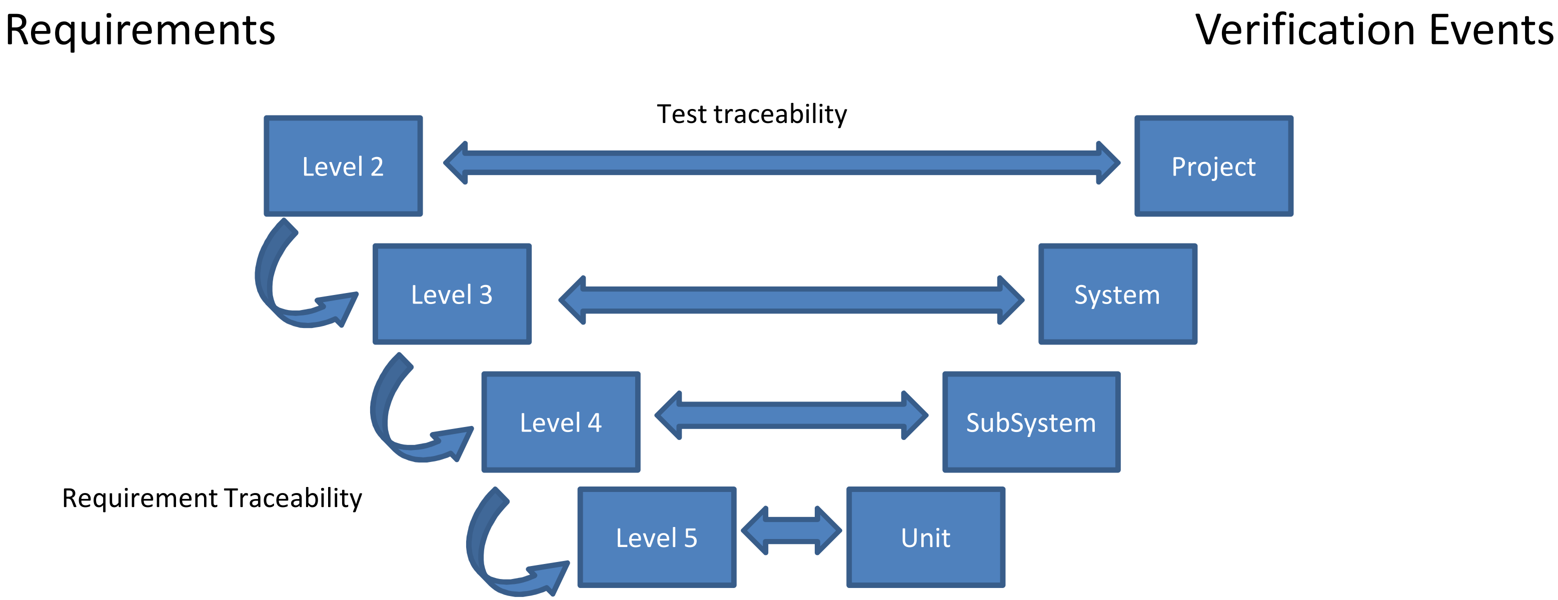
Application of Graph Theory to Requirements Traceability

A methodology for visualization of large requirements sets

Sam Brown
L-3 Communications

Traceability in the Large



Traceability is key to both requirements development and requirements verification
Each project has unique approaches to traceability and verification

Motivation for a Visualization Methodology

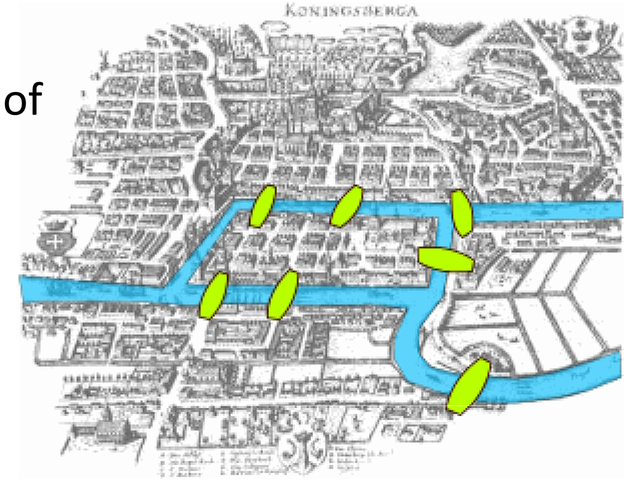
Studying characteristics of information flow in large Requirements sets

>20 documents
>10,000 requirements
>7,000 linkages

LMA	Requirement	FSW	Requirement	TLCM	Requirement
387	The flight system shall support the DOR tone capability in the SDST, including wideband DOR tones at X-band.	1317	The flight software shall command the transponder as defined by the transponder documentation	706	The flight software shall configure the transponder telemetry inputs in accordance with the active FS side whenever the trnsponder is powered ON. Reference transponder ICD for selection table.
1997	The flight system shall accommodate PCM / PSK / PM modulation for the X band downlink.			711	The flight software shall provide the capability to command an “active” telecom side which determines the “active” transponder in use and the uplink channel.
				712	The flight software shall propagate the power state and configuration of the transponder when changing the “active” telecom side.
				713	The flight software shall provide a default “active” telecom side upon initialization. The default “active” telecom side, in the absence of faults or obituary table entries, will be Side 1 (Telecom Side 1 uses SDST 1, and Telecom Side 2 uses SDST 2).
				714	The flight software shall only perform the necessary SDST initializations if a commanded “active” telecom side is different from the currently “active” telecom side.
				725	The flight software, upon initial application of transponder power ON, shall provide for a configurable default state. Subsequent power ON transitions will default to last commanded state.
				1200	The flight software shall provide the capability to enable or disable the X-Band exciter for the active transponder.
				1201	The flight software shall provide the capability to enable or disable c mode for the active transponder.
				1202	The flight software shall provide the capability to enable or disable X-Band Ranging for the active transponder.
				1203	The flight software shall provide the capability to set the Ranging Modulation Index for the active transponder.
				1204	The flight software shall provide the capability to enable or disable X-Band Differential One-Way Ranging (DOR) Mode for the active transponder.
				1205	The flight software shall provide the capability to command an X-Band convolutional encoding mode of TLM_OFF, rate 7 1/2, or BYPASS for the active transponder.
				1206	The flight software shall provide the capability to to command the Ranging Mode to BASEBAND or EXTERNAL for the active transponder.
				1207	The flight software shall provide the capability to command the X-Band Subcarrier for the active transponder to one of the following frequencies: 281.25 Khz squarewave, 281.25 Khz sinewave, 25 Khz squarewave, or 25Khz sinewave.
				1208	The flight software shall provide the capability to command the X-Band Squarewave Telemetry Modulation Index to one of 128 discrete values (0x00 to 0x7F) for the active SDST.
				1209	The flight software shall provide the capability to command the X-Band Sinewave Telemetry Modulation Index to one of 16 discrete values (0x0 to 0xF) for the active SDST.
				1211	The flight software shall provide the capability to command the X-Band telemetry modulation mode to SUBCARRIER or BPSK for the active transponder.

Quickly communicate regarding patterns involving hundreds or thousands of requirements

Graph Theory History



Leonhard Euler: The seven bridges problem
Publication in 1736 as the first description of graph theory, and is generally regarded as the origin of topology

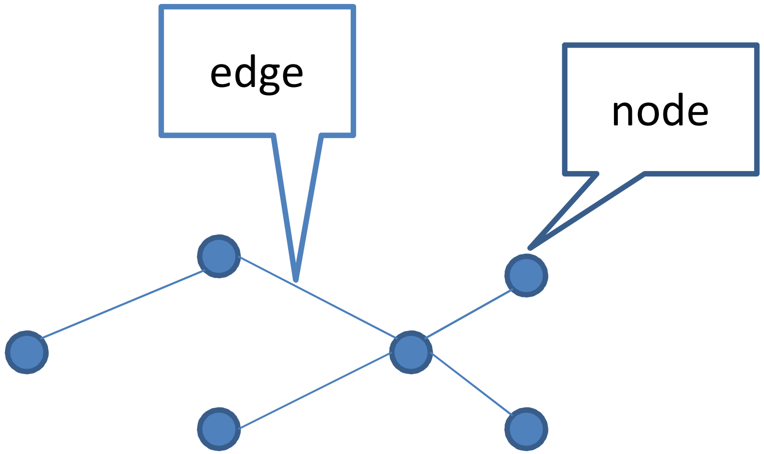
Vanermonde: The knights tour problem

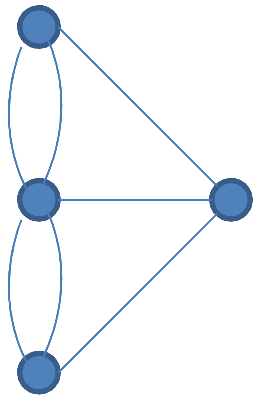
Cauchy and L'Hullier: Relationships between faces, edges, and vertices of convex polyhedrons

Study of pair-wise relationships between objects
Graphs are the parent family to a variety of topologies:
directed graphs
trees – Cayley and differential calculus
coloring problem

What is a graph?

- Graph theory is the study of mathematical structures used to model relationships between objects in finite collections.
- A graph is composed of nodes and edges
- Graphs can be classified as undirected, directed, tree, planar, etc depending upon the nature of the connections.



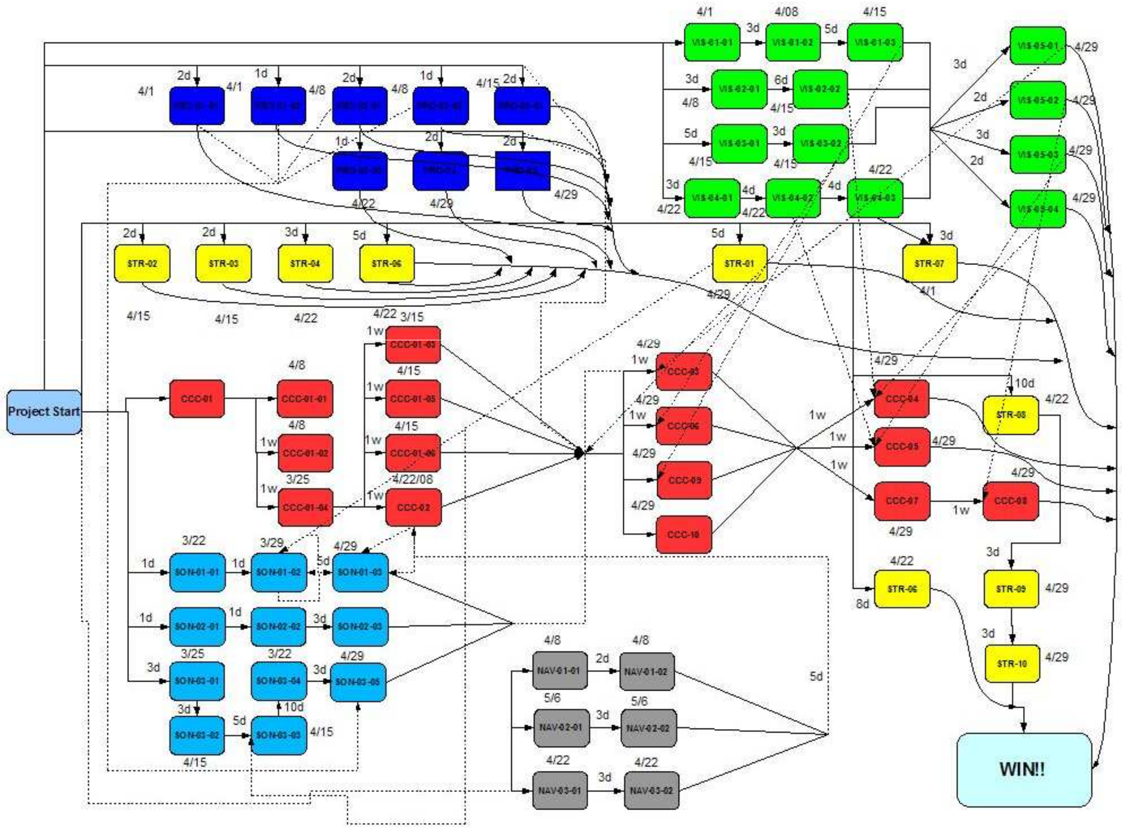


The Seven Bridges Problem
Four nodes, seven edges

Graphs all around us

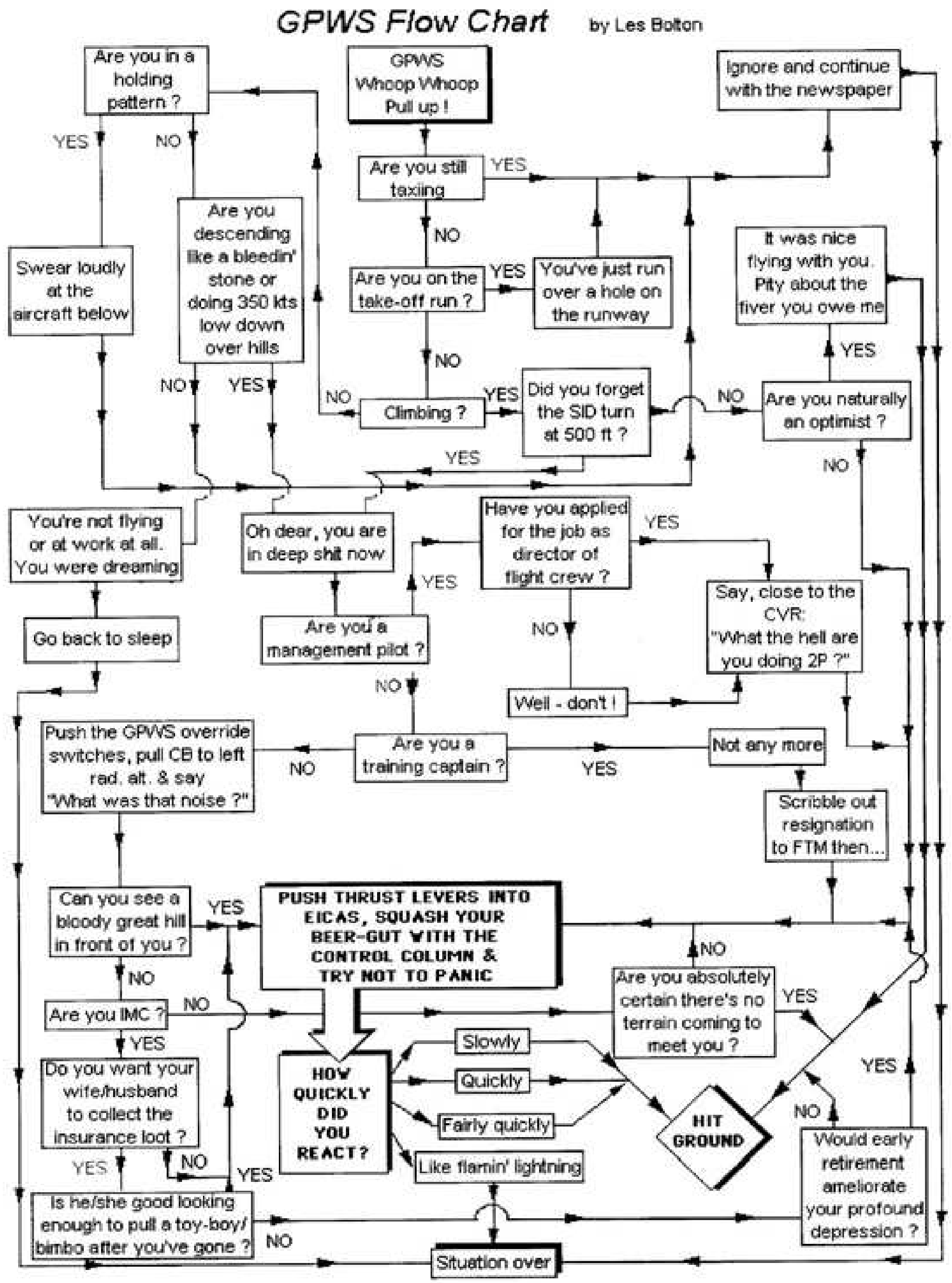
- PERT Chart
  - Directed graph
  - Acyclic (no loops)

Subsystem Pert Chart

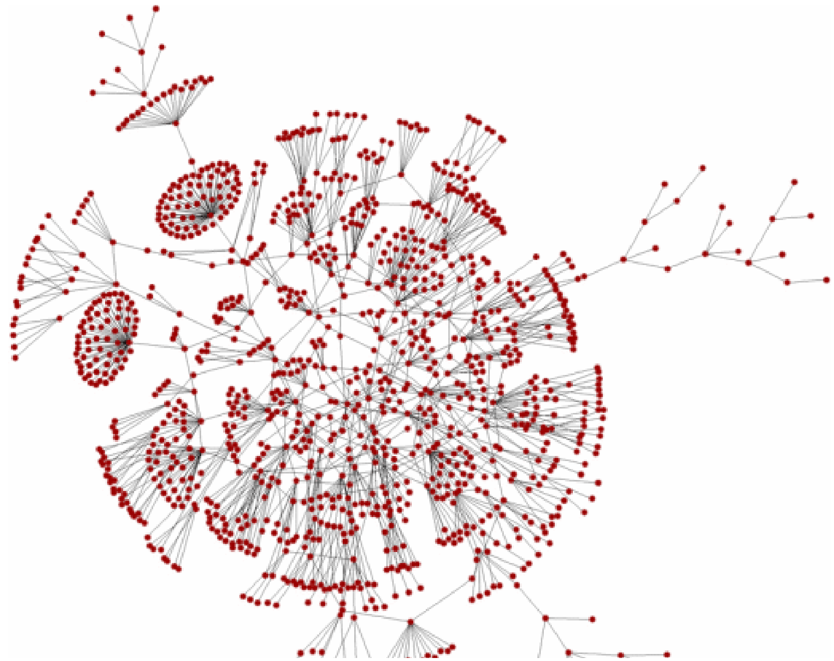


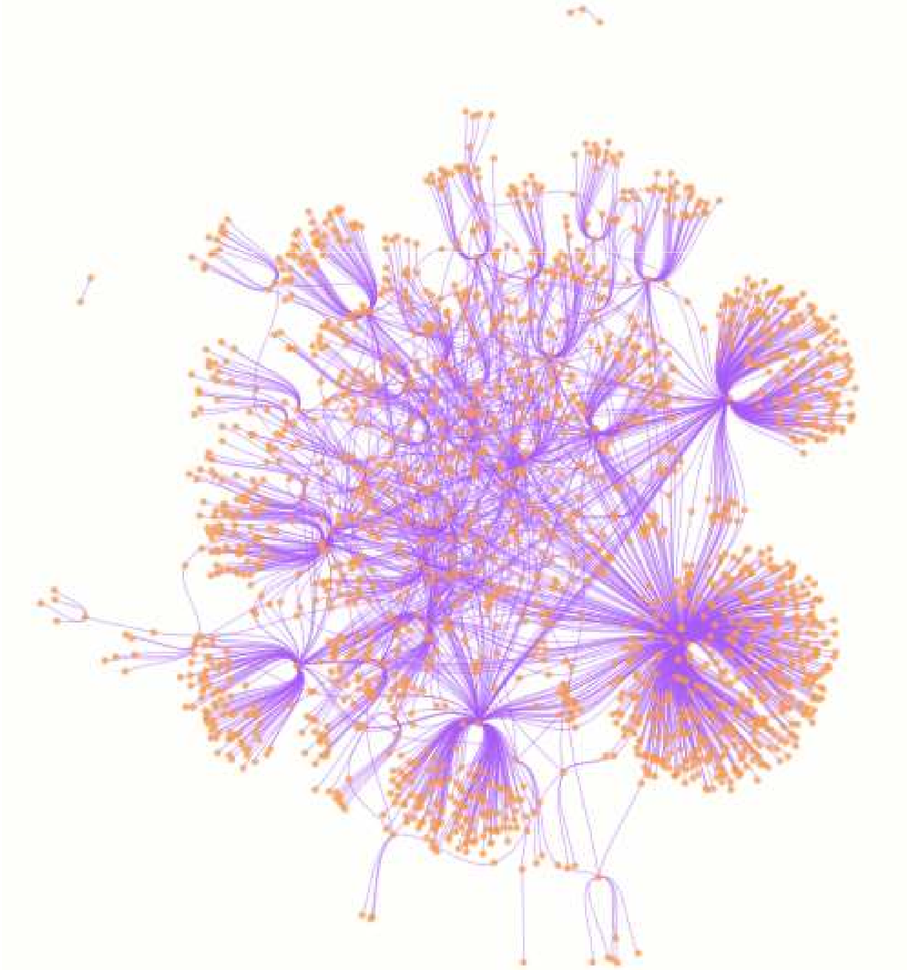
Flow Charts as Graphs

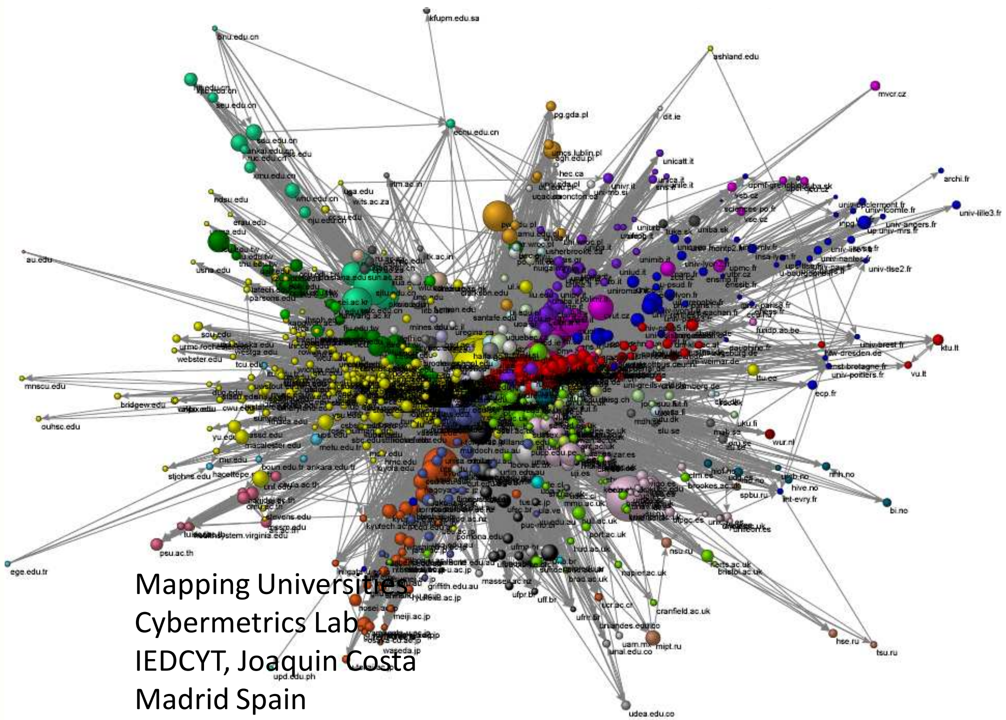
Directed graph
Sometimes cyclic



Network Exploration Graphs







Mapping Universities
Cybermetrics Lab
IEDCYT, Joaquin Costa
Madrid Spain

Graphs of Requirements Sets

Getting to the good stuff soon now…

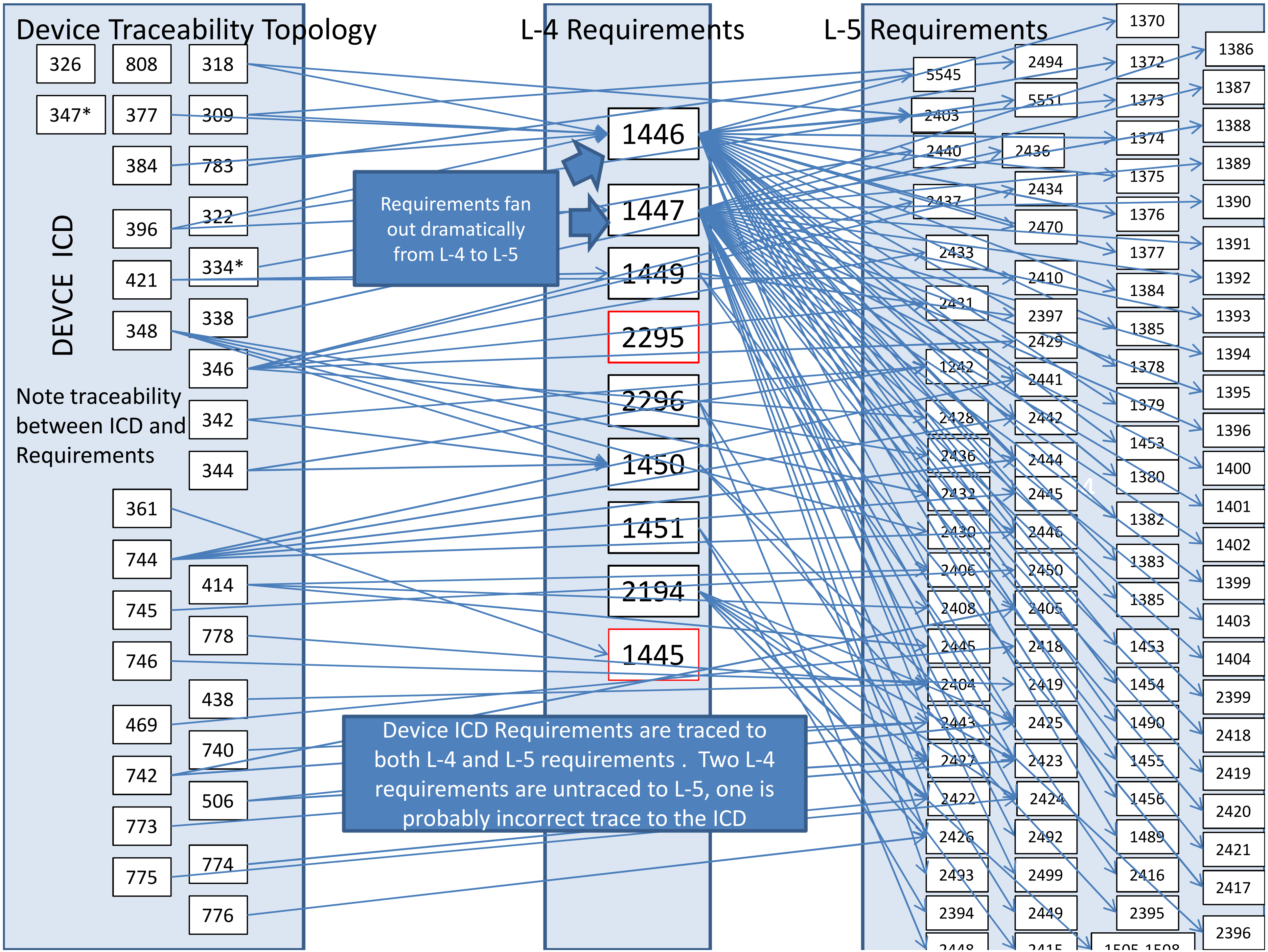
Types of Graphs

- Simple graph – nodes and edges
- Directed graph – nodes and edges with direction (digraph)
- Acyclic graph – no cycles (loops)
- Connected graph – every node is reachable from any other node
- Tree – connected acyclic graph
- Forest – acyclic graph but unconnected

In the general case, requirements traceability forms an acyclic digraph, or forest

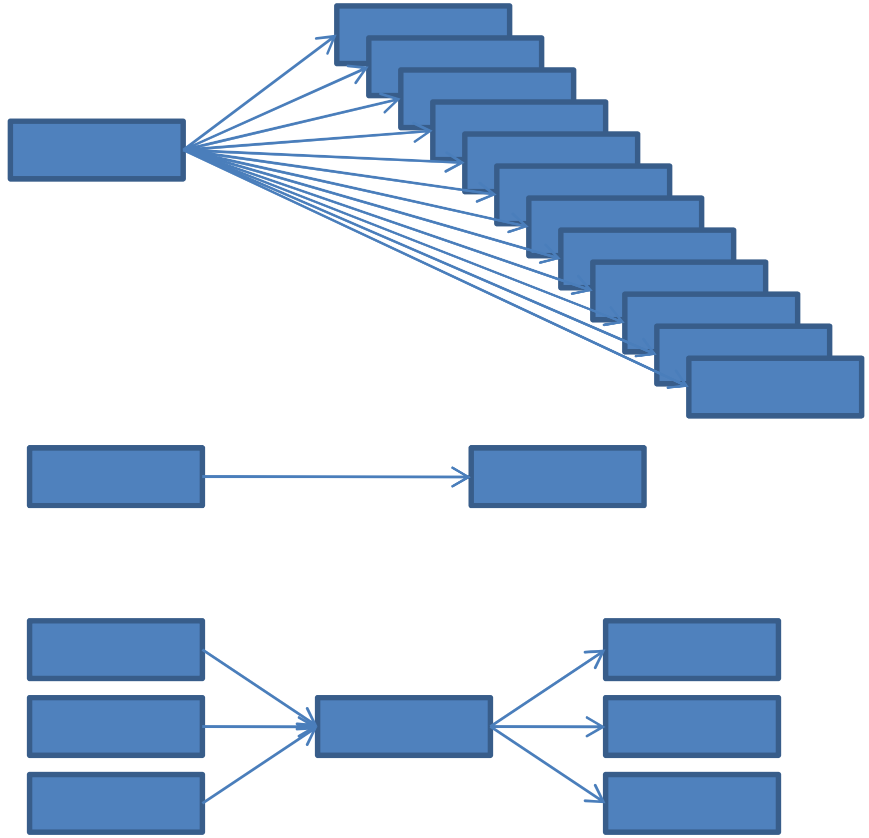
- Generally no single top-level node
- Generally not connected
- Almost always acyclic
- Directed

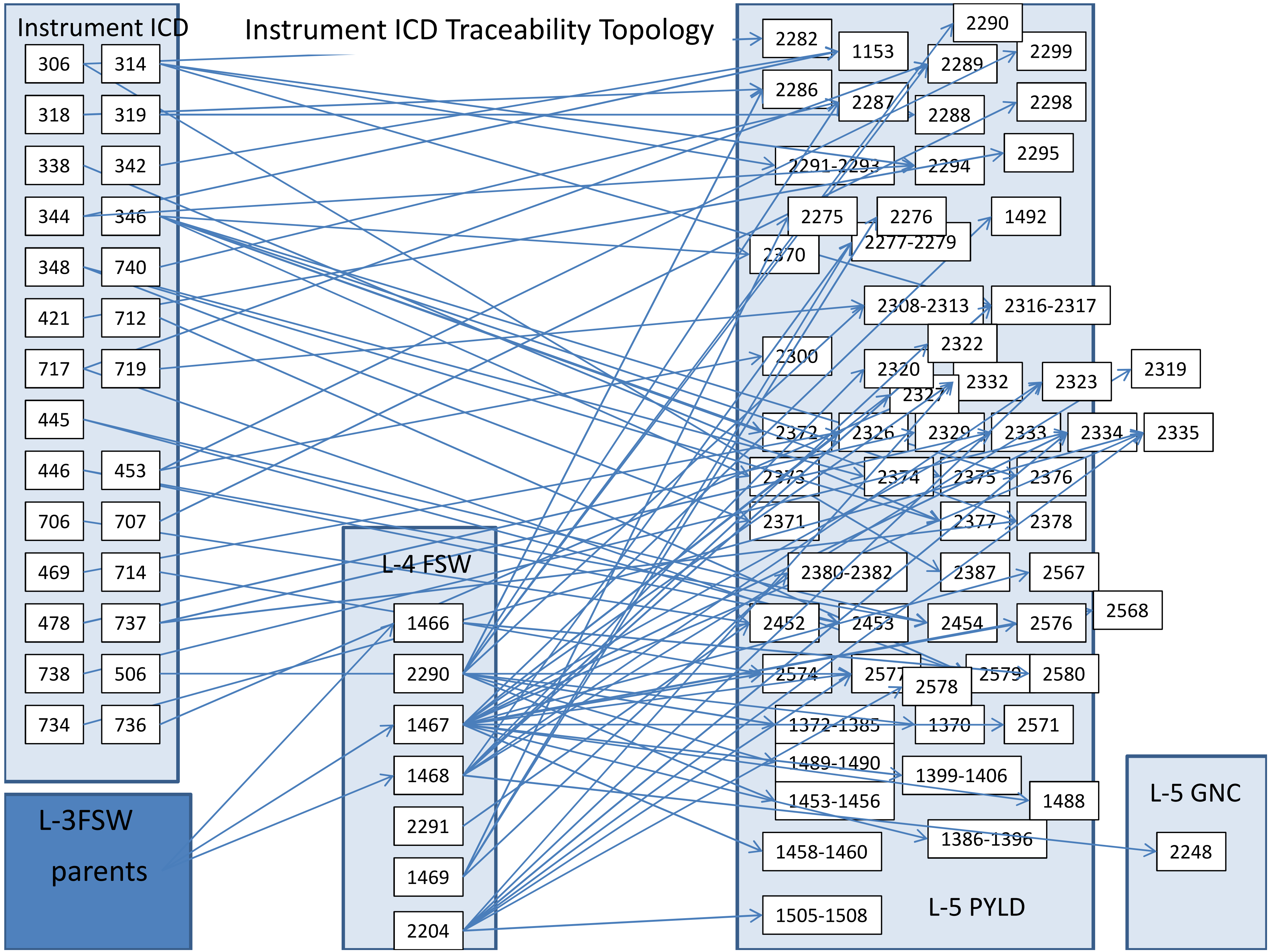
In the following examples of real system requirements graphs, the graphs are drawn as digraphs with the arrow pointing from the parent to the child. Untraced requirements are shown with red borders. We use boxes to denote the nodes simply because they fit the numbers better. These examples show a subnet of the full requirements net for clarity.

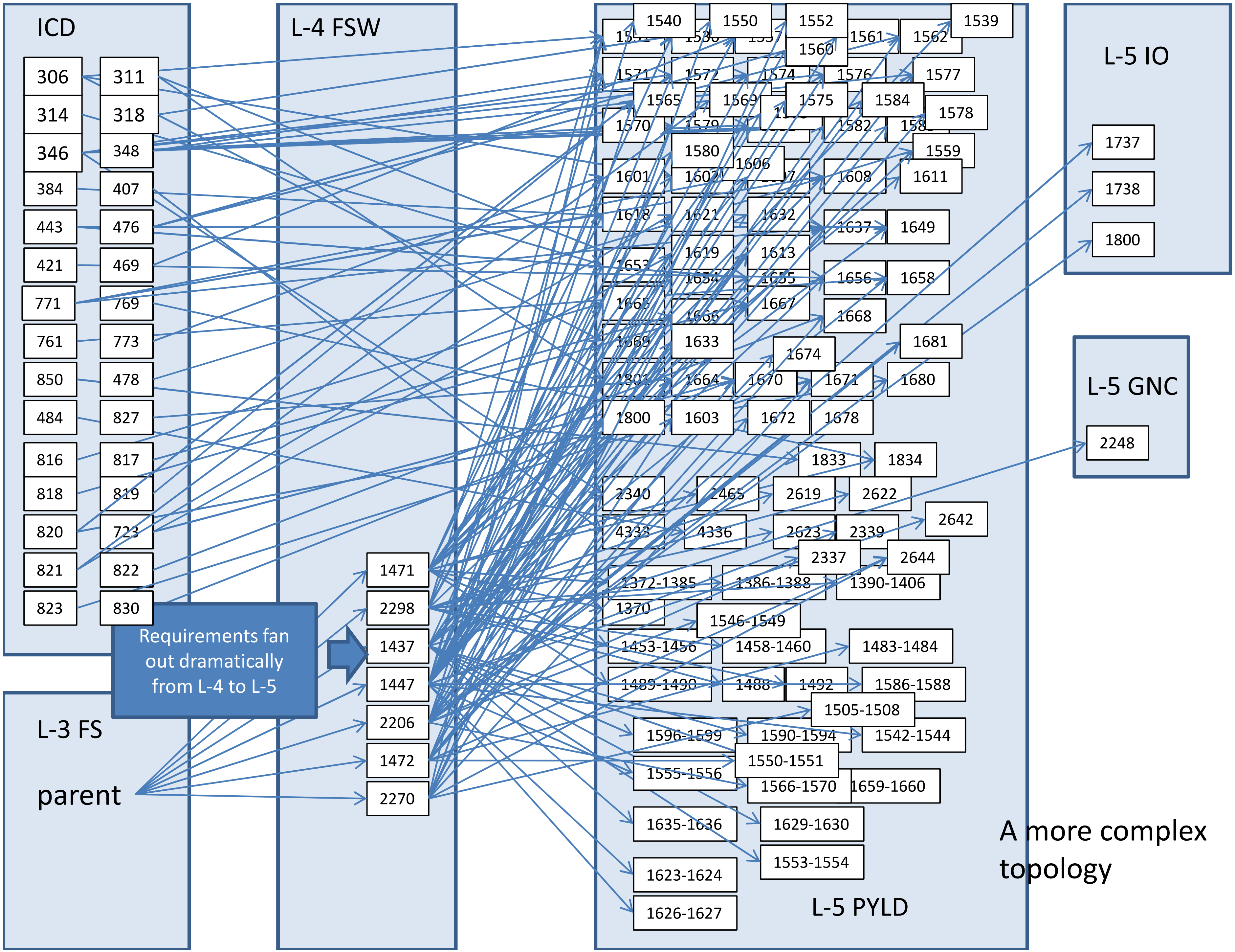

Device Traceability Topology
L-4 Requirements
L-5 Requirements
DEVCE ICD
Note traceability between ICD and Requirements
Requirements fan out dramatically from L-4 to L-5
Device ICD Requirements are traced to both L-4 and L-5 requirements . Two L-4 requirements are untraced to L-5, one is probably incorrect trace to the ICD
326
808
318
347*
377
309
384
783
396
322
421
334*
348
338
346
342
344
361
744
414
745
778
746
438
469
740
742
506
773
774
775
776
1446
1447
1449
2295
2296
1450
1451
2194
1445

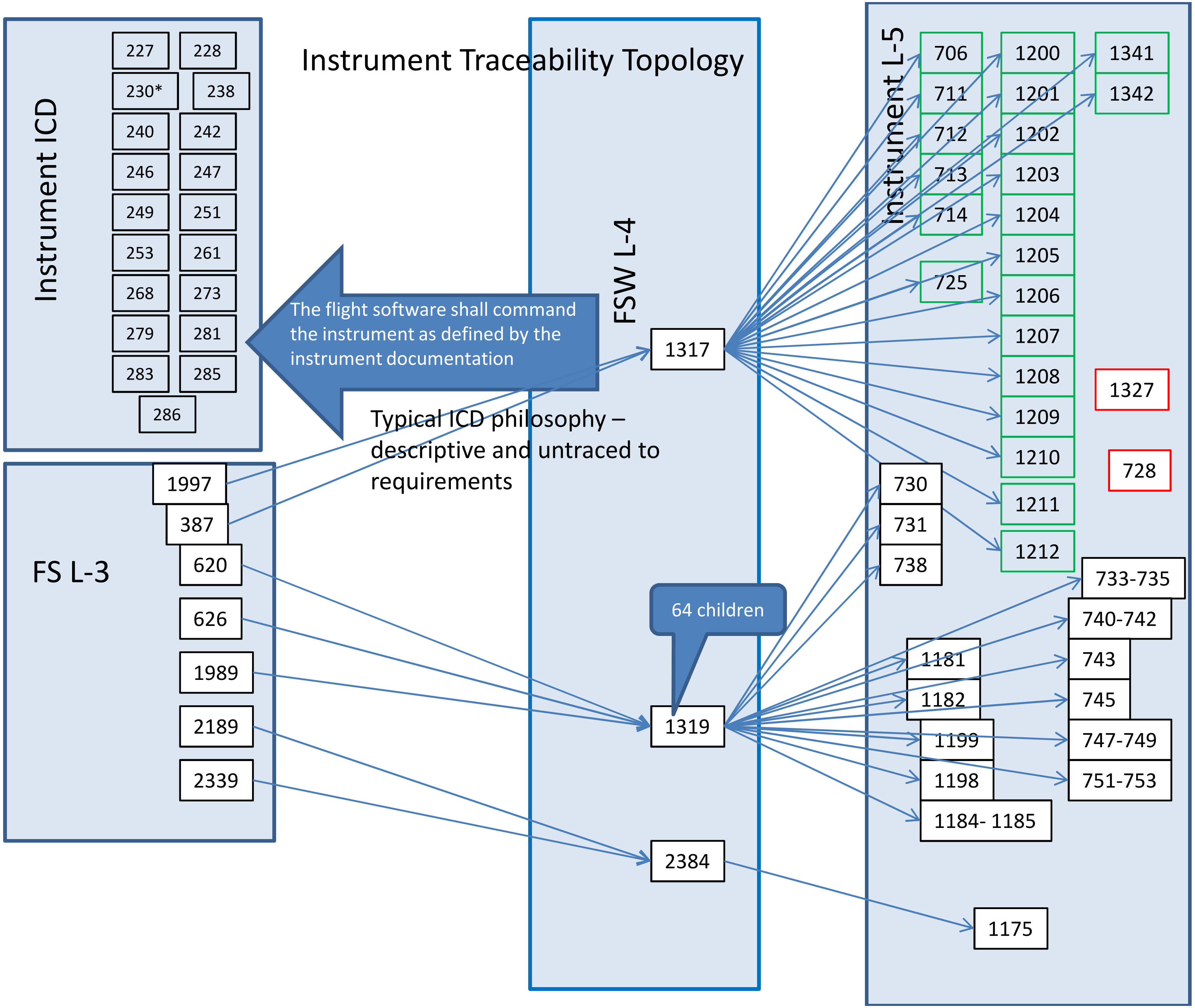
Traceability Patterns

- Large fan-out from parent to child suggest a large change in level of abstraction.
- One-to-one suggests under-specified lower-level requirements.
- Hour-glass traces seem to indicate serious problems in the intermediate requirements document; traceability event-horizon. May indicate verification difficulties.




Instrument ICD Traceability Topology
Instrument ICD
306
314
318
319
338
342
344
346
348
740
421
712
717
719
445
446
453
706
707
469
714
478
737
738
506
734
736
L-3FSW
parents
L-4 FSW
1466
2290
1467
1468
2291
1469
2204
2282
1153
2290
2299
2289
2286
2287
2288
2298
2291-2293
2294
2295
2275
2276
2277-2279
1492
2370
2308-2313
2316-2317
2300
2322
2320
2332
2323
2319
2327
2372
2326
2329
2333
2334
2335
2373
2374
2375
2376
2371
2377
2378
2380-2382
2387
2567
2568
2452
2453
2454
2576
2574
2577
2578
2579
2580
1372-1385
1370
2571
1489-1490
1399-1406
1453-1456
1488
1386-1396
1458-1460
1505-1508
L-5 PYLD
L-5 GNC
2248

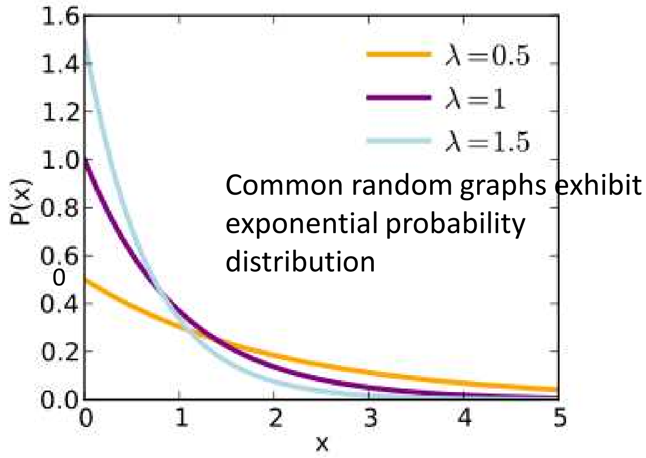

ICD
306
311
314
318
346
348
384
407
443
476
421
469
771
769
761
773
850
478
484
827
816
817
818
819
820
723
821
822
823
830
L-4 FSW
1471
2298
1437
1447
2206
1472
2270
Requirements fan out dramatically from L-4 to L-5
L-3 FS
parent
1540
1550
1552
1539
1560
1561
1562
1571
1577
1565
1569
1575
1584
1578
1570
1582
1580
1559
1606
1601
1608
1611
1618
1621
1632
1637
1649
1619
1613
1656
1658
1667
1668
1633
1681
1674
1670
1671
1680
1664
1800
1603
1672
1678
1833
1834
2340
2465
2619
2622
2642
4333
4336
2623
2339
2337
2644
1372-1385
1386-1388
1390-1406
1370
1546-1549
1453-1456
1458-1460
1483-1484
1489-1490
1488
1492
1586-1588
1505-1508
1596-1599
1590-1594
1542-1544
1550-1551
1555-1556
1566-1570
1659-1660
1635-1636
1629-1630
1553-1554
1623-1624
1626-1627
L-5 PYLD
L-5 IO
1737
1738
1800
L-5 GNC
2248
A more complex topology

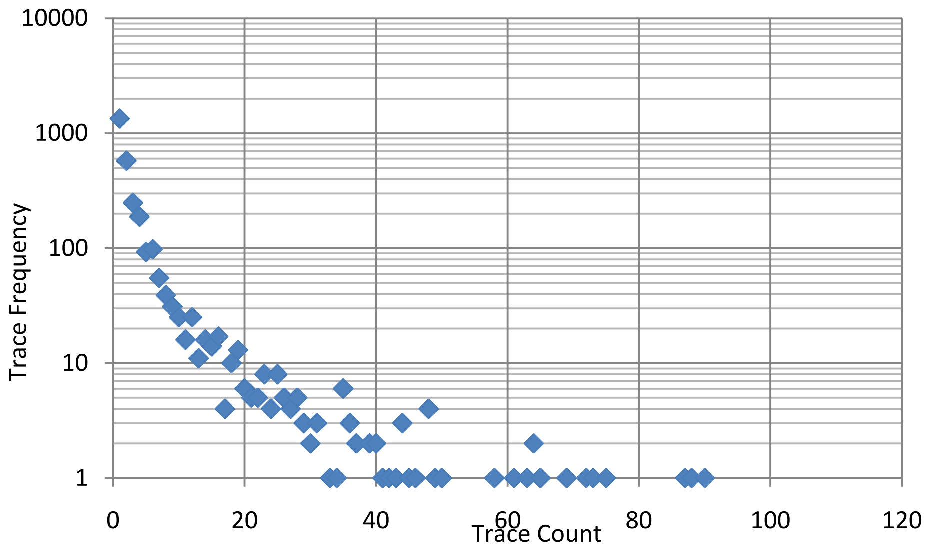


Instrument Traceability Topology
Instrument ICD
227
228
230*
238
240
242
246
247
249
251
253
261
268
273
279
281
283
285
286
The flight software shall command the instrument as defined by the instrument documentation
Typical ICD philosophy – descriptive and untraced to requirements
FSW L-4
1317
1319
64 children
2384
FS L-3
1997
387
620
626
1989
2189
2339
Instrument L-5
706
711
712
713
714
725
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1341
1342
1327
728
730
731
738
733-735
740-742
743
745
747-749
751-753
1181
1182
1199
1198
1184- 1185
1175

Graphs as Traceability Diagnostics

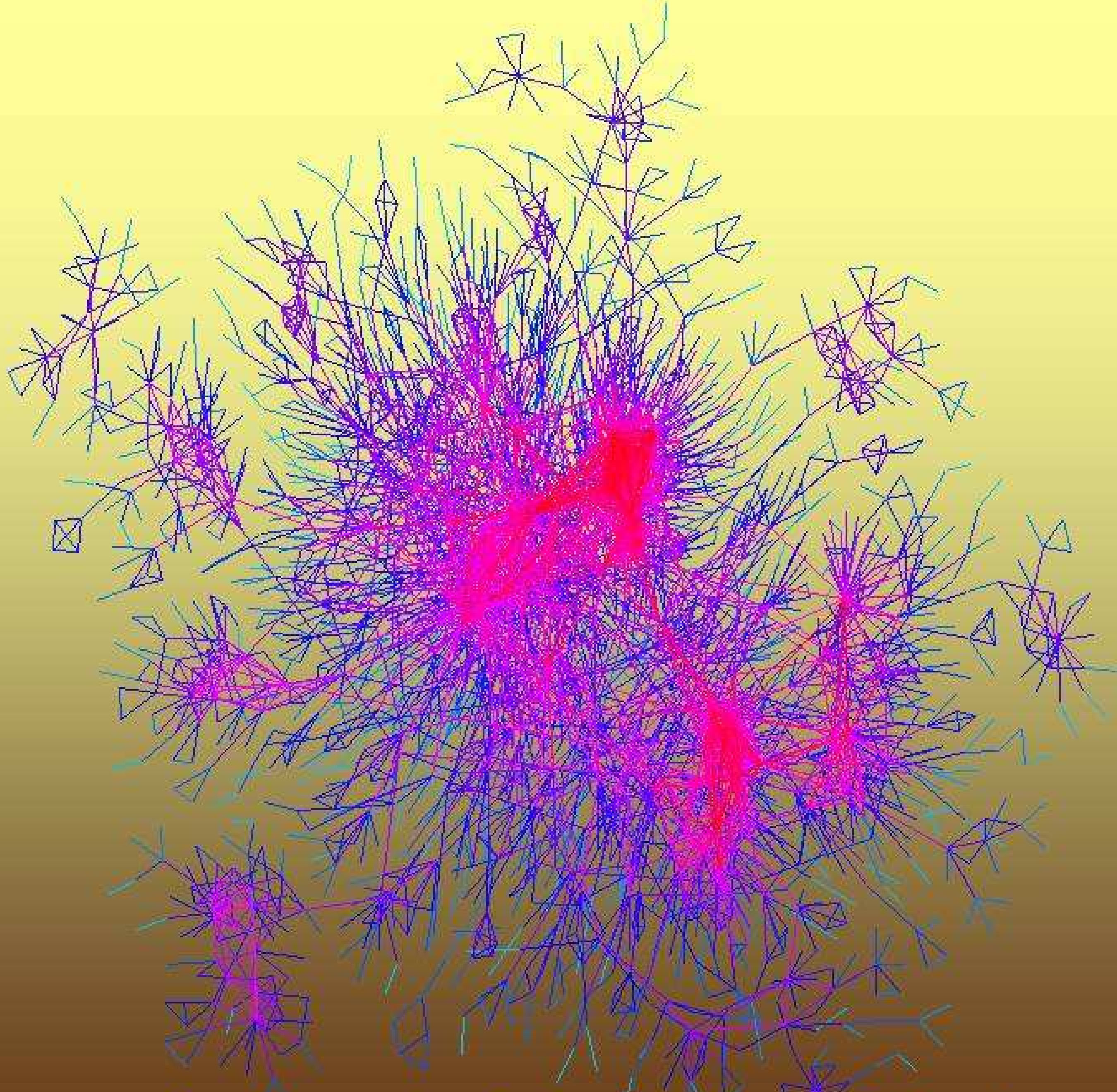
- Histograms of connection counts:
  - Statistics of connection counts may suggest decomposition problems
  - Distributions are typically exponential (Internet, Kevin Bacon – movie graph)





Example real project connection histogram

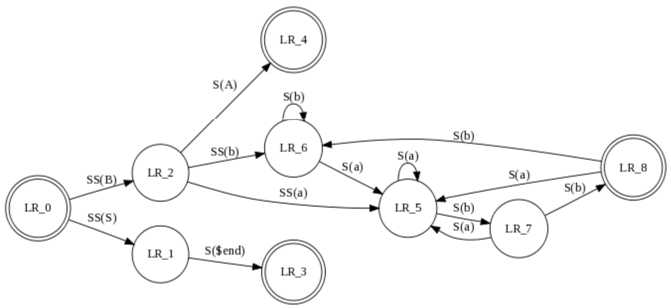
Conjecture: Exponent may be relatable to the overall degree of abstraction change between linked requirements:
High values mean small change

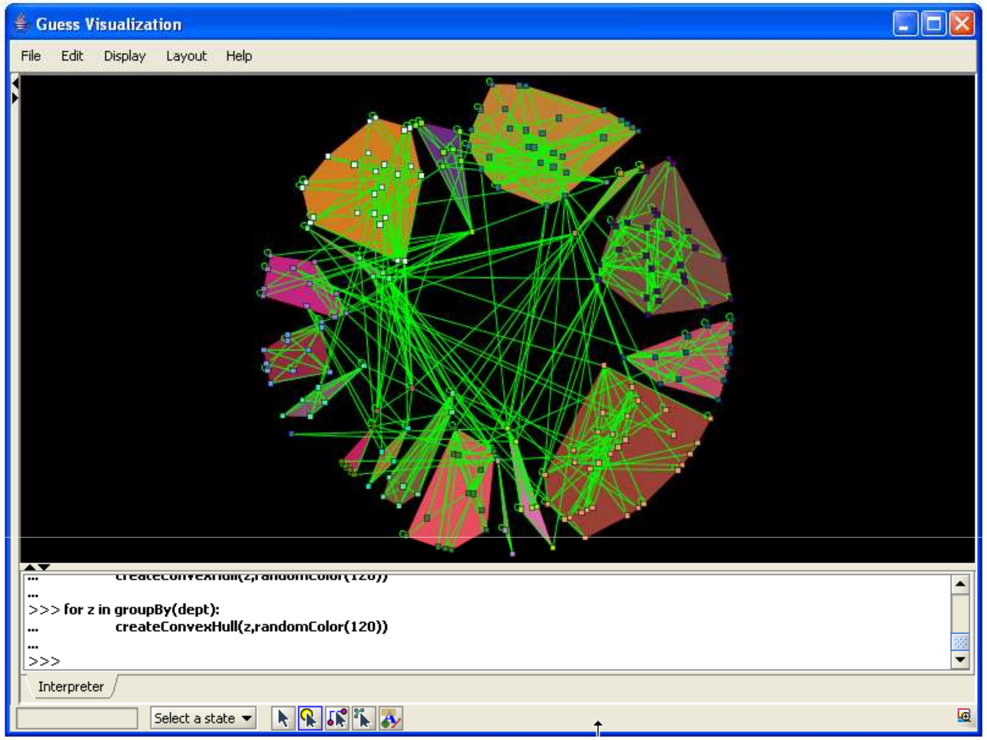


A subgraph of the Hollywood graph.

Automation of the Graphing Process

PowerPoint is NOT the best tool for analysis





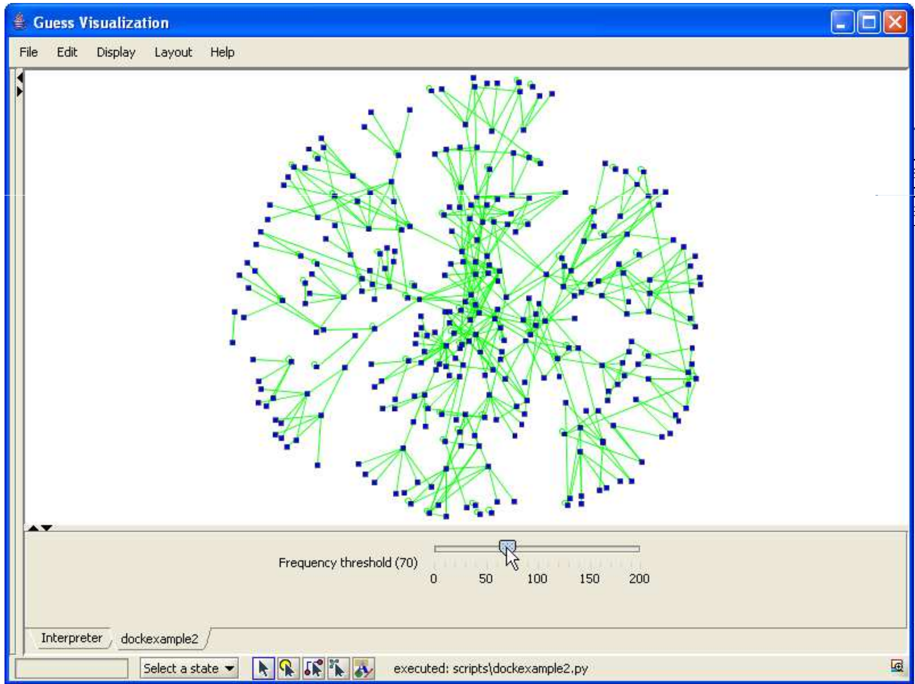
Automatic graph generation from A matrix and specification of groups

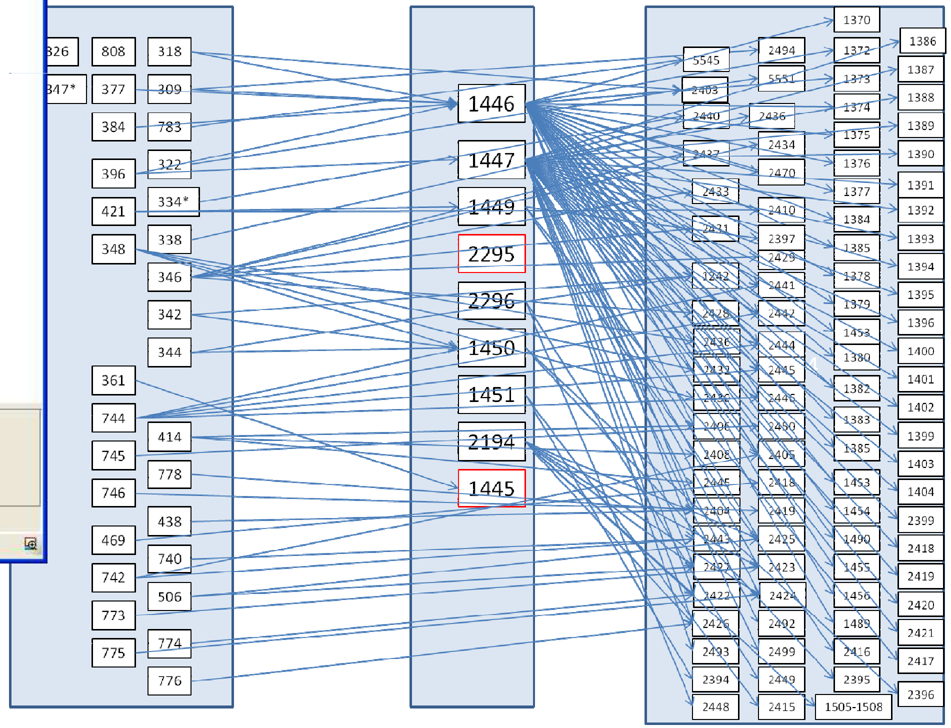
Numerous applications available

- VCG - http://rw4.cs.uni-sb.de/~sander/html/gsvcg1.html
- Graphviz http://www.graphviz.org/
- Jgraph – www.jgraph.com
- Guess - http://graphexploration.cond.org/

How Connected is a Graph?

Separability of subnets -> modularity of requirements to limit propagation of change





Expressing Graphs as Mathematical Structures - Vocabulary

Vertex: Endpoint (or connection point or node)
Edge: Connection between vertices
Incidence List : Array of pairs (tuples if directed) of vertices or connections
Adjacency List: List of pairs of vertices as a list (2x n array)
Incidence Matrix: Vertices by Edges matrix where each entry contains the
endpoint data (1 = incident, 0 = not incident)
Adjacency matrix (A): N by N matrix where N = the number of vertices in the
graph. Entries are either 0 if not connected, 1 if connected.
If there is an edge from vertex k to vertex j then $A(j,k)=1$
Degree: Matrix of connection counts on the diagonal (D)
Laplacian matrix: $L=D-A$, where D= the diagonal degree matrix

Danger: Math Ahead

Connectivity and Graphing

Here comes the math

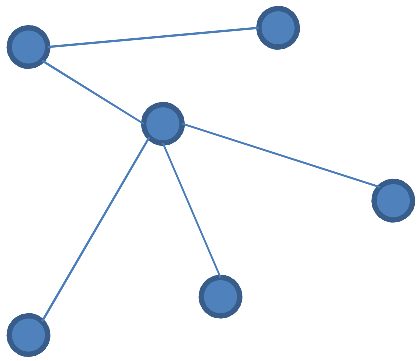
Graph	Fiedler Value
Path	1/n**2
Grid	1/n
3D Grid	n**2/3
Expander	1
Binary tree	1/n
dumbell	1/n

The smallest nonzero eigenvalue of the Laplacian matrix is called the Fiedler value (or spectral gap).

Small values of the Fiedler number mean the graph is easier to cut into two subnets. If the number is large, then every cut of the graph must cut many edges.

Conjecture: Would a large Fiedler number for a requirements graph indicate a system that was difficult to partition into subnets, thus difficult to change?

A Simple Graph and Spectral Analysis



A Matrix

0	1				
1	0	1			
	1	0	1	1	1
		1	0		
		1		0	
		1			0

Laplacian (D-A)

1	-1				
-1	2	-1			
	-1	4	-1	-1	-1
		-1	1		
		-1		1	
		-1			1

Eigenvalue:
From linear algebra
$Lx=\lambda x$ where λ is an eigenvalue
And x is a non-null eigenvector
Because L is symmetric the
eigenvalues are all real

$$\lambda=\{0, 0.486, 1, 1, 2.428, 5.086\}$$

Fiedler number = 0.486
implying somewhere between an
expander (1) and a tree form (1/6)

Summary

- Graphs can be useful visualization tools for large requirements sets
  - Big picture viewpoint
  - Patterns easily recognized
  - Multi-level tracing
  - Identification of subnets
- Potential for analysis
  - Relationship between connection histogram and requirement decomposition
  - Ability to quantify interconnectedness by spectral analysis