



Forschungsprofil des Lehrstuhls „Datenbank- und Informationssysteme“

Am Lehrstuhl gibt es die folgenden zwei Forschungsschwerpunkte:

1. Integration von Retrieval-Suchtechniken in ein Datenbankanfragesystem und
2. Metrikindexe zur effizienten Ähnlichkeitssuche.

Datenbanksysteme verwalten heutzutage neben herkömmlichen alphanumerischen, relational strukturierten Daten XML-Daten, Multimedia-Objekte, Texte sowie räumliche und zeitbehaftete Daten. Bisherige Datenbankanfragesprachen geraten an ihre Grenzen, wenn unscharfe Suchbedingungen, wie sie etwa von Retrieval-Systemen her bekannt sind, in eine Anfrage integriert werden sollen. So sollte etwa die Suche in XML-Dokumenten eine Text-Retrieval-ähnliche Suche unterstützen. Auch eine inhaltsbasierte Suche in Multimedia-Daten erfordert die Integration von Retrieval-Mechanismen in Datenbankanfragesprachen. Da die meisten Datenbankanfragesprachen auf der Booleschen Logik aufbauen, können diese nicht adäquat mit unscharfen Suchbedingungen umgehen. Statt dessen werden solche Bedingungen in der Regel durch Vergleich mit einem Schwellenwert in Boolesche Bedingungen unter Informationsverlust umgeformt. Eine adäquate Integration von Retrieval-Bedingungen erfordert aber die Möglichkeit der freien Kombination beider Arten von Bedingungen mittels einer Logik, welche die Boolesche Logik selbst nicht sein kann. Gesucht ist also ein Formalismus, welcher eine geeignete Erweiterung der Booleschen Logik bildet. An der Universität Magdeburg wurde daher ein Fuzzy-Logik-basiertes Anfragesystem entwickelt. Die Gesetze der Booleschen Algebra sind dort jedoch nur mit der t-Norm/Conorm \min/\max erfüllbar. Diese Norm weist jedoch Nachteile auf, etwa ist das Minimum zweier Werte unsensibel bezüglich Werteänderungen des größeren Eingangswertes. Dies widerspricht der menschlichen Intuition für eine Konjunktion. Ein neuer Ansatz liegt in der Verwendung des Formalismus der Quantenlogik. Dieser behebt die bestehenden Probleme und ist aus datenbanktheoretischer Sicht von Interesse, da sich die Anfrageverarbeitung direkt in das Gebiet der linearen Algebra überführen lässt und sich daher neue Impulse für theoretische Datenbankprobleme ergeben. An der BTU Cottbus soll dieser Ansatz in Richtung einsetzbare Anfragesysteme weiterentwickelt werden. Dabei stehen Probleme der Anbindung an bestehende Sprachen, Optimierung und Anfrageformulierung als zukünftige Forschungsthemen im Vordergrund.

Ein zweiter Forschungsschwerpunkt bildet die Indexunterstützung für Metriken zur schnellen Ähnlichkeitsberechnung. Bei einer Metrik wird von einer Menge von Objekten mit einer Distanzfunktion ausgegangen. Für den Index ist die genaue Kenntnis der Distanzfunktion nicht erforderlich. Ein Metrikindeks ist damit flexibel bezüglich der Distanzfunktion und der beteiligten Objekte, ist also im Gegensatz zu vielen Indexstrukturen nicht auf die Euklidische Distanzfunktion zur Ähnlichkeitsberechnung beschränkt. Eine neue Idee liegt im Ansatz von

selbstverfeinernden Distanzen bei der Nächsten-Nachbarsuche. In Zukunft soll am Lehrstuhl diese Idee ausgebaut werden.

Ingo Schmitt, Cottbus am 29.03.2007