Basic Concepts of a Database System

- Database

  A collection of interrelated data accessible by multiple users or multiple purposes

- Database Management System(DBMS)

  Software that allows one or many persons to use and /or modify data

- Database system = Database + DBMS

Schemes and Instances

- Scheme of a database

  => structure of a database

  (structural skeleton)
- Instance of a database

  =>current content of the database
- Programming Language analogy

  type customer = Record {…..}

  var customer1: customer

Data Model

- A formalism for describing the logical structure of a database and operation on the database
- Programing language analogy:

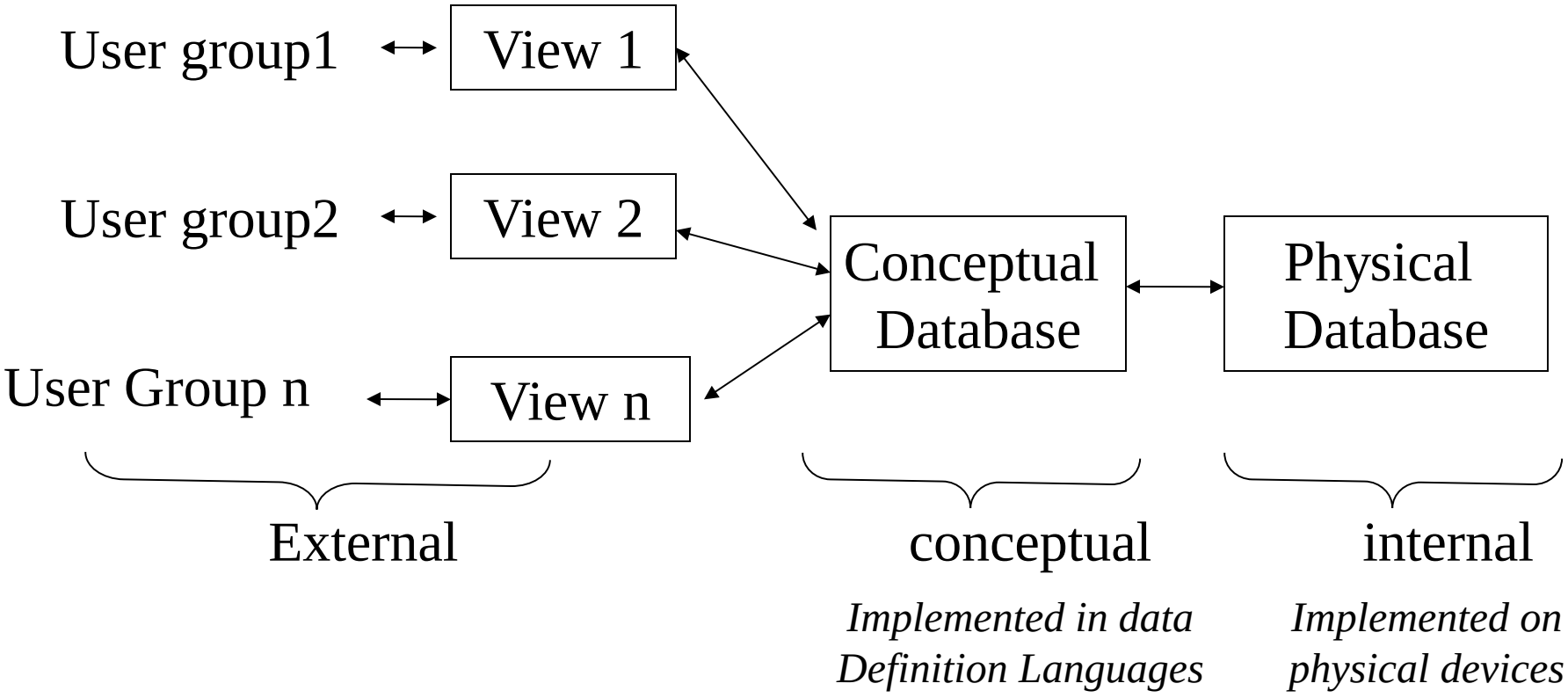
  var customer1: customer;

  ↑ ↑

  instance scheme

  PASCAL(programming Language)≡data model

Levels of Abstraction in DBMS



Database Languages

- Data Definition Language (DDL)

  to describe a scheme of a database
- Data Manipulation Language(DML)

  to manipulate (retrieve, insert, delete & modify) a database instance

  →non-procedural (declarative)

  →procedural
- Query Language

  Interactive DML

- Host Language

  Programming language in which Statements in a DML can be embedded (e.g. C)

Classification of DBMS

- From viewpoint of Data Models:
  - simple flat tables
  - Hierarchical DBMS
  - Network DBMS
  - Relational DBMS
- From a viewpoint of Control:
  - Centralized DBMS
  - Distributed DBMS (DDBMS)
    - heterogeneous DDBMS
    - homogeneous DDBMS

Important Database Properties

- Data Abstraction

  (hide storage detail from user)

  $\rightarrow$ current trend : operation abstraction in object-oriented systems

- Security and Authorization

  (file systems ok for all/nothing access)

  But often want to grant selected field access

  Grant student.advisor ACCESS(+R) to student.grade

- Control of Redundancy

1. Duplication of effort — repetition of same data in multiple
2. Waste of storage space — places
3. May lead to inconsistencies

change (phote #) one place, change everywhere?

(Date of Birth)

situation with

Rosie Donnaldson	8/26/65	TA
Rose Donnaldson	8/26/65	Student Record

Controlled Redundancy may be useful.

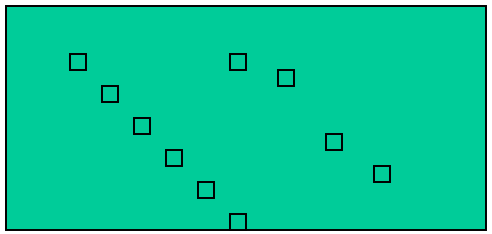
$\rightarrow$case where difference values of field for different recs

A Brief history of Database Technology

- Flat Databases

early 1900's – The Punch card

- fixed fields for storing data
- (initially) 1 record per card

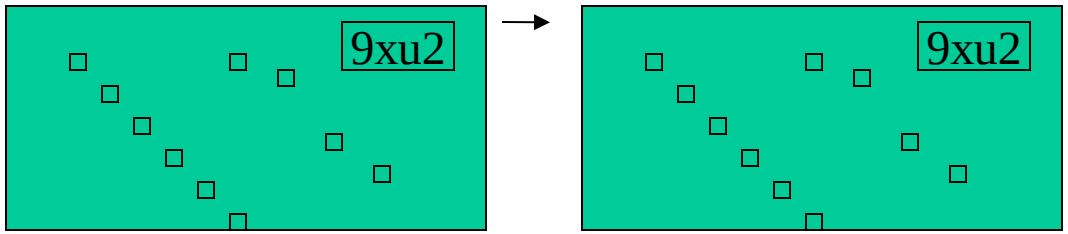


1945 – Magnetic Tape

(Punchcards on tape, but allows fasted search + sorting)

Flat databases motivated by Punch cards

- The Record key – allowed flat records to continue on multiple punchcards (facilitated sorting)



(Wider records on magnetic tape)

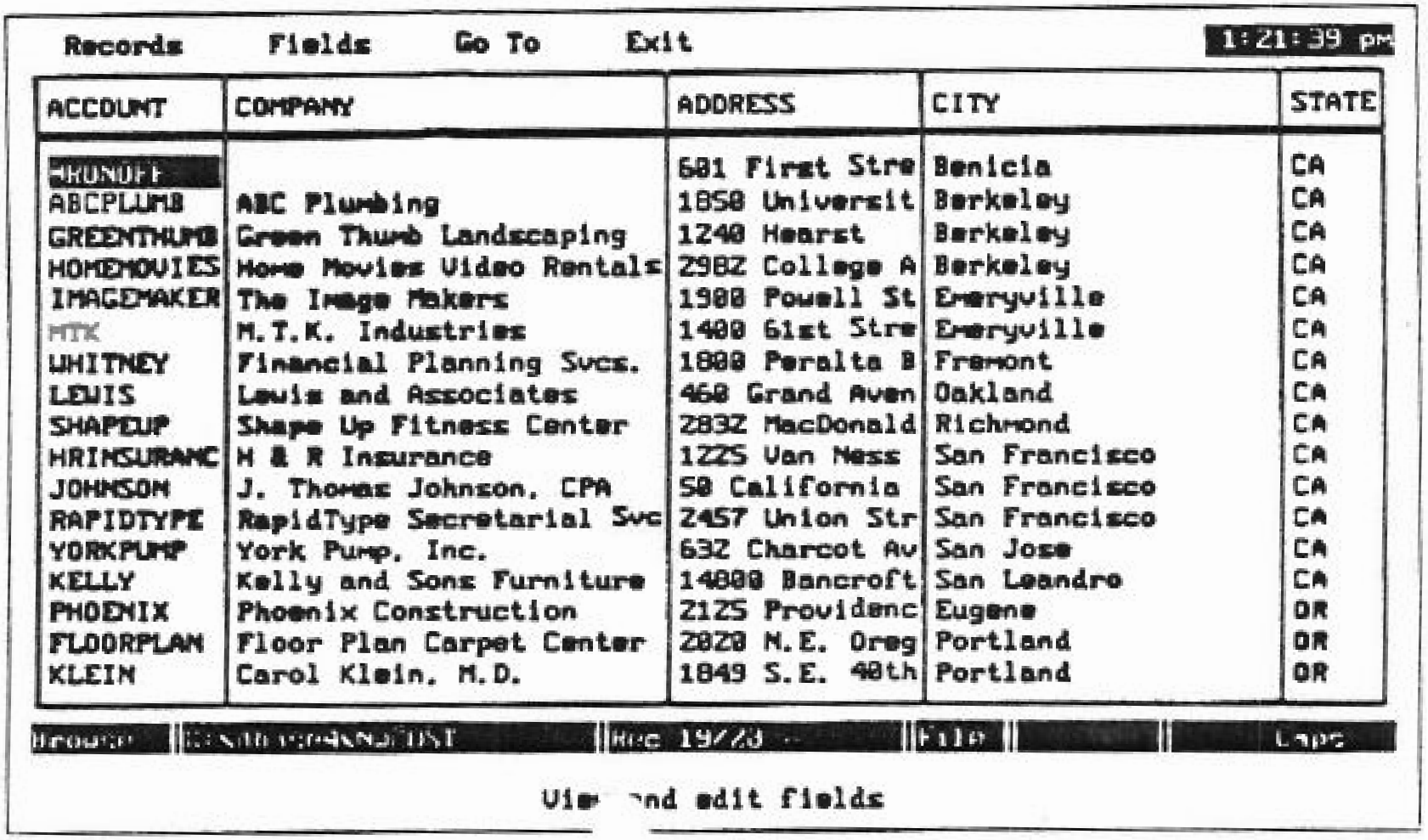
Flat Databases are still with us

- Popular PC database programs of 70’s/80’s

  Dbase (fixed field widths)

  DB2

Flat Databases are still with us



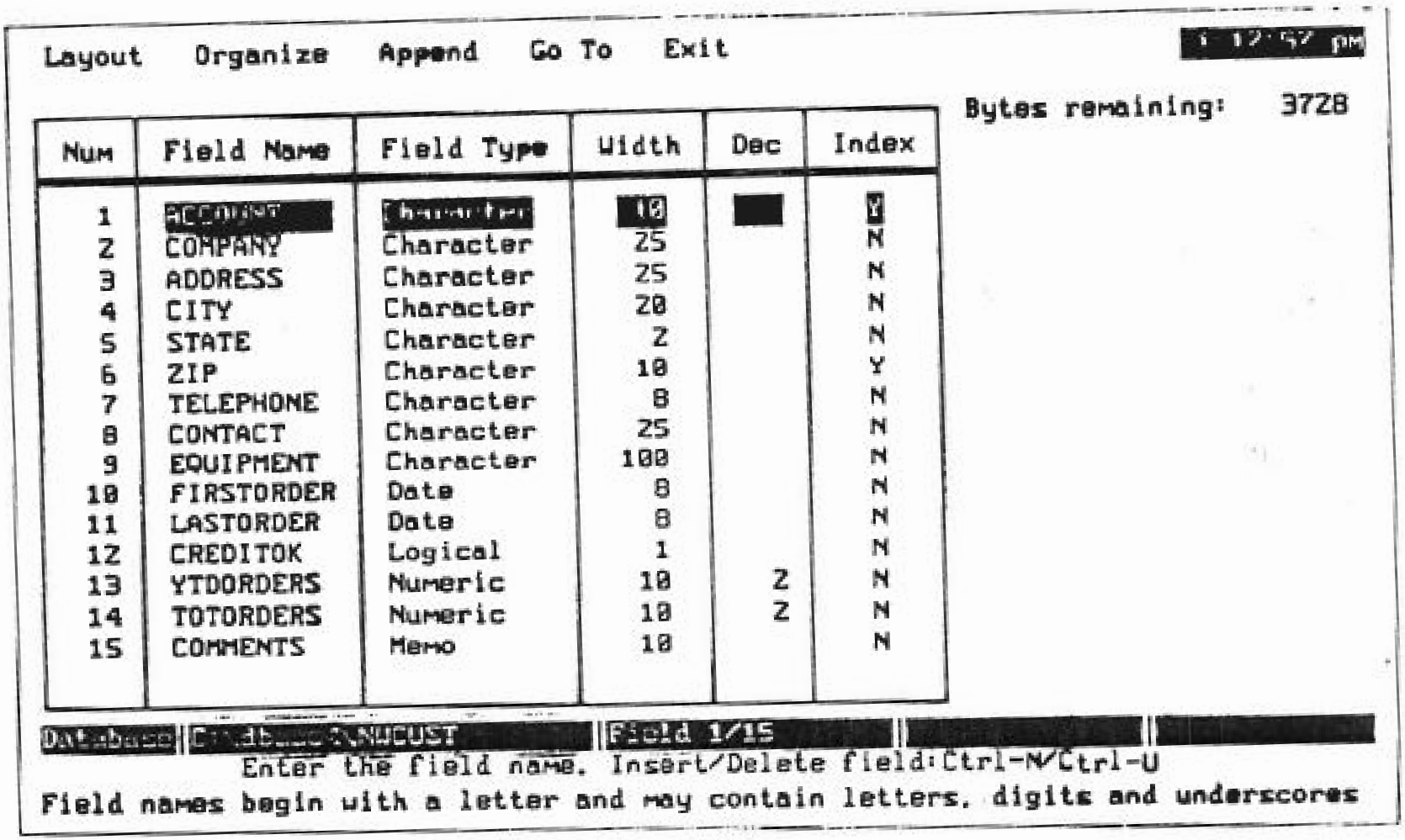
Records Fields Go To Exit 1:21:39 pm

ACCOUNT	COMPANY	ADDRESS	CITY	STATE
BRUNOFF		681 First Stre	Benicia	CA
ABCPLUMB	ABC Plumbing	1858 Universit	Berkeley	CA
GREENTHUMB	Green Thumb Landscaping	1240 Hearst	Berkeley	CA
HOMEMOVIES	Home Movies Video Rentals	2982 College A	Berkeley	CA
IMAGEMAKER	The Image Makers	1988 Powell St	Emeryville	CA
MTK	M.T.K. Industries	1400 61st Stre	Emeryville	CA
WHITNEY	Financial Planning Svcs.	1888 Peralta B	Fremont	CA
LEWIS	Lewis and Associates	468 Grand Aven	Oakland	CA
SHAPEUP	Shape Up Fitness Center	2832 MacDonald	Richmond	CA
HRINSURANC	H & R Insurance	1225 Van Ness	San Francisco	CA
JOHNSON	J. Thomas Johnson, CPA	58 California	San Francisco	CA
RAPIDTYPE	RapidType Secretarial Svc	2457 Union Str	San Francisco	CA
YORKPUMP	York Pump, Inc.	632 Charcot Av	San Jose	CA
KELLY	Kelly and Sons Furniture	14888 Bancroft	San Leandro	CA
PHOENIX	Phoenix Construction	2125 Providenc	Eugene	OR
FLOORPLAN	Floor Plan Carpet Center	2828 N.E. Oreg	Portland	OR
KLEIN	Carol Klein, M.D.	1849 S.E. 48th	Portland	OR

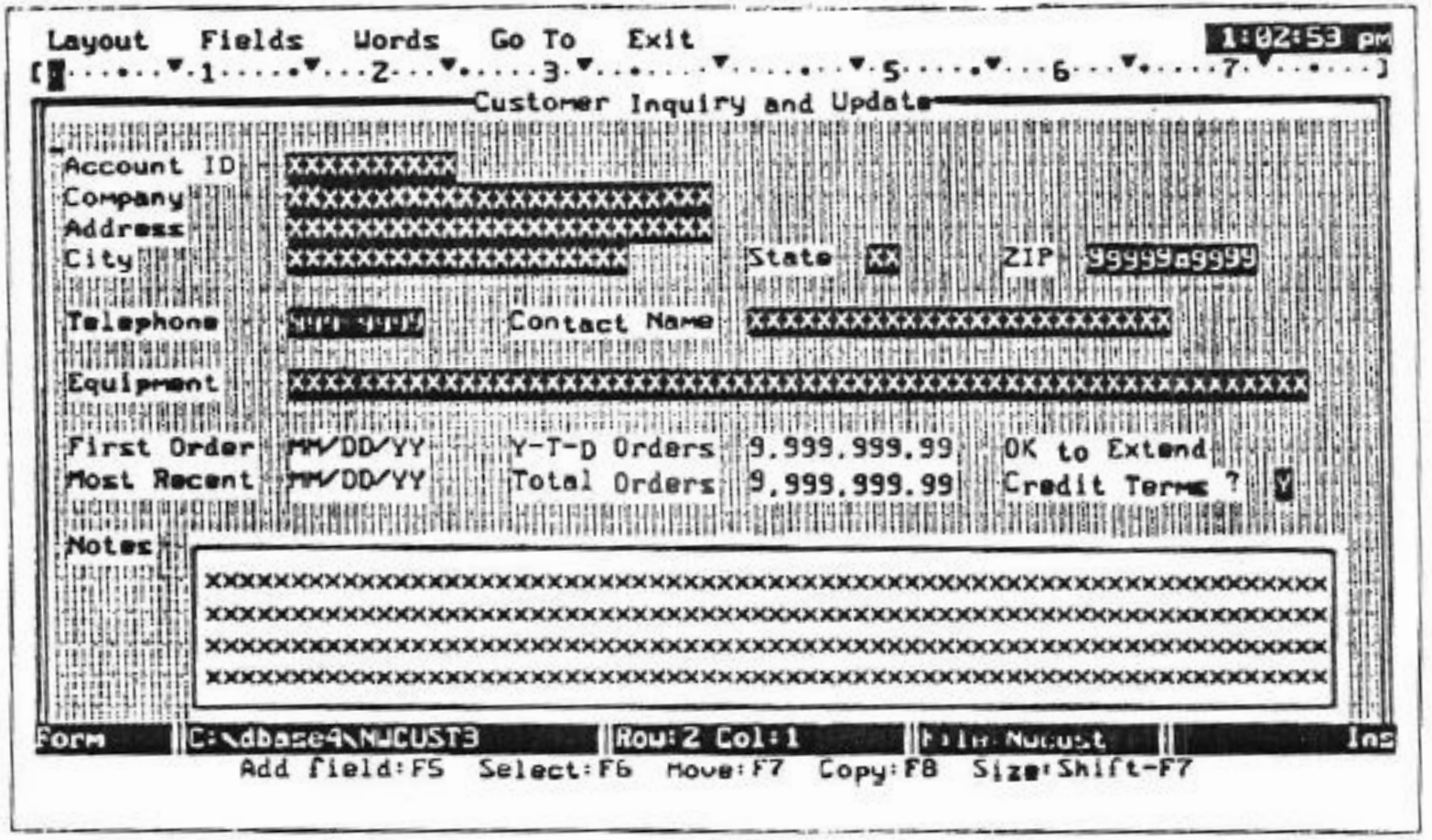
Browse [illegible] Rec 19/20 File Caps

View and edit fields

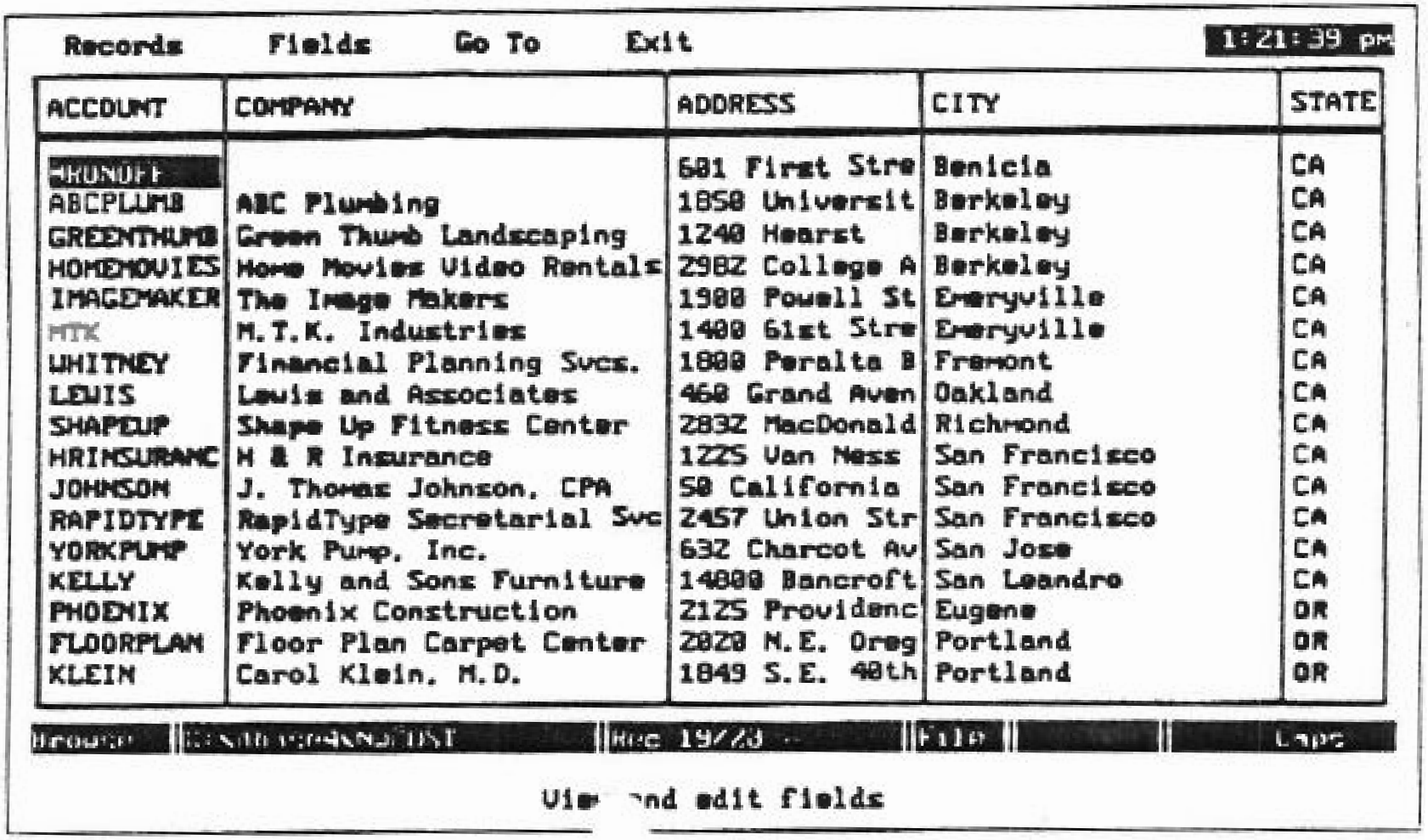
Flat Databases are still with us



Flat Databases are still with us



Flat Databases are still with us



Records Fields Go To Exit 1:21:39 pm

ACCOUNT	COMPANY	ADDRESS	CITY	STATE
BRUNOFF		681 First Stre	Benicia	CA
ABCPLUMB	ABC Plumbing	1858 Universit	Berkeley	CA
GREENTHUMB	Green Thumb Landscaping	1240 Hearst	Berkeley	CA
HOMEMOVIES	Home Movies Video Rentals	2982 College A	Berkeley	CA
IMAGEMAKER	The Image Makers	1988 Powell St	Emeryville	CA
MTK	M.T.K. Industries	1400 61st Stre	Emeryville	CA
WHITNEY	Financial Planning Svcs.	1888 Peralta B	Fremont	CA
LEWIS	Lewis and Associates	468 Grand Aven	Oakland	CA
SHAPEUP	Shape Up Fitness Center	2832 MacDonald	Richmond	CA
HRINSURANC	H & R Insurance	1225 Van Ness	San Francisco	CA
JOHNSON	J. Thomas Johnson, CPA	58 California	San Francisco	CA
RAPIDTYPE	RapidType Secretarial Svc	2457 Union Str	San Francisco	CA
YORKPUMP	York Pump, Inc.	632 Charcot Av	San Jose	CA
KELLY	Kelly and Sons Furniture	14888 Bancroft	San Leandro	CA
PHOENIX	Phoenix Construction	2125 Providenc	Eugene	OR
FLOORPLAN	Floor Plan Carpet Center	2828 N.E. Oreg	Portland	OR
KLEIN	Carol Klein, M.D.	1849 S.E. 48th	Portland	OR

Browse [illegible] Rec 19/20 File Caps

View and edit fields

Flat Databases – where else?

…

Flat Databases – where else?

Name	SSN	HW1	HW2	HW3	FP	FE
…	….	99	83	58	22	92
….	….	78	84	92	33	91
…..						

- AWK/GAWK/Perl
  - Variable field widths
  - separated by tab character (or equiv.)

Advantages:

easy to code(simple data model)

efficient to index + access (data all in one place)

deletion/insertion easy (if fixed width)

→ fewer dependencies

Disadvantages:

- consistency management
- Redundancy

History (continued)

- 1970 Relational Data Model

Ted Codd, IBM research fellow

square
Sequel(SQL)
QBE
Quel

} relational languages

System R – IBM
INGRES - Berkeley

} Relational Research projects

History (continued)

- 1964 – GE Integrated Data Store (IDS)

  Bachman, Network. Data Model

- 1965 – IBM Information Management System(IMS)

  Hierarchical Data Mode

- Late 60's – SABRE (IBM + American Airline)

  First large distributed database with intense concurrency and transaction control needs