

Prof. Dr. T. Härder  
 Fachbereich Informatik  
 AG Datenbanken und Informationssysteme  
 Universität Kaiserslautern

## Übungsblatt 7

für die freiwillige Übung

**Unterlagen zur Vorlesung:** „[www.dvs.informatik.uni-kl.de/courses/DBSREAL/](http://www.dvs.informatik.uni-kl.de/courses/DBSREAL/)“

**Aufgabe 1: Analyse einer Hash-Funktion** Die Schlüssel zur Identifikation der Datensätze einer Anwendung seien als CHAR(8) im EBCDIC-Code definiert. Sie werden typischerweise als Zahlen in festen Abständen, also 10, 20, 30, ..., vergeben. Vor Verwendung in einer Hash-Funktion werden diese Schlüssel erst nach folgendem Schema gefaltet und dabei mit EXOR verknüpft. In hexadezimaler Darstellung (im EBCDIC-Code) ergibt sich beispielsweise nach Faltung für den Schlüssel  $S_1$  mit 0000 0010 der Wert von  $K_1$  mit

$$\begin{array}{r}
 \text{F 0 F 0 F 0 F 0} \\
 \text{F 0 F 0 F 1 F 0} \\
 \hline
 \text{0 0 0 0 0 1 0 0}_{16}
 \end{array}
 \oplus
 \hat{=}
 256_{10}$$

Danach werde das Divisionsrest-Verfahren mit

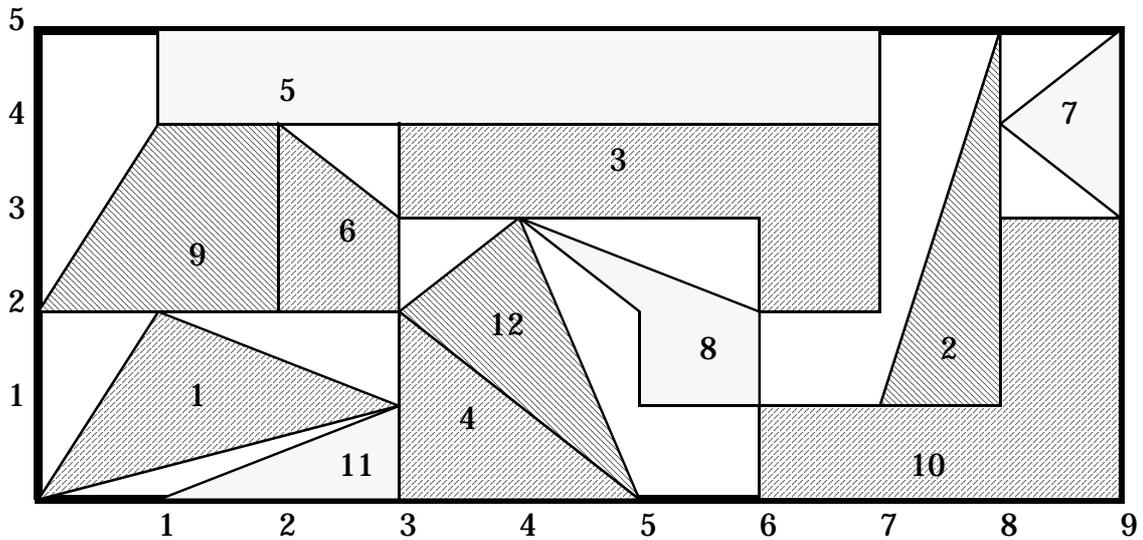
$$h(K_i) = K_i \bmod n$$

angewendet.

Analysieren Sie das Kollisionsverhalten dieser Funktion bei obiger Schlüsselvergabe und einem Hash-Bereich von  $n = 576$  Buckets.

- Nach welchen Abständen von  $j$  vergebenen Schlüsseln, die sich jeweils mit einem Inkrement von  $\Delta k$  (was dem Wert von  $K_1$  entspricht) erhöhen, treten Kollisionen auf?
- Was ist eine effektive Abhilfemaßnahme?





b.) Schreiben Sie die Funktionen

FINDE\_ALLE\_VOLLSTAENDIG\_ENTHALTENEN\_FLAECHEN ( Rechteck )

FINDE\_ALLE\_TEILWEISE\_ENTHALTENEN\_FLAECHEN ( Rechteck )

die alle Flächen findet, die vollständig resp. teilweise in dem gegebenen Rechteck enthalten sind.

c.) Testen Sie die Funktionen mit dem Rechteck, daß durch die Eckpunkte (3,1) und (7,4) beschrieben wird.

d.) Welche Verbesserung des R-Baums erlaubt eine einfachere Suche?