hat sich für Geodaten in vielen Bereichen als Standard durchgesetzt, weshalb hier und im Folgenden immer von Vektordaten die Rede ist, auch wenn dies nicht explizit erwähnt wird.

3.1 Modellierung raumbezogener Daten eines Strukturierungsgrades

Bei der Modellierung von raumbezogenen Daten wird üblicherweise jedes Realweltdatum höchstens einmal abgebildet. Dies dient der Redundanz- und damit der Fehlervermeidung. Außerdem werden alle Daten im Allgemeinen in einer bestimmten Auflösung (für die Darstellung in Karten) oder (allgemeiner) in einem bestimmten Strukturierungsgrad modelliert. Die Modellierung von Daten eines Strukturierungsgrades ist das Thema dieses Abschnittes. Erst wenn man Daten in verschiedenen Auflösungen darstellen möchte, wenn also ungeneralisierte und generalisierte Daten in einer Datenbasis vorhanden sind oder erzeugt werden sollen, dann benötigt man ein Datenmodell, in dem mehrere Strukturierungsgrade derselben Realweltdaten zugleich abgebildet werden können. Diese Modelle werden im zweiten Teil dieses Kapitels (Abschnitt 3.2) behandelt.

Hier wird zunächst in Abschnitt 3.1.1 der Begriff *Datenmodell* differenzierter betrachtet und seine Anwendung in den hier zusammentreffenden wissenschaftlichen Disziplinen – in der Kartographie und in der Informatik – betrachtet. Der darauf folgenden Abschnitt (3.1.2) beschreibt drei aufeinander aufbauende Grundtypen von raumbezogenen Datenmodellen. Das dritte Modell dieser Grundtypen, das hierarchische Datenmodell, wird in Abschnitt 3.1.3 um nicht-räumliche Informationen und in Abschnitt 3.1.4 um die rekursiven Aggregationen erweitert. Das resultierende Datenmodell ist für die folgenden Teile dieser Arbeit maßgeblich. Abschnitt 3.1.5 erläutert anhand eines Beispiels, wie Geodaten in dieses Modell abgebildet werden können. Abschnitt 3.1.6 schließlich vergleicht das hier dargestellte Modell mit der Modellierung in der amtlichen Kartographie.

3.1.1 Datenbankbezug räumlicher Datenmodelle

Bei den drei im folgenden Abschnitt 3.1.2 beschriebenen Datenmodellen handelt es sich um das *Spaghetti-Datenmodell*, das *topologische Datenmodell* und das *hierarchische Datenmodell*. Alle drei Datenmodelle gehen zurück auf Peucker und Chrisman (1975), wobei dort die Beschreibung und Visualisierung dieser Datenmodelle dem damaligen Forschungsstand entspricht und aus heutiger Sicht unübersichtlich erscheint. Die Ursache hierfür liegt darin begründet, dass für die Beschreibung und den Vergleich unterschiedlicher raumbezogene Datenmodelle der *Datenbankbezug* betrachtet werden muss. Im Kontext von Datenbanken und damit im Bereich der Informatik spricht man ebenfalls von Datenmodellen, versteht hierunter jedoch eine *Notation zur Beschreibung der Daten* zusammen mit einer Menge von Operationen zu ihrer Manipulation (Ullman 1988). Hake u. a. (2002, S. 155) bezeichnen diese Datenmodelle aus Sicht der Kartographie als *Modellbeschreibungssprachen*. Um Missverständnissen entgegen zu wirken, wird im Folgenden der Begriff *Modellbeschreibungssprache* für Datenmodelle mit Datenbankbezug benutzt. Beispiele für Modellbeschreibungssprachen sind das Entity-Relationship-Modell, das relationale Datenmodell, das Netzwerk-Datenmodell, das hierarchische Datenmodell (nicht

3.1 Modellierung raumbezogener Daten eines Strukturierungsgrades

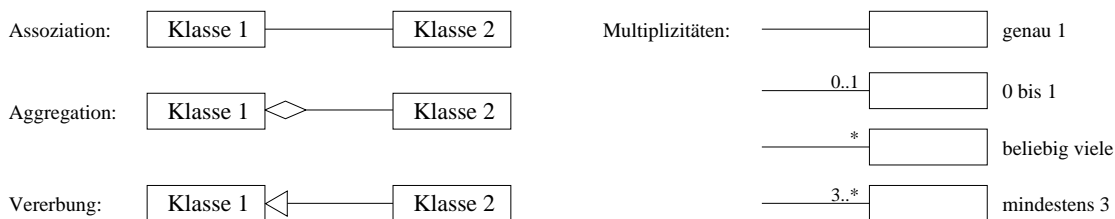


Abbildung 3.1: Wichtige Elemente der Unified Modelling Language (nach Balzert 2001).

zu verwechseln mit dem gleichnamigen raumbezogenen Datenmodell) sowie objektorientierte Datenmodelle (Ullman 1988).

Die raumbezogenen Datenmodelle und die Modellbeschreibungssprachen verhalten sich orthogonal zueinander. Jedes raumbezogene Datenmodell lässt sich also im Prinzip durch jede Modellbeschreibungssprache darstellen. Um einen Vergleich der unterschiedlichen raumbezogenen Datenmodelle zu ermöglichen, sollte man sich auf eine ausgewählte Modellbeschreibungssprache beschränken.

Die objektorientierten Datenmodelle sind die am weitesten entwickelten Datenmodelle und kommen dem intuitiven Verständnis von der Modellierung der Landschaft sehr nahe. Geodaten werden hier in *Klassen* eingeteilt, die Elemente der Klassen sind die *Objekte*. Klassen besitzen *Attribute* und *Methoden* (lokale Funktionen). Abhängigkeiten zwischen Klassen können mit *Assoziationen* beschrieben werden. Assoziationen können bestimmte *Multiplizitäten* besitzen. Des Weiteren können Objekte *aggregiert* (zusammengefasst) werden und Attribute und Methoden können über Klassen *vererbt* werden. Wegen der vielfältigen Möglichkeiten, die objektorientierte Datenmodelle bieten, sowie aus dem Grund, dass die objektorientierten Datenmodelle den *state-of-the-art* bilden, wird zur folgenden Beschreibung von raumbezogenen Datenmodellen ein objektorientiertes Datenmodell benutzt.

Die *Unified Modelling Language (UML)*, entwickelt von James Rumbaugh, Grady Booch und Ivar Jacobson (Booch u. a. 1999) ist entstanden aus verschiedenen früher entwickelten Modellbeschreibungssprachen (Rumbaugh u. a. 1999). Einen kurzen Überblick über wichtige Elemente von UML gibt Abbildung 3.1. UML bildet heute den De-facto-Standard für objektorientierte Modellbeschreibungssprachen. Im Folgenden wird UML als Modellbeschreibungssprache für raumbezogene Datenmodelle verwendet.

3.1.2 Grundtypen raumbezogener Datenmodelle

Die hier behandelten raumbezogenen Datenmodelle gehen zurück auf Peucker und Chrisman (1975). Deren Datenmodelle wurden in vielen Publikationen zitiert und weiterentwickelt. Sie können daher mit Recht als Grundtypen bezeichnet werden.

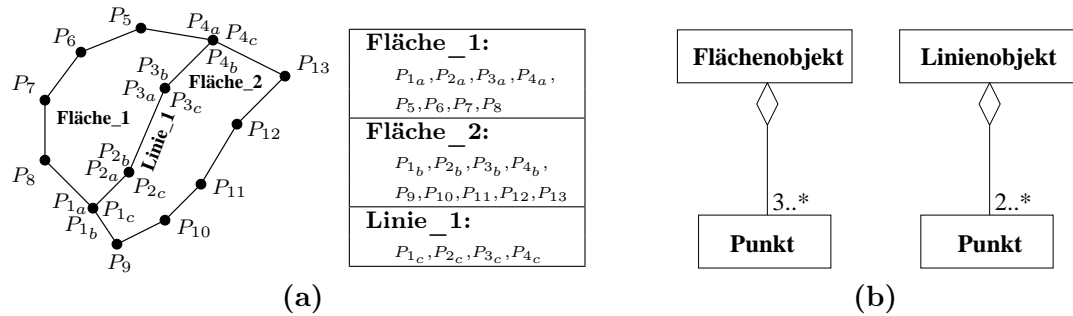


Abbildung 3.2: Beispiel für ein Spaghetti-Datenmodell: (a) Situation, (b) das dazugehörige UML-Diagramm.

Die hier im Folgenden beschriebenen drei Grundtypen von raumbezogenen Datenmodellen bauen aufeinander auf. Ihre Komplexität nimmt von Modell zu Modell zu.

Spaghetti-Datenmodell

Das Spaghetti-Datenmodell (Hake u. a. 2002), auch als Spaghetti-Datenstrukturen (Findeisen 1990, Worboys 2004) oder Spaghettistrukturen (Brassel 1983) bezeichnet, enthält zu jedem Geo-Objekt eine Liste von Koordinaten, die seine Geometrie beschreiben. Brassel (1983) unterscheidet noch zwischen Spaghettistrukturen als linienkonsistenten Strukturen, mit denen man nur linienförmige Objekte beschreiben kann, und umrisskonsistenten Strukturen, mit denen sich Flächen beschreiben lassen. Findeisen (1990) sowie Hake u. a. (2002) geben diese Unterscheidung auf und sprechen einheitlich von Spaghetti-Datenmodell beziehungsweise Spaghetti-Datenstrukturen, mit denen sich sowohl linien- als auch flächenförmige Objekte beschreiben lassen. Abbildung 3.2 (a) zeigt ein Beispiel für eine Situation, die mit dem Spaghetti-Datenmodell modelliert ist. Zu jedem flächenförmigen und linienförmigen Objekt wird eine Liste von Punkten gespeichert. Abbildung 3.2 (b) zeigt eine mögliche Modellierung in UML. Weitere Modellierungen findet man bei Findeisen (1990).

Das Spaghetti-Datenmodell erlaubt eine sehr einfache Speicherung von Geodaten. Diese Simplität ist allerdings mit vielen Nachteilen verbunden: Das Spaghetti-Datenmodell ist redundant, da viele Punkte mehrfach gespeichert werden müssen. Im Beispiel von Abbildung 3.2 (a) sind dies die Punkte, die zu beiden Flächengrenzen und zu der Linie gehören. Diese Redundanz macht das Modell sehr fehleranfällig für Splitterpolygone, Overshoots und Undershoots, die in Folge von Digitalisierung, Fortführung oder Weiterverarbeitung von Geodaten auftreten können. Ein weiterer Nachteil dieses Modells ist die schwierige Analyse von räumlichen Beziehungen. Um zum Beispiel die Nachbarschaft der beiden flächenförmigen Objekten aus Abbildung 3.2 (a) zu zeigen, müssen solange Punkte der beiden Punktlisten paarweise auf Gleichheit getestet werden, bis ein benachbartes Paar von Punkten gefunden wird, das in beiden Punktlisten enthalten ist.

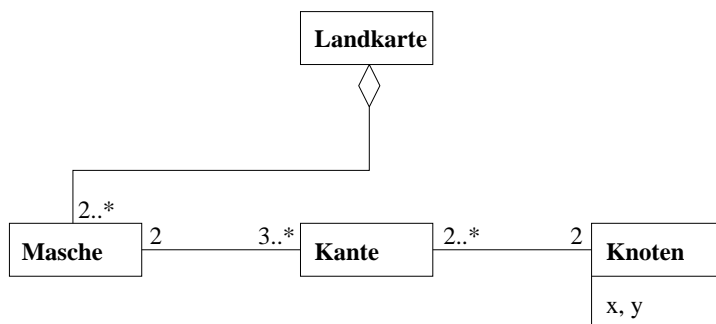


Abbildung 3.3: Topologisches Datenmodell (nach Plümer und Gröger 1997).

Das Spaghetti-Datenmodell wird trotz seiner Nachteile in der Praxis häufig eingesetzt. Bekannte Beispiele sind das *Shapefile* (ESRI 1998) und die *Geodatabase* (Zeiler 1999) der Firma ESRI, sowie *Simple Features* des Open GIS Consortium (Open GIS Consortium 1999).

Topologisches Datenmodell

Unter topologischen Datenmodellen versteht man Modelle, in denen topologische Beziehungen explizit modelliert werden. Die meisten Modelle stützen sich dabei auf die graphentheoretischen Elemente *Knoten*, *Kante* und *Masche*.² Die modellierten Beziehungen umfassen die *Inzidenz* (natürliche Verbundenheit) zwischen Elementen unterschiedlichen Typs, und die *Adjazenz* (Nachbarschaft) zwischen Elementen des gleichen Typs. Eine ausführliche Beschreibung der möglichen Beziehungen findet man bei Findeisen (1990, S. 54 f.). Mit topologisch modellierten Daten lassen sich topologische Relationen effizient analysieren.

Abbildung 3.3 zeigt ein mögliches topologisches Datenmodell, das minimale Landkartenmodell (Plümer und Gröger 1997). Eine Landkarte als Element aus der Graphentheorie ist eine Aggregation von Maschen. Die Maschen bilden eine flächendeckende Tessellation der Erdoberfläche unter Einbeziehung einer ausgezeichneten Masche *Außen*, die die Fläche außerhalb des betrachteten Gebietes abdeckt. Jede Masche ist mit mindestens drei Kanten verknüpft, während eine Kante zu genau zwei Maschen inzident ist. Eine Kante wiederum hat genau zwei Endknoten, jeder Knoten ist mit mindestens zwei Kanten verbunden. Nur an den Knoten werden Koordinaten gespeichert. Die Abbildung der Daten erfolgt somit redundanzfrei und ermöglicht eine effiziente Überprüfung der Konsistenz der Datenmenge (Plümer und Gröger 1997, Gröger 2000).

Das in Abbildung 3.3 dargestellte Landkartenmodell lässt keine Inseln und Löcher, außerdem keine Sackgassen und keine isolierte Knoten zu. Erweiterungen um einige dieser

²Zur Graphentheorie siehe zum Beispiel Wagner (1970), Harary (1972), Even (1979), Aigner (1984).

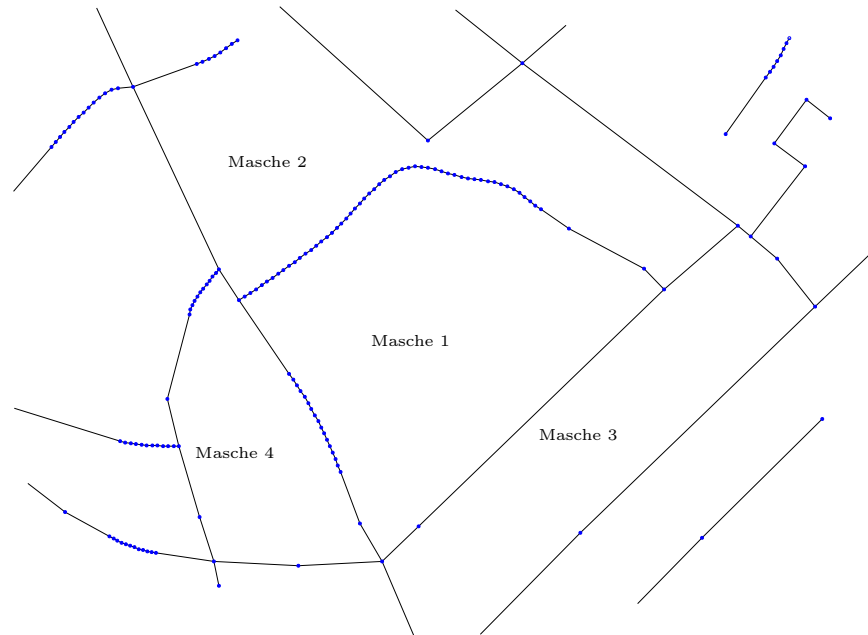


Abbildung 3.4: Modellierung von Kurven durch Polygonzüge (Beispieldaten aus dem TOPIS-DLM höchster Auflösung, abgeleitet aus DYNAMO-Daten).

Eigenschaften findet man in Veröffentlichungen zu verschachtelten Landkarten in (Plümer und Gröger 1996, 1997, Gröger 2000). Diese Problematik wird in Abschnitt 3.1.5 ausführlicher behandelt.

Das topologische Datenmodell besitzt durch seine Redundanzfreiheit einige Vorzüge im Vergleich zum Spaghetti-Datenmodell, was besonders bei raumbezogenen Analysen zum Tragen kommt. Am Beispiel aus dem vorigen Abschnitt kann dies deutlich gemacht werden: Das Beispiel aus Abbildung 3.2 (a) würde mit dem topologischen Datenmodell durch zwei Maschen, 14 Kanten und 13 Knoten modelliert. Um die Nachbarschaft der Flächenobjekte 1 und 2 zu zeigen, muss jetzt nur überprüft werden, ob die Schnittmenge der beiden Kantenmengen, die mit den beiden Maschen verknüpft sind, nicht leer ist.

Ein entscheidender Nachteil von topologischen Datenmodellen ist jedoch der Rechenaufwand, der bei topologischen Analysen auftritt, wenn die Zahl der Knoten und Kanten im Vergleich zur Zahl der Maschen sehr hoch ist – anders ausgedrückt: wenn digitalisierte Liniendaten eine hohe Stützpunktdichte haben. Dies ist meistens dann der Fall, wenn ein GIS keine Möglichkeit hat, Kurven durch Splines zu modellieren, sondern Kurven durch Polygonzüge annähert (siehe Abbildung 3.4). Ein Beispiel hierfür ist das kommerzielle GIS *DYNAMO* der Firma *Intergraph*.

Das folgende Beispiel soll die Problematik topologischer Analysen mit im topologischen Datenmodell modellierten Daten verdeutlichen (Abbildung 3.4): Ziel ist es, alle Nachbarmaschen von Masche 1 zu finden (Maschen-Adjazenz). Ein möglicher Algorithmus sucht

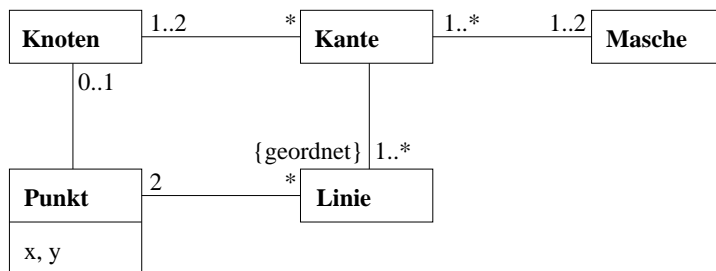


Abbildung 3.5: Hierarchisches Datenmodell (nach Bobzien 2000).

zunächst alle zu Masche 1 inzidenten Kanten. Diese werden über die Relation zwischen der Masche und den Kanten gefunden. Im Beispiel wird Masche 1 durch 77 Kanten begrenzt. Die hohe Anzahl der begrenzenden Kanten ergibt sich aus der Tatsache, dass jede Kante ein gerades Liniensegment darstellt. Für jede dieser Kanten muss dann die komplementäre Masche gefunden werden. Dazu werden zu jeder Kante die inzidenten Maschen gesucht, und diejenige Masche, welche nicht Masche 1 ist, ist dann eine zu Masche 1 adjazente Masche. Im Beispiel müssen also 2×77 Inzidenzen überprüft werden, um die drei zu Masche 1 adjazenten Maschen zu finden (Maschen 2, 3 und 4). Diese Art von Suche ist hochgradig redundant und wird durch das im nachfolgenden Abschnitt vorgestellte Modell verbessert. Ähnliche Probleme wie bei der Suche von Nachbarmaschen ergeben sich bei der Pfadsuche.

Hierarchisches Datenmodell

Der beschriebene Nachteil des topologischen Datenmodells wird durch das hierarchische Datenmodell beseitigt (Hake u. a. 2002). Ursprünglich beschrieben es Peucker und Chrisman (1975). Sie bezeichnen Kanten als *Ketten*, die aus beliebig vielen Punkten bestehen können. Außerdem unterschieden sie zwischen Punkten, die topologische Knoten bilden und damit zu mehreren Kanten gehören können, und Punkten, die innerhalb einer Kante bestehen. Letztere werden im Allgemeinen als *Stützpunkte* bezeichnet (zum Beispiel Hake u. a. 2002).

Eine objektorientierte Modellierung des hierarchischen Datenmodells ist in Abbildung 3.5 dargestellt. Sie ist ein Extrakt aus dem generalisierungsspezifischen Datenmodell aus (Bobzien 2000), dessen weitergehende Beschreibung in den Abschnitten 3.1.3 und 3.1.4 erfolgt. Das Modell ist so konzipiert, dass man die Daten in unterschiedliche *Ebenen* unterteilen kann. Abbildung 3.5 zeigt die topologische und die geometrische Ebene. Die topologische Ebene besteht aus Masche, Kante und Knoten, so wie sie auch im topologischen Datenmodell enthalten sind. Auch hier bilden die Maschen eine Tesselation der Erdoberfläche. Im Unterschied zum topologischen Modell sind jedoch im hierarchischen Modell die Koordinaten nicht als Attribute an den Knoten modelliert, sondern sie sind

Attribute von Punkten, welche zusammen mit den Linien die geometrische Ebene bilden.³ Linien enthalten die Verbindungsinformation zwischen den Punkten. Die Kardinalitäten sind so gewählt, dass auch komplexere Sachverhalte (wie Inseln, Sackgassen und isolierte Knoten) modelliert werden können, was jedoch gleichzeitig zu einer Erschwerung der Integritätsprüfung im Sinne von Gröger (2000) führt.

Die Kardinalitäten auf der topologischen Ebene werden nun detaillierter erläutert: Eine Masche besteht aus mindestens einer Kante. Besteht sie aus genau einer Kante, so bildet die Kante einen *Ring*. Eine Kante kann zu einer oder zwei Maschen inzident sein. Die Inzidenz zu zwei Maschen ist der übliche Fall. Inzidenz zu nur einer Masche bedeutet, dass die Kante zu einer Sackgasse gehört oder isoliert in einer Masche liegt. Eine Kante ist zu ein oder zwei Knoten inzident (Endknoten der Kante). Hat die Kante nur einen Endknoten, so bildet sie einen Ring. Ein Knoten kann zu beliebig vielen Kanten inzident sein. Ist er zu keiner Kante inzident, ist er ein *isolierter Knoten*, ist er zu nur einer Kante inzident, so ist die Kante eine Sackgasse oder sie bildet einen isolierten Ring und damit ein Loch oder eine Insel.

Auf der geometrischen Ebene gibt es Punkte und Linien. Punkte haben Koordinaten als Attribute und können Endpunkte zu beliebig vielen Linien sein (analog zu der Beziehung zwischen Knoten und Kanten). Keine zwei Punkte haben dieselben Koordinaten. Eine Linie ist ein gerades Segment, das genau zwei Endpunkte haben muss. Eine Ringbildung ist für eine Linie somit nicht möglich.

Die Verknüpfung zwischen Topologie und Geometrie findet zwischen den punktförmigen und den linienförmigen Elementen statt. Jeder Knoten ist mit genau einem Punkt verbunden. Ein Punkt kann die Geometrie zu einem oder zu keinem Knoten bilden. Im letzteren Fall handelt es sich bei dem Punkt um einen Stützpunkt. Bei den linienförmigen Elementen bestehen folgende Beziehungen: Jede Kante besteht aus einer geordneten Liste von mindestens einer Linie. Die Ordnung ergibt sich zwar implizit aus den Verknüpfungen zwischen Linien und Punkten, ist aber nur aufwändig durch Sortieren zu berechnen.⁴ Eine Linie muss zu genau einer Kante gehören.

Mit dem hierarchischen Datenmodell reduziert sich der Aufwand bei der Suche nach topologisch verbundenen Objekten im Gegensatz zum topologischen Datenmodell erheblich, wenn viele Stützpunkte verwendet werden. Betrachtet man das Beispiel aus dem letzten Abschnitt (Abbildung 3.4), in dem die Nachbarmaschen von Masche 1 gesucht werden sollen, findet man, dass sich der Aufwand erheblich reduziert: Wenn das Beispiel durch das hierarchische Datenmodell abgebildet wird, ist Masche 1 jetzt nicht mehr durch 77 Kanten begrenzt, sondern nur noch durch drei Kanten. Jede Kante wird aus mehreren Linien gebildet, die Kante zwischen Masche 1 und Masche 2 aus 56 Linien, die Kante zwischen Masche 1 und Masche 3 aus zwei Linien und die Kante zwischen Masche 1 und

³Bei Gröger (2000) werden die Koordinaten ebenfalls in einer eigenen Klasse „Punkt“ modelliert. Jedoch ist dort jedem Punkt genau ein Knoten zugeordnet.

⁴Aufwandsabschätzung für das Sortieren: $O(n \log n)$ mit n als Anzahl der Linien (zum Beispiel Knuth 1973).

Masche 4 aus 19 Linien. Da jetzt nicht mehr jede Linie zur Adjazenzanalyse herangezogen werden muss, sondern nur noch die drei Kanten, reduziert sich der Aufwand des im vorigen Abschnitts skizzierten Algorithmus von 154 Inzidenzanalysen auf sechs Inzidenzanalysen. Zur genauen Abgrenzung von Kanten und Linien siehe Abschnitt 3.1.5.

3.1.3 Nicht-räumliche Informationen: Thematik

Ausschließlich mit der Geometrie und Topologie lassen sich Geo-Objekte nicht modellieren. Es fehlen Informationen, die die Art der Objekte und die Eigenschaften der Objekte beschreiben. Im Kontext von topographischen DLM und speziell im Kontext der amtlichen Kartographie in Deutschland spricht man von *Objektarten* und deren *Attributen* (AdV 1989). Die Objektarten bilden Klassen innerhalb der Menge aller Geo-Objekte, die für ein DLM relevant sind. Das bedeutet, jedes erfasste Geo-Objekt wird genau einer Klasse zugeordnet. Übliche Klassen in topographischen DLM sind zum Beispiel Wald, bebautes Gebiet oder Straße. Die Attribute beschreiben ein Geo-Objekt genauer. Die Objektarten und Attribute bilden die *nicht-räumlichen Informationen* zu einem Geo-Objekt, auch *Semantik* eines Objektes genannt (AdV 1989). Die Bezeichnung Semantik für die nicht-räumlichen Informationen ist insofern problematisch, als natürlich auch die Geometrie und die Topologie eines Objektes eine Semantik enthält, nämlich wo und in welchem Verhältnis zu anderen Objekten dieses Objekt sich befindet. In der amtlichen Kartographie in Deutschland gibt es eine Unterteilung in Semantik, Semantik/Topologie und Geometrie (AdV 1989). Die Problematik dieser Einteilung wird in Abschnitt 3.1.6 untersucht. In dieser Arbeit wird der Begriff *Thematik* bevorzugt, wie er auch von Nyerges (1991), Weber (1991) und Bartelme (2005) in diesem Kontext verwendet wird.

Geo-Objekte haben im Allgemeinen genau eine geometrische Ausprägung für ihre zweidimensionale Darstellung: Sie werden entweder punktförmig (zum Beispiel Sendemast oder Festpunkt), linienförmig (zum Beispiel Straße oder Eisenbahnlinie) oder flächenförmig (zum Beispiel Wald oder Wohngebiet) modelliert. Diese geometrische Ausprägung wird auch *Geometrietyt* (Schoppmeyer und Heisser 1995, Schürer 2002) oder *Objekttyp* (AdV 1989) genannt. In seltenen Fällen kann man sich die Modellierung eines Objekts durch verschiedene geometrische Ausprägungen vorstellen (zum Beispiel kann ein Fluss an breiten Stellen flächenförmig, an schmalen Stellen linienförmig modelliert werden); für eine systematische Erfassung und Modellierung wird im Allgemeinen solch ein Geo-Objekt in mehrere Geo-Objekte mit je einer Ausprägung unterteilt. Sind Geo-Objekte mit einer eindeutigen Ausprägung modelliert, so können sie mit den Topologie-Elementen verknüpft werden, wie im Folgenden erläutert wird.

In Abbildung 3.6 ist das Hierarchische Datenmodell aus Abbildung 3.5 um die thematische Ebene ergänzt. Ein *punktförmiges Objekt* ist mit genau einem Knoten verbunden, umgekehrt kann ein Knoten mit beliebig vielen punktförmigen Objekten verbunden sein. In den meisten Fällen ist ein Knoten jedoch mit keinem Objekt verbunden, weil die meisten Knoten nur die topologische Funktion eines Kreuzungsknotens haben. Ein *linienförmiges Objekt* ist mit mindestens einer Kante verknüpft. Umgekehrt kann eine Kante

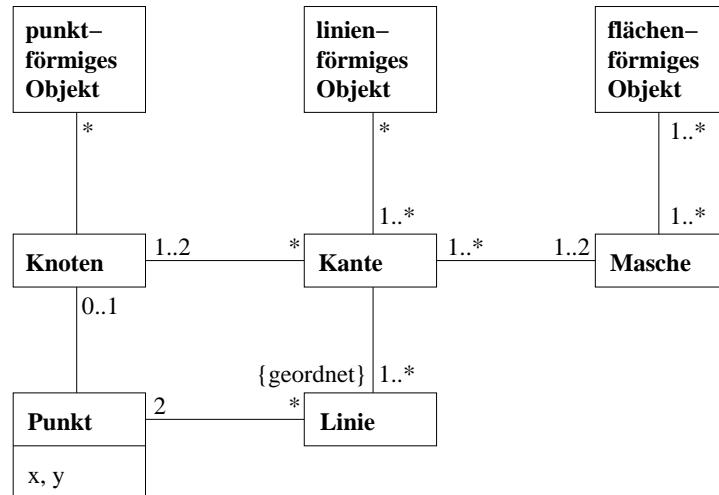


Abbildung 3.6: Hierarchisches Datenmodell erweitert um die thematische Ebene (nach Bobzien 2000).

mit beliebig vielen linienförmigen Objekten verbunden sein. Ist eine Kante mit keinem linienförmigen Objekt verbunden, so repräsentiert sie die topologische Funktion einer Grenze von Maschen. Ist sie mit mehr als einem linienförmigen Objekt verbunden, so überlagern sich diese Objekte. Zum Beispiel kann eine Grenze in der Mitte eines Flusses verlaufen. Ein *flächenförmiges Objekt* besteht aus mindestens einer Masche. Umgekehrt ist eine Masche mit mindestens einem Objekt verbunden. Würde eine Masche mit keinem Objekt verknüpft sein, so würde an dieser Stelle eine Erfassungslücke bestehen. Wenn keine Erfassungslücke besteht, so erfüllen die Daten die Bedingung der *Flächendeckung*. Die Flächendeckung wird zum Beispiel im Bereich der amtlichen Kartographie (AdV 1989) gefordert. Auf die Bedingung der Flächendeckung wird im Abschnitt 5.5.2 detailliert eingegangen. Ist eine Masche mit mehr als einem flächenförmigen Objekt verbunden, so findet eine Überlagerung statt. Schürer (2002, Kap. 3.3) unterteilt flächenförmige Objekte in vier Ausgestaltungs-klassen und untersucht die erlaubten Überlagerungen.

Die Objekte haben in der thematischen Ebene keine Beziehungen zueinander. Solche Beziehungen werden manchmal gefordert, können aber meistens aus der Topologie unter Berücksichtigung der Thematik abgeleitet werden. Zum Beispiel könnte es gewünscht sein, die Verbindung von Hochspannungsmasten zu Hochspannungsleitungen explizit zu modellieren. Diese Verbindung kann aber auch durch Analyse der Beziehungen auf der topologischen Ebene abgeleitet werden. In besonderen Fällen kann eine explizite Modellierung von Beziehungen zwischen Objekten in der thematischen Ebene notwendig sein, wie das Beispiel von Über- und Unterführungen zeigt: Wenn eine Straße und ein Schienenweg sich nicht niveaugleich kreuzen, ist es zweckmäßig, zu speichern, welches Objekt oberhalb des anderen verläuft. Diese Informationen lassen sich nicht aus der hier verwendeten zweidimensionalen Topologie ableiten.

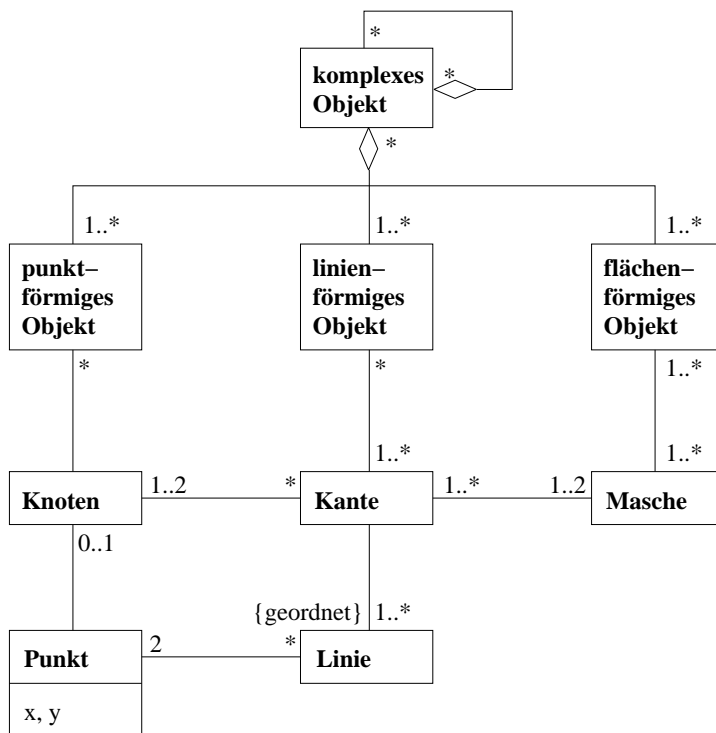


Abbildung 3.7: Erweitertes hierarchisches Datenmodell.

3.1.4 Komplexe Objekte

Für eine Reihe von Anwendungen sind die bisher vorgestellten Modellierungen zu starr. Um komplizierte Situationen abbilden zu können, möchte man oftmals hierarchische Beziehungen zwischen Landschaftsobjekten modellieren. Zum Beispiel kann eine Autobahn aus einem Straßenkörper und zwei Fahrbahnen bestehend modelliert werden, wie dies in AdV (1989) der Fall ist. Viele andere Aggregationsbeziehungen sind möglich: Industrieanlagen bestehen aus Straßen und Gebäuden; Brücken bestehen aus Fahrbahnen, Pfeilern und Geländern und so fort. Auch für die automatische Durchführung einer Generalisierung ist es oftmals nötig, Geo-Objekte in größeren Zusammenhängen zu betrachten (Bobzien 2000, Schürer 2002). Zum Beispiel kann eine Straße oder eine Bahnstrecke durch mehrere linienförmige Objekte modelliert sein. Für die Generalisierung der Geometrie mag es aber notwendig sein, die Straße oder die Bahnstrecke als Ganzes zu betrachten. Ähnliches gilt für flächenförmige Objekte: Um benachbarte flächenförmige Objekte zusammenfassen zu können, sollten Gebiete abgegrenzt werden, in denen primär zusammengefasst wird. Selbst für punktförmige Objekte ist die Notwendigkeit einer Zusammenfassung denkbar: Bäume zum Beispiel können in einem Maßstabsbereich als einzelne punktförmige Objekte modelliert sein, während sie bei der Generalisierung zu einem flächenförmigen Objekt wie Plantage oder Wald zusammengefasst werden.

In Abbildung 3.7 ist das Modell aus Abbildung 3.6 nochmals dargestellt, erweitert um *komplexe Objekte*⁵. Ein komplexes Objekt ist zusammengesetzt aus beliebig vielen punkt-, linien- oder flächenförmigen Objekten, dargestellt durch die Aggregationsbeziehung. Ein komplexes Objekt kann selber auch aus einem oder mehreren komplexen Objekten bestehen, wie durch die rekursive Aggregationsbeziehung zu sich selbst dargestellt ist. Das hier vorgestellte Modell wird im Folgenden *erweitertes hierarchisches Datenmodell* genannt. Es ist die Grundlage für alle weiteren Betrachtungen über Modellierung und Formalisierung der Generalisierung.

Abbildung 3.7 stellt eine sehr allgemeine Form der Modellierung dar. In den meisten Fällen wird die Aggregation des komplexen Objektes von weiteren Regeln bestimmt. In der ATKIS-Modellierung zum Beispiel besteht ein Damm aus genau einer Dammkrone und aus einer Böschung oder einem Kliff (AdV 1989). In den oben erwähnten Beispielen im Zusammenhang mit der Generalisierung wird die Aggregation von komplexen Objekten beschränkt auf die Aggregation aus Objekten eines bestimmten Geometrietyps, so wird eine komplexe Straße nur aus linienförmigen Objekten bestehen, ein Generalisierungsblock aus flächenförmigen Objekten und eine Baumgruppe aus punktförmigen Objekten.

In (Bobzien 2000) wird zwischen komplexen und generalisierungsspezifischen Objekten unterschieden; Schürer (2002) unterscheidet zwischen komplexen Objekten und komplexen Generalisierungsobjekten⁶. Beide Modelle können in das erweiterte hierarchische Datenmodell integriert werden. Selbiges gilt für die in der amtlichen Kartographie genutzten Modelle (ATKIS/TOPIS, siehe Abschnitt 3.1.6).

3.1.5 Modellierungsbeispiele

Die in den vorigen Abschnitten unternommene Unterteilung der Modellierung in Geometrie, Topologie und Thematik (oder Semantik) wird in diesem Abschnitt an einer beispielhaften Modellierung (Abbildung 3.8) näher erläutert. Insbesondere die Abgrenzung von topologischen Elementen (hier: Knoten und Kanten) zu geometrischen Elementen (Punkte und Linien) sowie die Abgrenzung von modellierten Landschaftsobjekten zur Topologie soll hier verdeutlicht werden. Des Weiteren werden topologische Sonderfälle wie einzelstehende Knoten, Sackgassen und Löcher behandelt.

In Abbildung 3.8 (a) ist eine einfache topographische Situation dargestellt. Man erkennt einen Fluss und einen kleinen See (beide in blau), Grünland beiderseits des Flusses (grün), bebautes Gebiet im rechten oberen Viertel (pink), des Weiteren Straßen (schwarz), von denen eine über eine Brücke (rot) über den Fluss führt. Links sind eine Seilbahn durch

⁵Die Bezeichnung komplexes Objekt ist angelehnt an die Terminologie der amtlichen Kartographie. Ebenso treffend wären die Begriffe zusammengesetztes oder aggregiertes Objekt.

⁶Generalisierungsspezifische Objekte und komplexe Generalisierungsobjekte stellen in beiden Quellen spezielle, zur Durchführung der Generalisierung benötigte Aggregationen von Geo-Objekten dar.

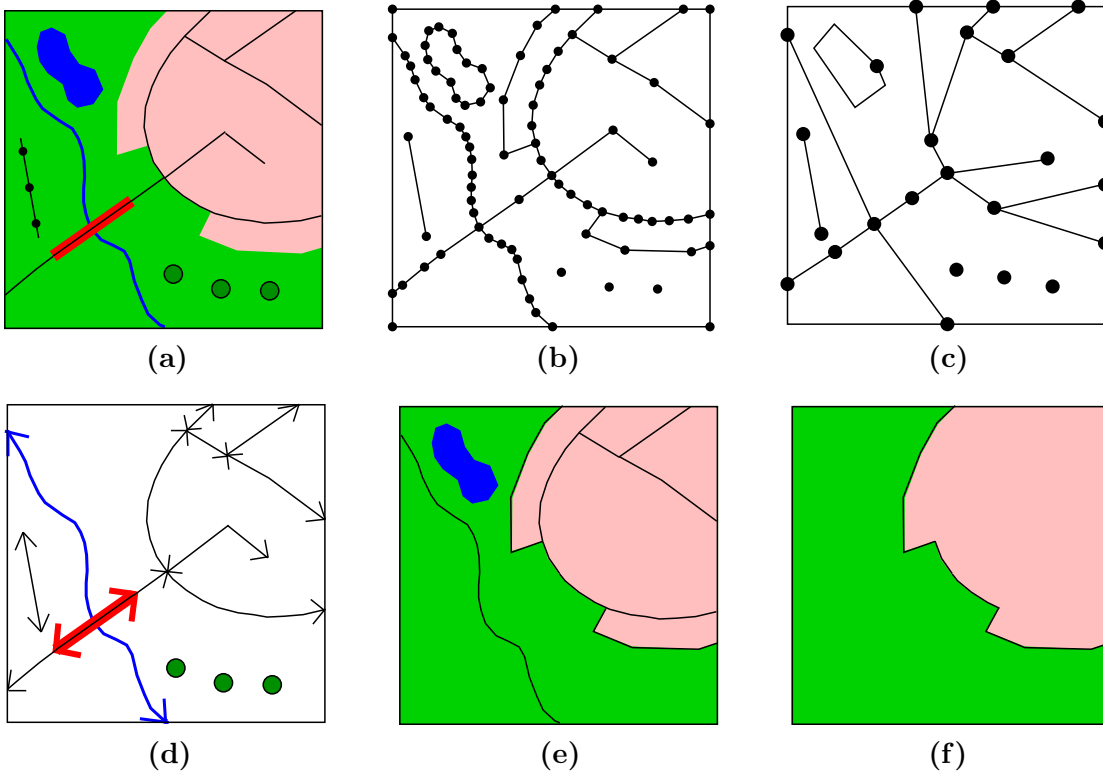


Abbildung 3.8: *Beispielhafte Modellierung mittels erweitertem hierarchischen Datenmodell: (a) Situation, (b) geometrische Ebene, (c) topologische Ebene, (d) thematische Ebene (Linien- und Punktobjekte), (e) thematische Ebene (Flächenobjekte), (f) komplexe Objekte (flächenförmig).*

eine linienförmige Signatur und im rechten unteren Viertel drei Bäume durch kreisförmige Signaturen dargestellt.

Abbildung 3.8 (b) stellt die geometrischen Elemente Punkte und Linien zu dieser Situation dar. Jede Linie wird durch ein gerades Teilstück gebildet und ist durch genau zwei Punkte begrenzt. Ein Punkt hingegen hat in diesem Beispiel zwischen null und vier abgehende Linien. Die Baumsignaturen sind durch alleinstehende Punkte abgebildet, die Sackgasse im Wohngebiet sowie die Seilbahn enden jeweils in Punkten mit nur einer verbundenen Linie. Die meisten Punkte haben zwei verbundene Linien, einige Punkte haben drei oder vier verbundene Linien. An der Kardinalität der verbundenen Linien kann man jedoch nicht immer ablesen, ob es sich um Stützpunkte handelt. Dies wird erst deutlich, wenn man die topologische Ebene betrachtet:

Die topologische Ebene mit den Elementen Knoten, Kanten und Maschen zu der Beispielsituation ist in Abbildung 3.8 (c) dargestellt. Jeder ausgefüllte Kreis stellt einen Knoten dar, jeder Linienabschnitt eine Kante und jede abgegrenzte Fläche eine Masche. Die Maschen sind überlappungsfrei und flächendeckend. Sie werden von mindestens einer

Kante begrenzt, im Allgemeinen jedoch von mehr als einer Kante. Im Beispiel wird die Masche zum See nur von einer Kante begrenzt. Diese Masche bildet gleichzeitig ein *Loch* zu der den See umgebenden Masche. Kanten werden im Allgemeinen von zwei Knoten begrenzt. Der Umring des Sees wird jedoch nur von einem Knoten begrenzt, diese Kante bildet einen Ring. Ebenso begrenzen Kanten im Allgemeinen immer genau zwei Maschen. Ausnahmen sind Sackgassen und isolierte Kanten, im Beispiel die Straße in der Mitte des bebauten Gebietes und die Seilbahn. Knoten haben beliebig viele verbundene Kanten. Im Beispiel gibt es drei isolierte Knoten (die Bäume) und vier Knoten mit einer verbundenen Kante (die Sackgasse im Wohngebiet, die Seilbahn und der Knoten zu der den See umgebenden Kante). Die meisten Knoten haben mehr als drei abgehende Kanten, dies sind klassische Kreuzungsknoten. Eine Ausnahme bilden die beiden Knoten, die die Enden der Brücke markieren: Dies sind die einzigen Knoten, die genau zwei verbundene Kanten haben. Im Allgemeinfall würden diese Knoten nicht modelliert, weil die zugehörigen Punkte Stützpunkte wären. Hier jedoch ist die Modellierung dieser Knoten nötig, weil die Thematik eine Trennung zwischen den Kanten fordert, die mit der Brücke *und* der Straße verknüpft sind, und den Kanten, die nur mit der Straße verknüpft sind.

In Abbildung 3.8 (d) sind die linien- und punktförmigen Objekte aus der thematischen Ebene dargestellt. Punktförmige Objekte sind die Bäume, linienförmige Objekte Straßen, der Fluss, die Brücke und die Seilbahn. Die Abgrenzung der Objekte ist durch Pfeilspitzen an deren jeweiligen Enden angedeutet. Im Beispiel ist die Modellierung so gewählt, dass neue Straßenobjekte durch jede Straßenkreuzung gebildet werden. Man beachte, dass die Straße, die über die Brücke führt, nicht am Brückenanfang oder -ende und auch nicht an der Kreuzung mit dem Fluss in zwei oder mehrere Objekte aufgetrennt wird. Ebenso wird der Fluss als ein Objekt modelliert, obwohl er topologisch von der Brücke beziehungsweise der Straße getrennt wird.⁷ Die Auftrennung des Flusses, der Straße und der Brücke findet hingegen auf der topologischen Ebene statt, so dass die Straße aus vier Kanten, der Fluss und die Brücke jeweils aus zwei Kanten bestehen. Hier findet auch eine Überlagerung statt: Zwei Kanten sind mit denselben zwei linienförmigen Objekten verbunden (Straße und Brücke).

Die flächenförmigen Objekte aus der thematischen Ebene sind in Abbildung 3.8 (e) dargestellt. In dieser Modellierung ist das bebauten Gebiet durch fünf Objekte abgebildet, es wurde also durch die Straßen aufgetrennt. Hingegen wurde das Grünland nur wegen des Flusses in mehrere Objekte abgebildet, nicht jedoch wegen der Straße.⁸ In der Abbildung kann man nicht erkennen, ob flächenförmige Objekte sich überlagern. Im Beispiel kann der Teich entweder als ein Grundobjekt ein Loch in dem Grünland bilden, oder der Teich kann als Ausgestaltungsobjekt auf dem Grünland liegen, wobei die verknüpfte Masche dann mit beiden Objekten assoziiert wird. Zu der Problematik von Überlagerungen siehe Schürer (2002, Kap. 3.3).

In Abbildung 3.8 (f) schließlich wird eine mögliche Bildung von komplexen (flächenförmigen) Objekten dargestellt. Das eine komplexes Objekt besteht aus den bebauten

⁷Die hier angewandte Modellierung entspricht der in ATKIS (AdV 1989) verwendeten Modellierung.

⁸Diese Modellierung entspricht *nicht* der in ATKIS (AdV 1989) verwendeten Modellierung.

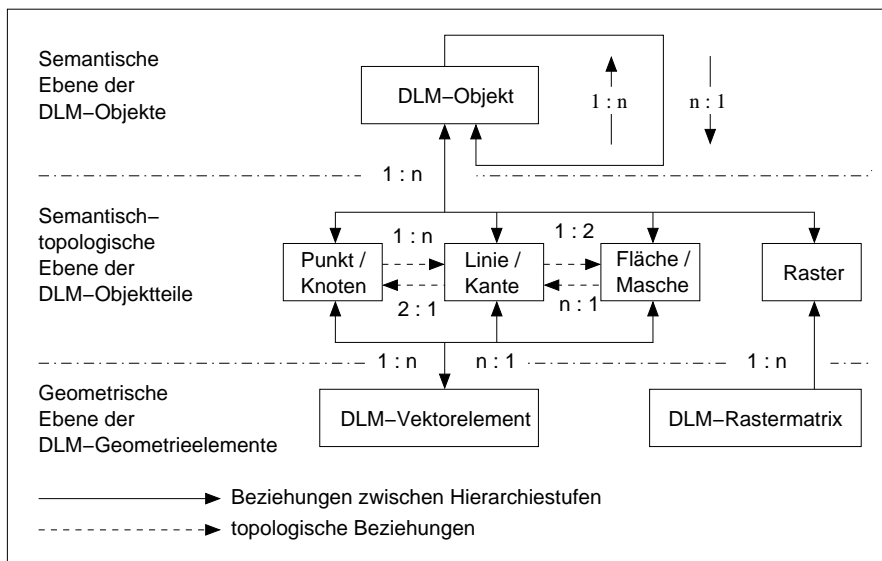


Abbildung 3.9: Struktur des ATKIS-Datenmodells (nach AdV 1989, Teil C).

Gebieten, das andere aus den beiden Grünflächen und dem See. Das Kriterium für die Bildung komplexer Objekte war hier offensichtlich die Grenze zwischen bebautem und unbebautem Gebiet. Die hier erfolgte Aggregation kann zum Beispiel für die Flächenzusammenfassung während einer Generalisierung verwendet werden.

3.1.6 Raumbezogene Datenmodelle in der amtlichen topographischen Kartographie

In diesem Abschnitt wird die Modellierung topographischer Daten in der amtlichen Kartographie in Deutschland und ihr Verhältnis zu dem hier entwickelten erweiterten hierarchischen Datenmodell dargestellt. Von der Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (AdV) wurde das Projekt ATKIS (Amtliches Topographisch-Kartographisches Informationssystem) entwickelt. ATKIS enthält einen bundeseinheitlichen digitalen topographischen Datenbestand über die Topographie der Erdoberfläche (AdV 1989).

Parallel dazu entwickelte das Amt für Militärisches Geowesen der deutschen Bundeswehr das *Topographische Informationssystem der Bundeswehr* (TOPIS), dessen Aufbau im MilGeo-Objektartenkatalog beschrieben ist (MilGeo-OK 1993). Dieses System baut auf der Struktur von FACC, einem Teil des Digital Geographic Information Exchange Standard (DIGEST) (Digest 2000), auf, berücksichtigt aber für den Bereich der Bundesrepublik Deutschland auch den Inhalt von ATKIS (MilGeo-OK 1993).

Die nächsten beiden Unterabschnitte behandeln die Modellierung von ATKIS und TOPIS und ihre Abbildungen auf das oben entwickelte erweiterte hierarchische Datenmodell.

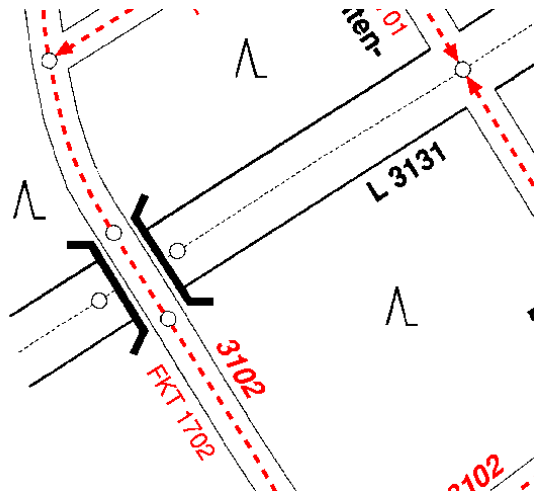


Abbildung 3.10: Die Kanten der Objektteile der Straße (L3131) und des Weges (3102) kreuzen sich, bilden aber keinen Knoten (Ausschnitt aus einem Modellierungsbeispiel für die Objektart 3102 Weg, AdV 2003).

Das ATKIS-Datenmodell

Das ATKIS-Datenmodell setzt sich aus drei Ebenen zusammen (vgl. Abb. 3.9), aus der *semantischen Ebene der DLM-Objekte*, der *semantisch-topologischen Ebene der DLM-Objektteile* und der *geometrischen Ebene der DLM-Geometrielemente*.

Es liegt nahe, die drei Ebenen dieses Modells auf die drei Ebenen des erweiterten hierarchischen Modells (Abschnitt 3.1.3) abzubilden. Dieser Versuch scheitert aber, weil die semantisch-topologische Ebene der Objektteile *nicht* die Topologie im Sinne der Graphentheorie widerspiegelt: Die Kanten im ATKIS-Modell dürfen sich in bestimmten Fällen schneiden, ohne dass an diesen Schnittpunkten ein Knoten gebildet werden muss. Diese Fälle treten auf, wenn linienförmige Objekte sich über- oder unterlagern, im Allgemeinen dann, wenn Verkehrswege sich nicht niveaugleich kreuzen. (Ein Beispiel ist in Abbildung 3.10 gezeigt.) Damit bilden die Maschen in ATKIS keine Tesselation der Erdoberfläche. Ausführlich wird diese Unvereinbarkeit von ATKIS-Topologie und graphentheoretischer Topologie von Gröger (2000, Kap. 4.4) behandelt.

Ein weiteres Problem, das sich aus dem ATKIS-Datenmodell aus objektorientierter Sicht ergibt, ist das Verhältnis zwischen *Objekten* und *Objektteilen*. Die Dokumentation zum ATKIS-Datenmodell beschreibt das Verhältnis von DLM-Objekt und DLM-Objektteilen wie folgt: „Ein DLM-Objekt besteht aus einer Menge von DLM-Objektteilen“ (AdV 1989, Teil C, S. 7). Dies legt eine objektorientierte Modellierung durch eine Aggregation nahe, die eine ‚besteht-aus‘-Relation ist. Zur Bildung von DLM-Objekten und DLM-Objektteilen werden jedoch topographische Unterscheidungen herangezogen, die in den Attributen zum Ausdruck kommen (ebd.). Attribute wiederum können der Objektebene *und* der Objektteilebene zugeordnet werden (AdV 1989, Teil D0, S. 9). Entscheidend für die

objektorientierte Modellierung ist jedoch, dass „der einem Attributtyp auf Objektebene zugeordnete Wert für alle Objektteile eines Objektes gilt.“ (ebd.). Diese Aussage spiegelt sich jedoch nicht in dem ATKIS-Datenmodell wider und lässt sich auch durch eine objektorientierte Modellierung nicht ohne weiteres ausdrücken. Eine mögliche Modellierung bieten die Arbeiten von Bender (1996) und Averdung (1996), in denen die Autoren die Attribute durch eine eigene Klasse modellieren. Diese Klasse assoziieren sie sowohl mit den Objekten als auch mit den Objektteilen; Bender (1996) auch mit den komplexen Objekten. Die Zugehörigkeit von Attributwerten der Objekte zu den Objektteilen ist aber auch hier nur implizit vorhanden. In dem hier gezeigten Kontext der Geodatenmodellierung wird die zuerst erwähnte Möglichkeit der Modellierung durch Aggregation, ohne eigene Attributklassen, gewählt.

Eine mögliche Lösung der Zuordnung zwischen ATKIS-Datenmodell und erweitertem hierarchischen Datenmodell ist in Abbildung 3.11 schematisch dargestellt. Dabei ist die semantisch-topologische Ebene der Objektteile aus dem ATKIS-Datenmodell auf die thematische Ebene des erweiterten hierarchischen Modells abgebildet. Die semantische Ebene der DLM-Objekte muss damit der Ebene der komplexen Objekte entsprechen, welche aus „einfachen“ Objekten aggregiert werden. Die geometrischen Ebenen beider Modelle entsprechen einander eins zu eins, wobei die Vektorelemente aus dem ATKIS-Modell geordneten Mengen von Linien entsprechen. Für die topologische Ebene des erweiterten hierarchischen Modells gibt es damit keine Entsprechung im ATKIS-Datenmodell.

Die topologische Ebene des erweiterten hierarchischen Datenmodells erzeugt man, indem die Topologie, die im ATKIS-Datenmodell in der semantisch-topologischen Ebene enthalten ist, kopiert und um die fehlenden Elemente ergänzt wird, damit sie der zweidimensionalen Topologie im graphentheoretischen Sinne entspricht. Im Beispiel aus Abbildung 3.10 muss man den geometrischen Schnittpunkt zwischen Straße und Weg durch einen zusätzlichen Knoten modellieren und die Kanten der Objektteile an dieser Stelle aufspalten. Damit ergibt sich auch für die angrenzenden Flächen (drei Nadelwälder und ein unbestimmtes Gebiet) eine Modellierung durch Maschen, die konsistent mit den Kanten verknüpft werden können, und die damit eine konsistente Tessellation bilden.

Das TOPIS-Datenmodell

Das TOPIS-Datenmodell kennt im Gegensatz zum ATKIS-Datenmodell keine Unterteilung von Objekten in Objektteile. Schürer (2002, Kap. 3.2.5) befasst sich ausführlich mit den Unterschieden in der Modellierung von ATKIS- und TOPIS-Datenmodell. Er überträgt die Ebenendefinition des ATKIS-Datenmodells auf das TOPIS-Datenmodell und bildet das TOPIS-Datenmodell aus den zwei Ebenen der Geometrieelemente und der semantisch-topologischen Elemente. Die semantische Ebene des ATKIS-Datenmodells entfällt. Diese Herangehensweise wird gestützt durch den MilGeo-OK (1993, Teil 1, S. 17), in dem es zur Objektbildung heißt: „...jede geometrisch abgrenzbare Einheit [führt] zur Objektbildung.“ Schürer (2002) spricht bei TOPIS von einem *flachen* Datenmodell im Gegensatz zu einem *stärker strukturierten* Datenmodell bei ATKIS.

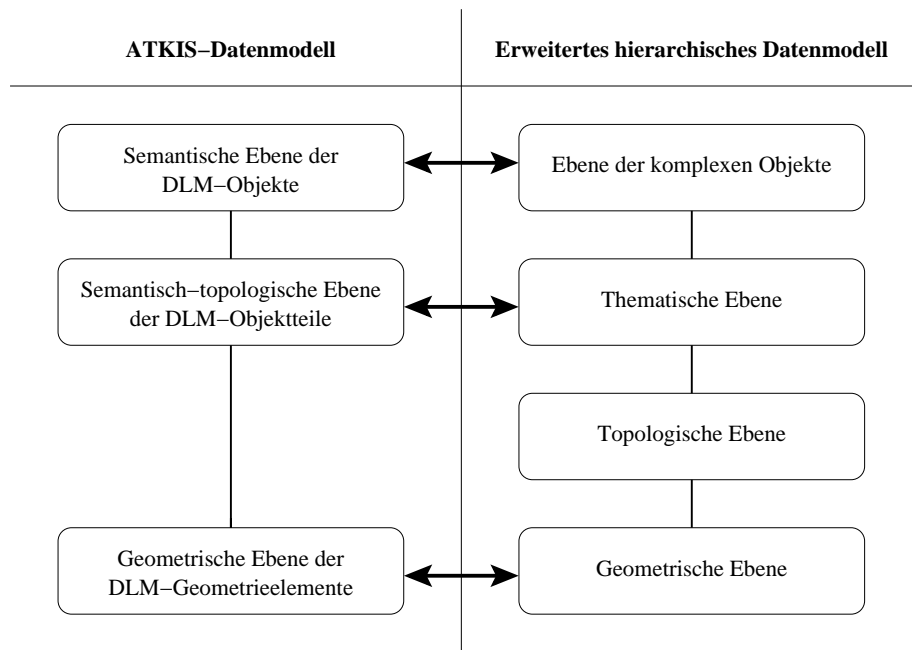


Abbildung 3.11: Ebenenzuordnung zwischen *ATKIS-Datenmodell* und *erweitertem hierarchischen Datenmodell*.

Die Zuordnung des TOPIS-Datenmodells zum erweiterten hierarchischen Datenmodell ist somit einfacher als beim ATKIS-Datenmodell (Abbildung 3.12). Die semantisch-topologische Ebene der DLM-Objekte des TOPIS-Datenmodells wird der thematischen Ebene zugeordnet, die geometrische Ebene der DLM-Geometrielemente der geometrischen Ebene des erweiterten hierarchischen Modells. Die topologische Ebene wird ebenso wie bei der Überführung des ATKIS-Datenmodells mit topologischen Elementen ergänzt, während die Ebene der komplexen Objekte im TOPIS-Datenmodell kein Äquivalent findet.

3.2 Modellierung raumbezogener Daten in mehreren Strukturierungsgraden

Ging es im ersten Abschnitt dieses Kapitels um die Modellierung von Geodaten, die im selben Strukturierungsgrad vorliegen, so ist in diesem Abschnitt die Modellierung von Geodaten in verschiedenen Strukturierungsgraden das Thema. Diese Modellierung ist insbesondere im Kontext der Generalisierung wichtig, weil es hierbei um die Abbildung von Daten in (mindestens) zwei Strukturierungsgraden geht.

Wie in Abschnitt 2.2.3 bereits erläutert, ergeben sich für die automatisierte Generalisierung und für die Fortführung von Daten erhebliche Vorteile, wenn die Beziehung von

3.2 Modellierung raumbezogener Daten in mehreren Strukturierungsgraden

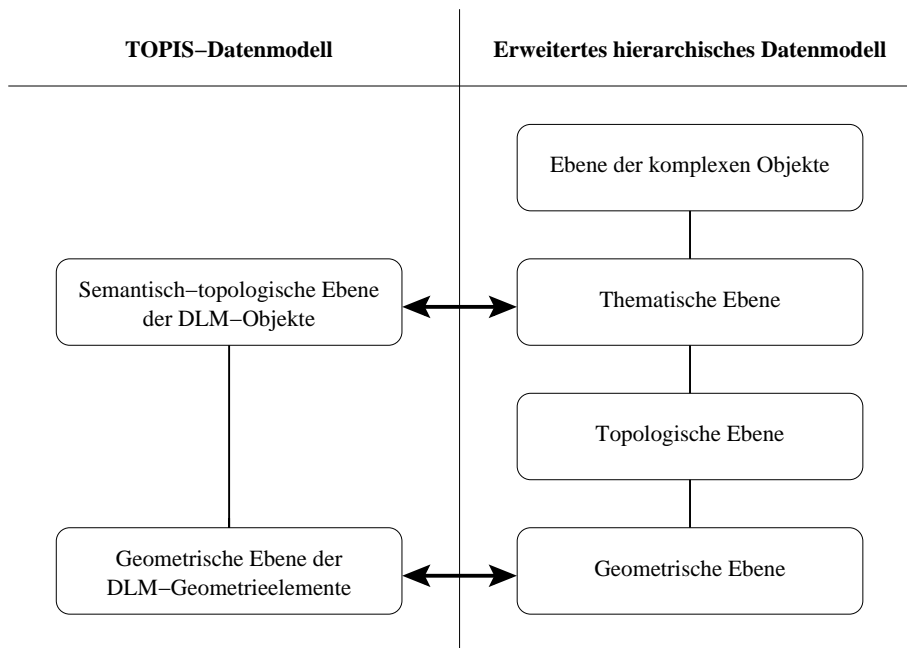


Abbildung 3.12: Ebenenzuordnung zwischen TOPIS-Datenmodell und erweitertem hierarchischen Datenmodell.

generalisierten Daten zu den ungeneralisierten Daten erhalten bleibt. Zu diesem Zweck ist es notwendig, ein Datenmodell zu schaffen, das diese Beziehungen abbildet und speichert.

Der folgende Unterabschnitt behandelt Möglichkeiten und Ansätze, wie mit Geodaten in verschiedenen Strukturierungsgraden verfahren werden kann. Die Weiterentwicklung des im vorigen Abschnitt (3.1) entwickelten erweiterten hierarchischen Datenmodells und die Übertragung auf die Modellierung zweier miteinander verknüpfter Datensätze unterschiedlicher Strukturierungsgrade ist Ziel des darauf folgenden Unterabschnitts.

3.2.1 Überblick

Wie in der Einleitung zu diesem Kapitel beschrieben, ist der Zweck eines Datenmodells die Festlegung einer Abbildung von Geo-Objekten in einen Rechner (Seite 39). Bei den abgebildeten Daten handelt es sich also um eine *Repräsentation* von Geo-Objekten. Der erste Teil dieses Kapitels (3.1) behandelt ausführlich Möglichkeiten einer solchen Abbildung. Die so abgebildeten Geodaten können für eine Vielzahl von Anwendungen benutzt werden. Viele dieser Anwendungen verlangen aber eine vorhergehende Generalisierung der Daten. Ideal wäre eine Datenbank, in der die Geodaten hochpräzise gespeichert sind, und aus der man alle gewünschten Anwendungsdaten automatisch ableiten könnte (Spacapietra u. a. 2000). Diese Automatismen existieren aber (zur Zeit) nicht, und es ist

fraglich, ob alle Ableitungen jemals vollautomatisch hergestellt werden können.

Im Gegensatz dazu werden in heute praktizierten Ansätzen dieselben Geodaten mehrfach in eine oder mehrere Datenbanken abgebildet, meistens in verschiedenen Auflösungen, also für die Darstellung in Karten mit verschiedenen Skalierungen. Für diese multiplen Abbildungen existieren in der Fachliteratur mehrere Begriffe, die allerdings nicht klar von einander abgegrenzt sind. Die Begriffe, um die es sich handelt, sind:

- *Multi-Scale Database* (Multiskalige Datenbanken), verwendet zum Beispiel von Hardy (2000), Müller u. a. (1995) und Jones (1991)
- *Multi-Resolution Database* (Datenbanken mit multiplen Auflösungen), verwendet zum Beispiel von Hampe u. a. (2003) und Anders u. a. (2003)
- *Mehrfacher Strukturierungsgrad*, verwendet zum Beispiel von Schürer (2002), Bobzien (2000, 2001)
- *Multi-Representation Database* (Multirepräsentationsdatenbank), verwendet zum Beispiel von Harrie (2001), Bédard u. a. (2002) und Kreiter (2002b).

Auch Mischformen der Begriffe tauchen auf, wie zum Beispiel *Multi-Scale Representations of Spatial Data* (ICA-ISPRS 2002) oder die annähernde Gleichsetzung der Begriffe *multiple representation database* und *multiple resolution database* (Hampe u. a. 2003). Die Bedeutung des Themas in der aktuellen Forschung, insbesondere im Kontext von GIS und Generalisierung, kann man zum Beispiel daran sehen, dass die *ICA⁹ Commission on Map Generalization* (Kommission für Kartengeneralisierung) im Jahre 2003 umbenannt wurde in *ICA Commission on Generalisation and Multiple representation* (Kommission für Generalisierung und multipler Repräsentation) (Ruas 2003).

Folgende Einordnung dieser Begriffe wird hier vorgeschlagen: Die Begriffe *multiskalig*, *multiple Auflösungen* und *mehrfache Strukturierungsgrade* beziehen sich alle auf die Darstellung von Daten in verschiedenen Präzisionen oder Genauigkeiten. Hierbei ist der Begriff ‚multiskalig‘ am meisten mit der graphischen Darstellung verknüpft, während der Begriff ‚Auflösung‘ hiervon ein wenig abstrahiert und der Begriff ‚Strukturierungsgrad‘ am weitesten von der graphischen Darstellung entfernt ist. Im Gegensatz zu den drei vorgenannten Begriffen nimmt der Begriff *Multirepräsentation* eine Sonderstellung ein: Multirepräsentation abstrahiert vollständig von einer Präzision oder Genauigkeit der Geodaten. Es geht hierbei nur um eine *mehrfache Abbildung von Geodaten* in die Datenbank. Insofern kann man diesen Begriff als Oberbegriff zu den vorhergenannten bezeichnen (siehe Abbildung 3.13).

Einen sehr guten Überblick über die verschiedenen Formen der Multirepräsentation bieten Spaccapietra u. a. (2000). Hier wird eine Systematisierung des Begriffes vorgenommen, und es werden unterschiedliche Aspekte des Begriffes in Bezug auf Datenbanken untersucht. Spaccapietra u. a. (2000) unterscheiden verschiedene *Dimensionen*, nämlich

⁹ICA = *International Cartographic Association*, Internationale Kartographische Vereinigung.

3.2 Modellierung raumbezogener Daten in mehreren Strukturierungsgraden

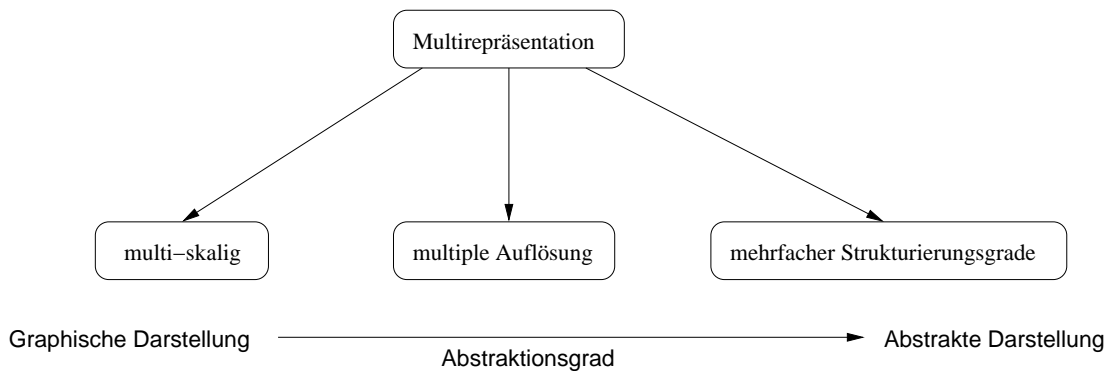


Abbildung 3.13: Beziehungen zwischen den Begriffen zur mehrfachen Abbildung von Geodaten.

die räumliche Dimension der Auflösung, die Dimension der Sichten und die Klassifizierungsdimension dieses Begriffes. Im Kontext dieser Arbeit ist die räumliche Dimension der Auflösung entscheidend, obwohl man bei der Erstellung eines konkreten Datenmodells auch die anderen Dimensionen berücksichtigen muss. Spaccapietra u. a. (2000) gebrauchen den Begriff ‚Auflösung‘ als Synonym zu den Begriffen Skala, Auflösung und Strukturierungsgrad in unserem Sinne. Innerhalb der räumlichen Dimension der Auflösung unterscheidet Spaccapietra u. a. (2000) zwischen zwei grundsätzlich verschiedenen Ansätzen der Repräsentation in Datenbanken: Zu einem Geo-Objekt in der Landschaft kann es

a) eine (Datenbank-)Instanz mit *mehrfacher* Auflösung
oder

b) *mehrere* (Datenbank-)Instanzen mit jeweils *einer* Auflösung

geben. In der ersten Variante wird zu jedem Geo-Objekt genau eine Instanz in der Datenbank geschaffen. Diese Instanz kann entsprechend den Anforderungen mit mehreren Geometrien (für unterschiedliche Karten) versehen sein. Diese Variante der Multirepräsentationsdarstellung wird zum Beispiel von Timpf und Frank (1995), Bédard u. a. (2002) und Vangenot (1998) vorgeschlagen. Die zweite Variante sieht vor, zu jedem Geo-Objekt mehrere Instanzen in der Datenbank zu schaffen, von der jede aber nur mit genau einer Geometrie versehen ist. Diese Variante wird zum Beispiel von Kilpeläinen (2000), Kreiter (2002a) und Bobzien (2000) vertreten. Die beiden Varianten sind in Abbildung 3.14 veranschaulicht.

Vorteil der Variante *a* ist, dass die Zuordnung zwischen Geo-Objekt und Datenbankinstanz eindeutig ist. In Variante *b* müssen Integritätsbedingungen sicher stellen, dass alle Instanzen zu einem Geo-Objekt dieselben auflösungsunabhängigen Eigenschaften besitzen, wie zum Beispiel attributive Eigenschaften (Spaccapietra u. a. 2000). Dies gilt nur für den Fall, dass ein Geo-Objekt in den verschiedenen Repräsentationen gleich modelliert sein soll. Soll das Geo-Objekt hingegen in einer Repräsentation mit anderen Geo-Objekten aggregiert modelliert werden, so entfällt dieser Nachteil.

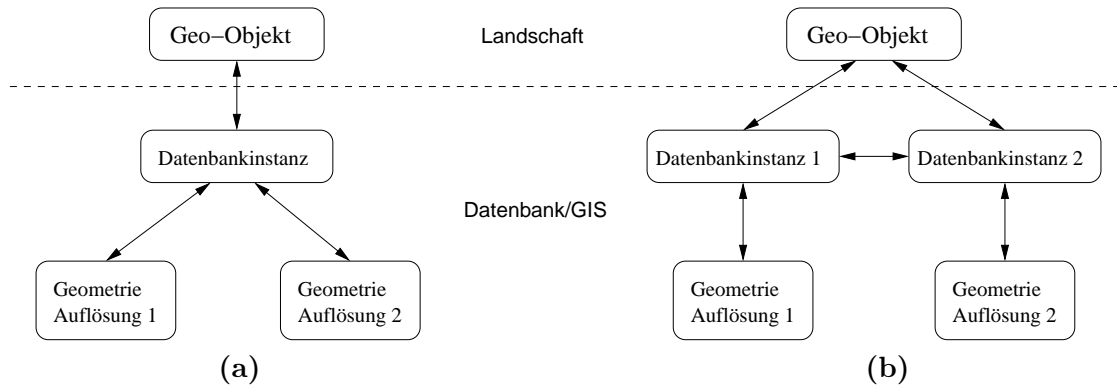


Abbildung 3.14: Repräsentation von Landschaftsobjekten in einer Datenbank: (a) eine Instanz mit multipler Auflösung, (b) mehrere Instanzen mit jeweils einer Auflösung.

Bei der Änderung des Strukturierungsgrades tritt häufig der Fall auf, dass Instanzen von Geo-Objekten infolge der Generalisierung fortgelassen oder zusammengefasst werden müssen. Für die Variante *a* bedeutet dies, dass im Fall des Fortlassens eine Instanz für eine bestimmte Auflösung keine Geometrie mehr besitzt, im Fall des Zusammenfassens, dass eine Geometrie mit mehreren Datenbankinstanzen verknüpft wird. Diese Fälle erfordern bei Variante *a* Spezialfälle in der Modellierung. Mit Variante *b* kann man Aggregationen problemlos modellieren, und auch das Fortlassen von Objekten ist leicht abzubilden. Hier hat Variante *b* einen nicht unerheblichen Vorteil gegenüber Variante *a* aufzuweisen. Die Behandlung von Spezialfällen entfällt, und die Modellierung bekommt einen übersichtlicheren Charakter.

Der folgende Unterabschnitt behandelt deshalb eine Modellierung nach Variante *b*. Mit dem in Abschnitt 3.1 vorgestelltem erweiterten hierarchischen Datenmodell als Grundlage ergibt sich eine symmetrische Modellierung von Geodaten in zwei Repräsentationen, zusammen mit einer Verknüpfung der einzelnen Elemente auf den unterschiedlichen Ebenen.

3.2.2 Erweitertes hierarchisches Datenmodell für mehrere Strukturierungsgrade

Das erweiterte hierarchische Datenmodell aus Abschnitt 3.1.4 bildet die Grundlage für die Modellierung eines Datenmodells für zwei (und mehr) Strukturierungsgrade. Für jeden Strukturierungsgrad wird das komplette Datenmodell übernommen. Neu hinzu kommen die Verknüpfungen zwischen den Datenmodellen, deren Modellierung explizit erfolgt. Jeder modellierte Datensatz ist in sich strukturiert und muss dementsprechend konsistent sein. Jeder Datensatz kann auch für sich existieren.

3.2 Modellierung raumbezogener Daten in mehreren Strukturierungsgraden

Im Kontext der Generalisierung ergibt sich eine *Ableitungskette* (Bobzien 2000) zwischen den Datenmodellen eines jeden Strukturierungsgrades, so wie es zum Beispiel von der amtlichen Kartographie in Deutschland (Harbeck 2000) oder in der Schweiz (Kreiter 2002a) für digitale Landschaftsmodelle gewünscht oder gefordert wird. Den Beginn der Ableitungskette bildet das Modell der ungeneralisierten Daten. Die Modelle der generalisierten Daten sind Folgedatenmodelle mit abgeleiteten Daten. Die Verknüpfungen zwischen den Datensätzen erlauben eine einfache Fortführung der Daten und ebenso eine einfache Analyse der Generalisierungsbeziehungen.

Das resultierende Gesamtdatenmodell kann beliebig viele Strukturierungsgrade umfassen. Um das Prinzip zu verstehen, genügt jedoch die Darstellung von zwei Strukturierungsgraden, der (ungeneralisierten) *Ausgangsdaten* und der (generalisierten) *Zieldaten*. Diese Begriffe sind den Begriffen Ausgangs-DLM und Ziel-DLM (Bobzien 2000) entlehnt.

Abbildung 3.15 zeigt zwei miteinander verknüpfte Datenmodelle. Im linken Teil der Abbildung sind die Ausgangsdaten modelliert, im rechten Teil die Zieldaten.¹⁰ Die Elemente des erweiterten hierarchischen Modells (vergleiche Abbildung 3.7) sind in der Abbildung vertikal ein wenig verschoben, um die Referenzen zwischen den Modellen besser darstellen zu können. Außerdem ist das rechte Modell eine Spiegelung des linken Modells. Die Assoziationen innerhalb beider Modelle sind punktiert dargestellt, da sie mit den Assoziationen des Modells aus Abbildung 3.7 übereinstimmen. An der Semantik der einzelnen Modelle hat sich nichts geändert.

Entscheidend an der dargestellten Modellierung sind die Verknüpfungen zwischen den Modellen der einzelnen Strukturierungsgrade. Verknüpfungen gibt es auf der Ebene der komplexen Objekte, auf der thematischen Ebene und auf der topologischen Ebene. Auf der geometrischen Ebene existieren keine Verknüpfungen.

Die Kardinalitäten der Assoziationen zu den Zieldaten sind im Allgemeinen durch Null oder Eins gegeben. Null entspricht dem Fortlassen eines Objektes oder eines Topologieelementes, Eins entspricht der Übernahme des Objektes oder das Aufgehen in einem Objekt. Auf diese Weise werden durch die Generalisierung auftretende Aggregationen modelliert. Auf der topologischen und auf der thematischen Ebene werden außerdem Geometriotypwechsel modelliert. Möglich sind nur *kleinerwertige* Geometriypwechsel, das sind Übergänge von einer flächenförmigen zu einer linienförmigen oder punktförmigen Ausprägung sowie Übergänge von einer linienförmigen zu einer punktförmigen Ausprägung. Übergänge von punktförmiger zu linien- oder flächenförmiger Ausprägung, sowie von linienförmiger zu flächenförmiger Ausprägung werden *größerwertige* Geometriypwechsel genannt und können nicht automatisch berechnet werden (Schürer 1992, Schoppmeyer und Heisser 1995, Schürer 2002). Für die automatisierte Generalisierung sind diese Übergänge deshalb nicht von Belang und werden hier nicht modelliert.

Im Einzelnen sehen die Verknüpfungen zwischen den beiden Modellen wie folgt aus: Ein komplexes Objekt kann entweder zu einem komplexen Objekt oder zu einem punkt-

¹⁰Zur Verdeutlichung der Abgrenzung sind beide Modelle durch Farbe unterlegt. Diese Kodierung gehört nicht zum UML-Sprachumfang.

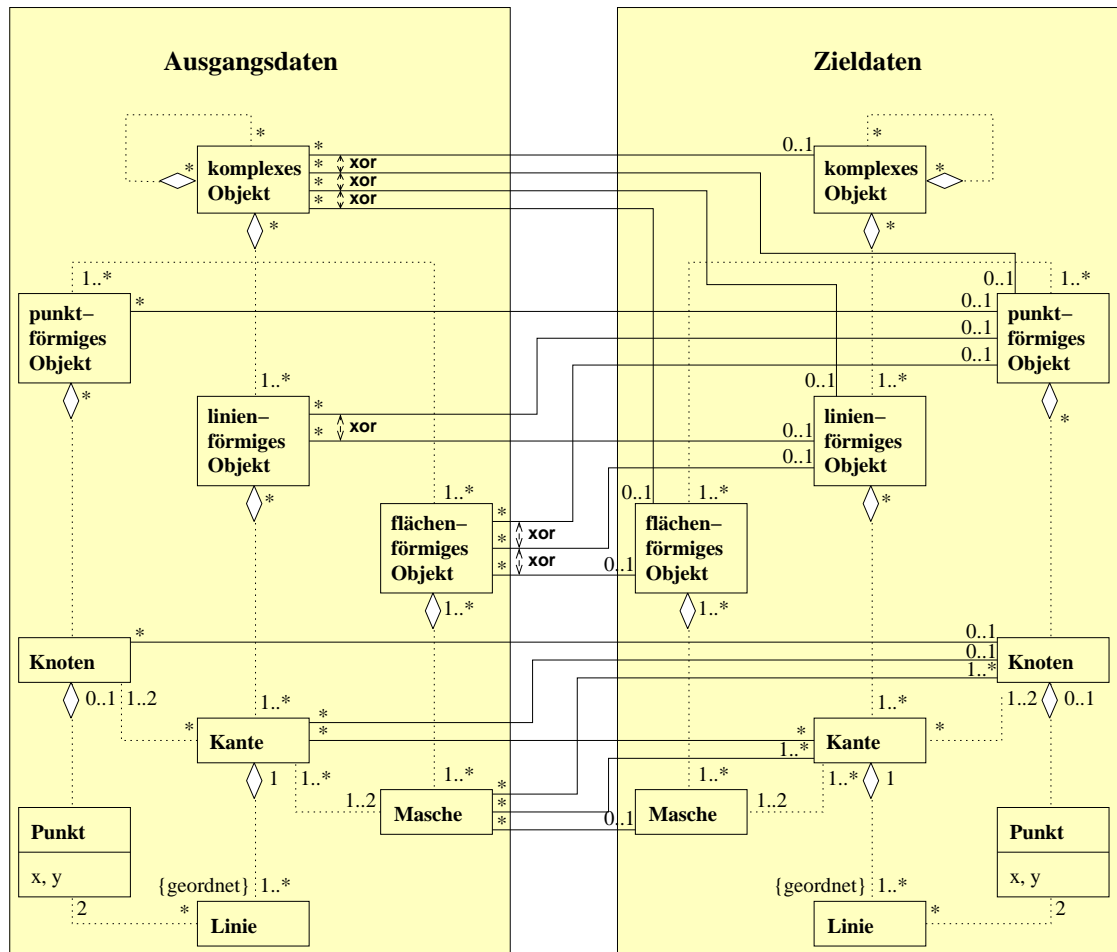


Abbildung 3.15: Verknüpfung zweier erweiterter hierarchischer Datenmodelle in zwei Strukturierungsgraden (nach Bobzien 2000, angepasst an das erweiterte hierarchische Datenmodell, siehe auch Abbildung 3.7).

linien- oder flächenförmigen Objekt generalisiert werden, oder es kann wegfallen. Die entweder-oder Beziehung wird durch die Bedingung xor (exklusives Oder) zwischen den Assoziationen formuliert.¹¹ Ein komplexes Objekt kann also seine „Komplexität“ durch die Generalisierung verlieren. Zum Beispiel wird eine Autobahn im ATKIS Basis-DLM komplex modelliert und besteht aus einem Straßenkörper und mehreren Fahrbahnen. Im ATKIS DLM50 hingegen wird eine Autobahn nicht komplex modelliert; bei der Generalisierung einer Autobahn zwischen diesen Landschaftsmodellen entsteht also aus einem komplexen Objekt ein linienförmiges Objekt.

¹¹Korrekt müsste die Bedingung, dass höchstens eine der vier Assoziationen belegt sein kann, durch einen logischen Ausdruck modelliert werden. Hierfür bietet sich die Object Constraint Language (OCL, Warmer und Kleppe 2003) an. Um die Graphik übersichtlicher zu machen, sind diese Bedingungen durch xor angedeutet. Selbiges gilt für die ausgehenden Assoziationen des flächenförmigen Objekts.

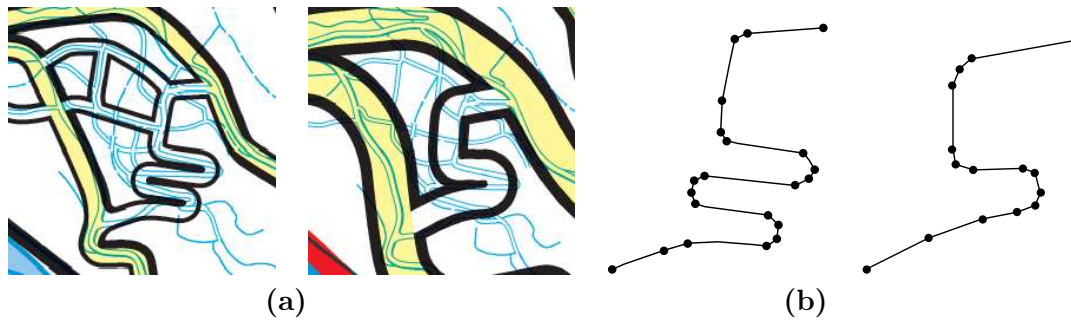


Abbildung 3.16: *Unmöglichkeit der Zuordnung von Geometrieelementen: (a) Beispiel der Generalisierung von Serpentine (aus SGK 2002), (b) Die entsprechende Darstellung durch Geometrieelemente.*

Auf der thematischen Ebene kann ein punktförmiges Objekt zu einem punktförmigen Objekt überführt werden oder es fällt weg. Aus einem linienförmigen Objekt kann ein linien- oder ein punktförmiges Objekt entstehen oder es entfällt. Der zweite dieser drei Fälle impliziert einen Geometriertypwechsel. Ein flächenförmiges Objekt kann zu einem flächen-, linien- oder punktförmigen Objekt generalisiert werden oder es entfällt. Die mittleren beiden Fälle bedeuten einen Geometriertypwechsel.

Auf der topologischen Ebene sehen die Assoziationen ähnlich aus wie auf der thematischen Ebene: Es kann Assoziationen geben zwischen Knoten und Knoten, zwischen Kante und Kante oder Knoten, und zwischen Masche und Masche, Kante oder Knoten. Es stellt sich die Frage, ob Geometriertypwechsel auf der topologischen Ebene abgebildet werden sollen und können. Bei Bobzien und Morgenstern (2002) werden thematisch-topologische Algorithmen zu Geometriertypwechseln entwickelt. Diese Algorithmen legen eine Verknüpfung auf der Ebene der Topologie nahe, weshalb hier bereits bei der Datenmodellierung die Möglichkeit der Verknüpfung geschaffen wird. Bei der Generalisierung von Maschen und Kanten wird bewusst auf die Bedingungen des exklusiven Oders verzichtet, welches man in der thematischen Ebene findet. Es ist also zum Beispiel zulässig, dass eine Kante gleichzeitig zu zwei Kanten und zu einem Knoten generalisiert wird. Es gibt jedoch folgende Einschränkung bei Maschen, die nicht in der Abbildung aufgeführt wird: Eine Masche wird *entweder* zu einer Masche generalisiert, *oder* sie wird zu Kanten oder einem Knoten oder zu Kanten und einem Knoten generalisiert. Diese Modellierung hängt eng mit dem in Bobzien und Morgenstern (2002) dargestellten Algorithmus zusammen, weswegen für Details auf diesen Artikel verwiesen wird.

Auf der geometrischen Ebene sind keine Verknüpfungen zwischen den einzelnen Modellen modelliert. Der Grund dafür ist Folgender: Geometrien, insbesondere Linienzüge, werden bei der Generalisierung vereinfacht und geglättet. Manuelle Generalisierung sowie einige Algorithmen zur Linienvereinfachung können die Linien soweit verändern, dass eine eindeutige Zuordnung von Punkten und Linien nicht mehr möglich ist. Ein Beispiel zeigt Abbildung 3.16. Nicht alle Stützpunkte der generalisierten Straße besitzen entsprechende Stützpunkte in der ungeneralisierten Straße. Bei der Generalisierung der Serpentine

wurden aus Platzgründen zwei Serpentinien zu einer zusammengefasst, so dass man für die Punkte der verbleibenden Kurve nicht mehr entscheiden kann, aus welcher der beiden Originalkurven sie entstanden sind. Gleiches gilt für die die Punkte verbindenden Linien.

3.3 Fazit

Dieses Kapitel hat die Modellierung von Geodaten als Thema. Der erste Abschnitt behandelt eine Modellierung für Daten eines Strukturierungsgrades. Diese Modellierung wird als *erweitertes hierarchisches Datenmodell* bezeichnet. Hauptmerkmal dieses Modells ist eine strikte Trennung zwischen geometrischen, topologischen und thematischen Elementen von Geodaten. Diese Trennung ermöglicht eine detaillierte Analyse von Eigenschaften der Geodaten, sowohl einzeln als auch im Zusammenspiel mit anderen Geodaten. Das Modell und die Trennung von Geodaten in ihre drei Bestandteile dient als Grundlage für die weiteren Untersuchungen in dieser Arbeit.

Der zweite Abschnitt befasst sich mit der Modellierung von Geodaten in mehreren Strukturierungsgraden. Die Analyse von bestehenden Ansätzen zeigt, dass dieses Thema in der Forschung aktuell sehr ausführlich, aber auch sehr uneinheitlich behandelt wird. Basierend auf dem erweiterten hierarchischen Datenmodell und der Multirepräsentation durch mehrere Datenbankinstanzen mit jeweils einer geometrischen Auflösung wird ein Datenmodell für zwei, durch Generalisierung auseinander hervorgegangene Strukturierungsgrade vorgestellt. Dieses Modell lässt sich auf mehr als zwei Strukturierungsgrade erweitern.

Die Verknüpfung von Datensätzen verschiedener Strukturierungsgrade ist auch das Thema des nächsten Kapitels. Dieses befasst sich mit der *Erzeugung* der generalisierten Daten aus den ungeneralisierten Daten und bietet dafür ein mathematisches Modell an, mit dem die Generalisierung auf eine mathematische Grundlage gestellt wird.

Kapitel 4

Generalisierung als Funktion

Während das Kapitel 3 die Modellierung der Generalisierung aus der Sicht der Datenerhaltung betrachtete, behandelt dieses Kapitel die Modellierung der Generalisierung aus mathematischer Sicht. Die Generalisierung erfährt eine Einordnung in die mathematischen Teilgebiete der Relationen und algebraischen Strukturen. Ziel dieses Kapitels ist die mathematische Beschreibung der Generalisierung in Form einer Funktion. Dabei erfolgt ein Rückgriff auf die im vorigen Kapitel erzielte Auftrennung von Geo-Objekten in Thematik, Topologie und Geometrie. Die Beziehungen zwischen zwei Datensätzen, die im Abschnitt 3.2 aufgestellt wurden, erhalten nun eine Einbettung in den Kontext der mathematischen Relationen und Funktionen. Das folgende Kapitel (Kapitel 5, Invariante Eigenschaften der Generalisierungsfunktion) nutzt die mathematische Formalisierung der Generalisierung, um Kriterien zur Beurteilung von Generalisierungsergebnissen aufzustellen und damit Entwicklern von automatisierten Generalisierungsfunktionen ein Mittel an die Hand zu geben, ihre Ergebnisse zu bewerten.

Der erste Abschnitt dieses Kapitels (Abschnitt 4.1) gibt eine ausführliche Motivation zur Formalisierung der Generalisierung. Diese zeigt, dass die Formalisierung kein Selbstzweck ist. Die Formalisierung beginnt im Abschnitt 4.2 zunächst mit der Formulierung der Generalisierung als Relation, was beim Entwickeln des Datenmodells (Abschnitt 3.2) und vorher schon in der Diskussion der Generalisierung als Konstruktion oder Reduktion (Abschnitt 2.2.3) angedeutet wurde. Ein entscheidender Punkt im Abschnitt über die Generalisierungsrelation ist die Betrachtung der Multiplizitäten dieser Relation, weil aus einer *Rechtseindeutigkeit* der Relation die Möglichkeit einer Formulierung als *Funktion* folgt. Dieser Abschnitt diskutiert ausführlich die Multiplizitäten der Relation sowie ihre Gestaltungsmöglichkeiten zur Rechtseindeutigkeit.

Aufbauend auf diesen Vorarbeiten wird im Abschnitt 4.3 die Generalisierung als Funktion definiert. Je nach Betrachtungsweise der Multiplizitäten der Generalisierungsrelation ergibt sich entweder eine spezielle Form der Generalisierungsfunktion (Abschnitt 4.3.2) oder eine von zwei möglichen allgemeinen Formen (Abschnitt 4.3.3). In Abschnitt 4.4 folgt eine erste Charakterisierung der Generalisierungsfunktion über die drei Eigenschaften *Nicht-Totalität* (oder *Partiellität*), *Nicht-Injektivität* und *Surjektivität*. Der Unterschied zwischen der speziellen und den allgemeinen Formen der Generalisierungsfunktion muss dabei jeweils berücksichtigt werden. Abschnitt 4.5 schließt das Kapitel ab mit einer

Diskussion über den Determinismus und Nicht-Determinismus der Generalisierung und über die Frage, wie der Determinismus und Nicht-Determinismus im Zusammenhang mit der Generalisierung mathematisch formuliert werden kann.

4.1 Motivation

Im Zuge der Bestrebungen zur Automatisierung der Generalisierung innerhalb der letzten Jahrzehnte war es nötig, dass die Prozesse der Generalisierung formalisiert werden. Erst damit wurde es möglich, diese Prozesse in die formale Umgebung von Rechnern und Programmierung zu transferieren. Es existieren bereits einige Ansätze zur Formalisierung der Generalisierungsprozesse. Die Bestrebungen richten sich auf die Formalisierung kartographischen Wissens (Buttenfield und McMaster 1991) zum Beispiel durch Erstellung von Regeln (Mark 1991, Heisser u. a. 1995), durch Formulierung von Bedingungen beziehungsweise *constraints* (Beard 1991) und durch Wissensklassifizierung (Armstrong 1991). Die Formalisierung der Generalisierungsprozesse selber wird zum Beispiel in Shea (1991) und Brassel und Weibel (1988) behandelt. Eine mathematische Beschreibung der Generalisierung wurde jedoch erst spät angegangen (Ai und van Oosterom 2001, Bobzien 2003).

Durch die Hinwendung zur Automatisierung kommt die Kartographie zwangsläufig mit der Informatik als Wissenschaft von der systematischen Verarbeitung von Informationen in Kontakt. Die Verarbeitung der Generalisierungsprozesse durch Algorithmen verlangt eine *Abstraktion*¹ von dem Problem, was es zu lösen gilt. Diese Abstraktion benötigen insbesondere Entwickler, die oftmals von einem konkreten Problem eines bestimmten Fachbereiches wie der Kartographie nur unvollständige Kenntnisse besitzen. Sie sind keine Fachkräfte dieses Bereiches und können nur mit einer Abstraktion des Problems arbeiten, um es in ein Programm umzuwandeln.

Auf der anderen Seite sehen Fachkräfte eines spezifischen Fachbereiches die Notwendigkeit zur Abstraktion erst dann, wenn die Automatisierung von vormals manuell ausgeführten Tätigkeiten benötigt wird. Das bedeutet, dass über lange Zeit entwickeltes Wissen in eine neue Form gebracht werden muss, für die es bisher in dieser Form keinen Bedarf gab.

Um die Lücke zwischen dem Bedarf an Abstraktion auf Seite der Programmierer und dem fehlenden historisch gewachsenem Verständnis von der Notwendigkeit der Abstraktion auf Seite der Kartographen zu schließen, bieten sich Elemente der Mathematik an. Wenn es also gelingt, das konkrete Problem der Generalisierung durch mathematische Mittel auszudrücken, so können Informatiker und Programmierer leichter ihr Wissen mit dem Problem verknüpfen. Die Informatiker sind dazu ausgebildet, mit Hilfe einer *Abstraktion* das Problem zu formalisieren, um es zu lösen. Die Programmierer benötigen die *Formalisierung* und können damit die Lösung implementieren.

¹Eine *Abstraktion* ist eine Verallgemeinerung zu einem Begriff (Wahrig 1994), hier also eine Verallgemeinerung vom zu lösenden Problem beziehungsweise von der Problemstellung.

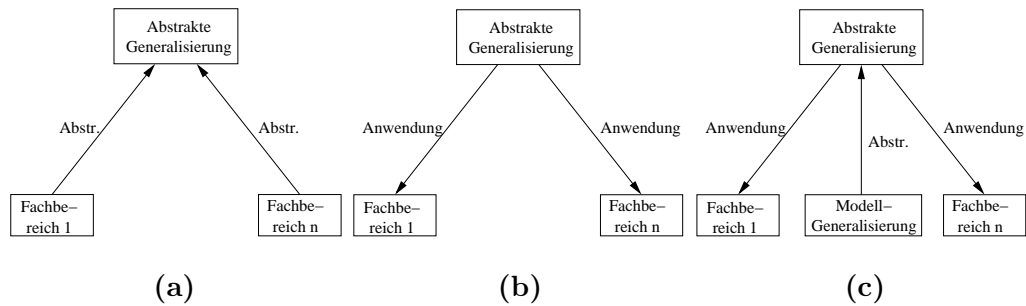


Abbildung 4.1: Mögliche Vorgehensweisen der Abstraktion und ihrer Anwendung.

In diesem Kapitel wird eine mathematisch beschreibende Abstraktion der Generalisierung entwickelt, gleichzeitig werden die nötigen Formalisierungen hinzugefügt. Es wird *kein* konkretes Verfahren beschrieben, mit dem die Generalisierung gelöst werden kann. Hingegen können bereits existierende Verfahren mit Hilfe der folgenden Abstraktion ausgedrückt werden. Die existierenden Verfahren sind größtenteils auf spezielle Probleme ausgerichtet. Mit Hilfe eines abstrahierten Verfahrens kann versucht werden, diese Verfahren auf andere Anwendungsbereiche zu übertragen.

Ein nicht unerheblicher zusätzlicher Vorteil der Abstraktion ist, dass Lösungen, die auf dieser Abstraktion basieren, nicht mehr an die Generalisierung in einem bestimmten Fachbereich gebunden sind, sondern auch für andere Generalisierungsanwendungen benutzt werden können. Zum Beispiel können Lösungen für die Modellgeneralisierung auf die Generalisierung geologischer Karten oder Straßenkarten übertragen werden. Diesen Beispielen wird in Kapitel 6 ausführlicher nachgegangen.

Abbildung 4.1 zeigt mögliche Vorgehensweisen, um zu einer Abstraktion zu gelangen und diese auf andere Anwendungen zu übertragen. Abbildung 4.1 (a) zeigt Generalisierungsprobleme aus unterschiedlichen Fachbereichen, die auf eine gemeinsame *abstrakte Generalisierung* verallgemeinert (generalisiert) werden. In Abbildung 4.1 (b) wird die abstrahierte Generalisierung auf verschiedene Fachbereiche angewendet (spezialisiert). Abbildung 4.1 (c) zeigt das in dieser Arbeit entwickelte Verfahren: Ein spezielles Generalisierungsproblem (hier die Modellgeneralisierung) wird verallgemeinert zur abstrakten Generalisierung und anschließend auf andere Generalisierungsprobleme angewendet.

Das nächste Kapitel (Kapitel 5) erweitert die in diesem Kapitel vorgestellten Formalisierungen auf konkrete Bereiche der Generalisierung und zeigt damit Möglichkeiten auf, wie bereits durchgeführte Generalisierungen beurteilt werden können.

Im übernächsten Kapitel (Kapitel 6) kommen die entwickelten Konzepte, die aus Erfahrungen mit der Modellgeneralisierung stammen, zur Anwendung auf Generalisierungsprobleme aus anderen Bereichen, indem die hier gewonnenen Konzepte auf diese übertragen werden.

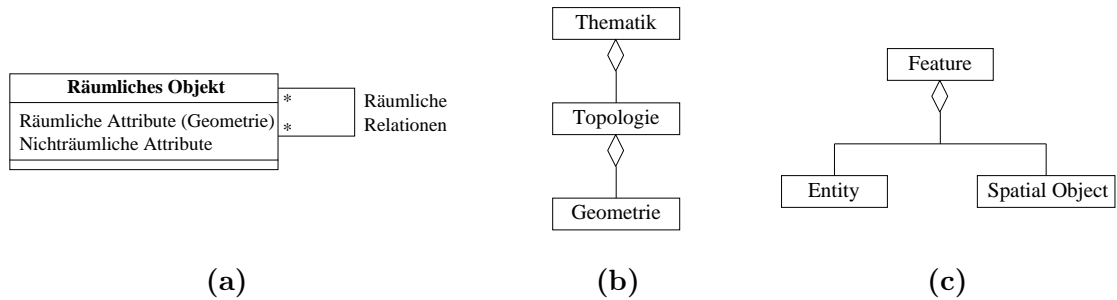


Abbildung 4.2: Der Begriff des Geodatoms und seine Modellierung: (a) Ausschnitt aus dem Datenmodell von Weber (1991) remodelliert in UML, (b) Trennung von Thematik, Topologie und Geometrie (siehe Kapitel 3), (c) Modellierung nach Moellering (2003), als UML-Diagramm.

4.2 Generalisierungsrelation

In den Abschnitten 2.2.3 und 3.2 wurde von *Verknüpfungen* und *Assoziationen* zwischen ungeneralisiertem und generalisiertem Datenbestand gesprochen. Verknüpfungen und Assoziationen sind Begriffe aus der Datenmodellierung. Im Bereich der Mathematik spricht man von *Zuordnungen* und *Relationen* zwischen Mengen. Mit Hilfe von Relationen lassen sich Zusammenhänge mathematisch beschreiben. Die Verknüpfungen zwischen den Ausgangsdaten und Zieldaten in Abschnitt 3.2 werden im Folgenden mathematisch als Relation interpretiert und in diesem Sinne formalisiert.

Relationen sind definiert als Teilmenge eines kartesischen Produktes von mindestens zwei Mengen (Reinhardt und Soeder 1974). Im Kontext der Generalisierung genügen zwei Mengen, die Menge der ungeneralisierten Geodaten und die Menge der generalisierten Geodaten. Eine Relation zwischen zwei Mengen wird *zweistellig* genannt. Geodaten, Mengen von Geodaten und Zuordnungen zwischen diesen Mengen sind Thema von Abschnitt 4.2.1. Der darauf folgende Abschnitt 4.2.2 formuliert die Generalisierungsrelation mathematisch. Abschnitt 4.2.3 bildet einen Exkurs über mögliche Schwierigkeiten beim Erstellen von Zuordnungen zwischen ungeneralisierten und generalisierten Daten. Der letzte Abschnitt zur Generalisierungsrelation (Abschnitt 4.2.4) diskutiert die Multiplizitäten der Generalisierungsrelation. Dieser Punkt ist entscheidend für die Interpretation der Generalisierungsrelation als Funktion (Abschnitt 4.3).

4.2.1 Geodaten, Mengen von Geodaten und Zuordnungen zwischen Mengen von Geodaten

Um *Mengen* von Geodaten und Zuordnungen zwischen diesen Mengen zu beschreiben, muss man zunächst den Begriff *Geodaten* – oder in der Einzahl *Geodatum* – noch einmal näher beleuchten. Im folgenden Absatz wird der Begriff Geodatum im Kontext von Vektordaten mit drei in der Literatur verwendeten Begriffen verglichen.

Hake u. a. (2002, S. 14 f.) sprechen von *Objekten mit Raumbezug*, die sich in *Gegenstände* und *Sachverhalte* einteilen lassen. Beide Arten können in Karten und digitalen Landschaftsmodellen dargestellt werden. Im Zusammenhang mit Datenmodellen sprechen Hake u. a. (2002, S. 156) von raumbezogenen Objekten und *Geo-Objekten*. Weber (1991) spricht von *räumlichen Objekten* und nennt Aggregationen davon *komplexe räumliche Objekte*. Räumliche Objekte besitzen nicht-räumliche und räumliche Attribute sowie räumliche Relationen (Abbildung 4.2 a). Abgebildet auf das erweiterte hierarchische Datenmodell aus Abschnitt 3.1 (abstrahiert dargestellt in Abbildung 4.2 b) ergibt sich in etwa folgende Entsprechung: Die nicht-räumlichen Attribute sind den Elementen der Thematik zugeordnet; die räumlichen Attribute der Geometrie und die räumlichen Relationen der Topologie.

Im angelsächsischen Sprachraum werden die Begriffe *entity*, *feature* und *(spatial) object* verwendet (Moellering 2003). Dabei entspricht *entity* (Entität, Instanz) am ehesten dem Begriff des Objektes mit Raumbezug als Gegenstand nach Hake u. a. (2002) und *spatial object* (raumbezogenes Objekt) einem geometrischen Element. Ein *feature* (Merkmal) wiederum ist nach Moellering eine Aggregation aus *entity* und *spatial object*, entspricht also weitgehend dem Begriff des räumlichen Objektes nach Weber (Abbildung 4.2 c).

An den unterschiedlichen Verständnissen des Begriffs Geodatum wird sichtbar, dass keine einheitliche Definition existiert, und auch, wie schwierig es ist, eine solche zu erstellen. Je nach Motivation wird die eine oder andere Sichtweise bevorzugt. Für die weiteren Betrachtungen wird hier unter dem Begriff *Geodatum* ein *räumliches Objekt in seiner Gesamtheit mit Thematik, Topologie und Geometrie* verstanden², so wie es in einem Rechner mit Hilfe eines Datenmodells repräsentiert ist. Der Begriff Repräsentation wurde bereits in Abschnitt 3.2 verwendet und zeigt, dass jedes Geodatum mit einem bestimmten Strukturierungsgrad verbunden ist. Dieser Begriff des Geodatums entspricht der Datenmodellierung aus Kapitel 3 und der Darstellung in Abbildung 4.2 (b).

Für die Formalisierung der Generalisierung als Relation muss als nächstes der Begriff *Menge von Geodaten* erläutert werden: Der mathematische Begriff *Menge* impliziert, dass jedes Element nur einmal in der Menge enthalten ist. Ein räumliches Objekt darf damit nur einmal auftreten. Ein Geodatum, als eine Repräsentation eines räumlichen Objektes, dürfte jedoch in verschiedenen Auflösungen oder Strukturierungsgraden mehrfach in einer Menge auftauchen. Dies würde aber eine Differenzierung der unterschiedlichen Strukturierungsgrade unnötig erschweren. Deshalb wird im Folgenden eine Menge von Geodaten immer nur Geodaten ein- und desselben Strukturierungsgrades enthalten.

Für die Generalisierung werden damit genau zwei Mengen von Geodaten in zwei unterschiedlichen Strukturierungsgraden für eine Menge von räumlichen Objekten benötigt. Die eine Menge ist die Menge der ungeneralisierten Geodaten, auch *Ausgangsmenge* oder *Urmenge* genannt, die andere Menge ist die Menge der generalisierten Geodaten, auch *Zielmenge* genannt.

²Der an anderen Stellen zusätzlich hinzugezogene *temporale Aspekt* wird hier nicht benötigt und deshalb nicht berücksichtigt.

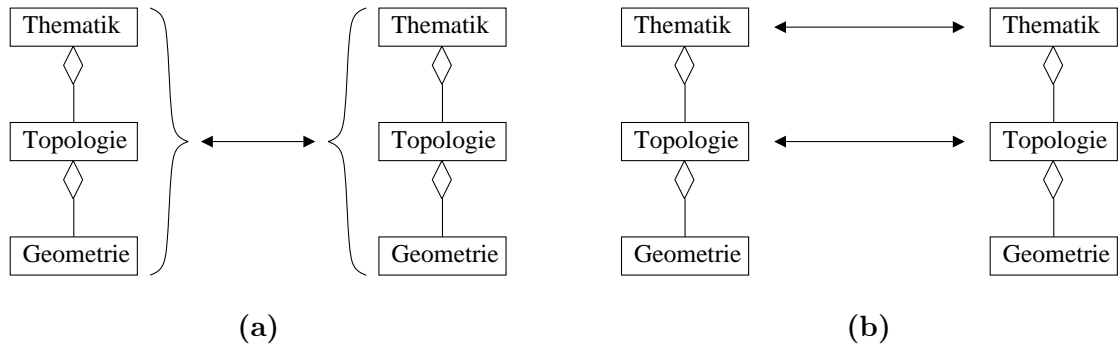


Abbildung 4.3: Relationen zwischen je zwei Geodaten. Die Klassen *Thematik*, *Topologie* und *Geometrie* bilden als Einheit das Geodatum ab.

Eine Relation ordnet Elemente aus zwei (oder mehreren) Mengen einander zu. Im Falle der Generalisierungsrelation ist dies eine Zuordnung zwischen Geodaten verschiedener Strukturierungsgrade. Mit der gewählten Modellierung von Geodaten aus Abbildung 4.2 (b) ergibt sich nun die Zuordnung, die in Abbildung 4.3 (a) gezeigt ist. Die Relation ist eine Zuordnung zwischen zwei Geodaten in ihrer Gesamtheit. Dies ist der kanonische Ansatz einer Relation zwischen Geodaten verschiedener Strukturierungsgrade.

In Abbildung 4.3 (b) hingegen sind zwei Relationen zwischen den Bestandteilen der Geodaten dargestellt, nämlich zwischen *Thematik* und *Topologie*. Auch eine Relation zwischen den Bestandteilen *Geometrie* wäre denkbar. In Anlehnung an die Diskussion in Abschnitt 3.2.2 wird diese Relation hier nicht eingeführt. Diese Modellierung der Relation verlangt jedoch eine zusätzliche Mengendefinition; nicht mehr Geodaten sind die Elemente der Mengen, sondern es existieren drei Mengen pro Strukturierungsgrad: Mengen, bestehend aus den thematischen Bestandteilen der Geodaten, Mengen, bestehend aus den topologischen Bestandteilen der Geodaten, und Mengen, bestehend aus den geometrischen Bestandteilen der Geodaten.

Diesen Ansatz benötigt man, wenn beim Generalisieren ein Konflikt zwischen den thematischen und topologischen Bestandteilen der Geodaten auftritt, die zu einem Widerspruch in der korrekten Repräsentation der Geodaten in einem geringeren Strukturierungsgrad führen können. In der Kartographie werden solche Widersprüche geduldet, da das Generalisieren das Entfernen von Spezifischem, Zufälligem und Unwichtigem zum Grundsatz hat (Brassel und Weibel 1988). In der Mathematik müssen solche Widersprüche aufgelöst werden. Der elementare Generalisierungsvorgang des Zusammenfassens (Aggregation) ist ein Beispiel, bei dem Widersprüche zwischen *Thematik* und *Topologie* auftreten können. Kapitel 5 und speziell Abschnitt 5.2.3 befassen sich unter anderem mit diesem Vorgang, bei dem die Auftrennung der Relation in zwei Relationen zwischen den Bestandteilen der Geodaten nötig wird.

Ein wichtiger Aspekt von Mengen von Geodaten ist die *Struktur* innerhalb der Mengen. In der Mathematik versteht man unter *strukturierten Mengen* solche Mengen, in denen

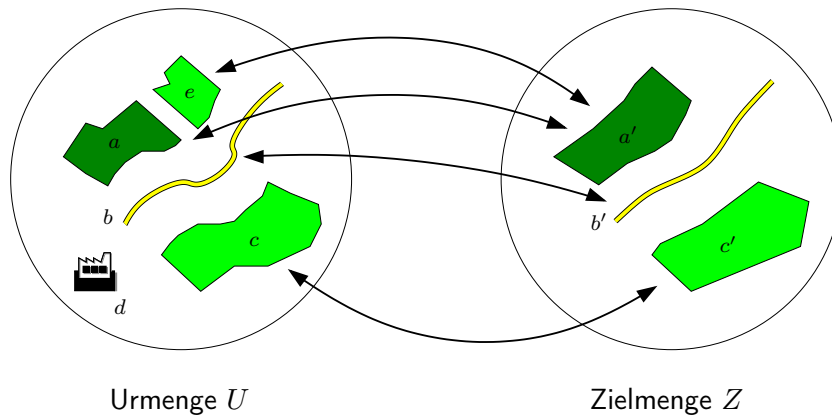


Abbildung 4.4: Beispiel für eine Generalisierungsrelation.

eine Familie von Relationen erklärt ist (Reinhardt und Soeder 1974, S. 37). Die Geodaten einer Menge und eines Strukturierungsgrades besitzen untereinander Zuordnungen, wie die räumlichen Relationen (geometrischer und topologischer Art) zwischen benachbarten räumlichen Objekten und die Relationen wie ‚verbunden mit‘ zwischen zum Beispiel Strommast und Stromleitung oder ‚Überführung/Unterführung‘ zwischen Brücken und Straßen oder Flüssen. Mit diesen und weiteren Relationen sind die Mengen von Geodaten strukturierte Mengen oder auch *Gebilde*.

Wenn auf den Mengen eine Algebra definiert ist, dann besitzen die Mengen eine *algebraische Struktur*. Diese Eigenschaft ist interessant, weil zwischen den Mengen eine Abbildung hergestellt werden soll (Abschnitt 4.3). Mit einer algebraischen Struktur bildet die Abbildung nämlich eine *Funktion* und kann mit algebraischen Methoden, zum Beispiel mit Hilfe von Morphismen, weiter untersucht werden. Algebren auf räumlichen Daten wurden zum Beispiel definiert von Egenhofer und Herring (1990a), Güting und Schneider (1993), Schneider (1997), Lee (1998).

4.2.2 Mathematische Formulierung der Generalisierungsrelation

Dieser Abschnitt bringt die beschriebenen Gedanken über die Generalisierungsrelation in einer mathematischen Form. Ausgangspunkt sind zwei Mengen von Geodaten, die ungeneralisierte *Urmenge* U und die generalisierte *Zielmenge* Z . Auf welche Weise die Generalisierung vollzogen wurde, ist an dieser Stelle zunächst unerheblich. Die *Generalisierungsrelation* R ist eine Teilmenge des kartesischen Produktes der Urmenge und der Zielmenge, also: $R \subseteq U \times Z$. Die Elemente der Relation sind Paare von Geodaten. Ein Paar besteht aus einem ungeneralisierten Geodatum und seinem generalisierten Gegenstück, zum Beispiel (Gebäude a , bebautes Gebiet b) oder (Waldweg x , Weg y).

Abbildung 4.4 verdeutlicht an einem Beispiel die Generalisierungsrelation. Die linke Menge ist ein Beispiel für eine kleine Menge von Ausgangsdaten, die rechte Menge ist die

dazugehörige Menge von generalisierten Geodaten, den Zieldaten. Die Urmenge U enthält fünf Elemente, ein Waldstück a , eine Straße b , zwei Grünflächen c und e und eine Fabrik d . Die Zielmenge Z enthält drei Elemente, das Waldstück a' , die Straße b' und die Grünfläche c' . Die Doppelpfeile visualisieren die Generalisierungsrelation. Das Waldstück a und die Grünfläche e aus der Urmenge sind dem Waldstück a' aus der Zielmenge zugeordnet. Sie wurden also bei der Generalisierung *zusammengefasst*. Die Straße b aus U ist der Straße b' aus Z zugeordnet und die Grünfläche c ist der Grünfläche c' zugeordnet. Die Fabrik d erhält zu keinem Element aus Z eine Zuordnung. Die Fabrik wurde *fortgelassen*. Formal besteht die Generalisierungsrelation R in diesem Beispiel aus folgenden Elementen: $R = \{(a, a'), (b, b'), (c, c'), (e, a')\}$.

4.2.3 Generalisierung versus relationale Zuordnung (Matching)

Dieser Abschnitt bildet einen Exkurs zu der Frage, wie die relationalen Zuordnungen zwischen zwei Mengen von Geodaten, ungeneralisiert und generalisiert, hergestellt werden können. Ein Anwendungsbeispiel ist die Digitalisierung von Karten in verschiedenen Maßstäben und die nachträgliche Zuordnung der digitalisierten Daten beider Maßstäbe. Um die Relation zu erstellen, müssen in den Datensätzen diejenigen Paare von Geodaten gefunden werden, die dieselben Realweltobjekte repräsentieren.

Das Problem der relationalen Zuordnung (englisch: *relational matching*) im Kontext von GIS wird ausführlich von Vosselmann (1992) behandelt, allerdings mit einem Fokus auf die Bildverarbeitung in der Photogrammetrie, mithin bei der Verarbeitung von Rasterdaten. Mit der Homogenisierung raumbezogener Vektordaten aus unterschiedlichen Quellen beschäftigt sich Illert (1995). Das spezielle Problem der Integration von ATKIS- und GDF-Daten ist Thema von Walter und Fritsch (1998). Mit der Integration von ungeneralisierten und generalisierten Daten aus unterschiedlichen Quellen beschäftigen sich Badard und Lemarié (2002). Der letzte Ansatz wird im Kontext von Multirepräsentationsdatenbanken geführt und kommt damit den hier angestellten Überlegungen über Generalisierung am nächsten. Das in dieser Arbeit entwickelte Datenmodell (siehe Kapitel 3) bietet eine Möglichkeit, diesen Ansatz zu erweitern.

Die relationale Zuordnung ist ein komplizierter Vorgang und kann in vielen Fällen nicht zufrieden stellend automatisch gelöst werden. Es ist auf jeden Fall vorzuziehen, diesen Vorgang zu vermeiden oder ihn überflüssig zu machen. Im Falle von Datensätzen aus verschiedenen Quellen (siehe Beispiele aus obiger Literatur) ist dies nicht möglich, wohl aber, wenn man Einfluss auf den Generalisierungsvorgang hat, beziehungsweise auf das Erzeugen des generalisierten Datensatzes. Dann nämlich kann mit Hilfe der Generalisierung als Konstruktion (Abschnitt 2.2.3) und des erweiterten hierarchischen Datenmodells in zwei Strukturierungsgraden (Abschnitt 3.2.2) die Beziehung zwischen den Datensätzen direkt gespeichert werden. Diese Vorgehensweise wird auch von Bobzien (2000) diskutiert.

4.2.4 Multiplizitäten der Generalisierungsrelation

Unter den Eigenschaften von zweistelligen Relationen ist die Frage nach der Multiplizität einer Relation am Interessantesten. Durch die Multiplizität wird ausgedrückt, wie viele Objekte der einen Menge wie vielen Objekten der anderen Menge zugeordnet sind und umgekehrt. Insbesondere die Eigenschaften der *Rechtseindeutigkeit* und der *Linkstotalität* sind wichtig, weil diese als Bedingungen gelten, damit eine Relation sich auch als Funktion ausdrücken lässt. Zu diesen beiden Eigenschaften gibt es weiter unten in diesem Abschnitt eine ausführliche Erläuterung und Diskussion. Ziel dieses Abschnittes ist, zu untersuchen, ob die Generalisierungsrelation die Voraussetzungen für die Interpretation als Funktion erfüllt.

Besondere Multiplizitäten werden gerne durch ein Verhältnis ausgedrückt, zum Beispiel $1:1$, $1:0$, $m:1$ oder $m:n$, wobei m und n natürliche Zahlen darstellen. Die drei ersten Beispiele sind hierbei Spezialfälle des letzten, allgemeinsten Falles. Dennoch verdienen sie eine gesonderte Erwähnung, weil sie gängige Multiplizitäten bei der Generalisierung darstellen.

Der erste Fall, die $1:1$ -Beziehung, tritt immer dann auf, wenn ein Geodatum aus der Urmenge in die Zielmenge übernommen wird, ohne es aufzuteilen oder mit einem anderen Objekt zusammenzufassen. Im Beispiel von Abbildung 4.4 entsprechen die Zuordnungen (b, b') und (c, c') dieser Form der Beziehung.

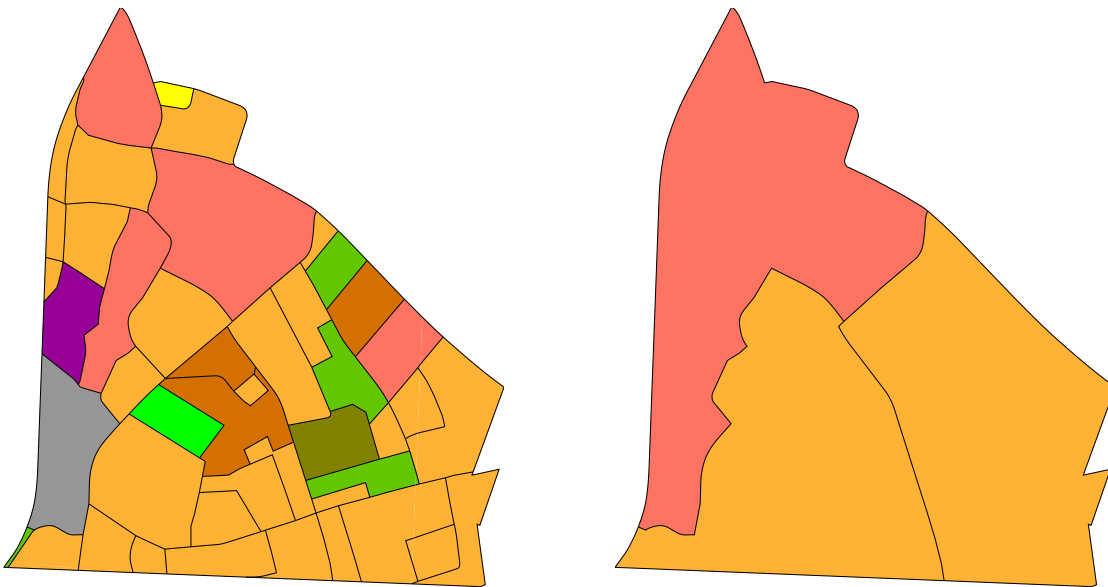


Abbildung 4.5: Beispiel für die Aggregation von Flächen bei der Generalisierung von Digitalen Landschaftsmodellen (aus Bobzien 2001).

Der zweite Fall, die 1:0-Beziehung, entspricht dem Fortlassen eines Objektes bei der Generalisierung. Die Fabrik in Abbildung 4.4 ist ein Beispiel für diese Art der Beziehung.

Der dritte Fall, die $m:1$ -Beziehung, mit $m > 1$, bedeutet ein Zusammenfassen (Aggregation) von *mehreren* Objekten aus der Urmenge zu *einem* Objekt der Zielmenge. In Abbildung 4.4 erfüllen die Paare (a, a') und (e, a') diese Form der Beziehung mit $m = 2$. Das Waldstück und die Grünfläche wurden zu einem Waldstück zusammengefasst.

In der Modellgeneralisierung tritt dieser Fall der Aggregation häufig auf, weil die Daten eines DLM im Allgemeinen flächendeckend sind, aber zu kleinflächig, um 1:1 in das generalisierte DLM aufgenommen zu werden. Abbildung 4.5 zeigt ein typisches Beispiel für die Generalisierung von Flächenobjekten innerhalb einer Ortslage. Links in der Abbildung die ungeneralisierten Daten aus dem Basis-DLM, rechts die aggregierten Flächen des DLM 250.

Die Aggregation von Flächenobjekten tritt auch bei der Generalisierung von Geodaten aus anderen Fachbereichen auf, zum Beispiel bei der Generalisierung von geologischen Karten (siehe Abschnitt 6.2) oder auch bei der Generalisierung topographischer Karten, zum Beispiel beim Zusammenfassen von Häusern zu Häuserblöcken.³

Der vierte Fall der Multiplizitäten, die $m:n$ -Beziehung, verdient eine besondere Betrachtung. Interessant sind hier besonders die Fälle, bei denen $m > 0$ und $n > 1$ ist. Dann nämlich besteht die Situation, in der mindestens ein Element der Urmenge mit mindestens zwei Elementen der Zielmenge verknüpft wird. Eine eindeutige Zuordnung der Elemente in die Zielmenge ist damit nicht mehr gegeben. Damit besteht auch keine Rechtseindeutigkeit in der Relation (siehe unten).

Abbildung 4.6 zeigt solch eine nicht-eindeutige Situation. Eine Häuserreihe, bestehend aus zehn Häusern (rot) im Maßstab 1:25 000, wird zahlenmäßig verringert zu einer Häuserreihe bestehend aus sechs Häusern (schwarz gestrichelt) im Maßstab 1:50 000. Auf den ersten Blick ist es nicht ersichtlich, welche Häuser zu welchen Häusern zusammengefasst wurden, oder welche Häuser bei der Generalisierung fortgelassen wurden. Man kann auch die Frage stellen, ob so eine Differenzierung hier überhaupt erlaubt ist, ob hier nicht vielmehr eine Häuserreihe als Ganzes generalisiert wurde und die einzelnen Häuser dabei ihre Identität verlieren.

Abbildung 4.7 zeigt vier mögliche und plausible Zuordnungen für die Generalisierung der Häuserreihe aus Abbildung 4.6. Eine Vielzahl von weiteren Möglichkeiten ist denkbar, ergibt aber keine grundsätzlich neuen Erkenntnisse. Die Zuordnung in (a) ist eine $m:n$ -Beziehung, mit $m \in \{1, 2, 3\}$ und $n \in \{1, 2\}$. Die weiteren Zuordnungen reduzieren die Anzahl der Paare aus der Relation: (b) ist eine $1:\{0, 1\}$ -Beziehung, also eine Mischung

³In der englischen Literatur wird beim Zusammenfassen unterschieden zwischen der *aggregation*, dem Zusammenfassen von punktförmigen Geodaten wie zum Beispiel Häusern, der *amalgamation*, dem Zusammenfassen von flächenförmigen Geodaten wie zum Beispiel Flächennutzungsarten, und dem *merging*, dem Zusammenfassen von linienförmigen Geodaten wie zum Beispiel Bahnstrecken (McMaster und Shea 1992).

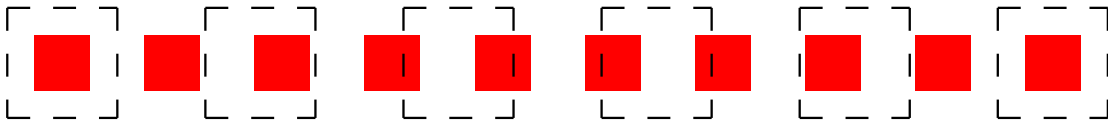


Abbildung 4.6: Generalisierung von Häuserreihen. Zehn Häuser (rot) werden zahlenmäßig zu sechs Häusern (schwarz) verringert (aus SGK 1980).

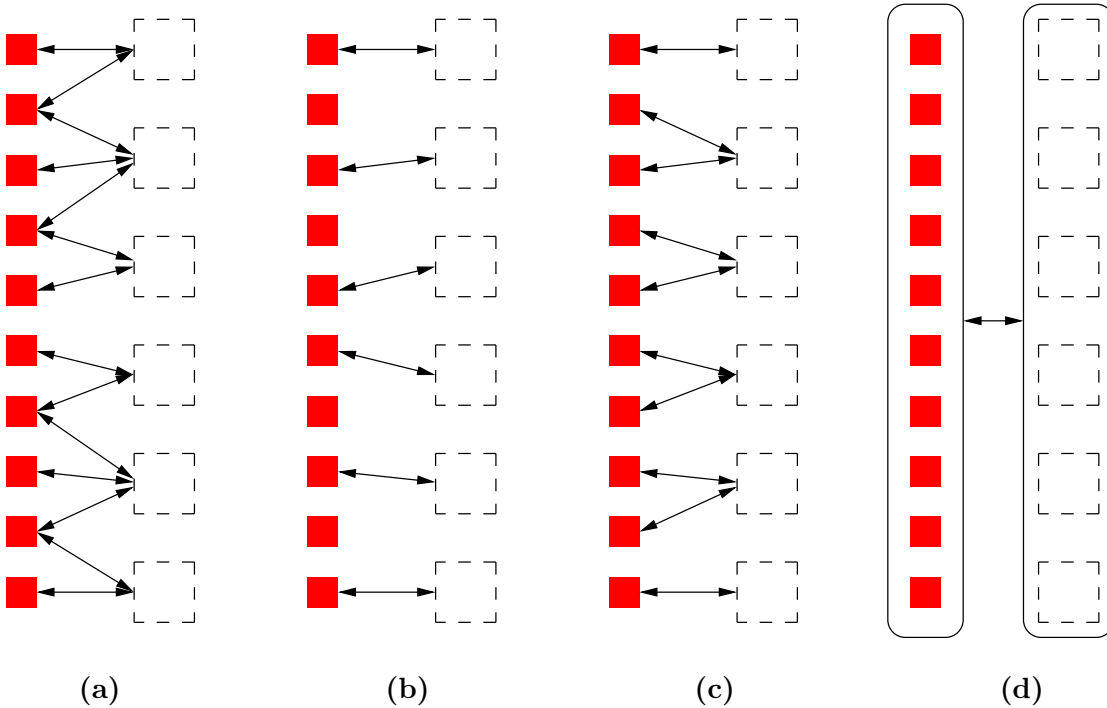


Abbildung 4.7: Mögliche Zuordnungen bei der Generalisierung der Häuserreihen aus Abbildung 4.6.

aus den beiden oben zuerst genannten Fällen einer 1 : 1- und 1 : 0-Beziehung. Die Möglichkeit in (c) ist eine $m : 1$ -Beziehung mit $m \in \{1, 2\}$. Viermal werden hier je zwei Häuser zu einem Haus aggregiert. Die Möglichkeit (d) ist ein Sonderfall: die Häuserreihen, ungeneralisiert und generalisiert, sind zu *Clustern* zusammengefasst, und die Relation besteht zwischen den Clustern. Dies erscheint wie eine 1 : 1-Beziehung, ist dies jedoch nur auf der Ebene der Cluster. Auf der Ebene der einzelnen Häuser entspricht diese Zuordnung einer $m : n$ -Beziehung, in der jedes ungeneralisierte Haus mit jedem generalisierten Haus verbunden ist, also eine 10 : 6-Beziehung. Cluster sind eine Art von Mengen; die Abbildung zwischen Mengen wird in Abschnitt 4.3.3 ausführlich behandelt.

Welche dieser vier Möglichkeiten der Zuordnung gewählt wird, oder ob noch eine andere Möglichkeit ausgesucht wird, hängt zum einen von der Interpretation des Generalisierungsvorganges ab, zum anderen aber auch, wie die Relation mathematisch ausgedrückt

werden soll. Mit dem Ziel, die Relation als Funktion zu verstehen, werden jetzt die beiden Eigenschaften *Rechtseindeutigkeit* und *Linkstotalität* eingeführt und anhand der Beispiele erläutert.

Rechtseindeutigkeit

Eine Relation $R \subseteq U \times Z$ ist *rechtseindeutig*, wenn für alle $x \in U$ und alle $y_1, y_2 \in Z$ gilt:

$$(x, y_1) \in R \wedge (x, y_2) \in R \implies y_1 = y_2.$$

Diese Definition sagt aus, dass jedes Element aus der Urmenge mit höchstens einem Element aus der Zielmenge durch die Generalisierungsrelation verknüpft sein darf. Der Ausdruck ‚höchstens ein‘ schließt natürlich mit ein, dass das Element auch mit *keinem* Element aus der Zielmenge verknüpft sein darf.

Das Beispiel aus Abbildung 4.4 ist rechtseindeutig, weil jedes Element aus U mit einem oder keinem Element aus Z verbunden ist.

Von den vier vorgestellten Möglichkeiten, die Relation zwischen den Häuserreihen in Abbildung 4.7 herzustellen, sind die beiden mittleren, (b) und (c), rechtseindeutig. Die vierte Möglichkeit (d) ist dann rechtseindeutig, wenn man sie auf der Ebene der Cluster betrachtet. Ganz klar nicht rechtseindeutig ist die Möglichkeit (a), weil hier einige ungeneralisierte Häuser mehreren generalisierten Häusern zugeordnet werden.

Es bleibt festzuhalten, dass es sich bei der Generalisierungsrelation in den *allgemeinen* Fällen um eine rechtseindeutige Relation handelt. Damit ist das Auswählen und Fortlassen von Objekten sowie die Aggregation von Objekten mit eingeschlossen. In komplizierteren Fällen wie in dem Beispiel der Häuserreihe kann es mehrere Möglichkeiten geben, die Relation zu bestimmen. Klar geworden jedoch sollte sein, dass es immer möglich, wenn auch vielleicht nicht immer kartographisch gewünscht sein kann, die Relation in eine rechtseindeutige Form zu bringen.

Linkstotalität

Eine Relation $R \subseteq U \times Z$ ist *linkstotal*, wenn für alle $x \in U$ ein $y \in Z$ existiert, so dass gilt:

$$(x, y) \in R.$$

Linkstotalität bedeutet also, dass jedes Element der Urmenge mindestens einem Element der Zielmenge zugeordnet wird.

Für die Generalisierungsrelation lässt sich sofort sagen, dass diese Eigenschaft im Allgemeinen *nicht* erfüllt ist. Denn wenn ein Geodatum durch die Generalisierung *fortgelassen* wird, so besitzt es kein zugeordnetes Geodatum in der Menge der generalisierten Daten.

Im Beispiel aus dem vorigen Abschnitt, Abbildung 4.4, wird die Fabrik d bei der Generalisierung fortgelassen und hat damit keine Entsprechung in der Zielmenge. Schon dieses kleine Beispiel ist eines für eine nicht-linkstotale Relation.

Im Beispiel der Häuserreihen aus Abbildung 4.7 sind die Zuordnungen (a), (c) und (d) linkstotal, letztere sowohl in der Interpretation als Cluster als auch als 10 : 6-Beziehung.

4.3 Generalisierungsfunktion

In diesem Abschnitt wird aus der Generalisierungsrelation eine Funktion entwickelt. Mathematisch ist eine Funktion ein Spezialfall einer Relation. Die Voraussetzungen für diesen Spezialfall werden im Unterabschnitt 4.3.1 diskutiert. Danach folgt die Formulierung der Generalisierungsrelation als Generalisierungsfunktion in der *speziellen* Form (Unterabschnitt 4.3.2). Wenn die Rechtseindeutigkeit nicht gegeben ist, so kann man auf eine von zwei *allgemeinen* Formen der Generalisierungsfunktion ausweichen (Unterabschnitt 4.3.3). An einem Beispiel wird in Unterabschnitt 4.3.4 der Unterschied zwischen der speziellen Form und den allgemeinen Formen der Generalisierungsfunktion deutlich gemacht.

In der Arbeit von Ai und van Oosterom (2001) ist ein ähnlicher Ansatz zur Formalisierung der Generalisierung vorgestellt worden. Unterabschnitt 4.3.5 beschäftigt sich ausführlich mit diesem Ansatz und vergleicht ihn mit dem hier entwickelten Konzept der Generalisierungsfunktion. Zum Abschluss folgt ein kleiner Unterabschnitt, der sich mit dem Unterschied zwischen Relation und Funktion aus kartographischer Sicht beschäftigt.

4.3.1 Voraussetzungen

Die Rechtseindeutigkeit einer Relation ist eine Bedingung dafür, dass die Relation eine Funktion bildet. In den Beispielen im vorigen Abschnitt wurde gezeigt, dass die Generalisierungsrelation im Allgemeinen rechtseindeutig ist. In den Fällen, in denen sie kartographisch nicht sofort als rechtseindeutig zu erkennen ist (Beispiel Häuserreihen), kann sie durch geeignete Wahl der Zuordnungen in eine rechtseindeutige Relation gewandelt werden.

Eine rechtseindeutige Relation kann als Funktion beschrieben werden, was hier als *spezielle* Form der Generalisierungsfunktion bezeichnet wird (siehe nächster Unterabschnitt). Wenn die kartographische Interpretation der Generalisierung eine Umwandlung in eine rechtseindeutige Relation nicht zulässt, so muss mit Hilfe der Potenzmengenbildung auf eine der *allgemeinen* Formen der Generalisierungsfunktion ausgewichen werden (siehe übernächster Unterabschnitt).

Oft wird gesagt, dass neben der Rechtseindeutigkeit auch die Linkstotalität eine Voraussetzung ist, um eine Relation *Funktion* zu nennen, zum Beispiel von Reinhardt und Soeder (1974). Es gibt aber auch Definitionen, insbesondere in der Informatik, die die

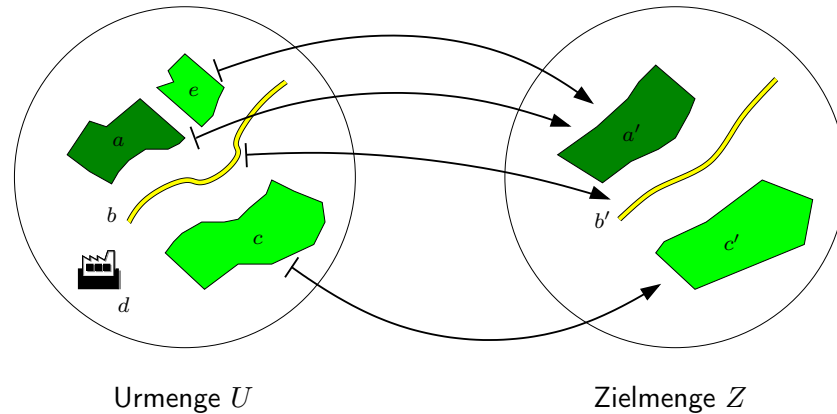


Abbildung 4.8: Beispiel für eine spezielle Generalisierungsfunktion.

Linkstotalität nicht als Voraussetzung für die Definition der Funktion nehmen, zum Beispiel die von Bauer und Goos (1992). Das Ergebnis wird *partielle* Funktion genannt, im Gegensatz zu einer *totalen* Funktion, die die Linkstotalität erfüllt. Mathematisch ist dieser Unterschied meistens unbedeutend, weil durch Anpassung der Definitionsmenge aus jeder partiellen Funktion eine totale Funktion entstehen kann. Im Falle der Generalisierung ist eine Änderung der Definitionsmenge jedoch nicht erwünscht. Sie würde eine Vorauswahl der zu generalisierenden Geodaten beinhalten, bei der alle Elemente, die bei der Generalisierung fortgelassen werden sollen, von vornherein aus der Menge entfernt werden. Dies kann nicht im Sinne der Kartographie sein. Deshalb wird hier ausdrücklich betont, dass die Generalisierungsrelation im Allgemeinen *nicht* linkstotal ist und die resultierende Generalisierungsfunktion eine *partielle Funktion* ist.

4.3.2 Spezielle Form der Generalisierungsfunktion

Gegeben seien die Urmenge U , die Zielmenge Z sowie die Generalisierungsrelation R . Wenn R rechtseindeutig ist, dann bildet R die *spezielle Generalisierungsfunktion* f_{spez} :

$$f_{\text{spez}} : U \longrightarrow Z.$$

Die spezielle Generalisierungsfunktion ist eine Abbildung *aus einer Menge auf eine Menge*. *Speziell* ist diese Form der Generalisierungsfunktion, weil die Abbildung jedes Geodatum einzeln behandelt und Zusammenhänge zwischen den Geodaten nur implizit enthält. Eine Aggregation kann zum Beispiel nur dadurch ausgedrückt werden, dass mehrere Elemente aus U auf ein und dasselbe Element aus Z abgebildet werden. Ein Abbilden eines Elements aus U auf mehrere Elemente aus Z ist nicht möglich.

Abbildung 4.8 zeigt das Beispiel aus Abbildung 4.4 als spezielle Generalisierungsfunktion: An die Stelle von Paaren von Geodaten, gekennzeichnet durch Doppelpfeile, sind Abbildungsvorschriften getreten, gekennzeichnet durch Pfeile mit Balken am anderen Ende.

Man schreibt $f(a) = a'$, $f(b) = b'$, $f(c) = c'$ und $f(e) = a'$. Die Abbildungsvorschrift für die Fabrik, $f(d)$, ist nicht definiert. Damit wird jedes Element von U auf höchstens ein Element aus Z abgebildet. Aus dem Beispiel der Häuserreihen aus Abbildung 4.7 können die Fälle (b) und (c) mit dieser Form der Abbildung ausgedrückt werden.

4.3.3 Allgemeine Formen der Generalisierungsfunktion

In diesem Unterabschnitt wird die spezielle Generalisierungsfunktion in zwei Stufen verallgemeinert. Die Verallgemeinerung findet durch den Übergang von der Abbildung je *eines* Elements aus U auf je *ein* Element aus Z auf die Abbildung von *mehreren* Elementen aus U auf *mehrere* Elemente aus Z statt. Die erste Verallgemeinerung betrachtet dabei nur mehrere Elemente aus U zugleich; in der zweiten Verallgemeinerung werden auch mehrere Elemente aus Z zugleich betrachtet.

Die Verallgemeinerung geschieht dabei mit Hilfe des Konzepts der Potenzmenge (Reinhardt und Soeder 1974). Eine *Potenzmenge* $\mathcal{P}(M)$ einer Menge M ist die Menge aller Teilmengen von M . Für $M = \{1, 2, 3\}$ ist $\mathcal{P}(M) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$. Man beachte, dass die leere Menge \emptyset und die Menge M selber Elemente von $\mathcal{P}(M)$ sind. Für eine endliche Menge mit n Elementen besteht ihre Potenzmenge aus 2^n Elementen (Bauer und Goos 1992).

Des Weiteren ist das Konzept der Partition nützlich zur Betrachtung der allgemeinen Generalisierungsfunktionen: Eine *Partition* (auch *Zerlegung*, *Klasseneinteilung* oder *Faserung*) P einer Menge M ist ein System von nichtleeren, paarweise disjunkten Teilmengen von M , deren Vereinigung M ist. Für eine Partition P einer Menge M muss also gelten:

$$M = \bigcup_{A \in P} A \quad \text{und} \quad \text{für alle } A_1, A_2 \in P \text{ mit } A_1 \neq A_2 \text{ ist } A_1 \cap A_2 = \emptyset.$$

Eine Partition teilt die Potenzmenge einer Menge in disjunkte Teilmengen so auf, dass jedes Element der Menge in genau einer dieser Teilmengen enthalten ist. Eine Partition P zu einer Menge M ist echte Teilmenge der Potenzmenge von M : $P \subset \mathcal{P}(M)$. Sowohl die Partition als auch die Potenzmenge haben Teilmengen von M als Elemente. Für die Menge $M = \{1, 2, 3\}$ lassen sich fünf verschiedene Partitionen bilden: $P_1 = \{\{1\}, \{2\}, \{3\}\}$, $P_2 = \{\{1, 2\}, \{3\}\}$, $P_3 = \{\{1, 3\}, \{2\}\}$, $P_4 = \{\{1\}, \{2, 3\}\}$, $P_5 = \{\{1, 2, 3\}\}$.

Erste allgemeine Form der Generalisierungsfunktion

Die *erste allgemeine Generalisierungsfunktion* ist eine Abbildung *aus der Potenzmenge von U auf Z* :

$$f_{\text{allg}_1} : \mathcal{P}(U) \longrightarrow Z.$$

Mit dieser Form der Generalisierungsfunktion können Aggregationen treffender ausgedrückt werden. Die Funktion bildet die Menge der zu aggregierenden Elemente auf das Aggregat ab. In unserem Beispiel aus Abbildung 4.8 werden das Waldstück a und das Grünland e aggregiert zu dem Waldstück a' . Die erste allgemeine Generalisierungsfunktion bildet dazu die Menge $\{a, e\}$ ab auf a' . Die übrigen Abbildungen aus diesem Beispiel sind Abbildungen aus einelementigen Teilmengen von U auf jeweils ein Element aus Z , insgesamt lautet die Abbildungsvorschrift also $f_{\text{allg}_1}(\{a, e\}) = a'$, $f_{\text{allg}_1}(\{b\}) = b'$ und $f_{\text{allg}_1}(\{c\}) = c'$.

Im Allgemeinen wird jedes Objekt aus U auch mit dieser Form der Abbildung nur einmal auf Z abgebildet, allerdings als Element einer Teilmenge von U . Das heißt, dass die abzubildenden Teilmengen im Allgemeinen disjunkt sind, mithin eine Partition von U bilden. Im Beispiel besteht die Partition aus folgenden Teilmengen von U : $\{a, e\}$, $\{b\}$, $\{c\}$ und $\{d\}$, wobei die letzte Teilmenge nicht abgebildet wird, weil die Fabrik bei der Generalisierung fortgelassen wird. Allerdings müssen die Teilmengen nicht zwingend eine Partition bilden. Im Unterabschnitt 4.3.4 findet sich hierzu ein Beispiel.

Von der ersten allgemeinen Generalisierungsfunktion gelangt man durch folgende Spezialisierung zur speziellen Generalisierungsfunktion: Man betrachte die Menge der einelementigen Teilmengen von U , die so genannten *Einermengen* (Halmos 1976), welche eine Partition von U bilden. Wenn man die Einermengen mit ihren Elementen identifiziert (einem in der Mengenlehre häufig anzutreffenden Vorgang), so erhält man die spezielle Generalisierungsfunktion.

Zweite allgemeine Form der Generalisierungsfunktion

Die *zweite allgemeine Generalisierungsfunktion* ist eine Abbildung *aus der Potenzmenge von U in die Potenzmenge von Z* :

$$f_{\text{allg}_2} : \mathcal{P}(U) \longrightarrow \mathcal{P}(Z).$$

Diese Form der Generalisierungsfunktion ermöglicht es, Situationen darzustellen, in denen nicht klar ist, welches Geodatum aus welchem Geodatum erzeugt wurde. Als Beispiel sei die Häuserreihe aus den Abbildungen 4.6 und 4.7 genannt. Die Clusterbildung in 4.7 (d) kann mit dieser Form der Generalisierungsfunktion ausgedrückt werden, indem die Menge der zehn ungeneralisierten Häuser auf die Menge der sechs generalisierten Häuser abgebildet wird.

Ebenso wie bei der ersten allgemeinen Generalisierungsfunktion gibt es hier im Allgemeinen eine Partitionierung der Urmenge, von der abgebildet wird. Zusätzlich wird es im Allgemeinen so sein, dass die Teilmengen aus Z , auf die abgebildet wird, eine Partitionierung von Z bilden. Das bedeutet, dass jedes generalisierte Geodatum in nur einer Teilmenge von Z enthalten ist, auf die mit f_{allg_2} abgebildet wird.

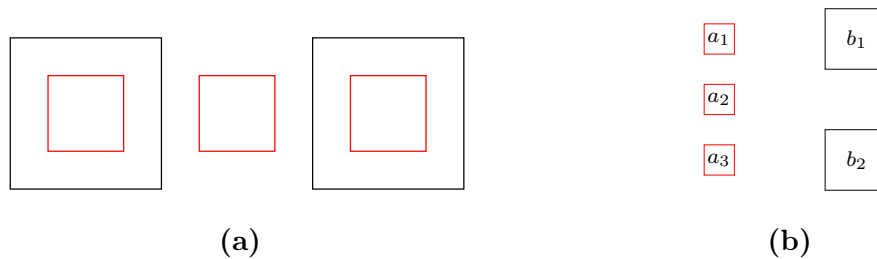


Abbildung 4.9: Zahlenmäßige Verringerung von drei Häusern zu zwei Häusern: (a) Originalbild mit Positionierung (aus SGK 1980), (b) Zur Verdeutlichung gedrehte und verschobene Darstellung mit Bezeichnung.

Das obige Beispiel der Häuserreihen zeigt eine Gefahr, die mit dieser Form der Generalisierung einhergeht, nämlich eine zu starke Verallgemeinerung. Es ist durchaus eine korrekte Vorgehensweise, die gesamte Urmenge auf die gesamte Zielmenge abzubilden. Im Beispiel aus Abbildung 4.8 könnte man die Funktion durch eine einzelne Abbildungsvorschrift beschreiben, in der die Menge $\{a, b, c, d, e\}$ (also die Urmenge U als Ganzes) auf die Menge $\{a', b', c', d'\}$ (also die Zielmenge Z als Ganzes) abgebildet wird. Damit hat die Funktion aber eine triviale Aussage bekommen. Ziel muss es hingegen sein, eine möglichst weitgehende Differenzierung, die kartographisch vertretbar ist, zu finden.

Die zweite allgemeine Generalisierungsfunktion kann zu der ersten allgemeinen Generalisierungsfunktion spezialisiert werden. Hierzu bildet man diejenige Partition zu Z , die aus den Einermengen von $\mathcal{P}(Z)$ besteht, und identifiziert diese Mengen mit ihren Elementen. Damit erhält man die erste allgemeine Form. Durch die weiter oben beschriebene Methode kann man durch weitere Spezialisierung zur speziellen Form der Generalisierungsfunktion gelangen.

4.3.4 Beispiel: Zahlenmäßige Verringerung dreier Häuser

Dieser Unterabschnitt verdeutlicht an einem Beispiel ausführlich die Unterschiede in den drei Modellierungen der Generalisierungsfunktion, der speziellen und den beiden allgemeinen Generalisierungsfunktionen. Das hier gewählte Beispiel der zahlenmäßigen Verringerung von drei Häusern zu zwei Häusern (Abbildung 4.9) ist ebenso wie das Beispiel der Häuserreihen aus Abbildung 4.6 der Publikation der Schweizer Kartographischen Gesellschaft (SGK 1980) entnommen. Bewusst wurde hier ein Beispiel mit nur drei zu generalisierenden Häusern gewählt, um die Anzahl der Möglichkeiten im überschaubaren Rahmen zu halten.

Die entscheidende Fragen in diesem Kontext aus kartographischer Sicht ist die Frage nach dem mittleren der drei Häuser: Wurde dieses Haus bei der Generalisierung fortgelassen oder wurde es dem linken oder rechten Haus oder auch beiden zugeschlagen? Wurden

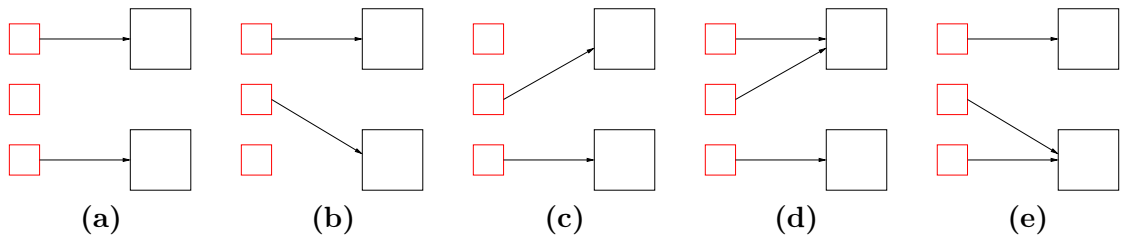


Abbildung 4.10: Beispiele für die spezielle Generalisierungsfunktion.

alle drei Häuser als eine Einheit zu zwei Häusern generalisiert? Je nach Antwort kann eines der folgenden Beispiele von Funktionsvorschriften gewählt werden.

Fünf mögliche Modellierungen der Generalisierung mit Hilfe der *speziellen Generalisierungsfunktion* zeigt Abbildung 4.10. Die erste Möglichkeit (a) bildet die äußeren beiden Häuser auf die generalisierten Häuser ab und lässt das mittlere Haus fort. Die zweite (b) und dritte (c) Möglichkeit bildet das mittlere Haus auf ein generalisiertes Haus ab und lässt jeweils ein äußeres Haus fortfallen. Die vierte (d) und fünfte (e) Möglichkeit fasst das mittlere Haus jeweils mit einem der beiden äußeren Häuser zusammen.

Die Abbildungsvorschriften dieser fünf Möglichkeiten finden sich in folgender Tabelle (die Abkürzung *n.d.* steht für: nicht definiert):

(a)	$a_1 \mapsto b_1$	$a_2 \mapsto \text{n.d.}$	$a_3 \mapsto b_2$
(b)	$a_1 \mapsto b_1$	$a_2 \mapsto b_2$	$a_3 \mapsto \text{n.d.}$
(c)	$a_1 \mapsto \text{n.d.}$	$a_2 \mapsto b_1$	$a_3 \mapsto b_2$
(d)	$a_1 \mapsto b_1$	$a_2 \mapsto b_1$	$a_3 \mapsto b_2$
(e)	$a_1 \mapsto b_1$	$a_2 \mapsto b_2$	$a_3 \mapsto b_2$

Die Modellierung der Generalisierung mit Hilfe der *ersten allgemeinen Generalisierungsfunktion* ist in Abbildung 4.11 anhand von vier Beispielen dargestellt. Zunächst können alle oben genannten Beispiele verwendet werden, indem man die drei Häuser a_1 , a_2 und a_3 mit ihren Einermengen $\{a_1\}$, $\{a_2\}$ und $\{a_3\}$ identifiziert. Für die Möglichkeit aus Abbildung 4.10 (a) ist dies in Abbildung 4.11 (a) beispielhaft dargestellt. (b) und (c) stellen das Zusammenfassen von je zwei benachbarten Häusern dar. Dies entspricht den Möglichkeiten aus Abbildung 4.10 (d) und (e), macht die Aggregation deutlicher, außerdem verringert sich damit die Zahl der Abbildungsvorschriften von jeweils drei auf zwei. Die drei Möglichkeiten (a), (b) und (c) zeigen deutlich, wie die Urmenge, bestehend aus a_1 , a_2 und a_3 jeweils partitioniert wurde. Möglichkeit (d) zeigt eine Abbildungsvorschrift, bei der die Urmenge *nicht* partitioniert wurde. Hier wird das mittlere Haus sowohl mit dem einen als auch dem anderen äußeren Haus aggregiert.

Die folgende Tabelle zeigt die Abbildungsvorschriften der vier vorgestellten Möglichkeiten

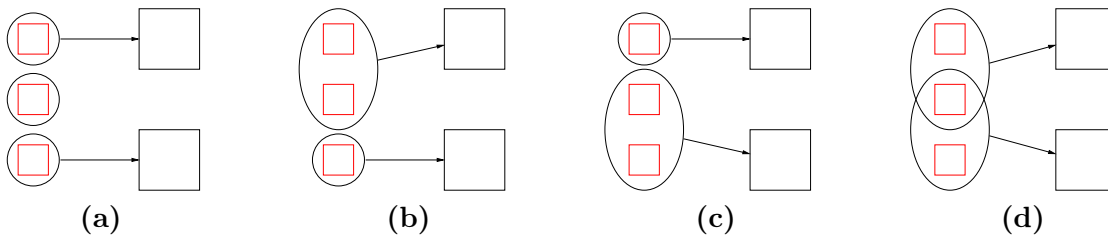


Abbildung 4.11: Beispiele für die erste allgemeine Generalisierungsfunktion.

aus Abbildung 4.11:

$$\begin{array}{lll}
 \text{(a)} & \{a_1\} \mapsto b_1 & \{a_2\} \mapsto \text{n.d.} & \{a_3\} \mapsto b_2 \\
 \text{(b)} & & \{a_1, a_2\} \mapsto b_1 & \{a_3\} \mapsto b_2 \\
 \text{(c)} & \{a_1\} \mapsto b_1 & & \{a_2, a_3\} \mapsto b_2 \\
 \text{(d)} & & \{a_1, a_2\} \mapsto b_1 & \{a_2, a_3\} \mapsto b_2
 \end{array}$$

Abbildung 4.12 zeigt vier Beispiele für die Modellierung der Generalisierung der drei Häuser durch die *zweite allgemeine Generalisierungsfunktion*. Zunächst können alle Möglichkeiten der ersten allgemeinen Generalisierungsfunktion auch mit der zweiten allgemeinen Generalisierungsfunktion ausgedrückt werden, indem die generalisierten Häuser b_1 und b_2 mit ihren Einermengen $\{b_1\}$ und $\{b_2\}$ identifiziert werden. Somit umschließt die zweite allgemeine Generalisierungsfunktion auch alle Möglichkeiten der speziellen Generalisierungsfunktion. Die Möglichkeit (a) zeigt das auf diese Weise verallgemeinerte Beispiel aus Abbildung 4.11 (b). In (b) ist die Möglichkeit visualisiert, die Häuser der Urmenge als Gesamtes in die Häuser der Zielmenge als Gesamtes zu generalisieren. Bei einem Beispiel mit drei Häusern ist dies sicherlich eine nahe liegende Sichtweise. Es sei an die Gefahr erinnert, die sich mit diesem gesamt betrachtenden Ansatz der Generalisierung verbindet, dass nämlich nicht mehr zwischen den Identitäten der Häuser (hier zum Beispiel der beiden äußeren) differenziert wird.

In den Varianten (a) und (b) sind beide möglichen Partitionierungen der Zielmenge Z aufgeführt. In Variante (c) ist die Zielmenge nicht partitioniert: Hier werden die beiden Einermengen der äußeren Häuser auf die Einermengen der generalisierten Häuser abgebildet, und das mittlere Haus auf die Menge beider generalisierter Häuser. Die letzte Variante, (d), zeigt eine kompliziertere Interpretation der Generalisierung der drei Häuser. Hier sind Abbildungsvorschriften sowohl für die drei Einermengen, als auch für zwei Mengen mit je zwei Elementen als auch für die Menge aller Häuser vorhanden. Diese Möglichkeit ist eher theoretischer Natur, zeigt aber die Mächtigkeit des hier vorgestellten Konzepts der zweiten allgemeinen Generalisierungsvorschrift.

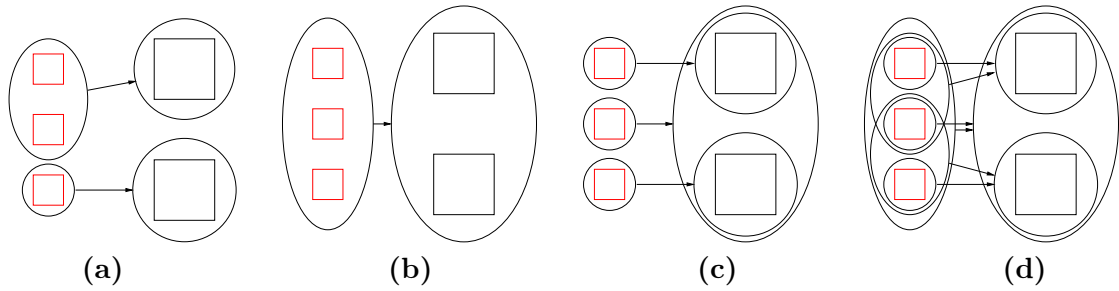


Abbildung 4.12: Beispiele für die zweite allgemeine Generalisierungsfunktion.

Die folgende Tabelle umfasst alle Abbildungsvorschriften aus den vier Beispielen aus Abbildung 4.12:

(a)	$\{a_1, a_2\} \mapsto \{b_1\}$	$\{a_3\} \mapsto \{b_2\}$	
(b)	$\{a_1, a_2, a_3\} \mapsto \{b_1, b_2\}$		
(c)	$\{a_1\} \mapsto \{b_1\}$	$\{a_2\} \mapsto \{b_1, b_2\}$	$\{a_3\} \mapsto \{b_2\}$
(d)	$\{a_1\} \mapsto \{b_1\}$	$\{a_2\} \mapsto \{b_1, b_2\}$	$\{a_3\} \mapsto \{b_2\}$
	$\{a_1, a_2\} \mapsto \{b_1\}$	$\{a_1, a_2, a_3\} \mapsto \{b_1, b_2\}$	$\{a_2, a_3\} \mapsto \{b_2\}$

Diskussion der drei Formen der Generalisierungsfunktion

Die drei Formen der Generalisierungsfunktionen bieten jede für sich einen in sich schlüssigen und konsistenten Ansatz zur Formalisierung der Generalisierung. Je allgemeiner die Form ist, desto mächtiger ist ihre Ausdrucksweise, das heißt, desto mehr unterschiedliche Arten gibt es, denselben Sachverhalt auszudrücken. Mit der Mächtigkeit steigt aber auch die Kompliziertheit in der Handhabung der Formalisierungen. So entspricht die spezielle Generalisierungsfunktion der „intuitiven“ Beschreibung der Generalisierung am ehesten, während die zweite allgemeine Generalisierungsfunktion eher wie ein künstliches Konstrukt erscheint. Dennoch hat jede Form der Generalisierungsfunktion ihre Daseinsberechtigung, weil mit jeder Form bestimmte Sachverhalte am besten ausgedrückt werden können.

Für das weitere Vorgehen in dieser Arbeit ist die *spezielle Generalisierungsfunktion* ausschlaggebend. Die weiteren Formalisierungen und Abstraktionen der Generalisierung greifen auf diese Form der Generalisierungsfunktion zurück, soweit es möglich und sinnvoll erscheint, auch, um die Formalisierungen so einfach wie möglich zu halten. In solchen Fällen, in denen die spezielle Generalisierungsfunktion nicht ausreicht, wird die Formalisierung auf die erste oder zweite allgemeine Generalisierungsfunktion erweitert.

4.3.5 Vergleich mit der Arbeit von Ai und van Oosterom

In der Literatur wurde bisher nur ein Ansatz vorgestellt, der eine ähnliche Ausrichtung hat wie der hier Vorgestellte. In der Arbeit von Ai und van Oosterom (2001) ist das Modellieren der Generalisierung mit der mathematischen Ausdrucksweise der Relationen und Strukturen als Teilgebiet der Algebra ebenso wie hier ein Teilziel. Dieser Abschnitt zeigt Gemeinsamkeiten und Unterschiede der beiden Ansätze in Hinblick auf die Generalisierungsfunktion auf.

Ai und van Oosterom (2001) teilen Mengen von Geodaten (*geographic data sets*) in zwei Bestandteile auf, in Geodaten (*spatial entities*) und Relationen zwischen Geodaten (*spatial entity relationships*). Die Funktion zwischen Mengen von Geodaten teilen sie analog auf in eine Funktion zwischen den Geodaten selber (*spatial entity mapping*) und eine weitere Funktion zwischen den Relationen zwischen den Geodaten (*spatial relationship mapping*). Die erstgenannte Funktion entspricht der hier vorgestellten Generalisierungsfunktion und wird im Folgenden genauer betrachtet:

Ai und van Oosterom bezeichnen die Generalisierungsfunktion ebenfalls als eine Funktion aus einer Menge von Geodaten in eine Menge von Geodaten, $f : E_1 \rightarrow E_2$. Sie teilen diese Funktion auf in drei Unterfunktionen f_1 , f_2 und f_3 , die Abbildungen entsprechen sollen, die $1:1$, $n:1$ und $n:m$ abbilden. Die Formalisierung drücken sie aus durch die Abbildungsvorschriften für jede dieser drei Unterfunktionen:

$$\begin{aligned} 1:1\text{-Abbildung} : & \quad e' = f_1(e) \\ n:1\text{-Abbildung} : & \quad e' = f_2(e_1, e_2, \dots, e_i) \\ n:m\text{-Abbildung} : & \quad (e'_1, e'_2, \dots, e'_j) = f_3(e_1, e_2, \dots, e_i) \end{aligned}$$

Auf den ersten Blick erscheint diese Definition über die Abbildungsvorschrift analog zu der hier entwickelten Einteilung in die spezielle Generalisierungsfunktion f_{spez} , die erste allgemeine Generalisierungsfunktion f_{allg_1} und die zweite allgemeine Generalisierungsfunktion f_{allg_2} . In drei Punkten unterscheiden sich jedoch diese Ansätze:

1. Ai und van Oosterom berücksichtigen nicht das *Fortlassen* von Mengenelementen, was in Abschnitt 4.2.4 auch mit $1:0$ -Beziehung bezeichnet wurde, ansonsten durch das Verständnis der Generalisierungsfunktion als *partielle Funktion* ausgedrückt wurde.
2. Ai und van Oosterom berücksichtigen ebenfalls nicht, dass mit der ersten Unterfunktion f_1 nicht nur $1:1$ -Relationen modelliert werden können, sondern auch $n:1$ -Relationen mit $n > 1$. Beispiele dafür sind die Aggregation von Häusern in Abbildung 4.7 (c), der Zusammenschluss von Wald- und Grünfläche in Abbildung 4.8 und die Aggregation von Häusern in den Abbildungen 4.10 (d) und (e), die alle mit der speziellen Generalisierungsfunktion modelliert werden können. Dies impliziert, dass die spezielle Generalisierungsfunktion nicht injektiv ist (siehe Ab-

schnitt 4.4.2). Wenn man die Injektivität der Funktion fordert, dann ist f_1 eine echte 1:1-Abbildung.

3. Die Beschreibung der Funktion durch die Abbildungsvorschriften, insbesondere die Vorschriften für f_2 und f_3 entsprechen nicht den mathematischen Formalisierungen für Funktionen. Ai und van Oosterom definieren mehrstellige Funktionen, ohne die Stelligkeit anzugeben. Die Definition der Unterfunktionen lautet implizit:

$$\begin{aligned}f_1 &: E_1 \longrightarrow E_2 \\f_2 &: E_1 \times E_1 \times \dots \times E_1 \longrightarrow E_2 \\f_3 &: E_1 \times E_1 \times \dots \times E_1 \longrightarrow E_2 \times E_2 \times \dots \times E_2\end{aligned}$$

Allerdings ist die Stelligkeit der kartesischen Produkte nicht fix. Damit ist eine solche Ausdrucksweise formal nicht korrekt. Dieses Problem wurde in dieser Arbeit ebenfalls erkannt und durch das Konzept der Potenzmengen gelöst.

Weitere Unterschiede bestehen in der Behandlung der Relationen und Strukturen innerhalb der Mengen von Geodaten. Diese werden in Abschnitt 5.1.1 weiter unten behandelt.

4.3.6 Vergleich von Relation und Funktion

In der Mathematik ist die Funktion ein Spezialfall einer Relation, somit ist jede Funktion gleichzeitig auch eine Relation. Eine Relation bildet eine *Zuordnung* zwischen zwei (oder mehr) Mengen, während eine Funktion Elemente der einen Menge auf Elemente der anderen Menge *abbildet*.

Bezogen auf die Generalisierung und die Betrachtung von zwei Mengen von Geodaten, einer ungeneralisierten und einer generalisierten Menge, ist eine Relation zwischen diesen beiden Mengen zunächst auch eine Zuordnung zwischen den Elementen beider Mengen. Damit beschreibt die Relation etwas *Bestehendes* zwischen den beiden Mengen. Die Funktion hingegen ordnet die Elemente zwischen den Mengen nicht nur zu, sondern bildet die Elemente auch ab, und zwar in einer bestimmten Richtung, nämlich von der ungeneralisierten Menge zur generalisierten Menge. Damit hat die Funktion einen *hinweisenden* und *erzeugenden* Charakter, sie *berechnet* quasi die generalisierten Elemente der Zielmenge aus den ungeneralisierten Elementen der Urmenge.

Auch aus diesem Grund gibt eine Funktion den Charakter der Generalisierung besser wieder als eine Relation und wird hier für die Beschreibung der Generalisierung bevorzugt.

4.4 Mathematische Charakterisierung der Funktion

In diesem Abschnitt werden einige einfache Eigenschaften der verschiedenen Formen der Generalisierungsfunktion diskutiert. Abgeleitet sind diese Eigenschaften aus Beobachtungen. Sie werden zunächst mathematisch formuliert und erläutert, um dann kartographisch interpretiert zu werden.

Die Eigenschaften werden jeweils zunächst für die spezielle Generalisierungsfunktion untersucht und danach, soweit möglich, auf die allgemeinen Generalisierungsfunktionen übertragen.

In diesem Abschnitt werden folgende drei Eigenschaften der Generalisierungsfunktion diskutiert:

- Nicht-Totalität oder Partiellität (Unterabschnitt 4.4.1)
- Nicht-Injektivität (Unterabschnitt 4.4.2)
- Surjektivität (Unterabschnitt 4.4.3)

4.4.1 Nicht-Totalität

Bereits in Abschnitt 4.2.4 wurde von der *Linkstotalität* von Relationen gesprochen. Für Funktionen ist die äquivalente Eigenschaft die *Totalität*. Eine Funktion $f : U \rightarrow Z$ ist *total*, wenn gilt:

$$\forall x \in U : \exists y \in Z : f(x) = y.$$

Eine Funktion, die diese Bedingung *nicht* erfüllt, heißt *partiell*.

Eine Funktion ist also dann partiell, wenn es Elemente von U gibt, die nicht durch f auf ein Element aus Z abgebildet werden.

Die Funktion aus dem Beispiel in Abbildung 4.8 ist deshalb partiell, weil die Fabrik d nicht in die Zielmenge abgebildet wird. Von den fünf Möglichkeiten der zahlenmäßigen Verringerung von Häusern in Abbildung 4.10 sind die ersten drei Funktionsvorschriften (a-c) partiell, die letzten beiden (d-e) total.

Die Entsprechung in der Kartographie ist hierbei leicht zu sehen: Wenn mindestens ein Objekt bei der Generalisierung fortgelassen wird, so ist die Funktion partiell, andernfalls total. Da das Wesen der Generalisierung aber die Auswahl, mithin das Fortlassen von Objekten beinhaltet, ist eine Generalisierungsfunktion im Allgemeinen immer partiell.

Betrachtet man die allgemeinen Formen der Generalisierungsfunktion, so ist die Charakterisierung über die Totalität zunächst trivial: Jede Generalisierung ist hier eine partielle Funktion, denn für eine totale Funktion müsste *jede* Teilmenge der Urmenge auf die Zielmenge abgebildet werden, und damit auch die leere Menge. Diese Menge jedoch abzubilden, ergibt kartographisch keinen Sinn.

Bei Abbildungen aus der Potenzmenge einer Menge kann man die Totalität jedoch auch anders verstehen, indem man nur fordert, dass jedes Element der Urmenge in mindestens einer Teilmenge der Urmenge enthalten sein muss, die auf die Zielmenge beziehungsweise die Potenzmenge der Zielmenge abgebildet wird, dass also gilt:

$$\forall x \in U : \exists X \subseteq U, y \in Z : x \in X \wedge f_{\text{allg}_1}(X) = y$$

beziehungsweise

$$\forall x \in U : \exists X \subseteq U, Y \subseteq Z : x \in X \wedge f_{\text{allg}_2}(X) = Y.$$

Mit dieser Formulierung der Totalität für Funktionen, die aus einer Potenzmenge abbilden, gelten die gleichen Betrachtungen wie für die spezielle Generalisierungsfunktion. Denn ein Element der Urmenge, das bei der Generalisierung fortgelassen wird, sollte in keiner abgebildeten Teilmenge der Urmenge vorhanden sein.

4.4.2 Nicht-Injektivität

Injektivität einer Funktion entspricht der Linkseindeutigkeit bei Relationen. Formal heißt eine Funktion $f : U \rightarrow Z$ *injektiv*, wenn gilt:

$$\text{Für alle } x_1, x_2 \in U \text{ gilt: Aus } f(x_1) = f(x_2) \text{ folgt } x_1 = x_2.$$

Injektivität bedeutet also, dass zu jedem Element y aus der Zielmenge höchstens ein Element der Urmenge existiert, das auf y abgebildet wird. Anders ausgedrückt dürfen nicht mehr als ein Element der Urmenge auf ein und dasselbe Element der Zielmenge abgebildet werden.

In den bisher behandelten Beispielen ist diese Bedingung jedoch meistens nicht erfüllt. In Abbildung 4.8 auf Seite 78 ist $f(a) = f(e)$, aber $a \neq e$. In Abbildung 4.10 ist in (d) und (e) das Haus a_2 jeweils auf dasselbe Haus abgebildet wie a_1 beziehungsweise a_3 . Die Generalisierungsfunktion ist im Allgemeinen also nicht injektiv.

Aus kartographischer Sicht ist die Nicht-Injektivität immer dann gegeben, wenn Objekte zusammengefasst (aggregiert) werden. Im ersten Beispiel sind es das Waldstück und die Grünfläche, die aggregiert werden, im zweiten Beispiel sind es jeweils zwei Häuser, die zusammengefasst werden.

Für die erste allgemeine Generalisierungsfunktion kann der Injektivitätsbegriff übernommen werden. Der einzige Unterschied ist, dass die Urmenge aus Mengen von Geodaten besteht, an Stelle von Geodaten selber. Aggregationen werden bei dieser Funktion durch Abbilden von Teilmengen modelliert, siehe Abbildung 4.11 (b) und (c). Somit können alle Aggregationen, die bei der speziellen Generalisierungsfunktion zur Nicht-Injektivität führen, mit Hilfe der ersten allgemeinen Generalisierungsfunktion zu einer injektiven Form der Funktion führen. Dies ist nicht zwingend so, denn, wie in Abschnitt 4.3.4

dargestellt, kann durch Identifikation der Instanzen mit ihren Einermengen jede Abbildungsvorschrift der speziellen Generalisierungsfunktion in die erste allgemeine Form der Generalisierungsfunktion überführt werden. Damit können auch die *nicht-injektiven* Abbildungsvorschriften in die erste allgemeine Form überführt werden und bleiben dabei natürlich nicht-injektiv. Die oben genannte Variante durch Abbilden der Teilmengen ist jedoch weitaus eleganter.

Bei der zweiten allgemeinen Generalisierungsfunktion gilt für die Aggregation gleiches wie bei der ersten allgemeinen Form. Jedoch werden hier Teilmengen der Zielmenge betrachtet, und man kann nun schärfer fordern, dass jedes Element der Zielmenge zu höchstens einer Teilmenge gehört, auf die abgebildet wird, also:

Für alle $X_1, X_2 \subseteq U$ gilt: Aus $y \in f(X_1)$ und $y \in f(X_2)$ folgt $X_1 = X_2$.

Diese schärfere Form der Injektivität wird in Abbildung 4.12 nur von den Beispielen (a) und (b) erfüllt. Beispiel (c) ist in diesem Sinne nicht injektiv; und Beispiel (d) ist weder in diesem noch im herkömmlichen Sinne injektiv.

4.4.3 Surjektivität

Eine Funktion $f : U \longrightarrow Z$ ist *surjektiv*, wenn gilt:

$$\forall y \in Z : \exists x \in U : f(x) = y.$$

Surjektivität bedeutet, dass für alle Elemente der Zielmenge ein Element der Urmenge existiert, das auf das Zielelement abgebildet wird. Somit ist die Zielmenge gleich der *Bildmenge* der Funktion.

Mit dem Verständnis der Funktion als etwas Produzierendes (siehe Unterabschnitt 4.3.6) heißt dies, dass alle Elemente der Zielmenge durch Abbildung aus der Urmenge entstanden sind.

In den Beispielen der Abbildungen 4.8 und 4.10 hat jedes Element in Z ein Urbild in U . Die Abbildungsvorschriften der Beispiele sind somit surjektiv.

Kartographisch interpretiert bedeutet dies, dass in einem generalisierten Datenbestand nur Geodaten enthalten sind, die in dem ungeneralisierten Datenbestand bereits vorhanden waren oder von diesen abgeleitet wurden. Eine Generalisierung ist damit eine surjektive Abbildung.

Die Aussage des letzten Abschnittes erscheint vielleicht auf den ersten Bild trivial, da eine Generalisierung immer von einem bestehenden Datenbestand ausgeht und aus diesem einen neuen Datenbestand erzeugt. Bei einer automatischen Generalisierung trifft dies zu, weil nur die Ausgangsdaten als Eingabe für die Generalisierung herhalten. Für die manuelle Generalisierung zum Beispiel von topographischen Karten muss die Surjektivität jedoch nicht zwingend gelten, denn hierbei können für die Generalisierung auch

Kapitel 4 Generalisierung als Funktion

zusätzliche Daten hinzugezogen werden, die nicht in den Ausgangskarten enthalten sind. Ein Beispiel hierfür zeigt Abbildung 4.13: Links ist ein Ausschnitt aus einer topographischen Karte aus Rheinland-Pfalz im Maßstab 1:50 000 gezeigt, der einen Teil der Eifel darstellt. Weder in diesem Ausschnitt noch in irgend einem anderen Ausschnitt einer topographischen Karte dieses Maßstabes findet man die Bezeichnung EIFEL für das gleichnamige Mittelgebirge. Aus den topographischen Karten 1:50 000 sind jedoch Karten kleineren Maßstabes abgeleitet, so auch in mehreren Schritten die Karte Bundesrepublik 1:1 Mio. Diese enthält aber die Bezeichnung EIFEL, obwohl diese Information in keiner entsprechenden Karte des Ausgangsmaßstabes enthalten ist. Es wurde also in dem Schritt der Generalisierung Information *hinzugefügt*. Diese Generalisierung ist damit nicht surjektiv. Für die Surjektivität der automatischen Generalisierung ist dies im Weiteren jedoch nicht von Belang.

Für die erste allgemeine Generalisierungsfunktion gilt in Bezug auf die Surjektivität dasselbe wie für die spezielle Generalisierungsfunktion. Dies liegt darin begründet, dass sowohl die spezielle als auch die erste allgemeine Generalisierungsfunktion auf Z abbilden.

Die zweite allgemeine Generalisierungsfunktion ist im Allgemeinen nicht surjektiv im oben genannten Sinne, denn dazu müssten alle Teilmengen der Zielmenge Z ein Urbild besitzen, auch zum Beispiel die leere Menge, was kartographisch keinen Sinn ergeben würde. Verschärft man die Definition der Surjektivität jedoch ähnlich wie im vorigen Unterabschnitt die Definition der Injektivität, dann ergibt sich, dass diese Form der Generalisierungsfunktion ebenfalls surjektiv ist.

Die schärfere Form der Surjektivität einer Funktion mit der Potenzmenge einer Men-

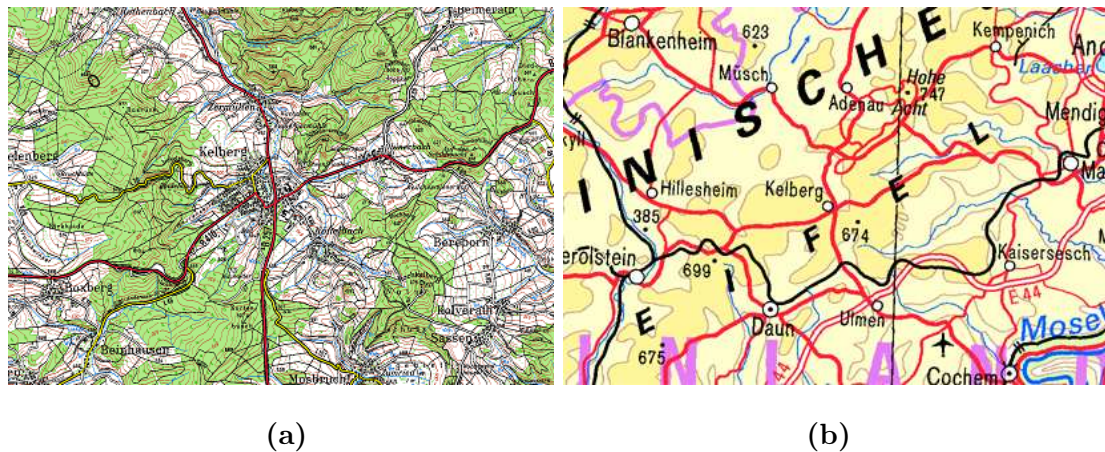


Abbildung 4.13: Nicht-Surjektivität bei manueller Generalisierung: Die Bezeichnung EIFEL findet sich erst in Karten mit kleinem Maßstab: (a) Ausschnitt aus einer TK50 (verkleinert), (b) Ausschnitt aus der Karte Bundesrepublik 1:1 Mio (verkleinert).

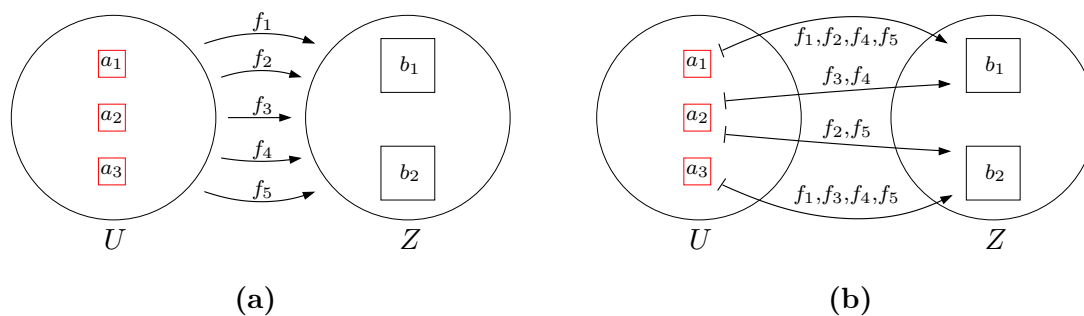


Abbildung 4.14: Die fünf Beispiele der speziellen Generalisierungsfunktion aus Abbildung 4.10 zusammengefasst in einer Abbildung: (a) die fünf Funktionen f_1 bis f_5 , (b) die zugehörigen Abbildungsvorschriften.

ge Z als Zielmenge fordert, dass jedes Element der Zielmenge Element von mindestens einer Teilmenge der Zielmenge ist, auf die abgebildet wird. Formal lautet die Bedingung dann:

$$\forall y \in Z : \exists Y \subseteq Z, \exists X \subseteq U : y \in Y \wedge f(X) = Y.$$

4.5 Nichtdeterminiertheit der Generalisierung

Dieses Kapitel behandelte bisher genau eine Funktion, zusammen mit ihren möglichen Ausprägungen und Eigenschaften. In diesem Abschnitt wird die Betrachtung auf mehrere Funktionen erweitert, wobei jede dieser Funktionen aus derselben Urmenge abbildet, aber jede Funktion eine andere Zielmenge haben kann (aber nicht muss). Anschaulich betrachtet bedeutet dies die Untersuchung von unterschiedlichen Generalisierungen derselben Menge von Geodaten.

Wenn man von *der* Generalisierungsfunktion oder von *der* Abbildungsvorschrift spricht, so meint man, dass sie die Elemente aus einer Urmenge auf Elemente einer Zielmenge abbildet. Diese Betrachtung wird im Folgenden verallgemeinert, indem mehrere Funktionen mit derselben Urmenge gleichzeitig untersucht werden. Zunächst behandeln wir ein Beispiel, in dem mehrere, unterschiedliche Funktionen aus derselben Urmenge in dieselbe Zielmenge abbilden. In dem Beispiel von Abbildung 4.10 auf Seite 82 wurden fünf unterschiedliche Abbildungsvorschriften vorgestellt, mit denen man drei (ungeneralisierte) Häuser auf zwei (generalisierte) Häuser abbilden kann. Jede dieser Abbildungsvorschriften bildet eine Funktion. Nennen wir diese Funktionen f_1 bis f_5 für die Möglichkeiten (a) bis (e). Jede dieser Funktionen bildet von derselben Urmenge U , bestehend aus drei Häusern, auf dieselbe Zielmenge Z , bestehend aus zwei Häusern, ab. Wir können diese fünf Funktionen in einem Bild zusammenfassen: Abbildung 4.14 (a) zeigt diese fünf Funktionen. In Abbildung 4.14 (b) sind alle fünf Abbildungsvorschriften dargestellt, wobei die Beschriftung an jedem Pfeil diejenige Funktion bezeichnet, für die dieser Pfeil Gültigkeit besitzt.

Das Ergebnis einer Generalisierung liegt aber im Allgemeinen nicht im vornherein fest, wie es in diesem Beispiel der Fall ist. Vielmehr ist es Aufgabe eines Kartographen (bei der manuellen Generalisierung) oder eines Programmes (bei der automatischen Generalisierung), aus der Vielzahl von Möglichkeiten, eine Situation zu generalisieren, diejenige Möglichkeit herauszusuchen, die dem gewünschten Ergebnis am nächsten kommt. Anhand von Kriterien wird beurteilt, ob eine Generalisierung richtig oder falsch, gut oder schlecht ist. Die Kriterien zur Beurteilung sind Thema von Kapitel 5. Hier ist jedoch zunächst wichtig, zu erkennen, dass eine Generalisierung derselben Geodaten zu unterschiedlichen Ergebnissen führen kann und dass dies in die Modellierung eingehen muss.

Bei der manuellen Generalisierung liegt es im Ermessen des Kartographen, die richtige oder beste Generalisierungsmöglichkeit zu wählen. Unterschiedliche Personen werden in vielen Fällen auch zu unterschiedlichen Ergebnissen kommen, die dennoch alle korrekt im kartographischen Sinne sein können. Bei der automatischen Generalisierung besteht im Grunde die gleiche Situation: Von einem Programm erwartet man, dass es zu einer Ausgangssituation immer dasselbe Ergebnis liefert, dass also die ausgeführte Funktion *determiniert* ist.⁴ Unterschiedliche Programme, oder ein Programm, das in verschiedenen Läufen unterschiedlich parametrisiert wird, können zu denselben Ausgangsdaten unterschiedliche Ergebnisse liefern. Man kann also sagen, dass die Generalisierung von vornherein *nicht determiniert* ist. Beispiele für unterschiedliche Generalisierungen finden sich besonders im Kontext der Modellgeneralisierung, zum Beispiel bei der Anpassung der Geometrieauflösung (Schürer 2002) oder bei der Flächenzusammenfassung (Bobzien 2001).

Abbildung 4.15 wiederholt das Beispiel aus Abbildung 4.8 von Seite 78 und fügt zu der bereits bekannten Generalisierung zwei weitere mögliche Generalisierungen hinzu. Die Funktion f_1 bildet auf die Zielmenge Z_1 ab, bei der im Vergleich zur Urmenge U keine Fortlassungen und keine Zusammenfassungen auftreten. Die Generalisierung besteht in diesem Fall nur aus einer geringen Vereinfachung der Geometrie. Funktion f_i ist die Funktion aus Abbildung 4.8, ebenso entspricht die Zielmenge Z_i der Zielmenge aus dieser Abbildung. Die Funktion f_k bildet ab auf die Zielmenge Z_k , bei der alle flächenförmigen Objekte zu einer Grünfläche c' zusammengefasst sind. Die Beurteilung dieser unterschiedlichen Generalisierungsfunktionen hängt ab von den Kriterien, die mit dem Maßstab oder dem Strukturierungsgrad der Zielmenge verknüpft sind. Kapitel 5 diskutiert diese Thematik. Die Abbildung verdeutlicht, dass unterschiedliche Zielmengen, und da die Funktionen surjektiv sind, auch unterschiedliche Bildmengen möglich sind. Durch die Punkte wird angedeutet, dass eine weitere, im Allgemeinen unendliche Anzahl von Ziel- beziehungsweise Bildmengen und damit von Generalisierungen möglich ist.

In Abbildung 4.16 erfolgte eine Vereinigung der Zielmengen aus Abbildung 4.15 in eine Menge. Diese Menge $Z_{\text{ges}} := \bigcup_{i \in \mathbb{N}} Z_i$ bezeichnet die Menge aller Geodaten eines Strukturierungsgrades. In der Abbildung kann diese (unendlich große) Menge nur angedeutet

⁴Es existieren auch nichtdeterminierte Programme, die zu gleichen Eingaben und Startbedingungen unterschiedliche Ergebnisse liefern (Claus und Schwill 2001). Solche Programme basieren auf nichtdeterministischen Algorithmen, wie es zum Beispiel stochastische Algorithmen sind. Diese Möglichkeiten werden hier jedoch nicht weiter behandelt.

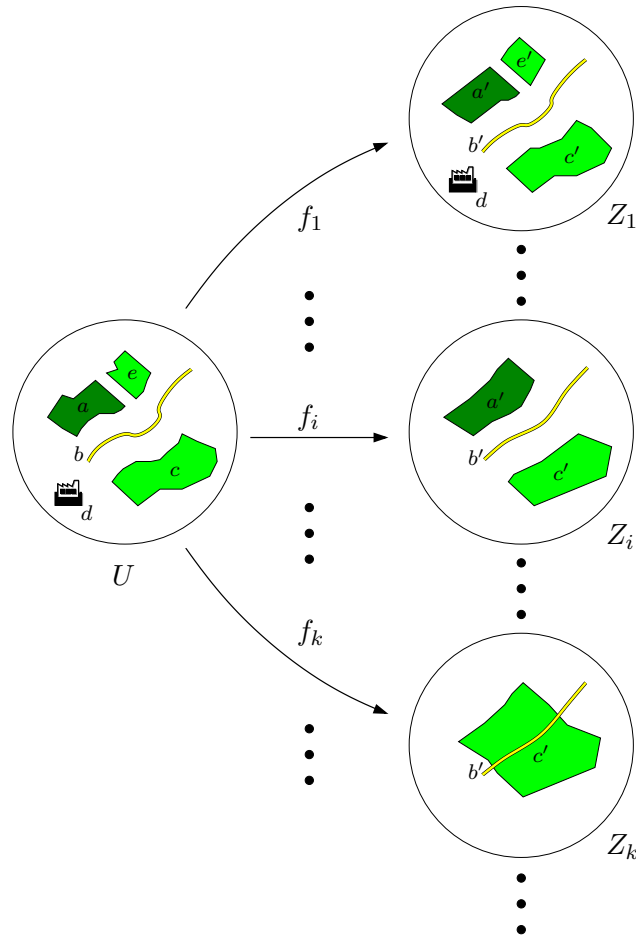


Abbildung 4.15: Das Beispiel für eine Generalisierungsfunktion aus Abbildung 4.8, hier durch f_i gekennzeichnet, erweitert um zwei Generalisierungsfunktionen f_1 und f_k mit unterschiedlichen Ziel- beziehungsweise Bildmengen.

werden. Die Elemente dieser Menge repräsentieren die Realweltobjekte mehrfach. So wird das in der Urmenge durch a repräsentierte Realweltobjekt „Waldstück“ in Z_{ges} unter anderem durch die Elemente a'_1, a'_i und c'_k repräsentiert. Die Geometrien dieser Elemente überlappen sich fast vollständig, was diese Abbildung wiederum nur andeuten kann. Die Funktionen $f_1 \dots f_i \dots f_k \dots$ bilden eine *Familie von Funktionen*. Da diese Funktionen in dieser Interpretation nicht mehr auf eine Zielmenge abbilden, sondern in die Vereinigungsmenge aller Zielmengen, sind sie nicht surjektiv.

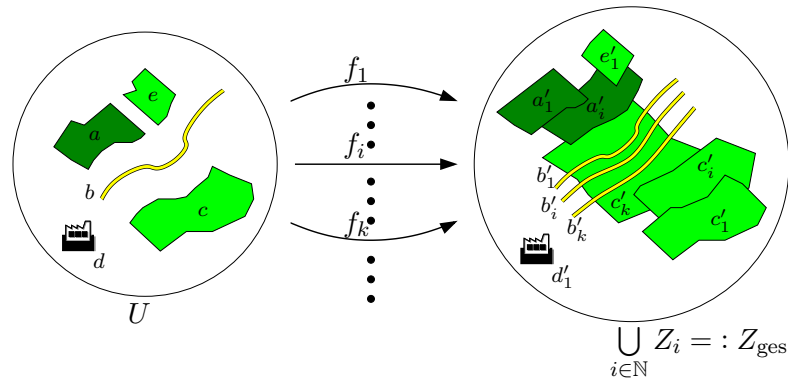


Abbildung 4.16: Die Generalisierungsfunktionen aus Abbildung 4.15, zusammengefasst in der Darstellung mit einer Zielmenge als Vereinigung aller Zielmengen aus Abbildung 4.15.

4.6 Fazit

Dieses Kapitel formalisiert die Generalisierung in Form einer Funktion aus einer ungeneralisierten Menge von Geodaten auf eine generalisierte Menge von Geodaten, ausgehend von einer Relation. Die Generalisierungsfunktion existiert in drei unterschiedlichen Ausprägungen: als spezielle, als erste und als zweite allgemeine Generalisierungsfunktion. Jede dieser drei Generalisierungsfunktionen bietet andere Möglichkeiten zur Modellierung der Generalisierung. Die Funktionen werden durch die Eigenschaften der Nicht-Totalität, der Nicht-Injektivität und der Surjektivität näher charakterisiert, wobei jede dieser Eigenschaften eine Entsprechung in einer kartographischen Interpretation findet.

Tabelle 4.1 stellt die drei Formen der Generalisierungsfunktion einander gegenüber und hebt ihre Unterschiede hervor. In der ersten Spalte findet sich die Bezeichnung der Form, die zweite Spalte enthält ihre mathematische Formulierung.

In der dritten Spalte sind Vor- und Nachteile kurz zusammengefasst. Hierbei wird die Modellierung von Aggregationen und die Möglichkeit, mehrdeutige Abbildungsvorschriften zu modellieren, beurteilt. In dieser Spalte nicht gesondert erwähnt ist, dass die Komplexität der Generalisierungsfunktionen zunimmt, je allgemeiner die Funktion ist. Die Modellierung von Aggregationen gelingt zwar mit allen drei Formen, jedoch bieten die allgemeinen Formen eine elegante Lösung durch Mengenbildung der zu aggregierenden Elemente an, während die spezielle Form die Aggregationen nur implizit modelliert. Mehrdeutige Abbildungen sind nur mit der zweiten allgemeinen Form der Generalisierungsfunktion möglich. Wenn man auf die Eindeutigkeit in der Abbildung verzichten möchte, weil sich in der Situation bei der Generalisierung keine Eindeutigkeit der Zuordnung anbietet, was besonders in der kartographischen Generalisierung auftreten mag, dann muss auf diese Form der Generalisierungsfunktion zurückgegriffen werden.

Die letzte Spalte der Tabelle stellt die mathematischen Charakteristiken der unterschied-

lichen Formen einander gegenüber. Für die Formen der Generalisierungsfunktion, die aus oder auf Potenzmengen abbilden, werden zusätzlich zu den gängigen Definitionen jeweils eine strengere Definition von Nicht-Totalität, Injektivität und Surjektivität gegeben, die sich speziell auf die Eigenschaften der Potenzmenge beziehen. Diese Definitionen sind in der Tabelle *kursiv* und in Klammern berücksichtigt. Die in dieser Spalte angegebenen Eigenschaften gelten jeweils für den „Allgemeinfall“. Wie in Abschnitt 4.4 gezeigt wird, sind bei der Generalisierung für alle aufgezeigten Eigenschaften Ausnahmen denkbar. Diese werden jedoch üblicherweise nur in besonderen Fällen auftreten und geben nicht die „typische“ Generalisierung wieder. Beispiele für solche Ausnahmen sind Generalisierungen, bei der kein Geodatum wegfällt, die Generalisierung also total ist, oder Generalisierungen, bei der keine Geodaten zusammengefasst werden, wodurch die Generalisierung auch in der speziellen Form injektiv ist.

Abschnitt 4.5 behandelt die Frage, wie man unterschiedliche Generalisierungsfunktionen aus derselben Urmenge zu bewerten hat, das heißt, wie die jeweiligen Bildmengen und Abbildungsvorschriften auf ihre Korrektheit und ihre Güte hin beurteilt werden können. Für diese Bewertung und Beurteilung benötigt man bestimmte Kriterien. Diese Kriterien sind unter anderem Thema des folgenden Kapitels.

Bezeichnung	mathematische Formulierung	Vor- und Nachteile	Charakteristiken
spezielle Form	$f_{\text{spez}} : U \longrightarrow Z$	Aggregationen: <i>implizit</i> mehrdeutige Abbildung: <i>nein</i>	partiell nicht injektiv surjektiv
erste allgemeine Form	$f_{\text{allg}_1} : \mathcal{P}(U) \longrightarrow Z$	Aggregationen: <i>explizit</i> mehrdeutige Abbildung: <i>nein</i>	partiell (<i>partiell</i>) injektiv surjektiv
zweite allgemeine Form	$f_{\text{allg}_2} : \mathcal{P}(U) \longrightarrow \mathcal{P}(Z)$	Aggregationen: <i>explizit</i> mehrdeutige Abbildung: <i>ja</i>	partiell (<i>partiell</i>) injektiv (<i>nicht injektiv</i>) nicht surjektiv (<i>surjektiv</i>)

Tabelle 4.1: Die drei Formen der Generalisierungsfunktion im Vergleich: Definition, Vorteile, Nachteile und mathematische Charakteristiken. (Die *kursiv* und in Klammern hervorgehobenen Charakteristiken beziehen sich auf die strengeren Formen der Eigenschaften.)

Kapitel 5

Invariante Eigenschaften der Generalisierungsfunktion

In Kapitel 3 wurde ein für die Generalisierung geeignetes Datenmodell entwickelt, mit dem Geodaten formalisiert in eine rechnergestützte Umgebung abgebildet werden können. Insbesondere die Unterteilung der Geodaten in die Ebenen von Thematik, Topologie und Geometrie sei hier noch einmal erwähnt, weil diese es ermöglicht, bestimmte Eigenschaften von Geodaten getrennt zu betrachten. Die dort vorgenommene Unterteilung wird im Folgenden für die Analyse der Generalisierung weiter verwendet.

Das vorige Kapitel führte die auf mathematischen Grundlagen basierende Generalisierungsfunktion ein. Die Entwicklung von Verfahren zur Generalisierung kann dadurch auf die Suche nach einer geeigneten Funktion reduziert werden. Im vorletzten Abschnitt des Kapitels wurden unterschiedliche Generalisierungsfunktionen, die aus derselben Urmenge abbilden, untersucht und formalisiert. Auf eine Bewertung der unterschiedlichen Generalisierungsfunktionen wurde an dieser Stelle verzichtet.

Aufbauend auf den beiden vorigen Kapitel behandelt dieses Kapitel die *Beurteilung* der Ergebnisse von verschiedenen Generalisierungen ein und derselben Menge von Geodaten. Hierzu werden *Eigenschaften* (mathematisch *Strukturen*) der Mengen von Geodaten, der ungeneralisierten Urmenge und der generalisierten Zielmenge, herausgearbeitet. Die Strukturen basieren auf *Relationen* innerhalb der beiden Mengen. Wenn diese Relationen durch die Generalisierungsfunktion nicht – oder nur unwesentlich – verändert werden, so ist die Generalisierungsfunktion *invariant* bezüglich dieser Relationen beziehungsweise der darauf basierenden Strukturen. Die Überprüfung dieser Relationen vor und nach der Generalisierung führt zu einem *Qualitätsmaß*, das eine Beurteilung der Generalisierung ermöglicht. Anders ausgedrückt bilden die zu erhaltenden Strukturen *Bedingungen*, die an die Generalisierungsfunktion gestellt werden sollen.

Grundfragen, die mit Hilfe dieser Formalisierung beantwortet werden können, sind zum einen die Frage nach der *Korrektheit* einer Generalisierung. Zu jedem Zielmaßstab oder -strukturierungsgrad existieren Kriterien, die erfüllt sein müssen, damit eine Generalisierung als korrekt eingestuft werden kann. Diese Kriterien entstammen den Anforderungen,

die an die generalisierten Daten gestellt werden, wie sie zum Beispiel in den Objektartenkatalogen für Digitale Landschaftsmodelle formuliert sind oder sich aus den kartographischen Prinzipien der Lesbarkeit, geometrischen Richtigkeit und Vollständigkeit ergeben (Hake u. a. 2002). Erfüllt ein generalisierter Datenbestand diese Kriterien nicht oder nur ungenügend, so ist die Generalisierung *falsch* erfolgt. Diese Kriterien bilden *kategorische Bedingungen* an die Generalisierung und damit an die Generalisierungsfunktion.

Ausgehend von einem Ausgangsdatenbestand kann es zum anderen zu jedem Zielmaßstab oder -strukturierungsgrad nicht nur eine einzige korrekte Generalisierung geben. Im Gegenteil gibt es zu einem Ausgangsdatenbestand im Allgemeinen eine Vielzahl, wenn nicht sogar eine unendliche Anzahl von korrekten Generalisierungen.¹ Aus kartographischem Blickwinkel sind diese Generalisierungen im Allgemeinen nicht alle gleich gut – es gibt *bessere* und *schlechtere* Generalisierungen in der Menge der korrekten Generalisierungen. Es stellt sich damit die Frage nach der *Güte* einer Generalisierung. Die Kriterien zur Güte einer Generalisierung gründen oftmals im Erfahrungsschatz des Kartographen und sind deshalb schwierig zu formalisieren. Sie bilden keine absoluten Kriterien, die erfüllt sein *müssen*, sondern Kriterien, die zu einem möglichst hohen Grad erfüllt sein *sollten*. Damit bilden diese Kriterien *weiche Bedingungen*, die an die Generalisierung und damit an die Generalisierungsfunktion gestellt werden.

Ein Ziel dieses Kapitels ist die Formulierung von kategorischen und weichen Bedingungen für die Generalisierung. Die formulierten Bedingungen sind allgemein gehalten und damit nicht auf eine bestimmte Art der Generalisierung, wie zum Beispiel Modellgeneralisierung oder kartographische Generalisierung beschränkt. Sie sollen vielmehr auf so viele Generalisierungsarten wie möglich anwendbar sein. Kapitel 6 gibt Anwendungsbeispiele der hier entwickelten Bedingungen auf verschiedene Arten von Generalisierungen. Dennoch sind die in diesem Kapitel formulierten und formalisierten Bedingungen aus der Analyse von Generalisierungen im Bereich der Modellgeneralisierung entstanden und werden daher anhand von Beispielen aus der Modellgeneralisierung erläutert.

Wie in den Betrachtungen zur Methodik dieser Arbeit (Abschnitt 2.4) erwähnt wurde, geht es hier nicht darum, ein konkretes Verfahren zur Generalisierung zu entwickeln. Die hier angestellten Untersuchungen bieten jedoch mindestens zwei Möglichkeiten der Anwendung:

1. *Bestehende Generalisierungsverfahren* und ihre Ergebnisse können anhand der hier entwickelten Bedingungen auf ihre Korrektheit und Güte hin untersucht werden.
2. *Neu zu entwickelnde Verfahren* können sich die hier entwickelte Formalisierung zu Nutze machen und die Bedingungen in die Verfahren integrieren. Bei erfolgreicher Integration ist damit gewährleistet, dass das entwickelte Verfahren korrekte und gute Ergebnisse liefert.

Der nächste Abschnitt (Abschnitt 5.1) behandelt die mathematischen Grundlagen, die für die Formalisierung der Strukturen und Bedingungen notwendig sind. Dort wird auch der

¹Die unendliche Anzahl der Generalisierungsmöglichkeiten resultiert aus der Kontinuität des zu Grunde liegenden Raumes, hier der Ebene.

Begriff *Morphismus* eingeführt, der den Begriff der invarianten Strukturen mathematisch beschreibt. Die übrigen vier Abschnitte dieses Kapitels (Abschnitte 5.2 bis 5.5) arbeiten einzelne invariante Eigenschaften der Generalisierungsfunktion heraus und formulieren Bedingungen zu diesen Eigenschaften. Die Abschnitte sind unterteilt in Abschnitte über thematische, topologische, geometrische und kombinierte Eigenschaften. Im Folgenden wird in erster Linie auf die spezielle Generalisierungsfunktion Bezug genommen. Für die allgemeinen Formen der Generalisierungsfunktion sind die Formalisierungen entsprechend anzupassen.

5.1 Mathematische Grundlagen: Strukturen und Morphismen

In Kapitel 4, Abschnitt 4.2.1 wurde der Begriff der *Menge von Geodaten* eingeführt. Dabei wurde auch erwähnt, dass solche Mengen strukturiert sind. Die *Strukturen von Mengen* sind das zentrale Thema dieses und der folgenden Abschnitte. Mathematische *Grundstrukturen* sind algebraische Strukturen, Ordnungsstrukturen und topologische Strukturen. Eine *algebraische Struktur* ist zum Beispiel die Addition auf den reellen Zahlen. Ein Beispiel für eine *Ordnungsstruktur* ist die Größer-Relation ebenfalls auf den reellen Zahlen. Abgeschlossene Punktfolgen in der Ebene bilden zum Beispiel eine *topologische Struktur*. Allen Grundstrukturen ist gemeinsam, dass sie auf *Relationen* basieren oder sich auf *Relationen* zurückführen lassen (Reinhardt und Soeder 1974). Diese Relationen sind nicht, wie die in Kapitel 3 und 4 verwendeten Relationen, zwischen *unterschiedlichen* Mengen definiert, sondern sie bestehen *innerhalb* einer Menge.

Beispiele für Strukturen innerhalb einer Menge von Geodaten können auf den unterschiedlichen Ebenen des Datenmodells gefunden werden: Auf der thematischen Ebene gehört zum Beispiel ein Geodatum, das eine Straße in der realen Welt repräsentiert, zu einer Objektart in einem DLM, in ATKIS zum Beispiel zur Objektart 3101 **Straße**. Auf der topologischen Ebene besteht die Relation „*kreuzt*“ zwischen linienförmigen Objekten wie Straße und Eisenbahnstrecke. Auf der geometrischen Ebene existieren zum Beispiel Relationen, die die Entfernung zwischen Geodaten bezeichnen wie etwa die Relation, die alle Paare von Geodaten enthält, die eine Entfernung von weniger als 10 m zueinander besitzen. Natürlich gibt es auch Strukturen, die sich nicht eindeutig einer der drei Ebenen des Datenmodells zuordnen lassen. Diese werden im Folgenden *kombinierte Strukturen* genannt. Die hier genannten vier Arten von Strukturen werden in den Abschnitten 5.2 bis 5.5 behandelt.

Relationen beschreiben *Eigenschaften*, die der Menge von Geodaten inhärent sind. Mathematisch gesprochen werden die Relationen einer Menge von Geodaten *aufgeprägt* und geben der Menge damit eine *Struktur*. Eine *strukturierte Menge* ist also eine Menge M mit einer Familie $\{R_i\}$ von Relationen. Notiert wird $(M; R_1, \dots, R_n)$, wenn M mit den Relationen R_1, \dots, R_n versehen ist. Das geordnete Tupel $(M; R_1, \dots, R_n)$ wird *Gebilde* genannt (Reinhardt und Soeder 1974).

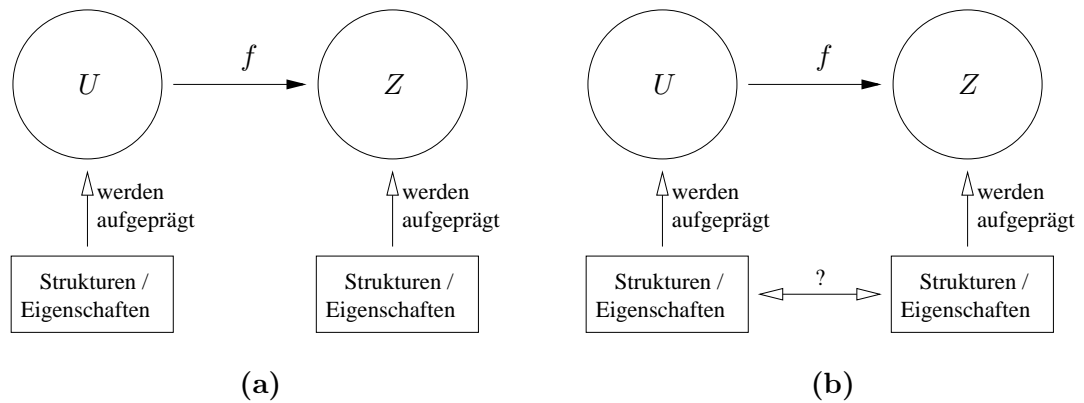


Abbildung 5.1: Generalisierungsfunktion f zwischen Ur- und Zielmenge, sowie Strukturen und Eigenschaften auf diesen Mengen: **(a)** unabhängige Strukturen, **(b)** Frage nach der Abhängigkeit dieser Strukturen.

Übertragen auf die Generalisierung existieren Strukturen, die auf die Mengen von Geodaten aufgeprägt sind. Abbildung 5.1 (a) zeigt die Situation der Urmenge und Zielmenge einer Generalisierung: die Generalisierungsfunktion f bildet ab aus der Urmenge U auf die Zielmenge Z . Sowohl auf U als auch auf Z sind Strukturen aufgeprägt; beide Mengen besitzen bestimmte Eigenschaften. Die Frage, die sich nun stellt, ist, in welcher Beziehung die Strukturen und Eigenschaften der Urmenge zu den Strukturen und Eigenschaften der Zielmenge stehen. Diese Frage ist in Abbildung 5.1 (b) angedeutet. Um diese Frage eingehend zu beantworten, muss man die Strukturen und Eigenschaften im Einzelnen betrachten, wie dies in den Abschnitten 5.2 bis 5.5 geschieht. Dass jedoch inhaltliche Beziehungen zwischen den Strukturen der ungeneralisierten und generalisierten Menge bestehen, ist offensichtlich, wie am oben genannten Beispiel deutlich wird: Wenn eine Straße und eine Eisenbahnstrecke sich im Ausgangsdatenbestand kreuzen, was einer topologischen Relation entspricht, und wenn beide Geodaten durch die Generalisierung in den generalisierten Datenbestand übernommen werden, so müssen sich die generalisierte Straße und die generalisierte Eisenbahnstrecke auch im Zieldatenbestand kreuzen.

5.1.1 Erweiterte Generalisierungsfunktion

Eine Möglichkeit, den Zusammenhang zwischen den Strukturen beider Mengen auszudrücken, besteht in der Einführung einer zusätzlichen Familie von Funktionen $\{g_i\}$, die die strukturprägenden Relationen von U auf die strukturprägenden Relationen von Z abbildet. Diese Möglichkeit ist in Abbildung 5.2 (a) skizziert. Zurückgehend auf obiges Beispiel wird damit die topologische Relation, die in der *Urmenge* das Kreuzen von linienförmigen Objekten beschreibt, auf die topologische Relation, die in der *Zielmenge* das Kreuzen von linienförmigen Objekten bezeichnet, abgebildet. Damit wird die Generalisierung, die bisher durch die Generalisierungsfunktion beschrieben wurde (siehe Kapitel 4),

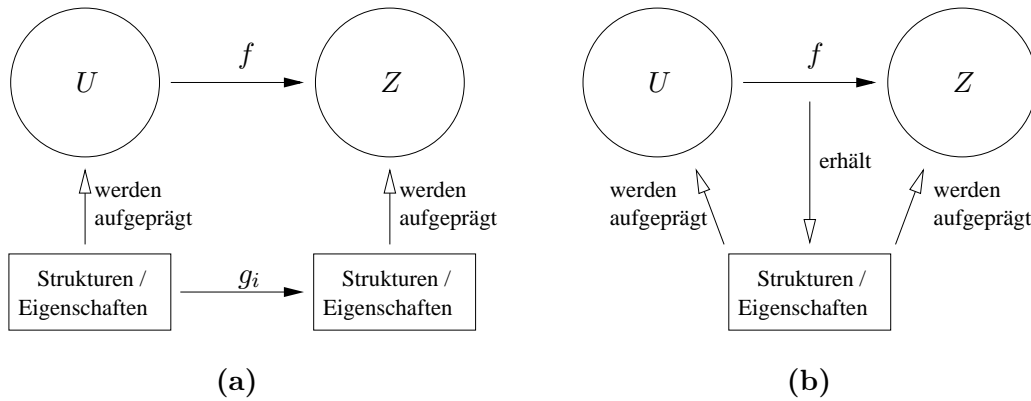


Abbildung 5.2: Generalisierungsfunktion f zwischen Ur- und Zielmenge, sowie Strukturen und Eigenschaften auf diesen Mengen: (a) strukturverträgliche Überföhrungsfunktionen g_i für die Strukturen und Eigenschaften, (b) Generalisierungsfunktion f als strukturhaltende Abbildung (Morphismus).

erweitert auf ein *Tupel von Funktionen* $(f; g_1, \dots, g_n)$. Im Folgenden wird diese Art der Erweiterung die *erweiterte Generalisierungsfunktion* genannt. Die erste Funktion bildet die Geodaten aufeinander ab und die weiteren Funktionen $\{g_i\}$ bilden die Relationen zwischen den Geodaten aufeinander ab.

Diese skizzierte Möglichkeit der Generalisierungsmodellierung durch mehrere Funktionen wurde verwendet von Ai und van Oosterom (2001). Allerdings benutzen Ai und van Oosterom die Unterscheidung der Funktionen, um diese hintereinander zu schalten und damit ein fehlerhaftes Generalisieren zu korrigieren. Konkret werden dort zunächst die Geodaten durch die Funktion f generalisiert, insbesondere geometrisch vereinfacht. In einem zweiten Schritt, mit Hilfe einer nachgeschalteten Funktion g_i werden die Geodaten geometrisch verschoben oder verdrängt, um topologische Fehler, die durch die geometrische Vereinfachung entstanden sind, zu korrigieren.

Da im Kontext dieser Arbeit die Geometrie *und* die Topologie als Teil eines Geodatum verstanden wird (siehe Abschnitt 4.2.1), lässt sich der Ansatz von Ai und van Oosterom (2001) nicht übertragen. Die oben skizzierte Möglichkeit einer Abbildung der Relationen bedeutet eine gleichzeitige Generalisierung sowohl der Geodaten als auch der Relationen zwischen ihnen, und nicht eine nachgeschaltete Generalisierung, wie Ai und van Oosterom sie propagieren.

Die Generalisierungsfunktion f ist in Zusammenhang mit den Relationenabbildungen *strukturverträglich*, da sie die Strukturen (oder die zu Grunde liegenden Relationen) im Rahmen der Generalisierung verträglich aufeinander abbildet. In einzelnen Fällen wird in dieser Arbeit auf die erweiterte Generalisierungsfunktion zurückgegriffen. Wenn es jedoch möglich ist, wird die im nächsten Unterabschnitt vorgestellte Alternative verwendet.

5.1.2 Alternative zur erweiterten Generalisierungsfunktion

In diesem Abschnitt wird eine andere Herangehensweise an die oben aufgeworfene Frage nach der Beziehung zwischen den Strukturen der Ur- und Zielmenge vorgeschlagen. Als Motivation dient wieder obiges Beispiel der sich mit einer Eisenbahnstrecke kreuzenden Straße: Wie schon erwähnt, muss sich die Relation „kreuzen“ zwischen einer Straße und einer Eisenbahnstrecke auch nach der Generalisierung im Zieldatenbestand wiederfinden. Man kann auch sagen, die Relation „kreuzt“ ist der Urmenge ebenso aufgeprägt wie der Zielmenge. In anderen Worten ausgedrückt, ist die Relation *dieselbe Relation* auf der Ur- wie auf der Zielmenge. Die Identität der Relation betrifft die *Bedeutung* oder die *Interpretation* der Relation. Mathematisch ist die Identität schwieriger auszudrücken, dies geschieht weiter unten in Abschnitt 5.1.4.

Die Behauptung, die hier aufgestellt wird, ist, dass die Relationen auf der Urmenge durch die Generalisierungsfunktion keine wesentlichen Änderungen in ihrer Bedeutung erfahren. Die Bedeutung bleibt gleich, oder sie wird nur innerhalb bestimmter Grenzen verändert.

Wenn die Relationen aber dem Wesen oder dem Sinn nach gleich bleiben, dann ist die Einführung einer zusätzlichen Funktionen $\{g_i\}$ zur Abbildung unnötig kompliziert. Denn die $\{g_i\}$ sind dann nichts anderes als Identitätsabbildungen zwischen Relationen der Urmenge und den entsprechenden Relationen in der Zielmenge.

Vorgeschlagen wird nun eine Modellierung der Struktur- und Eigenschaftsbeziehungen, wie es in Abbildung 5.2 (b) gezeigt ist. Die Familie von Relationen wird *beiden* Mengen von Geodaten, der Urmenge *und* der Zielmenge aufgeprägt. Das bedeutet für die Generalisierungsfunktion, dass sie die Relationen, beziehungsweise die durch sie ausgedrückten Strukturen *erhält*; sie ist *strukturertretend*. In anderen Worten ist die Generalisierungsfunktion *invariant* bezüglich der von ihr erhaltenen Strukturen und Relationen.

5.1.3 Morphismen

Wenn eine Funktion die Strukturen der Urmenge auf die Zielmenge überträgt, so nennt man diese Funktion eine *strukturverträgliche* (siehe Unterabschnitt 5.1.1) oder *strukturertretende Funktion* (siehe Unterabschnitt 5.1.2) *bezüglich der Strukturen*, die auf beiden Mengen gelten. Strukturverträgliche und strukturertretende Abbildungen werden auch mit *Morphismus* bezeichnet (Reinhardt und Soeder 1974). Morphismen sind Thema eines eigenen Teilgebietes der Mathematik, nämlich der *Kategorientheorie*. Mengen, ihre Strukturen und Abbildungen zwischen den Mengen sind die zentralen Themen der Kategorientheorie. Die Beschreibung von Überführungen einer mathematischen Struktur in eine andere mathematische Struktur fällt somit auch darunter. Ausführliche Auseinandersetzungen mit der Kategorientheorie finden sich bei Barr und Wells (1990), Mac Lane (1998), Arbib und Manes (1975).

Herring (1991) überträgt die Methoden der Kategorientheorie auf die Modellierung von räumlichen und nicht-räumlichen Daten in GIS. Er untersucht strukturverträgliche Abbildungen (Morphismen) unter anderem zwischen räumlichen Datenmodellen, um Rückschlüsse über die Eigenschaften eines Modelles aus den Eigenschaften eines anderen Modelles zu ziehen. Seine Motivation ist die Erkenntnis, dass dieselben Realweltdaten in der Praxis auf verschiedene Modelle abgebildet werden, und der Bedarf besteht, diese Modelle miteinander zu verknüpfen.

Die Generalisierung ist diesem Ansatz ähnlich, da auch hier zwei Modelle derselben Realweltdaten beziehungsweise entstehen und miteinander verknüpft werden. Im Unterschied zu den Ausführungen von Herring wird jedoch bei der Generalisierung das eine Modell aus dem anderen Modell abgeleitet, also erst erzeugt. Dennoch bildet die Generalisierungsfunktion einen *Morphismus* bezüglich der auf Ur- und Zielmenge aufgeprägten Strukturen. Ziel ist es nun, diese Strukturen zu identifizieren, die für eine im kartographischen Sinne korrekte und gute Generalisierung erhalten werden müssen.

Die Identifikation solcher Strukturen selber gründet auf der Analyse von Vorgängen bei der Generalisierung, mithin auf der kartographischen Erfahrung und dem kartographischen Wissen. Viele Regeln, die von Kartographen bei der Generalisierung intuitiv angewendet werden, entspringen dem gesunden Menschenverstand und sind deshalb nicht schriftlich fixiert. Zum Beispiel darf eine Straße in einer Folgekarte nicht in eine Eisenbahnstrecke umgewandelt werden. Für die automatisierte Generalisierung müssen solche Selbstverständlichkeiten erst formuliert werden. Andere Regeln hingegen sind sehr subtil, wenn man zum Beispiel an die Vorgänge bei der geometrischen Verdrängung von Objekten denkt. Diese und andere Aspekte müssen alle bei der Identifikation von invarianten Strukturen berücksichtigt werden.

5.1.4 Mathematische Formulierung der Strukturhaltung

In diesem Abschnitt werden die Strukturverträglichkeit und die Strukturhaltung, die in Abbildung 5.2 visualisiert sind, mathematisch formuliert. Dabei besteht das Problem, dass man von den *gleichen* Relationen auf der Ur- und der Zielmenge sprechen möchte, dies sich mathematisch jedoch nicht direkt ausdrücken lässt. Am Beispiel einer zweistelligen Relation soll dies verdeutlicht werden:

Eine Relation ist definiert als kartesisches Produkt mehrerer Mengen. Für eine zweistellige Relation R auf der Urmenge U gilt $R \subseteq U \times U$. Das Erhalten einer Relation durch die Generalisierung bedeutet, dass für alle Elemente $a, b \in U$ gelten muss, dass, wenn $R(a, b)$ gilt, und $f(a) = a'$ und $f(b) = b'$, auch $R(a', b')$ gelten muss. Im oben genannten Beispiel entsprechen a und b der Straße und der Eisenbahnstrecke. Die Relation R ist in diesem Fall die Beziehung „*kreuzt*“. $R(a, b)$ bedeutet, dass die Straße und die Eisenbahnstrecke sich kreuzen. Wenn nun beide Geodaten, Straße und Eisenbahnstrecke, in die Zielmenge übernommen werden, so muss auch die generalisierte Straße a' die generalisierte Eisenbahnstrecke b' kreuzen, mithin muss $R(a', b')$ gelten.

Hier besteht jedoch mathematisch ein Problem, denn a' und b' sind Elemente der Zielmenge Z , während R Teilmenge von $U \times U$ ist. Also kann (a', b') nicht Element der Relation R sein.

Zur Lösung dieses Dilemmas führen wir eine *Metamenge* \mathcal{E} als Menge *aller* Geodaten ein, so dass Urmenge und Zielmenge Teilmengen der Metamenge sind, also $U, Z \subset \mathcal{E}$. Auf dieser Metamenge definiert man nun die *Metarelation* „kreuzen“ als die Menge aller Paare von Geodaten, die sich kreuzen. Dann ist $R \subseteq \mathcal{E} \times \mathcal{E}$. Des Weiteren führen wir zwei Relationen R_U und R_Z ein als die Menge der Paare von Elementen aus U beziehungsweise Z , die sich kreuzen. Damit ist $R_U \subseteq U \times U$ und $R_Z \subseteq Z \times Z$. Damit sind diese beiden Relationen jeweils Teilmenge der Metarelation R , also $R_U, R_Z \subset R$. Insgesamt ergibt sich daraus folgendes Diagramm:

$$\begin{array}{rcccl} R & \subseteq & \mathcal{E} & \times & \mathcal{E} \\ \cup & & \cup & & \cup \\ R_U & \subseteq & U & \times & U \\ R_Z & \subseteq & Z & \times & Z \end{array}$$

Mit Hilfe dieser Konstruktion lassen sich Relationen, die gleich interpretiert werden (zum Beispiel „kreuzen“) auf verschiedene Mengen von Geodaten anwenden. Im Folgenden werden wir auf die Indizes U beziehungsweise Z an den Relationen verzichten, wenn Verwechslungen ausgeschlossen sind. Damit entspricht unser mathematisches Modell dem Morphismus, wie er in Abbildung 5.2 (b) abgebildet ist. Dieselben Strukturen, auf denselben Relationen beruhend, sind den unterschiedlichen Mengen U und Z aufgeprägt.

Seien nun die invarianten Relationen der Generalisierungsfunktion R_1, \dots, R_n , dann ist die Generalisierungsfunktion ein Morphismus, der von dem Gebilde der Urmenge U und der aufgeprägten Relationen R_1, \dots, R_n auf das Gebilde der Zielmenge Z mit denselben aufgeprägten Relationen R_1, \dots, R_n abbildet. Formell lautet die Generalisierungsfunktion dann

$$f : (U; R_1, \dots, R_n) \longrightarrow (Z; R_1, \dots, R_n).$$

In manchen Fällen (siehe Abschnitt 5.2) können die Relationen allerdings nicht in dieser Form aufeinander abgebildet werden. Dann benötigt man eine Unterscheidung der Relationen auf Ur- und Zielmenge, die in der folgenden Formulierung der Generalisierungsfunktion durch Hochindizes gekennzeichnet sind:

$$f : (U; R_1^U, \dots, R_n^U) \longrightarrow (Z; R_1^Z, \dots, R_n^Z).$$

Die folgenden Abschnitte befassen sich mit der Identifikation und Formalisierung der Eigenschaften, Strukturen und Relationen auf der Ur- und Zielmenge, die durch die Generalisierungsfunktion invariant gehalten werden. Außerdem erfolgt die Formulierung von Bedingungen, die diese Invarianz sicher stellen.

5.2 Invariante thematische Eigenschaften

Im Zusammenhang mit dem Datenmodell wurde in Abschnitt 3.1.3 der Begriff der Thematik eingeführt. Die Thematik eines Geodatums umfasst die nicht-räumlichen Informationen, die diesem Datum zugeordnet sind. Im Datenmodell ist für die Thematik eine eigene Ebene vorgesehen, die die nicht-räumlichen Informationen von der Topologie und der Geometrie des Geodatums trennt.

In Digitalen Landschaftsmodellen werden die thematischen Informationen eines Geodatums zum einen gebildet durch die Zugehörigkeit zu einer bestimmten *Objektart* und zum anderen durch eine detailliertere Beschreibung über *Attribute*. Diese Einteilung der Thematik von Geodaten wird im Folgenden beibehalten. Die Begriffe Objektart und Attribut sowie die speziellen Auswirkungen von Aggregationen hierauf werden zunächst hier eingeführt und dann in den folgenden Unterabschnitten auf ihre invarianten Eigenschaften hin untersucht. Dies führt zur Formulierung von *kategorischen* Bedingungen an eine Generalisierung.

Objektarten

Objektarten dienen im Kontext von Digitalen Landschaftsmodellen dazu, die Landschaft grob zu gliedern (AdV 1989). Die Zuordnung von Landschaftsobjekten oder Realweltobjekten zu einer bestimmten Objektart bedeutet mathematisch gesehen eine *Klassifizierung* der Landschaft. Jedes Landschaftsobjekt wird eindeutig einer bestimmten Objektart und damit einer bestimmten *Klasse* zugeordnet.² Zum Beispiel gehört ein Fußweg in ATKIS zu der Klasse 3102 *Weg*. Festgelegt ist diese Zuordnung durch die *Definition* der Objektarten in den Objektartenkatalogen. In *Digitalen Kartographischen Modellen* (DKM) und *Digitalen Topographischen Karten* (DTK) wird die Klassifizierung mit Hilfe von Signaturenkatalogen durchgeführt.

Für die *topographischen Kartenwerke* von Deutschland gibt es den Begriff Objektart nicht. Die Einteilung der Landschaft wird festgelegt durch die *Musterblätter*, die für jeden verwendeten Maßstab erhältlich sind. Anhand von Zeichenmustern und Kartenbeispielen wird eine Einteilung der Landschaft vorgeschrieben. Die Klassifizierung kann auch ersehen werden in der *Zeichenerklärung* (*Legende*), die den Karten beigelegt ist und eine Auswahl aus dem Musterblatt darstellt.

In *thematischen Karten* werden die Landschaftsobjekte über die Legende oder Zeichenschlüssel klassifiziert. In geologischen Karten zum Beispiel wird die Klassifizierung von Flächen über die chronostratigraphischen Kategorien (Rangordnung der Zeiteinheiten)

²Eine Klassenbildung ist mathematisch gleich bedeutend zu einer Partition (siehe Abschnitt 4.3.3). Hier werden jedoch die Begriffe Klassenbildung oder Klassifizierung benutzt, da sie den kartographischen Begriffen nahe kommen. Im Gegensatz hierzu ist die Partitionierung aus Abschnitt 4.3.3 nicht mit der kartographischen Klassifizierung verbunden.

durchgeführt. Die Erdoberfläche wird grob klassifiziert durch die Erdzeitalter. Diese Klassifizierung kann, in dieser Reihenfolge, verfeinert werden durch die Formation, die Abteilung, die Stufe und die Zone (Vossmerbäumer 1976). Ein Ausschnitt aus einer geologischen Zeittafel zeigt Tabelle 6.2 auf Seite 161. In Abschnitt 6.2 werden die in diesem Kapitel entwickelten Bedingungen zu invarianten thematischen Eigenschaften auf geologische Karten übertragen.

Die Klassenbildung wird in Abschnitt 5.2.1 formalisiert. Dabei wird Bezug genommen auf die Digitalen Landschaftsmodelle ATKIS (AdV 2003) und TOPIS (MilGeo-OK 1993). Davon ausgehend werden die invariant gehaltenen Eigenschaften der Objektarten herausgearbeitet.

Attribute

Im Gegensatz zur groben Gliederung durch Objektarten wird die Landschaft durch Attribute feiner gegliedert (AdV 1989). Damit wird die Klassifizierung durch Objektarten erweitert, indem durch Zuordnung von unterschiedlichen Attributwerten an Landschaftsobjekte Unterklassen gebildet werden. Im oben verwendeten Beispiel wurde der Fußweg der Objektart (Klasse) 3102 Weg zugeordnet. Zu dieser Objektart gehört im Objektartenkatalog unter anderem das Attribut FKT Funktion. Durch Zuweisung des Attributwertes 1703 Fußweg, Pfad wird der Fußweg einer Unterklasse der Klasse 3102 Weg zugeordnet.

Attribute können nach unterschiedlichen Kriterien eingeteilt werden. Schürer (2002) unterteilt Attribute unter anderem in *objektbildende* und *nicht-objektbildende Attribute*. Zum einen versteht Schürer (2002, S. 50) unter objektbildenden Attributen diejenigen Attribute, die bei unterschiedlichen Werten benachbarter Realwertobjekte die Bildung eines neuen Objektes bewirken. Damit entspricht diese Definition derjenigen von *herausgehobenen* Attributen aus ATKIS (AdV 1989). In ATKIS gibt es jedoch die Trennung zwischen Objekten und Objektteilen. Wertwechsel bei herausgehobenen Attribute zeigen an, dass ein neues *Objekt* gebildet werden muss; Wertwechsel bei sonstigen Attributen zeigen an, dass ein *Objektteil* gebildet werden muss. Diese Auffassung von objektbildenden Attributen bezieht sich damit auf die *Objektbildung innerhalb eines DLM*. Da in dieser Arbeit von einem flachen Modell ausgegangen wird wie bei TOPIS (siehe Abschnitt 3.1.6), wird diese Unterscheidung hier nicht benötigt.

Zum anderen versteht Schürer (2002, S. 100) unter objektbildenden Attributen aber auch diejenigen Attribute, „die verschiedene Realweltobjekte in einer Objektart über die Variation des Attributwertes modellieren“. Wenn solch ein definierender Attributwert im DLM niedrigerer Auflösung nicht vorgesehen ist, so kann ein Geo-Objekt mit diesem Attributwert nicht in den Zieldatenbestand übernommen werden. Damit ist ein objektbildendes Attribut auch ein Kriterium für das Fortlassen eines Geo-Objektes bei der Generalisierung. Diese Auffassung von objektbildenden Attributen bezieht sich also auf die *Objektbildung zwischen zwei DLM* unterschiedlicher Strukturierungsgrade. Das oben



Abbildung 5.3: *Unterklassenbildung in topographischen Karten an Hand des quantitativen Attributs Breite für die Klasse Kanal (nach dem Musterblatt TK 25, 1981).*

angeführte Beispiel des Funktionsattributes bei Wegen ist in diesem Sinne objektbildend, da zum Beispiel der Attributwert von 1703 Fußweg, Pfad im OK 25 und OK 250, jedoch nicht im OK 1000 vorgesehen ist.

Abschnitt 5.2.2 befasst sich mit der Formalisierung der Attributüberführung bei der Generalisierung. Die Gemeinsamkeiten mit der Strukturhaltung auf Ebene der Objektarten werden dabei herausgehoben.

Eine andere Unterteilung der Attribute ist in der Dokumentation zu ATKIS (AdV 1989) festgelegt: Dort werden *quantitative* und *qualitative* Attribute unterschieden. Quantitative Attribute enthalten als Wert eine Größe. Dies muss nicht unbedingt nur eine geometrische Größe (wie Höhe, Länge oder Fläche) sein, wie es Schürer (2002, S. 49+51) schreibt, sondern kann auch andere Größen wie zum Beispiel zulässiges Gesamtgewicht oder Einwohnerzahl umfassen. Qualitative Attribute enthalten als Wert einen Code, der einen bestimmten Sachverhalt bei dem Geo-Objekt ausdrückt (siehe das Beispiel vom Fußweg oben). Des Weiteren sind in ATKIS verschiedentlich *Namen* für die Geo-Objekte vorgesehen. In ATKIS sowie in TOPIS sind diese ähnlich wie Attribute kodiert. Sie werden im Folgenden *bezeichnende Attribute* genannt.³

Attribute finden nicht nur in DLM Verwendung. Auch in anderen Kartenarten gibt es Unterscheidungen durch Attribute. Abbildung 5.3 zeigt ein Beispiel für eine Unterklassenbildung in topographischen Karten des Maßstabes 1:25 000 durch unterschiedliche Attributwerte: An Hand des Wertes vom Attribut „Breite“ wird entschieden, welcher Zeichenschlüssel auf das Landschaftsobjekt „Kanal“ angewendet wird. Auch bei der Signaturierung von zum Beispiel Straßen findet sich die Unterteilung von Klassen durch Attribute, wie die Unterklassenbildung von Straßen in Autobahnen, Bundes-, Haupt- und Nebenstraßen etc. (siehe Tabelle 5.4 auf Seite 116).

Auswirkung von Aggregationen

Wenn Geodaten bei der Generalisierung aggregiert werden – ein besonders bei DLM häufig vorkommendes Phänomen, weil dort ein Flächenschluss vorgegeben ist –, dann

³Worboys (2004) unterscheidet zwischen nominalen, ordinalen, intervallartigen und verhältnisartigen Attributen. Die nominalen Attribute fassen die hier genannten qualitativen und bezeichnenden Attribute zusammen, während die übrigen drei Attributarten Unterteilungen der quantitativen Attribute bilden.

muss dem aggregierten Geodatum genau *eine* Objektart und für jedes Attribut genau *ein* Attributwert zugeordnet werden, obwohl die Ausgangsdaten, die zu diesem Geodatum aggregiert wurden, durchaus verschiedenen Objektarten angehören und unterschiedliche Attributwerte besitzen können. Abschnitt 5.2.3 befasst sich mit den Phänomenen, die bei der Aggregation auf der thematischen Ebene auftreten.

Die im letzten Abschnitt getroffene Unterscheidung in quantitative, qualitative und bezeichnende Attribute ist von Bedeutung auch bei der Aggregation von Objekten. Abhängig von der Art des Attributes müssen verschiedene Verfahren bei der Generalisierung zur Anwendung kommen. In Abschnitt 5.2.3 kommt diese Unterscheidung in Bezug auf Attribute ebenfalls zur Sprache.

5.2.1 Objektarten

In diesem Unterabschnitt wird die Klassenbildung durch Objektarten formalisiert und daraus eine kategorische Bedingung an die Generalisierungsfunktion abgeleitet. In Digitalen Landschaftsmodellen werden die Objektarten durch Objektartenkataloge definiert. Pro Strukturierungsgrad gibt es genau einen Objektartenkatalog (abgekürzt OK). Für ATKIS sind dies zum Beispiel die der Basis-OK (OK 25) und die OK 50, OK 250 und OK 1000 (AdV 2003). Die Bezeichnung der Objektarten in den unterschiedlichen Katalogen kann gleich sein; zum Beispiel finden sich die Objektarten mit Code und Namen 2101 *Ortslage* in allen vier genannten Katalogen. Diese Objektarten klassifizieren dieselben Realweltobjekte. Dennoch müssen die Objektarten unterschiedlicher Strukturierungsgrade unterschieden werden, da die Erfassungskriterien und die Bestimmung des Geometrietyps (Objekttyps) in jedem OK unterschiedlich definiert sein können (und es im Allgemeinen auch sind).

Jeder Objektartenkatalog enthält eine endliche Anzahl von Objektarten. Dies seien die Objektarten O_1, \dots, O_n , wobei n die Anzahl der Objektarten für einen bestimmten Strukturierungsgrad ist. Die Menge aller Objektarten eines Strukturierungsgrades sei $\mathcal{O} = \{O_1, \dots, O_n\}$. Um Objektartenkataloge verschiedener Strukturierungsgrade unterscheiden zu können, erhalten die Objektarten und die Mengen von Objektarten einen Hochindex, der den Strukturierungsgrad ausdrückt. Für den OK 25 aus ATKIS wäre die Menge aller Objektarten also $\mathcal{O}^{25} = \{O_1^{25}, \dots, O_n^{25}\}$.

Auf die Ur- und Zielmenge bezogen ist allgemeiner die Menge der Objektarten $\mathcal{O}^U = \{O_1^U, \dots, O_n^U\}$ für den Strukturierungsgrad der Urmenge U , und die Menge der Objektarten $\mathcal{O}^Z = \{O_1^Z, \dots, O_m^Z\}$ für den Strukturierungsgrad der Zielmenge Z . Im Allgemeinen ist die Anzahl der Objektarten von \mathcal{O}^U und \mathcal{O}^Z nicht gleich, weshalb die Kardinalitäten der Mengen auf n beziehungsweise m gesetzt worden sind.

Jedes Geodatum wird genau einer Objektart des entsprechenden Strukturierungsgrades zugeordnet. Formell lässt sich das durch eine Funktion o ausdrücken, die von der Urbeziehungsweise Zielmenge auf die Menge der zugehörigen Objektarten abbildet. Die Funktionen lauten also $o_U : U \rightarrow \mathcal{O}^U$ und $o_Z : Z \rightarrow \mathcal{O}^Z$.

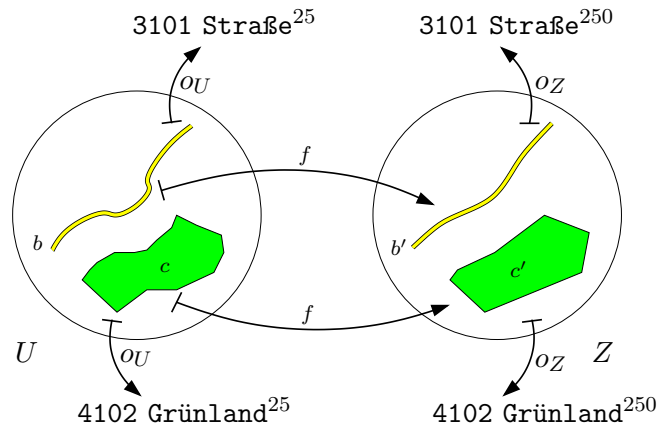


Abbildung 5.4: Zuordnung von ungeneralisierten und generalisierten Daten in die Menge der Objektarten durch die Funktionen o_U und o_Z .

Ein Beispiel soll dies erläutern. Abbildung 5.4 zeigt die Generalisierung eines Ausschnittes aus der Abbildung 4.8 aus Kapitel 4, reduziert auf die Straße b und die Grünfläche c . Im Kontext von ATKIS sei die Urmenge aus dem Basis-DLM (beschrieben durch den OK 25) und die Zielmenge aus dem DLM 250 (beschrieben durch den OK 250). Die Zuordnung der Geodaten auf die Objektarten geschieht also auf der Urmenge durch die Funktion o_{25} , die die Straße b auf die Objektart 3101 Straße^{25} und das Grünland c auf die Objektart $4102 \text{ Grünland}^{25}$ abbildet, formell: $o_{25}(b) = 3101 \text{ Straße}^{25}$ und $o_{25}(c) = 4102 \text{ Grünland}^{25}$. Die Generalisierungsfunktion sei die Abbildung von Basis-DLM zum DLM 250. Dann wird die generalisierte Straße $f(b) = b'$ abgebildet auf die Objektart 3101 Straße im DLM 250 und die generalisierte Grünfläche $f(c) = c'$ auf die Objektart 4102 Grünland im DLM 250. Formell lautet dies also $o_{250}(b') = 3101 \text{ Straße}^{250}$ und $o_{250}(c') = 4102 \text{ Grünland}^{250}$.

Die Funktionen o_{25} und o_{250} sind die thematischen Strukturen betreffend der Objektarten der Ur- und Zielmenge in diesem Beispiel. Allgemeiner werden die thematischen Strukturen für die Objektarten durch die Zuordnungen durch die Funktionen o_U und o_Z ausgedrückt.⁴ Mit dieser Ausdrucksweise haben wir Strukturen eingeführt, die die Zugehörigkeit von Geodaten zu Objektarten beschreiben. Die Frage, die in Abschnitt 5.1 gestellt wurde und die die zentrale Motivation für dieses Kapitel bildet, ist die Frage nach der Beziehung der Strukturen zwischen den verschiedenen Strukturierungsgraden.

Die Relevanz dieser Frage im Zusammenhang mit den Objektarten wird an Hand eines

⁴In Abschnitt 5.1 wurde erläutert, dass alle Strukturen sich auf *Relationen* zurückführen lassen. So auch in diesem Fall: Jede Objektart aus \mathcal{O}^U beziehungsweise \mathcal{O}^Z kann als einstellige Relation auf U beziehungsweise Z verstanden werden. Die O_i^U beziehungsweise O_i^Z bezeichnen dann disjunkte Teilmengen von U beziehungsweise Z . Jedes Geodatum ist Element von genau einem O_i^U beziehungsweise O_i^Z . Als einstellige Relation ausgedrückt wäre dies zum Beispiel $O_i^U(x)$ mit der Bedeutung, dass das Geodatum x aus U der Objektart O_i angehört. Für das Beispiel der Straße b würde dann gelten: $3101 \text{ Straße}^{25}(b)$. Der Anschaulichkeit halber wurde jedoch auf diesen Formalismus verzichtet.

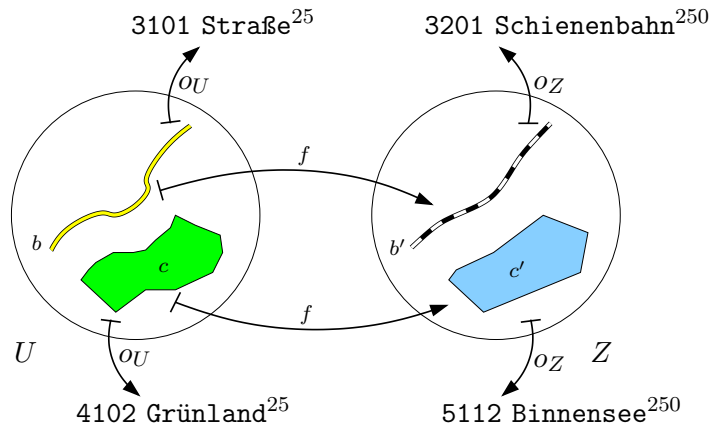


Abbildung 5.5: Beispiel für eine inkorrekte Generalisierung, erkennbar an der Zuordnung von ungeneralisierten und generalisierten Daten in die Menge der Objektarten.

Negativbeispiele in Abbildung 5.5 deutlich. Die Urmenge U entspricht der Urmenge aus Abbildung 5.4. Jedoch bildet in diesem Beispiel die Generalisierungsfunktion die Straße b auf eine Eisenbahnstrecke b' und die Grünfläche c auf einen See c' ab. Ein Mensch würde die Unsinnigkeit dieser Generalisierung sofort erkennen. Dieses Beispiel verdeutlicht, dass die Zugehörigkeit zu einer Objektart bei der Generalisierung bestimmten Regeln folgen muss. Ein Mensch „weiß“, dass eine Straße nicht in eine Eisenbahnstrecke umgewandelt werden darf, sondern als Straße weiter bestehen muss. Für einen Rechner muss diese Selbstverständlichkeit formalisiert werden. Benötigt wird also eine Formalisierung der Beziehungen zwischen den Objektarten.

Idealerweise lässt sich auch diese Beziehung zwischen den Objektarten der beiden Strukturierungsgrade durch eine Funktion ausdrücken, die hier g_O genannt wird. In Abschnitt 5.1.1 wurde der Buchstabe g für die Funktion zwischen den Strukturen (Relationen) der Mengen von Geodaten gewählt; der Index O zeigt, dass diese Funktion zwischen den Objektarten abbildet. Das folgende Diagramm zeigt nun die Beziehungen für das Beispiel der Straße zwischen den Geodaten und den Objektarten:

$$\begin{array}{ccc}
 b & \xrightarrow{o_{25}} & 3101 \text{ Straße}^{25} \\
 \downarrow f & & \downarrow g_O \\
 b' & \xrightarrow{o_{250}} & 3101 \text{ Straße}^{250}
 \end{array}$$

Die Straße b wird durch die Generalisierungsfunktion f auf die Straße b' abgebildet. Beide Geodaten werden durch die Funktionen o_{25} und o_{250} auf die zugehörigen Objektarten 3101 Straße^{25} und $3101 \text{ Straße}^{250}$ abgebildet. Die neu eingeführte Funktion g_O bildet diese beiden Objektarten aufeinander ab. In diesem Diagramm gibt es damit zwei

„Pfade“ von der Straße b zur Objektart **3101 Straße**²⁵⁰, den einen über die Objektart zur Straße b , den anderen über die generalisierte Straße b' . Die Strukturverträglichkeit der Funktion f fordert nun, dass beide „Pfade“ zum gleichen „Ziel“ führen. Man sagt, das Diagramm ist *kommutativ*. Formell ergibt sich daraus die Bedingung, dass folgende Gleichung erfüllt sein muss:

$$g_O(o_{25}(b)) = o_{250}(f(b))$$

Dies ist die Bedingung für ein konkretes Geodatum, hier der Straße b . Auf der Ebene der Mengen ergibt sich allgemeiner folgendes Diagramm:

$$\begin{array}{ccc} U & \xrightarrow{o_U} & \mathcal{O}^U \\ \downarrow f & & \downarrow g_O \\ Z & \xrightarrow{o_Z} & \mathcal{O}^Z \end{array}$$

Links ist die Generalisierungsfunktion f notiert, die von der Urmenge U auf die Zielmenge Z abbildet. Von links nach rechts bilden die Funktionen o_U und o_Z die Elemente der Ur- und der Zielmenge ab auf die Mengen der Objektarten \mathcal{O}^U und \mathcal{O}^Z des jeweiligen Strukturierungsgrades. Die Funktion g_O bildet die Mengen der Objektarten aufeinander ab. Die Strukturverträglichkeit der Funktion f lässt sich nun wieder durch die Kommutativität des Diagramms zu einer Bedingung formulieren:

Für alle $x \in U$ gilt: $g_O(o_U(x)) = o_Z(f(x))$	(5.1)
--	-------

Umgangssprachlich bedeutet dies, dass die Objektart eines generalisierten Geodatums gleich der Abbildung der Objektart des ungeneralisierten Geodatums sein muss. Diese Bedingung ist die Formulierung der invarianten Eigenschaft der Objektarten. Sie kann automatisch auf Erfüllung überprüft werden und ist damit eine *kategorische Bedingung*.

Beschreibung der Objektartenzuordnung

Die Funktion g_O der Objektartenzuordnung lässt sich durch eine Tabelle beschreiben. Dabei ist zu berücksichtigen, dass nicht alle Objektarten des Strukturierungsgrades der Urmenge eine Entsprechung im Strukturierungsgrad der Zielmenge haben müssen. Die Funktion g_O ist somit *partiell*. Außerdem können mehrere Objektarten aus \mathcal{O}^U auf dieselbe Objektart aus \mathcal{O}^Z abgebildet werden. g_O muss also nicht zwingend injektiv sein. Die folgende Tabelle gibt ein (fiktives) Beispiel für eine Funktion zur Objektartenzuordnung, in der die vorgenannten Fälle als Beispiele auftreten:

Objektart Urmenge	Objektart Zielmenge
Wald	Wald
Einzelbaum	—
Bundesstraße	Straße
Landesstraße	Straße

Schürer (2002, S. 89 ff.) weist darauf hin, dass in der Praxis auch uneindeutige Zuordnungen auftreten können, in denen eine Objektart des Ausgangsstrukturierungsgrades nicht eindeutig einer Objektart des Zielstrukturierungsgrades zugeordnet werden kann. Als Beispiel nennt er die Objektart AK120 **Park** aus dem TOPIS OK hoher Auflösung. Ein Geodatum dieser Objektart kann – je nach Belegung des Attributes US1 **Nutzung1** – entweder zu einem Objekt der Objektart AK120 **Park** oder der Objektart EC015 **Wald/Forst** des TOPIS OK mittlerer Auflösung generalisiert werden.

Zwei weitere Beispiele seien hier genannt, bei denen die Objektartenzuordnung nicht eindeutig ist:

1. Bei der Generalisierung und gleichzeitigen Umwandlung von TOPIS-Daten in ATKIS-Daten werden Geodaten der Objektart AL020 **Bebautes Gebiet** (TOPIS) abhängig vom Wert des Attributes US2 **Nutzung Funktion** zu Geodaten der Objektarten 2101 **Ortslage**, 2111 **Wohnbaufläche**, 2112 **Industrie- und Gewerbefläche**, 2113 **Fläche gemischter Nutzung** oder 2114 **Fläche besonderer funktionaler Prägung** generalisiert.

2. Die kartographische Darstellung aus einem DLM heraus erfordert eine Zuordnung von Signaturen zu den DLM-Objekten. Dieser Vorgang ist Teil der kartographischen Generalisierung im Umfeld der digitalen Modelle (Hake u. a. 2002). Im Zusammenhang mit dem zur Zuordnung verwendeten Signaturenkatalog (SK) wird nicht mehr der Begriff Objektart gebraucht. Die Einteilung bildet dennoch eine Klassifizierung der erfassten Objekte, und deshalb können die obigen Überlegungen auf die kartographische Darstellung übertragen werden. Im folgenden Unterabschnitt über Attribute ist ein Beispiel hierzu zu finden (Tabelle 5.4 auf Seite 116).

Zur Lösung dieser uneindeutigen Zuordnungen im Kontext der hier behandelten invarianten thematischen Eigenschaften auf Objektartenebene bieten sich zwei unterschiedliche Vorgehensweisen an. Für die erste Lösung betrachtet man die Zuordnungsfunktion g_O nicht mehr als *Funktion*, sondern als *Relation* $R_O \subseteq \mathcal{O}^U \times \mathcal{O}^Z$. Das oben gezeigte Diagramm erhält dann folgende Ausprägung:

$$\begin{array}{ccc}
 U & \xrightarrow{o_U} & \mathcal{O}^U \\
 \downarrow f & & \uparrow R_O \\
 Z & \xrightarrow{o_Z} & \mathcal{O}^Z
 \end{array}$$

FKT Funktion	
1701	Hauptwirtschaftsweg, Verbindungsweg (Fahrweg)
1702	Wirtschaftsweg (Feld-, Waldweg)
1703	Fußweg
1704	Park-, Friedhofsweg
1705	Karren- und Ziehweg
1706	Radweg
1707	Reitweg
1709	(Kletter-)Steig im Gebirge
1710	Rad- und Fußweg
1711	Skaterstrecke
9999	sonstige

Tabelle 5.1: *Vorgesehene Attributwerte für das Attribut FKT Funktion der Objektart 3102 Weg im ATKIS-OK 25. Jeder Attributwert definiert eine Unterklasse zur Klasse der Wege.*

An Stelle der Zuordnungsfunktion g_O ist hier die Zuordnungsrelation R_O getreten. Die Formulierung der invarianten Eigenschaften der Objektarten wird damit von einer Gleichung zu folgender Bedingung umgewandelt:

$$\forall x \in U \text{ gilt: } (o_U(x), o_Z(f(x))) \in R_O \quad (5.2)$$

Diese Lösung ist jedoch unbefriedigend, weil sie zwar die Attribute berücksichtigt, aber nicht explizit macht. Eine explizite Berücksichtigung der Attribute findet sich in der im folgenden Unterabschnitt über die invarianten Eigenschaften der Generalisierungsfunktion bezüglich der Attribute vorgestellten Lösung.

5.2.2 Attribute

Die zweite Lösung verzichtet auf die Einführung einer Relation und behält die Funktion bei, weitet aber die Klassifizierung auf, indem sie die Attribute der Geodaten mit einbezieht.

Wie bereits oben erwähnt, führt die die Attributierung der Geodaten zu einer feineren Gliederung der Landschaft; feiner, als es mit den Objektarten allein gelingen könnte. Die Geodaten werden durch die Attribute, die mit einem Attributwert besetzt sind, genauer charakterisiert. Da ein Geodatum für ein Attribut nur einen Attributwert erhalten kann, wird durch ein Attribut die Klasse einer Objektart in *Unterklassen* eingeteilt. Das oben bereits genannte Beispiel der ATKIS-Objektart 3102 Weg aus dem OK 25 wird durch das Attribut FKT Funktion entsprechend Tabelle 5.1 in elf Unterklassen eingeteilt.

Dieses Beispiel nimmt die Unterklassenbildung an Hand eines *qualitativen Attributes* vor. Diese Art von Attributen eignen sich besonders zur Unterklassenbildung, da sie im Allgemeinen nur eine endliche, fest definierte Menge von Attributwerten vorsieht.

Aber auch *quantitative Attribute* definieren Unterklassen. Allerdings hat man es hier im Allgemeinen mit einer sehr viel höheren Anzahl von Unterklassen zu tun, weil quantitative Attribute aus einer kontinuierlichen Wertemenge gewählt werden können, wie zum Beispiel beim Attribut **BRF Breite der Fahrbahn** der Klasse **3101 Straße**. Hier erfolgt die Breitenangabe in Dezimeter und lässt so zumindest theoretisch eine beliebige Anzahl von Unterklassen zu, nämlich die Klassen der Straßen mit einer Breite von 0 dm, 1 dm, 2 dm und so fort. In der Praxis werden natürlich nur Klassen mit einer sinnvollen Maßzahl für die Breite entstehen. Bei anderen quantitativen Attributen werden *Bereiche* (Intervalle) für die Klassenbildung vorgegeben, wie zum Beispiel beim Attribut **BRV Breite des Verkehrsweges**, das nur Werte von 6, 9, 12, 15 und so fort annehmen kann, entsprechend einer Breite von 0 bis 6 m, 6 m bis 9 m und so fort.

Auch *bezeichnende Attribute* bilden Unterklassen. Die meisten dieser Unterklassen enthalten nur ein Element, wie zum Beispiel die Unterklasse aller Ortslagen in Deutschland mit geographischem Namen „Bonn“. Andererseits kann man auch die Unterklasse aller Straßen mit geographischem Namen „Lindenstraße“ bilden, die bezogen auf Deutschland sicherlich mehr als nur ein Element beherbergt.

Wenn man mehrere Attribute gleichzeitig betrachtet, kann man die Unterklassen kombinieren, was zu weiteren Unterklassen führt. Nimmt man bei Wegen zusätzlich zum Attribut **FKT Funktion** auch das Attribut **BEF Befestigung**, so ergeben sich die Unterklassen der befestigten Fahrwege und der unbefestigten Fahrwege, der befestigten und unbefestigten Wirtschaftswege, Fußwege und so weiter. Oder man betrachtet zum Beispiel die Unterklasse aller Straßen mit einer Breite des Fahrwegs unter 6 m, die den Namen „Lindenstraße“ tragen.

Es stellt sich natürlich die Frage nach dem Nutzen solcher Überlegungen. Sinn macht die Bildung solcher Unterklassen im Zusammenhang mit den invarianten Eigenschaften der Generalisierungsfunktion nur dann, wenn sie bei der Formulierung der Bedingungen für die Invarianz behilflich ist. Im vorigen Unterabschnitt wurden die Fälle angesprochen, bei denen eine Objektart des Ausgangsstrukturierungsgrades nicht eindeutig auf eine Objektart des Zielstrukturierungsgrades abgebildet werden konnte. In all diesen Fällen kann man die Eindeutigkeit der Abbildung aber wieder herstellen, indem man die Objektarten in Unterklassen gemäß der entscheidenden Attribute einteilt.

Die Objektart **AK120 Park** in TOPIS aus dem oben genannten Beispiel kann in vier Unterklassen unterteilt werden, die dann entsprechend durch die Objektartenzuordnungsfunktion g_O eindeutig auf bestimmte Objektarten abgebildet werden. Tabelle 5.2 zeigt die Zuordnungsfunktion für dieses Beispiel. Ebenso können die anderen im vorigen Unterabschnitt genannten Beispiele durch die Unterklassenbildung zur Eindeutigkeit der Zuordnungsfunktion beitragen. Tabelle 5.3 zeigt das Beispiel für die Generalisierung und gleichzeitige Überführung von TOPIS-Daten in ATKIS-Daten für die Objektart **AL020**

5.2 Invariante thematische Eigenschaften

(TOPIS-OK hoher Auflösung)	$\xrightarrow{g_O}$	(TOPIS-OK mittlerer Auflösung)
AK120 Park mit US1 = 000 unbekannt	\mapsto	AK120 Park
AK120 Park mit US1 = 004 national	\mapsto	AK120 Park
AK120 Park mit US1 = 006 privat	\mapsto	EC015 Wald/Forst
AK120 Park mit US1 = 270 öffentlich	\mapsto	EC015 Wald/Forst

Tabelle 5.2: Unterklassenbildung im TOPIS-OK hoher Auflösung zur Erhaltung der Eindeutigkeit der Objektartenzuordnungsfunktion g_O (Beispiel nach Schürer 2002).

(TOPIS-OK)	$\xrightarrow{g_O}$	(ATKIS-OK)
AL020 bebautes Gebiet mit US2 Nutzung2/Funktion = 000 Unbekannt	\mapsto	2101 Ortslage
AL020 bebautes Gebiet mit US2 Nutzung2/Funktion = 544 Wohngebiet	\mapsto	2111 Wohnbaufläche
AL020 bebautes Gebiet mit US2 Nutzung2/Funktion = 510 Industrie- und Gewerbefläche	\mapsto	2112 Industrie- und Gewerbefläche
AL020 bebautes Gebiet mit US2 Nutzung2/Funktion = 528 Fläche gemischter Nutzung	\mapsto	2113 Fläche gemischter Nutzung
AL020 bebautes Gebiet mit US2 Nutzung2/Funktion = 524 Fläche besonderer funktionaler Prägung	\mapsto	2114 Fläche besonderer funktionaler Prägung

Tabelle 5.3: Unterklassenbildung im TOPIS-OK zur Erhaltung der Eindeutigkeit der Objektartenzuordnungsfunktion g_O bei der Generalisierung nach ATKIS.





Objektart	Attribute und -werte	Signaturenklasse
3101 Straße	FKT Funktion = 2301 Straßenverkehr WDM Widmung = 1303 Bundesstraße ZUS Zustand = 1100 in Betrieb BRF Breite der Fahrbahn ≥ 12 m	
3101 Straße	FKT Funktion = 2301 Straßenverkehr WDM Widmung = 1303 Bundesstraße ZUS Zustand = 1300 im Bau BRF Breite der Fahrbahn ≥ 12 m	
3101 Straße	FKT Funktion = 2301 Straßenverkehr WDM Widmung = 1305 Landesstraße ZUS Zustand = 1100 in Betrieb BRF Breite der Fahrbahn ≥ 12 m	
3101 Straße	FKT Funktion = 2301 Straßenverkehr WDM Widmung = 1303 Bundesstraße ZUS Zustand = 1100 in Betrieb BRF Breite der Fahrbahn < 6 m	

Tabelle 5.4: Unterklassenbildung bei der Signaturierung von Straßen (nach dem ATKIS Signaturenkatalog 1 : 25 000, AdV 1998).

bebautes Gebiet und Tabelle 5.4 zeigt das Beispiel für die Zuordnung von Signaturen für die ATKIS-Objektart 3101 Straße zu den entsprechenden Signaturen des Signaturenkataloges.

Diese Unterklassenbildung ist die oben angesprochene zweite Lösung, mit der man bei der eventuellen Mehrdeutigkeit der Objektartenzuordnung auf der Ebene der Objektarten nicht auf die Relation zurückgreifen braucht, wie sie in Bedingung (5.2) gegeben wurde, sondern die ursprüngliche Bedingung (5.1) benutzen kann, die die Objektartenzuordnung durch die Funktion g_O ausdrückt.

Auch für die weiteren Attribute, die nicht zur Unterklassenbildung der Objektarten dienen, erhält man Funktionen, die von der Attributwertemenge des Ausgangsstrukturierungsgrades zur Attributwertemenge des Zielstrukturierungsgrades abbilden. Für jedes Attribut erhält man damit eine Funktion g_A , die analog zu der Objektartenzuordnungsfunktion aufgebaut ist. Schürer (2002, S. 105) gibt in Tabellenform Beispiele für die Zuordnung von Attributwerten zwischen unterschiedlichen Strukturierungsgraden. Er liefert Beispiele der 1 : 1-Überführung, also einer bijektiven Funktion, sowie einer n : 1-Beziehung, also einer nicht-injektiven Funktion. Unter letzterer kann man die Fälle zusammenfassen, in denen mehrere Attributwerte des Ausgangsstrukturierungsgrades zu einem Attributwert des Zielstrukturierungsgrades zusammengefasst werden, sowie die Fälle, in denen es für die Werte im Ausgangsstrukturierungsgrad im Zielstrukturierungsgrad keinen adäquaten Wert mehr gibt und ein allgemeiner Wert (in ATKIS meistens 9999 Sonstiges) zugewiesen wird.

5.2.3 Auswirkungen von Aggregationen

Bei der Aggregation von Geodaten im Zuge der Generalisierung können Zusammenfassungen von Geodaten verschiedener Objektarten auftreten. Insbesondere bei flächenförmigen Objekten tritt dieses Phänomen auf, weil durch die Generalisierung die kleinteilige Gliederung der ungeneralisierten Daten verringert wird. In digitalen Landschaftsmodellen, bei denen im Allgemeinen eine flächendeckende Erfassung vorliegt, aber auch bei thematischen Karten wie geologischen Karten, bei denen ebenfalls eine Flächendeckung vorliegt, ist eine Zusammenfassung von Geodaten verschiedener Objektarten unumgänglich.

Wenn Geodaten aus verschiedenen Klassen zusammengefasst werden, dann muss dem resultierenden Geodatum genau eine Objektart zugewiesen werden. Bei flächenförmigen Objekten orientiert man sich dabei im Allgemeinen an dem ungeneralisierten Geodatum, das flächenmäßig das größte ist, oder an der Objektart, die an den ungeneralisierten Geodaten den größten Anteil hat. Weitere Kriterien finden sich bei Bobzien (2001). Entsprechend der in den beiden vorangegangenen Unterabschnitten beschriebenen Klassifizierung und der Zuordnungsfunktion g_O ergibt sich daraus die Objektart und damit die Klasse des generalisierten Datums.

In diesem Zusammenhang stellt sich die Frage, wie die Verträglichkeit der Generalisierungsfunktion f im Zusammenspiel mit den aggregierten Geodaten der nicht dominierenden Objektart oder Klasse gewährleistet ist. Denn die aufgestellte Bedingung (5.1) soll ja für die Abbildung aller aggregierter Geodaten gelten.

Das Beispiel aus Abbildung 5.6 verdeutlicht schematisch die Problematik bei der Aggregation von zwei flächenförmigen Objekten. Das Beispiel kann ohne weiteres auf die Aggregation mehrere Objekte verallgemeinert werden. In der Abbildung wird ein Wohngebiet a mit einem See b zusammengefasst.

Diese Zusammenfassung kann man auf zwei unterschiedliche Arten interpretieren: Zum einen kann man eine Zusammenfassung auf der thematischen Ebene erkennen, bei der die Realweltobjekte „Wohngebiet“ und „See“ zu einem „bebauten Gebiet“ zusammengefasst

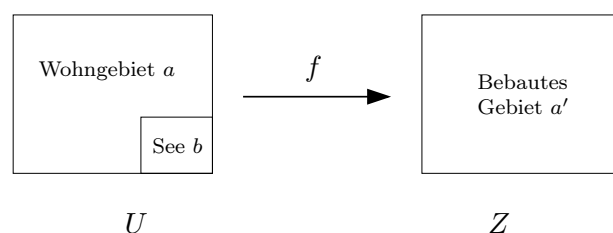


Abbildung 5.6: Schematisches Beispiel für eine Aggregation flächenförmiger Objekte: Ein Wohngebiet und ein See werden zu einem bebauten Gebiet zusammengefasst.

wurden. Von der Generalisierungsfunktion her betrachtet bedeutet dies, dass jeweils beide Objekte aus der Urmenge auf ein und dasselbe Objekt der Zielmenge abgebildet werden. In Kapitel 4 wurde auf diese Art die Aggregation beschrieben.

Für die in Unterabschnitt 5.2.1 eingeführte Objektartenzuordnungsrelation R_O heißt das, dass sie die Objektarten „Wohngebiet“ und „See“ mit der Objektart „Bebautes Gebiet“ verknüpfen muss. Führt man diesen Gedanken fort, so muss im Prinzip fast jede Objektart mit jeder anderen Objektart verknüpft werden, da immer Situationen auftreten können, bei denen ein kleines Objekt der einen Objektart einem benachbarten größeren Objekt einer anderen Objektart zugeschlagen werden muss. Damit würde die Zuordnungsrelation aber eine sehr große Teilmenge des kartesischen Produktes von \mathcal{O}^U und \mathcal{O}^Z abdecken und die formulierte Bedingung (5.2) würde ihre Bedeutung verlieren.

Aus diesem Grund wird im Folgenden eine andere Interpretation der Aggregation bevorzugt, bei der man die Zusammenfassung eher auf der topologischen und geometrischen Ebene sieht. Für das Beispiel heißt das, dass die *Flächen* des Wohngebietes und des Sees zusammengefasst wurden, zu einer Fläche, die mit der Thematik „Bebautes Gebiet“ verbunden ist. In dieser Interpretation fällt der See bei der Generalisierung weg, das heißt, die Generalisierungsfunktion ist für dieses Geodatum nicht definiert. Dennoch fallen die Flächen dieser Daten natürlich nicht weg, sondern sie werden mit den Flächen der anderen Objekte zusammengefasst.

Für die Generalisierungsfunktion bedeutet dies, dass nicht zwischen den Objekten als Ganzes abgebildet wird, sondern die Abbildung zwischen den Flächen, also nur auf der topologischen und geometrischen Ebene betrachtet wird. In Abschnitt 4.2.1 wurde dieses Problem bereits angedeutet und eine Aufteilung der Generalisierungsfunktion auf zwei Teilfunktionen, die eine zwischen den thematischen Ebenen und die andere zwischen den topologischen Ebenen beider Datenmodelle, vorgeschlagen. Abbildung 4.3 (b) auf Seite 70 verdeutlicht diesen Ansatz noch im Kontext der Relation zwischen den Datenmodellen. Dieser Ansatz wird kurz skizziert nun auf die Generalisierungsfunktion übertragen. Zunächst werden die Urmenge U und die Zielmenge Z aufgeteilt in ihre Bestandteile der Thematik, Topologie und Geometrie: $U = (U_{\text{Them}}, U_{\text{Topo}}, U_{\text{Geom}})$ und $Z = (Z_{\text{Them}}, Z_{\text{Topo}}, Z_{\text{Geom}})$. Die Generalisierungsfunktion f wird nun aufgeteilt auf zwei Teilfunktionen $f_{\text{Them}} : U_{\text{Them}} \rightarrow Z_{\text{Them}}$ und $f_{\text{Topo}} : U_{\text{Topo}} \rightarrow Z_{\text{Topo}}$. Auf der geometrischen Ebene findet wie in Abschnitt 4.2.1 erläutert keine Verknüpfung statt. Die ausführliche Begründung dazu findet sich in Abschnitt 3.2.2.

Angewendet auf das obige Beispiel ergibt sich, dass zunächst die Geodaten des Wohngebietes a und des Sees b aufgeteilt werden müssen in ihre Bestandteile der Thematik, Topologie und Geometrie, so dass $a = (a_{\text{Them}}, a_{\text{Topo}}, a_{\text{Geom}})$ und $b = (b_{\text{Them}}, b_{\text{Topo}}, b_{\text{Geom}})$ ist. Die Abbildung der beiden Geodaten wird in Abbildung 5.7 veranschaulicht und durch die folgende Abbildungsvorschriften formalisiert:

$$\begin{array}{ccc} a_{\text{Them}} & \xrightarrow{f_{\text{Them}}} & a'_{\text{Them}} & & b_{\text{Them}} & \xrightarrow{f_{\text{Them}}} & \text{nicht def.} \\ a_{\text{Topo}} & \xrightarrow{f_{\text{Topo}}} & a'_{\text{Topo}} & & b_{\text{Topo}} & \xrightarrow{f_{\text{Topo}}} & a'_{\text{Topo}} \end{array}$$

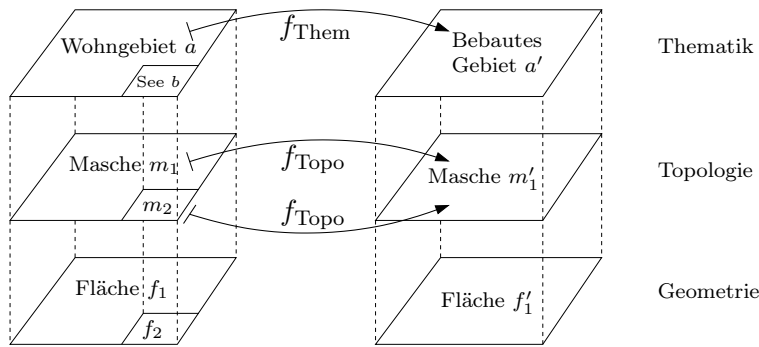


Abbildung 5.7: Aggregation zweier flächenförmiger Geodaten mit Aufteilung der Generalisierungsfunktion in zwei Teilfunktionen.

Diese Abbildungsvorschriften werden folgendermaßen interpretiert: Die Thematik des Geo-Objektes See wurde beim Generalisieren fortgelassen, aber seine Topologie (also die Masche) wurde mit der Topologie des Wohngebietes zusammengefasst. Die Aggregation findet demnach auf der *Ebene der Topologie* statt, nicht auf der *Ebene der Thematik*.

Auswirkungen von Aggregationen auf Attribute

Ebenso wie bei den Objektarten müssen auch die Attribute und ihre Werte im Zuge von Aggregationen gesondert behandelt werden. Schürer (2002, Kap. 5.1.3.2) behandelt die Überführung von Attributen bei der Zusammenfassung von Geodaten nur im gleichen Strukturierungsgrad. Für die Generalisierung bedeutet dies, dass die Geodaten im Ausgangsstrukturierungsgrad zusammengefasst und die Attribute überführt werden, und dass in einem zweiten Schritt die aggregierten Objekte generalisiert werden (Schürer 2002, ebd.). Eine zweite Möglichkeit bestünde in der einzelnen Überführung der Objekte in den Zielstrukturierungsgrad und einer in einem zweiten Schritt durchgeführten Zusammenfassung der Objekte und ihrer Attribute im Zielstrukturierungsgrad. Beide Vorgehensweisen erfordern für die Geodaten ein Zwischenmodell, entweder im Strukturierungsgrad der Ausgangsdaten oder im Strukturierungsgrad der Zieldaten.

In diesem Abschnitt wird ein Ansatz entwickelt, der die Aggregation und die dabei stattfindende Attributüberführung in einem Schritt durchführt und damit mit dem in Abschnitt 3.2.2 eingeführten Datenmodell kompatibel bleibt. Dabei wird auf die oben getroffene Einteilung in qualitative, quantitative und beschreibende Attribute zurückgegriffen. Die Einteilung wird in diesem Abschnitt weiter verfeinert. Sie stimmt in großen Teilen mit der von Schürer vorgeschlagenen Einteilung überein. Das von Schürer aufgestellte Regelwerk zur Ableitung der Attribute kann auch in weiten Teilen hierfür übernommen werden.

Für *qualitative Attribute* ergeben sich dieselben Betrachtungen wie für die Objektarten oben. Zum einen hat man für jedes Attribut eine Überführungsfunktion g_A , die die Attri-

butwerte der beiden Strukturierungsgrade einander zuordnet. Zum anderen darf ebenso wie bei den Objektarten die Interpretation der Zusammenfassung nur auf der topologischen Ebene stattfinden, da sonst die Überföhrungsfunktion zu einer Relation bestehend aus dem kartesischen Produkt der beiden Mengen von Attributwerten im Ausgangsbeziehungsweise Zielstrukturierungsgrades mutiert. Die Argumentation hierzu verläuft analog zu der oben Vorgestellten im Zusammenhang mit den Objektarten.

Bei den *quantitativen Attributen* erfolgt eine weitere Unterteilung in *dominierende*, *mittelnde* und *summierende Attribute*.⁵ *Dominierende Attribute* erhalten einen durch eine Ordnungsrelation (größer als/kleiner als) bestimmten Attributwert, und zwar entweder das Maximum oder das Minimum aller zu aggregierenden Werte. Als Beispiel nennt Schürer das Attribut **HHO Objekthöhe**, das unter anderem bei Bauwerken die höchste Höhe angibt und deshalb bei der Aggregation auf das *Maximum* der Werte gesetzt werden muss. Analog kann beim Attribut **DHU Durchfahrtshöhe** das *Minimum* aller Werte gesetzt werden.

Mittelnde Attribute „geben bei der Aggregation durchschnittliche Eigenschaften wieder“ (Schürer 2002). Für ihre Berechnung müssen alle Attributwerte des Ausgangsattributes berücksichtigt werden und (gegebenenfalls gewichtet) gemittelt werden. Schürer nennt als Beispiel das Attribut **BWD Bauwerksdichte** aus einer älteren ATKIS-Version (AdV 1989), bei dessen Berechnung die Werte aller aggregierten Objekte flächenmäßig gewichtet in den generalisierten Attributwert einfließen.

Summierende Attribute erhalten bei der Aggregation als Wert die Summe aller aggregierten Attributwerte. Ein Beispiel hierfür ist das Attribut **GLZ Anzahl der Gleise auf dem Bahnkörper**, dessen Werte bei der Aggregation von Schienenbahnen aufsummiert werden müssen. Die mittelnden und die summierenden Attribute erhalten ihre Werte durch eine algebraische *Berechnung* und können somit auf algebraische Strukturen zurückgeführt werden.

Die letzte Attributart, die bei der Aggregation berücksichtigt werden muss, ist die Gruppe der *bezeichnenden Attribute*. Diese Attribute lassen sich nur schwer automatisch aggregieren. Beim Zusammenfassen von zusammenhängenden Straßen – falls die Objektbildungsregeln dieses erlauben – können zum Beispiel geographische Namen nicht einfach wie oben berechnet werden. Ein heuristischer Ansatz zur automatischen Bearbeitung dieses Attributüberganges könnte eine Vorgehensweise wie bei den dominierenden Attributen sein, indem der Wert des längsten oder flächenmäßig größten Objektes übernommen wird.

⁵Schürer (2002) bezeichnet die mittelnden Attribute als *berechnende Attribute*. Summierende Attribute betrachtet er nicht. Diese müssten nach seiner Unterteilung ebenfalls unter die berechnenden Attribute fallen.

5.3 Invariante topologische Eigenschaften

Der vorige Abschnitt behandelte die invarianten Bedingungen, die sich vom Datenmodell aus auf der thematischen Ebene ableiten lassen. Dieser Abschnitt verlässt die thematische Ebene und formuliert die invarianten Bedingungen an die Generalisierungsfunktion auf der *topologischen Ebene* des Datenmodells. Diese Bedingungen sind – ebenso wie die Bedingungen auf der thematischen Ebene – *kategorische Bedingungen*, wie sie in der Einleitung zu diesem Kapitel beschrieben wurden.

Vorangestellt seien einige einleitende Gedanken über den Begriff der Topologie in der Geoinformation und in der Kartographie. Die Begriffe *topologischer Raum* und *Topologie* entstammen der Mathematik. Wie in der Einleitung zu dieser Arbeit erwähnt, beschränken wir uns auf Topologien in der euklidischen Ebene \mathbb{R}^2 . Generell kann unterschieden werden zwischen *Punktmengentopologie* (auch als *analytische Topologie* bezeichnet) auf der einen Seite als die Topologie, die sich mit Nachbarschaft und Nähe auf offenen Mengen beschäftigt, und *kombinatorischer Topologie* (auch *algebraische* oder *geometrische Topologie* genannt) auf der anderen Seite als die Topologie mit dem Thema der Zellkomplexe, Simplexes und Graphen (Worboys 2004).

Egenhofer et. al. gehen aus von der Punktmengentopologie und definieren mit Hilfe des *4-Schnitt-Modells* (Egenhofer und Herring 1990a, Egenhofer und Franzosa 1991) und des *9-Schnitt-Modells* (Egenhofer 1991) topologische Relationen zwischen punkt-, linien- und flächenförmigen Objekten in der Ebene. Sie interpretieren diese topologischen Relationen *geometrisch* und spannen damit den Bogen von der Punktmengentopologie zur kombinatorischen Topologie. Den so interpretierten Relationen geben sie umgangssprachlich gebräuchliche Namen. Eine Auswahl der benannten Relationen gibt Tabelle 5.5 wieder. Die Relationen *meet/touch* und *inside* werden in diesem Abschnitt im Zusammenhang mit der Generalisierung vertieft behandelt.

In dieser Arbeit wird die Topologie nicht von den Punktmengen her betrachtet, sondern von der Graphentheorie her (siehe Abschnitt 3.1.2 auf den Seiten 43 ff.). Damit entspricht diese Herangehensweise derjenigen der kombinatorischen Topologie. Die im

englisch	deutsch	behandelt in
disjoint	disjunkt	5.3.1 – Inzidenz und Adjazenz
meet/touch	treffend/berührend	
overlap	überlappend	
covers/covered by	überdeckend/überdeckt von	5.3.2 – Inklusion
inside/contains	innen liegend/enthaltend	
equal	ist gleich	

Tabelle 5.5: *Topologische Relationen in der kombinatorischen Topologie (nach Egenhofer und Franzosa 1991).*

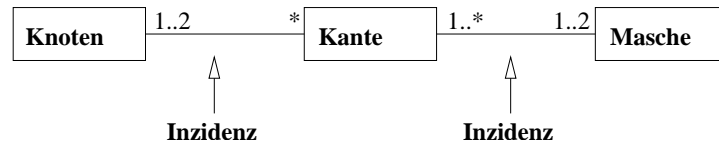


Abbildung 5.8: Die topologische Ebene des erweiterten hierarchischen Datenmodells (Ausschnitt aus Abbildung 3.5) sowie die explizit gespeicherten Inzidenzen.

Datenmodell *explizit* gespeicherten topologischen Beziehungen sind die Inzidenzen zwischen den Klassen Knoten und Kante und zwischen den Klassen Kante und Masche (siehe Abbildung 5.8). Aus diesen Beziehungen lassen sich weitere topologische Beziehungen der Inzidenz und der Adjazenz ableiten, diese sind also *implizit* gespeichert. Eine genauere Erläuterung zu Inzidenzen und Adjazenzen sowie zu den invarianten Eigenschaften bezüglich dieser beiden Arten von Relationen gibt Abschnitt 5.3.1.

Eine weitere Art von topologischer Relation, die auf Invarianz untersucht wird, ist die *Inklusionsrelation* (enthalten-in-Relation). Diese topologische Relation ist mit dem hier verwendeten erweiterten hierarchischen Datenmodell nicht allein durch Analyse der Beziehungen auf der topologischen Ebene ableitbar, sondern es muss auch die geometrische Ebene ausgewertet werden. Diese Relation wird dennoch in diesem Abschnitt der invarianten *topologischen* Eigenschaften behandelt, weil es eine rein topologische Relation ist (Egenhofer und Herring 1990a). Die Inklusionsrelation und die invarianten Eigenschaften bezüglich dieser Relation werden in Abschnitt 5.3.2 behandelt.

Weitere topologische Relationen, die in dieser Arbeit jedoch nicht weiter berücksichtigt werden, sind in Tabelle 5.5 aufgeführt. Die Ansätze zur Invarianz der hier behandelten topologischen Relationen können auf diese Relationen übertragen werden. Das Prinzip der Topologieerhaltung wird bereits aus den hier vertieft behandelten Relationen deutlich.

Anmerkend sei noch einmal darauf hingewiesen, dass die hier betrachtete Topologie nicht identisch ist mit der in der amtlichen topographischen Kartographie gebräuchlichen Topologie, obwohl teilweise Übereinstimmungen herrschen, die auf den ersten Blick auf Gleichheit schließen lassen. In ATKIS werden zum einen mehrere topologische Netze nebeneinander betrachtet, zum Beispiel das Schienennetz, das Straßennetz, das Gewässernetz und das Stromleitungsnetz, die sich zwar schneiden können aber nicht zwingend topologische Elemente bei den jeweils anderen Netzen ausbilden. Eine ausführliche Diskussion zur Objektbildung findet sich bei Schürer (2002, Abschnitt 3.3.2.3). Zum anderen sind diese Netze nicht im graphentheoretischen Sinne in die Ebene eingebettet. Wenn Straßen und/oder Schienenbahnstrecken sich nicht niveaugleich kreuzen, dann wird nicht zwingend ein topologischer Knoten gebildet (siehe Diskussion in Abschnitt 3.1.6, S. 24 f.).

5.3.1 Inzidenz und Adjazenz

Zunächst werden die Begriffe Inzidenz und Adjazenz definiert und formal erläutert. Danach wird die Bedingung der Invarianz formuliert und an Hand von Beispielen erklärt. Den Abschluss dieses Unterabschnittes bilden die Ausnahmen, die in Folge von Geometriypwechseln entstehen können.

Inzidenzen

In der Graphentheorie versteht man unter dem Begriff der *Inzidenz* den direkten Zusammenhang von topologischen Elementen *unterschiedlicher Ausprägung* (hier von unterschiedlichem Geometriyp, siehe Abschnitt 3.1.3). Mit den drei Arten von Topologieelementen, den Knoten, Kanten und Maschen, gibt es also drei Arten von Inzidenz, die Knoten-Kanten-Inzidenz, die Kanten-Maschen-Inzidenz und die Knoten-Maschen-Inzidenz (Aigner 1984, Findeisen 1990, Gröger 2000). Abbildung 5.9 zeigt Beispiele für die drei Arten der Inzidenzen. Im *erweiterten hierarchischem Datenmodell* (siehe Kapitel 3) sind genau zwei Inzidenzen direkt repräsentiert, nämlich die Knoten-Kanten-Inzidenz und die Kanten-Maschen-Inzidenz (siehe Abbildung 5.8). Findeisen (1990) bezeichnet diese Inzidenzen zusammen mit zwei Adjazenzen (Definition siehe unten) als *Grundbeziehungen*, aus denen sich alle weiteren Inzidenzen und Adjazenzen ableiten lassen. Tatsächlich genügen aber die beiden im Datenmodell explizit gespeicherten Inzidenzen, um auch die dritte Inzidenz und alle Adjazenzen daraus abzuleiten.

Formal bezeichnen wir die drei Ausprägungen der topologischen Elemente (bezogen auf einen Strukturierungsgrad) mit der Menge der Knoten V (englisch *vertices*), der Menge der Kanten E (englisch *edges*), und der Menge Maschen F (englisch *faces*). Die drei

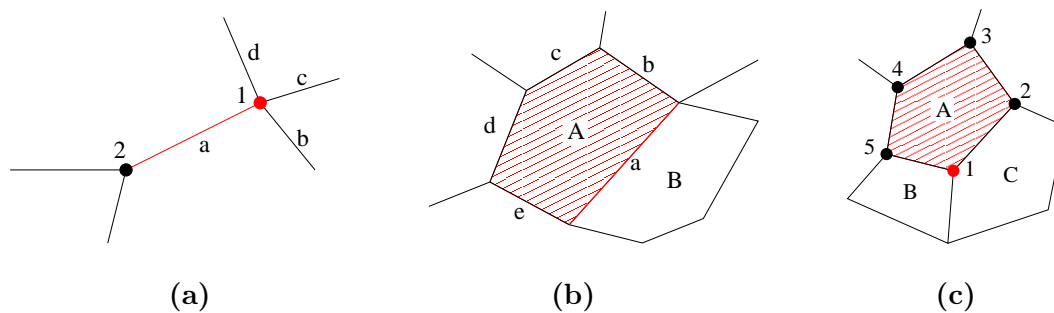


Abbildung 5.9: Beispiele für die drei Arten von Inzidenzen. Die bezeichneten Elemente sind zu dem rot markierten Element des jeweils anderen Geometriyps inzident: (a) Knoten-Kanten-Inzidenz, (b) Kanten-Maschen-Inzidenz, (c) Knoten-Maschen-Inzidenz.

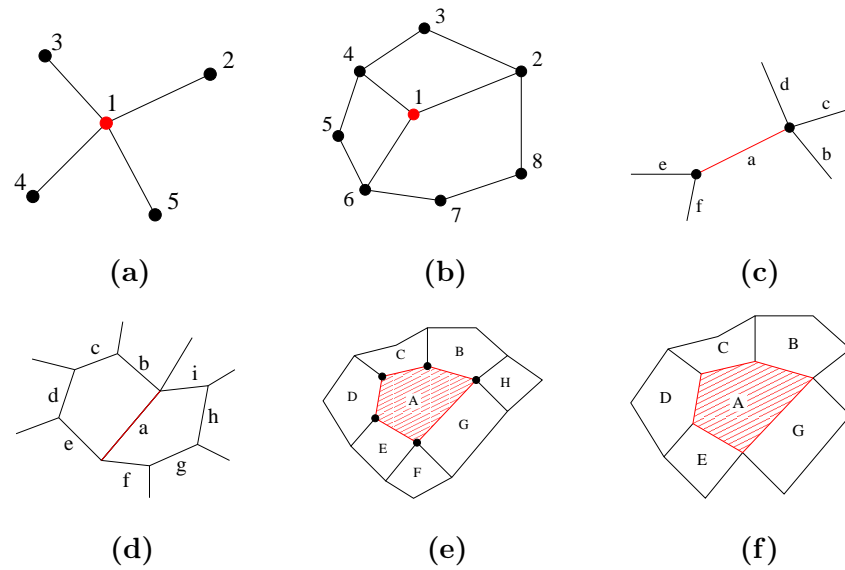


Abbildung 5.10: Beispiele für die sechs Arten von Adjazenzen. Die bezeichneten Elemente sind sämtliche zu dem rot markierten Element adjazenten Elemente: (a) Knoten-Adjazenz über Kanten, (b) Knoten-Adjazenz über Maschen, (c) Kanten-Adjazenz über Knoten, (d) Kanten-Adjazenz über Maschen, (e) Maschen-Adjazenz über Knoten, (f) Maschen-Adjazenz über Kanten.

Inzidenzrelationen, hier mit I_1 bis I_3 bezeichnet, lauten dann (Definition 5.1):

$$\begin{aligned}
 I_1 &\subseteq V \times E && \text{Knoten-Kanten-Inzidenz} \\
 I_2 &\subseteq E \times F && \text{Kanten-Maschen-Inzidenz} \\
 I_3 &\subseteq V \times F && \text{Knoten-Maschen-Inzidenz}
 \end{aligned}$$

Dabei sind die Relationen I_1 und I_2 direkt aus dem Datenmodell ableitbar. Die Relation I_3 ist durch eine Verknüpfung der beiden vorgenannten Relationen ableitbar.

Adjazenzen

Der Begriff *Adjazenz* bezeichnet im Gegensatz zur Inzidenz eine Relation zwischen Topologieelementen *gleicher Ausprägung* (also gleichen Geometrietyps) mit bestimmten Eigenschaften. In der Literatur findet man im Zusammenhang mit der Graphentheorie unterschiedliche Definitionen. Gröger (2000) unterscheidet drei Arten der Adjazenz, je eine für Knoten, Kanten und Maschen. Findeisen (1990) unterscheidet vier Arten der Adjazenz, er unterteilt die Kanten-Adjazenz abhängig davon, ob diese über *inzidente Knoten* oder *inzidente Maschen* definiert ist. Den Definitionen gemeinsam ist, dass zwei Topologieelemente dann als zueinander adjazent gelten, wenn sie zu *demselben* Topolo-

gieelement einer *anderen* Ausprägung *inzident* sind.⁶ Ausgehend von dieser Definition der Adjazenz ergeben sich insgesamt sechs mögliche Arten der Adjazenz, jeweils zwei für die Adjazenzen der Topologieelemente gleicher Ausprägung. Die Topologieelemente des anderen Geometrietyps werden im Folgenden *adjazenzbestimmend* genannt. Beispiele für die sechs möglichen Arten der Adjazenz sind in Abbildung 5.10 gezeigt.

Die Adjazenzen sind – ebenso wie die Inzidenzen – Relationen, die sich wie folgt ausdrücken lassen (Definition 5.2):

$A_1 \subseteq V \times V$ über E	Knoten-Adjazenz über inzidente Kanten
$A_2 \subseteq V \times V$ über F	Knoten-Adjazenz über inzidente Maschen
$A_3 \subseteq E \times E$ über V	Kanten-Adjazenz über inzidente Knoten
$A_4 \subseteq E \times E$ über F	Kanten-Adjazenz über inzidente Maschen
$A_5 \subseteq F \times F$ über V	Maschen-Adjazenz über inzidente Knoten
$A_6 \subseteq F \times F$ über E	Maschen-Adjazenz über inzidente Kanten

Es leuchtet ein, dass erstens $A_1 \subseteq A_2$ und zweitens $A_6 \subseteq A_5$ gilt, während $A_3 \subseteq A_4$ nicht zwingend gilt. Ein Gegenbeispiel wird gebildet von den Kanten c in Abbildung 5.10 (c) und (d).

Alle Adjazenzrelationen sind *symmetrisch*, das bedeutet, dass aus der Adjazenz von einem Element a zu einem Element b die Adjazenz von b zu a folgt.

Außerdem definieren wir die Adjazenzen als *reflexiv*, das heißt, dass jedes Topologieelement zu *sich selber* adjazent ist. Diese Definition wird für die Formulierung der Bedingungen weiter unten benötigt. Begründen lässt sich diese Definition aber auch mit weiteren Argumenten: Zum einen ist in der Definition der Adjazenz von zwei Topologieelementen die Rede, die jedoch nicht notwendigerweise unterschiedlich sein müssen. Wenn diese beiden Topologieelemente gleich sein dürfen, ergibt sich schon daraus die Reflexivität.

Zum anderen kann man die Reflexivität aus der Rechenvorschrift zum Finden von adjazenten Elementen begründen: Gegeben sei ein Topologieelement x . Die adjazenten Topologieelemente über die adjazenzbestimmenden Topologieelemente der Ausprägung A findet man, indem man zunächst alle inzidenten Topologieelemente der Ausprägung A zu x bestimmt. Diese Menge sei die Menge Y . Danach bestimmt man alle inzidenten Topologieelemente der Ausprägung von x zu jedem Element von Y . Dies sei die Menge X der zu x adjazenten Topologieelementen. Folgt man dieser Rechenvorschrift, so ist sofort einsichtig, dass $x \in X$ ist, und damit x zu sich selber adjazent ist.⁷

⁶In der Graphentheorie wurde die Adjazenz ursprünglich nur für Knoten und Kanten definiert. Erst die Einbettung von Graphen in die Ebene mittels planarer Graphen (Harary 1972, Even 1979) und ihre Anwendung in der automationsgestützten Kartographie und in der Geoinformation (Corbett 1975, Findeisen 1990, Plümer 1996, Gröger 2000) legen eine Übernahme der Definitionen für Maschen nahe.

⁷Die Reflexivität ergibt sich nach diesem Algorithmus nur, wenn das Topologieelement zu mindestens einem adjazenzbestimmenden Topologieelement inzident ist, was aber der Regelfall ist. Für Maschen trifft dies immer zu. Für Kanten trifft dies bei Adjazenz über Knoten immer zu. Ausnahmen von dem Regelfall sind folgende drei Fälle: Eine *isolierte* Kante ohne inzidente Masche ist nicht zu sich selber

Inzidenz und Adjazenz für Geodaten

Die Inzidenz- und Adjazenzrelationen können nicht nur auf der Topologieebene des Datenmodells definiert werden, sondern auch auf der Ebene der Thematik oder allgemeiner für die Gesamtheit des Geodatums. Hierzu definiert man, dass zwei Geodaten A_1 und A_2 genau dann inzident beziehungsweise adjazent sind, wenn ein Topologieelement t_1 von A_1 inzident beziehungsweise adjazent zu einem Topologieelement t_2 von A_2 ist.⁸

Da die im Folgenden betrachteten Eigenschaften für alle Inzidenz- und Adjazenzrelationen zugleich gelten sollen, werden wir die Inzidenz- und Adjazenzrelationen nicht mehr nur isoliert betrachten, sondern auch gemeinsam. Die Menge aller Topologieelemente eines Strukturierungsgrades sei die Menge $T = V \cup E \cup F$. Dann ist die Menge aller Inzidenzen und Adjazenzen $N = I_1 \cup \dots \cup I_3 \cup A_1 \cup \dots \cup A_6 \subseteq T \times T$. Wenn der Kontext eine eindeutige Zuordnung zu einem Strukturierungsgrad verlangt, kann man die Zugehörigkeit zu einem solchen wieder, wie im vorigen Abschnitt, mit Hochindizes U beziehungsweise Z ausdrücken. Die Relation N , die für die Topologieelemente gilt, kann damit auch auf die Mengen der ungeneralisierten und generalisierten Geodaten übertragen werden. Damit ist dann $N^U \subseteq U \times U$ und $N^Z \subseteq Z \times Z$. Im Folgenden werden wir auf die Hochindizes verzichten, da die Adjazenz und die Inzidenz in den beiden Mengen der ungeneralisierten und generalisierten Daten die gleiche Bedeutung haben (siehe Diskussion hierzu in Abschnitt 5.1.4).

Bedingung der Invarianz

Im Folgenden wird gezeigt, dass jede der neun Inzidenz- und Adjazenzrelationen (vergleiche Definitionen 5.1 und 5.2) durch eine korrekte Generalisierung erhalten bleiben wird. Auf die Gesamtmenge N all dieser Relationen bezogen bedeutet dies, dass auch die Relation N durch eine korrekte Generalisierung erhalten bleibt.

Zur Motivation dieser Behauptung dienen die folgenden Ausführungen: Umgangssprachlich gibt es viele Ausdrücke für Inzidenz und Adjazenz, die teilweise auch für spezielle topologische Beziehungen gebraucht werden, wie zum Beispiel die von Egenhofer und Franzosa (1991) durch die Punktmengentopologie eingeführten Relationen (siehe oben) oder die von Worboys (2004) verwendeten. Es handelt sich hierbei um die Relationen *touch* und *meet* (Berühren und Treffen).

Etwas unscharf lassen sich diese Beziehungen zusammenfassen unter den Begriffen „Aneinander grenzen“ und „Verbunden sein“. Zwei Maschen grenzen aneinander (Maschen-Adjazenz über Knoten oder Kanten), eine Kante grenzt an eine Masche (Kanten-Maschen-

adjazent über eine Masche. Für *isolierte* Knoten trifft die Reflexivität weder über eine inzidente Kante noch Masche zu; für *Endknoten isolierter* Kanten trifft die Reflexivität über inzidente Maschen nicht zu.

⁸Mit dieser Definition kann ein Geodatum auch ohne die Reflexivität der Adjazenz auf Topologieebene zu sich selber adjazent sein, nämlich dann, wenn ein Geodatum aus mehr als einem Topologieelement besteht, und wenn diese Topologieelemente zueinander adjazent sind.

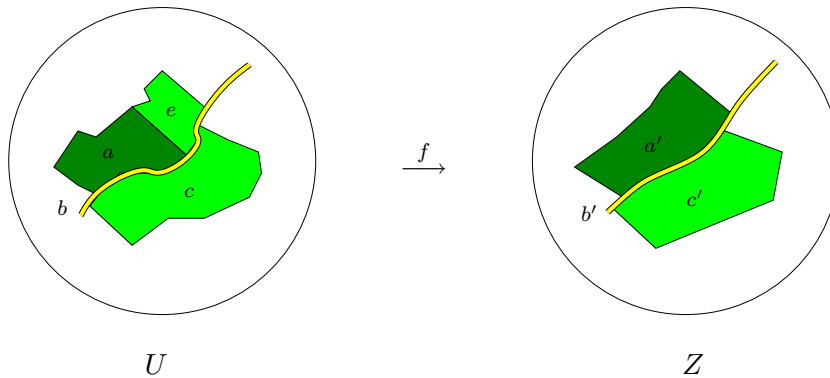


Abbildung 5.11: Die Inzidenz und die Adjazenz der Geodaten bleibt bei einer korrekten Generalisierung erhalten. Die Invarianzbedingung wird auch durch eine Aggregation nicht verletzt.

Inzidenz), zwei Kanten grenzen aneinander (Kanten-Adjazenz über Knoten), eine Kante grenzt an einen Knoten (hat Knoten als Anfangs- oder Endknoten, Kanten-Knoten-Inzidenz). Ein Knoten ist verbunden mit einer Masche (Knoten-Maschen-Inzidenz), zwei Knoten sind miteinander verbunden (Knoten-Adjazenz über Kanten und Maschen), zwei Kanten sind miteinander über eine Masche verbunden (Kanten-Adjazenz über Maschen). Diese umgangssprachlichen Ausdrücke lassen sich auch ohne weiteres auf Geodaten als Ganzes übertragen: Man sagt, dass *zwei Geodaten aneinander grenzen* oder *miteinander verbunden sind*. Diese Beziehung ist unabhängig vom Geometrietyt der Geodaten. Damit lässt sie sich auch für die Bedingung der Invarianz bei der Generalisierung verwenden: Wenn gesagt wird, dass zwei Geodaten aneinander grenzen oder miteinander verbunden sind, dann wird auch erwartet, dass diese Beziehung durch die Generalisierung nicht zerstört wird. Abbildung 5.11 verdeutlicht dies: Waldstück a grenzt an ein Grünland c , getrennt nur durch Straße b . Man erwartet, dass auch das generalisierte Waldstück a' an das generalisierte Grünland c' grenzt. Die Straße b , die an das Waldstück a und das Grünland c grenzt, soll auch nach der Generalisierung als Straße b' an das generalisierte Waldstück a' und das generalisierte Grünland c' grenzen. (Die topologischen Relationen zum bei der Generalisierung durch Aggregation wegfallenden Grünland e werden weiter unten behandelt.)

Die im letzten Absatz unscharf formulierte Bedingung an die Invarianz der topologischen Relationen wird jetzt scharf gefasst durch eine Bedingung, die an die Generalisierungsfunktion gestellt wird. Zunächst werden wir diese Bedingung für die topologischen Elemente formulieren, danach für die Geodaten allgemein.

Für Topologieelemente lautet die Bedingung wie folgt: Zwei Topologieelemente, die vor der Generalisierung aneinander grenzen (die also zueinander inzident oder adjazent sind), sollen auch in ihrer generalisierten Form aneinander grenzen (inzident oder adjazent sein). Voraussetzung ist, dass das entsprechende Topologieelement abgebildet wird, das heißt, dass es durch die Generalisierung nicht wegfällt. Denn wenn ein Topologieele-

ment wegfällt, dann entfällt natürlich auch die Inzidenz beziehungsweise Adjazenz mit diesem Topologieelement. In Abbildung 5.11 entfallen die topologischen Relationen betreffend Grünland e . Formal lässt sich die Bedingung für jede der neun Inzidenz- beziehungsweise Adjazenzrelationen einzeln ausdrücken: Sei die Relation R eine aus den neun Inzidenz- beziehungsweise Adjazenzrelationen, also $R \in \{I_1, \dots, I_3, A_1, \dots, A_6\}$. Dann ist also $R \subseteq T_1 \times T_2$ mit $T_{1,2} \in \{V, E, F\}$, und es gilt:

$$\boxed{\begin{array}{l} \text{Für alle } x \in T_1 \text{ und } y \in T_2, \text{ wobei } f(x), f(y) \text{ definiert sind, gilt:} \\ \text{Aus } (x, y) \in R \text{ folgt } (f(x), f(y)) \in R. \end{array}} \quad (5.3)$$

Diese Bedingung lautet, allgemeiner formuliert und nicht auf die einzelnen topologischen Elemente bezogen, sondern auf die Geodaten als Ganzes, folgendermaßen:

$$\boxed{\begin{array}{l} \text{Für alle } x, y \in U, \text{ wobei } f(x), f(y) \text{ definiert sind, gilt:} \\ \text{Aus } (x, y) \in R \text{ folgt } (f(x), f(y)) \in R. \end{array}} \quad (5.4)$$

Wenn also zwei Geodaten miteinander in Inzidenz- oder Adjazenzrelation stehen, dann gilt dies auch für die generalisierten Daten. Dies ist die invariante Eigenschaft der Generalisierungsfunktion bezüglich der topologischen Beziehungen der Inzidenz und Adjazenz. Diese Bedingung ist eine kategorische Bedingung an die Generalisierungsfunktion, das heißt, dass eine Generalisierung, die diese Eigenschaften nicht erfüllt, falsch erfolgt ist. Da die Bedingung formalisiert ist, kann sie durch ein Rechenprogramm automatisch überprüft werden.

Beispiele und Anmerkungen

In diesem Unterabschnitt werden drei Beispiele gegeben, an denen die Invarianz der Adjazenz ersichtlich wird. Die Invarianz der Inzidenzen ist – da Adjazenzen aus Inzidenzen abgeleitet werden – in diesen Beispielen enthalten. Zwei Besonderheiten der topologischen Invarianz werden an diesen Beispielen deutlich: Zum einen wird gezeigt, warum die Reflexivität der Adjazenz, die in der obigen Definition gefordert wurde, für die Erfüllung der Invarianzbedingung vonnöten ist. Zum anderen wird ersichtlich, dass die Implikation in den oben formulierten Bedingungen nicht umkehrbar ist. Aus der Inzidenz oder Adjazenz von generalisierten Daten kann nicht auf die Inzidenz oder Adjazenz der ungeneralisierten Daten geschlossen werden. Durch die Generalisierung können neue Relationen entstehen, die an den ungeneralisierten Daten nicht absehbar waren.

Abbildung 5.12 zeigt als erstes Beispiel die Erhaltung der Knoten-Adjazenz über Kanten, im Beispiel sind es Knoten des Straßennetzes. Knoten 1 und 2 liegen so nah beieinander, dass durch die Generalisierung beide Knoten zum Knoten $1'$ verschmelzen und die dazwischenliegende Kante entfällt. Damit ist $f(1) = f(2) = 1'$. Die Invarianzbedingung ist erfüllt, da aus $(1, 2) \in A_1$ folgt, dass $(f(1), f(2)) = (1', 1') \in A_1$ ist. Ohne Reflexivität

5.3 Invariante topologische Eigenschaften

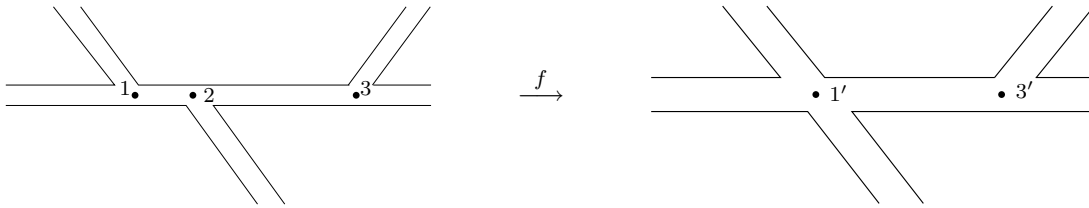


Abbildung 5.12: Beispiel für die Erhaltung der Knoten-Adjazenz über Kanten, sowie für die Reflexivität der Adjazenz und die Nicht-Umkehrbarkeit der Invarianzbedingung.

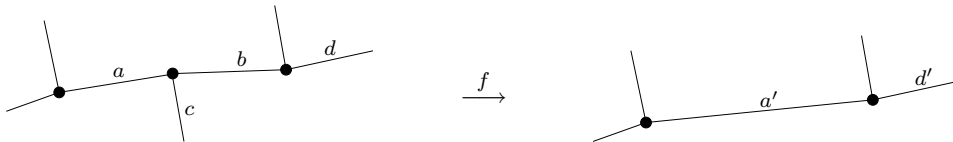


Abbildung 5.13: Beispiel für die Erhaltung der Kanten-Adjazenz über Knoten, sowie für die Reflexivität der Adjazenz und die Nicht-Umkehrbarkeit der Invarianzbedingung.

der Adjazenzrelation wäre die Bedingung nicht erfüllt. Des Weiteren ist die Invarianzbedingung nicht umkehrbar, denn aus $(1', 3') = (f(1), f(3)) \in A_1$ folgt nicht $(1, 3) \in A_1$.

Das zweite Beispiel zeigt die Invarianz der Kanten-Adjazenz über Knoten: In Abbildung 5.13 fällt bei der Generalisierung die Kante c und damit der zu den Kanten a , b und c inzidente Knoten fort. Die Kanten a und b werden zur Kante a' zusammengefasst. Hier wird wieder die Notwendigkeit der Reflexivität der Adjazenzrelation deutlich, denn aus $(a, b) \in A_3$ folgt $(f(a), f(b)) = (a', a') \in A_3$. Auch die Nicht-Umkehrbarkeit wird deutlich, denn aus $(a', d') = (f(a), f(d)) \in A_3$ folgt nicht $(a, d) \in A_3$.

Das dritte Beispiel in Abbildung 5.14 zeigt die Maschen-Adjazenz über Kanten. Die Maschen A und B werden bei der Generalisierung zur Masche A' zusammengefasst. Damit ist wiederum die Reflexivität der Adjazenzrelation notwendig, weil aus $(A, B) \in A_6$ folgt, dass $(f(A), f(B)) = (A', A') \in A_6$. Die Nicht-Umkehrbarkeit wird dadurch gezeigt, dass aus $(A', C') = (f(A), f(C)) \in A_6$ nicht folgt, dass $(A, C) \in A_6$ ist.

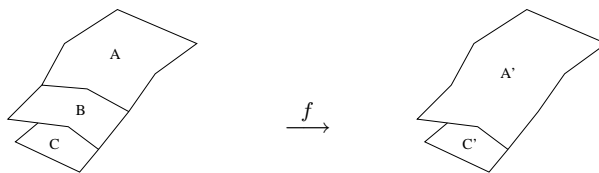


Abbildung 5.14: Beispiel für die Erhaltung der Maschen-Adjazenz über Kanten, sowie für die Reflexivität der Adjazenz und die Nicht-Umkehrbarkeit der Invarianzbedingung.

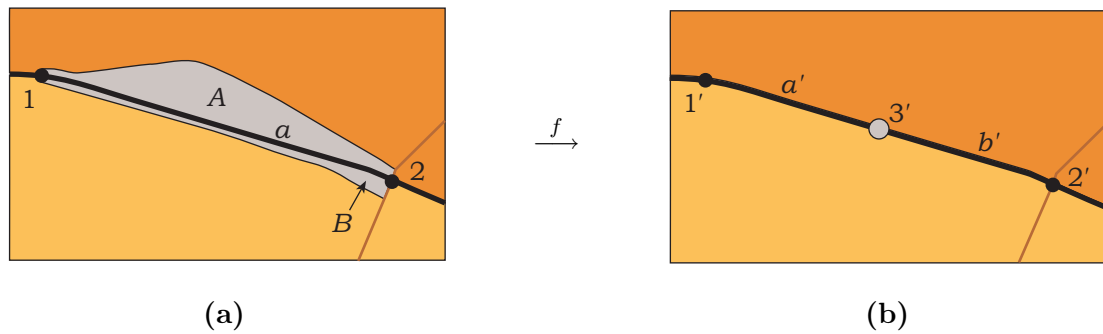


Abbildung 5.15: Verlust der Knotenadjazenz über Kanten durch Geometrietypwechsel. Der flächenförmige Bahnhof (grau) in (a) erfährt einen Geometrietypwechsel zu einem punktförmigen Bahnhof in (b). Ausschnitt aus dem Basis-DLM Nordrhein-Westfalen und des automatisch abgeleiteten DLM 250 (aus Bobzien und Morgenstern 2002).

Ausnahmen in Folge von Geometrietypwechseln

Die oben gestellten Bedingungen gelten nicht zwingend für den Fall, dass die beteiligten Geodaten *Geometrietypwechsel* erfahren. In diesen Fällen können Ausnahmen auftreten, die im Folgenden an Hand von drei Beispielen erläutert werden. Für die Behandlung von Geometrietypwechseln siehe zum Beispiel Bobzien und Morgenstern (2002), Schürer (2002), Schoppmeyer und Heisser (1995).

Das erste Beispiel (Abbildung 5.15) zeigt den Übergang eines flächenförmig modellierten Bahnhofs im ungeneralisierten Datenbestand zu einem punktförmig modellierten Bahnhof im generalisierten Datenbestand. Der Bahnhof (grau) wird in (a) durch die Maschen A und B modelliert. An den Punkten, an denen die Bahnstrecke den Bahnhof verlässt, befinden sich die Knoten 1 und 2. Diese Knoten sind adjazent zueinander über die Kante a , es ist also $(1,2) \in A_1$. Durch den Geometrietypwechsel im Zuge der Generalisierung kollabieren die beiden Maschen zum Knoten $3'$, der auf der Bahnstrecke liegt. Die Kante a wird geteilt in die Kanten a' und b' .⁹ Die Abbildungen der Knoten 1 und 2, das sind die Knoten $1'$ und $2'$, sind nicht mehr enthalten in der Adjazenzrelation A_1 und widersprechen damit der Bedingung 5.3.

Das zweite Beispiel (Abbildung 5.16) zeigt den Geometrietypwechsel von flächenförmig modellierten Ortslagen hin zu punktförmig modellierten Ortslagen. Die zwei Ortslagen (pink und orange) in (a) sind vor der Generalisierung zueinander adjazent. Die Generalisierungsvorschriften verlangen einen Geometrietypwechsel hin zu einer punktförmigen Modellierung (b). Die generalisierten Ortslagen (rote Kreise) sind nicht mehr zueinander adjazent, die Paare von Elementen sind nicht mehr in der Adjazenzrelation A_6 enthalten und widersprechen damit sowohl der Bedingung 5.3 als auch der Bedingung 5.4.

⁹Zum Algorithmus des Geometrietypwechsels siehe Bobzien und Morgenstern (2002).

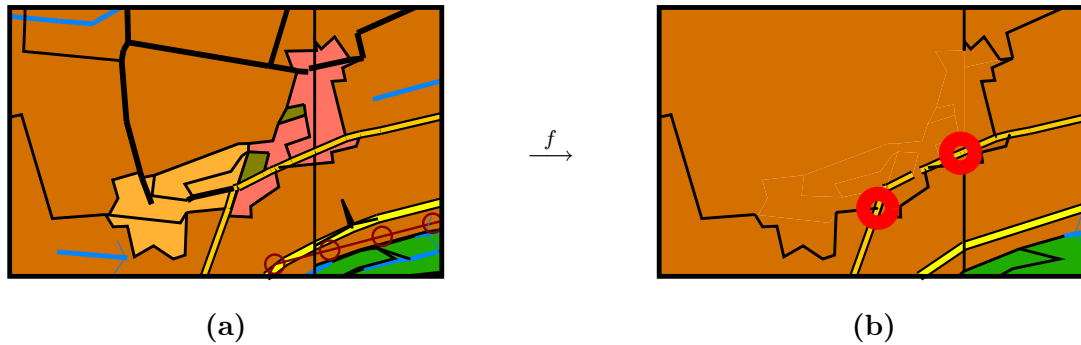


Abbildung 5.16: Verlust der Maschenadjazenz durch Geometriertypwechsel. Zwei adjazente flächenförmige Ortslagen (pink und orange) in (a) erfahren einen Geometriertypwechsel in (b). Ausschnitt aus dem DLM 250 Nordrhein-Westfalen und dem automatisch abgeleiteten DLM 1000.

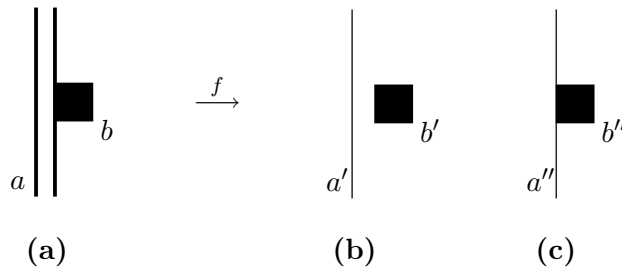


Abbildung 5.17: Verlust der Adjazenz beziehungsweise Umwandlung zur Inzidenz durch Geometriertypwechsel der flächenförmig modellierten Straße (a) zu einer linienförmig modellierten Straße (b) und (c).

Das letzte Beispiel (Abbildung 5.17) zeigt die Situation eines Gebäudes, das direkt an einer Straße gelegen ist. In (a) ist die Ausgangssituation gezeigt, wie sie zum Beispiel im Allgemeinen Liegenschaftskataster (ALK) erfasst sein kann. Die Straße a ist flächenförmig modelliert und das Gebäude b grenzt direkt an die Straße, ist also mit ihr topologisch verbunden. Bei der Generalisierung dieser Daten während der Übernahme in das Basis-DLM erfährt die Straße einen Geometriertypwechsel hin zu einer linienförmigen Modellierung. Im Allgemeinen wird für die Geometrie hierzu die Mittelachse der Fläche der Straße herangezogen. Für das Gebäude ergeben sich nun zwei Möglichkeiten der Erfassung: In (b) behält das Gebäude b' die ursprüngliche Lage bei, ist dann in der Folge aber nicht mehr topologisch mit der Straße verbunden. Die obige Bedingung der Topologieerhaltung ist hier verletzt. Für eine spätere kartographische Darstellung dieser Situation müssen zur korrekten Darstellung des Verhältnisses von der Straße zum Gebäude zusätzliche Informationen herangezogen werden, wie zum Beispiel der Wert des Attributs „Breite“ oder ein Mindestabstandskriterium. In (c) ist das Gebäude b'' an die Straße heran verschoben worden. Damit kann die geometrische Genauigkeit verletzt werden (siehe Abschnitt 5.4),

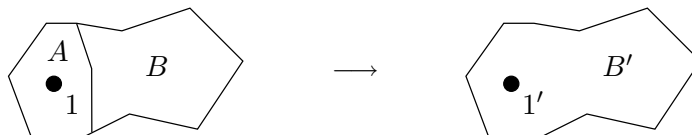


Abbildung 5.18: *Erhaltung der Inklusionsrelation bei Aggregation von flächenförmigen Objekten.*

aber die topologischen Relationen bleiben erhalten; und damit ist die Wiedergabe der topologische Beziehung zwischen den Objekten korrekt. Bei einer späteren kartographischen Darstellung dieser Situation ist aus der Topologie ersichtlich, dass das Gebäude an der Straße liegen muss. Geometrisch muss es dann abhängig von der Signaturenbreite der Straße verdrängt werden.

5.3.2 Inklusion

Mit dem Begriff *Inklusion* wird die topologische Relation des *Enthalten-Seins* bezeichnet (siehe Einleitung von Abschnitt 5.3). Sie kann betrachtet werden für Beziehungen zwischen Elementen aller drei geometrischen Ausprägungen, Knoten, Kanten und Maschen. Eine ausführliche Auseinandersetzung hierzu findet man bei Egenhofer und Herring (1990b). In diesem Abschnitt wird die Inklusion nur am Beispiel von Knoten und Masche betrachtet. Die Ergebnisse sind analog zu übertragen auf die Relationen zwischen Paaren topologischer Elemente anderer Ausprägung.

Wie bereits oben erwähnt, kann die Inklusionsbeziehung zweier Elemente zueinander im erweiterten hierarchischen Datenmodell nicht allein aus den Daten der topologischen Ebene abgeleitet werden, weil auf der topologischen Ebene nur Inzidenzen gespeichert sind. Benötigt werden zusätzlich Informationen aus der geometrischen Ebene. Insofern könnte man hier von geometrisch-topologischen Eigenschaften sprechen, die im Abschnitt 5.5.3 behandelt werden. Da die Inklusion im Allgemeinen aber als topologische Relation bezeichnet wird (Worboys 2004) und weil andere Datenmodelle als das hier genutzte die Inklusion explizit als topologische Relation modellieren können, wird sie auch hier unter den topologischen Eigenschaften behandelt.

In Abbildung 5.18 ist eine Situation dargestellt, in der ein punktförmig modelliertes Objekt – zum Beispiel eine Quelle – innerhalb eines flächenförmig modellierten Objektes liegt. Man sagt, das punktförmig modellierte Objekt ist *enthalten in* dem flächenförmig modellierten Objekt. Formell bildet diese Beziehung eine Relation zwischen den Knoten und Maschen auf der topologischen Ebene. Wir nennen diese Relation auf der topologischen Ebene IN_T und schreiben $IN_T \subseteq V \times F$. Ebenso wie die Inzidenzen und Adjazenzen im vorigen Unterabschnitt kann auch die Inklusion auf die Geodaten als Ganzes übertragen werden. Diese Relation umfasst dann die Geodaten, die topologisch vollständig in anderen Geodaten liegen. Wir bezeichnen diese Relation mit IN_G und

schreiben $IN_G^U \subseteq U \times U$ beziehungsweise $IN_G^Z \subseteq Z \times Z$ für die Inklusion innerhalb der Ausgangsdaten beziehungsweise innerhalb der Zieldaten.

Die Bedingung der Invarianz bezogen auf die Inklusionsrelation lautet umgangssprachlich formuliert, dass die Inklusion eines topologischen Elementes zu einem anderen topologischen Element durch die Generalisierung keine Veränderung erfährt. Diese Bedeutung lässt sich analog zur Invarianzbedingung der Adjazenz und Inzidenz verallgemeinern auf Geodaten: Ein Geodatum, das vor der Generalisierung innerhalb eines anderen Geodatoms liegt, muss auch nach der Generalisierung innerhalb des anderen generalisierten Geodatoms liegen. Voraussetzung für beide Bedingungen ist natürlich wieder, dass die topologischen Elemente beziehungsweise die Geodaten bei der Generalisierung nicht entfallen.

Formell lautet die Bedingung für die Invarianz der topologischen Elemente folgendermaßen:

$$\begin{array}{l} \text{Für alle } x \in V \text{ und } y \in F, \text{ wobei } f(x), f(y) \text{ definiert sind, gilt:} \\ \text{Aus } (x, y) \in IN_T \text{ folgt } (f(x), f(y)) \in IN_T. \end{array} \quad (5.5)$$

Für die Invarianz auf Ebene der Geodaten lautet sie:

$$\begin{array}{l} \text{Für alle } x, y \in U, \text{ wobei } f(x), f(y) \text{ definiert sind, gilt:} \\ \text{Aus } (x, y) \in IN_G^U \text{ folgt } (f(x), f(y)) \in IN_G^Z. \end{array} \quad (5.6)$$

Am Beispiel aus Abbildung 5.18 können diese Bedingungen verifiziert werden. Zu erwähnen ist, dass auch durch die Aggregation der beiden flächenförmigen Objekte A und B die Bedingung der Invarianz nicht verletzt wird. Vor der Generalisierung gilt $(1, A) \in IN_T$. Nach der Generalisierung gilt $(1', B') = (f(1), f(A)) \in IN_T$. Damit ist die Invarianzbedingung erfüllt.

5.4 Invariante geometrische Eigenschaften

In diesem Abschnitt werden die invarianten Eigenschaften der Generalisierung auf der geometrischen Ebene betrachtet. Im Gegensatz zu den semantischen und topologischen Eigenschaften sind die geometrischen Eigenschaften mit wenigen Ausnahmen „weiche“ Eigenschaften¹⁰. Während bei den vorgenannten Eigenschaften eine objektive Bewertung in „richtige“ und „falsche“ Generalisierungen möglich ist, können bei den geometrischen Eigenschaften graduelle Unterschiede in den Generalisierungen festgestellt werden. Man spricht also eher von „guten“ und „schlechten“ Generalisierungen, wie sie sich zum Beispiel

¹⁰Eine der wenigen Ausnahmen ist die Minimaldimension in der Kartendarstellung. Diese wird aber nicht wie die bereits behandelten „kategorischen“ Eigenschaften durch kartographische Regeln bestimmt, sondern durch die Wahrnehmungsfähigkeit des Menschen.

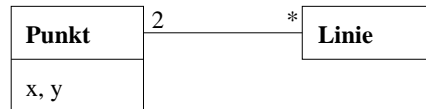


Abbildung 5.19: *Die geometrische Ebene des erweiterten hierarchischen Datenmodells (Ausschnitt aus Abbildung 3.5). Die Flächen werden nicht explizit in der geometrischen Ebene abgelegt, sondern ergeben sich aus den Beziehungen auf der topologischen Ebene (siehe Kapitel 3).*

in (SGK 1980) wiederfinden. Die Bewertungen sind in einem gewissen Grade subjektiv. Die Ursache für die weniger strikte Abgrenzung liegt in den Freiheitsgraden begründet, die sich durch den kontinuierlichen Raum der Ebene ergeben, in den die Geodaten eingebettet sind.

Ziel dieses Abschnittes ist es deshalb nicht, strikte Invarianzbedingungen wie in den vorhergegangenen Abschnitten zu entwickeln, sondern Kriterien aufzuzeigen, die dazu geeignet sind, unterschiedliche Generalisierungen derselben Situation auf der geometrischen Ebene zu beurteilen und miteinander zu vergleichen. In der Folge entsteht ein mehrdimensionaler (objektiver) Merkmalsraum, innerhalb dessen eine Generalisierung positioniert und bewertet werden muss. Dies ist ein aktuelles Forschungsthema und führte wegen der Komplexität der räumlichen Operationen und der Variationsmöglichkeiten in der geometrischen Generalisierung bislang noch zu keinem abschließendem Ergebnis (Ruas 2002).

Nach einem kurzen Überblick über die Grundlagen der geometrischen Darstellung von Geodaten und einer Vorstellung von einigen bestehenden Generalisierungsverfahren auf der geometrischen Ebene werden in diesem Abschnitt einige Kriterien zur Beurteilung von Generalisierungen erläutert. Für eine vertiefte Behandlung dieser komplexen Materie wird am Ende auf weiterführende Literatur verwiesen.

Grundlagen

Geo-Objekte werden auf der geometrischen Ebene im Allgemeinen durch Punkte, Linien und Flächen beschrieben (Worboys 2004). Flächen wiederum werden begrenzt von Linien, weshalb die geometrische Generalisierung von Flächen in großen Teilen durch die Generalisierung der Flächengrenzen abgedeckt werden kann.¹¹ Eine Entsprechung dieser Feststellung findet sich in der Modellierung der Geodaten im erweiterten hierarchischen Datenmodell auf der geometrischen Ebene. Abbildung 5.19 zeigt den Ausschnitt aus dem Datenmodell, der der geometrischen Ebene entspricht. Die geometrische Ebene wird ausschließlich durch Punkte und Linien beschrieben. Punkte werden durch ihre Koordinaten

¹¹Bei Schürer (2002) wird der Geometriotypwechsel auch als Teil der geometrischen Generalisierung gesehen. Mit der Einführung der topologischen Ebene fällt der Geometriotypwechsel aber ebenso in die Generalisierung der Topologie, weshalb er auch schon in Abschnitt 5.3 behandelt wurde.

in der euklidischen Ebene dargestellt, während Linien geraden Liniensegmenten zwischen genau zwei Punkten entsprechen.

In der Kartographie kommen im Allgemeinen nicht nur gerade Linien, sondern auch Kurven vor. Diese können jedoch in rechnergestützten Systemen nur sehr aufwändig modelliert werden. Deshalb behilft man sich häufig mit Interpolationen von Kurven durch Polynome innerhalb bestimmter Intervalle (Bartelme 2005), so genannten *Splines*. Gebräuchlich sind häufig Polynome dritten Grades, weil sie bereits eine sehr gute Annäherung an beliebige Kurven bieten, gegebenenfalls durch Aneinanderreihung mehrere Polynome, andererseits aber analytisch noch zu handhaben sind. Dennoch haben Polynome gravierende Nachteile bei der Analyse von geometrischen Informationen. Beispielsweise erfordert die Ermittlung von Schnitten zwischen zwei Splines oder das Berechnen des Flächeninhalts einer von Splines gebildeten Fläche aufwändige Kalkulationen.

Aus diesen Gründen verwenden die meisten heute verwendeten GIS keine Splines, sondern nähern Kurven durch Polylinien an. Polylinien bestehen aus einer Folge von geraden Liniensegmenten, bei denen der zweite Punkt des vorhergehenden Segments gleich dem ersten Punkt des folgenden Segments ist.¹² Um mit Polylinien Kurven anzunähern, benötigt man im Allgemeinen weitaus mehr Liniensegmente als bei einer Interpolation mit Polynomen höheren Grades. Der zusätzliche Speicherbedarf wird jedoch durch eine starke Vereinfachung von geometrischen Berechnungen kompensiert. Auch Forschungsarbeiten zur automatischen Generalisierung von Linien und entwickelte Algorithmen beziehen sich fast ausschließlich auf Polylinien. Eine der wenigen Ausnahmen ist der weiter unten erwähnte Algorithmus von Weber (1978).

Bestandteile der geometrischen Generalisierung von Linien

Die eigentliche geometrische Generalisierung von Linien beschränkt sich auf die zwei Vorgänge der *Linienvereinfachung* und der *Linienglättung*. Durch die geometrische Generalisierung können jedoch auch topologische und semantische Eigenschaften dieser Geodaten berührt werden. Auch Wechselwirkungen mit anderen Geodaten sind möglich. Ebenso können Darstellungsvorschriften – wie zum Beispiel die Minimaldimension oder die Signaturierung – für Karten von geometrischen Vereinfachungen berührt werden. Beispielsweise kann durch die Vereinfachung oder Glättung einer Linie ein topologischer Schnitt mit einer anderen Linie erfolgen. Diese Art der Bedingungsverletzung wird in Abschnitt 5.5 ausführlicher behandelt. Aber auch die Fläche eines Objektes kann durch Linienvereinfachung die Minimaldimension in der Darstellung einer Karte unterschreiten und damit eine semantische Nachbearbeitung des Objektes, durch Fortlassen oder Verschmelzen mit einem Nachbarobjekt, erfordern. Eine detaillierte Auflistung der durch geometrische Vereinfachung von Linien hervorgerufenen möglichen Wechselwirkungen findet man bei Weibel (1996). Dort werden mögliche Inkonsistenzen auf allen Ebenen der Geodaten betrachtet, die in Folge von Linienvereinfachung auftreten können. In Folge dieser

¹²Liniensegmente entsprechen Polynomen ersten Grades beschränkt auf ein geschlossenes Intervall.

Inkonsistenzen werden eventuell weitere Generalisierungsoperationen wie Verdrängung, Fortlassen oder Zusammenfassen von Objekten nötig. In diesem Abschnitt konzentrieren wir uns jedoch auf die rein geometrischen Eigenschaften und behandeln Inkonsistenzen in Verknüpfung mit anderen Ebenen des Datenmodells weiter unten.

Jede geometrische Generalisierung erzeugt eine geometrische Ungenauigkeit in der Geometrie der Geo-Objekte und damit in den Geodaten. Dieser Lagefehler sollte möglichst gering sein, damit die Aussagekraft der generalisierten Daten möglichst hoch ist. Andererseits ist durch das Wesen der Generalisierung ein Lagefehler nicht zu vermeiden; im Gegenteil ist ein gewisser Lagefehler in den generalisierten Daten einkalkuliert und wird in Kauf genommen. In topographischen Karten sind die kartographischen Darstellungen nur im Rahmen der Kartier- und Zeichengenauigkeit geometrisch richtig (Hake u. a. 2002). Durch Generalisierungen jedoch wird mit abnehmendem Maßstab der Lagefehler größer. In Digitalen Landschaftsmodellen ist eine Lagegenauigkeit in den Objektartenkatalogen festgesetzt. Diese liegt in ATKIS beim Basis-DLM für punkt- und linienförmige Objekte bei ± 3 m und beim DLM 250 bei ± 30 m (AdV 1989).

Linienvereinfachung und Linienglättung

Die *Linienvereinfachung* bezeichnet den Vorgang des Ausdünnens einer Polylinie durch Fortlassen von Stützpunkten. Die Anfangs- und der Endpunkte der Polylinie werden dabei nicht verändert. Die Linienvereinfachung wird auch *Punktreduktion* genannt. Bei der *Linienglättung* geht es darum, kleine Wellenformen, Ausbuchtungen und Zacken in den Linien durch Abschwächung dieser Formen zu glätten. Kleine Störungen im Linienverlauf sollen begradigt und nur die wichtigsten Trends der Linie eingefangen werden (McMaster und Shea 1992).

Bei der Linienvereinfachung ist gefordert, die geometrischen Charakteristika möglichst gut zu erhalten und dennoch eine adäquate Vereinfachung der Linie stattfinden zu lassen und dabei möglichst viele Punkte zu entfernen. Dies sind zwei sich widersprechende Ziele; es entsteht ein Zielkonflikt zwischen der Reduktion der Punktanzahl und der Erhaltung der Charakteristika der Linie. Weibel (1996) nennt diese beiden Ziele die „klassischen Kriterien der Punktreduktion“, zusammen mit dem dritten Kriterium der Minimierung der Berechnungskomplexität. Letzteres kann aber im Zuge von theoretischen Betrachtungen, als auch in dem Wissen, dass es sich erstens nicht um exponentielle Komplexitätssteigerung handelt, und zweitens, dass die Rechenleistung in der EDV ständig steigt und voraussichtlich weiter steigen wird, vernachlässigt werden.

Überblick über Linienvereinfachungs- und -glättungsalgorithmen

Einen umfassenden und heute immer noch aktuellen Überblick über entwickelte Algorithmen zur Linienvereinfachung und -glättung bieten McMaster (1987) und McMaster und Shea (1992). In diesem Abschnitt werden vier charakteristische Algorithmen kurz an

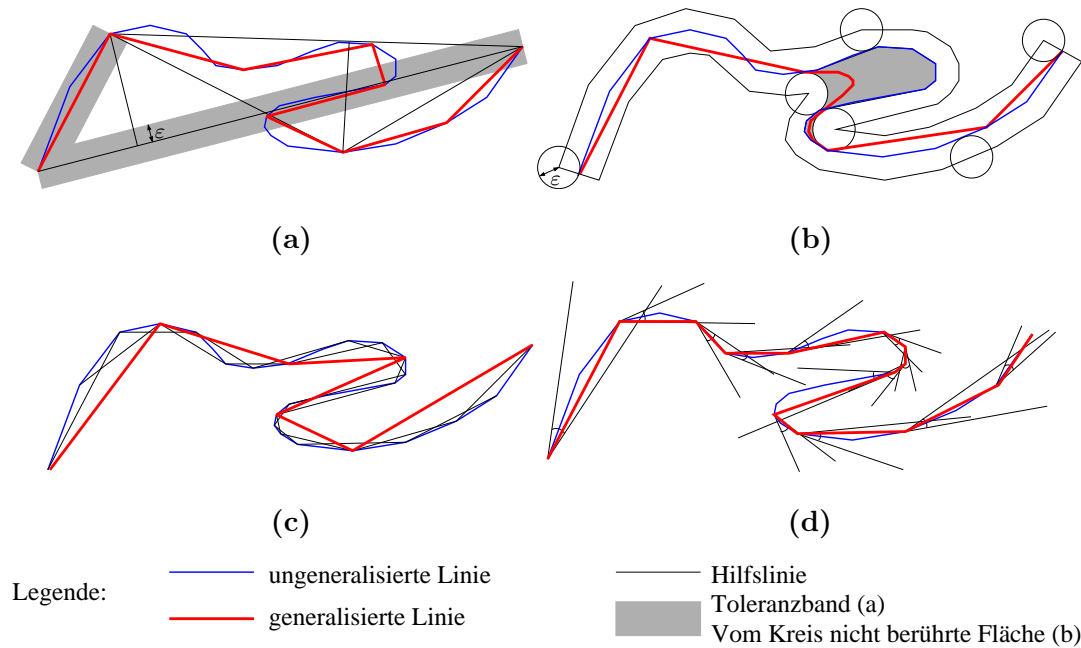


Abbildung 5.20: Vier Linienvereinfachungs- und -glättungsalgorithmen: **(a)** nach Douglas und Peucker, **(b)** nach Perkal, **(c)** nach Visvalingam und Whyatt, **(d)** nach Weber.

Hand von Beispielen skizziert, bevor im Weiteren auf die Kriterien zur Beurteilung von Generalisierungsergebnissen übergegangen wird.

Der wohl bekannteste Algorithmus zur Linienvereinfachung ist der Algorithmus von Douglas und Peucker (1973). Gleichzeitig vereinfachend und glättend arbeitet der Algorithmus von Perkal (1966a, b), der zum Beispiel von Christensen (1999) und Stenger (2003) implementiert wurde. Weitere, nicht in den oben erwähnten Publikationen behandelte Algorithmen sind die von Visvalingam und Whyatt (1993) und Weber (1978). Letzterer beschränkt sich nicht nur auf die Behandlung von Polylinien, sondern schließt Kurven mit ein. Die Vorgehensweisen der vier genannten Algorithmen sind in Abbildung 5.20 skizziert und in den nächsten Abschnitten kurz erläutert. Für eine vertiefte Behandlung mit diesen und weiteren Linienvereinfachungs- und Linienglättungsalgorithmen sei auf die oben erwähnte Literatur verwiesen.

Der Algorithmus von Douglas und Peucker (1973) verbindet die beiden Endpunkte einer Polylinie und sucht den Stützpunkt, der am weitesten von dieser Linie entfernt ist. Ist der Abstand zur Linie kleiner als ein gegebenes ε , so werden alle zwischen den Endpunkten liegenden Stützpunkte entfernt und der Algorithmus terminiert. Ist der Abstand größer, so bleibt der gefundene Punkt erhalten, die Polylinie wird an diesem Punkt geteilt, und der Algorithmus wird rekursiv für die beiden Teillinien fortgeführt.

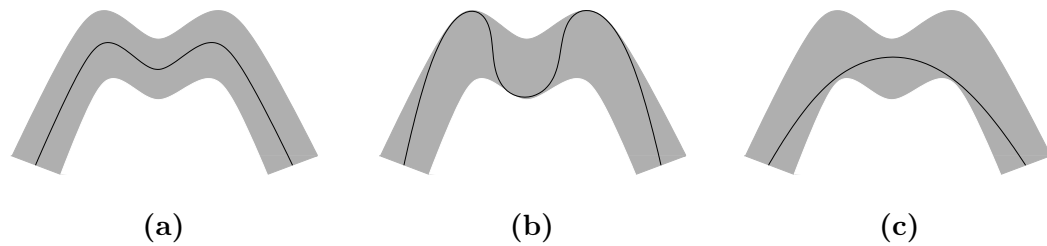


Abbildung 5.21: *Unterschiedliche Generalisierungen einer Linie innerhalb eines Toleranzbandes: (a) Ausgangsdaten, (b) betonte Generalisierung, (c) geglättete Generalisierung.*

Der Algorithmus von Perkal (1966b) lässt einen Kreis links und rechts um die Polylinie laufen. Innerhalb der Flächen, die nicht durch den Kreis abgedeckt werden, wird die Linie vereinfacht und geglättet, nach Christensen (1999) durch die Mittelachse der Fläche. Für die übrigen Linien schlägt Stenger (2003) vor, einen anderen Vereinfachungsalgorithmus anzuwenden.

Beim Algorithmus von Visvalingam und Whyatt (1993) wird jeder Punkt mit dem übernächsten Punkt der Polylinie verbunden. Die Fläche des so mit dem dazwischen liegenden Punkt entstandenen Dreiecks wird gemessen und der mittlere Punkt dieser drei Punkte des Dreiecks mit der geringsten Fläche wird entfernt. Die benachbarten Flächen werden neu berechnet und der Algorithmus wird iteriert, bis ein Abbruchkriterium erreicht wird, das aus einer Mindest- oder Höchstanzahl zu eliminierender Punkte oder aus einer Mindestflächengröße besteht.

Der Algorithmus von Weber (1978) arbeitet mit einer *gleitenden Blende*, die von einem Ende der Polylinie aus konstruiert wird. Es wird der letzte Punkt gesucht, der noch innerhalb dieser Blende liegt. Alle dazwischen liegenden Punkte werden eliminiert. Der Algorithmus wird iteriert, beginnend mit dem so gefundenen Punkt. Dieser Algorithmus arbeitet asymmetrisch; dies hat zur Folge, dass die Ergebnisse unterschiedlich ausfallen können, je nachdem von welcher Seite der Polylinie die Generalisierung begonnen wird.

Toleranzband und Lagegenauigkeit

Der Algorithmus von Douglas und Peucker steht stellvertretend für eine ganze Reihe von Algorithmen, die ein *Toleranzband* verwenden, das in einem Abstand ε um die Originallinie gelegt wird. An Stelle des Begriffs Toleranzband finden sich in der Literatur auch die Begriffe *tolerance corridor* (McMaster und Shea 1992), *Fehlersaum* (Heissler 1962) oder *Lagegenauigkeitsband* (Schürer 2002). Die Algorithmen garantieren durch ihre Vorgehensweise, dass das Ergebnis der Generalisierung innerhalb dieses Toleranzbandes liegt, dass also der Lagefehler sich in einem festen Rahmen bewegt. Dieses Kriterium ist häufig das einzig verwendete Kriterium zur Bestimmung der Güte einer geometrischen Generalisierung.

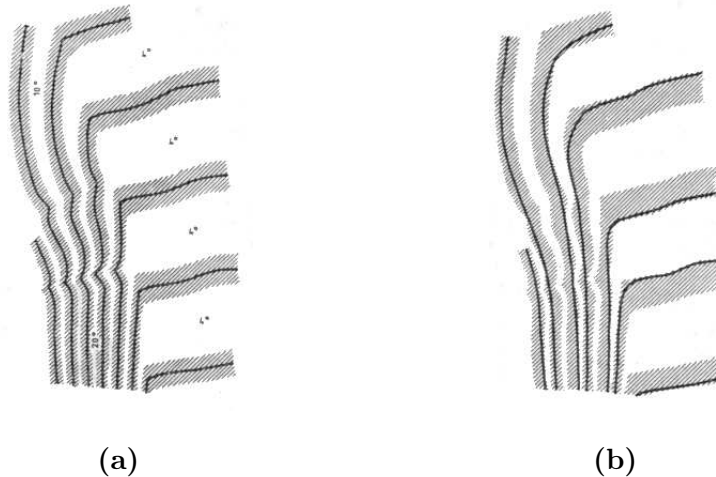


Abbildung 5.22: Höhenliniengeneralisierung und Toleranzband: (a) ungeneralisierte Höhenlinien und ihre Toleranzbänder, (b) eine Generalisierung innerhalb der Toleranzbänder, die morphologisch nicht korrekt ist (aus Heissler 1962, Heissler und Hake 1970, Hake 1982, Hake und Grünreich 1994, Hake u. a. 2002).

Der Begriff des Toleranzbandes korreliert mit dem Begriff der Lagegenauigkeit in der Kartographie. Die Lagegenauigkeit beschreibt den maximalen Lagefehler einer kartographischen Darstellung im Verhältnis zur wirklichen Lage eines Geo-Objektes. Das Toleranzband gibt analog dazu für einen Algorithmus einen maximalen Lagefehler vor, den eine Linie nach einer Generalisierung haben darf.

Aber das Toleranzband beziehungsweise die Lagegenauigkeit kann nicht als einziges Kriterium zur Beurteilung einer Generalisierung genügen. In Abbildung 5.21 ist dies an einem Beispiel verdeutlicht: Die ungeneralisierte Polylinie in (a) ist mit einem Toleranzband (in grau) versehen. Die Abbildungen (b) und (c) zeigen zwei (schlechte) Generalisierungen dieser Linie, die aber beide innerhalb des Toleranzbandes liegen, also die vorgegebene Lagegenauigkeit einhalten. Dennoch fallen diese Generalisierungen ausgesprochen unterschiedlich aus, die erste betont die Kurven in den Ausgangsdaten, die zweite glättet diese Kurven.

Bei Heissler (1962), Heissler und Hake (1970), Hake (1982), Hake und Grünreich (1994), Hake u. a. (2002) findet sich in einer Anwendung auf Höhenliniengeneralisierung ein weiteres Beispiel für die Notwendigkeit der Hinzuziehung anderer Kriterien außer dem des Toleranzbandes. Die morphologischen Anforderungen an die Generalisierung sind nicht automatisch dann erfüllt, wenn die generalisierten Höhenlinien innerhalb des Toleranzbandes der ungeneralisierten Linien liegen (siehe Abbildung 5.22).

Weitere Kriterien zur Beurteilung einer geometrischen Generalisierung

Aus der Literatur sind weitere Kriterien bekannt, die dem Erhalt des Charakters einer Linie dienen sollen. Zuerst zu nennen ist das bereits am Anfang dieses Abschnitts erwähnte Kriterium, die Anzahl der verbleibenden Punkte zu minimieren, um eine möglichst effektive Reduktion der Daten zu erhalten. Wie bereits beschrieben, kollidiert dieses Kriterium mit denjenigen Kriterien, die den Erhalt des Charakters einer Linie zum Ziel haben.

Von den Kriterien zum Erhalt des Charakters einer Linie sind die *metrischen Kriterien* diejenigen, die am ausgiebigsten untersucht wurden. Für metrische Kriterien wird der Fehler einer Generalisierung durch Längen- oder Flächenmessung bestimmt. McMaster (1986, 1987) nennt zwei metrische Kriterien, die im Folgenden kurz beschrieben werden:

Vektordifferenz: Die Entfernung der weggeneralisierten Punkte einer Linie zur generalisierten Linie werden aufsummiert. Diese Summe heißt *Vektordifferenz*. Je geringer sie ausfällt, desto näher ist die generalisierte Linie der ursprünglichen Linie.

Flächendifferenz: Die Flächen, die durch Überlagerung der generalisierten mit der ungeneralisierten Linie entstehen, werden berechnet. Die Größe dieser Fläche, die *Flächendifferenz* genannt wird, gibt ein weiteres Maß für die Nähe der generalisierten Linie zur Ursprungslinie. Je geringer die Flächendifferenz ausfällt, desto näher ist die Linie an der ursprünglichen.

Diese Kriterien lassen sich ins Verhältnis setzen zur Zahl der ursprünglichen und der verbliebenen Stützpunkte. Damit erhält man dann ein absolutes Maß zur Bestimmung der Güte einer Generalisierung. Weitergehende Untersuchungen zu metrischen Kriterien finden sich zum Beispiel bei Veregin (1999, 2000).

Metrische Kriterien sind jedoch nicht die einzigen Kriterien, die zur Beurteilung einer Generalisierung herangezogen werden können. In der neueren Forschung wird versucht, die *Form* und die *Gestalt*¹³ einer Linie, einer Fläche oder allgemeiner einer Gruppe von Objekten quantitativ und qualitativ zu erfassen. Bedeutende Ansätze zu dieser Forschung finden sich bei Ehrliholzer (1996), Weibel (1996), Jansen und van Kreveld (1998), Skopeliti und Tsoulos (1999), Weibel und Dutton (1999) und Bard (2004).

Das Herausarbeiten von einzelnen Merkmalen, die zur Form- und Gestaltbestimmung beitragen, sowie die Kombination mehrerer Merkmale sind Forschungsfelder, in denen zukünftig Fortschritte zu erwarten sind. Die Formalisierung dieser Merkmale, sowie der Abgleich zwischen den Merkmalen in Ausgangs- und Zieldaten sind aktuelle Fragen in der Forschung. An dieser Stelle sei auf die oben zitierte Literatur verwiesen, die sich mit der Ausformulierung solcher Merkmale beschäftigt.

Die in der Literatur untersuchten Kriterien beschränken sich nicht allein auf die Beurteilung der Generalisierung an Hand der Geometrie der Geo-Objekte, sondern beziehen auch

¹³Der Begriff *Gestalt* ist auch in der angelsächsischen Literatur in seiner deutschen Form gebräuchlich.

die Topologie und die Thematik mit ein. In dieser Arbeit wird versucht, diesen Schritt systematisch zu trennen von den Kriterien, die sich allein einer Ebene der Geodaten zuordnen lassen. Kriterien und Merkmale, die mehrere Ebenen der Geodaten berühren, werden deshalb nicht in diesem Abschnitt, sondern im folgenden Abschnitt über *invariante kombinierte Eigenschaften* behandelt. Die Trennung in die verschiedenen Ebenen sowie die Sonderbehandlung von kombinierten Eigenschaften ermöglicht eine Aufteilung auch der Programmteile für eine automatische Generalisierung und damit eine Aufteilung des komplexen Problems der Generalisierung in mehrere kleinere Teilprobleme, die auf diese Weise – hoffentlich – leichter gelöst werden können. Dies entspricht dem in der Informatik häufig angewendetem Prinzip des *divide and conquer*, dem Prinzip des *Teilens und Herrschens* (Knuth 1997).

5.5 Invariante kombinierte Eigenschaften

In diesem Abschnitt werden Beispiele für invariante Eigenschaften der Ur- und Zielmenge der Geodaten behandelt, die sich nicht eindeutig der thematischen, topologischen oder geometrischen Ebene des Datenmodells zuordnen lassen. Diese Eigenschaften entstehen aus der Kombination von zwei oder allen drei Ebenen des Datenmodells und werden deshalb invariante *kombinierte* Eigenschaften genannt.

Die ersten beiden Beispiele behandeln Kombinationen der thematischen mit der topologischen Ebene (Abschnitte 5.5.1 und 5.5.2), das dritte Beispiel eine Kombination der topologischen mit der geometrischen Ebene (Abschnitt 5.5.3). Denkbar sind auch Kombinationen aus allen drei Ebenen des erweiterten hierarchischen Datenmodells. Der theoretisch möglichen Kombination aus thematischer und geometrischer Ebene wird kein Beispiel zugeordnet, weil diese beiden Ebenen im hier verwendeten Datenmodell durch die topologische Ebene miteinander verbunden sind, und damit solch eine Kombination hier automatisch die topologische Ebene mit umfassen würde.

Alle vorgestellten Beispiele ergeben „kategorische“ Bedingungen, die an eine Generalisierung gestellt werden. Sie können durch Auswertung auf richtig oder falsch erkannt werden und so eine Generalisierung automatisch überprüfbar machen.

5.5.1 Erreichbarkeit und Netzzusammenhang

Die Erreichbarkeit und der Netzzusammenhang in Karten und Mengen von Geodaten beziehen sich in erster Linie auf Verkehrswege und -netze. Der Begriff Erreichbarkeit impliziert dabei die Möglichkeit, innerhalb einer Datenmenge eine ununterbrochene Verbindung von einem Punkt A zu einem Punkt B zu finden, der nur über Objekte mit bestimmter Thematik führt. Ein klassisches Beispiel ist die wechselseitige Erreichbarkeit zweier Punkten mittels eines PKW, der über einen Weg über (öffentliche) Straßen erfolgen soll. Der Netzzusammenhang verallgemeinert die Erreichbarkeit auf die Verbindung *beliebiger* Punkte (und nicht nur zweier Punkte), ebenfalls über Objekte mit bestimmter

Thematik. Alle so wechselseitig erreichbaren Punkte bilden ein *Netz*, das *zusammenhängend* ist.

In Digitalen Landschaftsmodellen existieren meist mehrere Netze nebeneinander. Jedes dieser Netze kann für sich zusammenhängend sein. Dies sind zum Beispiel das Gewässernetz, das Schienennetz, das Straßennetz und das Netz von Hochspannungsleitungen, zu dem auch Masten, Kraftwerke und Umspannstationen zählen. Innerhalb eines Datenausschnittes oder eines Kartenausschnittes kann es auch mehrere nicht miteinander verbundene Netze derselben Art geben; beispielsweise können, durch Wasserscheiden oder durch Ausschnittsbildung verursacht, in fest begrenzten Datenbank- oder Kartenausschnitten mehrere voneinander getrennte Gewässernetze bestehen.

Die Erreichbarkeit und der Netzzusammenhang sind Begriffe, die auch in der Graphentheorie bekannt sind. Für Graphen existieren bekannte Standardalgorithmen, die die Erreichbarkeit und den Netzzusammenhang überprüfen können (Harary 1972, Even 1979). Für die Erreichbarkeit zwischen zwei Punkten wird im Allgemeinen der Algorithmus von Dijkstra (1959) genutzt, für die Feststellung des Netzzusammenhangs der Algorithmus von Floyd (1962). Diese Algorithmen arbeiten auf einer Graphenstruktur mit Knoten und Kanten. Für die Erreichbarkeit und den Netzzusammenhang in den hier verwendeten Mengen von Geodaten müssen aber auch flächenförmige Objekte – topologisch modelliert durch Maschen – berücksichtigt werden. Beispielsweise kann ein mit dem PKW zurückgelegter Weg über einen (flächenförmig modellierten) Platz in einer Innenstadt führen, und der Zusammenhang eines Gewässernetzes soll natürlich auch nicht durch Seen oder flächenförmig modellierte Flüsse unterbrochen werden. Die erwähnten Algorithmen können auf den Einschluss von flächenförmig modellierten Objekten erweitert werden, indem beispielsweise die Flächen durch Skelettbildung an das ansonsten durch Kanten und Knoten gebildete Netz angeschlossen werden. Zur Skelettbildung siehe zum Beispiel (Eppstein und Erickson 1998).

Für die Generalisierung bilden die Erreichbarkeit und der Netzzusammenhang zwei Eigenschaften auf Datenmengen, die im Allgemeinen durch die Generalisierung nicht verändert werden sollen. Nicht-Veränderbarkeit heißt dabei, dass für zwei wechselseitig erreichbare Punkte im ungeneralisierten Datenbestand die korrespondierenden Punkte im generalisierten Datenbestand ebenfalls wechselseitig erreichbar sein sollen.

Abbildung 5.23 zeigt ein Beispiel, in dem die im ungeneralisierten Datenbestand bestehende Erreichbarkeit durch die Generalisierung nicht mehr zwischen allen Punkten eines Netzes gegeben ist. Die Ortschaft a ist vor der Generalisierung über die Straße b mit dem überörtlichen Straßennetz, symbolisiert durch die Bundesstraße c , verbunden. Nach der Generalisierung, bei der die geringerwertige Straße b fortgelassen wurde, ist die Erreichbarkeit zwischen a' und c' nicht mehr gegeben. In diesem konkreten Fall liegt ein Konflikt zwischen zwei unterschiedlichen Anforderungen an eine Generalisierung vor. Zum einen soll die Erreichbarkeit über das Straßennetz gewährleistet sein, zum anderen sollen weniger wichtige Straßen nicht in den Zieldatenbestand übernommen werden. Hier müssen die Kriterien an die Generalisierung aufeinander abgestimmt werden und eine Gewichtung der Bedingungen eingeführt werden. Eine Möglichkeit ist die Modifizierung

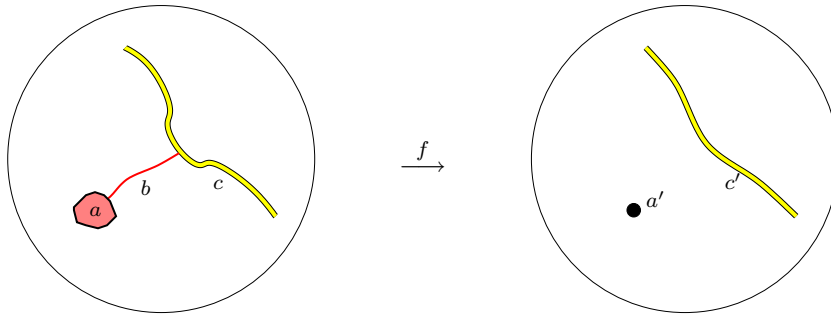


Abbildung 5.23: Netzzusammenhang und seine Verletzung durch eine Generalisierung.

der Regeln zur Generalisierung, zum Beispiel resultierend in der Bedingung, dass geringwertige Straßen bei der Generalisierung nur dann entfallen dürfen, wenn dadurch der Netzzusammenhang nicht verletzt wird.

Mathematisch ausdrücken lassen sich die Bedingungen der Erreichbarkeit und des Netzzusammenhangs durch eine Relation zwischen Geo-Objekten eines Netzes, die die Erreichbarkeit beschreibt. Für die Beschreibung eines Netzes werden die thematischen Eigenschaften der Geo-Objekte benötigt. Das Straßennetz wird beispielsweise aus allen Straßen, zusammen mit Plätzen, Brücken und Straßentunnel, gebildet. Auch Ortschaften könnten, wie im obigen Beispiel, zu diesem Netz gezählt werden. Für die Analyse der Erreichbarkeit zwischen den zu diesen Geo-Objekten gehörenden topologischen Elementen genügen dann die Beziehungen auf der topologischen Ebene.

Im Folgenden wird diese Bedingung formalisiert: Seien \mathcal{N}^U und \mathcal{N}^Z die Mengen der Netze innerhalb der Ausgangsdaten U und der Zieldaten Z , wobei vereinfachend angenommen wird, dass die Bedeutung der Netze in den Ausgangs- und Zieldaten die gleiche ist (siehe Abschnitt 5.1). Ein Netz $N \in \mathcal{N}^U$ ist eine Teilmenge von U , die bestimmten thematischen Kriterien genügt, wie zum Beispiel Zugehörigkeit zu bestimmten Objektarten oder eine Belegung mit bestimmten Attributwerten. Die Erreichbarkeit innerhalb eines Netzes $N \in \mathcal{N}$ für zwei Objekte $x, y \in N$ wird ausgedrückt durch die Erreichbarkeitsrelation E_N , die alle Paare von wechselseitig erreichbaren Objekten enthält. In der obigen Abbildung ist $E_{\text{Straßennetz}} = \{a, b, c\} \times \{a, b, c\}$; alle Elemente der Menge sind also wechselseitig erreichbar. Zu bemerken ist, dass die Erreichbarkeitsrelation *symmetrisch*¹⁴ und *transitiv* ist. Symmetrie bedeutet, dass aus der Erreichbarkeit von einem Geodatum x zu einem Geodatum y die umgekehrte Erreichbarkeit von y zu x folgt. Transitivität bedeutet, dass aus der wechselseitigen Erreichbarkeit zweier Geodaten x und y zusammen mit der Erreichbarkeit zweier Geodaten y und z die Erreichbarkeit von x und z folgt.

¹⁴Die Symmetrie gilt nur für *ungerichtete Graphen*, wovon in dieser Arbeit ausgegangen wird. Für andere Zusammenhänge kann die Datenstruktur auf *gerichtete Graphen* entsprechend erweitert werden. Damit lassen sich zum Beispiel Einbahnstraßen und Fließrichtungen modellieren. Die Erreichbarkeitsrelation ist damit nicht mehr symmetrisch.

Die Invarianz in der Erreichbarkeit kann mit Hilfe dieser Formalisierung und der Generalisierungsfunktion f folgendermaßen ausgedrückt werden:

Für alle $N \in \mathcal{N}$ gilt: für alle $x, y \in N$, für die $f(x), f(y)$ definiert sind, gilt: Aus $(x, y) \in E_N$ folgt $(f(x), f(y)) \in E_N$.	(5.7)
--	-------

Eine Verletzung dieser Bedingung bedeutet eine Verletzung des Netzzusammenhanges durch die Generalisierung. Weiterführende Bemerkungen zur Erreichbarkeit und zum Netzzusammenhang, die auch kompliziertere thematische Sachverhalte, wie die Reduktion des Straßennetzes an Grenzen bebauter Gebiete, sowie geometrische Sachverhalte, wie die Kontinuität einer Linie, mit einbezieht, finden sich zum Beispiel bei Edwardes und Mackaness (2006).

5.5.2 Flächendeckung

Das in diesem Abschnitt behandelte Beispiel für eine kombinierte Eigenschaft befasst sich mit der *Flächendeckung* in Digitalen Landschaftsmodellen. Negativ definiert beschreibt die Flächendeckung das Fehlen von flächenförmigen Erfassungslücken innerhalb eines erfassten Gebietes. Positiv ausgedrückt ist jedem Abschnitt der Fläche mindestens ein Geodatum zugeordnet. Die Forderung nach Flächendeckung kommt der Vorstellung nach, dass jeder Abschnitt der Erdoberfläche durch Geodaten modelliert werden kann.

Die Digitalen Landschaftsmodelle ATKIS der AdV (AdV 1989), TOPIS der Bundeswehr (MilGeo-OK 1993) und VECTOR25 von SwissTopo (SwissTopo 2004) haben die Flächendeckung als Bedingung in die Modellbeschreibung aufgenommen. In der Praxis ist diese Bedingung jedoch – zumindest in den Anfangsstadien der Erfassung – oftmals nicht erfüllt worden. Autobahnen zum Beispiel werden in ATKIS als komplexe Objekte modelliert, bestehend aus im Allgemeinen drei linienförmigen Objekten, einem Straßenkörper und zwei Fahrbahnen parallel dazu rechts und links. Die Fläche zwischen diesen drei linienförmigen Objekten wurde in den ersten Stadien von ATKIS nicht erfasst. Ähnliches galt für Wegebegleitgrün an Straßen, Böschungen oder Flüssen.

Die Flächendeckung findet ihr Pendant in der Graphentheorie bei dem Konzept der (graphentheoretischen) Landkarte (siehe Abschnitt 3.1.2). In diesen Landkarten ist die gesamte Fläche der Ebene mit Maschen abgedeckt, wobei die unbegrenzte Masche „Außen“ die Fläche in die Unendlichkeit abdeckt. Alle anderen Maschen sind abgeschlossen und entsprechen hier dem erfassten Teil der Landschaft.

Die Flächendeckung ist auf der topologischen Ebene also bereits gegeben. Für ein im Rahmen eines DLM erfasstes Gebiet bedeutet die Flächendeckung, dass jede Masche mit mindestens einem flächenförmigen Objekt verbunden ist. Dies ist im erweiterten hierarchischen Datenmodell durch die Kardinalität der Beziehung zwischen Masche und flächenförmigen Objekt durch die Kardinalität von > 1 ausgedrückt (siehe Abbildung 5.24).

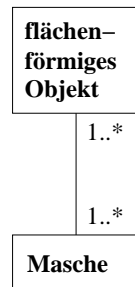


Abbildung 5.24: Die Klassen *Masche* und *flächenförmiges Objekt* des erweiterten hierarchischen Datenmodells (Ausschnitt aus Abbildung 3.5). Durch die Kardinalität der Beziehung wird die Flächendeckung garantiert.

Formell ausgedrückt ergibt sich folgende Bedingung sowohl für die Urmenge U als auch für die Zielmenge Z , wobei F die Menge der Maschen ist (siehe Abschnitt 5.3.1):

Für alle Maschen $m \in F^{\{U,Z\}}$ gilt: Es existiert ein $o \in \{U, Z\}_{\text{Them}}$, so dass m assoziiert ist mit o .	(5.8)
--	-------

Ist diese Bedingung in der Urmenge nicht erfüllt, so ist bereits der Ausgangsdatenbestand bezüglich der Flächendeckung inkonsistent. Ist die Bedingung nur für die Zielmenge nicht erfüllt, so ist die Generalisierung bezüglich der Flächendeckung fehlerhaft verlaufen.

Wenn flächenförmige Objekte nach ihrer Ausgestaltung unterschieden werden in Grund-, Ausgestaltungs- und Hintergrundobjekte (Bobzien 2001), beziehungsweise in Grund-, Ausgestaltungs-, übergeordnete und neutrale Objekte (Schürer 2002), dann können sich bezüglich obiger Bedingung Einschränkungen ergeben, oder die Bedingung kann spezifiziert werden. Zum Beispiel kann gefordert werden, dass die Flächendeckung schon durch Grundobjekte ausgefüllt werden muss, oder dass Verwaltungsflächen die Fläche komplett abdecken müssen. Die Formulierung dieser Bedingungen ist in hohem Grade von der Modellierung und der Interpretation des jeweiligen Modells abhängig und wird hier deshalb nicht weiter spezifiziert.

5.5.3 Topologische Konsistenz und geometrische Generalisierung

Das dritte und letzte Beispiel für eine invariante kombinierte Eigenschaft beschreibt eine Kombination von topologischen und geometrischen Eigenschaften. Wie bereits in Abschnitt 5.4 erwähnt wurde, kann die Linienvereinfachung und Linienglättung zu topologischen Veränderungen der generalisierten Daten im Vergleich zu den ungeneralisierten Daten führen. Einige dieser Veränderungen haben Inkonsistenzen zur Folge, die die Interpretation des generalisierten Datenbestandes verfälschen können.

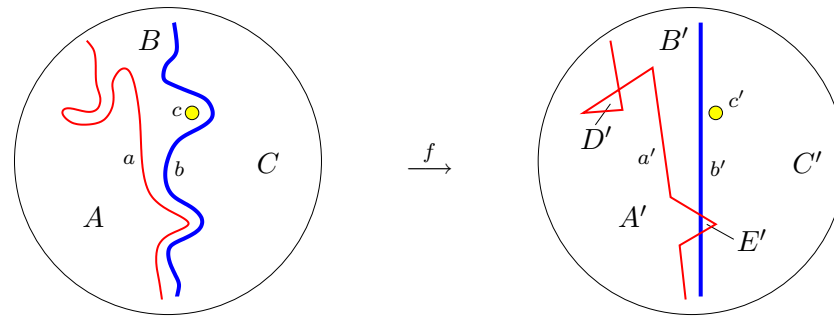


Abbildung 5.25: Mögliche topologische Inkonsistenzen in Folge von geometrischen Vereinfachungen.

Saalfeld (1999) hat mögliche Inkonsistenzen untersucht und die folgenden drei Fälle beschrieben:

1. **Selbstüberschneidung einer Linie:** Durch ungünstige Parametrisierung oder Anwendung eines ungeeigneten Linienvereinfachungsalgorithmus kann eine vereinfachte Linie nach der Generalisierung sich selbst überschneiden. Anfällig für solche Selbstüberschneidung ist zum Beispiel der Algorithmus von Douglas und Peucker (Saalfeld 1999). Die Straße a' in Abbildung 5.25 überschneidet sich im oberen Teil der Abbildung nach der Generalisierung selbst und erzeugt damit einen topologischen Fehler. In der Karteninterpretation würde dieses entweder als (nicht dargestellte) Über- oder Unterführung interpretiert oder böte die Möglichkeit einer in Wirklichkeit nicht vorhandenen Abkürzung. Einen Ansatz zur Formalisierung dieser Inkonsistenz bietet das verwendete Datenmodell, in dem ein solcher Schnitt ein Einfügen eines Knoten erfordern würde, der wiederum die Erzeugung von neuen Kanten und einer von dem Ring umschlossenen Masche erfordern würde.
2. **Schnitt mit einer anderen Linie:** Im ungeneralisierten Datenbestand nahe beieinander liegende Linien können durch Generalisierung unterschiedlich vereinfacht oder geglättet werden und sich dadurch nach der Generalisierung gegenseitig schneiden. Die Straße a und der Fluss b in Abbildung 5.25 kreuzen sich vor der Generalisierung nicht. Nach der geometrischen Vereinfachung gibt es zwei Überschneidungen zwischen der Straße a' und dem Fluss b' , die einen topologischen Fehler darstellen. In der kartographischen Interpretation ergibt diese Generalisierung zwei Querungen der Straße über den Fluss, die in dem ungeneralisierten Datenbestand (und damit vermutlich auch in der Wirklichkeit) nicht existieren. Die Formalisierung dieses Sachverhaltes geschieht analog zu der Formalisierung der Invarianz der topologischen Eigenschaften in Abschnitt 5.3.1. An Stelle der Inzidenz und der Adjazenz tritt hier die Disjunktheit.
3. **Relativer Lagewechsel / Wechsel der umschließenden Masche:** Die Vereinfachung oder Glättung einer Linie kann auch topologische Inkonsistenzen bei

Objekten hervorrufen, die mit der vereinfachten oder geglätteten Linie in keinem direkten Zusammenhang stehen, sie also nicht berühren oder in ihr enthalten sind. Durch die Vereinfachung oder Glättung kann ein Objekt, das innerhalb einer Masche liegt, in das Innere einer anderen Masche gelangen. In Abbildung 5.25 liegt das punktförmige Objekt c vor der Generalisierung zwischen Straße und Fluss, und damit in der Masche B , die unter anderem durch die Kanten der Straße und des Flusses gebildet wird. Nach der Generalisierung liegt es in der Masche C' , die unter anderem von b' gebildet wird. Damit ist die topologische Relation zwischen Knoten c zu den Maschen A und B verändert worden. In der Interpretation der Abbildung würde man sagen, dass das Objekt c durch die Generalisierung im Verhältnis zum Fluss b' „die Seite gewechselt“ hat. Die Formalisierung entspricht der Bedingung der Erhaltung der Inklusionsrelation, die in Abschnitt 5.3.2 dargestellt wurde.

Es existieren bereits mehrere Ansätze im Zusammenhang mit der geometrischen Generalisierung, die bestehende Algorithmen zur Linienvereinfachung so modifizieren, dass topologische Inkonsistenzen als Folge der Vereinfachung vermieden, verhindert oder nachträglich rückgängig gemacht werden. Im Folgenden werden deshalb nur einige bestehende Methoden und Algorithmen auf grundsätzlicher Ebene angesprochen.

Saalfeld (1999) unterscheidet drei Strategien zur Behandlung von topologischen Inkonsistenzen im Zuge der geometrischen Generalisierung:

1. **Generalisierung *en masse***: Die Menge der zu generalisierenden Objekte wird als Ganzes betrachtet; gleichzeitig mit der geometrischen Generalisierung werden die topologischen Relationen zwischen allen Elementen berücksichtigt und fließen in die Generalisierung mit ein.
2. **Generalisierung *en suite***: Geo-Objekte werden einzeln betrachtet und nacheinander geometrisch generalisiert; die topologischen Beziehungen zu anderen Objekten werden jedoch dabei berücksichtigt, und topologische Inkonsistenzen während der Generalisierung vermieden.
3. **Generalisierung *in vacuo***: Die Geo-Objekte werden isoliert betrachtet und generalisiert. Eventuell auftretende topologische Fehler werden durch nachträgliche Analyse erkannt und behoben, zum Beispiel durch partielle Umkehrung der Generalisierung.

Abbildung 5.26 verdeutlicht die Auswirkungen dieser drei Strategien am Beispiel zweier im Ausgangsdatenbestand nahezu parallel verlaufenden linienförmigen Objekte (a). Die Generalisierung *en masse* wird in (b) verdeutlicht: Da beide Objekte „gleichzeitig“ generalisiert werden, kann die parallele Linienführung in den generalisierten Datenbestand übernommen werden. Abbildung 5.26 (c) zeigt ein mögliches Ergebnis der Generalisierung *en suite*. Zunächst wird der äußere Bogen vereinfacht. Dabei wird die Lage des inneren Bogens berücksichtigt, weshalb nur wenige Punkte des Bogens entfallen können, da ansonsten ein Schnitt mit dem inneren Bogen auftreten würde. Im zweiten Schritt wird der innere Bogen vereinfacht. Dabei geht die parallele Linienführung zum äußeren

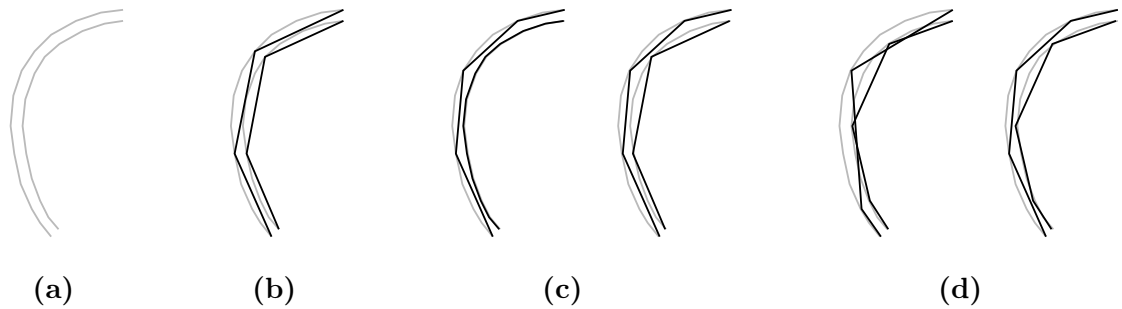


Abbildung 5.26: Grundsätzliche Strategien zur Vermeidung von topologischer Inkonsistenz in Folge von geometrischer Generalisierung: (a) Ausgangszustand, (b) Generalisierung en masse, (c) Generalisierung en suite, (d) Generalisierung in vacuo.

Bogen verloren. Abbildung 5.26 (d) zeigt die Generalisierung *in vacuo*. Dabei werden zunächst beide Linien unabhängig voneinander vereinfacht. Dabei entstehen topologische Inkonsistenzen, die durch Einfügen bereits entfernter Punkte wieder beseitigt werden. Im Ergebnis ist wiederum die Parallelität der Ausgangslinien nicht berücksichtigt.

Ansätze zur Generalisierung von Linien unter Berücksichtigung des Kontextes finden sich bei Jones u. a. (1995), de Berg u. a. (1998) und Saalfeld (1999).

Für die Formulierung der Konsistenzhaltung ist es schließlich unerheblich, welche der oben genannten Strategien zur Liniengeneralisierung angewendet werden. Formuliert werden muss für die Bedingung, dass nach der Generalisierung und der eventuell nachfolgenden Korrektur keine topologischen Inkonsistenzen bestehen. Zur Formulierung kann man zwei Ansätze wählen, die hier jedoch nur skizziert werden:

- Ausgehend von der *Topologie* formuliert man topologische Relationen im Ausgangsdatenbestand und ordnet die geometrisch generalisierten Objekte den topologischen Elementen in den Ausgangsdaten zu. Das hier verwendete erweiterte hierarchische Datenmodell und die Multirepräsentation (siehe Kapitel 3) unterstützen diese Vorgehensweise. Finden sich in den generalisierten Daten topologische Elemente oder Beziehungen, die sich nicht auf die Ursprungsdaten zuordnen lassen, ist eine Inkonsistenz aufgedeckt worden.
- Ausgehend von der *Geometrie* testet man zum einen auf Schnitte nach der Generalisierung. Schnitte deuten auf neue topologische Elemente hin, also auf Inkonsistenzen. Zum anderen testet man auf Enthalten-In-Relationen zu Maschen. So können relative Lagewechsel und Wechsel der umschließenden Masche erkannt werden.

5.6 Fazit

Dieser Abschnitt schließt das Kapitel über invariante Eigenschaften der Generalisierungsfunktion. Der Beginn dieses Kapitels enthält mathematische Grundlagen für die Formalisierung der invarianten Eigenschaften, speziell die Begriffe *Struktur* und *Morphismus*. Die folgenden Abschnitte stellen invariante Eigenschaften bezüglich der Generalisierung vor und entwickeln diese. Die invarianten Eigenschaften werden untergliedert nach den drei Ebenen des erweiterten hierarchischen Datenmodells: der thematischen, der topologischen und der geometrischen Ebene. Außerdem werden Eigenschaften vorgestellt, die auf Kombinationen der Ebenen basieren, und zwar sowohl auf der Kombination der thematischen mit der topologischen Ebene, als auch auf der Kombination der topologischen mit der geometrischen Ebene.

Die meisten dieser invarianten Eigenschaften können kategorisch durch logische Ausdrücke in Form von Bedingungen formuliert werden. Sie bilden damit Bedingungen an die Generalisierung, die leicht automatisch überprüft werden können. Dies betrifft die Bedingungen auf der thematischen und topologischen Ebene sowie die Kombination dieser beiden Ebenen.

Andere Bedingungen werden weich formuliert und nicht konkret formalisiert. Dies betrifft die Bedingungen auf der geometrischen Ebene sowie die Kombination zwischen topologischer und geometrischer Ebene. Die detaillierte Beschreibung dieser Bedingungen liegt nicht im Fokus dieser Arbeit. An dieser Stelle werden offene Fragen aufgeworfen. Dieses Kapitel versucht auch, eine Richtung für künftige Forschungsarbeiten in diesem Bereich zu skizzieren.

Kapitel 6

Anwendungen

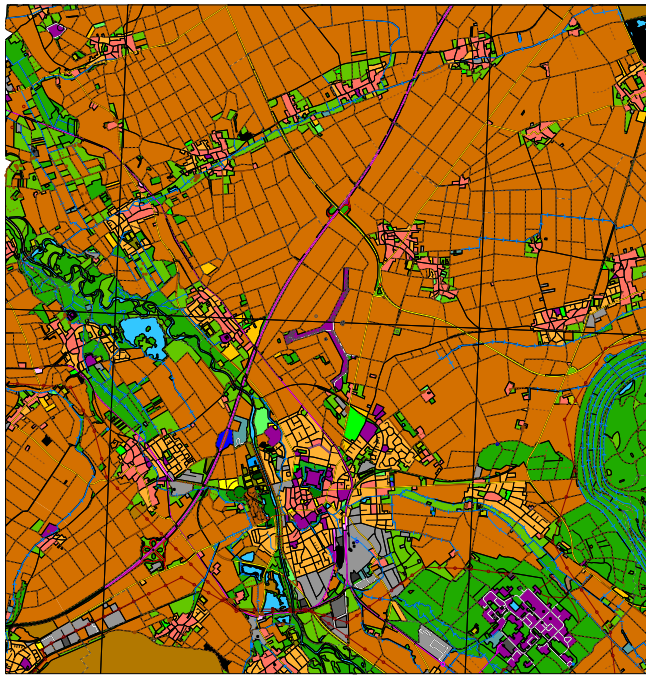
Die in den vorangegangenen Kapiteln entwickelten Konzepte werden in diesem Kapitel auf reale Daten aus einem Digitalen Landschaftsmodell angewendet, und die Ergebnisse dieser Anwendung werden untersucht. Wie bereits in der Einleitung erwähnt, ist diese Arbeit aus der Automatisierung der Modellgeneralisierung entstanden. Folgerichtig beschäftigt sich die erste Anwendung mit der Modellgeneralisierung (Abschnitt 6.1). An Hand von zwei Beispielen zur Modellgeneralisierung, die Übergänge zwischen unterschiedlichen Strukturierungsgraden behandeln, werden die im Kapitel 5 aufgestellten invarianten Kriterien dargestellt und erläutert.

Die weiteren Abschnitte beschäftigen sich mit Beispielen für Generalisierungsanwendungen aus den Bereichen geologischer Karten und Straßenkarten. Die aus der Modellgeneralisierung heraus entwickelten Konzepte und Formalisierungen werden an Hand passender Beispiele und Kriterien auf diese übertragen. Der Fokus in diesem Kapitel liegt auf der Darstellung der Möglichkeiten, die sich aus der Formalisierung der Modellgeneralisierung für diese kartographischen Anwendungen ergeben. Abschnitt 6.2 befasst sich mit der Generalisierung von *geologischen Karten* und Abschnitt 6.3 mit der Generalisierung von *Straßenkarten*.

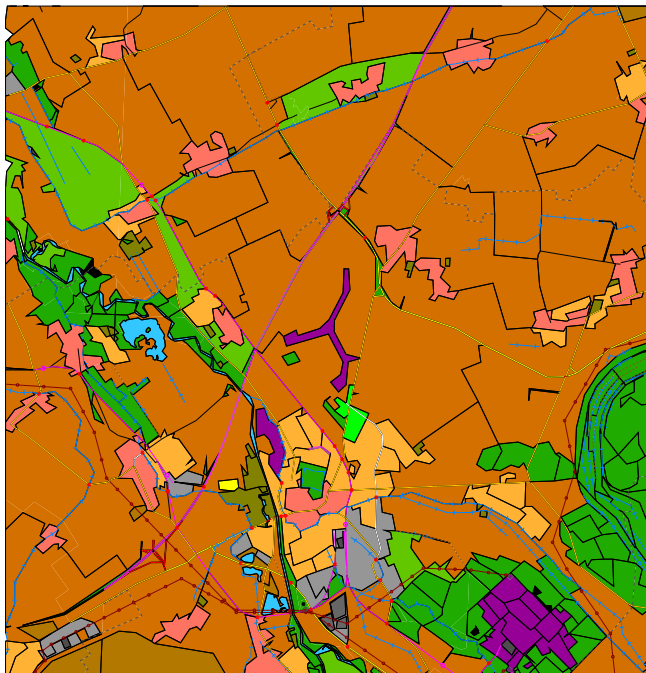
Nicht alle aus der Modellgeneralisierung heraus entwickelten Konzepte lassen sich auf jede kartographische Anwendung übertragen. Für eine bestimmte Generalisierungsanwendung muss sorgfältig ausgewählt werden, welche invarianten Eigenschaften für diese Anwendung gelten müssen und welche nicht. An Beispielen wird gezeigt, welche Kriterien aus Kapitel 5 sich auf welche Anwendung übertragen lassen.

6.1 Modellgeneralisierung

Im Rahmen dieser Arbeit wurden exemplarisch Teile eines Digitalen Landschaftsmodells generalisiert. Es handelt sich um Daten aus Deutschland, genauer aus Nordrhein-Westfalen. Der Datensatz entstammt dem TOPIS DLM 25. Diese Daten wiederum stammen ursprünglich aus dem ATKIS Basis-DLM. Sie sind also dem ATKIS Basis-DLM



DLM 25



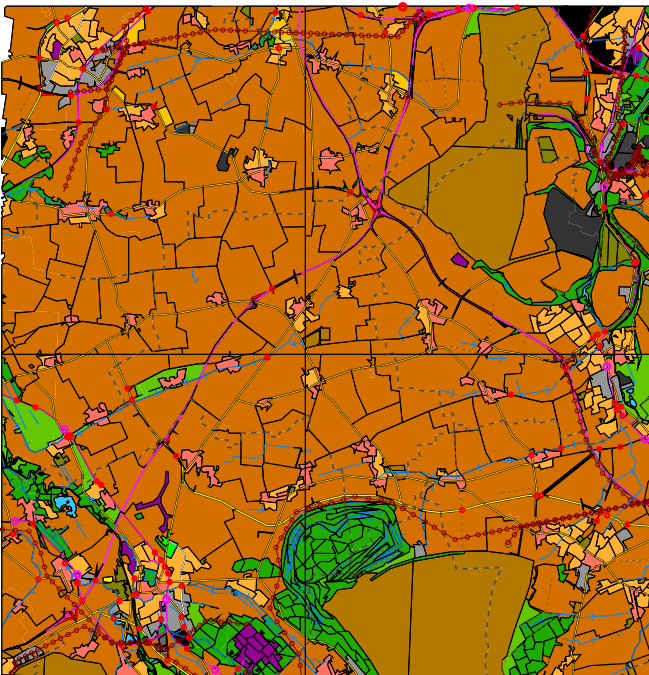
DLM 250

Legende zu den Abbildungen 6.1 und 6.2:

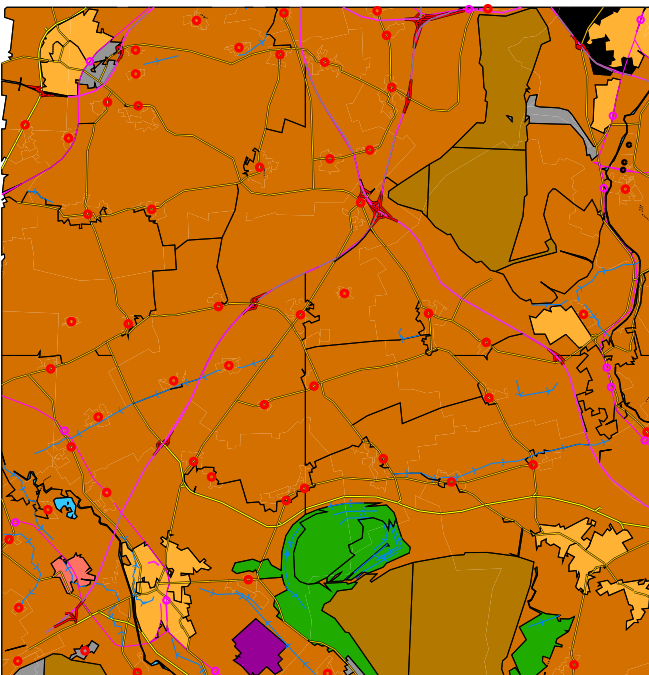
- Bergbauanlage
- Abfalldeponie
- Absetzbecken
- Wasserwerk
- Kläranlage
- Kraftwerk
- Umspannwerk
- Sportplatz
- Park
- Stadion
- Schwimmbecken
- Zoo
- Grünanlage
- Sportanlage / Freizeitanlage
- Freibad
- Fläche bes. funkt. Prägung
- Fläche gemischter Nutzung
- Industrie- und Gewerbefläche
- Wohngebiet
- Friedhof
- Platz
- Brücke
- Station
- Rastanlage
- Binnengewässer
- Erdoberflächenbeschaffenheit
- Ackerland
- Baumschule / Gärtnerei
- Gartenland
- Grünland
- Wald / Forst
- Flugplatz
- Nicht erfasstes Gebiet

(Fortsetzung in Abbildung 6.2)

Abbildung 6.1: Übergang vom DLM 25 zum DLM 250. Die Darstellung ist im Maßstab 1 : 125 000.



DLM 250



DLM 1000

Legende zu den Abbildungen 6.1 und 6.2:
(Fortsetzung aus Abbildung 6.1)

- Schienenweg
- Bahnstrecke
- Autobahn
- Straßenkörper
- Fahrbahn
- Bundesstraße
- Landesstraße
- Kreisstraße
- Gemeindestraße
- Sonstige Straße
- Feld/Waldweg
- Pfad/Fussweg
- Sonstiger Weg
- Brücke
- Freileitung
- Fluss/Kanal
- Wallhecke
- Baumreihe
- Verwaltungsgrenze
- Nutzungsartengrenze
- Sonstige Grenze
- Brunnen
- Kraftwerk
- Umspannwerk
- Schornstein
- Ortslage
- Turm
- Wasserturm
- Verkehrsknoten
- Brücke
- Station
- Stromleitungsmast
- Fernmeldeturm
- Baum
- Hubschrauberlandeplatz

Abbildung 6.2: Übergang vom DLM 250 zum DLM 1000. Die Darstellung ist im Maßstab 1:250 000. Der Ausschnitt aus Abbildung 6.1 entspricht in diesen Graphiken jeweils dem unteren linken Viertel.

vergleichbar.¹ Die Daten des TOPIS DLM 25 bilden die Ausgangsdaten für die im Rahmen dieser Arbeit durchgeführte Modellgeneralisierung.

Die blattschnittfreien Daten des DLM 25 wurden vor der Generalisierung in quadratische „Kacheln“ der Größe $12 \times 12 \text{ km}^2$ aufgeteilt, so dass benachbarte Kacheln sich in einem Bereich von 1 km Breite gegenseitig überdecken. Diese Aufteilung wurde gewählt, um die Generalisierung auf einen mit den technischen Mitteln gut zu beherrschenden Bereich zu beschränken. Im Vordergrund stand das Ziel der Entwicklung eines Prototyps, der die prinzipielle Machbarkeit der Implementation einer automatisierten Modellgeneralisierung zeigt. Die Beschränkung des Bearbeitungsraumes ermöglicht eine schnellere Implementation und Durchführung von Tests, ohne sich zu sehr auf die technisch orientierte Verwaltung von Massendaten und die Optimierung von Algorithmen zur Speicherung- und Laufzeitminimierung konzentrieren zu müssen.

Die Kacheln des DLM 25 wurden getrennt generalisiert in das DLM 250. Abbildung 6.1 zeigt eine Kachel im DLM 25 ungeneralisiert (oben) und vollautomatisch generalisiert im DLM 250 (unten). Die generalisierten Kacheln wurden an den gemeinsamen Kanten um die Hälfte des Überlappungsbereiches (also $\frac{1}{2}$ km) beschnitten. Damit wurde sichergestellt, dass die von der Randlage einer Kachel beeinflussten Ergebnisse nicht in den generalisierten Datensatz übernommen werden. Die Kacheln wurden danach aneinandergefügt und ergaben das Ergebnis im DLM 250.²

In einer zweiten Modellgeneralisierung wurden die so gewonnenen Daten für die Ableitung eines DLM 1000 generalisiert. Abbildung 6.2 zeigt oben vier aneinandergefügte Kacheln des DLM 250. Deutlich zu erkennen sind die Schnittlinien der einzelnen Kacheln. Im unteren Teil der Abbildung sieht man das generalisierte Ergebnis im DLM 1000. An einigen Stellen sind die Schnittlinien noch zu erkennen.

Anzumerken ist, dass die Darstellungen nur eine Visualisierung der DLM-Datenbestände sind und dass weder eine kartographische Signaturierung noch kartographische Generalisierung stattgefunden hat. Bei der Generalisierung stand die *Korrektheit* des Ergebnisses – so wie in Kapitel 5 erläutert – im Vordergrund. Aus dieser Priorisierung ergeben sich Situationen, die für den kartographisch geprägten Betrachter unschön erscheinen. Die Ergebnisse sind jedoch korrekt im Sinne der Modellgeneralisierung und entsprechen den Vorgaben der Objektartenkataloge.

6.1.1 Invariante thematische Eigenschaften

Der Erhalt von thematischen Eigenschaften während der Generalisierung wird im wesentlichen durch Klassenzugehörigkeit der Geo-Objekte bestimmt (siehe Abschnitt 5.2). Die

¹Zum Unterschied zwischen den Digitalen Landschaftsmodellen ATKIS und TOPIS siehe Abschnitt 3.1.6 auf Seite 53.

²Es war jedoch eine Randanpassung nötig, die manuell durchgeführt wurde. Pro Grenze einer Kachel mussten ca. fünf bis zehn Objekte manuell angepasst werden. Mit einem anderen (größeren) Überlappungsbereich könnte dieser Grad der Randanpassung sicherlich noch verringert werden.

Objektart	DLM 250	DLM 1000
	(Flächenanteil in %)	
Ackerland	67,28	70,44
Industrie- und Gewerbefläche	12,47	12,01
Wald / Forst	6,40	6,41
Wohngebiet	5,05	6,04
Fläche gemischter Nutzung	3,03	2,97
Grünland	1,49	0,07
Nicht erfasstes Gebiet	1,48	1,48
Buschwerk	1,05	0,11
Sportanlage	0,63	0,00
Fläche bes. funkt. Prägung	0,57	0,43
Baumschule / Gärtnerei	0,37	0,00
Friedhof	0,05	0,00
Freizeitanlage	0,10	0,03
Platz	0,02	0,00
Grünanlage	0,01	0,01

Tabelle 6.1: *Prozentualer Anteil der Grundflächen aus den Daten aus Abbildung 6.2. (Bergbauanlagen fallen unter Industrie- und Gewerbeflächen.)*

Klassenzugehörigkeit wiederum wird in der amtlichen Kartographie durch Zugehörigkeit eines Geo-Objektes zu einer Objektart beschrieben. Eine feinere Klasseneinteilung erhält man durch die Attributierung der Geo-Objekte (siehe Abschnitt 5.2.2). Der Erhalt der thematischen Eigenschaften wird mathematisch beschrieben durch die Kommutativität des Diagramms mit der Generalisierungsfunktion f , der Objektartenzuordnungsfunktionen o_U und o_Z , sowie der Funktion zwischen den Mengen der Objektarten g_O , so wie in der Bedingung 5.1 auf Seite 111 beschrieben. Die Kodierung der letzteren Funktion kann durch eine Tabelle geschehen. Bei Schürer (2002, Seite 93) findet sich ein Ausschnitt der für den in dieser Arbeit implementierten Algorithmus verwendeten Tabelle. Insbesondere spielt die Aggregation für die großen Strukturierungsgradübergänge, wie die in Abbildung 6.1 und 6.2 gezeigten, eine wichtige Rolle. In Abschnitt 5.2.3 wurde auf diese Problematik hinsichtlich der Modellierung der Aggregation ausführlich eingegangen. Für die Implementation der Aggregationsoperationen wurde auf die von Bobzien (2001) und Schürer (2002) beschriebenen Verfahren und Bewertungs- sowie Priorisierungsvorschriften zurückgegriffen.

Eine statistische Auswertung der in Abbildung 6.2 dargestellten Modellgeneralisierung findet sich in Tabelle 6.1. Die Güte der Flächenzuordnungen zwischen ungeneralisiertem und generalisiertem Datenbestand kann zusätzlich durch den optischen Eindruck vermittelt werden.

6.1.2 Invariante topologische Eigenschaften

Für den Erhalt der topologischen Eigenschaften kommen hier die in Abschnitt 5.3 entwickelten Bedingungen zum Tragen, also insbesondere die Erhaltung von Inzidenz-, Adjazenz- und Inklusionsrelationen. Die Algorithmen zur Generalisierung sind so angelegt, dass die topologischen Relationen weitgehend erhalten bleiben. Ausnahmen bilden Relationsverluste in Folge von Geometriertypwechseln. In Abschnitt 5.3.1 finden sich Beispiele für Geometriertypwechsel und von ihnen verursachte Inkonsistenzen bezüglich der topologischen Relationen. Für die Kontrolle der Konsistenz wurde eine Prüfroutine entwickelt, die für alle Paare von topologisch in direkter Relation stehenden³ Geo-Objekten prüft, ob die Bilder dieser Geo-Objekte nach der Generalisierung dieselbe topologische Relation erfüllen. Auf diese Weise wurden die oben erwähnten Inkonsistenzen aufgedeckt, und es wurde nach einer Analyse festgestellt, dass die Ursache im Wechsel des Geometriertyps lag. Die Algorithmen zum Geometriertypwechsel sind in Bobzien und Morgenstern (2003) beschrieben.

6.1.3 Invariante geometrische Eigenschaften

Zur geometrischen Vereinfachung im Zuge der Modellgeneralisierung wurden drei verschiedene Algorithmen verwendet, die alle drei im Abschnitt 5.4 beschrieben sind. Es handelt sich um den Algorithmus von Douglas und Peucker, den Algorithmus der gleitenden Blende (nach Weber) und den Algorithmus von Perkal. Die ersten beiden Algorithmen wurden von Schürer (2002) ausführlich im Zusammenhang mit der Modellgeneralisierung untersucht und beschrieben. Schürer (2002) schlägt vor, beim Übergang vom DLM 25 zum DLM 250 den Algorithmus der gleitenden Blende für Straßen anzuwenden und den Algorithmus von Douglas und Peucker für alle anderen Objektarten. Diese Vorschläge wurden für die Umsetzung der Modellgeneralisierung im Rahmen dieser Arbeit übernommen und größtenteils, jedoch mit anderer Parametrisierung, auch für den Übergang vom DLM 250 zum DLM 1000 verwendet.

Für den Übergang vom DLM 250 zum DLM 1000 hat Stenger (2003) den Algorithmus von Perkal untersucht und insbesondere auf Verwaltungsgrenzen und Begrenzungen von Ortslagen angewendet. Er ist zu dem Schluss gekommen, dass der Algorithmus sich zur Glättung, insbesondere von *spikes* (schmalen, spitzen Ausbuchtungen, „Dornen“), sehr gut eignet.

Die Ergebnisse von Stenger (2003) wurden genutzt, und der Algorithmus von Perkal wurde für die genannten Objektarten mit einer entsprechenden Parametrisierung verwendet. Probleme ergaben sich im Zusammenhang mit dem Algorithmus von Perkal jedoch bei der Erhaltung der topologischen Konsistenz bei gleichzeitiger geometrischer Vereinfachung. Dieses ist in Abschnitt 6.1.4 im Absatz über topologische Konsistenz und geometrische Generalisierung beschrieben.

³Dies sind alle topologischen Relationen außer der Disjunktheit.

6.1.4 Invariante kombinierte Eigenschaften

In Abschnitt 5.5 wurden drei Beispiele für kombinierte Eigenschaften von Geodatenmen-gen vorgestellt, die bei der Generalisierung invariant gehalten werden müssen. Alle drei Eigenschaften werden im Folgenden am Beispiel der Modellgeneralisierung erläutert.

Erreichbarkeit und Netzzusammenhang

Der Netzschluss wurde vom Generalisierungsalgorithmus in zweierlei Hinsicht unter-stützt. Zum einen wurden aneinander grenzende linienförmige Objekte, die nach der Generalisierung dieselben Attributwerte erhalten würden, schon vor der Generalisierung zu einem Objekt zusammengefasst (Datenvorverarbeitung). Dies ist insofern von Be-deutung, weil vielerorts Erfassungskriterien über die Länge von linienförmigen Objekten definiert sind. Wenn also schon vor der Generalisierung mehrere linienförmige Objek-te zu einem Objekt zusammengefasst werden können, so sinkt die Wahrscheinlichkeit, dass wegen einem Längenkriterium ein Objekt entfällt und dadurch eine Lücke im Netz entsteht.

Dennoch kann der Algorithmus nicht alle Netzerstückelungen vermeiden. Deshalb wur-de zum anderen ein Netzzusammenhangstest angefügt, der durch die Generalisierung verursachte Lücken entdeckt und markiert, und so für eine automatische oder manuelle Nachbearbeitung bereitstellt.

Die Erreichbarkeit von flächenförmigen oder punktförmigen Objekten über ein linien-förmiges Netz ist im Objektartenkatalog nicht explizit spezifiziert. Da der Netzzusam-menhang aber aus kartographischer Sicht eine große Bedeutung besitzt, wurde in den Generalisierungsalgorithmus ein Mechanismus zur automatischen Wiederherstellung des Netzzusammenhangs implementiert. Das Beispiel einer Ortslage, die nur über (kurze) Nebenstraßen erreichbar ist, wird an dem Ausschnitt in Abbildung 6.3 (a) an Hand von zwei von mehreren möglichen Generalisierungen gezeigt. Jede Straße, die zu dieser Orts-lage führt, ist für sich genommen zu kurz, um die Erfassungskriterien zu erfüllen. Folgt man diesen strikt, so ergibt sich das Ergebnis aus Abbildung 6.3 (b): Die Ortslage ist mit dem Straßennetz nicht mehr verbunden. Dies kann durchaus ein gewünschtes Ergebnis sein (vergleiche topographische Übersichtskarten im Maßstab 1 : 500 000 oder 1 : 1 Mio). Misst man jedoch der Erreichbarkeit eine höhere Bedeutung bei, wie sie zum Beispiel für Routenberechnung in Straßenkarten benötigt wird, so können nach Testen der Er-reichbarkeit und einer Aufdeckung einer Inkonsistenz bezüglich dieser eine oder mehrere Straßen wieder eingefügt werden, um die Erreichbarkeit der Ortslage wieder herzustellen. In dem Fall von Abbildung 6.3 (c) wurde die kürzeste Straße als Verbindung zum überörtlichen Straßennetz wieder hergestellt.

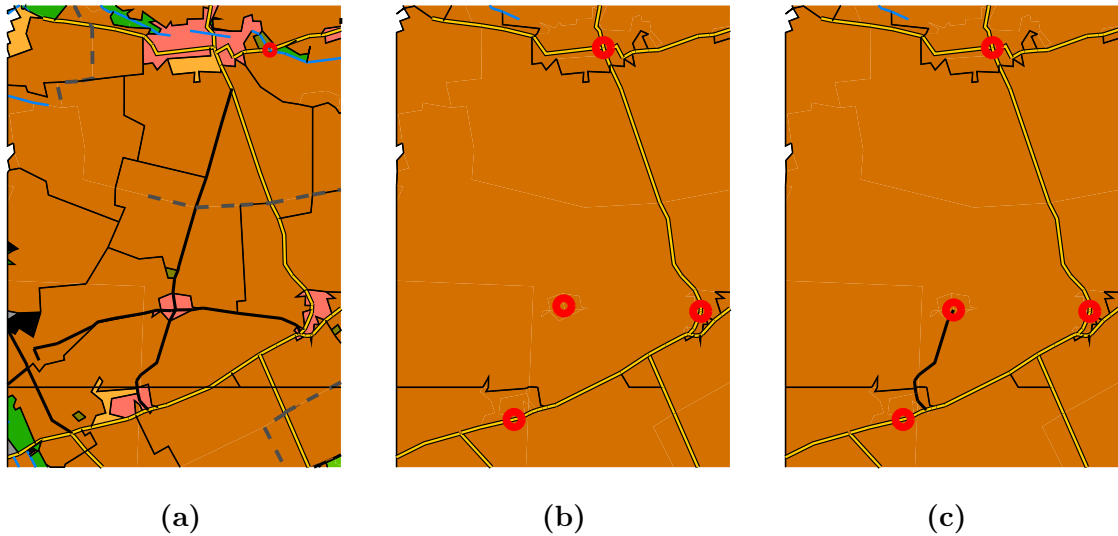


Abbildung 6.3: Erreichbarkeit einer Ortslage (Vergrößerter Ausschnitt aus Abbildung 6.2): (a) Ausgangsdaten DLM 250, (b) DLM 1000, generalisiert ohne Berücksichtigung der Erreichbarkeit, (c) DLM 1000, generalisiert mit Berücksichtigung der Erreichbarkeit. (Vergleiche auch Abbildung 5.23 auf Seite 143.)

Flächendeckung

Durch die Implementierung der Algorithmen zur Flächenzusammenfassung auf der Ebene der Grundobjekte aus (Bobzien 2001) und aus (Jakobitz 1997) und unter der Voraussetzung, dass die Ursprungsdaten flächendeckend sind, wurde die Flächendeckung vom Flächenzusammenfassungsalgorithmus sicher gestellt.

Topologische Konsistenz und geometrische Generalisierung

Als Vorbemerkung zu diesem Thema sei daran erinnert, dass im hier verwendeten Datenmodell zwei DLM vollständig nebeneinander existieren, sowohl auf thematischer, als auch auf topologischer und geometrischer Ebene (siehe Abschnitt 3.2.2). Die geometrische Generalisierung erfolgt demnach *nicht* durch ein Entfernen von Punkten aus einer Linie, sondern durch den Aufbau einer neuen Geometrie (abgeleitet aus der ungeneralisierten Geometrie) im generalisierten Datenbestand. Dementsprechend können neu erzeugte Geometrien nur mit anderen bereits erzeugten Geometrien auf topologische Konsistenz geprüft werden. Keinesfalls jedoch kann die gesamte Geometrie gleichzeitig generalisiert werden, was einer Generalisierung *en masse* entspräche (siehe Abschnitt 5.5.3). Stattdessen wurde eine Generalisierung *in vacuo* gewählt, bei der jede neu erzeugte Geometrie kumulativ mit bereits erzeugten Geometrien auf topologische Inkonsistenz hin geprüft

wird. Das Aufdecken einer Inkonsistenz hat *nicht* eine sofortige Behebung zur Folge, sondern die beteiligten beiden Geometrien werden als Paar in eine separate Datenstruktur abgelegt (und somit temporär gespeichert).

Erst nach vollständiger Erzeugung sämtlicher generalisierter Geometrien werden die gespeicherten Paare von Geometrien, die topologische Inkonsistenzen hervorriefen, nacheinander abgearbeitet. In die entsprechenden Liniengeometrien der generalisierten Daten werden solange Ausgangspunkte eingefügt, bis der Topologiefehler nicht mehr auftritt.

An zwei Stellen konnte dieses Vorgehen allerdings zu auf diese Weise nicht lösbareren Problemen führen:

1. **Geometriotypwechsel:** In Folge von Geometriotypwechseln neu berechnete Geometrie hat keine Ursprungsgeometrie, aus der Stützpunkte wieder eingefügt werden können. Hier muss im Zweifelsfall manuell eingegriffen werden.
2. **Algorithmus von Perkal:** Wird der Algorithmus von Perkal zur Vereinfachung und Glättung einer Linie herangezogen, so kann die generalisierte Linie Stützpunkte enthalten, die in der Ursprungslinie nicht enthalten waren. (Dieser Algorithmus berechnet neue Stützpunkte). Dementsprechend würde ein Einfügen von ursprünglichen Stützpunkten zu unvorhersehbaren Ergebnissen führen. Gibt es hier einen nicht automatisch behebbaren Konflikt, so werden die Daten mit einem anderen Algorithmus erneut generalisiert und das Prozedere mit den so entstandenen Daten wiederholt.

Bei dem oben beschriebenen Vorgehen des Wiedereinfügens von vorher bereits gelöschten Stützpunkten ist zu beachten, dass durch die Korrektur der generalisierten Daten erneut topologische Fehler entstehen können. Folglich muss nach dem Einfügen wiederum mit anderen Objekten im Zieldatenbestand auf topologische Konsistenz getestet, und die Methode zur Inkonsistenzbehebung gegebenenfalls iteriert werden.

6.2 Generalisierung geologischer Karten

Geologische Karten dienen zur regionalgeologischen und stratigraphischen Darstellung des unmittelbaren Untergrundes. Die geologische Karte veranschaulicht damit die Verteilung der verschiedenen Gesteine und Schichten auf der Erdoberfläche (Wagenbreth 1958).

Die geologische Karte basiert meist auf einer topographischen Karte. So existieren zum Beispiel in Deutschland als Kartenwerke die *geologische Spezialkarte* im Maßstab 1 : 25 000 und die *geologische Übersichtskarte* im Maßstab 1 : 200 000, die beide auf den topographischen Karten im gleichen Maßstab basieren (TK25 beziehungsweise TÜK200), sowie die Geologische Karte der Bundesrepublik Deutschland im Maßstab 1 : 1 Mio. Letztere Karte wird zur Zeit neu bearbeitet und an die Topographie des DLM 1000 angepasst (BGR 2000).

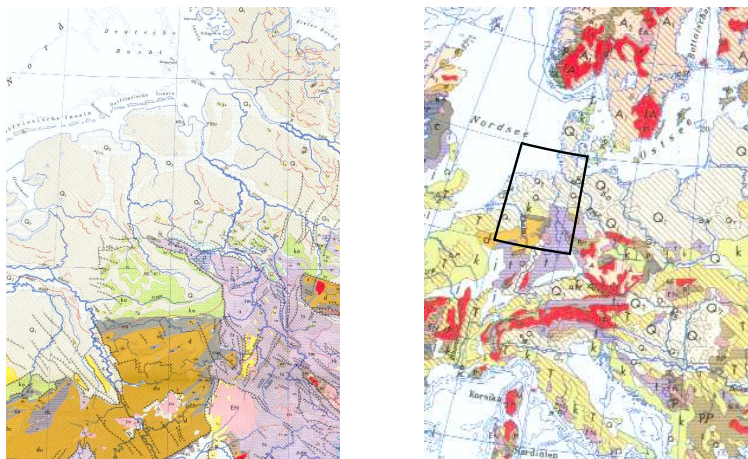


Abbildung 6.4: *Geologische Karten in zwei Maßstäben. Der Ausschnitt der linken Karte entspricht in etwa dem schwarz umrandeten Ausschnitt in der rechten Karte (Ausschnitte aus Bederke und Wunderlich 1968). Die Maßstäbe der Originalkarten sind 1:2,5 Mio und 1:25 Mio, beide Karten sind hier verkleinert dargestellt*

Einzelne Elemente der topographischen Karte werden in die geologische Karte übernommen und dort dargestellt. Dies umfasst die Gewässer, einzelne Verkehrswege sowie Ortslagen. Sie dienen der Orientierung und wie im Falle der Gewässer auch zur Erläuterung des geologischen Karteninhalts. In einigen Fällen wird eine topographische Karte als Kartengrund für eine geologische Karte genutzt. In solchen Fällen muss die generalisierte geologische Karte an den (durch kartographische Generalisierung erstellten) generalisierten Kartengrund angepasst werden. Die Generalisierung des Kartengrundes wird hier aber nicht weiter verfolgt, sondern es erfolgt eine Konzentration auf die Generalisierung der geologischen Kartenelemente.

Die geologischen Inhalte umfassen flächen-, linien- und punktförmige Signaturen sowie Textsignaturen. Die *flächenförmigen Signaturen* bezeichnen verschiedene Gesteine und stratigraphische Einheiten (Wagenbreth 1958). Sie sind das wichtigste Darstellungselement der geologischen Karte, weil durch sie die Verteilung der Gesteine und Schichten auf der Erdoberfläche dargestellt wird. Diese thematischen Einheiten sind hierarchisch unterteilt und bilden damit einen wichtigen Anknüpfungspunkt für die Übertragung der an Hand der Modellgeneralisierung gewonnenen invarianten Eigenschaften auf die Generalisierung von geologischen Karten. Eine weitere wichtige Eigenschaft der flächenförmigen Signaturen in geologischen Karten ist, dass sie die Kartenfläche vollständig ausfüllen – sofern geologische Erkenntnisse über dieses Gebiet vorliegen. Mithin ergeben die flächenförmigen Signaturen eine *flächendeckende Darstellung* des erfassten Untergrundes. Die Flächendeckung ist in diesem Abschnitt ein weiterer Anknüpfungspunkt zwischen Modellgeneralisierung und Generalisierung geologischer Karten.

Erdzeitalter	System / Formation	Serie / Abteilung
Mesozoikum	Jura	Malm Dogger Lias
	Trias	Keuper Muschelkalk Buntsandstein
Paläozoikum	Perm	Zechstein Rotliegendes

Tabelle 6.2: Ausschnitt aus einer geologischen Zeittafel (nach Vossmerbäumer 1976).

Des Weiteren kommen in geologischen Karten *linienförmige Signaturen* vor, von denen die Darstellung von *Verwerfungen* die charakteristischste für geologische Karten ist. Diese Signaturen können mit den Grenzen der Flächensignaturen übereinstimmen. Dies ist jedoch nicht zwingend. Sie werden in kleinmaßstäbigen Karten, wie zum Beispiel der Geologischen Karte der Bundesrepublik Deutschland, nicht mehr dargestellt.

Punktförmige Signaturen in geologischen Karten umfassen auffällige Gesteinsblöcke, Einfallen und Streichen von Schichten, sowie Bohrpunkte und Zeichen für besondere Aufschlüsse (Wagenbreth 1958). Die meisten Punktsignaturen finden nur in großmaßstäbigen Karten Verwendung und fallen bei einer Generalisierung in eine Geologische Übersichtskarte weg.

Textsignaturen dienen in geologischen Karten zumeist der Erläuterung von flächenförmigen Signaturen und bezeichnen durch ein Kürzel die stratigraphische Einheit der darunter liegenden Flächensignatur.

In Abbildung 6.4 sind zwei kleinmaßstäbige geologische Karten dargestellt, wobei der gemeinsame Teil den Nordwesten Deutschlands umfasst. In beiden Karten dominieren die flächenförmigen Signaturen, unterstützt von Textsignaturen. Linienförmige Signaturen (Verwerfungen) sind in der linken Karte noch dargestellt, in der rechten nicht mehr.

Wie oben erwähnt, nehmen die Flächensignaturen in geologischen Karten die wichtigste Rolle ein. Die Einteilung der stratigraphischen Einheiten wird in der Geologie mit Hilfe einer geologischen Zeittafel dargestellt. Diese ist hierarchisch aufgebaut, mit meist drei, manchmal auch vier oder Hierarchiestufen. Die Bezeichnung der Hierarchiestufen ist je nach Quelle unterschiedlich. Wir orientieren uns an der geologischen Zeittafel nach Vossmerbäumer (1976), die in Tabelle 6.2 ausschnittsweise wiedergegeben ist. Die dort verwendeten Hierarchiestufen sind das *Erdzeitalter*, das *System* beziehungsweise die *Formation* und die *Serie* beziehungsweise die *Abteilung*. Diese Hierarchie spiegelt sich oft auch in der Signaturierung wieder. Erdzeitalter werden einer Grundfarbe zugeordnet, Systeme

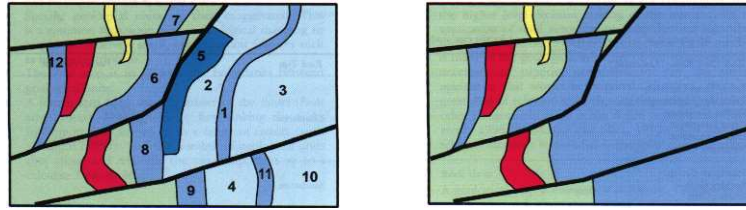


Abbildung 6.5: Automatische Flächenzusammenfassung in geologischen Karten (aus Downs und Mackaness 2002).

erhalten unterschiedliche Farbintensitäten dieser Grundfarbe, und Serien werden durch unterschiedliche Schraffuren unterschieden.⁴

Die *Automation* der Generalisierung geologischer Karten ist ein Thema, das in der Forschung immer hinter der Generalisierung topographischer Karten zurückgeblieben ist. Aus der kleinen Menge von Veröffentlichungen, die sich mit diesem Thema beschäftigen, sei die Arbeit von Downs und Mackaness (2002) herausgehoben. Dort wird die Problematik der Generalisierung von geologischen Karten ausführlich untersucht und eine Implementation für die automatische Generalisierung vorgestellt, zusammen mit einigen Beispielen von generalisierten Karten. Deutlich herausgehoben wird die Bedeutung von zum einen *kategorischen Bedingungen*, die sich in der eindeutigen Zuordnung einer Fläche zu einer stratigraphischen Einheit wieder findet. Abbildung 6.5 ist dieser Arbeit entnommen und zeigt die Flächenzusammenfassung von geologischen Einheiten. Zum anderen heben Downs und Mackaness die Bedeutung von Verwerfungen, also von linienförmigen Signaturen, heraus und untersuchen die Anforderungen an ihre Generalisierung.

Im Folgenden sollen die Ähnlichkeiten der Generalisierung von geologischen Karten zur Modellgeneralisierung aufgezeigt werden. Zuerst fällt auf, dass die kategorische Zuordnung von Flächen zu stratigraphischen Einheiten der thematischen Zuordnung von Flächen zu einer Objektart in Digitalen Landschaftsmodellen entspricht. Selbst die Zuordnung einer Fläche zu zwei oder mehr Objektarten in einem DLM (zum Beispiel die Überlagerung von Industriefläche und Gebäude, oder die Überlagerung von Naturschutzgebiet und Wald) findet sich in geologischen Karten wieder. Wenn die oberste Gesteinsschicht nicht sehr mächtig ist, und man auch die darunter liegende Gesteinsschicht darstellen möchte, so wird eine neue Signatur durch breite, senkrechte Striche als Kombination der beiden Signaturen erzeugt (Wagenbreth 1958) und damit die Fläche beiden stratigraphischen Einheiten zugeordnet.

Bei der Generalisierung der flächenförmigen Signaturen ist zuallererst die Minimaldimension ausschlaggebend. Wird eine Fläche in der Darstellung zu klein, so muss sie

⁴Gelegentlich werden in geologischen Karten auch mehrere sich überlagernde Schichten dargestellt, indem zweifarbige Schraffuren oder eine Grundfarbe und eine Signatur verwendet werden (Vossmerbäumer 1983). Diese Darstellungsart integriert die dritte Dimension in die zweidimensionale Kartendarstellung. Diese Fälle werden im Folgenden jedoch nicht weiter behandelt. Sie werfen im Zusammenhang mit der Generalisierung aber interessante Fragestellungen auf.

Abteilung	Formation
Malm	Jura
Dogger	Jura
Lias	Jura
Keuper	Trias
Muschelkalk	Trias
...	...

Tabelle 6.3: Ausschnitt aus der Zuordnungsfunktion zwischen geologischen Karten unterschiedlicher Auflösung.

mit einer benachbarten Fläche zusammengefasst werden. Die Entscheidung, mit welcher benachbarten Fläche zusammengefasst wird, wird ähnlich wie bei der Modellgeneralisierung getroffen: Zum einen können thematische Gründe ausschlaggebend sein, also die Antwort auf die Frage, welche Nachbarfläche thematisch am nächsten ist (Downs und Mackaness 2002). Zum Beispiel sind Lias und Malm dem Dogger am ähnlichsten. Dies entspricht der *Priorisierungstabelle* wie sie in Abschnitt 6.1.1 beschrieben ist. Zum anderen können aber auch geometrische Gründe die Zusammenfassung beeinflussen. Dies findet sich auch in der Modellgeneralisierung, wie zum Beispiel bei der Kompaktheit von Flächen, wieder (Bobzien 2001). Ein drittes Kriterium ist das Zusammenspiel von Flächen und Verwerfungen. Dieses Zusammenspiel ist spezifisch für geologische Karten und findet kein Äquivalent in der Modellgeneralisierung.

Abgesehen von den Kriterien, die bei der Flächenzusammenfassung eine Rolle spielen, muss auch vor der Generalisierung bestimmt werden, welchen Detaillierungsgrad die generalisierte Karte erhalten soll, also welche geologischen Hierarchiestufen dargestellt werden sollen. In geologischen Spezialkarten zum Beispiel möchte man alle geologischen Abteilungen darstellen, in der geologischen Übersichtskarte möchte man nur noch Formationen darstellen, in der Geologischen Karte der Bundesrepublik Deutschland nur noch die Erdzeitalter (siehe Abbildung 6.4).

Für eine Generalisierung können dafür entsprechend der Hierarchie der geologischen Zeittafel alle Abteilungen, die zu einer Formation gehören, durch diese Abteilung ersetzt werden, beziehungsweise alle Formationen, die zu einem Erdzeitalter gehören, durch dieses Erdzeitalter.

Es ergibt sich damit kanonisch eine Abbildung der Abteilungen in die Formationen und von dort in die Erdzeitalter. Diese Abbildung entspricht formell der Objektartenzuordnungsfunktion g_O aus Abschnitt 5.2. Dort wurde die Zuordnungsfunktion durch eine Tabelle ausgedrückt. Für die geologischen Einheiten ist in Tabelle 6.3 an Hand von Stufen und Abteilungen diese Zuordnung aufgezeigt, abgeleitet aus der geologischen Zeittafel aus Tabelle 6.2.

Die Bedingung 5.1 von Seite 111 besagt, übertragen auf die Generalisierung von geolo-

gischen Karten, dass die Zugehörigkeit einer Fläche zu einer Hierarchiestufe gleich der Abbildung der Hierarchiestufe der ungeneralisierten Fläche sein muss. Diese Bedingung gilt damit auch hier und bildet damit ein Kriterium, dass von einer Generalisierung erfüllt werden muss.

Ebenso wie die thematischen invarianten Eigenschaften können auch die topologischen und geometrischen Eigenschaften auf die Generalisierung geologischer Karten übertragen werden. Für die *Topologie* gilt insbesondere, dass die Adjazenzen (siehe Abschnitt 5.3.1) erhalten bleiben müssen. Am Beispiel der Flächen heißt das, dass benachbarte Flächen auch nach der Generalisierung benachbart sein müssen. Sie dürfen nicht durch eine andere Fläche getrennt werden, wenn dies nicht schon vor der Generalisierung der Fall war. Aber auch die Inzidenzen von Verwerfungen zu benachbarten Flächen müssen bei der Generalisierung erhalten bleiben. Ebenso muss die Inklusionsrelation von Punktssignaturen (sofern sie nicht weggeneralisiert werden) zu den umgebenden Flächen gewahrt bleiben.

Die *geometrische Vereinfachung* von Linien beziehungsweise Flächengrenzen ist auch bei der Generalisierung geologischer Karten von Bedeutung, da auch hier durch zu enge geometrische Konstrukte die Deutlichkeit einer Karte leiden würde. Dementsprechend können die Ausführungen für die Modellgeneralisierung in Abschnitt 5.4 auf die Generalisierung geologischer Karten übertragen werden.

Nicht übertragen werden hingegen kann die invariante kombinierte Eigenschaft der *Erreichbarkeit und des Netzzusammenhangs*. In geologischen Karten kommt den Liniensignaturen keine *verbindende* Bedeutung zu. Im Gegenteil hat gerade die Darstellung der Verwerfungen eine *trennende* Bedeutung. Die Invarianzbedingung zur Erreichbarkeit und zum Netzzusammenhang aus der Modellgeneralisierung gilt hier also nicht.

Anders verhält es sich jedoch bei der *Flächendeckung*. Gerade diese Bedingung findet sich in geologischen Karten wieder, da naturgemäß die gesamte Erdoberfläche stratigraphischen Einheiten zugeordnet werden kann. In der Darstellung gibt es natürlich die Ausnahme von nicht erfassten Gebieten, denen keine geologische Einheit zugeordnet ist. Durch den heutigen hohen Erkenntnisstand kann dies jedoch vernachlässigt werden, oder durch eine besondere Zuordnung gekennzeichnet werden. Dementsprechend kann die Flächendeckung auch bei geologischen Karten gefordert werden, und für die Automation durch die Überprüfung der Bedingung 5.8 auf Seite 145 sicher gestellt werden.

Die invariante kombinierte Eigenschaft der *topologischen Konsistenz bei geometrischer Generalisierung* kann ebenfalls von der Modellgeneralisierung auf die Generalisierung von geologischen Karten übertragen werden. Denn hier ebenso wie dort dürfen keine Splitterpolygone und keine zusätzlichen Kreuzungen durch die geometrische Vereinfachung von Flächenbegrenzungslinien entstehen.

Im Anschluss an den nachfolgenden Abschnitt über die Übertragung der invarianten Eigenschaften von der Modellgeneralisierung auf die Generalisierung von Straßenkarten werden in Tabelle 6.5 die Ergebnisse dieser beiden Abschnitte zusammengefasst.

6.3 Generalisierung von Straßenkarten

Straßenkarten dienen der Orientierung auf der Reise und der Routenplanung und -findung. Sie enthalten in erster Linie Informationen über Verkehrswege, zuvorderst über Straßen, aber gelegentlich auch über andere Verkehrswege wie zum Beispiel Eisenbahnen. Des Weiteren können sie einige geomorphologische oder topographische Inhalte wie Gewässer, Wald oder Gebirgsschummerung enthalten.

Dem Zweck von Straßenkarten entsprechend muss ihre Generalisierung in erster Linie die Verkehrswege und insbesondere die Straßen beachten. Aus den bisher betrachteten Eigenschaften, die bei einer Generalisierung invariant gehalten werden sollen, sticht im Zusammenhang mit Straßenkarten die Erreichbarkeit beziehungsweise der Netzzusammenhang heraus. Aber auch andere Eigenschaften müssen bei der Generalisierung von Straßenkarten invariant sein, was im Folgenden gezeigt wird.

Die Abbildungen 6.6 und 6.7 zeigen Ausschnitte aus Straßenkarten, jeweils mit zwei unterschiedlichen Auflösungen (Autoroute 2004, Hallwag International 2003). Deutlich ist zu erkennen, dass Straßen höherer Ordnung (wie Autobahnen) im Allgemeinen erhalten bleiben, während Straßen geringerer Ordnung nur erhalten bleiben, wenn sie wichtige Erschließungsfunktionen besitzen. Dies liefert einen ersten Hinweis auf thematische Eigenschaften von Straßenkarten und ihre Generalisierung. Aber auch andere Eigenschaften von Straßenkarten werden im Folgenden betrachtet.

Die *thematischen Eigenschaften* von Straßenkarten betreffen in erster Linie die Einteilung von Straßen in *Straßenklassen*, wie zum Beispiel Autobahnen, Autostraßen, Hauptstraßen sowie regionale und lokale Verbindungsstraßen (siehe Legende eines Straßenatlas in

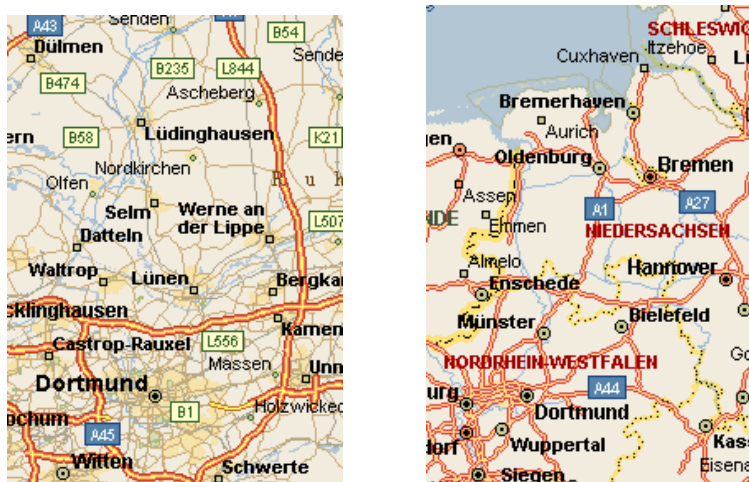


Abbildung 6.6: Straßenkarten in zwei Maßstäben aus einem elektronischen Straßenatlas (aus Autoroute 2004).

Kapitel 6 Anwendungen

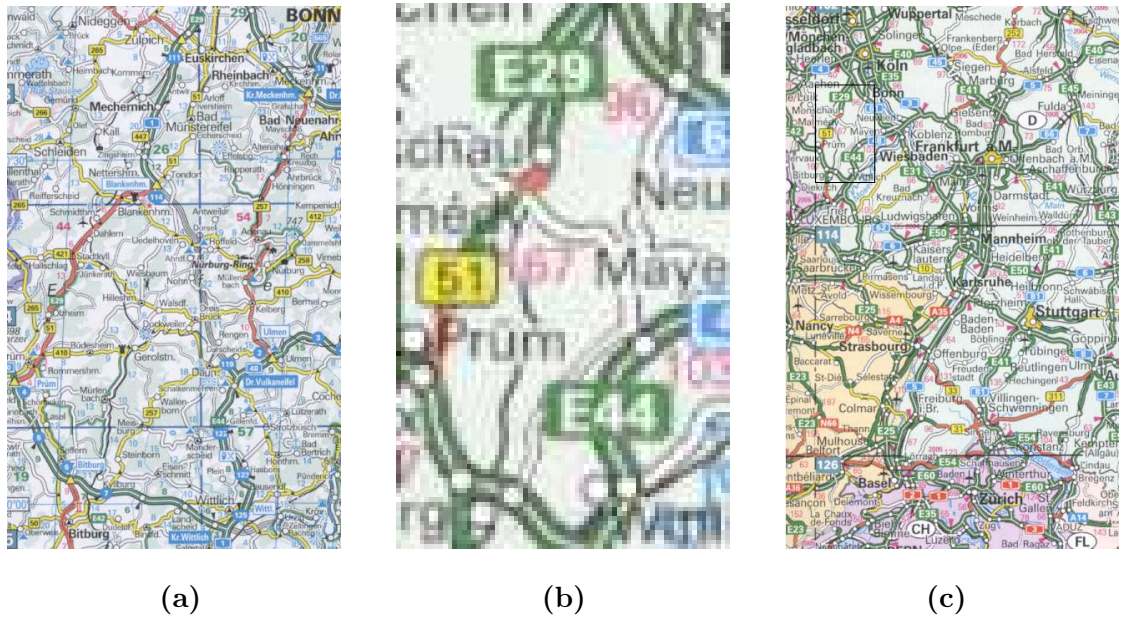


Abbildung 6.7: Straßenkarten in zwei Maßstäben: (a) Ausgangskarte im Maßstab 1:800 000, (b) und (c) im Maßstab 1:4,5 Mio, in (b) auf den Maßstab 1:800 000 vergrößert. Alle drei Darstellungen sind im gleichen Verhältnis verkleinert (aus Hallwag International 2003).

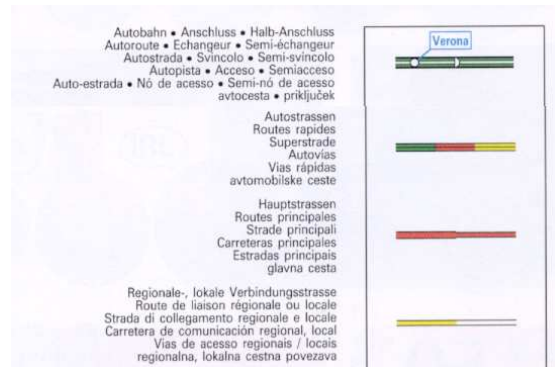


Abbildung 6.8: Legende eines Straßenatlas, gültig für beide in Abbildung 6.7 verwendeten Auflösungen (Ausschnitt aus Hallwag International 2003).

höhere Aufl.	niedr. Aufl.	höhere Aufl.	niedr. Aufl.
Autobahn	Autobahn	Autobahn	Autobahn
Autostraße	Autostraße	Autostraße	Autostraße
Hauptstraße	Hauptstraße	Hauptstraße	Hauptstraße
Reg. Verb.-str.	Reg. Verb.-str.	Hauptstraße	Reg. Verb.-str.
Lok. Verb.-str.	Lok. Verb.-str.	Hauptstraße	Lok. Verb.-str.
		Reg. Verb.-str.	Reg. Verb.-str.
		Reg. Verb.-str.	Lok. Verb.-str.
		Lok. Verb.-str.	Lok. Verb.-str.

(a)

(b)

Tabelle 6.4: Zwei mögliche Zuordnungsrelationen für den Straßenatlas aus den Abbildungen 6.7 und 6.8: (a) 1:1-Zuordnung, (b) $m:n$ -Zuordnung.

Abbildung 6.8). Im Allgemeinen geht man davon aus, dass eine Straße bei der Generalisierung keine Bedeutungsverschiebung erfährt, dass zum Beispiel eine Autobahn immer zu einer Autobahn, eine Hauptstraße immer zu einer Hauptstraße generalisiert wird. Eine genauere Betrachtung der Legende in Abbildung 6.8 und der Ausschnitte aus dem zugehörigen Straßenatlas in Abbildung 6.7 zeigt jedoch, dass hier Differenzierungen nötig sind. Zum einen bietet die Legende für verschiedene Straßenklassen mehrere Signaturen an (für Autostraßen und Hauptstraßen), zum anderen finden sich Beispiele in den Karten, bei denen Straßen einem Klassenwechsel erfahren. Zum Beispiel ist die B257 zwischen Meckenheim und Ulmen in Abbildung 6.7 (a) als Hauptstraße signaturiert, während sie in (b) und (c) als lokale Verbindungsstraße signaturiert ist. Dies kann einerseits eine fehlerhafte Generalisierung bedeuten, aber andererseits auch eine so gewünschte Bedeutungsverschiebung.

Durch eine Verknüpfung der Thematik beider Kartenaufösungen kann im ersten Fall die Inkonsistenz aufgedeckt, im zweiten Fall die Bedeutungsverschiebung modelliert werden. Hierzu definiert man wie in Abschnitt 5.2.1 beschrieben eine Zuordnungsrelation R_O von den Straßenklassen der höheren Auflösung zu den Straßenklassen der niedrigeren Auflösung. In Tabelle 6.4 (a) sind die Zuordnungen im Verhältnis 1:1 angenommen, so dass das obige Beispiel eine Inkonsistenz darstellt. In (b) sind die Zuordnungen so gewählt, dass die Bedeutungsverschiebung modelliert wird. Mit Hilfe dieser Tabellen kann die Bedingung 5.2 von Seite 113 an Hand der Generalisierung der Straßenkarten überprüft werden. Im Fall (a) würde so die Inkonsistenz aufgedeckt, in Fall (b) die beabsichtigte Bedeutungsverschiebung als korrekt erkannt.

Die wichtigsten *topologischen Eigenschaften* einer Straßenkarte sind die Inzidenz zwischen Knoten (Orten und Städten) und Kanten (Straßen) sowie die Adjazenz zwischen Kanten (Straßen). Diese Beziehungen dürfen durch die Generalisierung nicht verloren gehen, sofern die beteiligten topologischen Elemente in der generalisierten Karte eine Entsprechung finden. Die in Abschnitt 5.3 formulierten Bedingungen können hier angewendet werden. Inklusionsrelationen sind in Straßenkarten hingegen weniger von Bedeutung, da

	Digitale Landschaftsmodelle	Geologische Karten	Straßenkarten
Thematische Eigenschaften	Objektarten Attribute	stratigraphische Einheiten —	Straßenklassen —
Topologische Eigenschaften	Inzidenz Adjazenz Inklusionsrelation	Flächen und Verwerfungen Flächen —	Straßen und Orte Straßen —
Geometrische Eigenschaften	ja	ja	ja
Kombinierte Eigenschaften	Erreichbarkeit und Netzzusammenhang Flächendeckung Topolog. Konsistenz und geom. Vereinfachung	— inhärent ja	inhärent — ja

Tabelle 6.5: Überblick über invariante Eigenschaften bei der Generalisierung von Digitalen Landschaftsmodellen, von geologischen Karten und von Straßenkarten.

die flächenförmigen Signaturen eine untergeordnete Rolle spielen.

Ebenso wie bei geologischen Karten sind auch bei Straßenkarten die *geometrischen Eigenschaften* der Objekte wichtig. Denn genau wie dort spielt die Lesbarkeit einer Karte eine wichtige Rolle. Zu enge geometrische Konstrukte können diese beeinträchtigen.

Bezogen auf die *kombinierten Eigenschaften* ist der *Netzzusammenhang* von herausragender Bedeutung, da der Sinn und Zweck einer Straßenkarte darin besteht, *Erreichbarkeiten* auszudrücken. Im Gegensatz zu Digitalen Landschaftsmodellen oder auch zu topographischen Karten im kleinen Maßstab, in denen Ortslagen oder Orte auch ohne Straßenanschluss dargestellt sein können, wird man in Straßenkarten – von wenigen Ausnahmen abgesehen – keine Orte ohne Straßenanschluss finden. Ein Verlust der Erreichbarkeit hat also bei Straßenkarten – im Gegensatz zur Generalisierung eines Digitalen Landschaftsmodell – ein Fortlassen von Orten zur Folge (vergleiche dazu Abschnitt 5.5.1).

Die *Flächendeckung* hingegen ist bei Straßenkarten – die sich in erster Linie auf linienförmige Objekte konzentriert – nicht zu finden. Es gibt im Gegenteil viele Restflächen, denen keine Flächensignatur zugeordnet sind. Die Karte hat anders ausgedrückt an diesen Orten weiße Flecken, die eben keine Inkonsistenz darstellen.

Die invariante Eigenschaft der *topologische Konsistenz bei der geometrischen Generalisierung* ist ebenso wie bei der Generalisierung der geologischen Karten auf die Generalisierung von Straßenkarten übertragbar. Gerade bei der Generalisierung der Straßen dürfen die bei der Linienvereinfachung möglicherweise resultierenden Inkonsistenzen, wie die Selbstüberschneidung einer Linie, der Schnitt mit einer anderen Linie und der relative Lagewechsel, nicht auftreten. Die im Abschnitt 5.5.3 skizzierten Bedingungen müssen auch bei der Generalisierung von Straßenkarten berücksichtigt werden.

6.4 Fazit

Dieses Kapitel erläutert die in Kapitel 5 entwickelten Bedingungen zunächst am Beispiel der Modellgeneralisierung. Die vorgestellten Ergebnisse der Generalisierung von Landschaftsmodellen erfolgen mit einer Implementation, die auf den aufgestellten Bedingungen basiert. Der weitere Verlauf dieses Kapitels zeigt an Hand von zwei anderen Generalisierungsanwendungen die Übertragung der Bedingungen aus der Modellgeneralisierung. Als Beispiele dienen zum einen die Generalisierung von geologischen Karten und zum anderen die Generalisierung von Straßenkarten.

Tabelle 6.5 fasst die Ergebnisse der Übertragung der invarianten Eigenschaften von der Modellgeneralisierung auf die Generalisierung von geologischen Karten und Straßenkarten zusammen. Erkennbar ist, dass große Teile der Modellgeneralisierung auf die Generalisierung von geologischen Karten und Straßenkarten übertragen werden können.

Nicht übertragen werden können die invarianten Eigenschaften der *Attribute*, weil die attributiven Eigenschaften in Karten anders als in Landschaftsmodellen nicht modelliert werden. Die Inklusionsrelation innerhalb der topologischen Eigenschaften wird auch nicht übertragen. Die Ursache dafür liegt bei den geologischen Karten darin begründet, dass nur in Ausnahmefällen solche Relationen zwischen punkt- und linienförmigen Signaturen auf der einen Seite und den Flächensignaturen auf der anderen Seite hergestellt werden können. Bei den Straßenkarten liegt die Ursache darin begründet, dass den Flächensignaturen eine untergeordnete Rolle zukommt.

Was die kombinierten Eigenschaften angeht, so erkennt man, dass die Erreichbarkeit bei den Straßenkarten entscheidend ist, während bei den geologischen Karten die Flächendeckung ein wichtiges Merkmal ist. Diese beiden Merkmale lassen sich auf die jeweilige andere Kartenart nicht übertragen. Die Eigenschaft der topologischen Konsistenz bei der geometrischen Vereinfachung hingegen findet wieder bei beiden Kartenarten Verwendung.

Insgesamt ist erkennbar, dass große Teile der invarianten Eigenschaften, die aus der Modellgeneralisierung abgeleitet wurden, auch bei anderen Generalisierungsanwendungen, die teilweise weit von der Modellgeneralisierung entfernt sind, von Bedeutung sind. Es lässt sich daher vermuten, dass auch weitere Generalisierungsanwendungen von den Erkenntnissen, die aus der Modellgeneralisierung gewonnen wurden, profitieren können. Beispiele für weitere Generalisierungsanwendungen sind Seekarten und Bodennutzungskarten. Auch bei der Generalisierung von topographischen Karten könnten die hier formulierten Erkenntnisse Fortschritte in der Forschung ergeben.

Kapitel 7

Zusammenfassung und Ausblick

Die vorliegende Arbeit befasst sich mit der Automation der Generalisierung von Geodaten. Ausgangspunkt ist die Implementation eines Verfahrens zur Modellgeneralisierung von Digitalen Landschaftsmodellen. Die Abstraktion und die Verallgemeinerung dieses Verfahrens ist allgemein wünschenswert, um auch andere Anwendungen von den Ergebnissen profitieren zu lassen.

Diese Arbeit zeigt, dass in bestimmten Bereichen eine Verallgemeinerung des gesamten Verfahrens prinzipiell nicht zu erreichen ist. Denn das Verfahren orientiert sich sehr speziell an den Anforderungen der Modellgeneralisierung und der Modellierung von Digitalen Landschaftsmodellen. Um der Verallgemeinerung der Modellgeneralisierung dennoch näher zu kommen, definiert diese Arbeit das Problem der Generalisierung auf einer abstrakteren Ebene.

Aus diesen Überlegungen ergibt sich eine differenziertere Zielsetzung dieser Arbeit:

Erstens formalisiert diese Arbeit die Anforderungen an die Modellgeneralisierung auf einer abstrakten Ebene und *formalisiert* damit die *Generalisierungsziele*. Diese Formalisierung geht einher mit einer Aufteilung der einzelnen Geodaten in thematische, topologische und geometrische Anteile. Dies führt zu einer klaren Strukturierung der Generalisierungsziele. Außerdem geht die Formalisierung einher mit einer allgemeinen, mathematischen Formulierung der Generalisierung, die es erleichtert, die Generalisierungsziele zu formalisieren.

Zweitens *abstrahiert und verallgemeinert* diese Arbeit *die formalisierten Generalisierungsziele*. Verkörpert wird diese Abstraktion in erster Linie durch die Formulierung der Generalisierung als mathematische Funktion. In Form von Bedingungen werden die abstrahierten Anforderungen auf andere Generalisierungsanwendungen übertragen und angewendet.

Drittens identifiziert und charakterisiert diese Arbeit *invariante Eigenschaften* von Geodaten und Mengen von Geodaten. Die Invarianzen gelten sowohl bei der Durchführung der Modellgeneralisierung als auch bei der Durchführung anderer Generalisierungsanwendungen. Aus der Modellgeneralisierung heraus werden sie identifiziert; auf andere Generalisierungsanwendungen werden sie übertragen. Die Invarianzen werden in mathematische und logische *Bedingungen* gefasst. Dies wird in der Arbeit an einer Reihe von

Beispielen durchgeführt. Die Formulierung der Invarianzen liefert Bedingungen, die als Kriterien für eine Beurteilung des Ergebnisses nach erfolgter Generalisierung dienen. Diese Beurteilung kann automatisch erstellt werden. Sie bewertet die Korrektheit und Güte der ausgeführten Generalisierung.

Mit der Formalisierung der Anforderungen und der Bedingungen ist es außerdem möglich, bereits vor der Generalisierung zu entscheiden, welche Daten gegeben und wie die Anforderungen an die generalisierten Daten formuliert sein müssen, damit eine Generalisierung korrekt und qualitativ hochwertig ausgeführt werden kann.

Im Einzelnen werden die Ziele mit folgenden Mitteln erreicht:

In Kapitel 3 wird das sogenannte *erweiterte hierarchische Datenmodell* entwickelt. Das Modell ist auf eine große Bandbreite unterschiedlicher Arten von Geodaten anwendbar. Das wesentliche Merkmal dieses Datenmodells ist die Dreiteilung der Geodaten in Thematik, Topologie und Geometrie. Diese Dreiteilung ermöglicht eine differenziertere Charakterisierung von Geodaten und ihrer Eigenschaften als es mit der Betrachtung von Geodaten als Einheit möglich wäre. Sie hilft bei der Formalisierung des Generalisierungszieles, indem die Formalisierung in kleinere Teile strukturiert wird und damit besser differenziert werden kann. Ein besonders wichtiger Punkt des Datenmodells ist bei der Erstellung von Bedingungen zu finden: Hier werden durch die Dreiteilung der Geodaten die Bedingungen in die Bereiche der Thematik, Topologie und Geometrie eingeordnet, beziehungsweise in Bereiche aus Kombinationen von zweien dieser Bereiche. Die Bedingungen werden damit klar strukturiert und können getrennt gruppiert und behandelt werden.

Für die Generalisierung ist weiterhin eine Multirepräsentation der Daten notwendig. Hierfür wird das erweiterte hierarchische Datenmodell durch Duplizierung und Selbstverknüpfung weiterentwickelt. Das Ergebnis ist ein Datenmodell *für zwei Auflösungen oder Strukturierungsgrade*, die durch Generalisierung auseinander hervorgegangen sind. Das Modell ist so ausgelegt, dass es sich ohne nennenswerten Aufwand auf mehr als zwei Auflösungen oder Strukturierungsgrade erweitern lässt. Dieses Datenmodell findet seine Entsprechung in der entwickelten Generalisierungsfunktion (siehe nächster Abschnitt) und unterstützt die Abstraktion und Verallgemeinerung der Generalisierung durch Strukturierung.

In Kapitel 4 wird die Generalisierung durch eine *mathematische Funktion* ausgedrückt. Die Funktion bildet aus der ungeneralisierten Menge von Geodaten auf die Menge der generalisierten Geodaten ab. Die Generalisierungsfunktion wird in drei unterschiedlichen Ausprägungen, in der speziellen und in zwei allgemeinen Formen, definiert. Je nach Generalisierungsanwendung oder Anforderung an die Generalisierungsanwendung kann aus den drei Ausprägungen eine geeignete Auswahl getroffen werden. Des Weiteren wird die Generalisierungsfunktion charakterisiert und dieser Charakterisierung entsprechend kartographisch interpretiert.

Die Formulierung der Generalisierung als mathematische Funktion bildet die Grundlage für die Abstraktion der Generalisierung. Außerdem stellt die Generalisierungsfunktion

die Verknüpfung zwischen ungeneralisierten und generalisierten Daten in einen mathematischen Zusammenhang. Hier besteht eine Analogie zum erweiterten hierarchischen Datenmodell für zwei oder mehr Auflösungen oder Strukturierungsgrade, welches die Verknüpfung auf Datenmodellebene herstellt. Die Darstellung der Verknüpfung ist notwendig für die Formalisierung der Generalisierungsziele, weil die Generalisierungsziele zwar für die generalisierten Daten gelten, aber die ungeneralisierten Daten berücksichtigen müssen. Ebenso ist die Darstellung der Verknüpfung notwendig für die Formalisierung der Bedingungen an die Generalisierung. Denn die Bedingungen formulieren Anforderungen an die generalisierten Daten, die sich aus Eigenschaften der ungeneralisierten Daten ergeben.

In Kapitel 5 erhält die in Kapitel 4 entwickelte Generalisierungsfunktion zusätzlich eine Formulierung als *Morphismus*. Streng mathematisch gesprochen geschieht dies durch die Identifizierung von Strukturen auf den Ausgangsdaten und der Feststellung der Invarianz dieser Strukturen bezüglich der Generalisierungsfunktion. Übertragen auf den Bereich der Kartographie bedeutet dies die Identifizierung von Eigenschaften auf den ungeneralisierten Geodaten, die durch eine Generalisierung nicht verändert werden dürfen oder zumindest so wenig wie möglich verändert werden sollten. Diese Eigenschaften bilden *invariante Eigenschaften der Generalisierung*. Die Formulierung der invarianten Eigenschaften vollzieht sich entlang der in Kapitel 3 getroffenen Dreiteilung der Geodaten in Thematik, Topologie und Geometrie.

Beispielsweise werden folgende invariante Eigenschaften in dieser Arbeit herausgearbeitet: Im Bereich der Thematik wird die Zugehörigkeit zu Klassen als Invarianz betrachtet, ebenso wird die Belegung von Attributen mit bestimmten Werten auf Invarianzen untersucht. Im Bereich der Topologie wird die Adjazenz und die Inzidenz als Invariante herausgearbeitet sowie die Inklusionsrelation. Im Bereich der Geometrie werden metrische Kriterien sowie Kriterien der Form und der Gestalt untersucht. Außerdem werden Invarianzen herausgearbeitet, die sich aus der Kombination von zwei Bereichen ergeben: Der Netzzusammenhang beziehungsweise die Erreichbarkeit sowie die Flächendeckung sind Invarianzen, die auf einer Kombination der Thematik mit der Topologie beruhen. Die Invarianz der topologischen Konsistenz bei geometrischer Generalisierung ergibt sich aus Kombination der Topologie mit der Geometrie, mithin aus dem gesamten Raumbezug.

Die Invarianz wird – wo möglich – kategorisch formuliert und durch logische Ausdrücke formalisiert. Diese Ausdrücke bilden Bedingungen, die die automatische Überprüfung einer erfolgten Generalisierung ermöglichen. Sie korrespondieren zum Konzept der bedingungs-basierten Generalisierung, das in der aktuellen Forschung vorangetrieben wird. Damit tragen sie auch zur Formalisierung der Generalisierungsziele bei. Des Weiteren sind die Bedingungen abstrahiert und verallgemeinert formuliert, so dass sie unabhängig von einer konkreten Generalisierungsanwendung wie der der Modellgeneralisierung sind. Sie können damit auf andere Generalisierungsanwendungen angewandt werden, wie es in Kapitel 6 beispielhaft durchgeführt ist.

In Kapitel 6 wird unter Beweis gestellt, dass die in dieser Arbeit erstellten Formalisierungen, Abstraktionen und Verallgemeinerungen *in der Praxis* auf reale Daten *anwendbar* sind. In einem ersten Teil wird dies durch die Implementation der Modellgeneralisierung gezeigt. Dies betrifft die Formalisierung der Geodaten, des Generalisierungsvorgangs und der invarianten Eigenschaften. Die formalisierten Bedingungen werden auf die automatisch durchgeführte Modellgeneralisierung angewendet und überprüft. In einem zweiten Teil wird die Anwendbarkeit der Abstraktion und Verallgemeinerung der Generalisierung gezeigt. An Hand von zwei weiteren Anwendungsbeispielen – der Generalisierung von geologischen Karten und der Generalisierung von Straßenkarten – wird gezeigt, wie die abstrahiert formulierten Bedingungen der Invarianzen auf andere Generalisierungsanwendungen übertragen werden können. Hierbei wird auch deutlich, dass eine Auswahl der Invarianzen getroffen werden muss, da nicht jede Invarianz sich auf alle Generalisierungsanwendungen übertragen lässt.

Zum Beispiel wird deutlich, dass die Invarianzen aus dem Bereich der Thematik sowohl bei der Generalisierung von geologischen Karten als auch von Straßenkarten gelten. Gleiches gilt für die invariante Eigenschaft der topologischen Konsistenz bei geometrischer Vereinfachung. Nicht ohne weiteres übertragen werden kann aber zum Beispiel die Invarianz der Inklusionsrelation auf der topologischen Ebene, weder auf die Generalisierung von geologischen Karten noch auf die Generalisierung von Straßenkarten. Die invariante Eigenschaft der Erreichbarkeit und des Netzzusammenhanges kann auf die Generalisierung von Straßenkarten, nicht jedoch auf die Generalisierung von geologischen Karten übertragen werden. Andererseits kann die invariante Eigenschaft der Flächendeckung auf die Generalisierung von geologischen Karten, nicht jedoch auf die Generalisierung von Straßenkarten übertragen werden.

Insgesamt wird mit dieser Arbeit die Generalisierung zum ersten Mal auf eine theoretische Grundlage im Zusammenhang mit mathematischen Strukturen gestellt. In den Bestrebungen zur Automation der Generalisierung ist in der Forschung eine Tendenz zur bedingungs-basierten Modellierung der Generalisierung zu erkennen. Hierzu leistet diese Arbeit durch die Formalisierung invarianter Eigenschaften einen wesentlichen Beitrag. Durch den Bezug zur Mathematik werden die Kenntnisse über das Wesen der Generalisierung erweitert. Die Verwendung von mathematischen Strukturen zur Beschreibung der Generalisierung schafft ein grundlegendes Gerüst, auf dem die Automation der Generalisierung weiter aufbauen kann. Ergänzend zu den theoretischen Anteilen dieser Arbeit werden die Erkenntnisse durch Anwendung auf reale Daten in ihrer Praktikabilität bestätigt.

Auf dieser Arbeit können in Zukunft sowohl Forscher als auch Entwickler aufbauen. Letzteren bietet sich die Möglichkeit, sich bei der Implementation künftiger Generalisierungsanwendungen auf das in dieser Arbeit entwickelte theoretische Fundament zu stützen. Für Forscher hingegen bietet sich die Möglichkeit, die Menge der herausgearbeiteten invarianten Eigenschaften zu ergänzen und zu erweitern. Insbesondere die Konkretisierung und Ausformulierung von geometrischen Bedingungen könnte die Formalisierung der Generalisierung auch über die Modellgeneralisierung und die Generalisierung von thematischen

Karten hinaus bis hin zur kartographischen Generalisierung vorantreiben. Dies ist eine Vorgehensweise, die zum Beispiel in der aktuellen Forschung bei der Modellierung der Generalisierung als bedingungsbasierter Prozess helfen kann. Werden die Erweiterungen nur für eine bestimmte Generalisierungsanwendung ergänzt und ausformuliert, könnte in dieser Anwendung ein vertiefter Einblick in die Bedingungen der Generalisierung gewonnen werden. So könnte diese Arbeit eine Grundlage für weitere Arbeiten darstellen, die sich nicht mit Modellgeneralisierung, sondern mit anderen speziellen Generalisierungsanwendungen befassen. Zusammen mit der vorliegenden Arbeit ergäbe sich ein umfassenderes Bild der Generalisierung als Ganzes.

Abbildungsverzeichnis

2.1	Notwendigkeit der Generalisierung	21
2.2	Elementare Vorgänge der Generalisierung	24
2.3	Zusammenhänge der elementaren Generalisierungsvorgänge	25
2.4	Generalisierung als Reduktion	28
2.5	Generalisierung als Konstruktion	29
2.6	Sichtbarkeit von Straßen ohne Signaturierung	30
3.1	Wichtige Elemente der Unified Modelling Language	41
3.2	Beispiel für ein Spaghetti-Datenmodell	42
3.3	Topologisches Datenmodell	43
3.4	Modellierung von Kurven durch Polygonzüge	44
3.5	Hierarchisches Datenmodell	45
3.6	Hierarchisches Datenmodell erweitert um die thematische Ebene	48
3.7	Erweitertes hierarchisches Datenmodell	49
3.8	Beispielhafte Modellierung mittels erweitertem hierarchischen Datenmodell	51
3.9	Struktur des ATKIS-Datenmodells	53
3.10	Unterschied zwischen ATKIS-Topologie und graphentheoretischer Topologie	54
3.11	Ebenenzuordnung zwischen ATKIS-Datenmodell und erweitertem hierar- chischen Datenmodell	56
3.12	Ebenenzuordnung zwischen TOPIS-Datenmodell und erweitertem hierar- chischen Datenmodell	57
3.13	Beziehungen zwischen den Begriffen zur mehrfachen Abbildung von Geo- daten	59
3.14	Repräsentation von Landschaftsobjekten in einer Datenbank	60
3.15	Verknüpfung zweier erweiterter hierarchischer Datenmodelle	62
3.16	Unmöglichkeit der Zuordnung von Geometrieelementen	63
4.1	Mögliche Vorgehensweisen der Abstraktion und ihrer Anwendung	67
4.2	Der Begriff des Geodatums und seine Modellierung	68
4.3	Relationen zwischen je zwei Geodaten	70
4.4	Beispiel für eine Generalisierungsrelation	71
4.5	Beispiel für die Aggregation von Flächen bei der Generalisierung von Di- gitalen Landschaftsmodellen	73
4.6	Generalisierung von Häuserreihen	75
4.7	Mögliche Zuordnungen	75
4.8	Beispiel für eine spezielle Generalisierungsfunktion	78

4.9	Zahlenmäßige Verringerung von drei Häusern zu zwei Häusern	81
4.10	Beispiele für die spezielle Generalisierungsfunktion	82
4.11	Beispiele für die erste allgemeine Generalisierungsfunktion	83
4.12	Beispiele für die zweite allgemeine Generalisierungsfunktion	84
4.13	Nicht-Surjektivität bei manueller Generalisierung	90
4.14	Nichtdeterminiertheit der Generalisierungsfunktion (1)	91
4.15	Nichtdeterminiertheit der Generalisierungsfunktion (2)	93
4.16	Nichtdeterminiertheit der Generalisierungsfunktion (3)	94
5.1	Generalisierungsfunktion, Strukturen und Eigenschaften (1)	100
5.2	Generalisierungsfunktion, Strukturen und Eigenschaften (2)	101
5.3	Unterklassenbildung in topographischen Karten	107
5.4	Zuordnung in die Menge der Objektarten	109
5.5	Beispiel für eine inkorrekte Generalisierung	110
5.6	Schematisches Beispiel für eine Aggregation flächenförmiger Objekte . . .	117
5.7	Aggregation zweier flächenförmiger Geodaten mit Aufteilung der Generalisierungsfunktion	119
5.8	Die topologische Ebene des erweiterten hierarchischen Datenmodells . . .	122
5.9	Beispiele für die drei Arten von Inzidenzen	123
5.10	Beispiele für die sechs Arten von Adjazenzen	124
5.11	Erhaltung von Inzidenz und Adjazenz	127
5.12	Beispiel für die Erhaltung der Knoten-Adjazenz über Kanten, sowie für die Reflexivität der Adjazenz und die Nicht-Umkehrbarkeit der Invarianzbedingung	129
5.13	Beispiel für die Erhaltung der Kanten-Adjazenz über Knoten, sowie für die Reflexivität der Adjazenz und die Nicht-Umkehrbarkeit der Invarianzbedingung	129
5.14	Beispiel für die Erhaltung der Maschen-Adjazenz über Kanten, sowie für die Reflexivität der Adjazenz und die Nicht-Umkehrbarkeit der Invarianzbedingung	129
5.15	Verlust der Knotenadjazenz über Kanten durch Geometriotypwechsel . . .	130
5.16	Verlust der Maschenadjazenz durch Geometriotypwechsel	131
5.17	Verlust der Adjazenz beziehungsweise Umwandlung zur Inzidenz durch Geometriotypwechsel	131
5.18	Erhaltung der Inklusionsrelation bei Aggregation von flächenförmigen Objekten	132
5.19	Die geometrische Ebene des erweiterten hierarchischen Datenmodells . . .	134
5.20	Vier Linienvereinfachungs- und -glättungsalgorithmen	137
5.21	Unterschiedliche Generalisierungen einer Linie innerhalb eines Toleranzbandes	138
5.22	Höhenliniengeneralisierung und Toleranzband	139
5.23	Netzzusammenhang und seine Verletzung durch eine Generalisierung . . .	143
5.24	Die Klassen Masche und flächenförmiges Objekt des erweiterten hierarchischen Datenmodells	145

5.25	Mögliche topologische Inkonsistenzen in Folge von geometrischen Vereinfachungen	146
5.26	Grundsätzliche Strategien zur Vermeidung von topologischer Inkonsistenz in Folge von geometrischer Generalisierung	148
6.1	Übergang vom DLM 25 zum DLM 250	152
6.2	Übergang vom DLM 250 zum DLM 1000	153
6.3	Erreichbarkeit einer Ortslage	158
6.4	Geologische Karten in zwei Maßstäben	160
6.5	Automatische Flächenzusammenfassung in geologischen Karten	162
6.6	Straßenkarten in zwei Maßstäben (1)	165
6.7	Straßenkarten in zwei Maßstäben (2)	166
6.8	Legende eines Straßenatlas	166

Abbildungsverzeichnis

Tabellenverzeichnis

2.1	Arten der Generalisierung (1)	27
2.2	Arten der Generalisierung (2)	27
4.1	Die drei Formen der Generalisierungsfunktion im Vergleich	95
5.1	Vorgesehene Attributwerte für das Attribut FKT Funktion	113
5.2	Unterklassenbildung im TOPIS-OK (1)	115
5.3	Unterklassenbildung im TOPIS-OK (2)	115
5.4	Unterklassenbildung bei der Signaturierung von Straßen	116
5.5	Topologische Relationen in der kombinatorischen Topologie	121
6.1	Prozentualer Anteil der Grundflächen (Beispiel)	155
6.2	Ausschnitt aus einer geologischen Zeittafel	161
6.3	Ausschnitt aus der Zuordnungsfunktion zwischen geologischen Karten unterschiedlicher Auflösung	163
6.4	Zwei mögliche Zuordnungsfunktionen für einen Straßenatlas	167
6.5	Überblick über invariante Eigenschaften	168

Tabellenverzeichnis

Literaturverzeichnis

- AdV 1989** ADV-ARBEITSGRUPPE ATKIS: *ATKIS-Gesamtdokumentation*. Hannover : Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (AdV), 1989
- AdV 1998** ADV-ARBEITSGRUPPE ATKIS: *ATKIS-Signaturenkatalog 1 : 25 000*. Hannover : Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (AdV), 1998. – Stand 1. 6. 2002
- AdV 2003** ADV-ARBEITSGRUPPE ATKIS: *ATKIS-Gesamtdokumentation*. Hannover : Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (AdV), 2003. – Stand 1. 7. 2003
- Ai und van Oosterom 2001** AI, Tinghua ; VAN OOSTEROM, Peter: A map generalization model based on algebra mapping transformation. In: AREF, Walid G. (Hrsg.): *Proceedings of the 9th ACM International Symposium on Advances in Geographic Information Systems*, 9.–10. November 2001, S. 21–27
- Aigner 1984** AIGNER, Martin: *Graphentheorie*. Stuttgart : Teubner, 1984
- Anders u. a. 2003** ANDERS, Karl-Heinrich ; BILDIRICI, İbrahim Ö. ; SESTER, Monika: 2003. – URL http://www.ikg.uni-hannover.de/aga/pdf-files-aga03/sester_aga03.pdf. – Vortrag auf der 40. Sitzung der Arbeitsgruppe „Automation in der Kartographie“ am 23. und 24. September 2003 in Erfurt. Letzter Zugriff am 19. August 2005
- Arbib und Manes 1975** ARBIB, Michael A. ; MANES, Ernest G.: *Arrows, Structures, and Functors*. New York, NY : Academic Press, 1975
- Armstrong 1991** ARMSTRONG, Marc P.: Knowledge classification and organization. In: BUTTENFIELD, Barbara P. (Hrsg.) ; MCMASTER, Robert B. (Hrsg.): *Map Generalization*. Longman Scientific & Technical, 1991, S. 86–102
- Autoroute 2004** MICROSOFT: *Autoroute 2004*. – URL <http://www.microsoft.com/germany/autoroute/>. – Letzter Zugriff am 26. November 2005
- Averdung 1996** AVERDUNG, Christoph: Wissensbasierte Planung auf der Basis von ATKIS. In: *Das Geoinformationssystem ATKIS und seine Nutzung in Wirtschaft und Verwaltung*, Landesvermessungsamt Rheinland-Pfalz, 1996, S. 59–67
- Badard und Lemarié 2002** BADARD, Thierry ; LEMARIÉ, Cécile: Associer des données: l'appariement. In: RUAS, Anne (Hrsg.): *Généralisation et représentation multiple*. Paris : Hermes, 2002, S. 163–183

- Balzert 2001** BALZERT, Heide: *UML kompakt*. Heidelberg : Spektrum-Verlag, 2001
- Bard 2004** BARD, Sylvain: Quality assessment of cartographic generalisation. In: *Transactions in GIS* 8 (2004), Nr. 1, S. 63–81
- Barr und Wells 1990** BARR, Michael ; WELLS, Charles: *Category Theory for Computing Science*. Prentice Hall, 1990
- Barrault u. a. 2001** BARRAULT, Mathieu ; REGNAULD, Nicolas ; DUCHÊNE, Cécile ; HAIRE, Kelvin ; BAEIJS, Christof ; DEMAZEAU, Yves ; HARDY, Paul ; MACKANESS, William A. ; RUAS, Anne ; WEIBEL, Robert: Integrating multi-agent, object-oriented, and algorithmic techniques for improved automated map generalization. In: *Proceedings of the 20th International Cartographic Conference*. Beijing, 6.–10. August 2001, S. 2210–2216
- Bartelme 2005** BARTELME, Norbert: *Geoinformatik*. 4., vollständig überarbeitete Auflage. Berlin : Springer, 2005
- Bauer und Goos 1992** BAUER, Friedrich L. ; GOOS, Gerhard: *Informatik 2*. 4. Auflage. Berlin : Springer, 1992
- Bédard u. a. 2002** BÉDARD, Yvan ; PROULX, Marie-Josée ; LARRIVÉE, Suzie ; BERNIER, Eveline: Modeling multiple representations into spatial data warehouses: A UML-based approach. In: *Joint International Symposium on Geospatial Theory, Processing and Applications*. Ottawa, 8.–12. Juli 2002
- Beard 1991** BEARD, M. Kate: A rule for describing line feature geometry. In: BUTTENFIELD, Barbara P. (Hrsg.) ; MCMASTER, Robert B. (Hrsg.): *Map Generalization*. Longman Scientific & Technical, 1991, S. 150–171
- Bederke und Wunderlich 1968** BEDERKE, Erich (Hrsg.) ; WUNDERLICH, Hans-Georg (Hrsg.): *Meyers Großer Physischer Weltatlas*. Bd. 2: *Atlas zur Geologie*. Bibliographisches Institut Mannheim / Kartographisches Institut Meyer, 1968
- Bender 1996** BENDER, Gaby: *Objektorientierte Modellierung des Amtlich Topographisch-Kartographischen Informationssystems (ATKIS) und Überführung ausgewählter Bereiche in ein relationales Datenbanksystem*. Bonn, Institut für Informatik III, Universität Bonn, Diplomarbeit, 1996
- BGR 2000** BUNDESANSTALT FÜR GEOWISSENSCHAFTEN UND ROHSTOFFE: *Geologische Karte der Bundesrepublik Deutschland 1:1 000 000*. September 2000. – URL <http://www.bgr.de/index.html?karten/gk1000/gk1000.htm>. – Letzter Zugriff am 19. August 2005
- Birth 2003** BIRTH, Konrad: ATKIS-Projekt Modell- und kartographische Generalisierung, und die Entwicklung geht weiter. In: *Kartographische Nachrichten* (2003), Nr. 3, S. 119–126

- Bobzien 2000** BOBZIEN, Matthias: Implementationsaspekte der Modellgeneralisierung. In: *Mitteilungen des Bundesamtes für Kartographie und Geodäsie*. Frankfurt am Main : Bundesamt für Kartographie und Geodäsie, 2000, S. 15–25
- Bobzien 2001** BOBZIEN, Matthias: Flächenzusammenfassung in der Modellgeneralisierung. In: *Mitteilungen des Bundesamtes für Kartographie und Geodäsie*. Frankfurt am Main : Bundesamt für Kartographie und Geodäsie, 2001, S. 19–29
- Bobzien 2003** BOBZIEN, Matthias: Verallgemeinerung der Modellgeneralisierung. In: *Mitteilungen des Bundesamtes für Kartographie und Geodäsie*. Frankfurt am Main : Bundesamt für Kartographie und Geodäsie, 2003, S. 17–26
- Bobzien und Morgenstern 2002** BOBZIEN, Matthias ; MORGENSTERN, Dieter: *Geometry-Type Change in Model Generalization – A Geometrical or a Topological Problem?* 2002. – URL http://www.ikg.uni-hannover.de/isprs/workshop/abstract_bobzien.pdf. – Unveröffentlichtes Arbeitspapier für den Joint ISPRS/ICA Workshop on Multi-Scale Representations of Spatial Data, 7.–8. Juli 2002. Letzter Zugriff am 19. August 2005
- Bobzien und Morgenstern 2003** BOBZIEN, Matthias ; MORGENSTERN, Dieter: Abstracting model generalization — A comprehensive approach towards automated generalization. In: *Proceedings of the 21st International Cartographic Conference*. Durban, 10.–16. August 2003, S. 136–144
- Bollmann und Koch 2001** BOLLMANN, Jürgen (Hrsg.) ; KOCH, Wolf G. (Hrsg.): *Lexikon der Kartographie und Geomatik*. Heidelberg, Berlin : Spektrum Akademischer Verlag, 2001
- Booch u. a. 1999** BOOCH, Grady ; RUMBAUGH, James ; JACOBSON, Ivar: *Unified Modelling Language User Guide*. Reading, MA : Addison-Wesley, 1999
- Brassel 1983** BRASSEL, Kurt E.: Grundkonzepte und technische Aspekte von Geographischen Informationssystemen. In: *Internationales Jahrbuch für Kartographie XXIII* (1983), S. 31–52
- Brassel 1990** BRASSEL, Kurt E.: Der Generalisierungsbegriff in der Kartographie und anderen Disziplinen. In: *Kartographisches Generalisieren*. Schweizerische Gesellschaft für Kartographie, 1990 (Kartographische Publikationsreihe 10), S. 5–16
- Brassel und Weibel 1988** BRASSEL, Kurt E. ; WEIBEL, Robert: A review and conceptual framework of automated map generalization. In: *International Journal of Geographical Information Systems* 2 (1988), Nr. 3, S. 229–244
- Burghardt und Meier 1997** BURGHARDT, Dirk ; MEIER, Siegfried: Cartographic displacement using the snakes concept. In: FÖRSTNER, Wolfgang (Hrsg.) ; PLÜMER, Lutz (Hrsg.): *Semantic Modeling for the Acquisition of Topographic Information from Images and Maps*. Basel : Birkhäuser Verlag, 1997, S. 59–71

- Burghardt u. a. 2005** BURGHARDT, Dirk ; NEUN, Moritz ; WEIBEL, Robert: Generalization services on the web — A classification and an initial prototype implementation. In: *Proceedings of the American Congress on Surveying and Mapping, Auto-Carto*. Las Vegas, NV, 2005
- Buttenfield und McMaster 1991** BUTTENFIELD, Barbara P. (Hrsg.) ; MCMMASTER, Robert B. (Hrsg.): *Map Generalization*. Longman Scientific & Technical, 1991
- Christensen 1999** CHRISTENSEN, Albert H. J.: Cartographic line generalization with waterlines and medial-axes. In: *Cartography and Geographic Information Science* 26 (1999), Nr. 1, S. 19–32
- Claus und Schwill 2001** CLAUS, Volker ; SCHWILL, Andreas: *Duden, Informatik*. 3. Auflage. Mannheim : Dudenverlag, 2001
- Corbett 1975** CORBETT, James P.: Topological principles in cartography. In: *Proceedings of the International Symposium on Computer Assisted Cartography (Auto-Carto II)*, 1975, S. 61–65
- de Berg u. a. 1998** DE BERG, Mark ; VAN KREVELD, Marc ; SCHIRRA, Stefan: Topologically correct subdivision simplification using the bandwidth criterion. In: *Cartography and Geographic Information Science* 25 (1998), Nr. 4, S. 243–257
- Digest 2000** *DIGEST Edition 2.1*. September 2000. – URL <http://www.digest.org/>. – Letzter Zugriff am 19. August 2005
- Dijkstra 1959** DIJKSTRA, Edsger W.: A note on two problems in connexion with graphs. In: *Numerische Mathematik* 1 (1959), S. 269–271
- Douglas und Peucker 1973** DOUGLAS, David H. ; PEUCKER, Thomas K.: Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. In: *The Canadian Cartographer* 10 (1973), Nr. 2, S. 112–122
- Downs und Mackaness 2002** DOWNS, Timothy C. ; MACKANESS, William A.: An integrated approach to the generalization of geological maps. In: *The Cartographic Journal* 39 (2002), December, Nr. 2, S. 137–152
- Edwardes u. a. 2003** EDWARDES, Alistair ; BURGHARDT, Dirk ; BOBZIEN, Matthias ; HARRIE, Lars ; LEHTO, Lassi ; REICHENBACHER, Tumasch ; SESTER, Monika ; WEIBEL, Robert: Map generalisation technology: Addressing the need for a common research platform. In: *Proceedings of the 21st International Cartographic Conference*. Durban, 10.–16. August 2003, S. 170–179
- Edwardes und Mackaness 2006** EDWARDES, Alistair ; MACKANESS, William A.: Road Network Simplification in Urban Areas. Wird erscheinen in: *Computers, Environment and Urban Systems (CEUS)* (2006)

- Egenhofer 1991** EGENHOFER, Max J.: Reasoning about binary topological relations. In: GÜNTHER, Oliver (Hrsg.) ; SCHEK, Hans-Jörg (Hrsg.): *Advances in Spatial Databases, Proceedings of the 2nd International Symposium On Large Spatial Databases* Bd. 525, Lecture Notes in Computer Science. Zürich : Springer, 28.–30. August 1991, S. 143–160
- Egenhofer und Franzosa 1991** EGENHOFER, Max J. ; FRANZOSA, Robert D.: Point-set topological spatial relations. In: *International Journal of Geographical Information Systems* 5 (1991), Nr. 2, S. 161–174
- Egenhofer und Herring 1990a** EGENHOFER, Max J. ; HERRING, John R.: A mathematical framework for the definition of topological relationships. In: *Proceedings of the 4th International Symposium on Spatial Data Handling*. Zürich, 1990, S. 803–813
- Egenhofer und Herring 1990b** EGENHOFER, Max J. ; HERRING, John R.: Categorizing Binary Topological Relationships Between Regions, Lines, and Points in Geographic Databases / Department of Surveying Engineering, University of Maine, Orono. 1990. – Forschungsbericht
- Ehrlholzer 1996** EHRLIHOZLER, Regula: *Methoden für die Bewertung der Qualität von Generalisierungslösungen*. Zürich, Geographisches Institut der Universität Zürich, Diplomarbeit, 1996
- Eppstein und Erickson 1998** EPPSTEIN, David ; ERICKSON, Jeff: Raising roofs, crashing cycles, and playing pool: Applications of a data structure for finding pairwise interactions. In: *Symposium on Computational Geometry*, 1998, S. 58–67
- ESRI 1998** ESRI: ESRI Shapefile Technical Description / ESRI Inc. URL http://downloads.esri.com/support/whitepapers/ao_/shapefile.pdf, 1. Juli 1998. – An ESRI White Paper. Letzter Zugriff am 19. August 2005
- Even 1979** EVEN, Shimon: *Graph Algorithms*. London : Pitman, 1979
- Findeisen 1990** FINDEISEN, Dirk: *Datenstruktur und Abfragesprachen für raumbezogene Informationen*, Institut für Kartographie und Geoinformation, Universität Bonn, Dissertation, 1990
- Floyd 1962** FLOYD, Robert W.: Algorithm 97: Shortest path. In: *Communications of the ACM* 5 (1962), S. 345
- Galanda 2003** GALANDA, Martin: *Automated Polygon Generalization in a Multi Agent System*, Universität Zürich, Institut für Geographie, Dissertation, 2003
- Gröger 2000** GRÖGER, Gerd: *Modellierung raumbezogener Objekte und Datenintegrität in GIS*. Heidelberg : Herbert Wichmann Verlag, 2000
- Güting und Schneider 1993** GÜTING, Ralf H. ; SCHNEIDER, Markus: Realms: A foundation for spatial data types in database systems. In: ABEL, David (Hrsg.) ; OOI, Beng C. (Hrsg.): *Advances in Spatial Databases, Proceedings of the 3rd International Symposium SSD '93* Bd. 692. Singapur : Springer, June 1993, S. 14–35

- Hake 1975** HAKE, Günter: Zum Begriffssystem der Generalisierung. In: *Nachrichten aus dem Karten- und Vermessungswesen* Sonderheft (1975), S. 53–62
- Hake 1982** HAKE, Günter: *Kartographie*. Bd. I. 6., neubearbeitete Auflage. Berlin : Walter de Gruyter, 1982
- Hake und Grünreich 1994** HAKE, Günter ; GRÜNREICH, Dietmar: *Kartographie*. 7. Auflage. Berlin : De Gruyter, 1994
- Hake u. a. 2002** HAKE, Günter ; GRÜNREICH, Dietmar ; MENG, Liqiu: *Kartographie*. 8. Auflage. Berlin : De Gruyter, 2002
- Hallwag International 2003** HALLWAG INTERNATIONAL (Hrsg.): *Europa und Schweiz Strassenatlas 1:800 000*. Hallwag Kümmerly+Frey, 2003
- Halmos 1976** HALMOS, Paul R.: *Naive Mengenlehre*. 4. Auflage. Göttingen : Vandenhoeck & Ruprecht, 1976
- Hampe u. a. 2003** HAMPE, Mark ; ANDERS, Karl-Heinrich ; SESTER, Monika: MRDB applications for data revision and real-time generalisation. In: *Proceedings of the 21st International Cartographic Conference*. Durban, 10.–16. August 2003, S. 192–201
- Harary 1972** HARARY, Frank: *Graph Theory*. Reading, MA : Addison-Wesley, 1972
- Harbeck 1987** HARBECK, Rolf: Das AdV-Vorhaben „Amtliches Topographisch-Kartographisches Informationssystem (ATKIS)“ – Inhaltliche Konzeption. In: *Nachrichten aus dem Karten- und Vermessungswesen* I (1987), Nr. 99, S. 7–14
- Harbeck 2000** HARBECK, Rolf: Eine geographische Basis für Europa – Utopie, Vision, Wirklichkeit? In: *Kartographische Nachrichten* (2000), Nr. 3, S. 103–112
- Hardy 2000** HARDY, Paul: Multi-scale database generalisation for topographic mapping, hydrography and web-mapping, using active object techniques. In: *Proceedings of the 19th ISPRS Congress*. Amsterdam, July 2000
- Harrie 2001** HARRIE, Lars: *An Optimisation Approach to Cartographic Generalisation*, Department of Technology and Society, Lund Institute of Technology, Lund University, Sweden, Dissertation, 2001
- Harrie und Weibel 2006** HARRIE, Lars ; WEIBEL, Robert: Modelling the Overall Process of Generalisation. Wird erscheinen in: MACKANESS, William A. (Hrsg.) ; RUAS, Anne (Hrsg.) ; SARJAKOSKI, L. Tiina (Hrsg.): *Generalisation of Geographic Information: Models and Applications*. Elsevier Science, 2006, Kap. 4
- Heisser 1996** HEISSER, Michael: Modellierung der Generalisierung für die Anwendung in digitalen Systemen. In: *Nachrichten aus dem Karten- und Vermessungswesen* I (1996), Nr. 115, S. 69–78

- Heisser u. a. 1995** HEISSER, Michael ; VICKUS, Georg ; SCHOPPMAYER, Johannes: Rule-oriented definition of the small area ‘selection’ and ‘combination’ steps of the generalization procedure. In: MÜLLER, Jean-Claude (Hrsg.) ; LAGRANGE, Jean-Philippe (Hrsg.) ; WEIBEL, Robert (Hrsg.): *GIS and Generalization – Methodology and Practice*. London : Taylor & Francis, 1995, S. 148–160
- Heissler 1962** HEISSLER, Viktor: *Kartographie*. Berlin : Walter de Gruyter & Co, 1962
- Heissler und Hake 1970** HEISSLER, Viktor ; HAKE, Günter: *Kartographie*. Bd. I – Kartenaufnahme, Netzentwürfe, Gestaltungsmerkmale, topographische Karten. 4. Auflage. Berlin : Walter de Gruyter & Co, 1970
- Herring 1991** HERRING, John R.: The mathematical modeling of spatial and non-spatial information in Geographic Information Systems. In: MARK, David M. (Hrsg.) ; FRANK, Andrew U. (Hrsg.): *Cognitive and Linguistic Aspects of Geographic Space* Bd. 63. Dordrecht : Kluwer Academic Publishers, 1991, S. 313–350
- ICA-ISPRS 2002** *Multi-Scale Representations of Spatial Data*. Joint ISPRS/ICA workshop. 7.–8. Juli 2002. – URL <http://www.ikg.uni-hannover.de/isprs/wg43-workshop.html>. – Letzter Zugriff am 19. August 2005
- IGN 2004** INSTITUT GÉOGRAPHIQUE NATIONAL (Hrsg.): *BD TOPO La composante topographique 3D du Référentiel à Grande Échelle (RGE)*. Mai 2004. – URL <http://www.ign.fr/telechargement/MPro/plaquettes/bdtopo.pdf>. – Letzter Zugriff am 19. August 2005
- Illert 1995** ILLERT, Andreas: Aspekte der Zusammenführung digitaler Datensätze unterschiedlicher Quellen. In: *Nachrichten aus dem Karten- und Vermessungswesen I* (1995), Nr. 113, S. 105–115
- Jakobitz 1997** JAKOBITZ, Uwe: *Untersuchungen und Programmkonzeption für die Behandlung flächenhafter Objekte bei der Modellgeneralisierung für den Übergang vom DLM 25 zum DLM 250*, Institut für Kartographie und Topographie, Universität Bonn, Diplomarbeit, 1997
- Jansen und van Kreveld 1998** JANSEN, Michiel ; KREVELD, Marc van: Evaluating the consistency of cartographic generalization. In: POIKER, Thomas K. (Hrsg.) ; CHRISMAN, Nicholas (Hrsg.): *Proceedings of the 8th International Symposium on Spatial Data Handling*, 1998, S. 668–678
- Jäger 2003** JÄGER, Ernst: ATKIS als Gemeinschaftsaufgabe der Länder und des Bundes. In: *Kartographische Nachrichten* (2003), Nr. 3, S. 113–119
- Jones 1991** JONES, Christopher B.: Database architecture for multi-scale GIS. In: *Proceedings Auto-Carto*. Baltimore, 1991, S. 1–14

- Jones u. a. 1995** JONES, Christopher B. ; BUNDY, Geraint L. ; WARE, J. Mark: Map generalization with a triangulated data structure. In: *Cartography and Geographic Information Systems* 22 (1995), Nr. 4, S. 317–331
- Kilpeläinen 2000** KILPELÄINEN, Tiina: Maintenance of multiple representation databases for topographic data. In: *The Cartographic Journal* 37 (2000), December, Nr. 2, S. 101–107
- Knuth 1973** KNUTH, Donald E.: *The Art of Computer Programming*. Bd. 3: *Sorting and Searching*. 2nd edition. Reading, MA : Addison-Wesley, 1973
- Knuth 1991** KNUTH, Donald E.: *Computers & Typesetting*. Bd. A: *The T_EX-Book*. 11th printing. Reading, MA : Addison-Wesley, 1991
- Knuth 1997** KNUTH, Donald E.: *The Art of Computer Programming*. Bd. 1: *Fundamental Algorithms*. 3rd edition. Reading, MA : Addison-Wesley, 1997
- Kreiter 2002a** KREITER, Novit: *Multiple representation database as basis for topographic maps*. 2002. – URL <http://www.ikg.uni-hannover.de/isprs/workshop/Abstract%20Kreiter%20ttawa02.pdf>. – Unveröffentlichtes working paper für den *Joint ISPRS/ICA Workshop Multi-Scale Representations of Spatial Data*, 7.–8. Juli 2002. Letzter Zugriff am 19. August 2005
- Kreiter 2002b** KREITER, Novit: *Multirepräsentationsdatenbank als Basis von topografischen Landeskarten*, Eidgenössische Technische Hochschule Zürich, Diplomarbeit, 2002
- Laubert u. a. 1982** LAUBERT, Hans ; WOSKA, Erwin ; HABEL, Rudolf: *Lehrbuch für Kartographiefacharbeiter*. Bd. 2. VEB Hermann Haack, Geographisch-Kartographische Anstalt Gotha, 1982
- Lee 2003** LEE, Dan: *Recent Generalization Development and Road Ahead*. 2003. – URL http://www.geo.unizh.ch/ICA/docs/paris2003/papers/lee_v0.pdf. – Unveröffentlichtes Arbeitspapier für den Fifth ICA Workshop on Progress in Automated Map Generalization, 28.–30. April 2003. Letzter Zugriff am 19. August 2005
- Lee 1998** LEE, Feng-Tyan: Many sorted algebraic data models for GIS. In: *International Journal of Geographical Information Science* 12 (1998), Nr. 8, S. 765–788
- Mac Lane 1998** MAC LANE, Saunders: *Categories for the Working Mathematician*. 2nd edition. Berlin, Heidelberg, New York : Springer, 1998
- Mark 1991** MARK, David M.: Object modelling and phenomenon-based generalization. In: BUTTENFIELD, Barbara P. (Hrsg.) ; MCMMASTER, Robert B. (Hrsg.): *Map Generalization*. Longman Scientific & Technical, 1991, S. 103–118
- McMaster und Shea 1992** MCMMASTER, Robert B. ; SHEA, K. S.: *Generalization in Digital Cartography*. Washington D.C. : Association of American Geographers, 1992

- McMaster 1986** MCMMASTER, Robert B.: A statistical analysis of mathematical measures for linear simplification. In: *The American Cartographer* 13 (1986), Nr. 2, S. 103–116
- McMaster 1987** MCMMASTER, Robert B.: Automated line generalization. In: *Cartographica* 24 (1987), Nr. 2, S. 74–111
- MilGeo-OK 1993** DEUTSCHER MILITÄRGEOGRAPHISCHER DIENST (Hrsg.): *MilGeo-Objektartenkatalog*. Euskirchen : Amt für Militärisches Geowesen, April 1993
- Müller u. a. 1995** MÜLLER, Jean-Claude ; WEIBEL, Robert ; LAGRANGE, Jean-Philippe ; SALGÉ, François: Generalization: State of the art and issues. In: MÜLLER, Jean-Claude (Hrsg.) ; LAGRANGE, Jean-Philippe (Hrsg.) ; WEIBEL, Robert (Hrsg.): *GIS and Generalization – Methodology and Practice*. London : Taylor & Francis, 1995, S. 3–17
- Moellering 2003** MOELLERING, H.: Possibilities for a general theory of spatial data. In: *Proceedings of the 21st International Cartographic Conference*. Durban, 10.–16. August 2003, S. 1895–1901
- Morgenstern und Schürer 1999** MORGENSTERN, Dieter ; SCHÜRER, Dietrich: A concept for model generalization of digital landscape models from finer to coarser resolution. In: *Proceedings of the 19th International Cartographic Conference*. Ottawa, 14.–21. August 1999, S. 1021–1028
- Musterblatt TK 25, 1981** LANDESVERMESSUNGSAMT NORDRHEIN-WESTFALEN (Hrsg.): *Musterblatt für die Topographische Karte 1 : 25 000*. 2. Ausgabe. 1981
- Nyerges 1991** NYERGES, Timothy L.: Representing geographical meaning. In: BUTTENFIELD, Barbara P. (Hrsg.) ; MCMMASTER, Robert B. (Hrsg.): *Map Generalization*. Longman Scientific & Technical, 1991, S. 59–85
- Open GIS Consortium 1999** OPEN GIS CONSORTIUM (Hrsg.): *OpenGIS Simple Features Specification for SQL*. Revision 1.1. Open GIS Consortium, Inc., May 1999. – OpenGIS Project Document 99-049
- Perkal 1966a** PERKAL, Julian: An attempt at objective generalization. In: NYSTUEN, John D. (Hrsg.): *Discussion Paper Number 10*. Ann Arbor, MI : University of Michigan, Department of Geography, 1966. – übersetzt von W. Jackowski von: Proba obiektywnej generalizacji. *Geodézia és Kartográfia*, Tom VII, Zeszyt 2(1958), S. 130–142
- Perkal 1966b** PERKAL, Julian: On the length of empirical curves. In: NYSTUEN, John D. (Hrsg.): *Discussion Paper Number 10*. Ann Arbor, MI : University of Michigan, Department of Geography, 1966. – übersetzt von W. Jackowski von: O Długości Krzywych Empirycznych. *Zastosowania Matematyki III*, 3–4 (1958), S. 257–186
- Peucker und Chrisman 1975** PEUCKER, Thomas K. ; CHRISMAN, Nicholas: Cartographic data structures. In: *The American Cartographer* 2 (1975), Nr. 1, S. 55–69

- Plümer 1996** PLÜMER, Lutz: Achieving integrity of geometry and topology in Geographical Information Systems. In: *Proceedings of the „Samos“ International Conference on Geographic Information Systems in Urban, Environmental and Regional Planning*, 19.–21. April 1996
- Plümer und Gröger 1996** PLÜMER, Lutz ; GRÖGER, Gerhard: Nested maps — A formal, provably correct object model for spatial aggregates. In: SHEKHAR, Shashi (Hrsg.) ; BERGOUGNOUX, Patrick (Hrsg.): *Proceedings of the 4th ACM Workshop on Advances in Geographic Information Systems*. Rockville, MD : ACM, November 1996, S. 77–84
- Plümer und Gröger 1997** PLÜMER, Lutz ; GRÖGER, Gerhard: Achieving integrity in Geographic Information Systems — Maps and nested maps. In: *GeoInformatica* 1 (1997), Nr. 4, S. 345–367
- Reinhardt und Soeder 1974** REINHARDT, Fritz ; SOEDER, Heinrich: *Dtv-Atlas zur Mathematik – Tafeln und Texte*. Bd. I: *Grundlagen, Algebra und Geometrie*. Deutscher Taschenbuch Verlag, 1974
- Rieger und Coulson 1993** RIEGER, Monika K. ; COULSON, Michael R. C.: Consensus or confusion: Cartographers' knowledge of generalization. In: *Cartographica* 30 (1993), Nr. 2 & 3, S. 69–80
- Robinson u. a. 1984** ROBINSON, Arthur H. ; SALE, Randall D. ; MORRISON, Joel L. ; MUEHRCKE, Phillip C.: *Elements of Cartography*. New York, NY : John Wiley & Sons, 1984
- Ruas 2002** RUAS, Anne (Hrsg.): *Généralisation et représentation multiple*. Paris : Hermes, 2002
- Ruas 2003** RUAS, Anne: *Minutes of the 2003 Business Meeting*. ICA Commission on Generalisation and Multiple Representation. September 2003. – URL <http://ica.ign.fr/Business/Business2003.pdf>. – Letzter Zugriff am 19. August 2005
- Rumbaugh u. a. 1999** RUMBAUGH, James ; JACOBSON, Ivar ; BOOCH, Grady: *The Unified Modeling Language Reference Manual*. Reading, MA : Addison-Wesley, 1999
- Saalfeld 1999** SAALFELD, Alan: Topologically consistent line simplification with the Douglas-Peucker algorithm. In: *Cartography and Geographic Information Science* 26 (1999), Nr. 1, S. 7–18
- Schneider 1997** SCHNEIDER, Markus: *Lecture Notes in Computer Science*. Bd. 1288: *Spatial Data Types for Database Systems: Finite Resolution Geometry for Geographic Information Systems*. Berlin : Springer, 1997
- Schoppmeyer und Beines 1991** SCHOPPMAYER, Johannes ; BEINES, Marion: Literaturdokumentation und -analyse im Rahmen der ATKIS-Generalisierung. In: *Nachrichten aus dem Karten- und Vermessungswesen* Reihe I (1991), Nr. 106, S. 95–104

- Schoppmeyer und Heisser 1995** SCHOPPMAYER, Johannes ; HEISSER, Michael: Behandlung von Geometriotypwechseln in GIS. In: *Nachrichten aus dem Karten- und Vermessungswesen* I (1995), Nr. 113, S. 209–224
- Schürer 1992** SCHÜRER, Dietrich: *Untersuchungen und Programmkonzeption zur Behandlung von Geometriotypwechseln zwischen den Objektabbildungen des ATKIS-DLM und -DKM*, Institut für Kartographie und Topographie, Universität Bonn, Diplomarbeit, 1992
- Schürer 2002** SCHÜRER, Dietrich: *Ableitung von digitalen Landschaftsmodellen mit geringerem Strukturierungsgrad durch Modellgeneralisierung*, Institut für Kartographie und Geoinformation, Universität Bonn, Dissertation, 2002
- Schuurman 1999** SCHUURMAN, Nadine: Critical GIS: Theorizing an emerging science. In: *Cartographica* 36 (1999), Nr. 4. – Chapter 4
- Sester u. a. 2004** SESTER, Monika ; SARJAKOSKI, Tapani ; HARRIE, Lars ; HAMPE, Mark ; KOIVULA, Tommi ; SARJAKOSKI, Tiina ; LEHTO, Lassi ; ELIAS, Birgit ; NIVALA, Annu-Maaria ; STIGMAR, Hanna: Real-Time Generalisation and Multiple Representation in the GiMoDig Mobile Service / GiMoDig-project. URL <http://gimodig.fgi.fi/deliverables.php>, 2004. – Deliverables D7.1.1*, D7.2.1* and D7.3.1, Public EC report. Letzter Zugriff am 19. August 2005
- SGK 1980** SCHWEIZERISCHE GESELLSCHAFT FÜR KARTOGRAPHIE (Hrsg.): *Kartographische Generalisierung*. 2. Auflage. Schweizerische Gesellschaft für Kartographie, 1980 (Kartographische Publikationsreihe 1)
- SGK 2002** SCHWEIZERISCHE GESELLSCHAFT FÜR KARTOGRAPHIE (Hrsg.): *Topografische Karten – Kartengrafik und Generalisierung*. Schweizerische Gesellschaft für Kartographie, 2002 (Kartographische Publikationsreihe 16)
- Shea 1991** SHEA, K. Stuart: Design considerations for an artificially intelligent system. In: BUTTENFIELD, Barbara P. (Hrsg.) ; MCMASTER, Robert B. (Hrsg.): *Map Generalization*. Longman Scientific & Technical, 1991, S. 3–20
- Skopeliti und Tsoulos 1999** SKOPELITI, Andriani ; TSOULOS, Lysandros: On the parametric description of the shape of the cartographic line. In: *Cartographica* 36 (1999), Nr. 3, S. 53–65
- Smith 2003** SMITH, Vince: *Generalization for Medium-Scale Mapping: Results and Statistics from One Production Implementation*. 2003. – URL <http://www.geo.unizh.ch/ICA/docs/paris2003/presentations/smith.pdf>. – Präsentation auf dem 5th ICA Workshop on Progress in Automated Map Generalization, 28.–30. April 2003. Letzter Zugriff am 19. August 2005
- Spaccapietra u. a. 2000** SPACCAPIETRA, Stefano ; PARENT, Christine ; VANGENOT, Christelle: GIS databases: From multiscale to multirepresentation. In: CHOEIRY, Berthe Y. (Hrsg.) ; WALCH, Toby (Hrsg.): *Abstraction, Reformulation, and Approximation* Bd. 1864. Springer, 2000, S. 57–70

- Stenger 2003** STENGER, Jens: *Entwicklung und Test eines Linienvereinfachungsalgorithmus für den Modellübergang von mittlerer Auflösung zu grober Auflösung unter besonderer Berücksichtigung der Erkennung und Bearbeitung von Eigenüberlappung*. Bonn, Rheinische Friedrich-Wilhelms-Universität Bonn, Institut für Kartographie und Geoinformation, Diplomarbeit, Juli 2003
- SwissTopo 2004** BUNDESAMT FÜR LANDESTOPOGRAFIE DER SCHWEIZ: *Vector 25 – Das digitale Landschaftsmodell der Schweiz*. April 2004. – URL http://www.swisstopo.ch/pub/down/products/digital/landscape/vector25_release2005_de.pdf. – Letzter Zugriff am 19. August 2005
- Timpf und Frank 1995** TIMPF, Sabine ; FRANK, Andrew U.: A multi-scale data structure for cartographic objects. In: *Proceedings of the 17th International Cartographic Conference*. Barcelona, 3.–9. September 1995, S. 1389–1396
- Ullman 1988** ULLMAN, Jeffrey D.: *Principles of Database and Knowledge-Base Systems*. Bd. 1: *Classical Database Systems*. Computer Science Press, 1988
- Vangenot 1998** VANGENOT, Christelle: Multiresolution representations. Concepts for the description of multiple representation databases. In: *International Journal of GIS and Spatial Analysis* 8 (1998), Nr. 1–2, S. 121–148
- Veregin 1999** VEREGIN, Howard: Line simplification, geometric distortion, and positional error. In: *Cartographica* 36 (1999), Nr. 1, S. 25–39
- Veregin 2000** VEREGIN, Howard: Quantifying positional error induced by line simplification. In: *International Journal of Geographical Information Science* 14 (2000), Nr. 2, S. 113–130
- Visvalingam und Whyatt 1993** VISVALINGAM, Maheswari ; WHYATT, J. Duncan: Line generalisation by repeated elimination of points. In: *The Cartographic Journal* 30 (1993), Nr. 1, S. 46–51
- Vosselmann 1992** VOSSELMANN, George: *Lecture Notes in Computer Science*. Bd. 628: *Relational Matching*. Springer, 1992
- Vossmerbäumer 1976** VOSSMERBÄUMER, Herbert: *Allgemeine Geologie – Ein Kompendium*. Stuttgart : E. Schweizerbart'sche Verlagsbuchhandlung (Nägele u. Obermiller), 1976
- Vossmerbäumer 1983** VOSSMERBÄUMER, Herbert: *Geologische Karten*. Stuttgart : E. Schweizerbart'sche Verlagsbuchhandlung (Nägele u. Obermiller), 1983
- Wagenbreth 1958** WAGENBRETH, Otfried: *Geologisches Kartenlesen und Profilzeichnen*. Leipzig : B. G. Teubner Verlagsgesellschaft, 1958
- Wagner 1970** WAGNER, Klaus: *Graphentheorie*. Mannheim : Bibliographisches Institut, 1970

- Wahrig 1994** WAHRIG, Gerhard: *Deutsches Wörterbuch*. Gütersloh : Bertelsmann Lexikon Verlag, 1994
- Walter und Fritsch 1998** WALTER, Volker ; FRITSCH, Dieter: Integration von ATKIS- und GDF-Daten mit Hilfe eines relationalen Zuordnungsansatzes. In: *Zeitschrift für Vermessungswesen* (1998), Nr. 1, S. 2–10
- Ware und Jones 1998** WARE, J. Mark ; JONES, Christopher B.: Conflict reduction in map generalization using iterative improvement. In: *GeoInformatica* 2 (1998), Nr. 4, S. 383–407
- Warmer und Kleppe 2003** WARMER, Jos ; KLEPPE, Anneke: *The Object Constraint Language – Getting Your Models Ready For MDA*. 2. Auflage. Addison-Wesley, 2003
- Weber 1978** WEBER, Wigand: Liniengeneralisierung und Datenreduktion unter dem Gesichtswinkel der mathematischen Optimierung. In: *Nachrichten aus dem Karten- und Vermessungswesen I* (1978), Nr. 74, S. 55–66
- Weber 1991** WEBER, Wigand: Geographische Datenmodelle – Ein Überblick. In: *Nachrichten aus dem Karten- und Vermessungswesen I* (1991), Nr. 106, S. 123–144
- Weibel 1996** WEIBEL, Robert: A typology of constraints to line simplification. In: KRAAK, Menno-Jan (Hrsg.) ; MOLENAAR, Martin (Hrsg.): *Proceedings of the 7th International Symposium on Spatial Data Handling* Bd. 2, 1996, S. 9A.1–9A.14
- Weibel und Dutton 1999** WEIBEL, Robert ; DUTTON, Geoffrey: Generalizing spatial data and dealing with multiple representations. In: LONGLEY, Paul (Hrsg.) ; GOODCHILD, Michael F. (Hrsg.) ; MAGUIRE, David J. (Hrsg.) ; RHIND, David W. (Hrsg.): *Geographical Information Systems: Principles, Techniques, Management and Applications* Bd. 1. 2nd edition. Cambridge, MA : John Wiley & Son, 1999, S. 125–155
- Worboys 2004** WORBOYS, Michael F.: *GIS: A Computing Perspective*. 2nd edition. London : Taylor & Francis, 2004
- Zeiler 1999** ZEILER, Michael: *Modeling Our World — The ESRI Guide to Geodatabase Design*. Redlands, CA : ESRI Press, 1999

Literaturverzeichnis

Danksagung

Mein herzlicher Dank gilt Professor Dr.-Ing. Dieter Morgenstern für die Betreuung der Dissertation. Seine vorbehaltlose Unterstützung gab mir die Freiheit, die für die Erstellung einer Dissertation nötig ist.

Besonderer Dank geht an Professor Dr. Lutz Plümer als Korreferent. Durch Anregungen und Ideen, die in intensiven Diskussionen entstanden, hat er einem Großteil der Arbeit ihre Ausrichtung gegeben.

Große Teile dieser Arbeit entstanden während zweier Forschungsprojekte, die am Institut für Kartographie und Geoinformation ausgeführt und vom Amt für Geowesen in der Bundeswehr (AGeoBw, vormals Amt für militärisches Geowesen, AMilGeo) finanziert und gefördert wurden.

Ein sehr großes Dankeschön geht an Dr. Gerhard Gröger, dem ich in einer Unzahl von Diskussionen Details dieser Arbeit darlegen durfte und der mich durch seine kompetente Meinung in vielen Punkten bestätigte und in anderen Punkten mir neue Anregungen gab.

Ebenso möchte ich ein sehr großes Dankeschön an Dr.-Ing. Ingo Petzold richten, dessen Aufmunterungen mich bei der Stange hielten, und der mich unermüdlich zur Fertigstellung anspornte. Seine Kommentare, auch zu unfertigen Gedanken, haben diese Arbeit wesentlich vorangebracht.

Außerdem möchte mich bei allen Mitarbeitern des Instituts für Kartographie und Geoinformation für die hervorragende Arbeitsatmosphäre und die große Unterstützung sowohl inhaltlicher als auch technischer Art bedanken.

Meinen Freunden und meiner Familie möchte ich dafür danken, dass sie die Zeit der Erstellung dieser Arbeit geduldig ertrugen und immer zu mir standen. Susanne danke ich für ihre Liebe, für ihre Unterstützung und für den Rückhalt, den sie mir immer gegeben hat.