Software Project Management

Durga Prasad Mohapatra
Professor
CSE Deptt.
NIT Rourkela

Project Estimation Techniques cont…

- **Halstead's Software Science**

Halstead's Software Science

- An analytical technique to measure:
  - size, development effort, and development time.

Halstead's Software Science

- Halstead used primitive program parameters:
  - developed expressions for:
    - over all program length,
    - potential minimum volume,
    - actual volume,
    - language level,
    - effort,
    - development time.

Halstead's Software Science cont…

- For some given program, let:
  - η_1 be the number of unique operators used in the program,
  - η_2 be the number of unique operands used in the program,
  - N1 be the total number of operators used in the program,
  - N2 be the total number of operands used in the program.

Halstead's Software Science cont…

- The terms operators and operands have intuitive meanings,
  - a precise definition of these terms is needed to avoid ambiguities.
  - Unfortunately there is no general agreement among researchers
    - on definition of operators and operands:

Operators

- Some general guidelines can be provided:
  - All assignment, arithmetic, and logical operators are operators.
  - A pair of parentheses,
    - as well as a block begin --- block end pair, are considered as single operators.
  - A label is considered to be an operator,
    - if it is used as the target of a GOTO statement.

Operators

- An `if ... then ... else ... endif` and a `while ... do` construct are single operators.
- A sequence (statement termination) operator ';' is a single operator.
- function call
  - Function name is an operator,
  - I/O parameters are considered as operands.

Halstead's Software Science cont…

- The set of operators for the ANSI C language: () [] . , -> * + - ~ ! ++ -- * / % + - << >> < > <= >= != == & ^ | && || = *= /= %= += -= <<= >>= &= ^= |= : ? { ; CASE DEFAULT IF ELSE SWITCH WHILE DO FOR GOTO CONTINUE BREAK RETURN and a function name in a function call

Halstead's Software Science cont…

- Operands are those variables and constants
  - which are being used with operators in expressions.
- Note that variable names appearing in declarations
  - are not considered as operands.

Examples

- In the expression `a = &b;`
  - {a, b} are operands and
  - { =, &} are operators.

Examples

- The function name in a function definition
  - not counted as an operator.
  - int func (int a, int b)
    {
    …
    }
    - the operators are: {}, ( )
    - We do not consider func, a, and b as operands.

Examples cont…

- In the function call statement: func (a, b);
  - {func ( ) ;} are considered as a operator
  - variables a, b are treated as operands.

Length and Vocabulary

- Length of a program quantifies
  - total usage of all operators and operands in the program:
  - Thus, length N=N1+N2.
- Program vocabulary:
  - number of unique operators and operands used in the program.
  - program vocabulary $\eta = \eta_1 + \eta_2$.

Program Volume

- The length of a program:
  - total number of operators and operands used in the code
  - depends on the choice of the operators and operands,
    - i.e. for the same program, the length depends on the style of programming.

Program Volume cont …

- We can have highly different measures of length
  - for essentially the same problem.
- To avoid this kind of problem,
  - the notion of program volume V is introduced:
  - $V = N \log_2 \eta$

Potential Minimum Volume

- Intuitively, program volume V denotes
  - minimum number of bits needed to encode the program.
- To represent η different identifiers,
  - we need at least log2 η bits (η is the program vocabulary)

Potential Minimum Volume cont …

- The potential minimum volume V*:
  - volume of the most succinct program in which the program can be coded.

Potential Minimum Volume cont …

- Minimum volume is obtained :
  - when the program can be expressed using a single source code instruction:
    - say a function call like foo().

Potential Minimum Volume cont…

- Lower bound on volume:
  - a program would have at least two operators
  - and no less than the requisite number of operands (i.e. input/output data items).

Potential Minimum Volume cont …

- If an algorithm operates on input/output data d1, d2,... dn,
  - the most succinct program is f(d1,d2, ...,dn);
  - for which $\eta_1 = 2$, $\eta_2 = n$
- Therefore, $V^* = (2+ \eta_2) \log2(2+ \eta_2)$

Potential Level

- The program level L is given by L=V*/V.
- L is a measure of the level of abstraction:
  - languages can be ranked into levels that appear intuitively correct.

Effort and Time

- Effort E=V/L, where
  - E is the number of mental discriminations required to write the program
  - also the effort required to read and understand the program.

Effort and Time cont …

- Thus, programming effort E = (V^2) / V*
  - since L= V*/V varies as the square of the volume.
- Experience shows
  - E is well correlated to the effort needed for maintenance.

Effort and Time cont …

- The programmer's time T=E/S,
  - where S is the speed of mental discriminations developed from psychological results due to Stroud,
  - the recommended value for software is 18.

Length Estimation

- Halstead assumed that it is quite unlikely that a program has several identical parts ---
  - or substrings of length greater than η (η being the program vocabulary).

Length Estimation cont …

- In fact, once a piece of code occurs identically in several places,
  - it is usually made into a procedure or a function.
- Thus, we can safely assume:
  - any program of length N consists of N/η) unique strings of length .

Length Estimation cont…

- It is a standard combinatorial result that for any given alphabet of size K,
  - there are exactly K^r different strings of length r.
- Thus, $N/\eta <= \eta^{\eta}$
- Or, $N <= \eta^{\eta+1}$

Length Estimation cont …

- Since operators and operands usually alternate in a program,
  - we can further refine the upper bound into $N <= \eta (\eta_1)^{\eta 1} (\eta_2)^{\eta 2}$.

Length Estimation cont…

- Also, N must include not only the ordered set of N elements,
  - but it must also include all possible subsets of that ordered set,
  - i.e. the power set of N strings.
  - Therefore, $2^N = \eta\ (\eta_1)^{\eta 1}\ (\eta_2)^{\eta 2}$.

Length Estimation cont …

- $N = \log_2 \eta + \log_2 (\eta_1^{\eta_1} \eta_2^{\eta_2})$
- So, we get, $N = \log_2 (\eta_1^{\eta_1} \eta_2^{\eta_2})$ (approx. by ignoring $\log_2 \eta$)
- Or, $N = \log_2 (\eta_1)^{\eta_1} + \log_2 (\eta_2)^{\eta_2}$
  $= \eta_1 \log_2 \eta_1 + \eta_2 \log_2 \eta_2$
- Experimental analysis of large number of programs suggests:
  - computed and actual lengths match very closely.

Example

```
main()
{
    int a,b,c,avg;
    scanf("%d %d %d",&a,&b,&c);
    avg=(a+b+c)/3;
    printf("avg= %d",avg);
}
```

Example:

- **The unique operators are:**
  main, (), {}, int, scanf,&, ",", ";", =, +, /, printf
- **The unique operands are:**
  a,b,c,&a,&b,&c,a+b+c,avg,3, "%d %d %d", "avg=%d"

Example cont…

- **Therefore, η_1=12, and η_2=11**
- **So, Estimated Length=(12*log12+11*log11)**
  =(12*3.58 + 11*3.45) =(43+38)=81

  Volume=Length*log(23)=81*4.52=366

Summary

- Discussed Halstead's software science for estimating
  - length
  - volume
  - effort
  - Time
- Worked out some examples using Halstead's software science

References:

1. B. Hughes, M. Cotterell, R. Mall, *Software Project Management*, Sixth Edition, McGraw Hill Education (India) Pvt. Ltd., 2018.
2. R. Mall, *Fundamentals of Software Engineering*, Fifth Edition, PHI Learning Pvt. Ltd., 2018.

Thank you