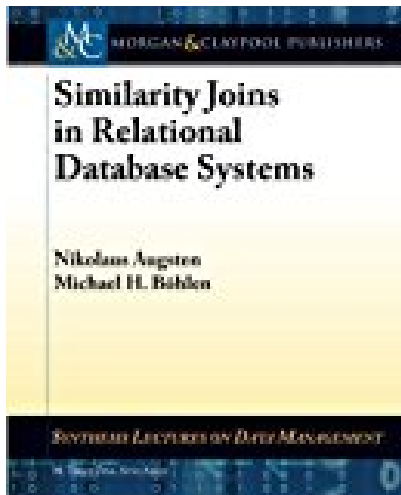


# Similarity Joins in Relational Database Systems Synthesis Lectures on Data Management

---



## BOOK DETAILS

- Author : Nikolaus Augsten
- Pages : 124 Pages
- Publisher : Morgan & Claypool Publishers
- Language : English
- ISBN : 1627050280

[↓ DOWNLOAD](#)

## **BOOK SYNOPSIS**

State-of-the-art database systems manage and process a variety of complex objects, including strings and trees. For such objects equality comparisons are often not meaningful and must be replaced by similarity comparisons. This book describes the concepts and techniques to incorporate similarity into database systems. We start out by discussing the properties of strings and trees, and identify the edit distance as the de facto standard for comparing complex objects. Since the edit distance is computationally expensive, token-based distances have been introduced to speed up edit distance computations. The basic idea is to decompose complex objects into sets of tokens that can be compared efficiently. Token-based distances are used to compute an approximation of the edit distance and prune expensive edit distance calculations. A key observation when computing similarity joins is that many of the object pairs, for which the similarity is computed, are very different from each other. Filters exploit this property to improve the performance of similarity joins. A filter preprocesses the input data sets and produces a set of candidate pairs. The distance function is evaluated on the candidate pairs only. We describe the essential query processing techniques for filters based on lower and upper bounds. For token equality joins we describe prefix, size, positional and partitioning filters, which can be used to avoid the computation of small intersections that are not needed since the similarity would be too low.

### **SIMILARITY JOINS IN RELATIONAL DATABASE SYSTEMS SYNTHESIS**

**LECTURES ON DATA MANAGEMENT** - Are you looking for Ebook Similarity Joins In Relational Database Systems Synthesis Lectures On Data Management ? You will be glad to know that right now Similarity Joins In Relational Database Systems Synthesis Lectures On Data Management is available on our online library. With our online resources, you can find Applied Numerical Methods With Matlab Solution Manual 3rd Edition or just about any type of ebooks, for any type of product.

Best of all, they are entirely free to find, use and download, so there is no cost or stress at all. Similarity Joins In Relational Database Systems Synthesis Lectures On Data Management may not make exciting reading, but Applied Numerical Methods With Matlab Solution Manual 3rd Edition is packed with valuable instructions, information and warnings. We also have many ebooks and user guide is also related with Similarity Joins In Relational Database Systems Synthesis Lectures On Data Management and many other ebooks.

We have made it easy for you to find a PDF Ebooks without any digging. And by having access to our ebooks online or by storing it on your computer, you have convenient answers with Similarity Joins In Relational Database Systems Synthesis Lectures On Data Management . To get started finding Similarity Joins In Relational Database Systems Synthesis Lectures On Data Management , you are right to find our website which has a comprehensive collection of manuals listed.