
Algorithmen und Datenstrukturen

Sortieren zusammengesetzter Objekte nach mehreren Kriterien

Sortieren großer Objekte: Radix Sort

Prof. Dr. Ralf Möller

Universität zu Lübeck

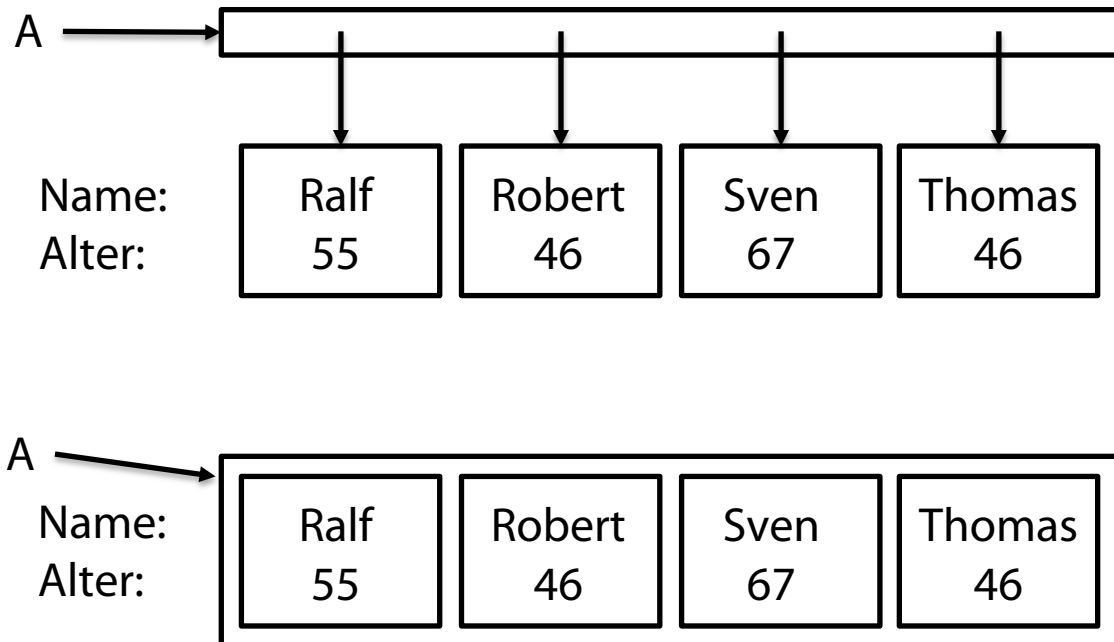
Institut für Informationssysteme

Felix Kuhr (Übungen)

sowie viele Tutoren

Sortieren nach mehreren Kriterien

- Zusammengesetzte Objekte in einem Feld



Name		Address		
Last	First	Street	City	State
Bayless	Andrew	West Ave	Houston	TX
Benitez	Michael	North Ave	Los Angeles	CA
Chu	Henry	East Ave	San Diego	CA
Dangelo	David	Third St	Detroit	MI
Dawood	Hussam	Lincoln Rd	Detroit	MI
Devineni	Soujanya	Northwestern Ave	Houston	TX
Dunne	Brendan	EAST AVE.	Dallas	TX
Edwards	Brian	De Zavala Rd	San Antonio	TX
Godfrey	Daryl	MAIN ST	Austin	TX
Guerra	John	DALLAS AVE.	Austin	TX
Guevara	Clovis	University Pkwy	San Antonio	TX
Halbeisen	Gerald	FOREST CIRCLE	Los Angeles	CA
Hawkins	Richard	RIVERSIDE ST	San Antonio	TX
Hernandez	Monica	COLLEGE PKWY	San Antonio	TX
Hohmann	Shawn	COLLEGE PKWY	Cleveland	OH
Honeycutt	Richard	Southwest Ave	San Antonio	TX
Martinez	Juan	OAK CLIFF	Pheonix	AZ
Mayo	Nathan	UTSA BLVD	San Antonio	TX
Mirabal	Renato	FIRST ST	Columbus	OH
Modebe	Nnaemeka	RIVERSIDE ST	Houston	TX
Moon	Ryan	EAST AVE.	Madison	WI

Name		Address		
Last	First	Street	City	State
Martinez	Juan	OAK CLIFF	Pheonix	AZ
Benitez	Michael	North Ave	Los Angeles	CA
Chu	Henry	East Ave	San Diego	CA
Halbeisen	Gerald	FOREST CIRCLE	Los Angeles	CA
Dangelo	David	Third St	Detroit	MI
Dawood	Hussam	Lincoln Rd	Detroit	MI
Hohmann	Shawn	COLLEGE PKWY	Cleveland	OH
Mirabal	Renato	FIRST ST	Columbus	OH
Bayless	Andrew	West Ave	Houston	TX
Devineni	Soujanya	Northwestern Ave	Houston	TX
Dunne	Brendan	EAST AVE.	Dallas	TX
Edwards	Brian	De Zavala Rd	San Antonio	TX
Godfrey	Daryl	MAIN ST	Austin	TX
Guerra	John	DALLAS AVE.	Austin	TX
Guevara	Clovis	University Pkwy	San Antonio	TX
Hawkins	Richard	RIVERSIDE ST	San Antonio	TX
Hernandez	Monica	COLLEGE PKWY	San Antonio	TX
Honeycutt	Richard	Southwest Ave	San Antonio	TX
Mayo	Nathan	UTSA BLVD	San Antonio	TX
Modebe	Nnaemeka	RIVERSIDE ST	Houston	TX
Moon	Ryan	EAST AVE.	Madison	WI

Name		Address		
Last	First	Street	City	State
Godfrey	Daryl	MAIN ST	Austin	TX
Guerra	John	DALLAS AVE.	Austin	TX
Hohmann	Shawn	COLLEGE PKWY	Cleveland	OH
Mirabal	Renato	FIRST ST	Columbus	OH
Dunne	Brendan	EAST AVE.	Dallas	TX
Dangelo	David	Third St	Detroit	MI
Dawood	Hussam	Lincoln Rd	Detroit	MI
Bayless	Andrew	West Ave	Houston	TX
Devineni	Soujanya	Northwestern Ave	Houston	TX
Modebe	Nnaemeka	RIVERSIDE ST	Houston	TX
Benitez	Michael	North Ave	Los Angeles	CA
Halbeisen	Gerald	FOREST CIRCLE	Los Angeles	CA
Moon	Ryan	EAST AVE.	Madison	WI
Martinez	Juan	OAK CLIFF	Pheonix	AZ
Edwards	Brian	De Zavala Rd	San Antonio	TX
Guevara	Clovis	University Pkwy	San Antonio	TX
Hawkins	Richard	RIVERSIDE ST	San Antonio	TX
Hernandez	Monica	COLLEGE PKWY	San Antonio	TX
Honeycutt	Richard	Southwest Ave	San Antonio	TX
Mayo	Nathan	UTSA BLVD	San Antonio	TX
Chu	Henry	East Ave	San Diego	CA

Name		Address		
Last	First	Street	City	State
Hohmann	Shawn	COLLEGE PKWY	Cleveland	OH
Hernandez	Monica	COLLEGE PKWY	San Antonio	TX
Guerra	John	DALLAS AVE.	Austin	TX
Edwards	Brian	De Zavala Rd	San Antonio	TX
Chu	Henry	East Ave	San Diego	CA
Dunne	Brendan	EAST AVE.	Dallas	TX
Moon	Ryan	EAST AVE.	Madison	WI
Mirabal	Renato	FIRST ST	Columbus	OH
Halbeisen	Gerald	FOREST CIRCLE	Los Angeles	CA
Dawood	Hussam	Lincoln Rd	Detroit	MI
Godfrey	Daryl	MAIN ST	Austin	TX
Benitez	Michael	North Ave	Los Angeles	CA
Devineni	Soujanya	Northwestern Ave	Houston	TX
Martinez	Juan	OAK CLIFF	Pheonix	AZ
Modebe	Nnaemeka	RIVERSIDE ST	Houston	TX
Hawkins	Richard	RIVERSIDE ST	San Antonio	TX
Honeycutt	Richard	Southwest Ave	San Antonio	TX
Dangelo	David	Third St	Detroit	MI
Guevara	Clovis	University Pkwy	San Antonio	TX
Mayo	Nathan	UTSA BLVD	San Antonio	TX
Bayless	Andrew	West Ave	Houston	TX

Name		Address		
Last	First	Street	City	State
Bayless	Andrew	West Ave	Houston	TX
Benitez	Michael	North Ave	Los Angeles	CA
Chu	Henry	East Ave	San Diego	CA
Dangelo	David	Third St	Detroit	MI
Dawood	Hussam	Lincoln Rd	Detroit	MI
Devineni	Soujanya	Northwestern Ave	Houston	TX
Dunne	Brendan	EAST AVE.	Dallas	TX
Edwards	Brian	De Zavala Rd	San Antonio	TX
Godfrey	Daryl	MAIN ST	Austin	TX
Guerra	John	DALLAS AVE.	Austin	TX
Guevara	Clovis	University Pkwy	San Antonio	TX
Halbeisen	Gerald	FOREST CIRCLE	Los Angeles	CA
Hawkins	Richard	RIVERSIDE ST	San Antonio	TX
Hernandez	Monica	COLLEGE PKWY	San Antonio	TX
Hohmann	Shawn	COLLEGE PKWY	Cleveland	OH
Honeycutt	Richard	Southwest Ave	San Antonio	TX
Martinez	Juan	OAK CLIFF	Pheonix	AZ
Mayo	Nathan	UTSA BLVD	San Antonio	TX
Mirabal	Renato	FIRST ST	Columbus	OH
Modebe	Nnaemeka	RIVERSIDE ST	Houston	TX
Moon	Ryan	EAST AVE.	Madison	WI

Name		Address		
Last	First	Street	City	State
Hernandez	Monica	COLLEGE PKWY	San Antonio	TX
Hohmann	Shawn	COLLEGE PKWY	Cleveland	OH
Guerra	John	DALLAS AVE.	Austin	TX
Edwards	Brian	De Zavala Rd	San Antonio	TX
Chu	Henry	East Ave	San Diego	CA
Dunne	Brendan	EAST AVE.	Dallas	TX
Moon	Ryan	EAST AVE.	Madison	WI
Mirabal	Renato	FIRST ST	Columbus	OH
Halbeisen	Gerald	FOREST CIRCLE	Los Angeles	CA
Dawood	Hussam	Lincoln Rd	Detroit	MI
Godfrey	Daryl	MAIN ST	Austin	TX
Benitez	Michael	North Ave	Los Angeles	CA
Devineni	Soujanya	Northwestern Ave	Houston	TX
Martinez	Juan	OAK CLIFF	Pheonix	AZ
Hawkins	Richard	RIVERSIDE ST	San Antonio	TX
Modebe	Nnaemeka	RIVERSIDE ST	Houston	TX
Honeycutt	Richard	Southwest Ave	San Antonio	TX
Dangelo	David	Third St	Detroit	MI
Guevara	Clovis	University Pkwy	San Antonio	TX
Mayo	Nathan	UTSA BLVD	San Antonio	TX
Bayless	Andrew	West Ave	Houston	TX

Name		Address		
Last	First	Street	City	State
Guerra	John	DALLAS AVE.	Austin	TX
Godfrey	Daryl	MAIN ST	Austin	TX
Hohmann	Shawn	COLLEGE PKWY	Cleveland	OH
Mirabal	Renato	FIRST ST	Columbus	OH
Dunne	Brendan	EAST AVE.	Dallas	TX
Dawood	Hussam	Lincoln Rd	Detroit	MI
Dangelo	David	Third St	Detroit	MI
Devineni	Soujanya	Northwestern Ave	Houston	TX
Modebe	Nnaemeka	RIVERSIDE ST	Houston	TX
Bayless	Andrew	West Ave	Houston	TX
Halbeisen	Gerald	FOREST CIRCLE	Los Angeles	CA
Benitez	Michael	North Ave	Los Angeles	CA
Moon	Ryan	EAST AVE.	Madison	WI
Martinez	Juan	OAK CLIFF	Pheonix	AZ
Hernandez	Monica	COLLEGE PKWY	San Antonio	TX
Edwards	Brian	De Zavala Rd	San Antonio	TX
Hawkins	Richard	RIVERSIDE ST	San Antonio	TX
Honeycutt	Richard	Southwest Ave	San Antonio	TX
Guevara	Clovis	University Pkwy	San Antonio	TX
Mayo	Nathan	UTSA BLVD	San Antonio	TX
Chu	Henry	East Ave	San Diego	CA

Name		Address		
Last	First	Street	City	State
Martinez	Juan	OAK CLIFF	Pheonix	AZ
Halbeisen	Gerald	FOREST CIRCLE	Los Angeles	CA
Benitez	Michael	North Ave	Los Angeles	CA
Chu	Henry	East Ave	San Diego	CA
Dawood	Hussam	Lincoln Rd	Detroit	MI
Dangelo	David	Third St	Detroit	MI
Hohmann	Shawn	COLLEGE PKWY	Cleveland	OH
Mirabal	Renato	FIRST ST	Columbus	OH
Guerra	John	DALLAS AVE.	Austin	TX
Godfrey	Daryl	MAIN ST	Austin	TX
Dunne	Brendan	EAST AVE.	Dallas	TX
Devineni	Soujanya	Northwestern Ave	Houston	TX
Modebe	Nnaemeka	RIVERSIDE ST	Houston	TX
Bayless	Andrew	West Ave	Houston	TX
Hernandez	Monica	COLLEGE PKWY	San Antonio	TX
Edwards	Brian	De Zavala Rd	San Antonio	TX
Hawkins	Richard	RIVERSIDE ST	San Antonio	TX
Honeycutt	Richard	Southwest Ave	San Antonio	TX
Guevara	Clovis	University Pkwy	San Antonio	TX
Mayo	Nathan	UTSA BLVD	San Antonio	TX
Moon	Ryan	EAST AVE.	Madison	WI

Aufgabe: Sortiere Studierende nach (state, city, street).

Name		Address			
Last	First	Street	City	State	
Martinez	Juan	OAK CLIFF	Pheonix	AZ	
Halbeisen	Gerald	FOREST CIRCLE	Los Angeles	CA	
Benitez	Michael	North Ave	Los Angeles	CA	
Chu	Henry	East Ave	San Diego	CA	
Dawood	Hussam	Lincoln Rd	Detroit	MI	
Dangelo	David	Third St	Detroit	MI	
Hohmann	Shawn	COLLEGE PKWY	Cleveland	OH	
Mirabal	Renato	FIRST ST	Columbus	OH	
Guerra	John	DALLAS AVE.	Austin	TX	
Godfrey	Daryl	MAIN ST	Austin	TX	
Dunne	Brendan	EAST AVE.	Dallas	TX	
Devineni	Soujanya	Northwestern Ave	Houston	TX	
Modebe	Nnaemeka	RIVERSIDE ST	Houston	TX	
Bayless	Andrew	West Ave	Houston	TX	
Hernandez	Monica	COLLEGE PKWY	San Antonio	TX	
Edwards	Brian	De Zavala Rd	San Antonio	TX	
Hawkins	Richard	RIVERSIDE ST	San Antonio	TX	
Honeycutt	Richard	Southwest Ave	San Antonio	TX	
Guevara	Clovis	University Pkwy	San Antonio	TX	
Mayo	Nathan	UTSA BLVD	San Antonio	TX	
Moon	Ryan	EAST AVE.	Madison	WI	

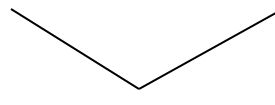
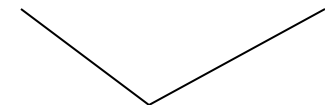
Aufgabe: Sortiere Studierende nach (state, city, street).

Stabiles Sortieren

- Die meisten $O(n^2)$ -Sortialgorithmen sind stabil
 - oder können “einfach” stabil gemacht werden
- $O(n \log n)$ -Sortialgorithmen sind nicht immer stabil
 - Ist Heap-Sort stabil?
- Generischer Ansatz, Stabilität zu erzeugen:
 - Verwende zwei Schlüssel,
der zweite ist der originale Index des Elements
 - Wenn zwei Elemente gleich, vergleiche zweite Schlüsselkomponenten

[5, 6, 5, 1, 2, 3, 2, 6]

[(5, 1), (6, 2), (5, 3), (1, 4), (2, 5), (3, 6), (2, 7), (6, 8)]



Wie kann man sehr große Zahlen sortieren?

198099109123518183599
340199540380128115295
384700101594539614696
382408360201039258538
614386507628681328936
738148652090990369197
987084087096653020299
185664124421234516454
785392075747859131885
530995223593137397354
267057490443618111767
795293581914837377527
815501764221221110674
142522204403312937607
718098797338329180836
856504702326654684056
982119770959427525245
528076153239047050820
305445639847201611168
478334240651199238019

Jede Zeile sei eine
"lange" Zahl, eine
Gensequenz oder....

Zahlen dieser Art sind zu groß für den Integer-Datentyp, sie werden als "Zeichenkette" repräsentiert

Verwende vergleichsbasiertes Sortieren mit einer Zeichenkettenvergleichsfunktion

if $A[i] < A[j]$ wird zu $\text{if vergleiche}(A[i], A[j]) < 0$ mit

```
function vergleiche(s, t)
  for i = 1 to min( { length(s), length(t) } ) do
    if s[i] < t[i] then
      return -1
    else if s[i] > t[i] then
      return 1
  if length(s) = length(t)
    then return 0
  else if length(s) < length(t)
    then return -1
  else return 1
```

Kosten des Vergleichs von zwei Zeichenketten der Länge d? $O(d)$
Gesamtkosten: $O(d \cdot n \log n)$

Radix-Sort

- Ähnlich wie das Sortieren von Adressbüchern
- Handle jede Zahl als Sortierschlüssel
- Starte vom Least-significant-Bit

Most significant



Least significant



198099109123518183599
340199540380128115295
384700101594539614696
382408360201039258538
614386507628681328936

Jede Zeile entspricht
einer "langen" Zahl

Radix-Sort wurde schon **1887** in Arbeiten von
Herman Hollerith zu Volkszählungsmaschinen verwendet

Radix-Sort: Illustration

- Hier ein vereinfachtes Beispiel:

Jede Zeile entspricht einer
langen Zahl

7 4 2
7 4 8
0 5 4
6 8 8
4 1 2
2 3 0
9 3 5
1 1 6
1 6 1
4 3 4
3 8 5
6 6 6
0 3 1
0 1 3
3 6 5
1 7 3
0 1 6

Radix-Sort: Illustration

- Sortiere nach letzter Zahl:



2 3 0
1 6 1
0 3 1
7 4 2
4 1 2
0 1 3
1 7 3
0 5 4
4 3 4
9 3 5
3 8 5
3 6 5
1 1 6
6 6 6
0 1 6
7 4 8
6 8 8

Radix-Sort: Illustration

- Sortiere nach zweitletzter Zahl:

↓

4	1	2
0	1	3
1	1	6
0	1	6
2	3	0
0	3	1
4	3	4
9	3	5
7	4	2
7	4	8
0	5	4
1	6	1
3	6	5
6	6	6
1	7	3
3	8	5
6	8	8

Radix-Sort: Illustration

- Sortiere nach erster Zahl:

↓

<u>0</u>	<u>1</u>	<u>3</u>
<u>0</u>	<u>1</u>	<u>6</u>
<u>0</u>	<u>3</u>	<u>1</u>
<u>0</u>	<u>5</u>	<u>4</u>
<u>1</u>	<u>1</u>	<u>6</u>
<u>1</u>	<u>6</u>	<u>1</u>
<u>1</u>	<u>7</u>	<u>3</u>
<u>2</u>	<u>3</u>	<u>0</u>
<u>3</u>	<u>6</u>	<u>5</u>
<u>3</u>	<u>8</u>	<u>5</u>
<u>4</u>	<u>1</u>	<u>2</u>
<u>4</u>	<u>3</u>	<u>4</u>
<u>6</u>	<u>6</u>	<u>6</u>
<u>6</u>	<u>8</u>	<u>8</u>
<u>7</u>	<u>4</u>	<u>2</u>
<u>7</u>	<u>4</u>	<u>8</u>
<u>9</u>	<u>3</u>	<u>5</u>

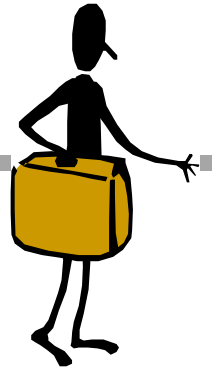
Zeitkomplexität

- Sortierung jeder der d Spalten mit Counting-Sort
- Gesamtkosten: $d(n + k)$
 - Wähle $k = 10$
 - Gesamtkosten: $\Theta(dn)$
- Partitionierung der d Zahlen in z.B. in Dreiergruppen
 - Gesamtkosten: $(n + 10^3)d/3$
- Wir arbeiten mit Binärzahlen anstelle von Dezimalen
 - Partitionierung der d Bits in Gruppen von r Bits
 - Gesamtkosten: $(n + 2^r)d/r$
 - Wähle $r = \log n$
 - Gesamtkosten: $O(dn / \log n)$
 - Vergleiche mit $O(dn \log n)$ für das einfache Verfahren
- Aber: Radix-Sort hat hohen konstanten Faktor

Platzkomplexität

- Verwendung von Counting-Sort
- Daher zusätzlicher Speicher nötig: $\theta(n)$

Zusammenfassung



- Bisher behandelt:
 - Sortieren zusammengesetzter Objekte nach mehreren Kriterien
 - Sortieren großer Objekte: Radix Sort