Graphs as Visualization Aids

- Graphs are mathematical structures with obvious visualizations that seem often to help many stakeholder communities to visualize key relations.
- A graph’s edges visually represent the ordered pairs that compose the relation.
- If the pairs in E are ordered, then G is a directed graph, and its edges are depicted with arrowheads.
- If not, the graph is called an undirected graph.

What is This Graph Specifying?

Annotations Provide Intuitions

What’s wrong with this diagram

There is ambiguity and misuse of notation here:
- one circle is a test, others are functions
- some edge annotations are data, some predicates
- are multiple arrows in and out “and” or “or”?

Differences in Graphs Result from Different Choices for Nodes & Relations

- Data Flow:
  - Nodes represent set of sites where data is generated/used.
  - Each edge is a (data generated, data used) node pair.
- Control Flow:
  - Nodes represent units of functionality.
  - \((n_1, n_2)\) is an edge in this graph if and only if unit \(n_2\) can execute immediately after \(n_1\) executes (ImmFol relation).
- Hierarchy:
  - Models “consists of” or “is a part of”.
  - Key to divide-and-conquer approaches to understanding.
- Finite State Machines:
  - Nodes represent all possible different execution states.
  - \((s_1, s_2)\) is an edge if and only if it is possible for state \(s_1\) to immediately succeed \(s_2\). Called a transition from \(s_1\) to \(s_2\).
  - Edges annotated with transition condition.
- Petri Nets:
  - Multiple node and edge types in the same diagram.
DATA FLOW DIAGRAMS

- Capture system functionality: What does system do? How?
- Basic components of a data flow diagram:
  --Nodes, represented by circles (boxes), are functional units
  --Edges, represented by arrows, are data flows between units
  --Both augmented by separate annotation relations
  --Boxes (sometimes circles), represent I/O data
<<This is actually yet another relation

Flowgraphs

Let $S = \{ \text{all statements s in a program, P} \}$
Let $\text{ImmFol} = \{ (s_i, s_j) | \text{the execution of s immediately follows the execution of s for some execution of P} \}$
Then: If $FG = (S, \text{ImmFol})$, FG is called the flowgraph of P
FG is an ordered graph
Every execution sequence (ie. the sequence in which the statements of P are executed) for a given execution of P corresponds to a path in FG.
However---the converse is not true. A path through FG may not correspond to an execution sequence for P
A loop in P appears as a cycle in FG

Hierarchy

- Enables incrementally adding detail
- Increased precision too
- Draws upon innate human mental capability
  -- Abstraction
  -- Encapsulation
- A typical solution to the problem of needing detail, but needing to avoid overload
- But creates potential problems

Consistency is a principal concern

- Are the diagrams consistent with each other?
- Top view consistent with elaborations?
  -- Arrows consistent
  -- Data flows consistent
  -- Other semantics?
- Invitation to subtle errors
Consistency is a principal concern

- Are the diagrams consistent with each other?
- Top view consistent with elaborations?
  - Arrows consistent
  - Data flows consistent
  - Other semantics?
- Invitation to subtle errors

Simple Hierarchical Elaboration

Hierarchical Elaboration (of “Select New Floor”)

Hierarchical Elaboration

Is this consistent with its parent?

Inconsistencies with parent
Hierarchical Elaboration (of “Select New Floor”)

Consistency definitions

- Let NodeAnnotation(n, somenode) for some n \in N, where ParentGraph = (N, E)
- Let somenode = (N', E')
- This is the DFG elaborating on "somenode"
- Some consistency properties
  - If \{(m, n)\} \neq \emptyset, m, n \in N, then I INPUT(somenode) I \neq 0
  - If \{(n, m)\} \neq \emptyset, m, n \in N, then I OUTPUT(somenode) I \neq 0
- Maybe some others(?)
  - If \{(m, n)\} = k, m, n \in N, then I INPUT(somenode) I = k
  - If \{(n, m)\} = k, m, n \in N, then I OUTPUT(somenode) I = k

Data Flow Diagram

Elaboration of “Select New Floor” Suggests Artifact Issues

Data needs precise specification too
- DFD’s focus on functionality, using data as a vehicle
- Data shown as unstructured atomic units—usually unrealistic
- Complex functions cannot be adequately defined without delving into the details of how they handle structured data
- Sub-DFD’s can show how the high level data that high level DFD’s deal with is decomposed
  - But this is implicit data definition
  - Can be hard to read/inconsistent
- Data specification is worth doing explicitly, carefully
- Usually using Disciplined Natural Language—eg. Templates
- Formal specification of data is an important future topic

IDEF0

- Commercial DFD formalism
- Some formality and rigor behind it
- Primarily pictorial
- In wide use
- Additional semantics in IDEF1, IDEF2, etc.
Note: The edges here do not comprise a set. They comprise a collection.

Broadening DFG Semantics

- Node cannot begin until data arrives along all in-edges
  - DFGNODE is set of nodes, \( \langle \text{dfgnode} \rangle \) such that \( \text{dfgnode} \) is a function defined on the set of all artifacts \( IAs = \{ia\} \), such that \( ia \) is an annotation on an edge (dfgnode, dfgone)
  - How to adapt this for
    - any semantics?
    - for output semantics?
    - exactly one Output node
    - Etc.

These sorts of constraints can support additional types of reasoning:

Eg. about parallelism
More Broadening

- Use of "open boxes" to indicate data store
  - A different set, with different semantics
  - Not a computation function
  - Methods are: put, get, search(?)

Still More Broadening

- Different shapes of boxes
- Different pictures instead of boxes
- Different colors
- Different lines
  - .......

Central questions: What are the semantics? Does this really help? Or confuse?

Kepler--Another DFG Technology

- Data Flow diagram notation
- Has hierarchical decomposition
- Capability for specifying DFD semantics
  - For each diagram
    - Can be different at each level of hierarchy (!)
- Based on Ptolemy II system

Kepler Elaboration

What kinds of questions are well addressed by DFGs?

- Overall structure of functional capabilities
  - What does this piece do?
- System outputs and inputs
- How might changes be made?
- What functions create what data entities
Given that Precision is essential

- What about the other three dimensions?
  - Detail
    - Gained from hierarchical elaboration
  - Breadth
    - Comes from different (sub)types of DFG
  - Clarity
    - Seems to be reduced by increased Detail and Breadth
    - With the need for Precision

Clarity helps appeal to diverse stakeholders

- Users
- Developers
- Managers
- Evolvers
- Bystanders (?)

Final observations

- Very primitive representation
  - Very limited semantics
- But actually more a family of model types
  - Different sets of semantics
- The actual relation(s) are rarely made clear and precise
- Powerful aid to intuition and efficiency of communication
  - Clear advantages over natural language
- But is intuition misled by ambiguity, misinterpretation?
- Does not help explain HOW things get done

Control Flow Diagrams

- Semantics are different from DFD’s semantics
  - Arrows represent flow of control, rather than flow of data
  - Different shapes/types of boxes with different semantics
- Basic control flow graph:
  - Boxes represent functions
  - Some other shape represents control flow alternation
  - Arrows represent control flow:
    - If there is an arrow from circle A to circle B, it means ImmFol: “the execution of B can immediately follow the execution of A for some execution”
    - Different semantics for arrows between different shapes
  - Still other shapes represent Start and Stop
  - Use of hierarchy for elaborating boxes
- Usual enhancements:
  - Annotate edges with predicates
  - Special symbols for branching, concurrency control....

Example Control Flow Graph Semantics

- Square boxes: Functions
  - Any number of inedges
  - One outedge
- Edges: ImmFol relation
  - Function at head executes immediately after function at tail concludes
- Round ovals: Branches/decisions
  - Each oval represents a Boolean function
  - For each oval there are exactly two outedges, labeled True and False

Control flow graphs also address questions like "what does this do" and "how does this do it"
Example Control Flow Graph Semantics

- FG = (FN, FE), where FN = OPS ∪ TEST and
  OPS = {ops_i} nodes where computation is done
  TEST = {test_i} nodes where tests are done
- Edges: ImmFol relation
  if (fe_i, fe_k) ∈ FE then there is an execution path for which
  the execution of fe_k immediately follows the execution of fe_i
- | {(fe_i, fe_k)} | where fe_i ∈ TEST = 2
- | {(fe_i, fe_k)} | where fe_i ∈ OPS = 1
- FG is a directed, connected graph
- Maybe(?): FN = OPS ∪ TEST ∪ INPUT ∪ OUTPUT where
  fn_i ∈ OUTPUT ⇒ ∄ (fn_i, fn_k) ∈ FE
  fn_i ∈ INPUT ⇒ ∄ (fn_i, fn_k) ∈ FE
  Other semantics?

More Detail through Hierarchical Elaboration

- The challenges and solutions here are similar to those for DFGs
- Need for consistency between levels
- Semantics can be hard to define, stick to

Concurrency graph with “fork and join”

More Breadth from more (sub)types of CFG

- Concurrency
- Timing
- Data flow augmentations
- Etc.

Rendezvous Graph

- Rendezvous graph: RG = (N, E), where N = COMP ∪ SYNC
- S ∈ SYNC ⇒ S executes only after the execution of all C ∈ COMP such that (C, S) ∈ E
- If (S, C) ∈ E, where S ∈ SYNC, then C cannot execute until S has executed
Trace Flow Graph

- Explicitly represents interleaved execution
- The green edges are the ImmFol relation
- Double-headed arrows represent one arrow in each direction
- Sync nodes are filled in black
- Sync edges are blue

Different variations to address different stakeholders and their needs

- More semantic issues
  - Different graph types for different issues
- More detail
  - Hierarchy helps here
- More clarity
  - Does more elaboration and more graph types help clarity or impede it?

Focus on Clarity

- CFG has more procedural detail
- Does this help?
- What about combining DFG and CFG?
  - Iconography must be clear (?)

Control Flow Graph with Data Flow

Control Flow Graph

Control Flow Graph with Data Flow
What is Control Flow good/not good for?

- Sense of what algorithms to use
- Constraints on data appearing
  - eg. Assuming that the Request List is maintained in sorted order
- Can estimate running speed
- Can reason about functionality
  - Possible to strand requesters
  - Possible to take riders in the wrong direction
- Drawbacks:
  - What about safety?
  - What about data?

What Stakeholders does this address?

- Developers
- ???

Control Flow vs. Data Flow Graphs

- Both shed light on similar questions
- One focuses on data evolution, the other on functional development
- Both are useful, neither removes the need for the other
- Control flow graphs map closely to implementation code written in procedural languages.
  - Good basis for determining consistency of code with ideas expressed as data flow
- Data flow graphs focus more on the product itself, seem better at helping understand if and how it gets evolved
  - Seem better adapted to studying earlier formulations of the problem to be solved, and ways of solving it

Finite State Machines (FSM's)

- FSM = (Q, I, δ), where
  - Q = {q}, the set of all possible system states
  - I = {i}, the set of all events that can affect the state of the system
  - δ is a function, δ: Q x I → Q such that if the system is in state q and event i occurs, then the system transitions to state δ(q, i)
- Use of hierarchy can help add detail
  - But presents familiar consistency problems
- Other definitions:
  - Start state
  - Accepting state
  - Trap state

Finite State Machines (FSM's)

- Primary appeal is visualizability—clarity
  - Circles represent states
  - (Curved) arrows represent transitions
  - Arrows are annotated with inputs
- Intuitively: Can "watch" a stream of inputs "drive" the behavior of the system as a sequence of movements from state to state

Kinds of Questions FSMs seem adept at helping answer:
- "What is a good way to think about the problem to be solved?"
- "What is the solution approach?"
- "How does this program work?"
Finite State Machine

Doors Closed, Elevator Idle
Request to Move to new floor
Doors closed, Elevator in Motion
Arrival at New Floor
Getting new floor request
Queue empty
Queue open, Elevator Stopped
Queue empty

Request to Move to new floor
Doors closed, Elevator in Motion
Request to Move to new floor
Arrival at New Floor
Getting new floor request
Queue empty
Queue open, Elevator Stopped
Queue empty

More Finite State Machine Details

Time display mode
Press button A
Alarm display mode
Press button B
Datebook mode
Phone book mode

More Complex FSMs

• State also specifies activities
  – Leans towards a CFG
• Transition may involve computation
  – Considered to be “instantaneous”
• Transition may be conditional
  – Event as well as a condition
• Transition may emit events
  – To drive other FSMs
• FSMs may be hierarchical
  – What exactly are semantics?

Statecharts: More Complex FSMs

Activity to be performed in this state

Statecharts: More Complex FSMs

- Initialize: Initialize course
- Assign professor to course
- Register a student
- Report course is full
- Send cancellation notices
- Add student / numStudents = 0
- Cancel course
- Registration complete
- Generate class roster
- Registration closed
- numStudents = 10
- numStudents ≥ 3
- numStudents < 3
- Unassigned

What is FSM good/not good for?

- Focus on specific issue: safety concern
- Model unsafe state
- Model state transitions
- Can unsafe state be reached?
- Drawbacks:
  - No sense of functionality
  - Unless additional semantics
  - No sense of how functionality achieved
  - Except perhaps hierarchy
  - Hard to deal with concurrency
  - Without additional semantics
  - Impossible to reason about timing

Petri Nets

- More powerful and intuitive depiction of control flow
- Strong on depiction of parallelism and concurrency
- A Petri Net structurally consists of
  - A finite number of places
  - A finite number of transitions
  - A finite set of arrows that connect places to transitions (or vice versa)
- If an arrow goes from a place to a transition, then place is said to be an input place of the transition.
- If an arrow goes from a transition to a place, then place is said to be an output place of the transition.

Marking and Firing Petri Nets

- A Petri Net place can be marked by the presence of a token
  - Any collection of places can be marked.
  - Any such marking is said to define a state of the Petri Net
- Petri Nets proceed from one state to another by means of a firing
  - Occurs only when every input place of a transition is marked with a token.
- The effect of the firing of a transition is to
  - Remove all of the tokens from the transition’s input places
  - Put tokens in all of the transition’s output places
Unmarked Petri Net

Petri Net Marked
With Up button pressed

Petri Net Marked
With Up button pressed

Petri Net Marked
With Up button pressed

Unmarked Petri Net--Two Elevator Cars

Petri Net--Two Elevator Cars
Car 1 Wins Race
Petri Net--Two Elevator Cars
Car 2 Wins Race

Car 1 approaching Floor n, going up
Up Button Pressed at Floor n
No Press
Stop at Floor
Go to Next Floor
Up

Car 2 approaching Floor n, going up
Up Button Pressed at Floor n
No Press
Stop at Floor
Go to Next Floor
Up

Car 2 approaching Floor n, going up
Up Button Pressed at Floor n
No Press
Stop at Floor
Go to Next Floor
Up

Car 1 approaching Floor n, going up
Up Button Pressed at Floor n
No Press
Stop at Floor
Go to Next Floor
Up

Up Button
Pressed at Floor n
Car 1 approaching
Floor n, going up
No Press
Stop at Floor
Go to Next Floor
Up

Car 2 approaching Floor n, going up
Up Button Pressed at Floor n
No Press
Stop at Floor
Go to Next Floor
Up

Marking for moving up to pick up a passenger
Move up to Nearest Up Button
Car going up
Up button higher
No Up Button Higher
Move to nearest Down Button
Visit floor, Turn off button
Down button lower
No Down Button Lower
Visit floor, Turn off button

Marking for moving up to pick up a passenger
Move up to Nearest Up Button
Car going up
Up button higher
No Up Button Higher
Move to nearest Down Button
Visit floor, Turn off button
Down button lower
No Down Button Lower
Visit floor, Turn off button

Marking for moving up to pick up a passenger

Car going up
Up button higher
Move up to Nearest Up Button
Visit floor, Turn off button
Move to nearest Down Button
Visit floor, Turn off button
Down button lower
No Down Button Lower
No Up Button Higher
Move up to Nearest Up Button
Car going down

Queues at places?

Marking when no passengers higher, but passenger lower

Car going up
Up button higher
Move up to Nearest Up Button
Visit floor, Turn off button
Move to nearest Down Button
Visit floor, Turn off button
Down button lower
No Down Button Lower
No Up Button Higher
Move up to Nearest Up Button
Car going down

Marking when no passengers higher, but passenger lower

Car going up
Up button higher
Move up to Nearest Up Button
Visit floor, Turn off button
Move to nearest Down Button
Visit floor, Turn off button
Down button lower
No Down Button Lower
No Up Button Higher
Move up to Nearest Up Button
Car going down
Petri Nets

• More powerful and intuitive depiction of control flow
  strong on depiction of parallelism and concurrency

• A Petri Net structurally consists of
  – A finite number of places
  – A finite number of transitions
  – A finite set of arrows that connect places to transitions
    (or vice versa)

• If an arrow goes from a place to a transition, then place is
  said to be an input place of the transition.

• If an arrow goes from a transition to a place, then place is
  said to be an output place of the transition.

Marking and Firing Petri Nets

• A Petri Net place can be marked by the presence of a token
  – Any collection of places can be marked.
  – Any such marking is said to define a state of the Petri Net

• Petri Nets proceed from one state to another by means of a firing
  – Occurs only when every input place of a transition is
    marked with a token.

  • The effect of the firing of a transition is to
    – Remove all of the tokens from the transition’s input places
    – Put tokens in all of the transition’s output places

Some Semantics

• PN = (Places, Transitions, Edges)
  – Places (PN) = \{place\}_i
  – Transitions (PN) = \{transition\}_j
  – Edges (PN) = \{inedges U outedges\}
    \{\text{inedges} (PN) = \{(place, transition)\}_i \times \text{outedges} (PN) = \{(transition, place)\}_j\}
  – Marked: PN X Places \rightarrow \{True, False\}
  – If Marked (PN, place) = True we say that place is marked
  – A transition, \text{transition} \in Transitions (PN) can fire if
    for all of its inedges, \text{Marked (place)} = True
    – After a transition fires, \text{Marked (place)} = True

Evaluation of Petri Nets

What are Petri Net representations good for?:

• How do things get done?
• Especially parallelism and nondeterminism
• Helps spot races and deadlocks

What kinds of things are Petri Nets not good for?

• Too little focus on product (tokens represent it)

• Many extensions:
  – Hierarchical Petri Nets
  – Colored tokens
  – “Or” transitions
  – Queues at places

• Too many extensions confuse the picture?