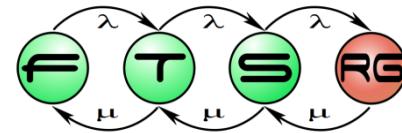


# Concrete Syntax Design for Domain-specific Languages

Zoltán Ujhelyi  
Model Driven Software Development  
Lecture 7



# Structure of DSMs

Graphical syntax



Abstract syntax



Well-formedness constraints



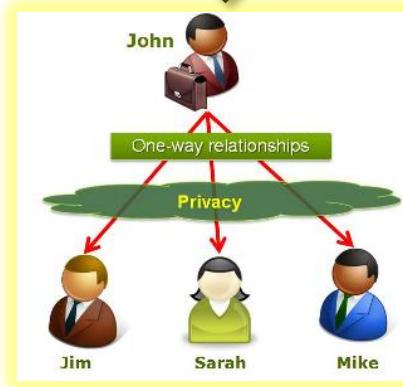
Behavioural semantics, simulation



Code generation

Textual syntax

```
test.socialnetwork
SocialNetwork {
    Person Ujhelyi {
        male
        memberships BME, VVEC
    }
    Person Horvath {
        male
        memberships FTSRG
    }
    Community BME {
        Community FTSRG {
            Community test
        }
    }
}
```



View

Code documentation, configuration

```
</membership>
<profile defaultProvider="Sitefinity">
    <providers>
        <clear/>
        <add name="Sitefinity" connectionS
    </providers>
    <properties>
        <add name="FirstName"/>
        <add name="LastName"/>
        <!-- SNP specific properties -->
        <add name="NickName" />
        <add name="Gender" />
    </properties>

```

Code documentation, configuration

# DSM aspects



# Concrete Syntax Design

- User-managed parts of a modeling language
  - Performance
  - Robustness
  - Usability issues
- Creating model editors
  - Similar problems at programming languages
  - IDE extensions needed

# Approaches

- Textual syntax
  - Character-based edit operations
  - Abstract syntax: traditional AST
- Graphical syntax
  - Editing operations: translated to abstract syntax
  - Abstract syntax: based on metamodel
- Form-based entry
  - Less common
  - Behaves similar to graphical syntax

# Advanced features

# Advanced features

## High level editing support

- Outline view
- Documentation display (e.g. Javadoc)
- Templates/snippets/examples
- Content assist
- Validation, automatic fixes

## Project-level integration

- Code generation
- Wizards to create projects/files
- Integration with manually written code in GPL

# Advanced features

## High level editing support

- Outline view
- Documentation display (e.g. Javadoc)
- Templates/snippets/examples
- Content assist
- Validation, automatic fixes

## Project-level integration

- Code generation
- Wizards to create projects/files
- Integration with manually written code in GPL

# Technology

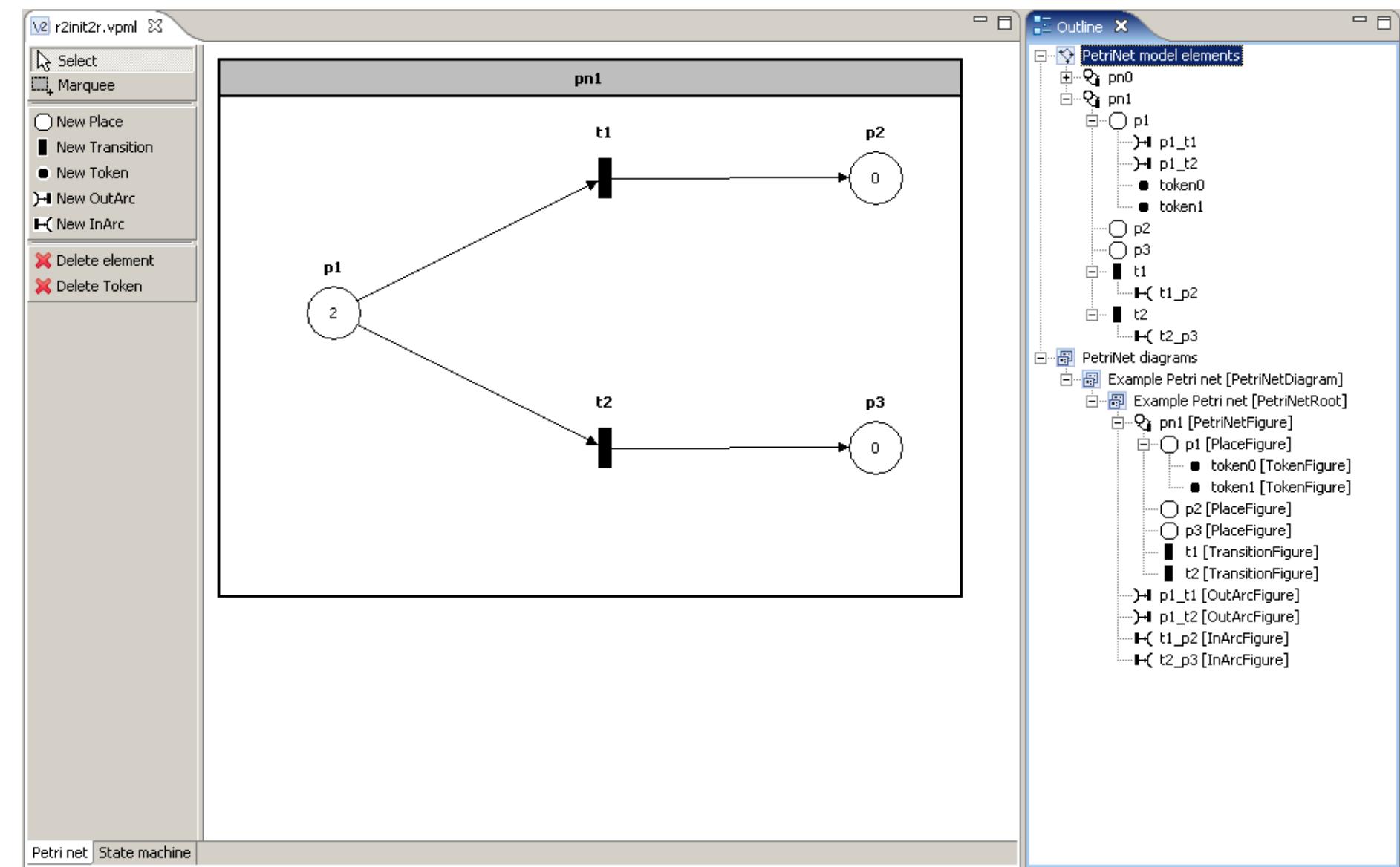
- Eclipse Modeling Tools
  - Several related subprojects
  - Each supports a single aspect
  - Examples of today
- Microsoft Visual Studio 2010 Visualization & Modeling SDK
  - DSL modeling framework from Microsoft
  - Own metamodeling core
  - Focuses on graphical modeling

# Graphical Editors

# Graphical Modeling

- Model
  - Typically graph-based modeling
  - In our case EMF
- Idea
  - Display and editing as a graph model

# Example: Petri net editor



# Example: Petri net editor

The screenshot shows a Petri net editor interface with two main panes. The left pane displays a Petri net diagram titled 'pn1'. It contains three places: p1 (with 2 tokens), p2 (with 0 tokens), and p3 (with 0 tokens). There are two transitions: t1 and t2. Arcs connect p1 to t1, t1 to p2, p1 to t2, and t2 to p3. A red callout bubble points from the text 'Tree-based outline view' to the right pane. The right pane is a 'PetriNet model elements' tree view. The root node is 'pn1', which has children 'p1', 'p2', 'p3', 't1', and 't2'. 'p1' has children 'p1\_t1' and 'p1\_t2', each with tokens 'token0' and 'token1'. 't1' has children 't1\_p2' and 't2', which has child 't2\_p3'. The right pane also lists 'PetriNet diagrams' under 'Example Petri net [PetriNetDiagram]' and 'PetriNetRoot'.

Tree-based outline view

```
graph TD; pn1[pn1] --- p1((p1)); pn1 --- t1[ ]; pn1 --- t2[ ]; p1 --- t1; t1 --- p2((p2)); p1 --- t2; t2 --- p3((p3));
```

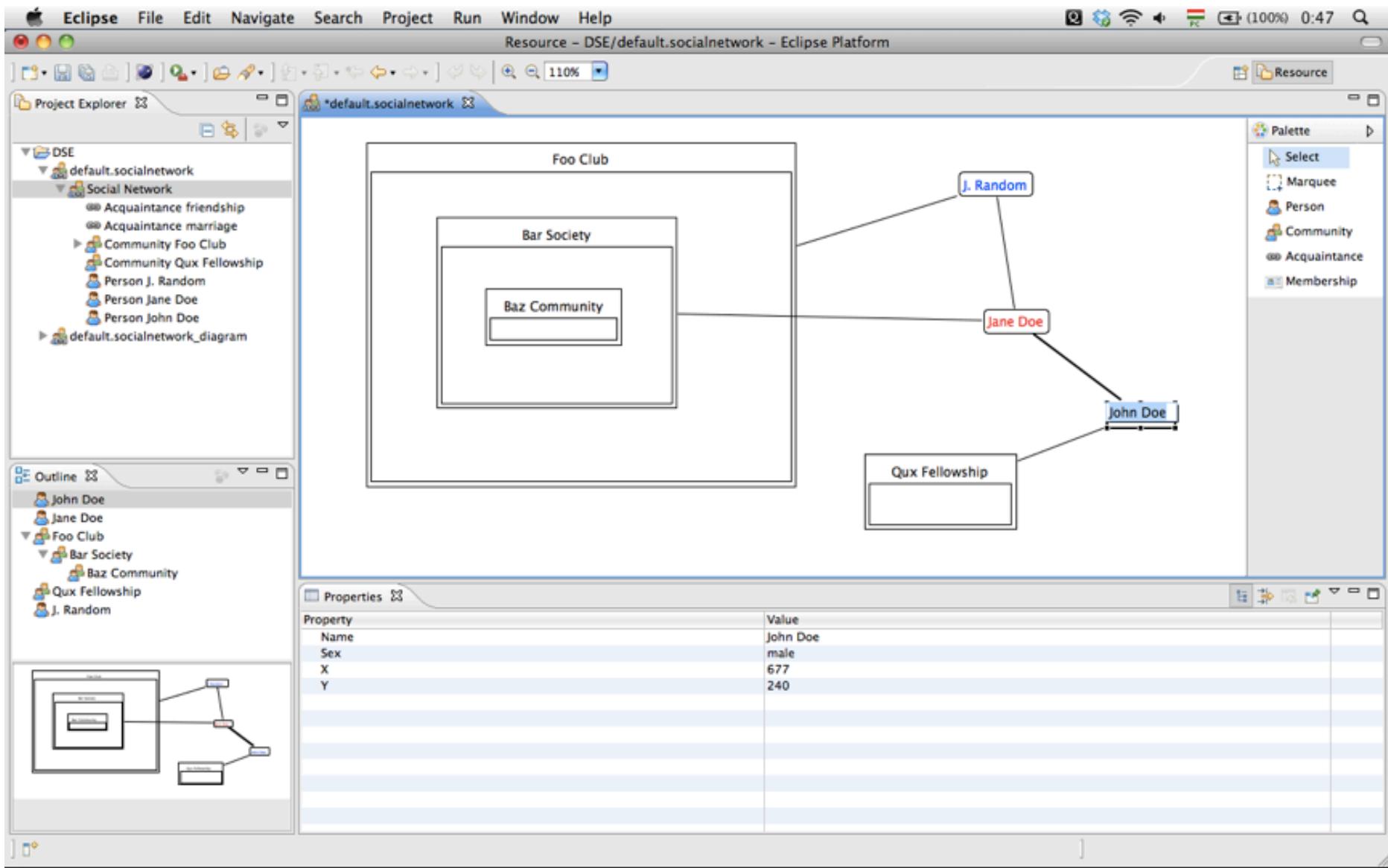
PetriNet model elements

- + PetriNet model elements
- pn0
- pn1
  - p1
    - p1\_t1
    - p1\_t2
    - token0
    - token1
  - p2
  - p3
  - t1
    - t1\_p2
  - t2
    - t2\_p3

PetriNet diagrams

- Example Petri net [PetriNetDiagram]
- Example Petri net [PetriNetRoot]
  - pn1 [PetriNetFigure]
    - p1 [PlaceFigure]
      - token0 [TokenFigure]
      - token1 [TokenFigure]
    - p2 [PlaceFigure]
    - p3 [PlaceFigure]
    - t1 [TransitionFigure]
    - t2 [TransitionFigure]
  - p1\_t1 [OutArcFigure]
  - p1\_t2 [OutArcFigure]
  - t1\_p2 [InArcFigure]
  - t2\_p3 [InArcFigure]

# Example: Social Network editor



# Example: Social Network editor

Project Explorer extensions

Eclipse File Edit Navigate Search Project Run Window Help

Resources DSE/default.socialnetwork - Eclipse Platform

Project Explorer

DSE

- default.socialnetwork
- Social Network
  - Acquaintance friendship
  - Acquaintance marriage
  - Community Foo Club
  - Community Qux Fellowship
  - Person J. Random
  - Person Jane Doe
  - Person John Doe
- default.socialnetwork\_diagram

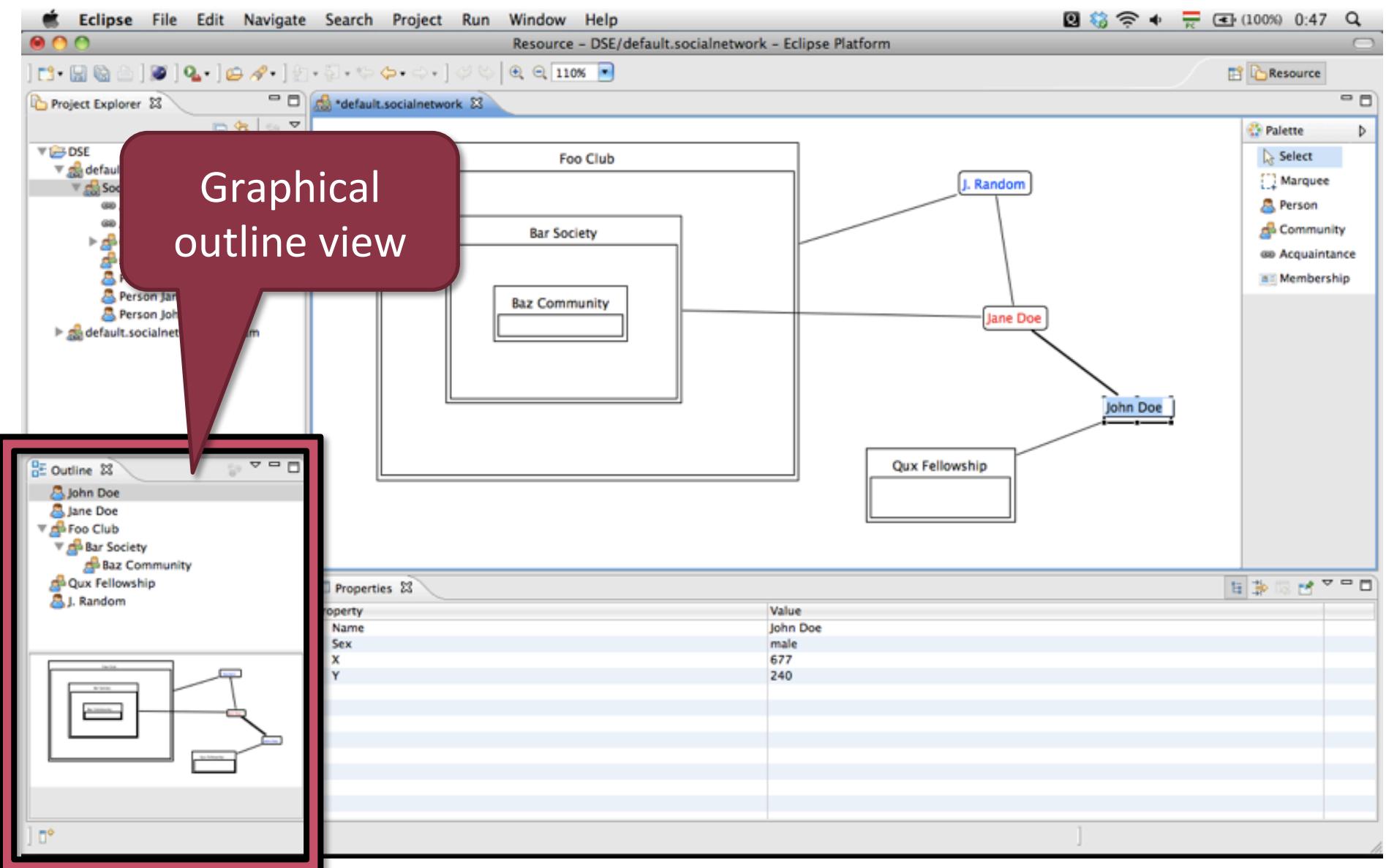
Outline

- John Doe
- Jane Doe
- Foo Club
  - Bar Society
    - Baz Community
  - Qux Fellowship
  - J. Random

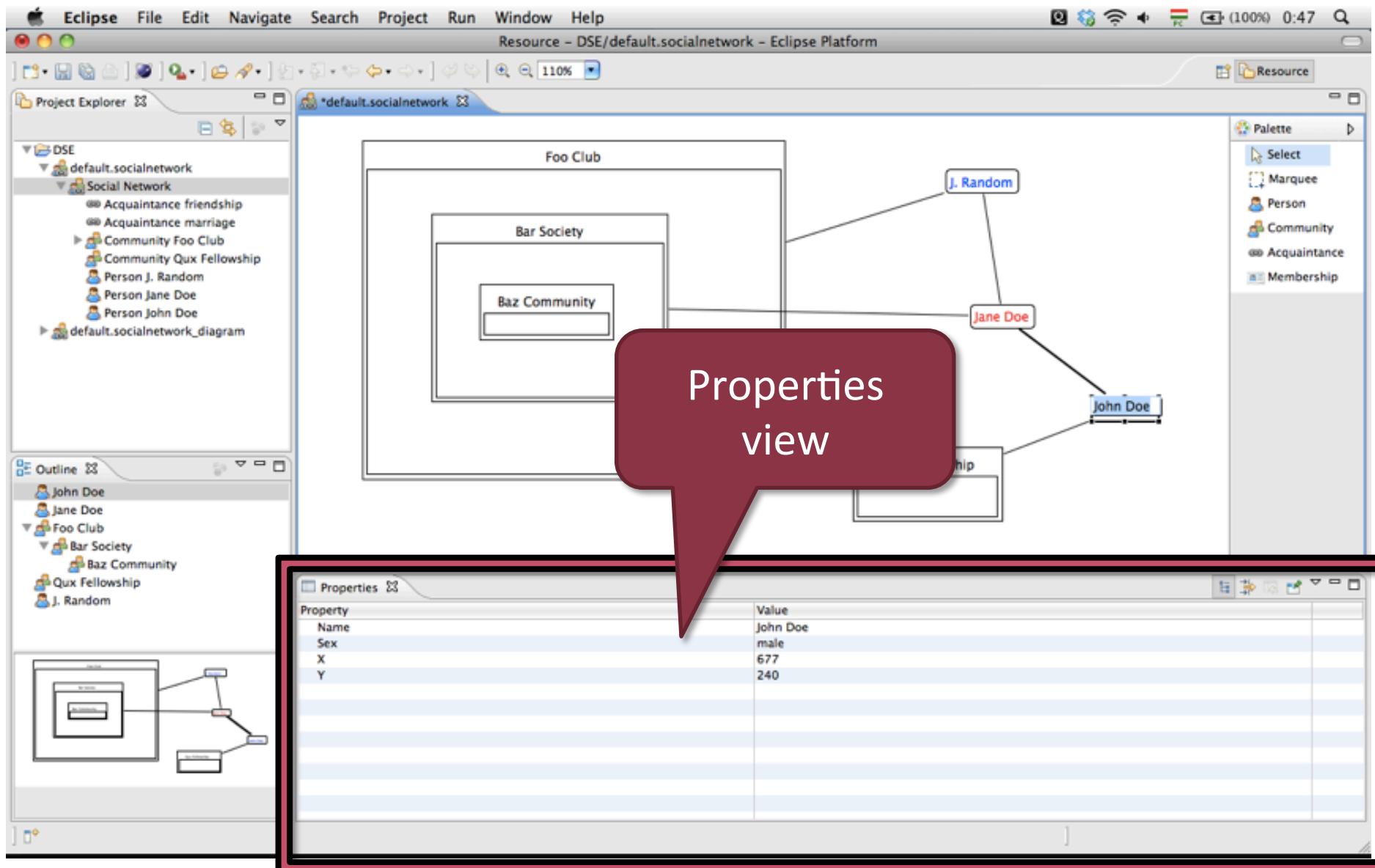
Properties

Property	Value
Name	John Doe
Sex	male
X	677
Y	240

# Example: Social Network editor



# Example: Social Network editor



# Implementation

- Presentation
  - Based on a Canvas
  - Using vector-graphic libraries (GEF/Draw2d)
- Model manipulation
  - EMF Edit model manipulation commands
    - Atomic operations: create/modify/remove node/edge
  - Transactional modifications
    - Undo/redo support
    - Replayability

# Implementation 2.

- View models
  - Modeling for view-specific information
    - Coordinates
    - Size
    - Colors and fonts
    - ...
  - Generic implementation in GMF and Graphiti
  - Often stored in external files
    - Separation of concerns
    - E.g. code generator not interested in view information

# Technologies 1. - GEF

- Graphical Editing Framework (GEF)
  - “Low level” editor framework
  - Not EMF-specific
- Model-View-Controller approach
- Generic graph-based editor framework
  - Including undo/redo support
  - Graphical outlines
- Manual coding for every possible element

# Technologies 2. – GMF

- Graphical Modeling Framework
- Based on GEF and EMF
- Well-separated view and domain models
  - Generic view model
  - Synchronization provided by GMF framework
- Relatively old technology
  - Widely used
  - Very complex to start

# Technologies 2. – GMF

- Model-driven development environment
  - Common model for graphical editors, using
    - Figure definition model
      - Basic symbol definition of the graphical language
    - Tooling model
      - Defining model manipulation commands
    - Mapping model
      - Mapping figures and tools to domain model
  - Fully functional editor can be generated
    - Problematic manual modifications
- Or a high-level editor framework
  - Manual coding

# Technologies 3. - Graphiti

- Newer high level graphical editor framework
  - Based on EMF and GEF
  - But: different approach then GMF
    - Simplified programmatic API
    - Manual coding
  - Idea
    - All Graphiti based editors should
      - Look similar
      - Behave similar

# Technologies 3. - Graphiti

- Development methodology
  - Coding over a high-level Java framework
    - Much simpler than GMF
    - Repetitive code needed
- Spray project
  - Textual modeling environment for graphical editors
  - Generates code over the Graphiti framework

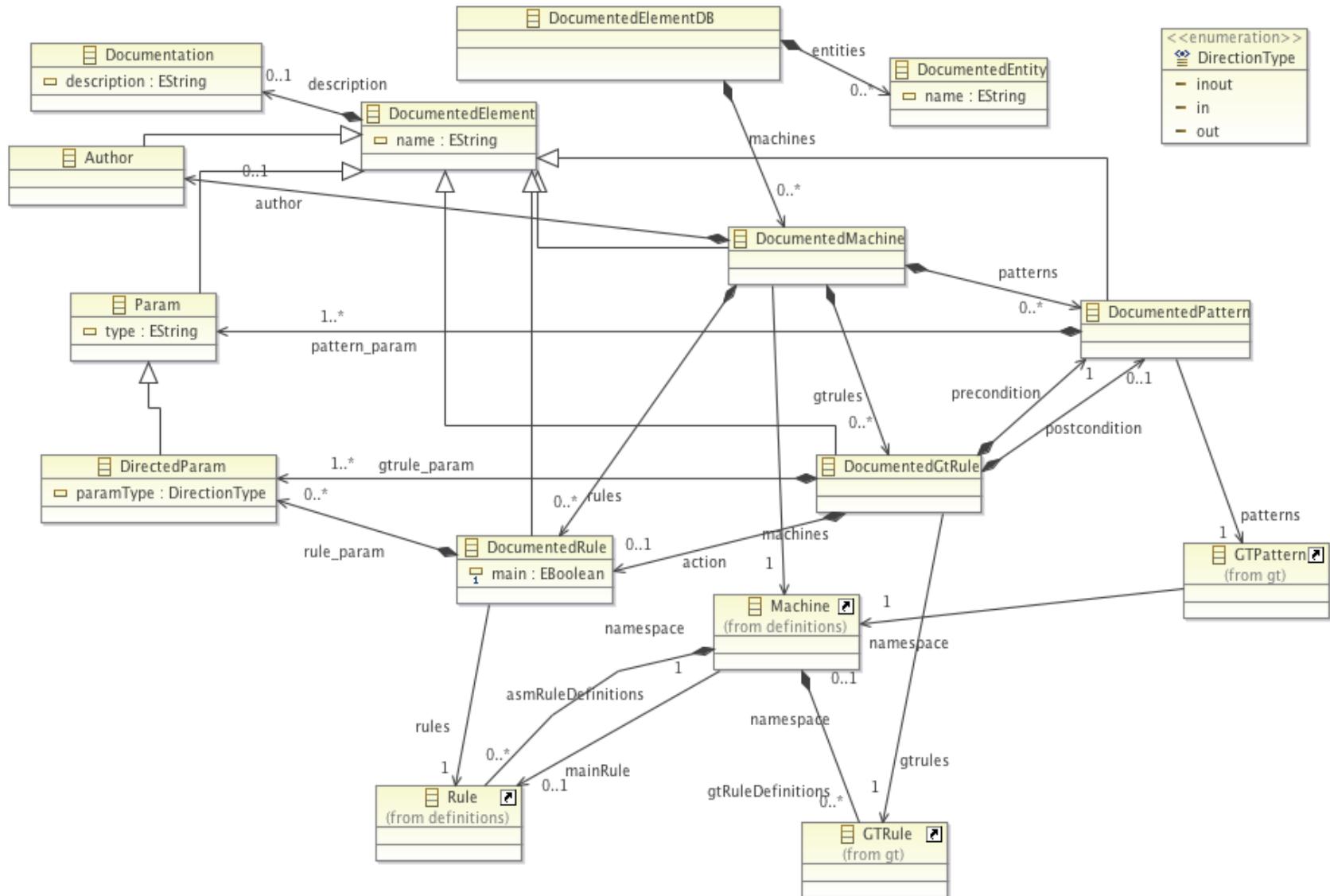
# Technology Comparison

	GEF	GMF	Graphiti
Model	Arbitrary	EMF	EMF
Non graph-based presentation	Manageable	Large amount of customization needed	Not supported
Code size	Large, repetitive code	Mostly modeling, some coding	Smaller amount, but repetitive code
Development workflow	Only coding	Modeling and coding	Coding

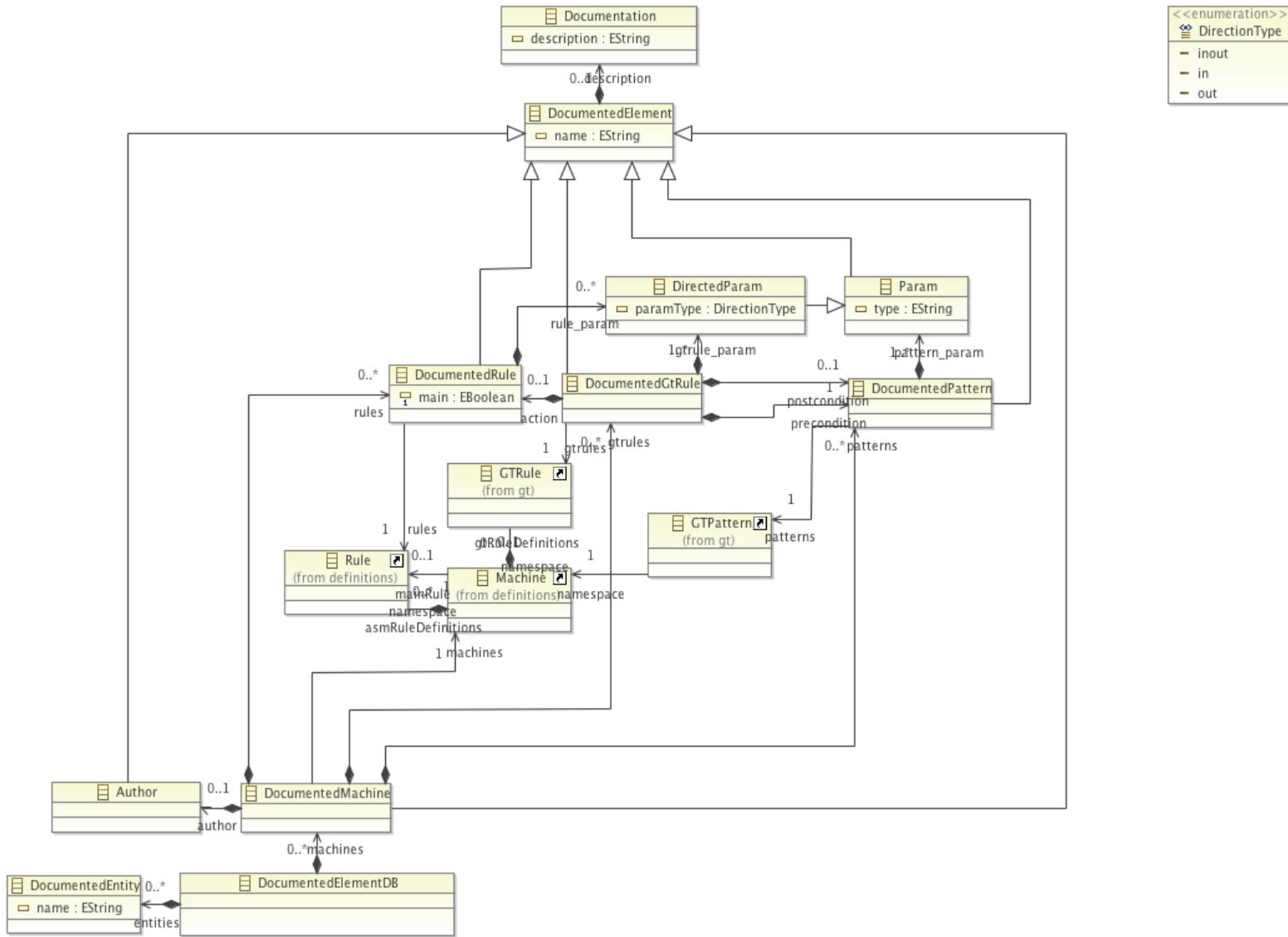
# Advanced issues

- Cumbersome editing
  - E.g., reorganization to instert a node to the middle
- Handling large models
  - 20+ nodes on a diagram:
    - Logical structure, readability possible
    - But needs human support
  - 100-1000+ nodes on a diagram
    - Technological limitations
    - Usability limitations

# Example: Layouting

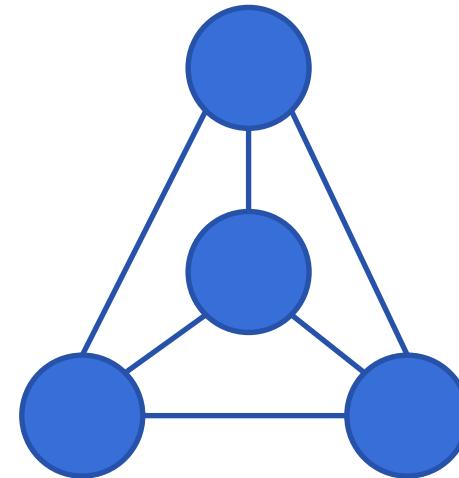
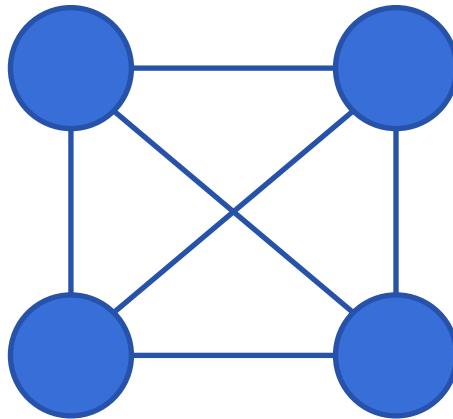


# Example: Layouting



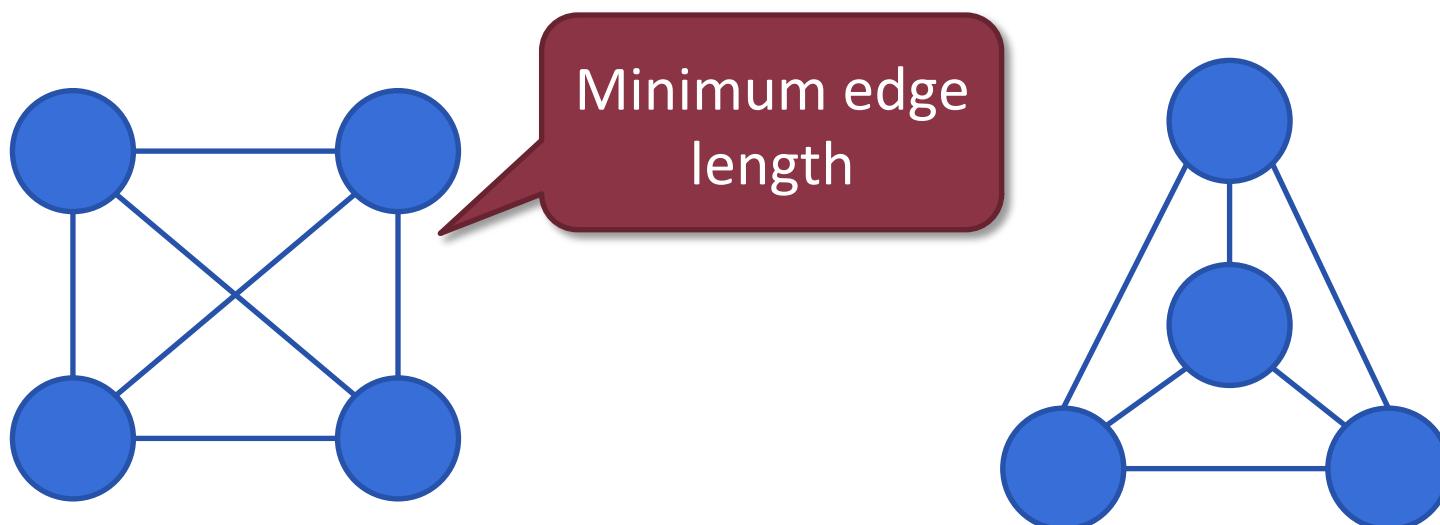
# Layouting Support for Graphical Editors

- Computation of the position of nodes
  - Possible to do automatically
  - For a given metamodel
    - No unified visual requirements possible
    - We have to decide what is important to show



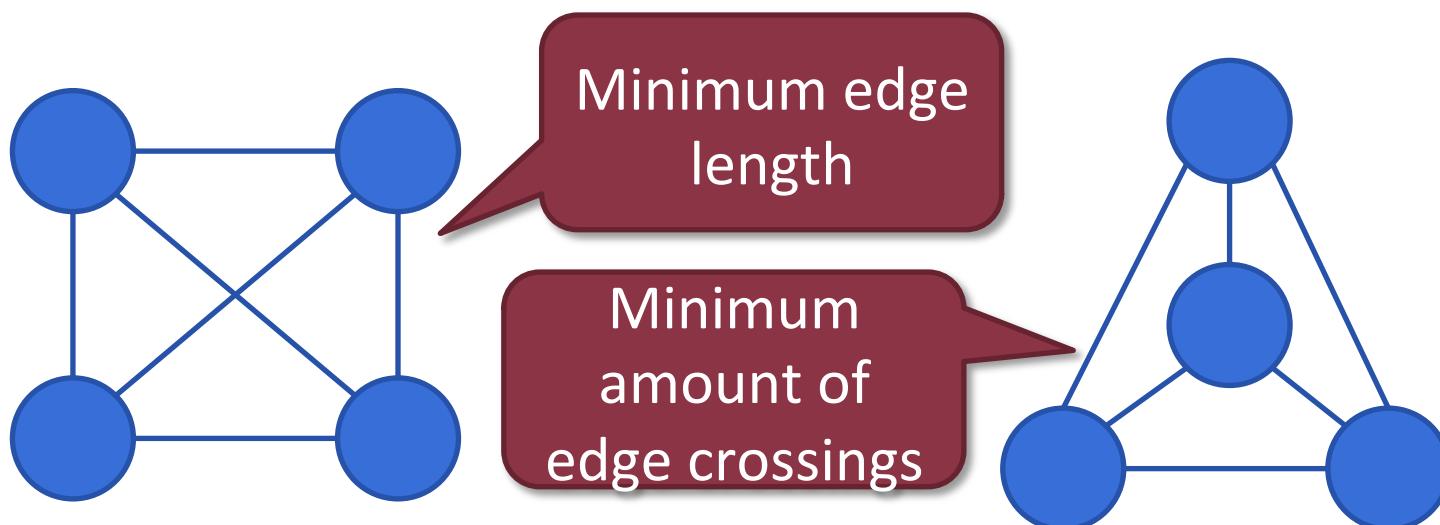
# Layouting Support for Graphical Editors

- Computation of the position of nodes
  - Possible to do automatically
  - For a given metamodel
    - No unified visual requirements possible
    - We have to decide what is important to show



# Layouting Support for Graphical Editors

- Computation of the position of nodes
  - Possible to do automatically
  - For a given metamodel
    - No unified visual requirements possible
    - We have to decide what is important to show



# Layouting Support for Graphical Editors

- **GraphViz** - <http://graphviz.org>
  - Layouting project with high quality layout algorithm
  - Hard to integrate into Eclipse applications
- **Zest** - <http://wiki.eclipse.org/index.php/Zest>
  - Graph widgets for SWT applications
  - Easily integratable into Eclipse applications
  - Not the best layout algorithms
- **KIELER** - <http://rtsys.informatik.uni-kiel.de/trac/kieler/>
  - Eclipse based tools
  - Built-in support for GMF layouting

# Textual modeling languages

# Textual Domain-specific Languages

- Idea
  - Describing models as text files
- Textual development
  - Long history (30+ years)
  - Well-researched theory
  - Mature tools

# Regular expressions

- Pattern matching for strings
  - Good support
    - Most programming languages
    - More or less the same syntax
  - Calculates and returns matches
- Usable as DSL parser?

# RegEx: Validation of email addresses

This regular expression will only validate addresses that have had any comments stripped and replaced with whitespace

# Additionally

- Output is a single boolean variable
  - Decides whether the string matches the language
- What is missing?

# Additionally

- Output is a single boolean variable
  - Decides whether the string matches the language
- What is missing?

# Error localization!

# Example: Grammar for describing name lists

- Terminal Symbols

- “*Dániel*”, “*István*”, “*Zoltán*”, “*and*”, “,”

- Non-terminal Symbols

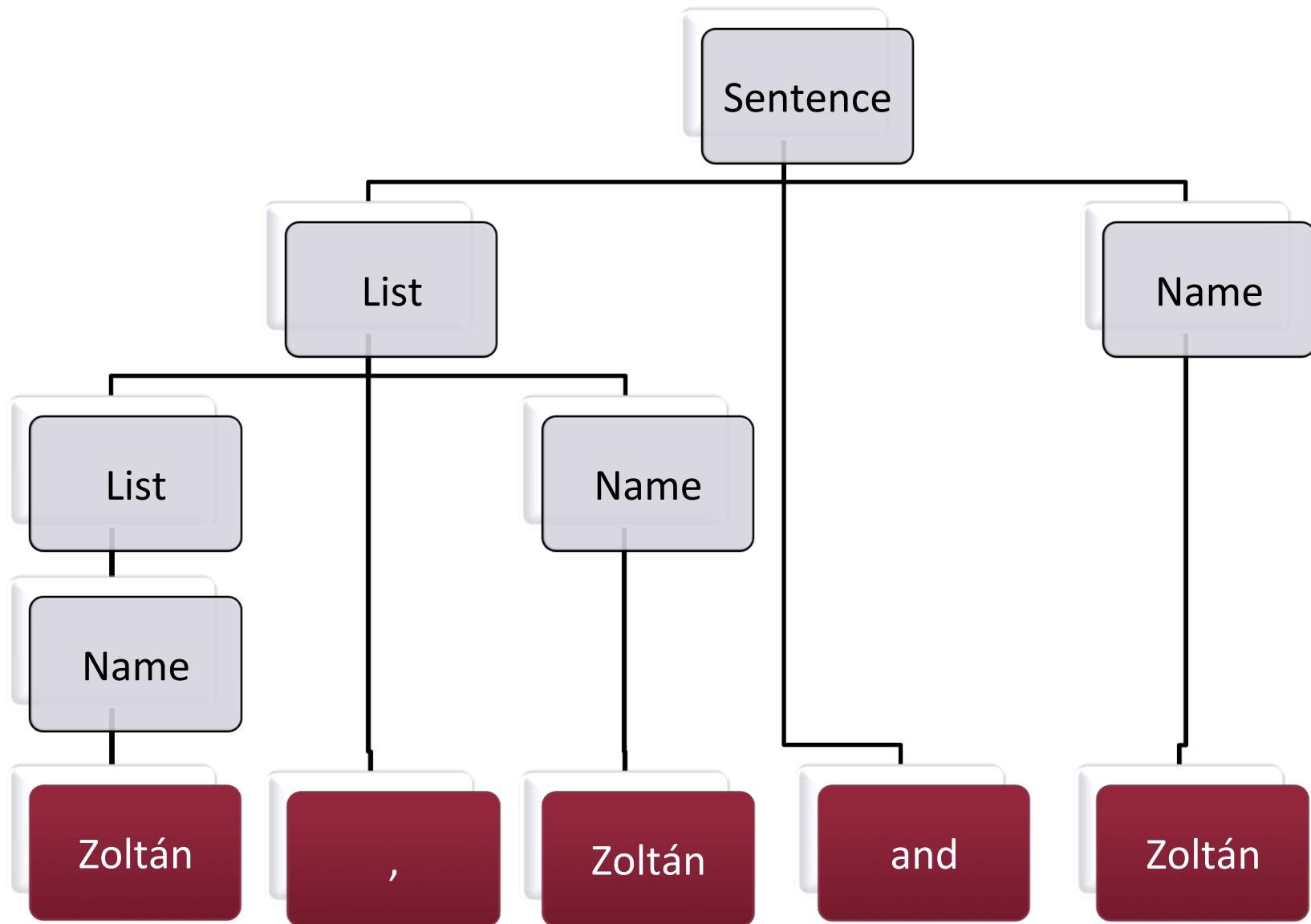
- «Name», «Sentence», «List»

«Name» ::= *Zoltán* | *István* | *Dániel*

«Sentence» ::= «Name» | «List» *and* «Name»

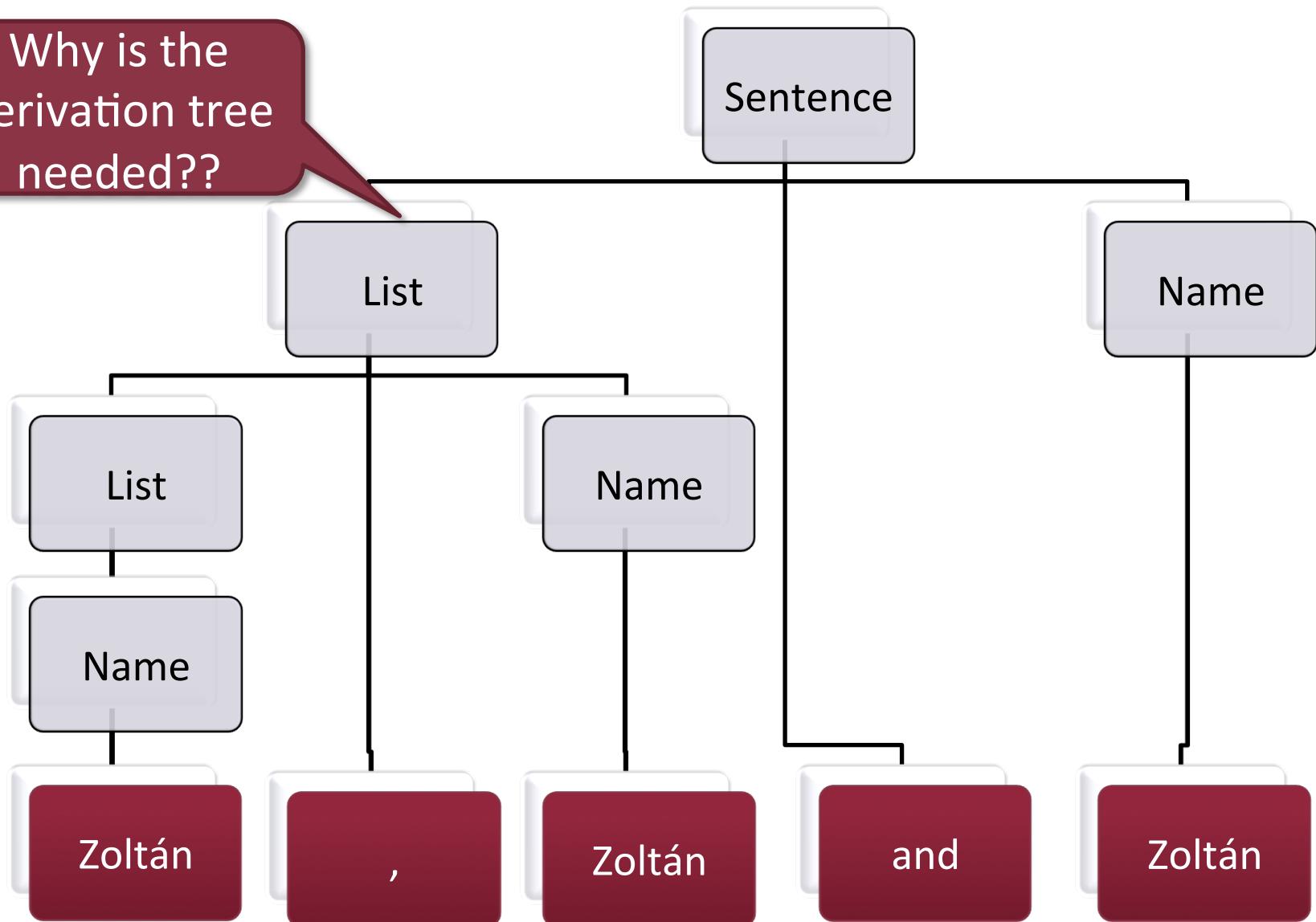
«List» ::= «List», «Name» | «Name»

# Derivation Tree

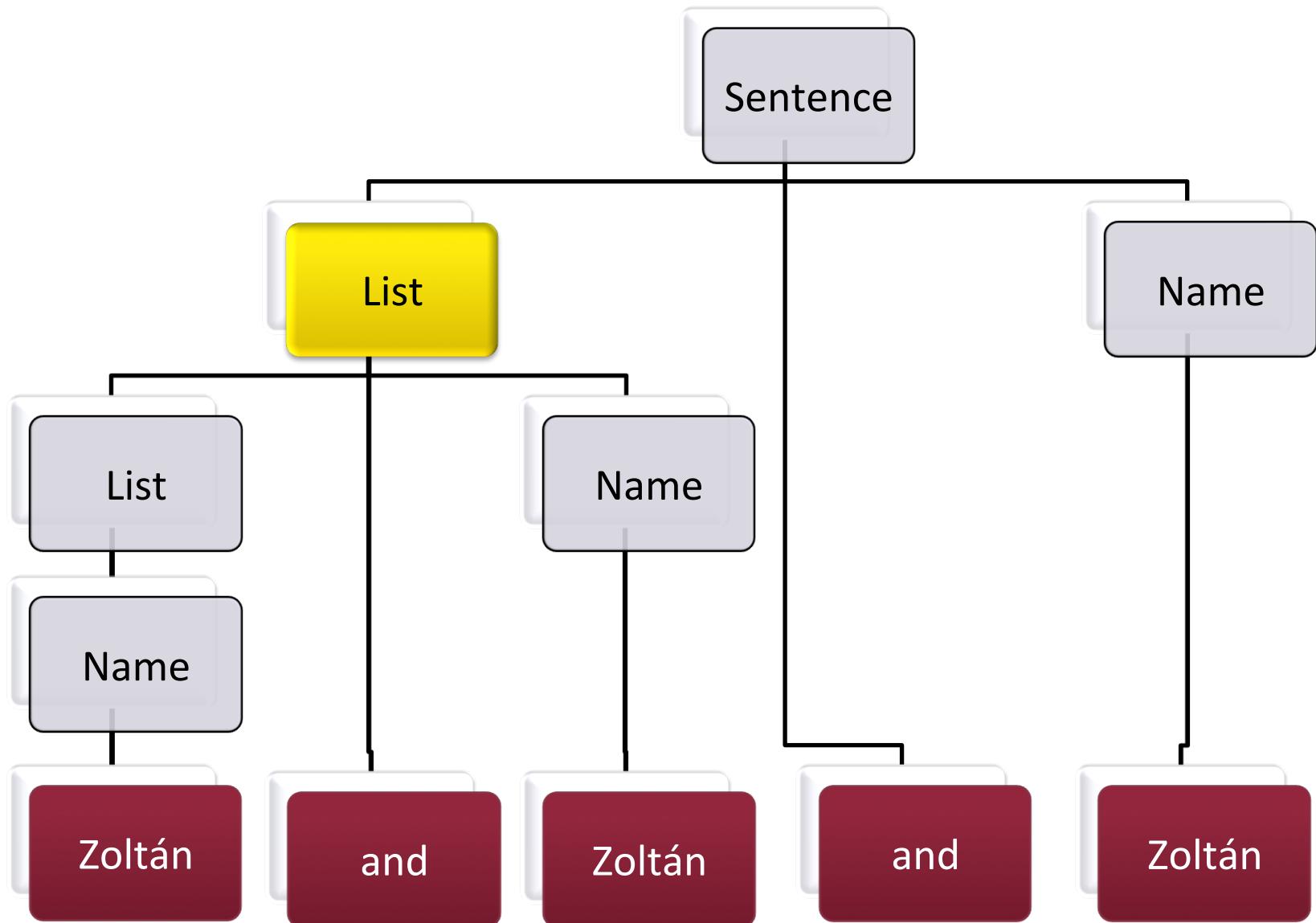


# Derivation Tree

Why is the derivation tree needed??



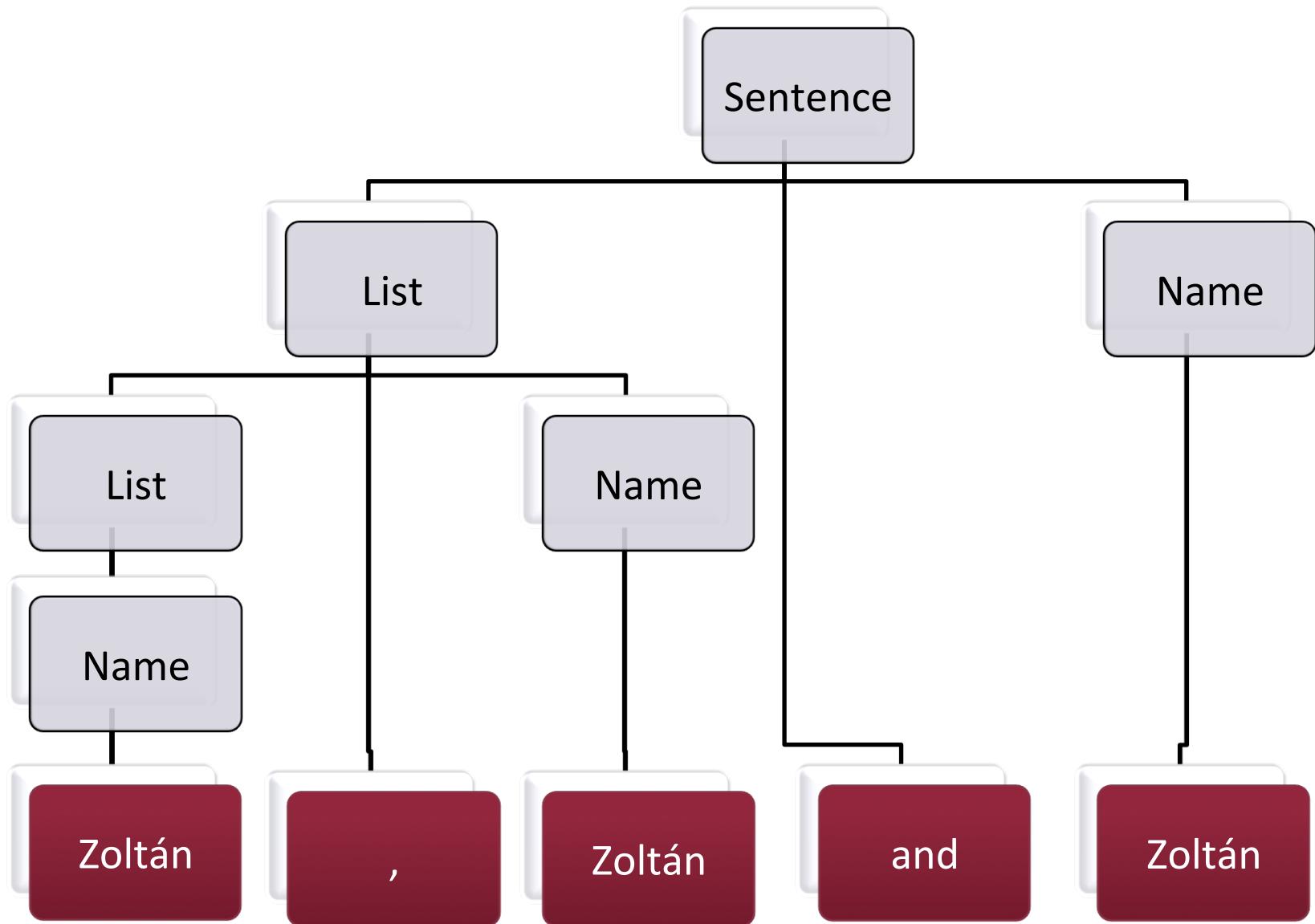
# Derivation Tree



# Additional practical tasks

- Handling input strings
- Variable handling
- High level analysis

# Input



# Input

Input string:

'Z' 'o' 'l' 't' 'á' 'n' ' ' ' ' ',' 'Z'  
'o' 'l' 't' 'á' 'n' ' ' ' 'a' 'n' 'd'  
' ' 'Z' 'o' 'l' 't' 'á' 'n'

Zoltán

,

Zoltán

and

Zoltán

# Input handling

Input:

Character stream

Parser  
input:

Higher level tokens

- «Name»
- ;
- “and”

Input gap

Filled by ‘lexer’

# Input handling

Input:

Character stream

Parser  
input:

- Higher level tokens
  - «Name»
  - ;
  - “and”

Input gap

Filled by ‘lexer’

- Why is this indirection useful?
  - Error handling
  - Performance
  - Problem decomposition

# Input handling

Input:

Character stream

Parser  
input:

Higher level tokens

- «Name»
- ;
- “and”

Input gap

Filled by ‘lexer’

- Why is this indirection useful?
  - Error handling
  - Performance
  - Problem decomposition

# Lexer

- Goal:
  - Tokenizing the input character stream
  - Similar to the parsing problem
    - But usually simpler – Typically regular expressions
    - Only word/token identification
    - Optional task: leaving out comments
  - Simplifies parsing significantly

# Variable Handling

```
a=3;
```

```
System.out.println(a);
```

## ■ Variables

- In runtime
  - Value calculation/substitution
- Editing/analysis time
  - References to other parts of the AST

# Variable Handling

- *Variable reference*
  - A use of an already defined variable
- *Variable definition*
  - A declaration of a variable
  - Unique naming
    - Variable definitions must be resolvable
    - Extra phase required after parsing

# Variable Handling

- Parser checks
  - “Can a variable named like ‘a’?”
- Reference resolution
  - “Is a variable defined?”
    - Scoping problem
  - “Is a variable uniquely defined?”

# Scoping Problem

```
private int value;
```

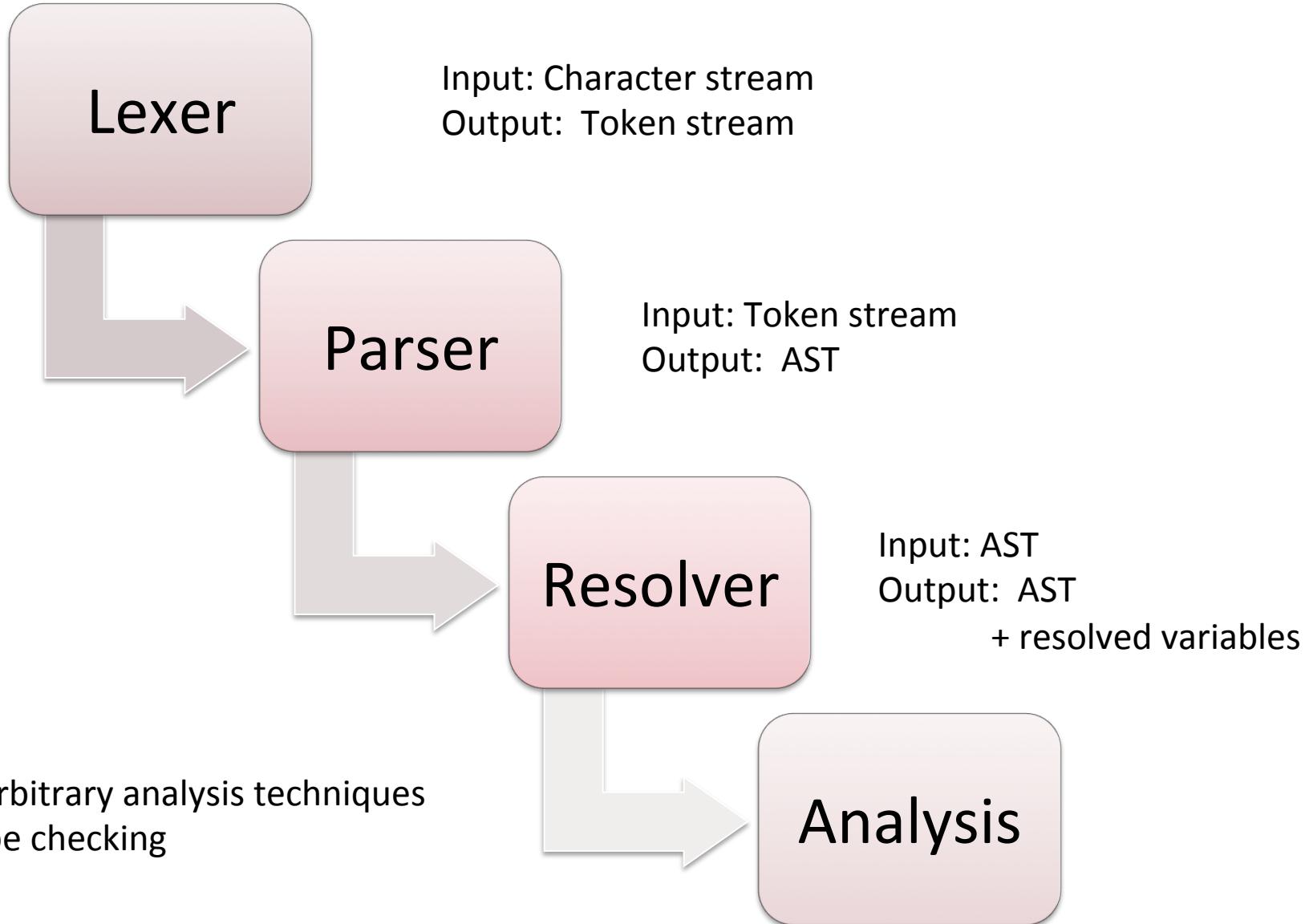
```
public void setValue(int value) {  
    this.value = value;  
}
```

Which variable  
declaration is referred  
by ‘value’?

# Scoping Problem

- Solution
  - Resolver should define this
- Possible approaches
  - Most specific declaration
  - Conflict is error
  - Qualified references
  - ...

# The Parsing Process



# Technologies 1. – IMP

- IDE Meta-tooling platform
  - Goal:
    - Language editor creation
  - Every parser-generator should be re-usable
  - Manual coding required
  - Project is close to dead

# Technologies 2. - EMFText

- Editor generation
  - Based on existing EMF model
  - Different generated grammar styles
    - Manually modifiable
    - Limited grammar support
- Syntax Zoo
  - ~100 different syntax examples available
  - Including a Java implementation (called JaMoPP)

# Technologies 3. – Xtext

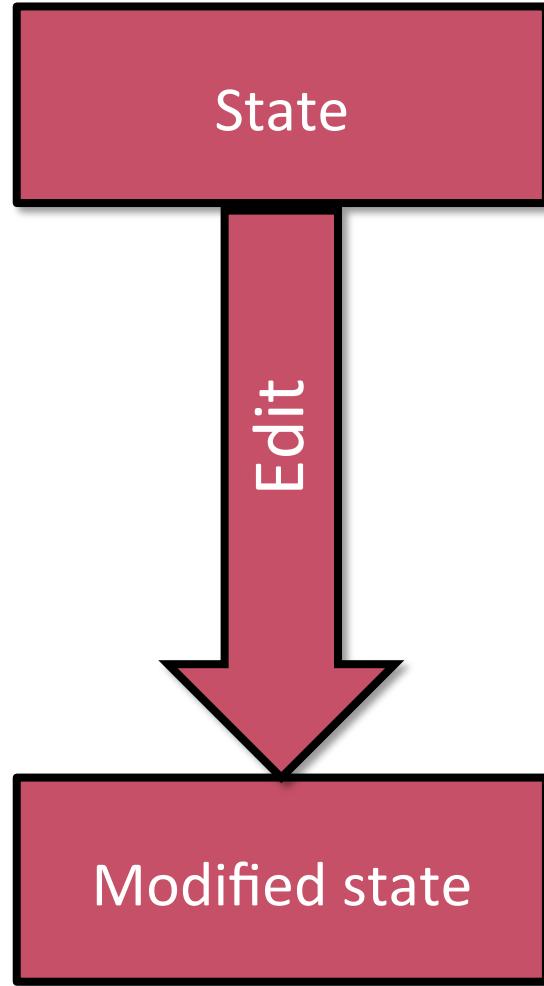
- Editor and EMF model generation
  - Based on grammar
  - Optionally can be initiated from an existing EMF model
- Easy to work with
  - But highly customizable
  - Both the grammar and the generated code
- New development – DSLs over JVM:
  - Xbase expression language
  - JVM Model Inference instead of code generation

# Textual or graphical?

# Comparison

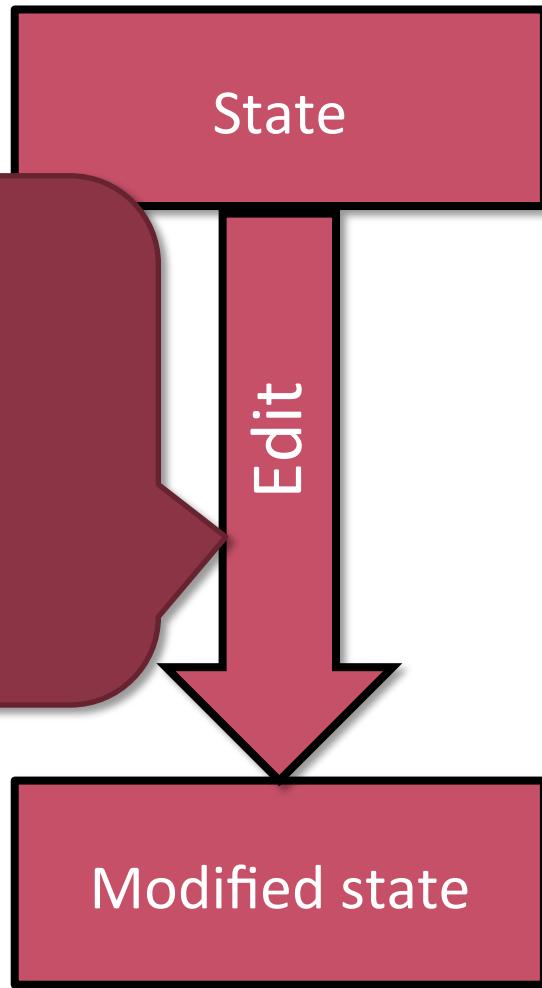
Textual Languages	Graphical Languages
Quick and simple editing	More cumbersome editing
References described as <i>string identifiers</i>	References displayed visually
Inconsistent models during editing	Models always syntactically correct
Automatic formatting	Automatic layouting
Content assist	Tool list to add nodes/edges
Displaying validation errors, offering quick fixes	
Both are supported with EMF-based technologies	

# Editing

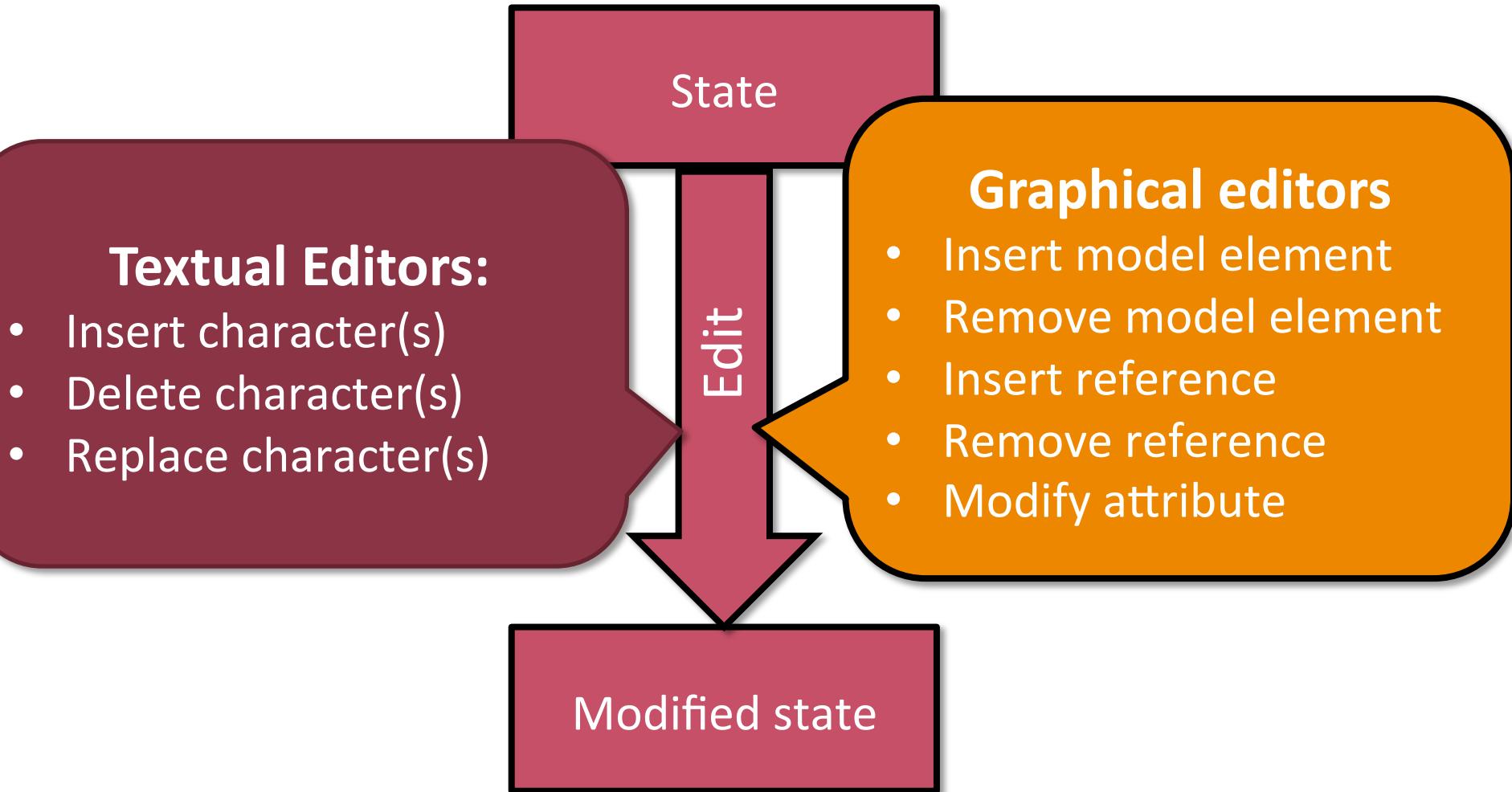


# Editing

- Textual Editors:**
- Insert character(s)
  - Delete character(s)
  - Replace character(s)



# Editing



# Question: textual or graphical?

- No clear choice
- Rules of thumb
  - Behaviour description is usually simpler in textual
  - For structural information graphical is often better
- For simple languages
  - Form-based editing might also be an alternative

# Xtext and GMF on the same instance model

The screenshot shows two views side-by-side. On the left is the Xtext editor view titled "test.socialnetwork" containing UML-like code for a social network. On the right is the GMF diagram view titled "test.socialnetwork\_diagram" showing the corresponding visual representation of the model.

**Xtext Editor View (test.socialnetwork):**

```
SocialNetwork {
    Person Ujhelyi {
        male
        memberships BME, VVEC
    }
    Person Horvath {
        male
        memberships FTSRG
    }
    Community BME {
        Community FTSRG {
            Community test
        }
    }
    Person Test {
        female
        memberships test
    }
    Community VVEC
    Person Proba {
        male
    }
    Community Pr2
    Person valaki {
        male
    }

    Ujhelyi is friend of Horvath
    Test is married to Ujhelyi
}
```

**GMF Diagram View (test.socialnetwork\_diagram):**

- Nodes:
  - Ujhelyi (Person)
  - Horvath (Person)
  - Test (Person)
  - Proba (Person)
  - valaki (Person)
  - BME (Community)
  - VVEC (Community)
  - Pr2 (Community)
  - FTSRG (Community)
  - test (Community)
- Relationships:
  - Ujhelyi is connected to Horvath.
  - Test is connected to Ujhelyi.

**Palette:**

- Community
- Person
- Acquaintance
- Membership

# Derived Graphical viewer support

- Xtext Generic Viewer component
  - Created by Xtext developers
  - Independent from the main Xtext development
  - Requires an extra language
    - to define uni-directional mapping
    - to define format
- See “A fresh look at graphical modeling” for details
  - <http://www.slideshare.net/schwurbel/a-fresh-look-at-graphical-editing-10068461>

# Concrete Syntax Design

Conclusion

# Concrete Syntax Design

- Multiple approaches
    - Textual and/or graphical syntaxes
    - Combinable
  - Large amount of development work needed
    - Directly used by users
    - Usability issues
  - Not everything is coded in an editor
    - Editor + corresponding views form the interface

