DOCUMENTATION: A GROWING NEED... A NEW TOOL

Marcia A. Metcalfe

ABSTRACT

The Software Design and Documentation Language (SDDL) has proven to be an effective, automated documentation tool. This paper presents the purpose, timing, and components of documentation. SDDL is introduced and related to the different timing scenarios; its capability to provide reasonable documentation is discussed and demonstrated by the use of several examples.

I. INTRODUCTION

Documentation has typically been viewed as the drudgery of software development. Weinberg refers to documentation as the castor oil of programming [1]; Kleine says that if design is Cinderella, then certainly documentation is her ugly sister [2]. And it logically follows that documentation became seen in this light since, usually, by the time the documentation stage was reached, the main characters were tired of the program and ready to move on to another project.

However, as more and more programs come into existence, and the need for meaningful documentation also increases, it will be necessary to correct this image if reasonable documentation is to be produced. Without it, the task of understanding, updating, and/or modifying a program becomes increasingly difficult (if not impossible) and costly.

This paper is organized in the following manner: the purpose, timing, and various components of 'documentation are discussed. Next, the automated tool, SDDL, is presented and related to the different timing scenarios; its capability to provide meaningful documentation is discussed and demonstrated through the use of several examples. Finally, conclusions are drawn.

A. PURPOSE OF DOCUMENTATION

The primary purpose of documentation is to provide undistorted communication between the parties interested in the software, both

present and future [3]. Documentation should convey such information as:

- What the program does; how it functions
- How the data/information is represented
- How the various segments/routines relate to one another

This information becomes increasingly valuable as modification is required, and the original developer(s) is no longer available or has forgotten program details [4].

B. TIMING OF DOCUMENTATION

Three documentation timing scenarios will be addressed. The first, "before the fact" documentation, usually presents the program specifications and/or design. The second scenario, "concurrent" documentation, occurs throughout the various phases of software development and provides a working vehicle to prevent distortion of ideas, promotes project control, captures design changes, and permits the orderly development of software. It is useful to the development programmer as well as to the maintenance programmer. The third scenario, "after the fact" documentation typically records the history of development, demonstrates that the program works, and provides a means for maintenance [5].

C. COMPONENTS OF DOCUMENTATION

Given that different users need different information, and that no one document could probably ever provide all the information which in the future may be required, the following documentation components have been identified to accommodate the changing requirements of documentation over the program's lifecycle.

From the very early stages of software design forward, several components become essential. One such component is a high-level description (overview) of the program in prose. It should include such information as what function the program is being designed to perform and what its limitations

and assumptions are. As the design progresses, the data structure becomes more defined and the algorithms/procedures which operate on the data are identified and developed. Therefore, data structure diagrams and a calling sequence diagram which shows the interrelationships of the procedures become two additional basic components of documentation.

As the design matures and coding begins, other components surface. These include the procedure (job control runstream) necessary to execute the program. It should contain a written explanation of the various steps, as well as the input/output declarations, definitions, and allocations. Another component includes descriptions of the data files which contain examples of reasonable data in terms of its mode and size/length. Source code listings, preferably with cross reference tables become an invaluable component. Machine specifications can also be documented at this time.

By the time the program code is completed and the usual documentation phase begins, current versions of the previously-mentioned components coupled with actual data files, a sample testcase with output, and references to other related documents (i.e., user's guide, design document, supporting technical papers) will provide a full complement of meaningful documentation.

II. SDDL: A TOOL

The Software Design and Documentation Language which was developed by Henry Kleine [6] at the Jet Propulsion Laboratory of the California Institute of Technology, has as its main objective communication. It facilitates communication between all the characters in the software development process (i.e., managers, customers, designers, development programmers, maintenance programmers, and the machine). SDDL automates the documentation task; it processes input (expressed in natural language or source code) and produces formatted, software documentation.

Further, methodologies have been developed for displaying representations of data, project management, and the direct processing of source code. Automatic features provided by SDDL include a table of contents, module reference tree (calling sequence diagram), and cross reference listing. User-defined features include specific cross reference listings and SDDL keyword definitions.

The remainder of this section will present SDDL's capability to address the changing documentation requirements in a program's lifecycle, and relate them through the various timing scenarios.

A. SDDL AS A DOCUMENTATION TOOL FOR THE DESIGN PHASE ("BEFORE THE FACT")

SDDL has been used successfully to design both SIMSCRIPT and FORTRAN programs. The processing of natural language statements allows a high-level description of the program's function, limitation, and assumptions, which does not have to meet typical programming language syntactical requirements. Figure 1 [7] demonstrates this capability. Capturing the physical representation of the data is also facilitated by SDDL. Figure 2 [7] illustrates this capability; Figure 3 [7] shows a refined design of a data structure.

Automatic features of the SDDL processor include:

- a table of contents
- a module reference tree (forward-calling sequence diagram)
- a module cross reference listing
- logic error detection

Figure 4 [7] is a segment of an automaticallygenerated module reference tree; it provides information regarding the interrelationships of the various, identified program procedures. A glossary of terms can also be facilitated by the SDDL processor.

B. SDDL AS A DOCUMENTATION TOOL FOR THE CODING PHASE ("CONCURRENT")

SDDL provides a working vehicle which facilitates the coding phase of software development (see Figure 5 [7]). This is the point at which coding conventions can be adopted to allow for the direct processing of source code; and the external data file representation can be documented. Procedures for flagging revisions can be instituted; and userspecified cross reference listings (for global variables, data files, footnotes, etc.) can assist the development programmer. Project management techniques (including a calendar of events, milestones, and progress charts) can be incorporated into the document for use during this phase.

C. SDDL AS A TOOL FOR THE DOCUMENTATION PHASE ("AFTER THE FACT")

Two methodologies for using SDDL during the documentation phase have been developed. The first is the direct processing of SIMSCRIPT, and other high-level languages, source code through SDDL; the second is using the SDDL processor to generate a supporting document to existing source code listings.

 The Direct Processing of SIMSCRIPT Source Code

Figure 6 is the result of processing SIMSCRIPT source code through SDDL. The document formatting features enhance both the clarity and flow of control in this routine. User-defined cross reference listings can be generated at this point to capture machine portability considerations and I/O devices. The automatically-generated SDDL features provide additional information about the source code.

 As a Supporting Document for an Existing Program

SDDL can be used to document existing software written in any programming language. Figure 7 [8] shows SDDL being used to capture the physical data representation of an existing FORTRAN program. Meaningful identifiers have been added to clarify the cryptic descriptors; and their instances can be gathered in a user-specified cross reference table. Additionally, variable mode and units of measure have been supplied.

Figure 8 [8] illustrates SDDL being applied to capture the structure/algorithm, at a high level, of an existing FORTRAN program. Natural language statements lend clarity to the routine description; automatic document formatting features lend flow of control visibility.

III. CONCLUSIONS

SDDL can take some of the "drudge" out of documentation by capturing meaningful information during the various phases of software development as well as by transfering the burden to the computer. It allows for details, usually recalled from someone's memory at the end of a project, to be recorded as they occur during the project. This single, automated medium facilitates the communication between the various members in the software development process over time, thereby providing a two-dimensional documentation tool (i.e., between people, over time).

Additionally, SDDL provides a framework for implementing documentation standards. It skews the documentation effort away from the end (when developers are very busy verifying, debugging, and testing) and toward the beginning of the project (when developers are less busy).

When used in a "concurrent" documentation mode, SDDL provides a working vehicle which begins as a designer's tool, then becomes a development programmer's tool, and finally emerges a a maintenance programmer's tool. When used in an "after the fact" documentation mode, SDDL can produce a document in support of existing source code which adds clarity and visibility into the program's actions. Further, SDDL can generate documentation directly from source code.

REFERENCES

- WEINBERG, Gerald M. (1971) <u>The Psychology of Computer Programming</u>. New York: Van Nostrand Reinhold Company.
- KLEINE, Henry (1977) "A Vehicle for Developing Standards for Simulation Programming." Presented at the Winter Simulation Conference, Gaithersburg, MD., December 5-7.
- COOLEY, Belva J. (1977) "Documenting Simulation Studies for Management." Presented at the Winter Simulation Conference, Gaithersburg, MD., December 5-7.
- 4. HARPER, William L. (1973) <u>Data Processing</u>
 <u>Documentation: Standards, Procedures and Applications</u>. <u>Englewood Cliffs: Prentice-Hall, Inc.</u>
- 5. TAUSWORTHE, Robert C. (1976) Standardized Development of Computer Software. Pasadena: Jet Propulsion Laboratory, California Institute of Technology.
- KLEINE, Henry (1977) <u>Software Design and Documentation Language</u>. Pasadena: Jet Propulsion Laboratory, California Institute of Technology.
- 7. CHAMBERLAIN, R.G., P.J. FIRNETT, D.A. HEIMBURGER, M.H. HORTON, B.L. KLEINE, M.A. METCALFE (1978) SAMIS III Design Document. Pasadena: Jet Propulsion Laboratory, Califfornia Institute of Technology.
- 8. Jet Propulsion Laboratory (1978) PARAMET (An Electric Vehicle Simulator) User's Manual.
 Pasadena: Jet Propulsion Laboratory, California Institute of Technology.

```
APRIL 27, 1978
                                                                                                                                                                                                                                              PAGE
                                                                                                                                                                                                                                                            8
RGC
 481 PROGRAM OBJECTIVES
                482
483
484
                * IT IS INTENDED THAT THE SAMIS III PROGRAM FACILITATE STANDARDIZED COMPARISON OF THE RELATIVE 
* ECONOMICS OF COMPETING MANUFACTURING PROCESSES. IT IS ALSO INTENDED THAT IT FACILITATE ASSESSMENT 
* OF COMPLETE SEQUENCES OF PROCESSES WITH RESPECT TO THE LOW-COST SOLAR ARRAY (LSA) PROJECT 
* GOALS. FURTHER, IT IS INTENDED TO PROVIDE INFORMATION THAT WILL HELP IN DETERMINING FRUITFUL 
* AREAS OF RESEARCH.
 486
487
 488
 489
490
 492
                            THE INPUTS TO THE SAMIS III MODEL FALL INTO SEVERAL GROUPS:
 493
                    A) DESCRIPTIONS OF THE ECONOMIC CHARACTERISTICS OF EACH MANUFACTURING PROCESS/MACHINE
1) PROCESS PARAMETERS (PRODUCT PRODUCED, RATE, DUTY CYCLE, ETC)
2) EQUIPMENT COST FACTORS
3) FACILITIES AND PERSONNEL REQUIREMENTS (PER MACHINE)
4) BYPRODUCTS PRODUCED AND UTILITIES AND COMMODITIES REQUIRED (PER MINUTE)
5) PRODUCTS USED IN THE PROCESS (AND THE ASSOCIATED YIELDS)
494
497
 498
499
                * B) DESCRIPTION OF THE TECHNOLOGICAL AND ECONOMIC STRUCTURE
* 1) OF FIRMS IN THE INDUSTRY
* 2) OF PROCESSES IN EACH FIRM
501
503
 504
                * C) STANDARD DATA

* 1) PRICES OF PERSONNEL, COMMODITIES, ETC. AS FUNCTIONS OF QUANTITIES

* 2) INDIRECT REQUIREMENTS AS FUNCTIONS OF QUANTITIES

* 3) RELATIONSHIPS FOR ESTIMATING INITIAL CAPITAL

* 4) INFLATION RATES AND OTHER ECONOMIC PARAMETERS
505
 506
507
 508
509
 51ó
                D) RUN TIME DATA

1) RANGE OF DEMANDS FOR PHOTOVOLTAIC POWER

2) RANGE OF ANOTHER PARAMETER TO BE VARIED [TO BE IMPLEMENTED IN A LATER RELEASE]

3) "SWITCH" SETTINGS (SUCH AS THE INTEGRAL.MACHINES.FLAG)
515
                FROM DESCRIPTIONS OF THE MANUFACTURING PROCESSES, DETERMINISTIC EQUATIONS DESCRIBING THE * MANUFACTURING COSTS OF EACH PROCESS, AND BUSINESS COSTS OF EACH FIRM HOUSING ONE OR MORE OF THESE * PROCESSES, THE MODEL ESTIMATES THE PRICES THAT MAY REASONABLY BE EXPECTED FOR SOLAR MODULES. AND * ANY RECOGNIZABLE INTERMEDIATE PRODUCTS THAT ARE USED IN THEIR MANUFACTURE. COST ELEMENTS ARE * ALLOCATED TO EVERY PROCESS.
516
517
518
519
520
521
522
                BY PERFORMING SENSITIVITY ANALYSES OF VARIOUS PARAMETERS INVOLVED IN THE MODEL, AND BY * ANALYZING THE EFFECT ON PRICE OF DIFFERENT INDUSTRY CONFIGURATIONS, INFERENCES CAN BE DRAWN WITH * RESPECT TO RESEARCH PRIORITIES AND THE EFFECTS OF GOVERNMENTAL AND INDUSTRIAL ACTIONS.
523
524
525
               526
527
528
529
547 FND
```

Figure 1. SDDL Processing High-level, Natural Language Statements

```
LINE

13 ATTAINSTRUCTURE NOMENCLATURE

14 DATA_STRUCTURE NOMENCLATURE

15 APPRIL 27, 1978

16 ADATA_STRUCTURE NOMENCLATURE

17 APPRIL 27, 1978

18 APPRIL 27, 1978

19 APPRIL 27, 1978

19 APPRIL 27, 1978

10 APPRIL 27, 1978

11 APPRIL 27, 1978

12 APPRIL 27, 1978

13 APPRIL 27, 1978

14 APPRIL 27, 1978

15 APPRIL 27, 1978

16 APPRIL 27, 1978

17 APPRIL 27, 1978

18 APPRIL 27, 1978

18 APPRIL 27, 1978

19 APPRIL 27, 1978

19 APPRIL 27, 1978

19 APPRIL 27, 1978

19 APPRIL 27, 1978

10 APPRIL 27, 1978

10 APPRIL 27, 1978

11 APPRIL 27, 1978

12 APPRIL 27, 1978

13 APPRIL 27, 1978

14 APPRIL 27, 1978

15 APPRIL 27, 1978

16 APPRIL 27, 1978

17 APPRIL 27, 1978

18 APPRIL 27, 1978

18 APPRIL 27, 1978

18 APPRIL 27, 1978

19 APPRIL 27, 1978

10 APPRIL 27, 1978

10
```

Figure 2. High-level Data Representation

```
APRIL 27, 1978
                                                         PAGE 15
3/24/78 1.9
  172
173
174
175
176
177
    183
184
185
186
187
188
189
190
193
195
198
199
200
201
202
203
204
205
206
207
209
211
212
213
214
215
217
218
219
220
221
224
225
226
227
229
230
236 END_DATA_STRUCTURE
```

Figure 3. Lower-level Data Structure Diagram

```
*********** MODULE REFERENCE TREE *****
                                                                                                                                                                   PAGE . 228
     PAGE
1
               TEAM_MEETINGS_AND_AGENDA
 2
               SCHEDULE_AND_MILESTONES
               PROGRESS_CHART
               MEMORANDA
               ACKNOWLEDGEMENTS
               OBJECTIVES
 6
          8
               READING_CONVENTIONS CALL_A_ROUTINE
 7
8
       200
               NOMENCLATURE
. COST_ACCOUNTS
. CURRENT_TECHNOLOGY
. AVAILABLE_COMPANIES
. CURRENT_CONFIGURATION
. DIRECTORIES
9
10
11
12
13
14
15
16
17
                    GLOBAL_VARIABLES PARSER
                    PARSER
TABLE_EXAMPLES
18
         28
               TOP_LEVEL_COMMANDS
               MANIPULATION_COMMANDS
19
         29
              22222222223333333333444444
       30
180
       180
181
182
183
184
188
127
129
       200
1223
1223
1223
1239
1230
1230
1230
1230
1230
1230
```

Figure 4. Automatically-generated Module Reference Tree .

```
PAGE 30
3/24/78 2.2
                               APRIL 27, 1978
     * THIS IS THE STARTING POINT AND "HOME BASE" FOR THE PROGRAM.
 96
97
    98
99
    NOW INITIALIZE_PROGRAM-----
LOOP UNTIL USER TERMINATES PROGRAM EXECUTION
      NOW GET_THE_NEXT_COMMAND-------*
* YIELDING; COMMAND NUMBER
       SELECT CASE PER COMMAND NUMBER
      CASE 1; "?"
NOW LIST_COMMAND_CHOICES----->(124)
CASE 2; "HELP" OR "EXPLAIN"
NOW HELP_WITH_TOP_LEVEL_COMMANDS----->(122)
CASE 3; "PROMPT"
      CASE 3; "PROMPT"
NOW SET PROMPT LEVEL-----CASE 4; "QUERY"
      CASE 5; "CREATE"
NOW CREATE_AND_MANIPULATE_THE_ENTITY----->( 32)
CASE 6; "FIND"
      12234
1224
1225
1226
1227
1231
1233
1233
1235
1235
1237
      ENDIF
      ENDSELECT
    REPEAT UNLESS USER HAS TERMINATED PROGRAM EXECUTION
139 END PROGRAM MAIN ROUTINE
```

Figure 5. Refined Design of a Routine with Project Management Information, Footnotes, and Various Automatic Formatting Features

```
85
6.1
    ROUTINE TO LIST. DATA. FOR. A. SYSTEM. ENTITY
                                                                                                         . .
 6
7
8
9
         DEFINE EXCEPTION. CODE AND COMMAND AS INTEGER VARIABLES
        NOW INTERPRET.THE.USERS.NEXT.WORD-----GIVEN 2 ''CONTEXT = ENTITY TYPES
AND 0 ''NO TEXT TO BE RETURNED
YIELDING EXCEPTION.CODE AND COMMAND
11123456789012234567890123
        COMMAND IS LESS THAN 1 OR COMMAND IS GREATER THAN 8, NOW PRINT. ERROR. MESSAGE GIVEN 3''OUT OF BOUNDS IN LIST. DATA----->( 23)
            NOW PRINILEARCH...
ELSE
SELECT CASE(COMMAND)'
CASE(1)'
NOW LIST.VEHICLE.DATA-----
CASE(2)'
NOW LIST.CHASSIS.DATA------
RETURN
CASE(3)'
NOW LIST.ENGINE.DATA-----
                33901234567
                    NOW LIST. DRIVING. SCHEDULE. DATA---
                 CASE(8)'
NOW LIST. ENVIRONMENT. DATA-----ENDSELECT
             ALWAYS
48 ALWAYS
49
50 <--RETURN
51 END
```

Figure 6. Actual SIMSCRIPT Source Code Enhanced by the SDDL Processor

LINE	•	PAGE 7
172	PROGRAM DATA_STRUCTURE	
174		
175	**** NOTE **** ALL VARIABLES ARE REAL. EXCEPT THOSE WITH AN I AFTER THE	
176		
177		•
178		•
179		
180	COMMON !BATRY! ALL PARAMETERS PERTAINING TO THE BATTERY	
181		(LB)
182	*CHLIM* BATTERY_CHARGING_LIMIT	(-)
183		(LB)
184		(LB)
185		(-)
186		(SEC)
187	*KDSCH* BATTERY_CHARGE_DISCHARGE_POLYNOMIAL_COEFFICIENTS	(-)
186		(2)
187	*REUNF * REGENERATIVE_BRAKING_FACTOR	(3)
190		(NI/HR)
192		
193		
194	· · · · · · · · · · · · · · · · · · ·	(SO FT)
195		(-)
196	*CD* VEHICLE.DRAG.COEFFICIENT *CDA* VEHICLE.DRAG.COEFFICIENT.FRONTAL.AREA.PRODUCT	(-)
197 198	***************************************	(SO FT)
199		(+)·
200		(LB)
202		1,50
203	COMMON !CONST! CONSTANTS AND FLAGS.	
204	*CFLAG* PRINT.CHANGE.OPTIONS.FLAG. I	(-)
205	*G* ACCELERATION.DUE.TO.GRAVITY	(FT/SEC/SEC)
206	"MHFS" MI.PER.HR.TO.FT.PER.SEC.CONVERSION.FACTOR	(FT-HR/SEC/HI)
207	*PFLAG* PERFORMANCE . REQUIREMENTS . PRINTING . FLAG . I	(-)
208		(-)
210		
211	COMMON :ETA! EFFICIENCY PARAMETERS	,
212	*ECOEF* EFFICIENCY.LINE.CALCULATION.ARRAY	(-)
213	*EFAVE* AVERAGE.EFFICIENCY	(2)
214	*EFFAC* CALCULATED.EFFICIENCY.POINT	(2)
215	*EFTIM* AVERAGE.EFFICIENCY.TIME.PERIOD. I	(SEC)
216	*ESPED* EFFICIENCY.SPEED.CALCULATION.BREAKPOINTS	(MI/HR)
217	*EVALU* EFFICIENCY.VALUE.CALCULATION.BREAKPOINTS	(\$)
219		•
220	COMMON !NOW! TRACE PARAMETERS AND INSTANTANEOUS VALUES	(SEC)
221 222	*PRHAX* TRACE.PRINTING.TERMINATE.TIHE. I *PRHIN* TRACE.PRINTING.START.TIHE. I	(SEC)
222	*PRMIN* TRACE.PRINTING.START.TIME, I *PTOT* INSTANTANEOUS.TOTAL.POWER	(HP)
223	*V* INSTANTANEOUS.SPEED	(MI/HR)
224	-4- Tubleulaurund/32250	441,441
227	COMMON SOUTPTS CALCULATED PARAMETERS WHICH WILL BE PRINTED	
228	· *COST* OPERATING.COST	(\$/NI)
229	*E* CONSUMED.ENERGY.OVER.CYCLE	(KU-HR)
230	*EPH* CONSUMED-ENERGY.PER.DISTANCE.OVER.CYCLE	(KW-HR/MI)
231	"HPWHL" AVERAGE.POWER.AT.WHEELS.OVER.CYCLE	(HP)
232	"RANGE" TOTAL RANGE ON A BATTERY CHARGE	(EM)
234		*****

Figure 7. SDDL Used to Document Data Structures in an Existing FORTRAN Program

```
LINE
                                                                                     PAGE
358 PROGRAM MAIN PROGRAM
359
360
      361
362
      S THE MAIN PROGRAM ESSENTIALLY IS USED TO CALL THE PRINCIPAL SUBROUTINES S
 364
      365
366
      COMMON !BATRY! !CAR! !CONST! !ETA! !NO.! !OUTPT! !POWER!
369
370
      ***** NOTE **** NO VARIABLES ARE TRANSFERRED IN THE MAIN LINE SUBROUTINES IN "CALL" STATEMENTS. ALL VARIABLES ARE TRANSFERRED IN THE LABELLED COMMONS
371
373 LABEL: START
374
      NOW INITIALIZE_VARIABLES (INLINE)---
375
      PRINT HELCOME MESSAGE
376
377
378
      CALL SUBROUTINE_QUERY TO INPUT CAR PARAMETERS---
379
380 LABEL: NEW-OR-RECALCULATE-DRIVING-CYCLE
381
382
      383
      IF CYCLE = 0. THAT IS CONSTANT SPEED IS SELECTED
386
387
         388
389
390
391
392
393
394
395
      ELSE ANY OF THE DRIVING CYCLES HAS BEEN SELECTED
         CALL SUBROUTINE_CCALC TO CALCULATE FOR DRIVING CYCLE------CALL SUBROUTINE_CRSLT TO PRINT RESULTS------
395
397
393
399
      400
      *YIELD USER CHOICE
402
403
      SELECT CASE PER USER CHOICE CASE 1: LOGOFF
        --EXITPROGRAM
404
405
      CASE 2: RESTART
      CASE 2: RESTAR!
----GO TO START
CASE 3: USER WANTS TO CHANGE DRIVING CYCLE
----GO TO NEW-OR-RECALCULATE-DRIVING-CYCLE TO GET A NEW CYCLE
CASE 4: USER CHANGED AN INPUT PARAMETER
-----GO TO NEW-OR-RECALCULATE-DRIVING-CYCLE TO RECALCULATE THE DRIVING CYCLE
PARAMETERS BECAUSE THE VEHICLE PARAMETERS WERE CHANGED
406
407
408
409
411
      ENDSELECT CASE
413
414 ENDPROGRAM MAIN PROGRAM
```

Figure 8. SDDL Used to Document a Routine/Algorithm in an Existing FORTRAN Program