Zeitschrift: Vermessung, Photogrammetrie, Kulturtechnik : VPK = Mensuration, photogrammétrie, génie rural

Herausgeber: Schweizerischer Verein für Vermessung und Kulturtechnik (SVVK) = Société suisse des mensurations et améliorations foncières (SSMAF)

Band: 77 (1979)

Heft: 7-8

Artikel: Datenbanken

Autor: Frank, A.

DOI: https://doi.org/10.5169/seals-229699

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. Find out more

Download PDF: 24.01.2026

ETH-Bibliothek Zürich, E-Periodica, https://www.e-periodica.ch

kuppelt die Forderung von gleich gross erscheinenden Messmarken, unabhängig vom Vergrösserungseinsatz, die Weiterentwicklung beeinflussen.

Formatfragen in Hinsicht auf 23 cm×46 cm Photographie können nicht ignoriert werden. Die Nachfrage in bezug auf solche Forderungen hängt weitgehend ab von der Freigabe von unorthodoxem Bildmaterial, wie solches mit militärischen Überwachungseinsätzen laufend erhalten wird, für zivile Zwecke.

Natürlich ist das Problem der Weiterentwicklung eng mit dem Begriff der automatischen Bildkorrelation verbunden. Hier ist wohl eine gewisse Skepsis am Platz. Diese richtet sich nicht auf den Wunsch nach angewandter Forschung auf diesem Gebiet, jedoch mag es nützlich sein, mit der Einführung derartiger Verfahren in die Praxis noch eine Weile hinzuwarten, bis die von der Photogrammetrie völlig unabhängige Entwicklung der digitalen Bildverarbeitung einerseits und der digitalen Steuertechnik anderseits weitere Fortschritte gemacht haben (vergleiche zum Beispiel [3]).

Abschliessend seien noch jene Massnahmen erwähnt, die zur Zeit am Institut für Geodäsie und Photogrammetrie der ETHZ im Lehrbereich Photogrammetrie in die Wege geleitet werden, um Lehre und Forschung der in den beschriebenen Gedankengängen angedeuteten Entwicklung der Photogrammetrie in ihren wesentlichsten Punkten auch durch entsprechende instrumentelle Komponenten anzupassen. Die Lösung dieses Problems wird mit einem in Abb. 9 mit seinen grundsätzlichen Komponenten gezeigten System angestrebt. Als Zentralkomponente dient ein Mini-Rechner der oberen Leistungsklasse mit Peripherie, wie sie in der untern Gruppe als Zusatzinstrumentation gezeigt ist. Mit dieser Figuration sollte es möglich sein, typische Probleme der rein numerischen Auswertung – also die numerische Umsetzung Komparator gemessener Bildkoordinaten in die Zielfunktion spezifischer Messanordnungen – auszuführen. Als ein weiteres Peripheriegerät ist ein mit Linear- beziehungsweise Dreh-Gebern digitalisiertes Analoggerät vorhanden, womit Probleme der computergestützten Analogauswertung (computergestützte Orientierungsverfahren) als auch auf Datenbank ausgerichtete Datenausgabe behandelt werden können. Schliesslich ist ein Analytischer Plotter vorhanden, der vor allem dazu dient, unter voller Ausnützung moderner Hardware-Kapazität durch sophistizierte Modellformulation und Echtzeitsteuerung die Anwendungsmöglichkeit der photogrammetrischen Methode weiter zu entwickeln.

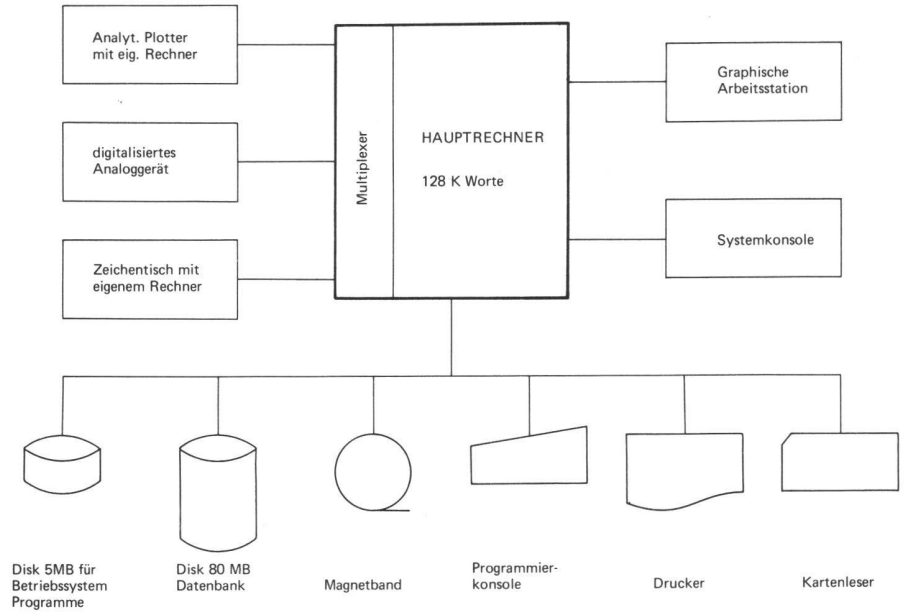

Abb. 9 Die wesentlichsten Komponenten eines computergestützten photogrammetrischen Auswertesystems

Der Datenpräsentation dient einmal ein «intelligenter» Zeichentisch und eine graphische Station. Wechselseitige Kombinationsmöglichkeiten der in Abb. 9 gezeigten Komponenten sollte eine Hardwareanpassung an eine beachtliche Anzahl von Problemstellungen im Sinne der in diesem Artikel angesprochenen Entwicklungstendenz in der Photogrammetrie ermöglichen.

Literatur

[1] Schmid, Hellmut, H.: Eine allgemeine analytische Lösung für die Aufgabe der Photogrammetrie. Bildmessung und Luftbildwesen Nr. 4, 1958 und Nr. 1, 1959.

[2] Schmid, Hellmut, H. und Heggli, Siegfried: Räumliche Koordinatentransformation Mitteilung Nr. 23, Juli 1978, des Instituts für Geodäsie und Photogrammetrie der Eidgenössischen Technischen Hochschule Zürich.

[3] Symposium on the Use of Digital Components in Photogrammetry. Technische Universität Hannover, Institut für Photogrammetrie und Ingenieurvermessungen, Hannover 1978.

Adresse des Verfassers:
Prof. Dr. Dr. h.c. Hellmut H. Schmid,
Institut für Geodäsie und Photogrammetrie,
ETH-Hönggerberg, CH-8093 Zürich

Datenbanken

A. Frank

La tendance actuelle des prix de revient dans le domaine de l'informatique peuvent etre résumée en quelques mots:

- *forte baisse des prix du matériel (jusqu'a 25% par an)*
- *augmentation incessante des frais de programmes.*

Les exigences sans cesse accrues imposées aux systèmes informatisés sont à l'origine de cette augmentation; elle a rendu nécessaire l'étude de solutions générales, faciles à entretenir et à adapter aux applications nouvelles; ceci a debouché sur l'élaboration de programmes modulaires et standardisés.

C'est ainsi que sont nées les banques de données, constituées des données elles-mêmes (base de données) et des programmes de gestion correspondants. Les programmes d'application peuvent être élaborés sans tenir compte de la forme sous laquelle sont stockées les données. Les programmes de ges-

tion de la base de données peuvent en outre assurer automatiquement l'exécution de certaines mesures de sécurité.

Il existe maintenant des langages de programmation très évolués, dont on pourrait mal se passer. On peut, par analogie, désigner les systèmes de banques de données par le terme de «langage de données». Elles offrent des possibilités immenses, pour la réalisation de programmes sûrs et d'entretien facile.

Le but de cet article est de préciser la notion de «banques de données», et d'évoquer quelques-uns des problèmes techniques relatifs à leur réalisation. Les possibilités d'applications dans les domaines touchant à la mensuration (Systèmes d'information du territoire) seront évoquées ultérieurement.

1. Einleitung

Eine neue Entwicklung der EDV-Technologie, das sog. Datenbankkonzept, wird auch im Vermessungswesen neue und bessere Lösungen verschiedener Probleme bringen.

Es sei gleich einleitend auf die enge Beziehung zum Aufgabenkreis «Landinformationssystem» verwiesen, die hier allerdings nicht behandelt werden soll. Ziel dieses Artikels ist vor allem die Klärung des Konzeptes «Datenbank» und der technischen Realisierung von Datenbanksystemen. Auf die sich daraus ergebenden Möglichkeiten der Anwendung wird später eingegangen.

1. Mechanische Bauteile werden immer mehr durch elektronische ersetzt.
2. Elektronische Bauteile werden immer billiger (zur Zeit etwa 25% pro Jahr!)
3. Ökonomische Grenzen werden daher kaum mehr von der Hardware gesetzt.
4. Es zeichnet sich ein Übergang von Batch-Verarbeitung[1] zu interaktivem timesharing Betrieb[2] ab.
5. Software (d. h. Programme aller Stufen) werden zum entscheidenden Faktor, sowohl was die Kosten als auch was die Entwicklungszeit betrifft.
6. Der Unterhalt von Programmen wird zu einem wichtigen Problem.

[1] Die Programme werden vorbereitet und laufen ohne Intervention des Benützers ab (Stapelverarbeitung).

[2] Mehrere Programme laufen nebeneinander (timesharing) und unter Kontrolle des Benützers (interaktiv) ab.

2. Zur Entwicklung der EDV

Die Entwicklung von Datenbanksystemen ist vor dem Hintergrund der allgemeinen Entwicklung der EDV gut zu verstehen.

Die angebotenen Systeme eröffnen ständig zusätzliche Anwendungsgebiete. Es ist aber kaum mehr möglich, die notwendigen Programme zu erstellen, die auch dem EDV-Laien erlauben, deren Vorteile zu nützen (sog. Software-Krise). Deshalb versucht man heute weniger als früher, die EDV-Hardware maximal auszunützen. Da die Hardware billiger geworden ist, kann man darauf verzichten, für jedes konkrete Problem jene Lösung zu suchen, die mit der minimalsten Hardware-Belastung auskommt. Um es etwas pointiert auszudrücken: es soll nicht mehr die Kernspeicher-Belegung oder die Laufzeit eines Programmes minimalisiert werden, wie noch vor wenigen Jahren notwendig, sondern heute ist der Aufwand für Programmerstellung und -unterhalt klein zu halten.

Eines der Mittel, um dies zu erreichen, ist die Verwendung *höherer Programmiersprachen* und vorbereiteter Standardprozeduren; dazu gehören auch *Datenbanksysteme*. Diese werden nicht nur eingesetzt, um grosse Datenbestände auf möglichst kleinem Raum speichern oder um möglichst rasch auf Daten zugreifen zu können, sondern vor allem, um Programme möglichst einfach und wartungsfreundlich zu gestalten. Datenbanksysteme können in Analogie zu den höheren Programmiersprachen als *«höhere Datensprachen»* [1] bezeichnet werden. Sie erlauben wie jene, Probleme auf einer dem Benützer näheren, logisch höheren Ebene zu beschreiben.

Die Probleme, die mit Datenbanksystemen gelöst werden, wären technisch an sich auch mit den bisherigen Mitteln zu bewältigen. Praktisch ist dies aber nicht mehr möglich, weil solche Programme zu erstellen zu lange dauerte und nicht zu bezahlen wäre, vom Unterhalt dieser Programme ganz zu schweigen.

3. Datenverwaltung

Ursprünglich wurde die EDV vor allem für schwierige Berechnungen oder umfangreiche Routinearbeiten eingesetzt. Dabei wurden aus Einzeldaten Ergebnisse berechnet und Zwischenergebnisse nach den Erfordernissen anschliessender Programme gespeichert. Die Datenorganisation geschah im Hinblick auf bestimmte Programme, und die notwendige Bearbeitung der Daten erfolgte durch diese Programme.

Die technische Entwicklung verlangt aber mehr und mehr *integrierte Lösungen*. Dabei bearbeiten verschiedene Programme immer wieder dieselben Daten. Deshalb ist es nicht mehr zweckmässig, die Daten in Funktion der Programme zu organisieren. Sie sollen von den Programmen getrennt werden, und einheitliche Zugriffsmethoden müssen sicherstellen, dass sich die verschiedenen Benützer gegenseitig nicht stören: zwischen Daten und Benützer tritt das Datenbankverwaltungssystem. Die Daten werden so von den Anwenderprogrammen unabhängig und stehen allen Anwendungen gleichartig zur Verfügung.

Es ist erkannt worden, dass Daten ein wertvoller «Rohstoff» für viele Verwaltungstätigkeiten sind. Dieser Stoff muss aber richtig bewirtschaftet werden. Man denke an die vielen Datensammlungen, die in Form von EDV-Dateien, aber auch als traditionelle Karteien in der Verwaltung geführt und für verschiedenste Aufgaben immer wieder benötigt werden. Da sie ständig nachgeführt werden müssen und doch immer wieder Fehler enthalten, ist es leicht verständlich, dass der Einsatz der EDV gerade hier zweckmässig wird.

Es wird immer wirtschaftlicher, statt Karteien von Hand nachzuführen, EDV-Mittel einzusetzen: der Sachbearbeiter kann über seinen Bildschirm den aktuellen Stand jederzeit abfragen. Moderne Betriebssysteme ermöglichen es, mehreren Benützern gleichzeitig Einblick in den gleichen Datenbestand zu geben. Man will Daten nur einmal erfassen und speichern, aber nach den verschiedensten Gesichtspunkten abfragen. Als Beispiel: eine Parzelle kann über den Eigentümer-Namen, aber auch über die Parzellennummer, die Postadresse oder die Koordinaten gefunden werden.

Da Datenverwaltung ein Grundproblem jeder Datenverarbeitung ist, tauchen diese Probleme immer wieder in ähnlicher Form auf. Deshalb wurde schon sehr bald begonnen, standardisierte Verwaltungs-Routinen zu entwikkeln und einzusetzen. 1971 wurde von der Data Base Task Group des CODASYL-Comittee, den Schöpfern von COBOL, ein Vorschlag zur Vereinheitlichung publiziert, der sich heute immer mehr durchsetzt [2]. Praktisch sind heute für alle grösseren Anlagen und zunehmend auch für Klein-Computer-CODASYL-Standard-Datenbanksysteme erhältlich.

4. Das Datenbank-Konzept

Im folgenden Kapitel soll erklärt werden, was man unter einer Datenbank versteht und welchen minimalen Anforderungen eine solche genügen muss.

4.1. Begriff

Wir unterscheiden begrifflich das Datenbanksystem, das Datenbankverwaltungssystem, die Datenbestände und die Datenbank. Das *Datenbanksystem*

enthält allgemeine Routinen, um Datenbankverwaltungssysteme für verschiedenste Zwecke zu erzeugen. Es wird im allgemeinen vom Hersteller gegen Gebühr geliefert und unterhalten. Mit diesem Datenbanksystem erstellt der Benützer ein bestimmtes *Datenbankverwaltungssystem*, ein für seine speziellen Zwecke zugeschnittenes Betriebs-Programm, das seinen *Datenbestand* verwaltet (vgl. Abb. 1).

Zu einer *Datenbank*[3] gehören also nicht nur die Daten (der Datenbestand), sondern auch das Verwaltungssystem, das den Datenverkehr regelt – genauso wie zu einer kommerziellen Bank nicht nur der Geldbestand im Tresor gehört, sondern auch der Schalterbeamte mit seiner Dienstanweisung, die festlegt, wem er nach welchen Verfahren Geld aushändigen darf.

4.2 Minimale Anforderungen

4.2.1 Trennung von Daten und Programmen

Datenbanken werden dann eingesetzt, wenn verschiedene Programme auf die gleichen Daten zugreifen sollen, und zwar vor allem damit später unabhängig von den Daten zusätzliche Programme erstellt und bestehende Programme geändert werden können. Entscheidend ist dabei, dass Änderungen in einem Programm nicht zu Anpassungen in andern, den gleichen Datenbestand verwendenden Programmen führen. Nur unter dieser Voraussetzung ist ein grösseres System von Programmen etwa für die Parzellarvermessung überhaupt zweckmässig zu unterhalten und fortzuführen.

Es muss aber auch möglich sein, die Form der Daten-Speicherung zu verändern, etwa einem wachsenden Datenbestand anzupassen oder für neue Programme neue Datensorten (z. B. für Fixpunkte neben den Koordinaten auch die Höhen) einzuführen, ohne dass die vorhandenen zugreifenden Programme geändert werden müssen.

Technisch wird das gelöst, indem die *Datenverarbeitung* in den Programmen und die *Datenverwaltung* in der Datenbank vollständig getrennt werden. Die benötigten Daten werden über standardisierte *Datenbankschnittstellen* ausgetauscht. Ein Programm, das den Inhalt eines bestimmten Datenfeldes benötigt, muss eine Abfrage in einer vorgeschriebenen Form an das Datenbankverwaltungsprogramm stellen, das dann den entsprechenden Wert bereitstellt (vgl. Abb. 1).

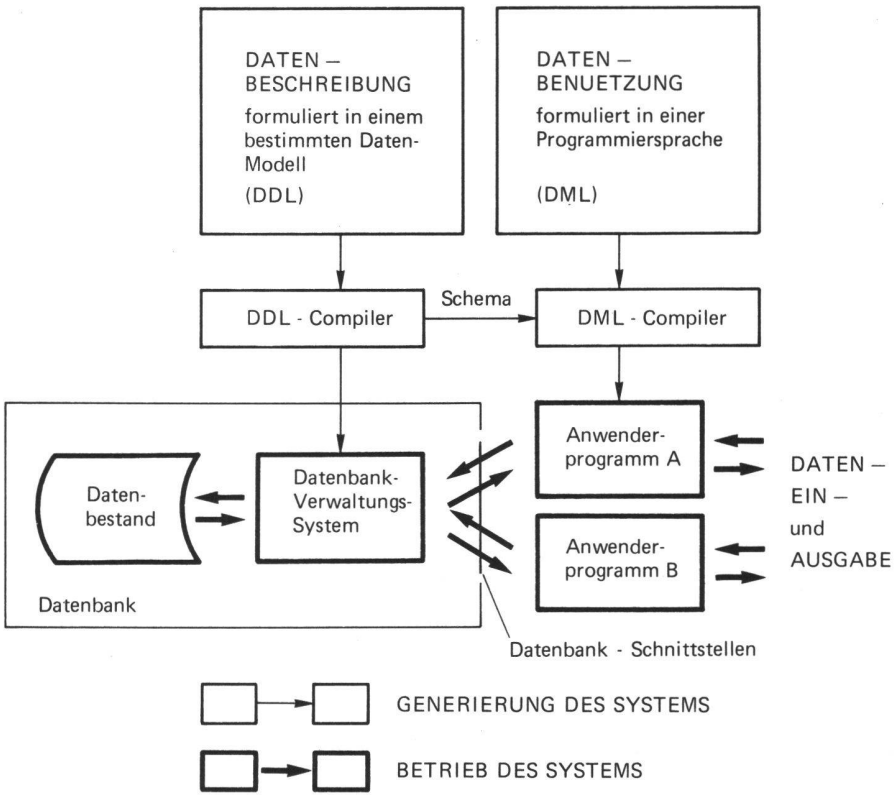

Abb. 1

Zum Beispiel will ein Geometer die Koordinaten eines Triangulationspunktes am Bildschirm abfragen. Dazu benützt er ein Programm, das ihn zur Eingabe der Punktnummer auffordert. Dieses Programm verlangt nun von der Datenbank die entsprechenden Werte, indem eine Abfrage in einer bestimmten Form (z. B. «FIND KOORDINATEN USING PUNKTNUMMER» und anschliessend «GET KOORDINATEN») gestellt wird. Das Datenbankverwaltungssystem beantwortet diese Abfrage, indem das Datenfeld «KOORDINATEN» die Werte der Koordinaten des gesuchten Punktes erhält.

Man wird die Prozedur mühsam finden; für den Programmierer aber, der nur diese Frage an das Datenbankverwaltungssystem in der von ihm verwendeten Programmiersprache schreiben muss, ist dies jedoch einfach, denn er muss sich nicht um die physische Speicherung kümmern und detailliert beschreiben, wie das Programm die verlangten Daten im Speicher finden soll (die einfältigen Dinge soll man dem Computer überlassen!).

4.2.2 Redundanzarme Speicherung der Daten

Mit «redundanzfreier» Speicherung der Daten meint man, dass jede Information nur einmal an einer Stelle gespeichert wird. Der Begriff stammt aus der Informationstheorie, wo unter Redundanz jener Teil einer übertragenen Nachricht verstanden wird, der nichts Neues zum Wissen des Empfängers beiträgt (d. h. das, was aus irgendeinem Grund bereits mitgeteilt oder selbstverständlich ist). Redundanz vergrössert einen Datenbestand, weil die selben Daten mehrmals gespeichert werden. Dennoch ist Redundanz nicht immer schädlich: Redundanz hat an vielen Stellen eine wichtige Kontrollfunktion. Als Beispiel dienen die allen bekannten schweizerischen Militärkoordinaten: jedermann weiss, dass sich mit weniger Ziffern die selben Angaben machen lassen. Die zusätzlichen, eigentlich redundanten Zeichen helfen Verwechslungen zwischen Y- und X-Werten zu vermeiden, haben also eine durchaus nützliche Funktion.

Für Datenbanken wird gefordert, dass die gespeicherten Daten redundanzarm seien. Einerseits ist dies wegen der Speicherökonomie angebracht (Redundanz braucht Speicherplatz), anderseits, und das scheint wichtiger, erschwert Redundanz die Nachführung von Datenbeständen erheblich: sind die selben Daten an verschiedenen Stellen gespeichert – und das liegt im Wesen der Redundanz – müssen sie auch an all diesen verschiedenen Stellen nachgeführt werden. Wären sie überall in gleicher Form gespeichert, so ergäben sich dabei keine speziellen Probleme. Meist tritt aber Redundanz in einer versteckten Form auf:

[3] Im Unterschied zur Terminologie von IBM wird im folgenden nicht von verschiedenen Datenbanken gesprochen, sondern der gesamte Datenbestand und das zugeordnete Datenbankverwaltungssystem wird als eine Datenbank angesprochen. Auf einem Computer-System können hingegen durchaus verschiedene Datenbanken existieren, die von verschiedenen Datenbankverwaltungssystemen verwaltet werden.

Sind in einer Datenbank beispielsweise die Parzellen, die Grenzpunkte mit den Koordinaten und die Parzellenflächen gespeichert, so müssen im Falle einer Mutation, die die Veränderung einer Grenze bewirkt, nicht nur Grenzpunkte und Koordinaten neu gespeichert werden, sondern die Fläche ist neu zu berechnen. Die Fläche ist neben den Grenzpunktkoordinaten eigentlich redundant. Bei der gleichen Mutation müssen aber auch die Nachbarparzellen angepasst werden. Auch hier tritt Redundanz auf, indem eine Aussage über eine Grenzlinie in zwei verschiedenen Zusammenhängen – den beiden benachbarten Parzellen – erfasst und gespeichert wird.

4.2.3 Trennung der verschiedenen Ebenen der Datenbeschreibung

Wenn die Datenbank von den zugreifenden Programmen unabhängig sein soll, muss auch die Beschreibung der Daten unabhängig von den Programmen erfolgen. Damit Änderungen möglichst einfach sind, wird verlangt, dass die Beschreibung der Daten und ihrer Speicherung und Verwendung auf folgende drei Ebenen aufgeteilt wird[3]:

1. Logische Datenbeschreibung
2. Beschreibung der physischen Datenspeicherung
3. Beschreibung der Daten im Hinblick auf die Verwendung durch die zugreifenden Programme.

4.2.3.1 Logische Datenbeschreibung

Die logische Datenbeschreibung enthält eine vollständige Aufzählung aller Datenfelder und deren Bezeichnungen. Dabei entspricht der CODASYL-Standard etwa der Syntax der Datendeklarationen in einem COBOL-Programm. Ferner müssen die Beziehungen zwischen den Daten (Datenstrukturen) dargestellt werden. Um für alle folgenden Erklärungen ein einheitliches Beispiel benutzen zu können, sei folgende Aufgabe gestellt:

Die auf einem Grundbuch-Plan dargestellten Parzellen und die darauf befindlichen Gebäude seien zu speichern (Abb. 2). Zuerst werden die Datenfelder (in Abb. 3 in den Kästchen dargestellt) und die entsprechenden Feldlängen angegeben. Damit kennt das System die Datensätze und Datenfelder, hingegen ist noch nicht erklärt, wie die Strassennamen mit den Gebäuden oder die Grenzpunkte mit den Parzellen zusammenhängen. Diese Zusammenhänge sind in der gleichen Figur grafisch dargestellt. Diese Beziehungen werden durch Anweisungen ausgedrückt; damit ist die logische Beschreibung, im allgemeinen (Datenbank-)Schema genannt, unseres Modells abgeschlossen.

Weil die Kodierung einfach ist, kann sich der Programmierer vermehrt auf das Erstellen des Modelles konzentrieren. Besonders bei Aufgaben aus dem Vermessungswesen tauchen schwierige Probleme auf, wenn raumbezogene Datenstrukturen in einem Modell abgebildet werden müssen; am Institut für Geodäsie und Photogrammetrie sind entsprechende Untersuchungen geplant.

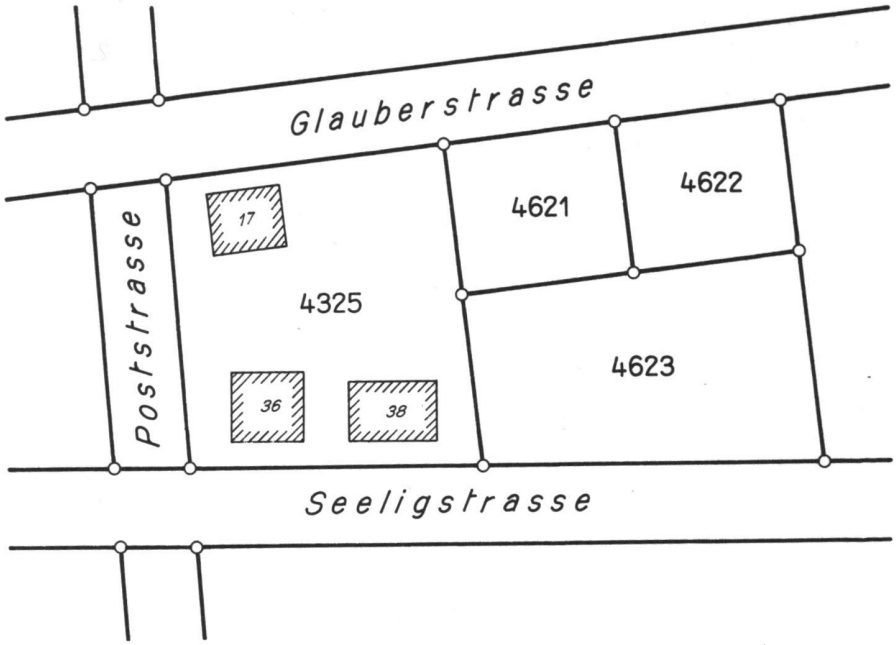

Abb. 2

4.2.3.2 Beschreibung der physischen Datenspeicherung

Neben der logischen Beschreibung der Daten muss dem EDV-System auch erklärt werden, wo diese Daten zu spei-

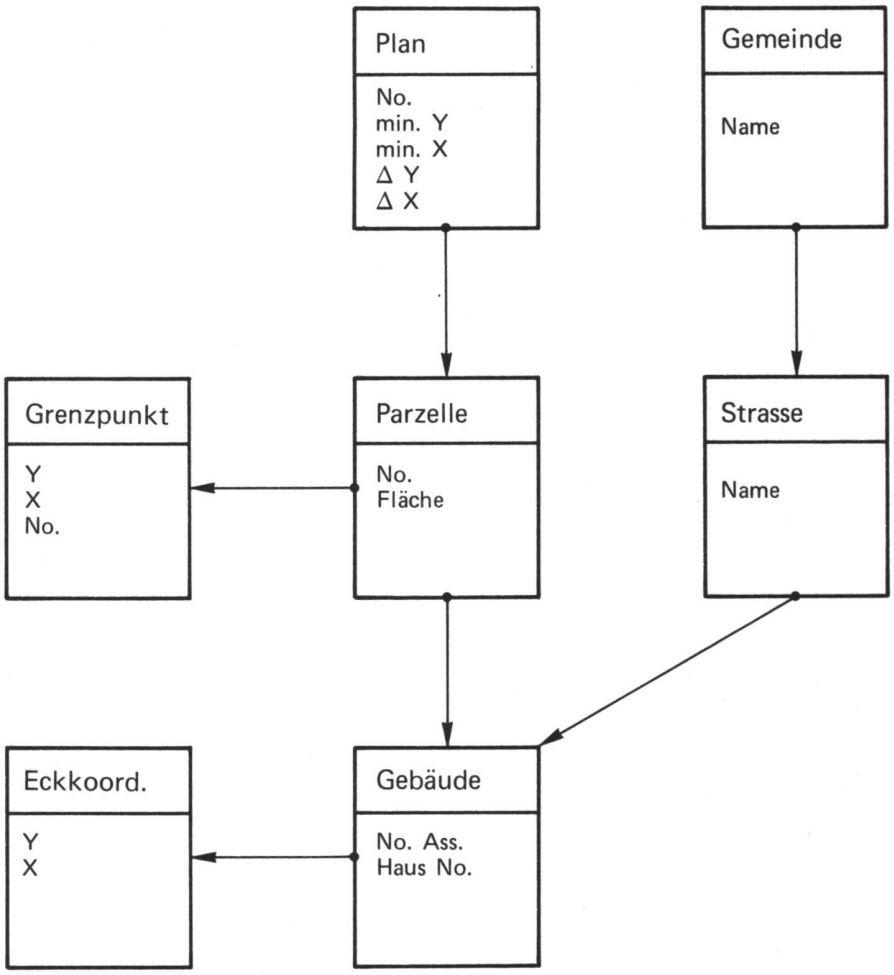

Abb. 3

chern sind. Das geschieht in einem speziellen Teil der Datenbank-Beschreibung. Dabei ist wichtig, dass dieser Teil angepasst werden kann, ohne dass die logische Datenbeschreibung verändert wird. Damit kann auf einfache Art mehr Platz für bestimmte Datensorten geschaffen werden; es kann aber auch erreicht werden, dass die Daten auf dem ihnen speziell angemessenen Speichermedium gespeichert werden. Verfügt eine Anlage über schnellere und langsamere Direktzugriffsspeicher, so werden die häufig gebrauchten Daten auf einem schnelleren (und im allgemeinen teureren) Speichermedium abgelegt als die weniger oft benötigten. Es sollte sogar möglich sein, ganze Datenbereiche auf andere Datenträger zu übertragen und aus dem Direktzugriffsspeicher zu entfernen, ohne dass die Funktion der Datenbank behindert wird. Dies etwa für Messdaten, die nach der Auswertung meist nicht mehr gebraucht werden, aber für besondere Fälle doch noch auf Magnetband aufbewahrt werden sollten.

4.2.3.3 Beschreibung im Hinblick auf den Benützer

Dem einzelnen Benützer müssen nicht immer alle Einzelheiten der ganzen Datenbeschreibung mitgeteilt werden. Er muss nur das für seine Anwendung Notwendige wissen. Datenfelder, zu denen er aus Datenschutzgründen ohnehin keinen Zugriff hat oder die für seine Anwendung ohne Bedeutung sind, sollen ihn nicht verwirren. Dieses Prinzip – man nennt es Informations-Hiding – ist auch im Hinblick auf den Programmunterhalt von Bedeutung: Die Auswirkungen von Änderungen werden damit auf jene Teile beschränkt, denen diese Informationen zugänglich sind.

Anderseits wird oft gewünscht, dass Datenfelder für bestimmte Kategorien von Benützern anders bezeichnet oder sogar anders eingeteilt werden. Es sollte auch möglich sein, «virtuelle» Datenfelder für die Benützer zu definieren. Beispielsweise werde ein Datenfeld «FLAECHE» im Datensatz «PARZELLE» eingefügt, dessen Wert gar nicht gespeichert wird, für das aber eine Berechnungsroutine vorhanden ist, die die Flächenrechnung im Bedarfsfall aus den gespeicherten Koordinaten vornimmt. Es wird dem Benützer also ermöglicht, auf einen Wert zuzugreifen, wie wenn er gespeichert wäre. Der Wert braucht aber keinen Speicherraum und stellt keine Nachführungsprobleme (vgl. 4.2.2).

4.2.4 Anforderungen der Datenintegrität

Hätte ein Datenbankverwaltungssystem nicht mehr zu leisten, als Daten zu speichern und wiederzufinden, so wäre dies eine recht einfache Angelegenheit. Der Benützer will aber meist wesentlich mehr: in vorderster Linie stehen die Anforderungen der Datenintegrität (Datenkonsistenz, Datenschutz und Datensicherheit).

4.2.4.1 Datenkonsistenz

Um zu verhindern, dass falsche Daten in die Datenbank aufgenommen werden, müssen diese vor der Speicherung geprüft werden. Es wird erstens geprüft, ob der angegebene Wert im *zulässigen Wertebereich* für dieses Datenfeld liegt (eine schweizerische Y-Koordinate muss zwischen 470 000 und 850 000 liegen). Dann wird geprüft, ob die eingegebenen Daten *vollständig und in sich widerspruchsfrei* sind (ein Punkt kann nicht eingegeben werden, ohne dass mindestens Y- und X-Koordinate gespeichert werden). Es wird drittens geprüft, ob nicht ein *Widerspruch zwischen den neu zu speichernden Daten und den bereits vorhandenen* entsteht (z. B. wäre beim Speichern eines Gebäudes zu kontrollieren, ob es auch tatsächlich innerhalb der angegebenen Parzelle liegt). Diese Bedingungen, abstrakt formuliert, sollen im logischen Schema enthalten sein; die heutigen Datenbanksysteme können noch nicht alle Möglichkeiten anbieten.

4.2.5 Datensicherheit

Diese Anforderung ist wohl fast am wichtigsten: wenn wir Daten speichern, müssen wir auch sicher sein, dass wir sie zu einem späteren Zeitpunkt wieder abrufen können und dass sie nicht verlorengehen oder gar versehentlich verfälscht werden. Es müssen also umfangreiche Sicherungsmassnahmen getroffen werden, damit nicht durch Fehler in der Hardware oder durch Fehler der Benützer Daten verlorengehen. Als zusätzliche Sicherung wird in beinahe allen Fällen eine laufende Aufzeichnung aller vorgenommenen Änderungen verlangt, und es werden Programme vorbereitet, die erlauben, aus den Sicherheitskopien, die periodisch zu erstellen sind, und diesen Änderungsprotokollen (sog. Log-Files) den neusten Stand wieder herzustellen. Dies für den Fall, dass z. B. durch physische Beschädigung des Speichermediums doch einmal Daten verlorengehen sollten.

4.2.6 Datenschutz

Der Datenschutz umfasst alle Massnahmen, die ergriffen werden, damit Unbefugte weder Daten verändern noch einsehen können. Im Bereich des Vermessungswesens ist sicher nicht nur der Schutz gegen versehentliche (vgl. Datensicherung), sondern auch gegen unbefugte Veränderungen wichtig. Deshalb müssen Daten auch gegen unberechtigten Zugriff geschützt werden. Dieser Schutz ist eine Hauptaufgabe der Datenbanksysteme und insbesondere ein Grund, die spezielle Benützersicht in der Datenbeschreibung (vgl. 4.2.3) einzuführen. Es wird mit Hilfe von Passwörtern, persönlichen Ausweisen oder speziellen Prozeduren gewährleistet, dass nur berechtigte Benützer und deren Programme auf Daten zugreifen oder diese verändern können; dabei kann beispielsweise das Lesen erlaubt, das Ändern hingegen verboten sein.

5. Technische Realisierung

Wie die Probleme der Datenspeicherung üblicherweise von modernen Datenbanksystemen gelöst werden, kann hier nur in den Grundzügen beschrieben werden. Für Einzelheiten sei auf die spezielle Literatur und insbesondere auf die Manuals der Herstellerfirmen verwiesen [4].

Hier soll von den prinzipiell möglichen Verfahren nur dasjenige beschrieben werden, das bei Realisierungen des CODASYL-Standards häufig angewendet wird. Andere Verfahren, wie sie etwa für relational aufgebaute Datenbanken oder für hierarchische Systeme wie IMS von IBM verwendet werden, unterscheiden sich davon z. T. erheblich.

Wir gehen bei dieser Beschreibung davon aus, dass die Datenfelder in Datensätzen (records) zusammengefasst werden. Vom Datenbankverwaltungssystem werden diese Datensätze mit einer eindeutigen Nummer versehen und im Speicher abgelegt, wobei zwischen der Speicheradresse und der Nummer, im allgemeinen Datenbankschlüssel genannt, ein eindeutiger Zusammenhang besteht. Jeder Datensatz kann im Speicher sofort gefunden werden, wenn sein Datenbankschlüssel bekannt ist.

Wie wir in 4.2.3.1 gesehen haben, müssen neben den Datensätzen auch deren Beziehungen zueinander festgehalten werden. Dies geschieht mit Hilfe von *verketteten Listen*. Alle Datensätze, etwa diejenigen der Häuser einer Strasse, die an beliebiger Stelle und in beliebiger Reihenfolge im Speicher stehen, werden untereinander mit Hilfe von Zeigern zu einer Kette verbunden, indem jeweils beim vorangehenden Datensatz der Datenbankschlüssel des nachfolgenden Satzes gespeichert wird. Nach dem gleichen Verfahren wird aber auch die logische Beziehung zwischen einer Parzelle und den darauf befindlichen Häusern angegeben.

Um nach Abbildung 2 zu speichern, dass das Haus «Glauberstrasse 17» (Schlüssel 4120) auf der Parzelle 4325 steht, wird beim Datensatz dieser Parzelle der Schlüssel des Hauses als Zeiger mitgespeichert (siehe Abb. 5). Würde z. B. beschlossen, das Haus «Glauberstrasse 17» neu als «Poststrasse 27» zu bezeichnen, so wäre nur der entsprechende einmal gespeicherte Datensatz für dieses Haus zu ändern. Am Daten-

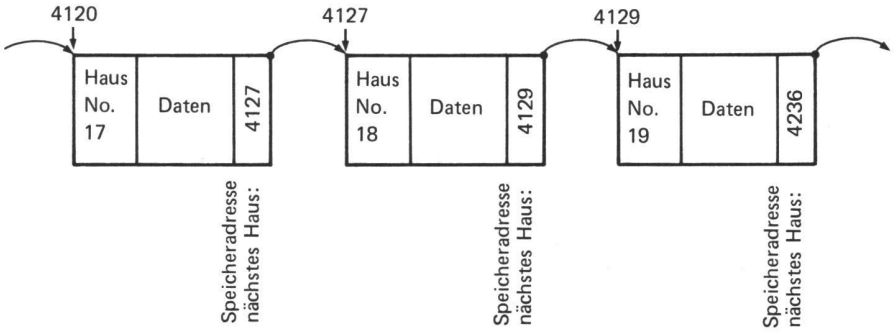

Abb. 4

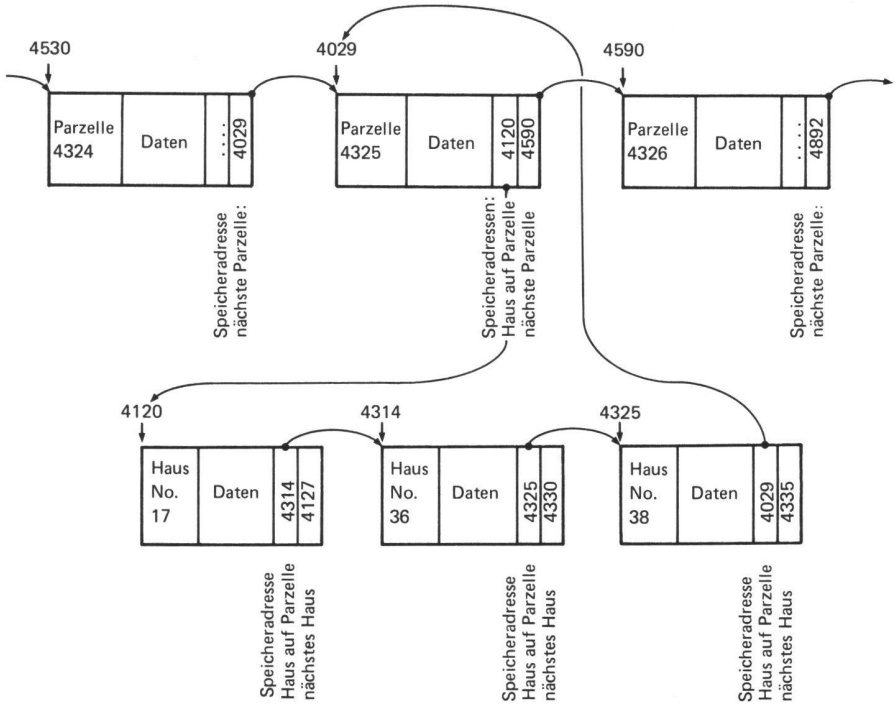

Abb. 5

satz der Parzelle ändert sich nichts, denn dort ist nur der Verweis auf den Datensatz des Hauses (der geändert wurde), nicht aber die Postadresse selbst gespeichert.

Mit diesem System wird die geforderte redundanzarme Speicherung erreicht und die Nachführung des Datenbestandes erleichtert. Bei der Verarbeitung müssen die einzelnen Datenelemente wieder zusammengesucht werden, was einen leistungsfähigen Rechner und Speicher mit schnellem Zugriff erfordert. Datenbanksysteme wurden deshalb im Anfang nur für Grossanlagen entwickelt und installiert; inzwischen hat sich aber ein solcher Leistungszuwachs bei Kleincomputern ergeben, dass Datenbanksysteme auch auf diesen installiert und damit diese Konzepte auch für kleinere Betriebe nutzbar werden.

6. Datenverteilung

Können Datenbanken auch auf Kleinrechnern betrieben werden, so stellt sich die Frage, wo die Daten gespeichert werden sollen, wenn Daten dezentral verwendet werden und bei räumlich verteilten Verwaltungstätigkeiten anfallen. Grundsätzlich sollte man davon ausgehen, dass die Daten am Ort ihrer Verwendung gespeichert werden sollen: «die Speicher gehen zu den Daten».

Dies lässt sich immer besser verwirklichen, da die elektronischen Speichermedien billiger werden, die Datenübertragung hingegen im Preis ungefähr gleich bleibt.

Aber wenn wir die Daten dezentral verwalten, entsteht früher oder später ein Bedürfnis, von einer andern Stelle aus auf die Daten zuzugreifen und Auswertungen von Datenbeständen anderer Rechner vorzunehmen. Der Ingenieur-Geometer wird beispielsweise auf die Fixpunktkoordinaten, die beim kantonalen Vermessungsamt nachgeführt werden, zugreifen wollen, oder eine Amtsstelle benötigt Angaben aus einem Leitungskataster. Damit die entsprechenden Rechner miteinander kommunizieren können, müssen einheitliche Schnittstellen und Protokolle für den Datenaustausch festgelegt werden.

In diesem Bereich sind noch Fragen offen, an deren Lösung die Hersteller von Computern intensiv arbeiten. Es scheint aber heute bereits festzustehen, dass Datenbanken in Zukunft vor allem auf *verteilten Kleinsystemen* angelegt und die zentralen Gross-Systeme eher Ausnahmeerscheinungen bilden werden.

7. Zusammenfassung

Die Forschung nach umfassenden EDV-Lösungen, die leicht zu unterhalten und wechselnden Ansprüchen anzupassen sind, verlangt einen modularen Aufbau der Programme und den Einsatz von Standardprogrammen. Dazu gehören insbesondere Datenbanken, bei denen die Datenbestände und ihre Verwaltung von den Anwenderprogrammen getrennt sind. Sie ermöglichen Daten und ihre Beziehungen untereinander für einen ganzen Aufgabenbereich so zu verwalten, dass für die Verarbeitung und Speicherung anpassungsfähige Systeme entstehen, deren Sicherheit weitgehend automatisch gewährleistet ist.

Analog zu den höheren Programmiersprachen, auf die heute kaum noch jemand verzichten möchte, sind Datenbanksysteme als *höhere Datensprachen* zu bezeichnen. Sie sind eine wichtige Voraussetzung, um zuverlässige, wartungsfreundliche und anpassungsfähige Programme mit minimalem Aufwand zu erstellen.

Literatur:
[1] Prof. C. A. Zehnder: Datenbank-Einsatz, ETH Zürich, Institut für Informatik, 1978
[2] CODASYL Data Base Task Group (DBTG) Report. IFIP Administrative Data Processing Group, Amsterdam 1971 (und spätere).
[3] ANSI/SPARC Interim Report of the Study Comittee on Data Base Management Systems. ACM Sigmod Newsletter 1975
[4] Als Beispiel: DBMS-10 Administrator's Procedures Manual, Digital Equipment Corp. Maynard (Mass.) 1977.

Allgemeine Literatur:
Wedekind, Helmut: Datenbanksysteme, BI Hochschultaschenbücher, Zürich 1974.
Lockemann, P. C., Mayr, H. C.: Rechnergestützte Informationssysteme, Springer-Verlag, Berlin 1978.

Adresse des Verfassers:
André Frank, dipl. Kult. Ing. ETH,
Institut für Geodäsie und Photogrammetrie,
ETH-Hönggerberg, 8093 Zürich