Path Testing and Test Coverage

Chapter 9

Structural Testing

- Also known as glass/white/open box testing
- Structural testing is based on using specific knowledge of the program source text to define test cases
  - Contrast with functional testing where the program text is not seen but only hypothesized

Structural Testing

- Structural testing methods are amenable to
  - Rigorous definitions
    - Control flow, data flow, coverage criteria
  - Mathematical analysis
    - Graphs, path analysis
  - Precise measurement
    - Metrics, coverage analysis

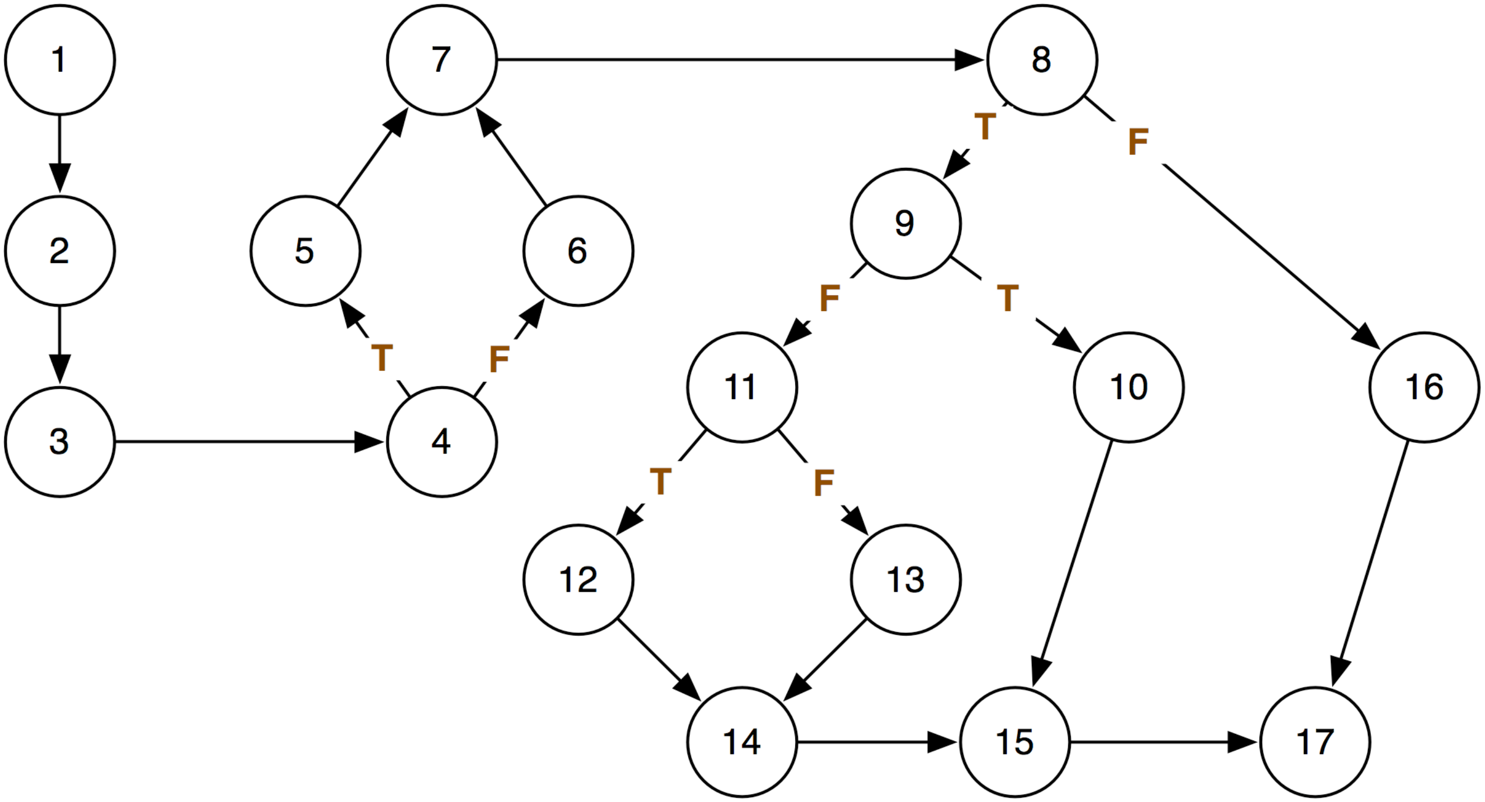
Program Graph - Definition

- Given a program written in an imperative programming language, its **program graph** is a directed graph in which nodes are statements and statement fragments, and edges represent flow of control

Triangle program text

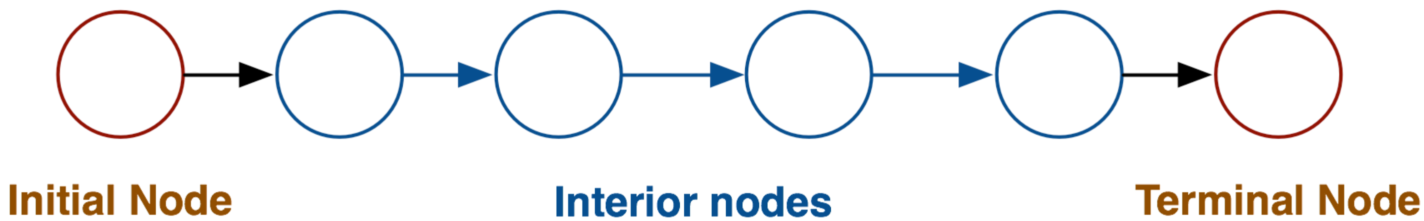
```
output ("Enter 3 integers")
input (a, b, c)
output("Side a,b c: ", a, b, c)
if (a < b) and (b < a+c) and (c < a+b)
then isTriangle ← true
else isTriangle ← false
fi
if isTriangle
then if (a = b) and (b = c)
        else output ("equilateral")
        else if (a ≠ b ) and ( a ≠ c ) and ( b ≠ c)
                then output ("scalene")
                else output("isosceles")
                fi
      fi
  else output ("not a triangle")
  fi
```

Program Graph - Example



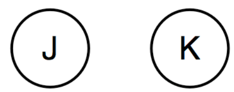
DD-Path – informal definition

- A **decision-to-decision** path (DD-Path) is a path chain in a program graph such that
  - Initial and terminal nodes are distinct
  - Every interior node has indeg =1 and outdeg = 1
    - The initial node is 2-connected to every other node in the path
    - No instances of 1- or 3-connected nodes occur

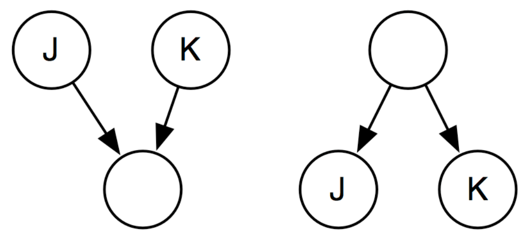


Connectedness definition

- Two nodes n_1 and n_2 in a directed graph are
  - 0-connected iff no path exists between them

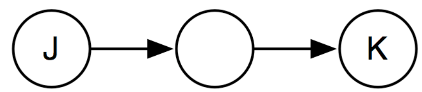


  - 1-connected iff a semi-path but no path exists between them

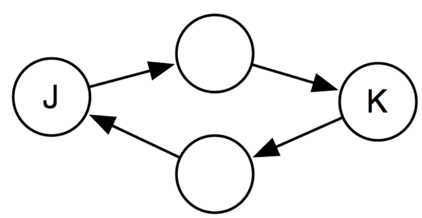


Connectedness definition – 2

- Two nodes n_1 and n_2 in a directed graph are
  - 2-connected iff a path exists between between them



  - 3-connected iff a path goes from n_1 to n_2 , and a path goes from n_2 to n_1



DD-Path – formal definition

- A **decision-to-decision** path (DD-Path) is a chain in a program graph such that:
  - Case 1: consists of a single node with indeg=0
  - Case 2: consists of a single node with outdeg=0
  - Case 3: consists of a single node with indeg $\geq$ 2 or outdeg $\geq$ 2
  - Case 4: consists of a single node with indeg =1, and outdeg = 1
  - Case 5: it is a maximal chain of length $\geq$ 1

- DD-Paths are also known as **segments**

Triangle program DD-paths

Nodes	Path	Case
1	First	1
2,3	A	5
4	B	3
5	C	4
6	D	4
7	E	3
8	F	3
9	G	3

Nodes	Path	Case
10	H	4
11	I	3
12	J	4
13	K	4
14	L	3
15	M	3
16	N	4
17	Last	2

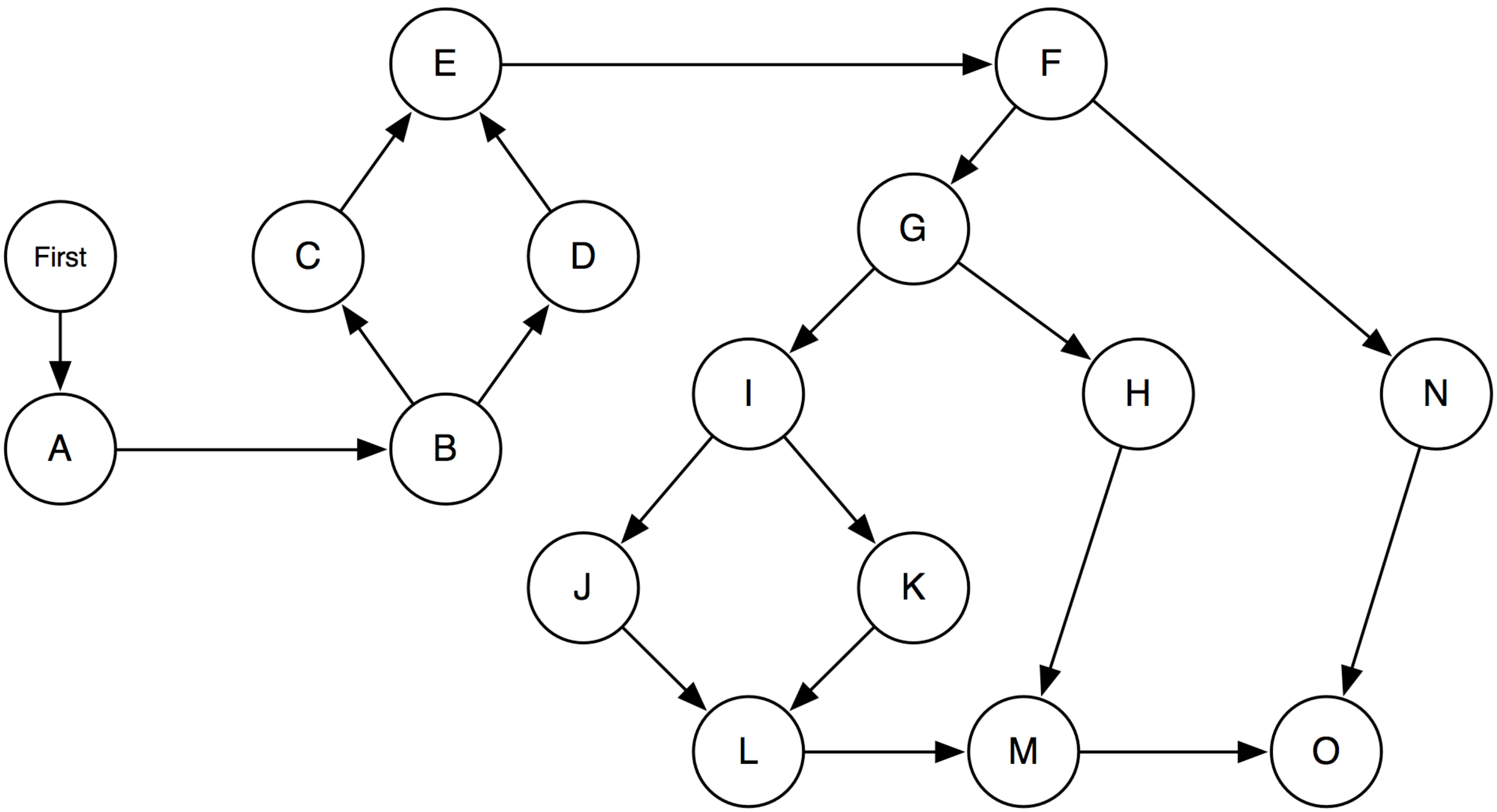
DD-Path Graph – informal definition

- Given a program written in an imperative language, its **DD-Path graph** is a directed graph, in which
  - nodes are DD-Paths of its program graph
  - edges represent control flow between successor DD-Paths.
- Also known as **Control Flow Graph**

Control Flow Graph Derivation

- Straightforward process
- Some judgment is required
- The last statement in a segment must be a predicate, a loop control, a break, or a method exit

Triangle program DD-path graph



Java example program

```
public int displayLastMsg(int nToPrint) {
  np = 0;
  if ((msgCounter > 0) && (nToPrint > 0)) {
    for (int j = lastMsg; (( j != 0) && (np < nToPrint)); --j) {
      System.out.println(messageBuffer[j]);
      ++np;
    }
    if (np < nToPrint) {
      for (int j = SIZE; ((j != 0) && (np < nToPrint)); --j) {
        System.out.println(messageBuffer[j]);
        ++np;
      }
    }
  }
  return np;
}
```

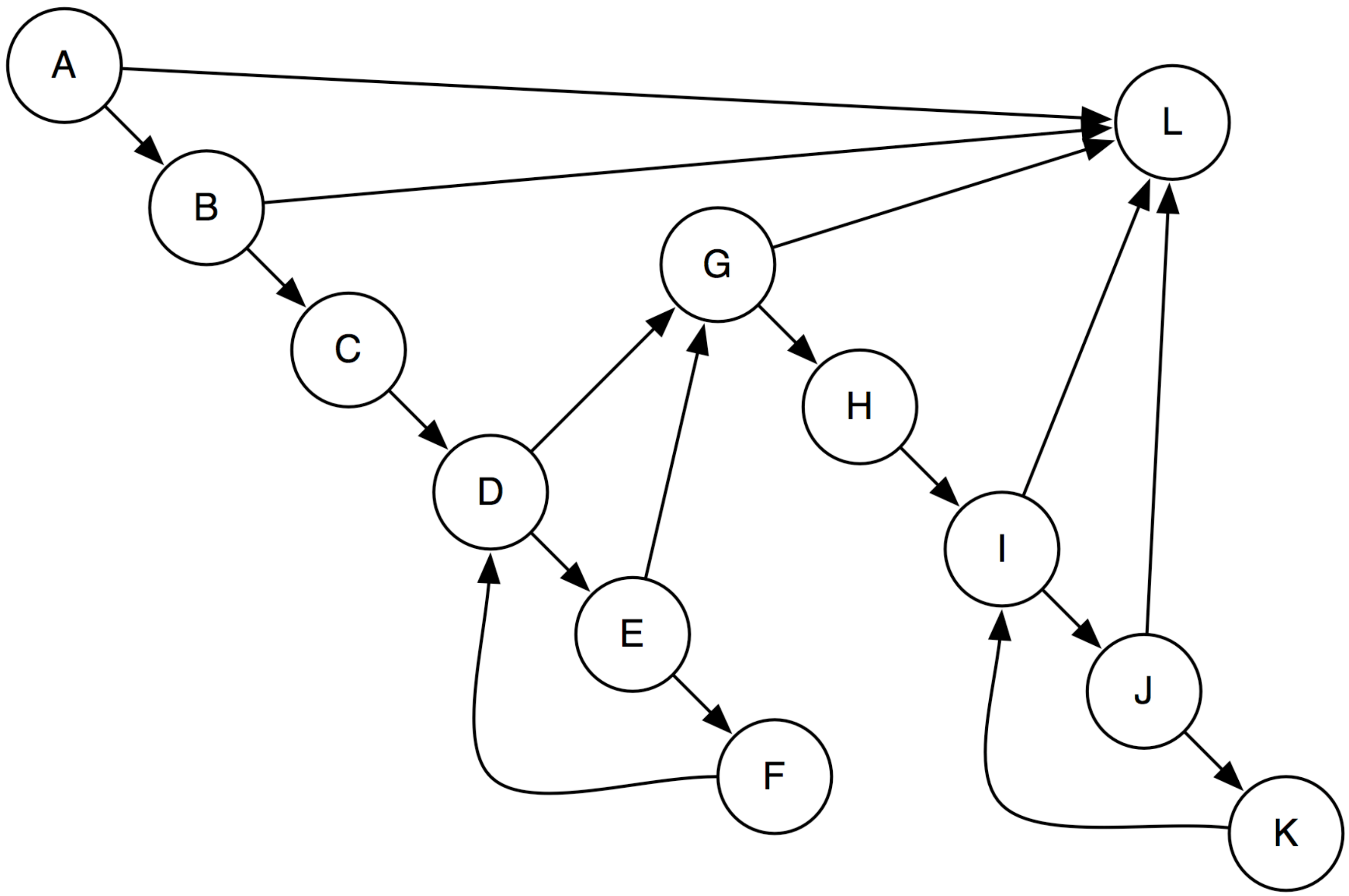
Java example program – Segments part 1

Line		Segment
1	`public int displayLastMsg(int nToPrint) {`	
2	`np = 0;`	A
3	`if (    (msgCounter > 0)`	A
4	`    && (nToPrint > 0))`	B
5	`{ for (int j = lastMsg;`	C
6	`        (    ( j != 0)`	D
7	`          && (np < nToPrint));`	E
8	`        --j)`	F
9	`   { System.out.println(messageBuffer[j]);`	F
10	`     ++np;`	F
11	`   }`	F

Java example program – Segments part 2

Line		Segment
12	`if (np < nToPrint)`	G
13	`{ for (int j = SIZE;`	H
14	`((j != 0) &&`	I
15	`(np < nToPrint));`	J
16	`--j)`	K
17	`{ System.out.println(messageBuffer[j]);`	K
18	`++np;`	K
19	`}`	L
20	`}`	L
21	`}`	L
22	return np;	L
23	`}`	L

Java example program displayLastMsg – DD-path graph



DD graphs definition – 1

- Depict which program segments may be followed by others
- A segment is a node in the CFG
- A conditional transfer of control is a **branch** represented by an edge
- An **entry node** (no inbound edges) represents the entry point to a method
- An **exit node** (no outbound edges) represents an exit point of a method

DD graphs definition – 2

- An **entry-exit path** is a path from the entry node to the exit node
- **Path expressions** represent paths as sequences of nodes
- Loops are represented as segments within parentheses followed by an asterisk
- There are 22 different path expressions in our example

Path expressions – part 1

Entry-Exit path

1	A L
2	A B L
3	A B C D G L
4	A B C D E G L
5	A B C (D E F)* D G L
6	A B C (D E F)* D E G L
7	A B C D G H I L
8	A B C D G H I J L
9	A B C D G H (I J K)* I L
10	A B C (D E F)* D E G H (I J K)* I J L
11	A B C D E G H I L

Path expressions – part 2

Entry-Exit path

12	A B C D E G H I J L
13	A B C D E G H (I J K)* I L
14	A B C D E G H (I J K)* I J L
15	A B C (D E F)* D G H I L
16	A B C (D E F)* D G H I J L
17	A B C (D E F)* D G H (I J K)* I L
18	A B C (D E F)* D G H (I J K)* I J L
19	A B C (D E F)* D E G H I L
20	A B C (D E F)* D E G H I J L
21	A B C (D E F)* D E G H (I J K)* I L
22	A B C (D E F)* D E G H (I J K)* I J L

Paths displayLastMsg – decision table – part 1

Path condition by Segment Name

	Entry/Exit Path	A	B	D	E	G	I	J
1	A L	F	–	–	–	–	–	–
2	A B L	T	F	–	–	–	–	–
3	A B C D G L	T	T	F	–	F	–	–
4	A B C D E G L	T	T	T	F	–	–	–
5	A B C (D E F)* D G L	T	T	T/F	T/–	F	–	–
6	A B C (D E F)* D E G L	T	T	T/T	T/F	F	–	–
7	A B C D G H I L	T	T	F	–	T	F	–
8	A B C D G H I J L	T	T	F	–	T	T	F
9	A B C D G H (I J K)* I L	T	T	F	–	T/F	T/–	T
10	A B C D G H (I J K)* I J L	T	T	F	–	T/T	T/F	T
11	A B C D E G H I L	T	T	T	F	T	F	–

x/x Conditions at loop entry and exit

Branch coverage – decision table example – part 2

Path condition by Segment Name

	Entry/Exit Path	A	B	D	E	G	I	J
12	A B C D E G H I J L	T	T	T	F	T	T	F
13	A B C D E G H (I J K)* I L	T	T	T	F	T	T/F	T/–
14	A B C D E G H (I J K)* I J L	T	T	T	F	T	T/T	T/F
15	A B C (D E F)* D G H I L	T	T	T/F	T/–	T	F	–
16	A B C (D E F)* D G H I J L	T	T	T/T	T/F	T	T	F
17	A B C (D E F)* D G H (I J K)* I L	T	T	T/F	T/–	T	T/F	T/–
18	A B C (D E F)* D G H (I J K)* I J L	T	T	T/F	T/–	T	T/T	T/F
19	A B C (D E F)* D E G H I L	T	T	T/T	T/F	T	F	–
20	A B C (D E F)* D E G H I J L	T	T	T/T	T/F	T	T	F
21	A B C (D E F)* D E G H (I J K)* I L	T	T	T/T	T/F	T	T	T
22	A B C (D E F)* D E G H (I J K)* I J L	T	T	T/T	T/F	T	T	T

x/x Conditions at loop entry and exit

Program text coverage Metrics

- C_0 Every Statement
- C_1 Every DD-path
- C_{1p} Every predicate to each outcome
- C_2 C_1 coverage + loop coverage
- C_d C_1 coverage + every dependent pair of DD-paths
- C_{MCC} Multiple condition coverage
- C_{ik} Every program path that contains k loop repetitions
- C_{stat} Statistically significant faction of the paths
- C_∞ Every executable path

Program text coverage models

- Statement Coverage
- Segment Coverage
- Branch Coverage
- Multiple-Condition Coverage

Statement coverage – C_0

- Achieved when all statements in a method have been executed at least once
- A test case that will follow the path expression below will achieve statement coverage in our example

A B C (D E F)* D G H (I J K)* I L

- One test case is enough to achieve statement coverage!

Segment coverage

- **Segment coverage** counts segments rather than statements
- May produce drastically different numbers
  - Assume two segments P and Q
  - P has one statement, Q has nine
  - Exercising only one of the segments will give 10% or 90% statement coverage
  - Segment coverage will be 50% in both cases

Statement coverage problems

- Predicate may be tested for only one value (misses many bugs)
- Loop bodies may only be iterated once
- Statement coverage can be achieved without branch coverage. Important cases may be missed

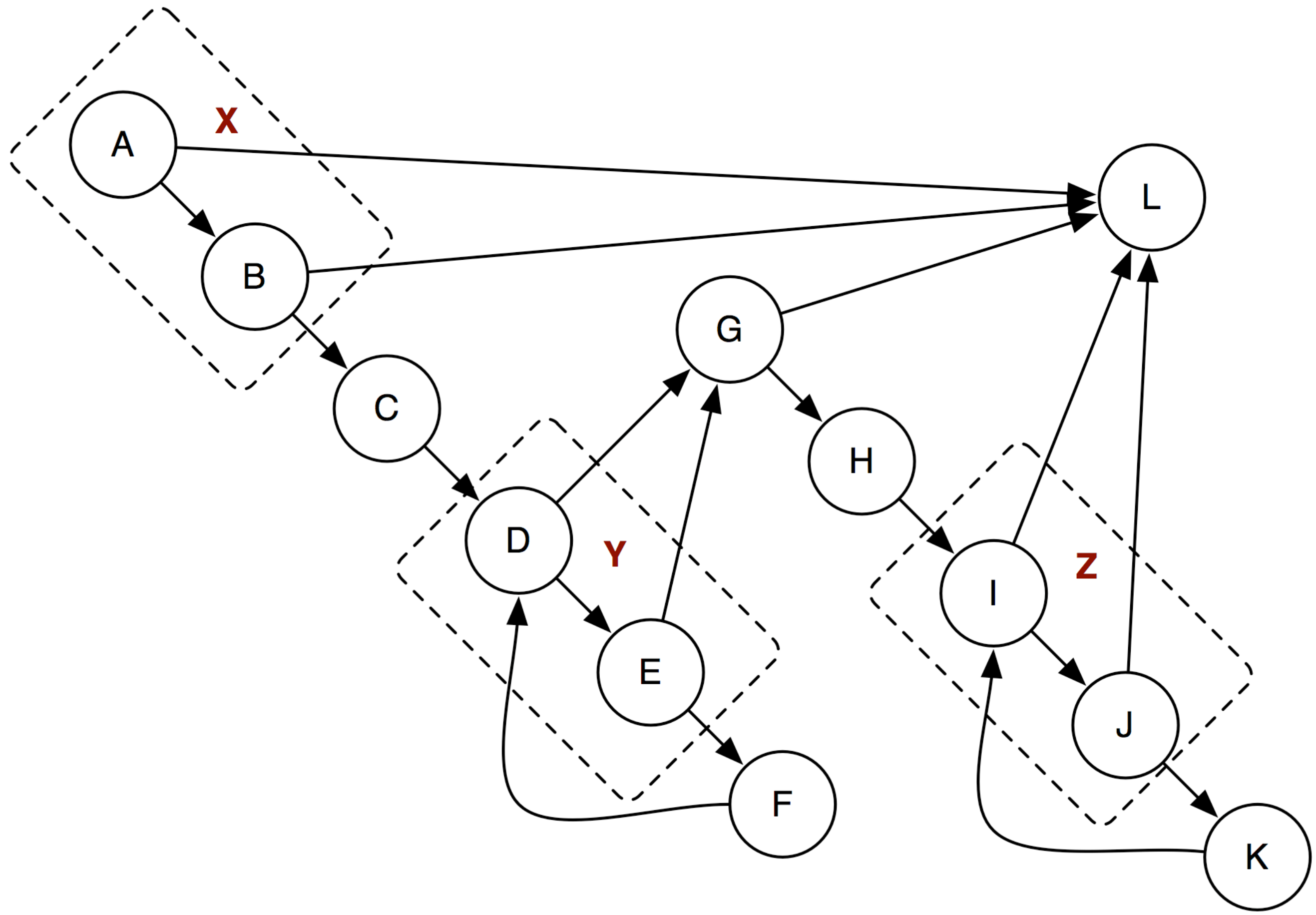
```
String s = null;
if (x != y) s = "Hi";
String s2 = s.substring(1);
```

Branch coverage – C_{1p}

- Achieved when every path from a node is executed at least once
- At least one true and one false evaluation for each predicate
- Can be achieved with D+1 paths in a control flow graph with D 2-way branching nodes and no loops
  - Even less if there are loops
- In the Java example displayLastMsg branch coverage is achieved with three paths – see next few slides

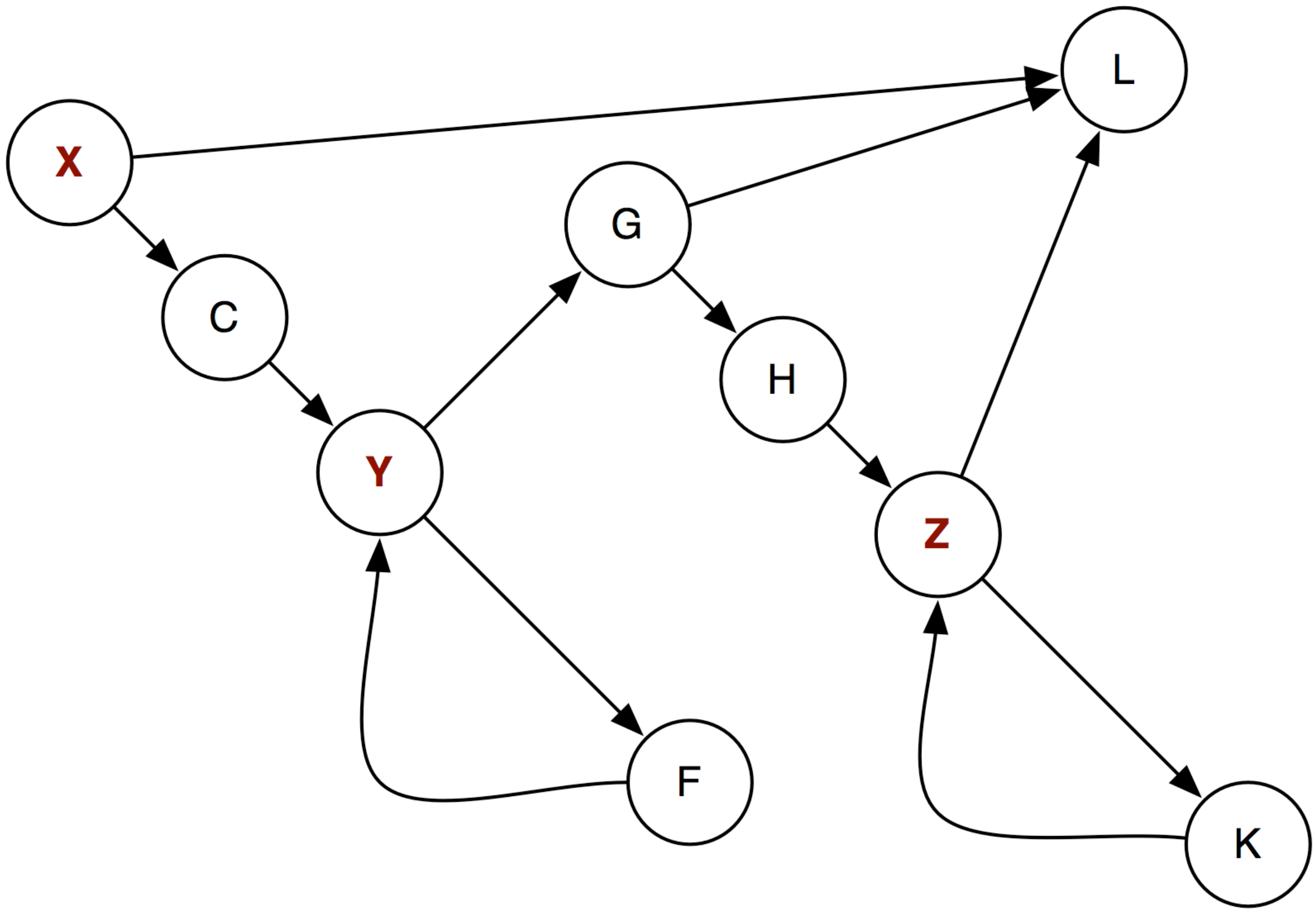
X L
X C (Y F)* Y G L
X C (Y F)* Y G H (Z K)* Z L

Java example program displayLastMsg – DD-path graph



X, Y & Z are shorthand for the nodes within the dotted boxes; used for branch testing

Java example program displastLastMsg – aggregate predicate DD-path graph



Paths aggregate – decision table – part 1

Path condition by Segment Name

	Branch Coverage	A	B	D	E	G	I	J
1	X L	F	–	–	–	–	–	–
2	X L	T	F	–	–	–	–	–
3	X C Y G L	T	T	F	–	F	–	–
4	X C Y G L	T	T	T	F	–	–	–
5	X C (Y F)* Y G L	T	T	T/F	T/–	F	–	–
6	X C (Y F)* Y G L	T	T	T/T	T/F	F	–	–
7	X C Y G H Z L	T	T	F	–	T	F	–
8	X C Y G H Z L	T	T	F	–	T	T	F
9	X C Y G H (Z K)* I L	T	T	F	–	T/F	T/–	T
10	X C Y G H (Z K)* I L	T	T	F	–	T/T	T/F	T
11	X C Y G H Z L	T	T	T	F	T	F	–

x/x Conditions at loop entry and exit

Branch coverage – decision table example – part 2

Path condition by Segment Name

	Branch Coverage	A	B	D	E	G	I	J
12	X C Y G H Z L	T	T	T	F	T	T	F
13	X C Y G H (Z K)* Z L	T	T	T	F	T	T/F	T/–
14	X C Y G H (Z K)* Z L	T	T	T	F	T	T/T	T/F
15	X C (Y F)* Y G H Z L	T	T	T/F	T/–	T	F	–
16	X C (Y F)*Y G H Z L	T	T	T/T	T/F	T	T	F
17	X C (Y F)* Y G H (Z K)* Z L	T	T	T/F	T/–	T	T/F	T/–
18	X C (Y F)* Y G H (Z K)* Z L	T	T	T/F	T/–	T	T/T	T/F
19	X C (Y F)* Y G H Z L	T	T	T/T	T/F	T	F	–
20	X C (Y F)* Y G H Z L	T	T	T/T	T/F	T	T	F
21	X C (Y F)* Y G H (Z K)* Z L	T	T	T/T	T/F	T	T	T
22	X C (Y F)* Y G H (Z K)* Z L	T	T	T/T	T/F	T	T	T

x/x Conditions at loop entry and exit

Branch coverage problems

- Ignores implicit paths from compound paths
  - 11 paths in aggregate model vs 22 in full model
- Short-circuit evaluation means that many predicates might not be evaluated
  - A compound predicate is treated as a single statement. If n clauses, 2^n combinations, but only 2 are tested
- Only a subset of all entry-exit paths is tested
  - Two tests for branch coverage vs 4 tests for path coverage
    - $a = b = x = y = 0$ and $a = x = 0 \wedge b = y = 1$

```
if (a == b) x++;
if (x == y) x--;
```

Multiple-condition coverage

- All true-false combinations of simple conditions in compound predicates are considered at least once
  - Guarantees statement, branch and predicate coverage
  - Does not guarantee path coverage
- A truth table may be necessary
- Not necessarily achievable due to lazy evaluation or mutually exclusive conditions

```
if ((x > 0) && (x < 5)) …
```

Dealing with Loops

- Loops are highly fault-prone, so they need to be tested carefully
- Simple view: Every loop involves a decision to traverse the loop or not
- A bit better: Boundary value analysis on the index variable
- Nested loops have to be tested separately starting with the innermost
- Once loops have been tested then can be condensed to a single node

Basis path testing

- For a vector space a basis set of vectors can be constructed
  - As a consequence every vector in the space is a linear combination of the basis vectors
- By analogy a basis set of paths can be constructed for a DD-path graph
- Problems
  - One cannot assume that testing the basis set is sufficient
  - Basis sets assume independence of members but program text paths are dependent
    - Analogous to variable dependencies causing problems for boundary value testing

Essential complexity

- The cyclomatic number for a graph is given by
  - $CN(G) = e - v + c$
    - e number of edges   v number of vertices
      c number of strongly connected components
      - **For strongly connected, need to add edges from every sink to every source**
- Condensation graphs are based on removing strong components or DD-paths
- For programs remove structured program constructs
  - One entry, one exit constructs for sequences, choices and loops
  - Each structured component once tested can be replaced by a single node when condensing its graph

Essential complexity – 2

- Program text that violates proper structure has
  - Branches either into or out of the middle of a loop
  - Branches either into or out of then and else phrases of if…then…else statements
  - This increases the cyclomatic number – i.e. the complexity of the program
- The higher the cyclomatic number the more tests are required.
  - If complexity is too high
    - Simplify the program rather than do more testing

Guidelines

- Functional testing is too far from the program text
- Path testing is too close to the program text
  - Obscures feasible and infeasible paths
    - Use dataflow testing to move out a bit
- Path testing
  - does not give good help in finding test cases
  - does give good measures of quality of testing through coverage analysis
  - Basis path testing gives a lower bound on the number of tests

Guidelines – 2

- Path testing
  - Provides set of metrics that cross-check functional testing
  - Use to resolve gap and redundancy questions
    - Missing DD-paths – have gaps
    - Repeated DD-paths – have redundancy
- Distinctions are made with the following types of paths
  - Feasible – infeasible
  - Specified – unspecified
  - Topologically possible – impossible

Guidelines – 3

- Re-examine the Venn diagram in the context of path testing