Schlagwort "Deduktive Datenbanken"

François Bry[1], Dietmar Seipel[2]
[1]Institut für Informatik, Ludwig–Maximilians–Universität München
[2]Institut für Informatik, Julius–Maximilians Universität Würzburg

Einleitung

Während der letzten 15 Jahre ist die Symbiose von Datenbanksystemen und *Logikprogrammierung* intensiv untersucht worden, mit dem Ziel traditionelle sowie neue Datenbankanwendungen besser zu behandeln. Dies führte zu Erweiterungen der herkömmlichen, relationalen Datenbankensysteme um *regelbasierte Sprachen* zur Wissensrepräsentation, die sowohl als Modellierungs- als auch als Anfragesprache verwendet werden, und um *Deduktionsmethoden* zur Wissensverarbeitung. Deduktive Datenbanksysteme ermöglichen eine wesentliche Erweiterung des Anwendungsbereiches von Datenbanksystemen.

Grundkonzepte

Eine deduktive Datenbank besteht aus Tupeln (auch *Fakten* genannt), wie dies auch für eine relationale Datenbank der Fall ist, sowie zusätzlich Ableitungsregeln und Integritätsbedingungen.

Ableitungsregeln. Eine Ableitungsregel ist eine allgemeine Definition, die es ermöglicht Daten intensional zu spezifizieren, wie etwa die Regel *"Alle Bücher des Springer–Verlages aus der Reihe LNCS sind auf Englisch verfaßt"* zur Verwaltung eines Informationssystems über eine Bibliothek. Zum Ausdruck von Ableitungsregeln werden Sprachen mit "Wenn–Dann–Regeln" verwendet, wie in den Bereichen der Künstlichen Intelligenz (vgl. manche Expertensysteme), der Logikprogrammierung sowie der automatischen Deduktion. Dabei wird ein Tupelkalkül, wie etwa in der bekannten relationalen Sprache SQL, verwendet (z.B. "*Buch.Nummer 1234 and Buch.Sprache Englisch*"), oder ein sogenannter Domänenkalkül (womit das obige SQL–Beispiel als *"Buch(1234, Englisch)"* ausgedrückt wird), wie in der Logik. Die oben erwähnte Ableitungsregel kann z.B. wie folgt im Domänenkalkül dargestellt werden:

$$\begin{aligned} \mathit{Buch(x,Englisch)} \leftarrow\ & \mathit{Verlag(x,Springer)}, \\ & \mathit{Reihe(x,LNCS)} \end{aligned}$$

Rekursive Regeln ermöglichen die intensionale Spezifikation von Daten, die auf Zusammenhänge von unbestimmtem Umfang beruhen, wie etwa die Definition einer *Stücklistenauflösung* (bill–of–materials) oder einer indirekten Flugverbindung mit möglicherweise mehreren Maschinenwechseln. Wie in relationalen Datenbanken wird die Negation als Scheitern interpretiert, d.h. ein negierter Ausdruck wie etwa "*not Buch(x,Englisch)*" gilt als wahr, wenn für *keine* Buchnummer "*x*" der positive Ausdruck "*Buch(x,Englisch)*" ableitbar ist.

Ableitungsregeln werden bei der Auswertung von Anfragen automatisch eingesetzt, um das intensional dargestellte Wissen effizient wiederzugeben, wie etwa zur Beantwortung folgender Anfrage über in Englisch geschriebene Bücher über Datenbanksysteme (DBS) unter Berücksichtigung der oben erwähnten Regel:

$$\mathit{Buch(x,Englisch),\ Thema(x,DBS)}$$

Ein Hauptthema der Forschung über deduktive Datenbanken war die Definition von terminierenden und vor allem effizienten *Anfrageauswertungsverfahren.*

Integritätsbedingungen. Integritätsbedingungen ermöglichen normative Bedingungen auszudrücken, wie z.B. folgende funktionale Abhängigkeit zwischen Attributen einer Relation: *"Die Nummer eines Buches ist ein eindeutiger Bezeichner des Buches"*, oder beliebig komplexe Zusammenhänge zwischen Daten: *"An jedem Arbeitstag gibt es mindestens einen direkten Flug aus einem der drei größten Flughäfen Deutschlands zur Ostküste der Vereinigten Staaten"*. Integritätsbedingen setzen also Normen, die die Datenbank nach jeder beliebigen Aktualisierung erfüllen soll. Im Gegensatz zu Ableitungsregeln können Integritätsbedingen in der Regel zur Beantwortung von Anfragen *nicht* verwendet werden, weil sie die Daten nicht notwendigerweise eindeutig spezifizieren.

Integritätsbedingungen können anschaulich mit Hilfe von Ableitungsregeln dargestellt werden, deren Konklusionen "falsch" sind, wie etwa in:

$$\begin{aligned} \mathit{false} \leftarrow \mathit{not}\ (\ & \mathit{Buch(x,Englisch),\ Reihe(x,LNCS)}, \\ & \mathit{Thema(x,DBS)}\ ) \end{aligned}$$

d.h. *"Es ist falsch, daß es kein Buch 'x' aus der LNCS–Reihe über Datenbanksysteme gibt"*. Weitere Hauptbeiträge der Forschung über deduktive Datenbanksysteme sind die Behandlung von *uneingeschränkten, allgemeinen Integritätsbedingungen* und die *Automatisierung der Integritätsprüfung*, d.h. der Spezifikation von effizienten, *inkrementellen* Verfahren zur Auswertung der Integritätsbedingungen nach jeder Aktualisierung.

Objektorientierte Datenmodellierung. Da die "flache" Wissensrepräsentation in der sogenannten "*ersten Normalform*" in Tupel- und Domänen–Kalkülen zu wenig Strukturierungsmöglichkeiten bietet, sind deduktive und objektorientierte Modellierungs- und Anfragesprachen kombiniert worden, so daß neben den herkömmlichen Daten auch Ableitungsregeln und Integritätsbedingungen in einem objektorientierten Kontext spezifiziert werden können. Man spricht dann von *DOOD–Systemen*, d.h. *deduktiven und objektorientierten Datenbanksystemen.*

Vorteile des deduktiven Ansatzes

In den meisten Datenbankanwendungen sind intensionale Spezifikationen konzeptuell allgegenwärtig. Sie werden in nicht-deduktiven Datenbanken durch extensionale, d.h. explizite, Spezifikationen von Tupeln oder Objekten sowie durch spezielle Anwendungsprogramme implementiert, die für jede Anwendung gesondert entwickelt werden müssen. Dabei zeigen sich folgende Nachteile, die von deduktiven Datenbankverwaltungssystemen vermieden werden:

- Abweichung von den tatsächlichen Spezifikationen,
- unsystematische Programmentwicklung,
- stark von der Programmierung abhängende Effizienz,
- Verwaltung der Programme außerhalb der Datenbank,
- imperative statt deklarative Spezifikationen.

Aus der Sicht relationaler Datenbanken ist die Erweiterung einer Datenbank um Ableitunsgregeln und (uneingeschränkte) Integritätsbedingungen aus mehereren Gründen natürlich. Zum einen stellen Ableitunsgregeln lediglich eine Verallgemeinerung des im relationalen Modell vorhandenen Begriffes der "Sicht" (view) dar. Zum anderen war ein allgemeiner Ansatz für Integritätsbedingungen bereits ein Ziel der relationalen Datenbanken, das bisher nur teilweise realisiert worden ist.

Intensionale Darstellungen haben nicht nur den Vorteil konzeptuell natürlicher zu sein als traditionelle, extensionale Datenbankspezifikationen. Sie haben den zusätzlichen Vorteil, oft eine wesentliche *Datenkompaktierung* zu ermöglichen. Kompaktierungsfaktoren von 10 bis mehrere 100 sind in Anwendungen nicht selten, die – wie etwa Fahrpläne – eine reguläre Stuktur aufweisen. Weniger Platz ist notwendig, um z.B. die folgende Regel zu spezifizieren:

> *"12 Minuten nach jeder vollen Stunde gibt es einen Flug von München nach Frankfurt/Main",*

als die Angaben zu allen solchen Flügen mehrfach zu speichern.

Anwendungsbereiche

Klassiche Datenbanken, wie etwa Verwaltungsdatenbanken, werden intern in Unternehmen eingesetzt, z.B. zur Verwaltung von Personaldaten, Bankkonten oder Versicherungspolicen. Während des letzten Jahrzehnts wurden zunehmend *Informationssysteme für ein breites Publikum* angeboten, z.B. öffentliche Verkehrsinformationssysteme. Weil sie eine *deklarative Modellierung* ermöglichen, sind für solche Anwendungen deduktive sowie relationale Datenbanksysteme sehr passend.

Neuerdings setzen sich Informationssysteme durch – auch übers Internet zugreifbar – die nicht nur Auskünfte geben, sondern auch *komplexe Transaktionen* wie etwa Buchungen von Flugscheinen oder Bestellungen bei einem Versandhaus ermöglichen. Für Anwendungen aus diesem Bereich erweist sich der *deklarative Ansatz* der deduktiven Datenbanksysteme als besonders geeignet.

Auch für neuere Datenbankanwendungen wie etwa im Bereich des *Computer–Aided–Designs* (CAD) oder der *Entscheidungsunterstützung* (z.B. *Diagnosesysteme* und *Expertensysteme*) sind deduktive Datenbanksysteme besonders geeignet.

Darüber hinaus sind wegen ihrer Behandlungung von uneingeschränkten Integritätsbedingungen deduktive Datenbanksysteme im traditionellen Bereich der Verwaltungdatenbanken auch sehr nützlich. Regeln erleichtern z.B. die Spezifikation der Aufnahme in Sperrlisten bei Kreditkartenbanken.

Letztlich erleichtern die deduktiven Datenbanksysteme die *Migration* einer Datenbank von einem – z.B. relationalen – Datenmodell in ein anderes – z.B. objektorientiertes – Modell, weil sich solche "Übersetzungen" besonders leicht und natürlich mit Regeln deklarativ spezifizieren lassen. Die Migration einer relationalen in eine deduktive Datenbank ist auch möglich, weil eine relationale Datenbank letztendlich ein Spezialfall einer deduktiven Datenbank ohne Regeln darstellt. Allerdings ermöglicht erst die Neumodellierung von relational, d.h. *extensional*, erfaßten Daten in deduktiv, d.h. *intensional*, spezifizierte Daten die volle Nutzung der Vorteile des deduktiven Datenbankansatzes.

Aktuelle Forschungsthemen

Neuere Anwendungen für deduktive Datenbanken erfordern erweiterte Möglichkeiten zur Repräsentation von Wissen. Zur Zeit wird deshalb intensiv an der Erweiterung deduktiver Datenbanken um eine uneingeschränkte *Negation als Scheitern* und um verschiedenen Formen von *unsicherem Wissen*, wie etwa in *disjunktiven Datenbanken*, gearbeitet. Für viele Anwendungen ist die *deklarative* Spezifikation von *Datenbankänderungen* sehr wichtig. Dies hat zu Forschungsaktivitäten im Bereich der *aktiven Datenbanken* geführt, in denen Regelformalismen zur Definition von Änderungen verwendet werden. Auch neueste Datenbankthemen, wie z.B. *Data Mining*, *Knowledge Discovery* und *Constraint–Datenbanken*, bei denen spezielle Theorien zur Spezifikationen von besonderen Daten wie etwa Intervalle oder numerische Werte eingesetzt werden, nutzen oder erweitern Methoden der deduktiven Datenbanken.

Die in deduktiven Datenbanken eingeführten Erweiterungen haben für die Datenbankforschung einen Sprung ins Unbekannte dargestellt. Zum ersten Mal wurde ein Datenbankkonzept vorgeschlagen, das genauso wie eine Programmiersprache uneingeschränkte Berechnungen ermöglicht, weil die aus Fakten und Ableitungsregeln bestehenden Sprachen *Turing–vollständig* sind. Daran mag es liegen, daß die neue, zukunftsweisende Technologie der deduktiven Datenbanken sich nur langsam außerhalb der Forschungskreise etabliert. Die kürzliche Aufnahme von *rekursiven Sichten* im neuen Standard SQL3 wird in Fachkreisen als Zeichen des Zusammenrückens von traditionellen und deduktiven Datenbanksystemen angesehen.

Literatur

[BMS 96] *F. Bry, R. Manthey, H. Schütz:* Deduktive Datenbanken, KI – Künstliche Intelligenz – Forschung, Entwicklung, Erfahrungen, wird erscheinen.

[CGH 94] *A. Cremers, U. Griefahn, R. Hinze:* Deduktive Datenbanken, Vieweg, 1994.

[Da 92] *S.K. Das:* Deductive Databases and Logic Programming, Addison–Wesley, 1992.

[KieGü 90] *W. Kießling, U. Güntzer:* Deduktive Datenbanken auf dem Weg zur Praxis, Informatik Forschung und Entwicklung, Heft 5, 1990, S. 177-187.

[VLDB 94] Special Issue on *Prototypes of Deductive Database Systems*, The VLDB Journal, vol. 3(2), 1994.