



SQLfast

Database processing made easy

Learning databases
Prototyping database applications
Developing light database applications

Volume 1 - Introduction to database programming

(draft version)

August 2018

Namur Digital Institute (NADI)
PReCISE Research Center - Data Engineering group
University of Namur
Faculté d'Informatique
Rue Grandgagnage, 21 - B-5000 Namur (Belgium)

Precise web site: <http://www.fundp.ac.be/en/precise/>
Data Engineering group: <http://www.info.fundp.ac.be/libd>
Contact: Jean-Luc Hainaut, jean-luc.hainaut@unamur.be

Foreword

SQLfast is a *language* and a *software environment* that allow novice or casual users to write and execute small scripts to create, modify, consult, manipulate and process relational databases.

This volume covers the main programming concepts of the SQLfast language: database structure, SQL, variables, graphical user interaction, programming structures (sequence, alternative, loops, procedure), expressions, data formatting, data presentation, large data objects, metadata, recursive programming, etc.)

As a companion to this book, a collection of representative database applications is also available. These applications encompass a wide range of problems from data management (e.g., a complete management system for a school library and a bill of material system) to system applications (directory management, SQL injection, temporal databases, large test data set generation, database prototyping), to decision systems (topological sorting, formal concept analysis, graph algorithms, text analysis) and to more playful applications (an adventure game, a novel generator, Conway's game of life, maze solver, writing a successful book).

This material can be downloaded from the SQLfast website. Most tutorial chapters are already available, though sometimes in a draft version. The others chapters still are being written. The code of these chapters is available in the **Script** folder.

Document *Dictionary of commands and parameters of SQLfast languages* contains a short description of all the SQLfast statements, functions, parameters and system variables. It is available on the SQLfast website (link [SQLfast-Commands.pdf](#)).

Jean-Luc Hainaut - Namur, August 2018

Table of contents

Chapter 1 - Introduction	17
1.1 What is SQLfast?	17
1.2 The target audiences of SQLfast	18
1.3 Installing and starting SQLfast	20
1.4 The levels of the SQLfast interface	20
1.5 Starting SQLfast with the Basic interface	20
1.6 Selecting another level	21
1.7 Starting SQLfast with the Expert interface (Level 1)	22
1.8 The example databases: ORDERS.db and CLICOM.db	23
1.9 SQLfast resources and documentation	25
1.9.1 Integrated help	25
1.9.2 Integrated tutorials	27
1.9.3 Scripts	28
1.9.4 The SQLfast e-book tutorial	29
Chapter 2 - Basic scripts	31
2.1 Preparing the work	31
2.2 Writing and running an SQLfast script	32
2.3 Saving the current script	33
2.4 Loading and modifying a script	34
2.5 Other ways to execute a script	34
2.6 Representing scripts and results in this tutorial	35
2.7 How to write SQLfast statements?	36
2.8 Questions on this first script?	37
2.9 A second script, just as simple	38
2.10 Using variables - First encounter	38
2.11 Variable's way of life	40
2.12 Interactive SQLfast scripts	41
2.13 To quote or not to quote	42
2.14 SQLfast sessions	43
2.15 Summary	44
Chapter 3 - Creating a database	45
3.1 How to create a new database	45
3.2 How (not) to loose data	47
3.3 Creating and loading the ORDERS database	49

3.4	Calling secondary scripts (first look)	50
3.5	The schema of the ORDERS.db database	52
3.6	A new skill: writing blank lines	53
3.7	Temporary tables	54
3.8	In memory databases	54
3.9	Upper case or lower case?	55
3.10	Summary	56
Chapter 4 - Updating data		59
4.1	Extending user-driven data selection	59
4.2	Data modification statements: a short reminder	61
4.3	Inserting rows into a table	63
4.4	About single quotes in SQL values	63
4.5	Updating rows of a table	64
4.6	Variables in statement customization	65
4.7	Variables in full statement specification	65
4.8	Improving data entry box layout	66
4.9	Dialogue box for updating values	67
4.10	Summary	68
Chapter 5 - SQL in a nutshell		71
5.1	Introduction	71
5.2	Creating and modifying the schema of a database	72
5.3	Single table extraction queries	75
5.4	Subqueries	76
5.5	SQL functions	77
5.6	Aggregate (statistical) functions	78
5.7	Joining tables	79
5.8	Set operators	80
5.9	Data grouping	81
5.10	Generalization of the select and from clauses	83
5.11	Data modification	83
5.12	SQL views	85
5.13	Stored procedures and triggers	85
5.14	About SQLite	87
Chapter 6 - More on variables		89
6.1	Time variables and statements	89
6.2	Measuring elapsed time	90

6.3	Oh! Wait a minute!	92
6.4	Variable delimiters - Reminder	92
6.5	We can change the delimiters	93
6.6	Sometimes variable substitution must be disabled	94
6.7	Beware substitution ambiguity	97
6.8	Numeric computation	98
6.9	Character string computation	99
6.10	Set or compute?	100
6.11	Multi-variables assignment	101
6.12	Simulating an array	102
6.13	Variables and SQLfast procedures	102
6.14	What is a variable value after all?	103
6.15	Summary	104
Chapter 7 - Taking decisions		105
7.1	Alternative execution of a script	105
7.2	Skipping statements	107
7.3	If-else-endif block structure	107
7.4	Better controlling user values	110
7.5	Conditions	111
7.6	If, goto, labels and variables	112
7.7	Complement: more on goto and label statements	113
7.7.1	Looping and counting	113
7.7.2	Looping again	114
7.7.3	Reading the buttons of the dialogue box	115
7.8	Summary	116
Chapter 8 - Reading data		117
8.1	Can we process individual rows?	117
8.2	Forcing a query to return one row only	118
8.3	Can we force a query to returns two rows?	119
8.4	Queries that return no row	120
8.5	Looping in the database	120
8.6	Embedded loops	122
8.7	A more elegant script	124
8.8	SQLfast arrays revisited	124
8.9	Empty sets, null values and statistics	126
8.10	Summary	128

Chapter 9 - Looping - The final chapter	131
9.1 Introduction	131
9.2 SQL for loop	132
9.3 Range for loop	132
9.4 File for loop	133
9.5 While loop	134
9.6 Forced loop iteration	135
9.7 Forced loop exit	136
9.8 While (True) loop	136
9.9 Correct loop structure	137
9.10 Variables in loops and if statements	138
9.11 How does a loop work, actually?	139
9.11.1 The while (True) loop	139
9.11.2 The while (condition) loop	140
9.11.3 The SQL for loop	140
9.11.4 The file for loop	140
9.11.5 The range for loop	141
9.12 Summary	142
Chapter 10 - SQLfast expressions	145
10.1 Introduction	145
10.2 Value of a variable	146
10.3 Binary values	146
10.4 Syntax of SQLfast expressions	147
10.5 Numeric functions	148
10.6 Character string functions	149
10.7 Logical functions	153
10.8 Temporal functions	154
10.9 File/web functions	157
10.10 General functions	157
10.11 User defined functions (UDF)	158
10.12 Functions from external libraries	159
10.13 Summary	160
Chapter 11 - Output channels, files and directories	161
11.1 The SQLfast output window: the default output channel	161
11.2 The (SQLfast) output window	163
11.3 User output files	163

11.4	The standard output file	165
11.5	Output variables	165
11.6	Closing an output channel	166
11.7	Setting the default output channel and output mode	166
11.8	Clearing the output window	166
11.9	Example: playing with output channels	167
11.10	Managing files and directories	167
11.10.1	About file and directory names	167
11.10.2	Selecting a file or a directory	169
11.10.3	Examining the contents of a directory	170
11.10.4	Examining the properties of a file or a directory	171
11.10.5	Renaming, copying and deleting files	171
11.11	Summary	172
Chapter 12 - Dialogues		175
12.1	The SQLfast GUI	175
12.2	Standard data entry box	176
12.3	Controlling user action	179
12.4	Update a data set	179
12.5	Text entry box	180
12.6	Selecting one item among several	184
12.7	Selecting items among several	185
12.8	Summary	187
Chapter 13 - Displaying data		189
13.1	Information message box	189
13.2	Displaying a list of values on the screen	191
13.3	Displaying a text	191
13.4	Displaying a picture	192
13.5	Displaying drawings	195
13.6	Won't you please, please help me	197
13.7	Summary	198
Chapter 14 - Advanced dialogues		201
14.1	Multicolumn data entry and selection	201
14.2	Logical field grouping	203
14.3	Logical button grouping	204
14.4	No label please!	206
14.5	Application: building a 2D menu	208

14.6 Data entry through selection in a predefined value list	209
14.7 Predefined values from the database	211
14.8 Displayed vs returned values	213
14.9 Customizing standard buttons	215
14.10 Composite dialogue boxes	216
14.11 Heterogeneous composite boxes	220
14.12 Of styles and colors	223
14.13 Example of complex composite box	224
14.14 Summary	224
Chapter 15 - Large data objects	227
15.1 Beyond elementary data items	227
15.2 Displaying large character strings	228
15.3 Reading and writing (large) text files	229
15.4 Reading and writing large binary files	230
15.5 Storing large objects in a database	231
15.6 About binary values	233
15.6.1 Text files: from cp850 to UTF-8	234
15.6.2 Binary files	236
15.6.3 Displaying extracted binary data	236
15.7 Binary data encoding/decoding	238
15.7.1 Hexadecimal encoding/decoding	238
15.7.2 base64 encoding/decoding	239
15.7.3 Secure hashing	240
15.8 A touch of cryptography	242
15.9 Where to store large objects	244
15.10 Loading large volumes of data	245
15.11 Summary	246
Chapter 16 - Writing data tables - Part 1	249
16.1 An in-depth look at output tables	249
16.2 Changing the layout of an output table	252
16.3 CSV output format	253
16.4 Key-Value output format	257
16.5 List output format	259
16.6 Summary	259
Chapter 17 - Writing data tables - Part 2	261
17.1 XML output format	261

17.2	HTML output format	264
17.3	Tuple output format	266
17.4	RTF output format	267
17.5	SQL output format	272
17.6	LateX output format	273
17.7	JSON output format	274
Chapter 18 - Importing and exporting data		277
18.1	External data import	277
18.2	CSV data import / export	278
18.2.1	CSV syntax rules	279
18.2.2	Exporting data in csv format	280
18.2.3	Refining export function	281
18.2.4	Values including double-quotes	281
18.2.5	Exporting binary values	281
18.2.6	Importing rows from a csv file	282
18.2.7	Example: save and restore relational data	285
18.2.8	Estimating column types through data analysis	285
18.3	dBase data import	287
18.3.1	DBF file format	288
18.3.2	Importing rows from a dbf file	288
18.3.3	Example: migrating a dBase database	290
18.4	SQL-DMO data import/export	290
18.5	JSON data export/import	290
18.6	Summary	290
Chapter 19 - Recursive programming		291
19.1	Introduction	291
19.2	SQL recursive queries	293
19.2.1	Common Table Expressions (CTE)	293
19.2.2	Recursive CTE	294
19.2.3	Union all or union?	295
19.3	SQL recursive queries: the Staff application	296
19.4	Recursive queries against cyclic data	300
19.4.1	Arbitrarily limiting the number of executions	302
19.4.2	Limiting the number of executions to the size of the problem	302
19.4.3	Preventing the query from visiting a person twice	302
19.4.4	Discarding duplicates in the union operator	303
19.5	Can recursive queries compute everything?	304

19.6	Tree traversal ordering	307
19.7	SQL recursive triggers	308
19.7.1	Factorial	309
19.7.2	The Staff application	309
19.8	Loop-based SQLfast implementation of recursive procedures	312
19.8.1	Computing the factorial of a number	312
19.8.2	The Staff application	313
19.9	SQLfast recursive scripts	315
19.9.1	Recursive programming in general	316
19.9.2	The Staff application - Version 1	317
19.9.3	Computing the factorial of a number	318
19.9.4	The Staff application - General approach	320
19.10	Comparison and conclusions	321
Chapter 20 - Metadata - Walking on the wild side		323
20.1	Introduction	323
20.2	The SQLfast dictionary tables	326
20.2.1	The SYS_TABLE and SYS_COLUMN tables	326
20.2.2	The SYS_KEY and SYS_KEYCOMP tables	328
20.2.3	The SYS_INTER_TABLE view	329
20.3	Dictionary or no dictionary?	330
20.4	Examining the schema of an unknown database	330
20.5	From metadata to data	334
20.6	Complement: SQL group_concat function	336
20.7	Summary	338
Chapter 21 - SQLdraw		339
21.1	Vector graphics in SQLfast	339
21.2	The SQLdraw scripting language	339
21.3	Displaying an SQLdraw script	339
21.4	The SQLdraw editor	339
21.5	SQLdraw project	343
21.6	Drawings	344
21.7	Drawing charts	344
21.8	Digitizing maps	344
21.9	The SQLdraw viewer	344
21.10	Animation and cartoon	344
21.11	Summary	344

Chapter 22 - Extending SQLfast	345
22.1 Introduction	345
22.2 SQLfast utilities	345
22.2.1 Service scripts	346
22.2.2 Functional scripts	346
22.3 Python external applications and libraries	346
22.3.1 External SQLfast applications	346
22.3.2 External SQLfast function libraries	347
22.4 User-defined functions (UDF)	347
22.5 SQLfast language extension	348
22.6 Foreign applications	348
22.7 Developing custom-made external applications and libraries	349
22.7.1 General procedure	349
22.8 Developing custom-made UDF for SQLfast and SQL	350
22.9 SQL queries in UDF	353
22.9.1 SQL statements in an SQL UDF	353
22.9.2 SQL statements in an SQLfast UDF	354
22.10 Complements: standard external SQLfast function libraries	355
22.10.1 The LStr library	355
22.10.2 The LFile library	356
22.10.3 The LWeb library	356
22.10.4 The LGeo library	357
22.10.5 Some recommendations for the development of libraries	358
Chapter 23 - Aid to SQLfast script development	359
23.1 Introduction	359
23.2 Summary	359
Chapter 24 - Storing transient data - The SQLfast lists	361
24.1 Transient data: temporary tables or SQLfast lists?	362
24.2 SQLfast lists	362
24.3 Creating an SQLfast list	363
24.4 Manipulation of SQLfast lists	363
24.5 Set-theoretic manipulations of SQLfast lists	365
24.6 List-based for loop (from version 3.0)	366
24.7 Standard versions of SQLfast list functions	366
24.8 Three application examples	367
Chapter 25 - Dynamic columns and schema-less tables	371

25.1	Introduction	372
25.2	Dynamic columns	372
25.3	Practical syntax	373
25.4	Metadata of dynamic columns	374
25.5	Technical note	376
25.6	Performance: static vs dynamic columns	377
25.7	Dynamic columns in DBMS	378
25.8	Summary	381
Appendix 1 - A basket of SQLfast examples		383
1.1	Introduction	386
1.2	Consulting customers	387
1.3	Interactive data loading	388
1.4	Comparing the contents of two tables	389
1.5	Displaying the contents of two tables in a single table	394
1.6	Dumping a database in XML format (version 1)	397
1.7	Dumping a database in XML format (version 2)	399
1.8	Data-driven multiple aggregations (table pivoting)	402
1.9	Data mining	406
1.9.1	The data	406
1.9.2	Studying coupled purchases	410
1.9.3	How to place products in the shelves of the store?	411
1.10	Condorcet voting paradox	413
1.11	Paged reading of data	415
1.12	Raising temperature	418
1.13	Managing a file of parameters	420
1.14	Dynamic branching	424
1.15	Simulating arrays in SQLfast	425
1.16	Smart writing of data table	428
1.17	Checking functional dependencies in a table	431
1.18	Image selection tool	434
1.19	Managing a photo album	437
1.20	A picture viewer	440
1.21	Collecting web statistics	441
1.22	Computing missing data in incomplete statistical series	443
1.23	The Travelling Salesman problem	448

1.24	Ordering	452
1.25	Numbering the rows of a table	454
1.25.1	First technique: applying a mathematical definition	454
1.25.2	Second technique: adding the source rowids	455
1.25.3	Third technique: adding the target rowids	455
1.25.4	Fourth technique: adding a auto incremente column	455
1.25.5	Fifth technique: recursive CTE	456
1.25.6	Comparing the numbering techniques	457
1.25.7	Pairing two tables	457
1.26	Creating empty arrays	458
1.26.1	Full arrays	458
1.27	Computing the powerset of a set	461
1.27.1	First technique	461
1.27.2	Second technique	462
1.27.3	Performances	463
1.27.4	Storing the powerset members in a table	463
1.28	The relational division	465
1.29	Checking set inclusion	467
1.29.1	The inclusion problem	467
1.29.2	The algorithms	468
1.29.3	Performance evaluation	470
1.29.4	The set equality problem	470
1.29.5	Transitive reduction	471
1.29.6	Transitive reduction with equality relationships	472
Volume index	423	