Theory and practice of computer technologies used for creating DEA software

Eugene P. MORGUNOV

Siberian State Aerospace University

Krasnoyarsk, Russia

Audience

- those who might want to develop DEA software
- scholars who would like to get an insight into how such software is being developed
- everyone who want to better understand DEA method by means of programming some simple DEA models

Some questions to answer

- choice of programming language
- choice of operating system (OS)
- choice of database management system (DBMS)
- use of Internet technologies
- use of special libraries of various mathematical sub-routines

Possible goals of creating DEA software

- to develop professional DEA software and to distribute it among DEA researchers and practitioners

- to study DEA method by means of programming some simple DEA models

Brief overview of components for creating DEA software

Programming language

- FORTRAN
- Visual Basic (with Microsoft Excel)
- Pascal
- C/C++ (with Borland C++Builder)
- MATLAB
- Perl

Operating system

- **Microsoft Windows**
  - ✓Windows 95/98/XP
  - ✓Windows NT/2000/2003
- **UNIX**
  - ✓Commercial UNIXes: Sun Solaris, SCO
  - ✓Non-commercial UNIXes: Linux, FreeBSD

Database management system (DBMS) (1)

Desktop DBMSs

- Microsoft Access
- Borland C++Builder and Delphi (local databases)
- Visual FoxPro

Server DBMSs

- Oracle
- Microsoft SQL Server
- Borland Interbase
- Informix
- DB2
- PostgreSQL
- MySQL

Database management system (DBMS) (2)

Non-commercial

- PostgreSQL
- MySQL

- **Commercial**
- All others

Advantages of DBMSs

1. Centralized management of all data
2. Reduced redundancy in data
3. Eliminating conflicts between fragments of data
4. Possibility of sharing data between users
5. Possibility of standardization
6. Ensuring consistency in data (transactions)
7. Easy manipulating data (SQL language)

Some notes on data

- Format of data stored in a database may not be the same as format of data displayed to a user

Relational databases – basic terms

- The main term is relation
- Relations are stored in a database in the form of so called tables
- Field (column) – elementary (non-divisible) fragment of data
- Record (row) – all fields that describe one object
- Key – unique identifier of a record (one or more fields)

Internet technologies

Internet-technologies can give

- easiness of centralized updating the software with its new versions
- possibility of solving large-scale problems for those users who don't have access to a powerful computer
- possibility of renting the software without buying it

Special libraries of various mathematical sub-routines

- GNU Scientific Library (GSL) – is a numerical library for C and C++ programmers (http://www.gnu.org)

Advantages for DEA software

- reduced time of developing
- higher reliability

Guidelines for a DEA user who would decide to program simple DEA models

Data for simple DEA problem

Coelli et al. (1998), pages 143–144

DMU	Y	X1	X2
1	1	2	5
2	2	2	4
3	3	6	6
4	1	3	2
5	2	6	2

Simple programs in MATLAB

1. CRS input-oriented DEA problem for one DMU
2. CRS/VRS input-oriented DEA problem for all DMUs
3. Reading data set from a separate data file with use of a special function (which is placed in a separate m-file)

Some tips for further developing this programs

- to add support for varying orientation of the model
- to add functionality for calculation slacks, radial movements, and projected values for every DMU

Brief discussion – MATLAB

- Manipulating with data is rather easy and a user can concentrate on essence of DEA method
- Relatively low speed of processing
- It is not easy to create graphical user interface

Brief discussion – C/C++

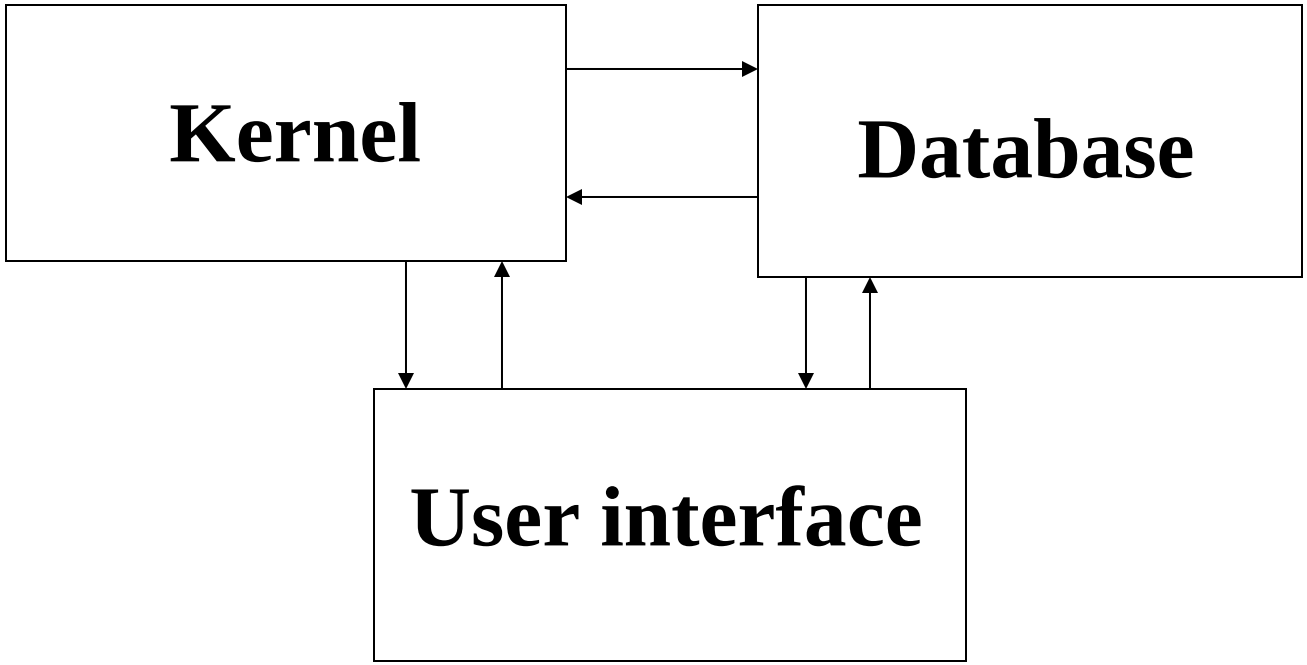
- High speed of processing
- Good user interface may be created using Borland C++Builder or Visual C
- Much time is needed for realizing various auxiliary functions such as allocating computer's memory for matrices, etc.

Guidelines for a computer programmer who would decide to develop DEA software

- **Desktop DEA software**
- **Internet DEA software**

Desktop DEA software

Architecture of DEA software



Kernel

Features

- realizing DEA models
- auxiliary mathematical processing, e.g., correlations, clusterizations, etc.

Language

ANSI C or C++ (because of portability, e.g. from Windows to UNIX system)

Database (1)

- An idea of repository of data
- A study – convenient abstraction of dataset
- Detailed info about any object or any variable is entered into the database only once
- So called 'cross-studying' or 'inter-studying' is possible
- Studying of multilevel hierarchies of objects is possible

Database (2)

Physical format of database –

may be recommended Borland Paradox

- Primary keys
- Indexes
- Rich set of data types
- Default values for fields in database tables

An example of database structure (1)

Table «Studies' descriptions»

- **Study's identifier**
- Study's name
- Study's description
- Date of implementing a study
- Name of a chief researcher

Table «Descriptions of all DMUs»

- **Object's identifier**
- Object's short name
- Object's full name

An example of database structure (2)

Table «List of all variables»

- **Variable's identifier**
- Variable's short name
- Variable's full name

Table «Objects included in studies»

- **Study's identifier**
- **Object's identifier**

An example of database structure (3)

Table «Variables included in studies»

- **Study's identifier**
- **Variable's identifier**
- Type of a variable (input or output)

Table «Repository of data for all objects»

- **Object's identifier**
- **Variable's identifier**
- Value of a variable
- **Number of a period this value is from**
- Date for this period

An example of database structure (4)

Table «Parameters of studies»

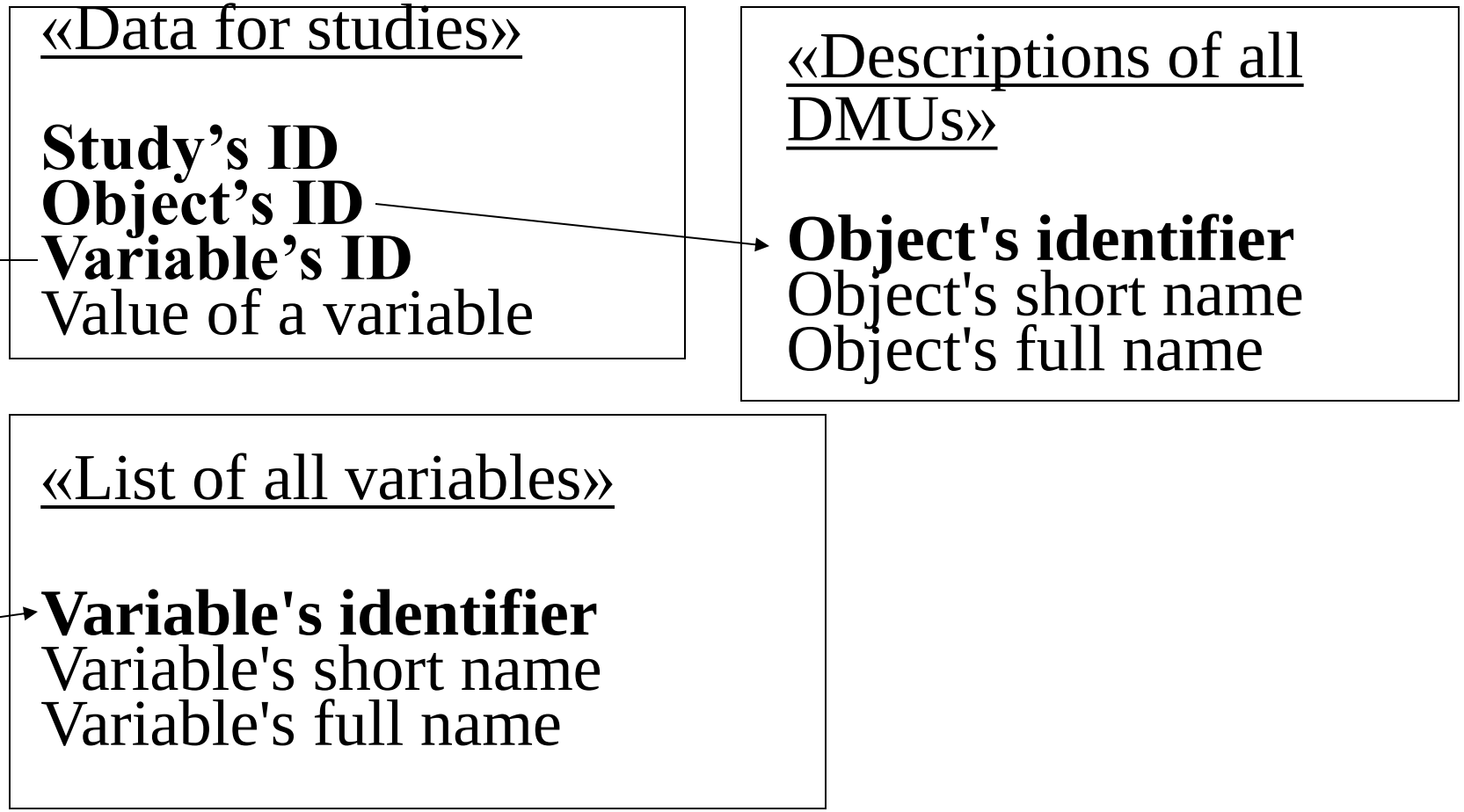
- **Study's identifier**
- Output file name
- Total count of DMUs
- Count of time periods
- Count of input variables
- Count of output variables
- Scale assumption (CRS, VRS, NIRS, NDRS)
- Orientation (input or output)

An example of database structure (5)

Table «Data for studies»

- **Study's identifier**
- **Object's identifier**
- **Variable's identifier**
- Raw value of a variable (from the repository)
- Value of a variable (may be pre-processed)
- Number of a time period
- Date this value was obtained for

An example of relations between database tables



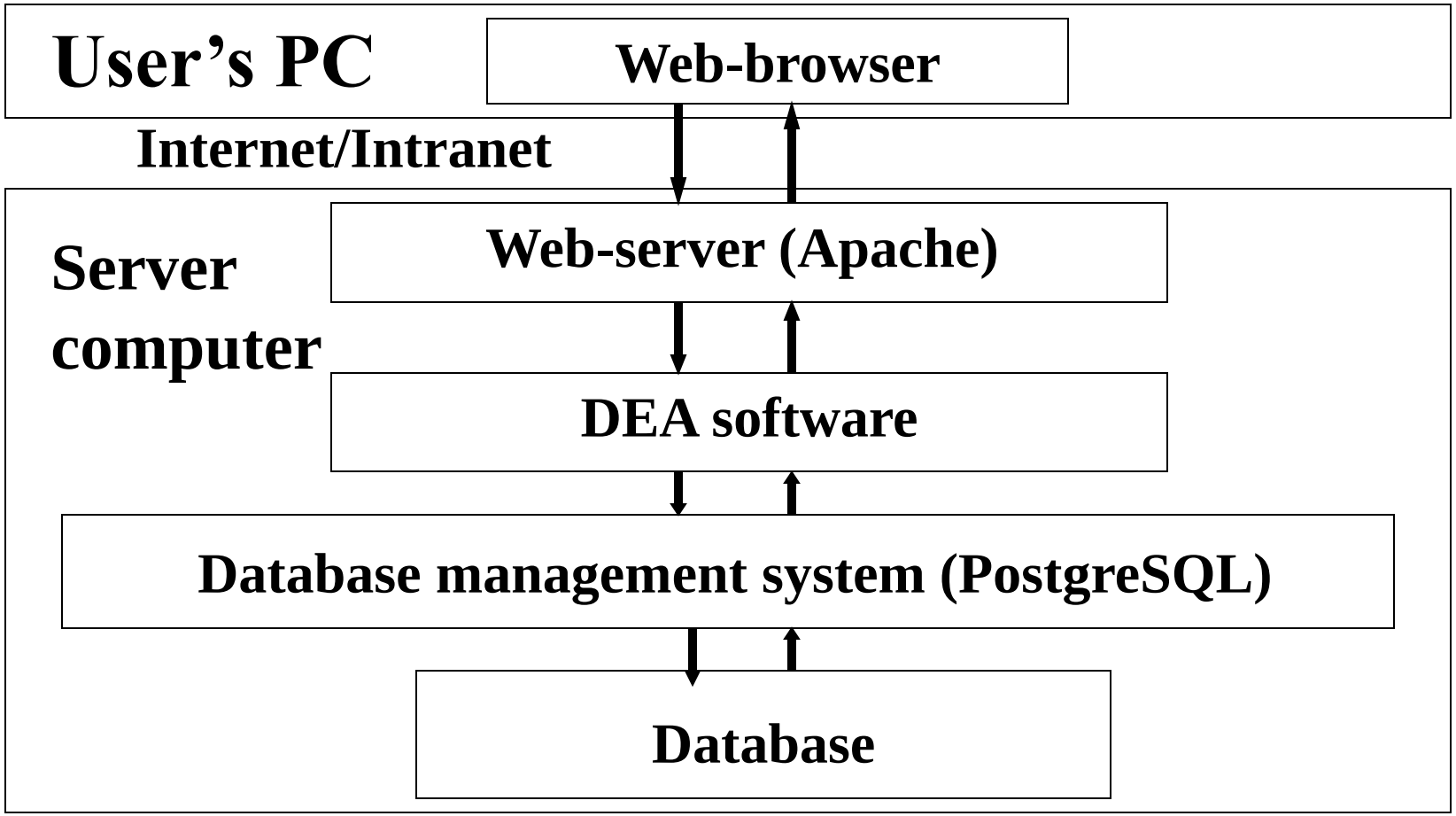
User interface

- Borland C++ Builder
- Borland C++ BuilderX (for UNIX)

- Borland Delphi
- Borland Kylix (for UNIX)

- Microsoft Visual C++

Internet DEA software (1)



Internet DEA software (2)

Kernel

- ANSI C or C++ languages

User interface

- Perl language
- PHP language

Database management system

- PostgreSQL
- MySQL