Conceptual model of an application and its use for application documentation

Martin Vonka
Faculty of Civil Engeneering
Czech Technical University in Prague
vonkama@gmail.com

Abstract: Following article proposes methodology for conceptual design of a software application. This form of design is suitable for dynamic development environment and agile principles of software development. Article discus the required scope and style used for description of the application. Unification of a documentation significantly reduces the time required for communication within the development team. Some part of the documentation are obtained using the method of reverse engineering, for example by analysis of the application structure or its source code.

Key words: software development, documentation, application concept

1. Introduction

This article deals with the design methodology on how to build a conceptual model of an application. Using this model in the development of the application itself and during the preparation of a documentation required for the final product. Therefore, not just the operational and maintenance documentation, but also the user documentation. Using the principles of application design described in this article it is possible to achieve significant time savings during the actual development as well as during documentation process. This should be the primary objective for every developer - effectiveness in the entire application lifecycle. Software development is undergoing in recent years a great innovation. There is a lot of new software development companies applying agile methodologies of software development. Clearly defined developments phases are fading out. - taken from [Weske, 2007]:

- requirements definition
- design and project documentation
- programming
- testing
- customer hand off, documentation updates
- maintenance and operations
- decommissioning

These tightly defined phases are being replaced by dynamic development methods based on the principles of extreme programming and the spiral model.

This article deals with basic principles that are necessary to build a conceptual model for the application development [Wikipedia 2] and the subsequent management of the entire software project. So the final outputs of the development process will be an actual application, administrative manual and user documentation. At the same time, this methodology takes into account the requirements and possibilities of modern development methods that can be used in a dynamic development environment.

The goal is to automatically assemble the documents, or document parts from the current state of the application. These documents can be divided into three categories, user documentation, documentation for developers and documentation for the administrator.

User documentation consists of the following parts:

- User documentation, or components for its compilation
- Documentation for individual parts of an information system using following breakdown:
  - Documentation for application = general description of the features of the application
Description of the implemented processes in the form of diagrams

Developer Documentation consists of the following parts [Makowski, 2010].

- Description of the application structure
- ER diagram
- Developer Documentation
- Description of functions including a description of the parameters and the methods of key values calculations.
- Analysis and control of an application
- Proper index settings
- Application function calls
- Authentication, to prevent attempts to call unreachable values

Administrator (operational documentation) - used for solving problems at runtime

- SQL queries execution time analysis
- Loading time of forms and tables presenting data to the user
- Execution time of multiple queries (join) and the use of joint columns

2. Solving the issue of documentation by reverse engineering

Software can be seen from different angles. The user sees the software or information system as a set of processes and functions [Řepa, 2006], which is used within the organization. The user often has no idea about architecture or method of functioning of the software. The important parts for the user are inputs and outputs of the functions that he/she is using.

The definition of reverse engineering by Elliot Chikofsky: "Reverse engineering is an analytical process of the system, which leads to the discovery of the system components and their interconnections, followed by description of the system in another way or by using a higher level of abstraction." [Eilam, 2005]. Methods of reverse engineering, approach the system as a whole and try to decompose it into subsystems. Relationships between system components are simplified and recorded using the available methodologies.

A prerequisite for a successful decomposition of the system is to understand how it is arranged within [Řepa, 2006].

For software programmers, the entire information system is defined by its components (subsystems), architecture and data flow between components. On the lower level (single component) is represents a specific set of software libraries, classes and their source codes which process data. In addition to individual components, then the programmer identifies each data source and method of communication with them. [PMI, 2004] Software structure could be divided into following layers according to their perception:

- Complex functions provided by information system to the user (inputs, outputs)
- The system architecture - components and data flows
- Feature specific components
- Functions source code
- Data – methods of data processing in an information system

If the operator of the information system has access to all these layers and has control over it, then he is able to modify and develop the system going forward. Otherwise operator uses information system functions and has no control over data processing. In order to gain this control it is necessary to examine the system to map the structure and functionality of the individual layers and data flow. This analysis is called reversing, or reverse engineering - by [Aalst, 2004].
2.1 Requirements definition

The goal of reverse engineering is to identify system components and their relationships, the initial steps are:
- create a representation of a system in another form (at a higher level of abstraction) - therefore any reality describing model
- inspection of the system – verification of links and bug finding
- preparing structures for further processing of data (data mining)
- documentation updates

2.2 Sources of study

For the study of “finished” applications we have following sources of information:
- source code (the systematic “top-down” study)
- architecture of the application
- database structure
- original documentation, which may be misleading in many aspects due to changes in project implementation

2.3 Tools for generating documentation from source code

Here I present list of available programs that are able to generate documentation from a properly commented application source code. In some special cases it is possible to generate at least the structure from objects and the already compiled executable file – or from dll library. These libraries are for example .NET binary executable files, which contains that .NET framework is using during the execution.

Next table is taken from [Wikipedia 1] and [Kerzner, 2003]. It contains a brief overview of programs designed for maintenance of project source code and documentation. This list contains only tools that can work with source code.

| Name                  | Creator                        | First public release date | Latest stable version | Software license          | Text | Bin. |
|-----------------------|--------------------------------|---------------------------|-----------------------|----------------------------|------|------|
| ApiGen                | David Grudl, Jaroslav Hanslík, Ondřej Nošpor | 2010                      | 2.5.0                 | Now BSD / GPL v2 / GPL v3 | Yes  | No   |
| Appledoc              | Gentle Bytes                   | 2009                      | 2.0.3 [1]             | Modified BSD               | Yes  | No   |
| Autoduck              | Eric Artzt                     | 1993                      | 2.10 [2]              | Freeware[3]                | No   | Yes  |
| BI Documenter         | Mienasoft                      | 2006                      | 1.5                   | Proprietary                | No   | Yes  |
| classdoc              | Jens Gulden                    | 2001                      | 1.0                   | GPL                        | Yes  | No   |
| Codenotes (format convention; not a generator) | Now Vintage Media              | 2012                      | 1.0                   | Omniliensce                | No   | Yes  |
| CppDoc                | Richard Feit                   | 1998/11/23                | 2.3.1                 | Proprietary                | No   | Yes  |
| DB Manual             | Blue Oceans Ltd                | 2006                      | 1.3                   | Proprietary                | No   | Yes  |
| DBDesc                | Logica2                        | 2005                      | 3.1                   | Proprietary                | Yes  | No   |
| DBScribe              | Leadum Software                | 2007                      | 1.1                   | Proprietary                | No   | Yes  |
| Ddoc                  | Walter Bright                  | 2005/09/19                | DMD 1.010 / GDC 0.14  | Freeware / GPL             | Yes  | No   |
## Conceptual Model of an Application and Its Use for Application Documentation

| Name           | Creator                  | First public release date | Latest stable version | Software license                  | Text | Bin. |
|----------------|--------------------------|---------------------------|-----------------------|-----------------------------------|------|------|
| devscribe      | Geoff Cox                | 2008                      | 2.00 [4]              | MSDN Code Gallery Licenses        | Yes  | No   |
| DOC++          | Dragos Acostachioale.    | 2002/12/22                | 3.4.10                | GPL                              | Yes  | Yes  |
| Doc-O-Matic    | toolsfactory software inc| 2000/11/--                | 6.1                   | Proprietary                      | Yes  | No   |
| Document! X    | Innovasys                | 1998                      | 2014.1                | Proprietary                      | Yes  | No   |
| Doxygen        | Dimitri van Heesch       | 1997/10/26                | 1.8.7                 | GPL                              | No   | Yes  |
| Epydoc         | Edward Loper             | 2002/01/--                | 3.0                   | MIT License                       | Yes  | No   |
| ForgeDoc       | Balazs Tihanyi           | 2012/02/21                | 1.0                   | Proprietary                      | Yes  | No   |
| fpdoc          | Sebastian Guenther       | 2005?                     | 2.4.0                 | GPL                              | Yes  | No   |
| GenHelp        | FraserSoft               | 1998/10/--                | 5.6.7                 | Proprietary                      | Yes  | No   |
| Haddock        | Simon Marlow             | 2002                      | 2.13.1 (2012)         | BSD                              | Yes  | Yes  |
| HeaderDoc      | Apple Inc.               | 2000/09/--                | 8.9.14                | APSL                             | Yes  | No   |
| Help Generator | Agora Software BV        | 2004/08/--                | 4.0                   | Proprietary                      | Yes  | No   |
| HyperSQL       | Randy Phillips, Itzchak Rehberg | 2001 | 3.7.5 | GPL | Yes | No |
| IDLdoc         | Michael Galloy           | 2002                      | 3.3.1                 | BSD                              | Yes  | No   |
| Imagix 4D      | Imagix Corp.             | 1995                      | 7.3                   | Proprietary                      | Yes  | No   |
| Javadoc        | Sun Microsystems         | 1995                      | 1.6                   | GPL                              | Yes  | No   |
| jGrouseDoc     | Denis Riabtchik          | 2007/04/--                | 1.5                   | BSD                              | Yes  | No   |
| JSDoc          | Michael Mathews          | 2001/07/--                | 1.10.2                | GPL                              | Yes  | No   |
| JsDoc Toolkit  | Michael Mathews          | 2007?                     | 2.0.0                 | MIT License                      | Yes  | No   |
| KDOC           | Sirtaj Singh Kang        | 1999                      | 3.0.1                 | ?                                 | Yes  | No   |
| mkd            | Jean-Paul Louyot         | 1989                      | 12.03 (2012)          | EUPL                             | No   | Yes  |
2.3.1 Inputs for an analysis

A very important parameter for evaluation of individual SW is the format of input data which will be processed. The last two columns of the table above indicates if the software allows the analysis of source codes - column "Text". Or if it allows the analysis of already compiled programs - column "Bin." In cases where the analysis is used during programing, source codes are available and it's possible to work with them. These types of analyzes are focused to support the team of developers. They keep the information about the structure of classes, object properties or manage documentation and description of functions written in the comments placed directly in the source code.
A second approach of the analysis - the analysis of the compiled binary files, is performed when for some reason we do not have access to the source code and we don’t know its structure and how it is programmed.

### 2.3.2 Supported Environments

The above listed programs for managing source code and supporting work on software development teams, works with almost all programming languages. Here is the list of the common one: C/C++, Java, C#, VB / VBScript, Delphi / Pascal, Ada, D, IDL, .NOT1, Access, Fortran, PHP, Perl, Python, Ruby, Java, JavaScript, ActionScript, PL/SQL, Tcl, Haskell, AppleScript, MIG, BourNo Shell, C shell.

### 2.3.3 Output format

The outputs of individual programs are always in HTML format, eventually in other formats such as CHM, RTF, PDF, LaTeX, PostScript, man pages, DocBook, XML, ePub.

A more detailed description and capabilities of the individual programs are described in [Wikipedia 1].

### 2.4 Generating the structure of database

Another part of the documentation is ER Diagram of the database. It is possible in some cases to generate ER from an existing database. This functionality is only available for certain database servers and only under certain circumstances. It is important that the tables have been linked at the level of the database. Setting up links in the database server can in some cases replace the identification of indexing columns that are used for bindings, or at least suggests to SQL Server what data it should include in the database cache. [Kerzner, 2003]

Relations between the tables, may also be maintained at the application level. In this case, database reverse engineering is more difficult, in some cases impossible. Tools for generating diagrams from DB structures usually expect relations to be defined on the database level. In this chapter is a brief overview of the tools which enable to generate this type of documentation. There is also description of conditions that must be met in order to be able to generate ER Diagram from database. The following paragraphs describe some tools that allow automated generating.

### 3. Design of methodology for creating conceptual model of an application

For application development is good to use a modern flexible development methods and tools. When using today's dynamic development tools it is a good practice to manage the development project using the spiral methodology and extreme programming practices.

It’s desirable to respect the mentioned methods when creating initial documentation - specification. The specification may not accurately describe all the details of applications at the lowest level. The level of simplification of application design must be chosen very carefully so that the basic model of application (either procedural or structural) contains and identifies all the parts (components) of the applications, all data flows and all relevant processes. The specification must also have clearly defined objectives of the project, according to which the application will be gradually developed and subsequently accepted. [Richards, 2007]

The initial proposal is therefore used mainly for relatively accurate estimate of the complexity of application development - work involved most often calculated in man-days and subsequently in money. Such a proposal is then attached as part of the contracts under which is the application being developed.

In other iterations of the project, the individual parts of the proposal are sequentially or in parallel (parallel - if it their character allows) described in more details to create specific tasks for development team.

During the detailing of the specification it is a good practice to think about it as documentation for the whole application, so that the individual parts can be integrated into the final product documentation without major modifications.

In iterative development of an application it can come to the changes to the fundamental design of the original logic of the entire applications. The amount of major changes depends on quality of the original conceptual design, eventually on the experience of the managing person of each iteration. It is important to keep the iterations within the contracted budget and timelines.
3.1 Structure of the conceptual model

Conceptual model should contain the following fundamental parts:

- **Goals**
  - the essential functionality provided by the application
- **Object schema**
  - the description and the links between different parts (components) of an application
  - Communication with other systems - data exchange
- **Basic ER Diagram**
- **Description of tables and their relationships**
- **Description of the processes that will be implemented in the program**
  - the list of functions that will be used by the process
- **Acceptance criteria and methods of verification**

3.2 Setting goals

Setting goals is based on the specification of the information system. It describes what inputs and outputs the program will have, what functions it will provide and what processes it will support.

**Form**

The form of the document is mostly text with diagrams illustrating the basic structure of the application or simplified processes.

3.3 Object schema of an application

Object schema complements a text description of the goals of targets of an application. It should provide a clear graphical representation of the basic components of the application and the underlying data flows. The emphasis is on clarity.

**Form**

Any graphical representation of the main processes and the structure of an application.

**Example**

An example of an application is a diagram solutions that provides data from maps. Data are presented as web and mobile application.
3.4 Basic ER Diagram

The ER diagram should contain fundamental tables on which the application is built. The diagram must include all relevant relations and a relations table. Diagram is used as the initial design. During the each iteration of development it may be partially modified or supplemented with other functions.

Diagram of a final design of an application, to reflect database structure changes during each iteration of development, it is no longer appropriate to incrementally update the diagram, but directly generate it from the final application. The current state of the application is reflected in the documentation - correct ER diagram.
3.5 Description of the basic processes that will be implemented in the program

The process description must include all the important decisions and states. Flowchart defines application functions being implemented that validate the various inputs and allow the user to continue the process further, or eventually stop it and request additional information. Automatic functions if they exist in the application are programmed based on these processes.

Form

Description of the process usually consists of two parts. The graphical representation – i.e. the flowchart and text description of the each function. For graphical representation as well as for a text description the "less is more". Thus, there is an emphasis on proper structuring of the text and simplicity of the whole process scheme.

Example

In the example is shown a graphical diagram of the process and part of a textual description of the functions used in the process.
Example text description of the functions of the above process diagram:

1:

**input:**
The user enters into the mobile phone information about fuel prices on the selected fuel station sends data it to a specific phone number.

**action:**
3rd party application receives the request.

**output:**
None

2:

**input:**
Timer initiates event every 2 minutes

**action:**
The application connects via SMS API and downloads all new data from SMS center.

**output:**
Raw SMS message in the application NET Genium.
- syntactically correct SMS

Example description of the process of the project for prices processing
3.6 Acceptance criteria and methodology for verification

Acceptance criteria are a very important part of the conceptual design, they offer a checklist of the required functionality and how the individual functions or support processes will be controlled during the acceptance phase of application deployment.

4. Time schedule and budget

When the conceptual design of the application is created and all goals are written, it is necessary to estimate the time required for the development of individual components and plan budget for the entire project implementation. In most cases the conceptual design of the application should be sufficient for budgeting and identification of other costs that may occur during project or other project expanses.

Examples of additional costs are licenses for operating systems or database, as well as other supporting software necessary.

For the purpose of the cost and time estimate it is good to divide the project into separate units, which will be subsequently assigned to individual developers. Such a structured estimate, is possible to use as assignment of tasks during the project realization.

For a project in the value of 100 hours is the estimated amount of work involved compiling the documentation about 10%, therefore 10 hours. If we use a new methodology of simplified application design then we can generate documentation with final formatting within 2 hours. It is necessary to also add time, which is spent by developers to enter the documentation into the application. In total we are looking at seven hours of work, comparing to the standard method. This represents about 30% less work involved in the preparation of application documentation. However, another area of a substantial savings is the application development itself. Developers are not forced to study the extensive project documentation. It is sufficient for them to get familiarity with the concept of an application. When solving individual tasks they use documentation embedded directly in the application and are only concerned with the parts that are directly related to their task.

5. Development resources planning

When using the spiral method of software development, it is necessary to carefully plan of each development iteration (design - development - testing - assessment - proposal2 - ...) and stick to the milestones established for the project and budget.

For the entire project is necessary to properly set the individual phases, which will correspond with individual part or components of developed software. Each iteration runs within these phases, of the spiral model.

For each development iteration in the given phase is necessary to compiled basic structures of specific tasks. In the analytical phase, requirements are being defined in details.

For each part of the application is necessary to prepare:

- Database structure and forms that will be used for data entry.
- In parallel, create functions that will be used to handle the forms (transformation and calculations with data for one record) and processing of database records (multiple records transformation, verification of states, privileges ...)
- Setting permissions, data access, creation of the data on which the implemented functions and process are verified.
- The plan for testing of the application. It is very important to document all possible combinations of inputs and resulting outputs. Furthermore, the test scenarios for all possible variants of the user steps through the flowchart of implemented processes.

6. Conclusion

The methodology described above for compiling the concept of an application is designed primarily for dynamic development environments. These environments allow you to minimize the time which is necessary for maintaining project documentation. Documentation in these environments are provided within the application itself. During the development of the individual parts of the application, based on models and principles described in this paper, are directly updated and further developed. The
information contained in the application during its development are then used for the preparation of final documentation. For the meaningful use of assembled documents in practice, it is very important to follow the rules described above. From a personal experience I can confirm that the time invested in training of individual developers and continuous control of the partial development results, contributes to the significant time savings during the management of documents that must be submitted along with the completed application before acceptance.

Another advantage of this approach is the unification of documents for developers. Especially when a new developer starts to work on the project. Then the developer can easily find all the documents for the analysis of the problem with the application or during the implementation of a new functionality. This systematic approach, according to my estimates, will save approximately 30% of the time required to maintain the documentation.

**Literature**

Aalst, W., Hee, K., 2004. *Workflow Management: Models, Methods and Systems*. The MIT Press, London, 384 pp. ISBN 978-0262720465

Makowski, L., 2010: *Specifikace a dokumentace projektu*, Zive.cz (Online resource): [http://www.zive.cz/clanky/specifikace-a-dokumentace-projektu-dil-2/sc-3-a-108611/default.aspx](http://www.zive.cz/clanky/specifikace-a-dokumentace-projektu-dil-2/sc-3-a-108611/default.aspx)

Eilam, E., 2005: *Reversing: Secrets of Reverse Engineering*, Wiley Publishing, ISBN: 978-0-7645-7481-8

Florac, W. A., Carleton, A.D., 1999: Measuring the software process:statistical process control for software process improvement. Addison Wesley. 250 pp. ISBN 0-201-60444-2

Kerzner, Herold, 2003: *Project Management: A Systems Approach to Planning, Scheduling, and Controlling*, John Wiley & Sons, 891 pp. ISBN 0471225770

Project Management Institute, 2004.: *A Guide to the Project Management Body of Knowledge (PMBOK® Guide)*, 3. rel. standard ANSI/PMI 99-001-2004. Newtown Square, Pennsylvania: Project Management Institute, 403 pp. ISBN 1-930699-50-60

Řepa, V., 2006, *Podnikové procesy: Procesní řízení a modelování*. Grada Publishing, a.s., Praha, 268 s. ISBN 80-247-1281-4

Richards, K., 2007: Integrating DSDM into an existing PRINCE2 environment. The Best Management Practices Knowledge Centre. 6 pp. [http://www.best-management-practice.com/gempdf/DSDM_White_Paper_v2.pdf](http://www.best-management-practice.com/gempdf/DSDM_White_Paper_v2.pdf)

Weske, M., 2007, *Business Process Management: Concepts, Languages, Architectures*. Springer, 368 pp. ISBN 978-3540735212

Wikipedie 1, Wikipedie.org, *Documentation generators*, (Online resource): [http://en.wikipedia.org/wiki/Comparison_of_documentation_generators](http://en.wikipedia.org/wiki/Comparison_of_documentation_generators)

Wikipedie 2, Wikipedie.org, *Methodology*, (Online resource): [http://en.wikipedia.org/wiki/Methodology](http://en.wikipedia.org/wiki/Methodology)

**JEL Classification: L86, M15**