Classification and Visualization of Contents using Real-Valued Attributes

Sushil K Bajracharya

Orion Auld

{sbajrach,oauld}@ics.uci.edu

Overview

- Problem Statement

- A Solution Strategy

- Our Application

- Platform (Libraries)

- Application Architecture

- Demo

Content Classification in small

- Content – can be anything. Collection of papers, images, songs, web pages, files etc.

- *Assumption*: we already have an implicit classification, like – 'papers', 'songs' etc.

- Classification in small – create folders and classify contents based on folder hierarchy

Issues

- Hierarchical organization is in essence a way of filtering content
- But, is Static
  - Only one explicit structure
    - Meaning tends to be subjective
  - What are the disadvantages ?
- *Assumption*:
  - We are focusing on classifying, categorizing and filtering, possibly navigating the contents
    - . i.e. seeking structure in the organization itself
  - Issues like versioning, updates, are not of our concern for this project

Solutions

- Associate meta-data (attributes) with contents and classify them according to those meta-data

- Define a schema with relationships between meta-data elements and filter/classify contents by querying on meta-data
  - Eg; get all 'papers' for 'Phase II' I studied in 'ICS 227'
  - DOCUMENT_TYPE = 'paper', EXAM='Phase II', COURSE = 'ICS 227'

- There are different ways of doing this

A Possible Solution

- Attach meta-data to the content itself and, do not explicitly define the schema
- Allow users to see the relationship between contents without defining all possible relationships among the attribute themselves
- Might not be feasible for all types of contents but here is an appropriate domain…

Example

- Contents = music files (mp3)
- Attributes
  - Nominal
    - Artists, Albums, Genre, Style (from ID3 tag)
    - Instrumentation
  - Real
    - …more on this later
      - Rating, Tempo, Mood, Ambience, Aggressiveness, Fidelity, Danceability, Year (can be nominal)
      - can be embedded or retrieved from external source, like a web-service
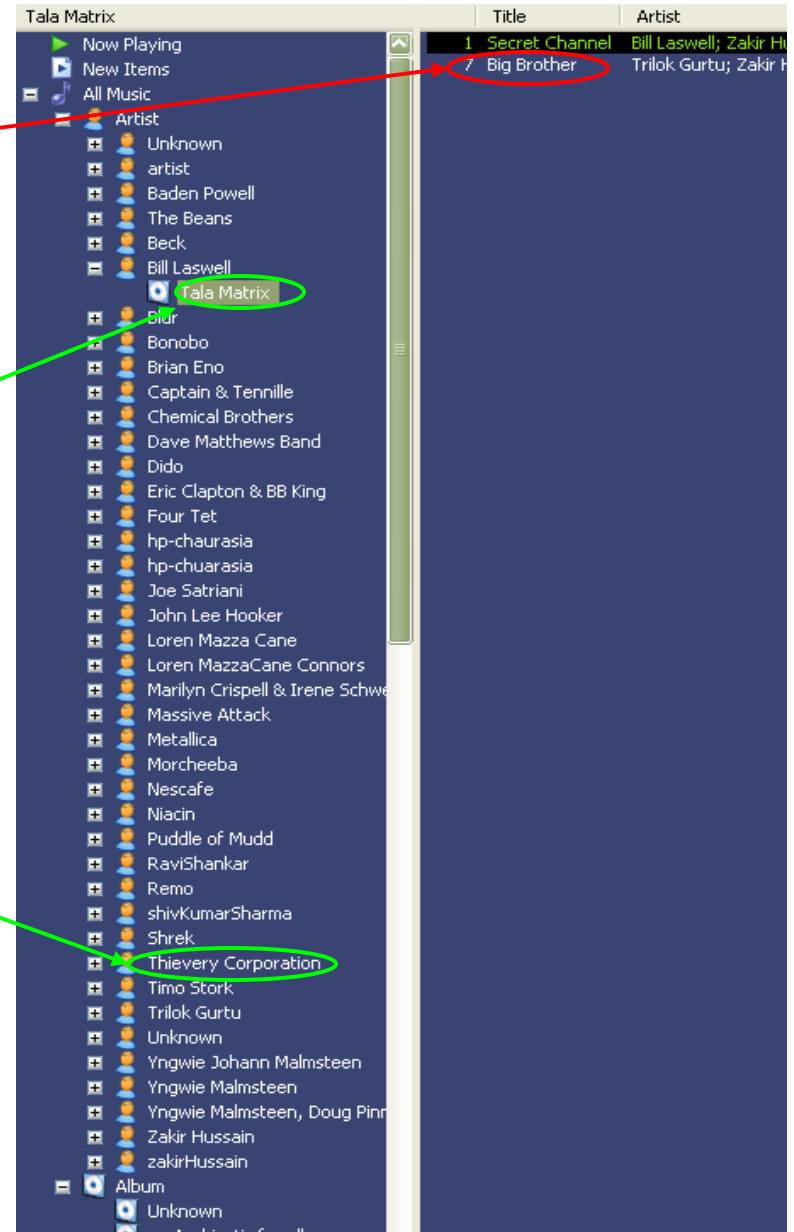
Hierarchies using nominal attributes

- **Organizing** into hierarchies is one way of *classifying* data

- **Navigating** a hierarchy is one way of *filtering* data

- *Demonstration*: music Libraries
  - next few slides >>
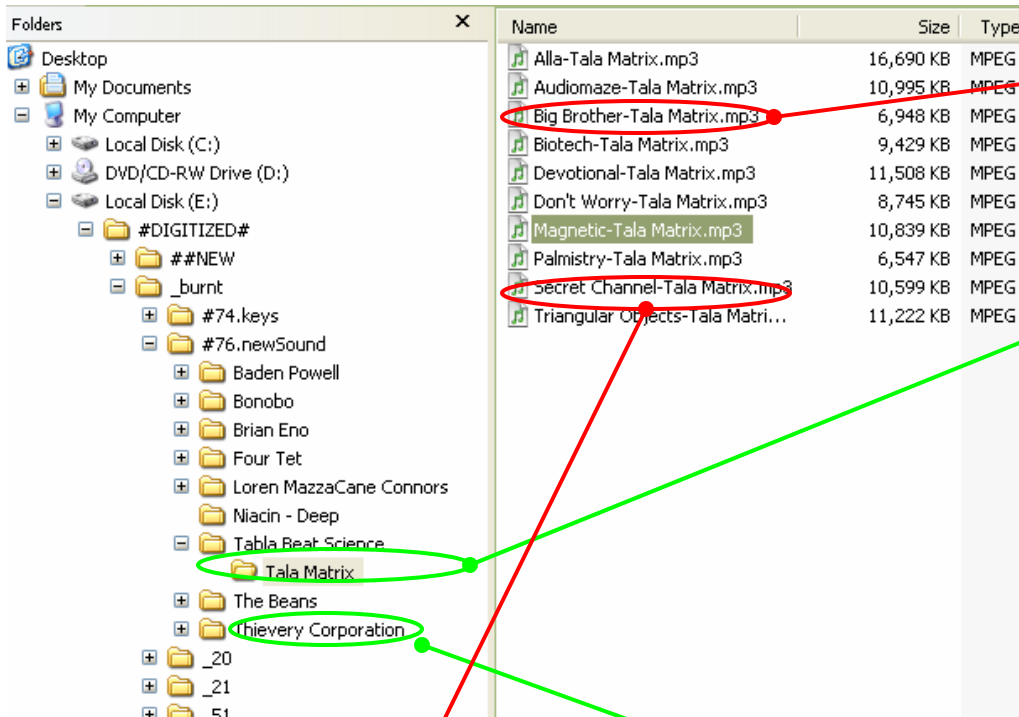
Simple Hierarchical Classification

Attribute Based Classification

Title	▲	Genre	Artist	Album	Rating	Length	Bit Rate	File Name
Part 3 - Pursuance/Part 4 -Psalm		Jazz	John Coltrane	A Love Supreme [Impulse]	★★★	17:52	192Kbps	E:\#DIGITIZED#\##NEW
My Favorite Things		Jazz	John Coltrane	My Favorite Things [Atlan...	★★★	13:43	192Kbps	E:\#DIGITIZED#\##NEW
Ev'ry Time We Say Goodbye		Jazz	John Coltrane	My Favorite Things [Atlan...	★★★	5:43	192Kbps	E:\#DIGITIZED#\##NEW
Summertime		Jazz	John Coltrane	My Favorite Things [Atlan...	★★★	11:35	192Kbps	E:\#DIGITIZED#\##NEW
But Not for Me		Jazz	John Coltrane	My Favorite Things [Atlan...	★★★	9:34	192Kbps	E:\#DIGITIZED#\##NEW
Goldie		Jazz	Lee Morgan	Planet Jazz	★★★	9:07	192Kbps	E:\#DIGITIZED#_burnt\
Whisper Not		Jazz	Lee Morgan	Planet Jazz	★★★	7:19	192Kbps	E:\#DIGITIZED#_burnt\
Close your eyes		Jazz	Lee Morgan	Planet Jazz	★★★	9:45	192Kbps	E:\#DIGITIZED#_burnt\
El-Lail (Night)		Jazz	Lee Morgan	Planet Jazz	★★★	4:24	192Kbps	E:\#DIGITIZED#_burnt\
Blue March For Europe no !		Jazz	Lee Morgan	Planet Jazz	★★★	10:14	192Kbps	E:\#DIGITIZED#_burnt\
Along came Manon		Jazz	Lee Morgan	Planet Jazz	★★★	11:19	192Kbps	E:\#DIGITIZED#_burnt\
Free Jazz		Jazz	Ornette Coleman	Free Jazz (A Collective I...	★★★	37:07	192Kbps	E:\#DIGITIZED#_burnt\
First Take [*] [#]		Jazz	Ornette Coleman	Free Jazz (A Collective I...	★★★	17:03	192Kbps	E:\#DIGITIZED#_burnt\
'Round Midnight		Jazz	Anthony Braxton; Ran Blake	A Memory of Vienna	★★★	11:06	192Kbps	E:\#DIGITIZED#\##NEW
Yardbird Suite		Jazz	Anthony Braxton; Ran Blake	A Memory of Vienna	★★★	5:37	192Kbps	E:\#DIGITIZED#\##NEW
You Go to My Head		Jazz	Anthony Braxton; Ran Blake	A Memory of Vienna	★★★	7:57	192Kbps	E:\#DIGITIZED#\##NEW
Just Friends		Jazz	Anthony Braxton; Ran Blake	A Memory of Vienna	★★★	6:23	192Kbps	E:\#DIGITIZED#\##NEW
Alone Together		Jazz	Anthony Braxton; Ran Blake	A Memory of Vienna	★★★	4:58	192Kbps	E:\#DIGITIZED#\##NEW
Four		Jazz	Anthony Braxton; Ran Blake	A Memory of Vienna	★★★	4:53	192Kbps	E:\#DIGITIZED#\##NEW
Soul Eyes		Jazz	Anthony Braxton; Ran Blake	A Memory of Vienna	★★★	7:33	192Kbps	E:\#DIGITIZED#\##NEW
I'm Getting Sentimental over You		Jazz	Anthony Braxton; Ran Blake	A Memory of Vienna	★★★	7:08	192Kbps	E:\#DIGITIZED#\##NEW
Sing		Jazz	Brus Trio; Roscoe Mitchell	After Fallen Leaves	★★★	15:23	192Kbps	E:\#DIGITIZED#\##NEW
A Lovely Day at the Point		Jazz	Brus Trio; Roscoe Mitchell	After Fallen Leaves	★★★	4:41	192Kbps	E:\#DIGITIZED#\##NEW
The Reverend Frank Wright		Jazz	Brus Trio; Roscoe Mitchell	After Fallen Leaves	★★★★	7:06	192Kbps	E:\#DIGITIZED#\##NEW
And Then There Was Peace		Jazz	Brus Trio; Roscoe Mitchell	After Fallen Leaves	★★★	4:51	192Kbps	E:\#DIGITIZED#\##NEW
The Two Faces of Everett Sloane		Jazz	Brus Trio; Roscoe Mitchell	After Fallen Leaves	★★★	3:30	192Kbps	E:\#DIGITIZED#\##NEW
After Fallen Leaves		Jazz	Brus Trio; Roscoe Mitchell	After Fallen Leaves	★★★	8:01	192Kbps	E:\#DIGITIZED#\##NEW
Mr. Freddie		Jazz	Brus Trio; Roscoe Mitchell	After Fallen Leaves	★★★	4:34	192Kbps	E:\#DIGITIZED#\##NEW
Come Gather Some Things		Jazz	Brus Trio; Roscoe Mitchell	After Fallen Leaves	★★★	11:58	192Kbps	E:\#DIGITIZED#\##NEW
Play With the Whistler		Jazz	Brus Trio; Roscoe Mitchell	After Fallen Leaves	★★★	7:25	192Kbps	E:\#DIGITIZED#\##NEW
Cool Struttin'		Jazz	Sonny Clark	Cool Struttin'	★★★	9:24	192Kbps	E:\#DIGITIZED#_burnt\
Blue Minor		Jazz	Sonny Clark	Cool Struttin'	★★★	10:19	192Kbps	E:\#DIGITIZED#_burnt\
Sippin' at Bells		Jazz	Miles Davis	Cool Struttin'	★★★	8:19	192Kbps	E:\#DIGITIZED#_burnt\
Deep Night		Jazz	Sonny Clark	Cool Struttin'	★★★	9:34	192Kbps	E:\#DIGITIZED#_burnt\
Royal Flush		Jazz	Sonny Clark	Cool Struttin'	★★★	9:02	192Kbps	E:\#DIGITIZED#_burnt\
Lover		Jazz	R.Rodgers-L.Hart	Cool Struttin'	★★★	7:02	192Kbps	E:\#DIGITIZED#_burnt\
With a Song in My Heart		Jazz	Art Taylor; Curtis Fuller; Donal...	Sonny's Crib	★★★	7:55	192Kbps	E:\#DIGITIZED#_burnt\
With a Song in My Heart [Alter...		Jazz	Art Taylor; Curtis Fuller; Donal...	Sonny's Crib	★★★	6:50	192Kbps	E:\#DIGITIZED#_burnt\
Speak Low {from the Musical Pr...		Jazz	Art Taylor; Curtis Fuller; Donal...	Sonny's Crib	★★★	7:29	192Kbps	E:\#DIGITIZED#_burnt\
Speak Low {from the Musical Pr...		Jazz	Art Taylor; Curtis Fuller; Donal...	Sonny's Crib	★★★	13:31	192Kbps	E:\#DIGITIZED#_burnt\
Come Rain or Come Shine		Jazz	Art Taylor; Curtis Fuller; Donal...	Sonny's Crib	★★★	8:34	192Kbps	E:\#DIGITIZED#_burnt\
Sonny's Crib		Jazz	Art Taylor; Curtis Fuller; Donal...	Sonny's Crib	★★★	8:48	192Kbps	E:\#DIGITIZED#_burnt\
Sonny's Crib [Alternate Take][*]		Jazz	Art Taylor; Curtis Fuller; Donal...	Sonny's Crib	★★★	6:57	192Kbps	E:\#DIGITIZED#_burnt\
News for Lulu		Jazz	Art Taylor; Curtis Fuller; Donal...	Sonny's Crib	★★★	9:53	192Kbps	E:\#DIGITIZED#_burnt\
Cool Struttin'		Jazz	Sonny Clark Memorial Quartet	Voodoo	★★★	5:30	192Kbps	E:\#DIGITIZED#_burnt\
Minor Meeting		Jazz	Sonny Clark Memorial Quartet	Voodoo	★★★	4:40	192Kbps	E:\#DIGITIZED#_burnt\
Nicely		Jazz	Sonny Clark Memorial Quartet	Voodoo	★★★	5:37	192Kbps	E:\#DIGITIZED#_burnt\
Something Special		Jazz	Sonny Clark Memorial Quartet	Voodoo	★★★	4:50	192Kbps	E:\#DIGITIZED#_burnt\

Classifying contents in a Hierarchical Structure based on Attributes

So…

- We saw classification based on nominal attributes
- What does it mean to say classification based on Real-Valued attributes and how can we visualize it
- Lets see…

Different Strokes

- Utility of nominal attributes
  - Comprehensibility
  - Comprehensivity

- Tyranny of nominality
  - Binary decisions only, or crude approximations
  - At its best in the presence of strong structure to embody rigid concepts and relationships

- Different Interfaces for different problems

Using Real-Valued attributes for classification
(Part I)

- ## Consider music domain again

- ## Quantitative values

  - Tempo, Rating, Bit rate

- ## Qualitative values

  - Familiar-Exotic, Conservative-Experimental, Sad-Happy, Classic-Contemporary

  - Mental Tyranny: Genre/Style is really real-valued!

…using Real Attributes (Part II)

- Application domains
  - Browsing commercial music
  - Creating playlists

- Applicable Strengths
  - Fluid specification
  - More easily visualized
  - Variable degree of importance

…using Real Attributes (Part III)

- How do all of these attributes get set?
  - By the user
  - By a third party
  - Groupware

- What do these qualitative attributes mean?
  - Must be consistent
  - Must be differentiating
  - Other than that, it doesn't matter
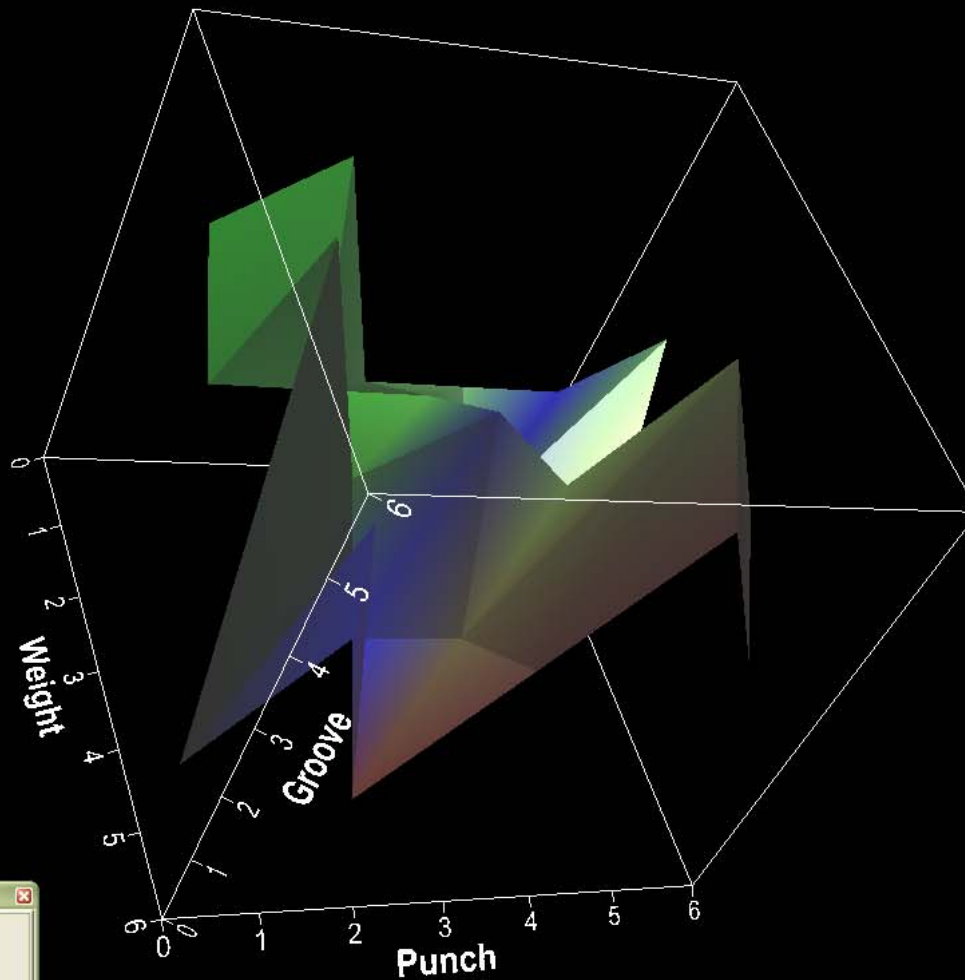
Our Application

- Clustering of library in several contexts
  - For playlists
    - Single centroid, user-specified
    - Single centroid, specified by song
  - For categorization
    - Multiple centroids
- Visualization of clusters
- *…next* >> a screenshot
  - [Note: application is partially complete]

ICS 227 - Advanced UI Architectures Presentaion, Spring 2004. UCI

Platform

- What we need
  - Toolkit for *classifying dataset* using ML
    - WEKA
  - *Visualizing* data points in 3D space
    - 3D Plot from FreeHEP java library
  - A platform to put these all together
    - (We will not talk about these in detail)
      - Java (AWT ☹, Swing)
      - AspectJ to localize/encapsulate event triggering message calls

Clustering Tools

- Weka
  - Easy to use machine learning toolkit in Java
  - Provides several clustering algorithms
    - K-Means
    - Gaussian mixture models
  - Some are only appropriate offline
  - For maximal effectiveness, we may need to implement our own online algorithm

3D Plotting Package

- FreeHEP visualization library
  - A Huge library for scientific visualization
    - www.freehep.org
  - We are using
    - Java 3D plotting package from FreeHEP
      - org.freehep.j3d.plot (available as a single .jar)
      - Uses Java3D underneath
      - For basic functionality we don't need to know intricacies of Java 3D
      - …[some details, next slide]

3D Plotting Package (org.freehep.j3d.plot)

- **Class Hierarchy**
- class java.lang.**Object**
  - class org.freehep.j3d.plot.**AbstractPlotBuilder**
    - class org.freehep.j3d.plot.**LegoBuilder**
    - class org.freehep.j3d.plot.**ScatterBuilder**
    - class org.freehep.j3d.plot.**SurfaceBuilder**
  - class org.freehep.j3d.plot.**AxisBuilder**
    - class org.freehep.j3d.plot.**XAxisBuilder**
    - class org.freehep.j3d.plot.**YAxisBuilder**
    - class org.freehep.j3d.plot.**ZAxisBuilder**
  - class org.freehep.j3d.plot.**AxisLabelCalculator**
  - class java.awt.**Component** (implements java.awt.image.ImageObserver, java.awt.MenuContainer, java.io.Serializable)
    - class java.awt.**Canvas** (implements javax.accessibility.Accessible)
      - class javax.media.j3d.Canvas3D
        - » class org.freehep.j3d.plot.**Plot3D**
        - » class org.freehep.j3d.plot.**LegoPlot**
        - » class org.freehep.j3d.plot.**SurfacePlot**
  - class org.freehep.j3d.plot.**KeyNavigator**
  - class org.freehep.j3d.plot.**Rainbow**
  - class javax.media.j3d.SceneGraphObject
    - class javax.media.j3d.Node
      - class javax.media.j3d.Leaf
        - » class javax.media.j3d.Behavior
        - » class org.freehep.j3d.plot.**KeyNavigatorBehavior**
        - » class org.freehep.j3d.plot.**PlotKeyNavigatorBehavior**
  - class org.freehep.j3d.plot.**TimeStamp**
- **Interface Hierarchy**
- interface org.freehep.j3d.plot.**Binned2DData**
- interface org.freehep.j3d.plot.**Unbinned3DData**

Extends Java3D

3D canvases to draw on (View)

Defines Mouse and Keyboard Behavior (Controller)
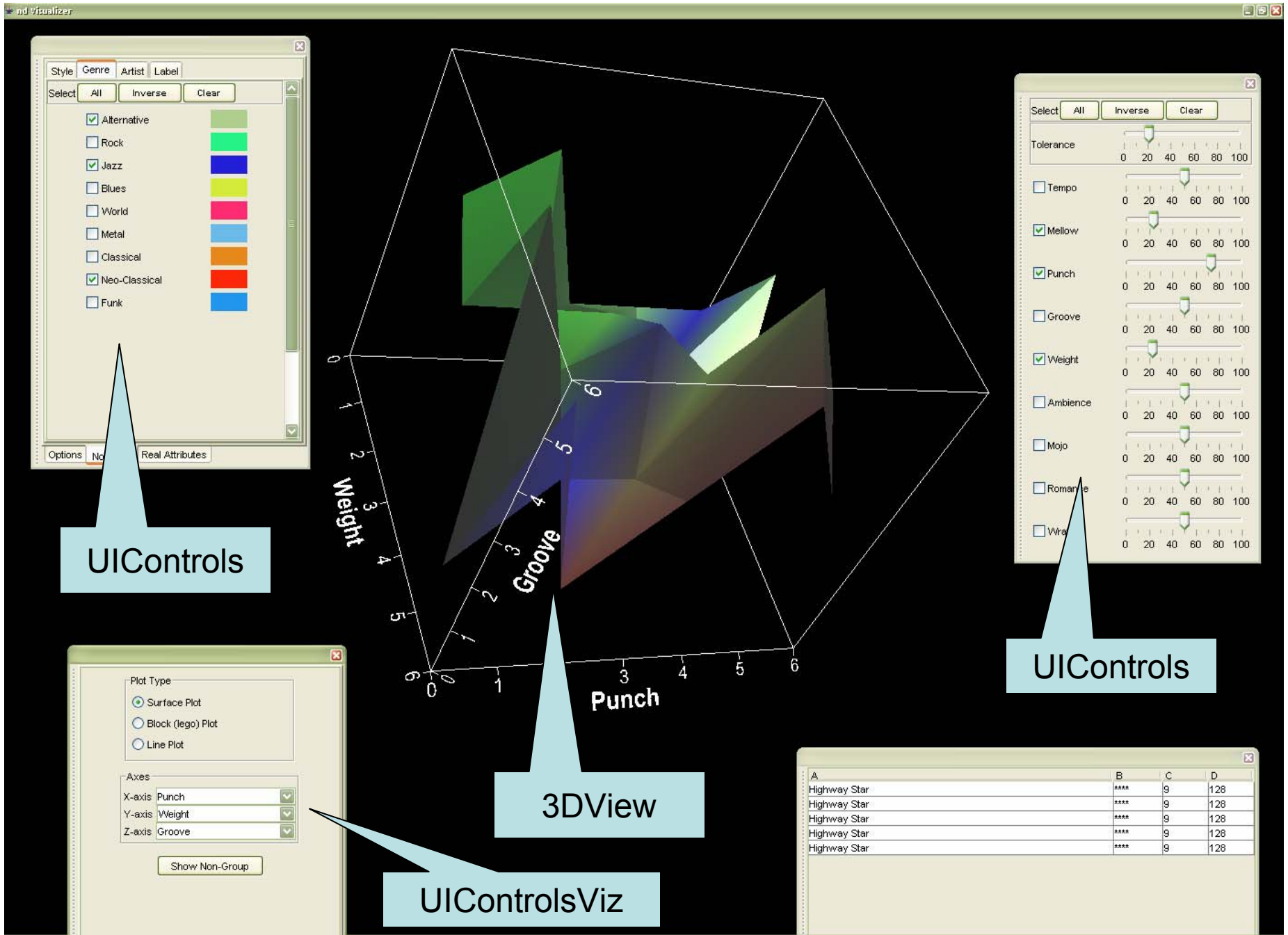
DataModel to feed the 3D Canvases (Model)

...using org.freehep.j3d.plot

- Provide a Data-Model by implementing
  - Binned2DData
    - interface org.freehep.j3d.plot.**Binned2DData**
- Implement classes and hook-up custom controllers to synchronize the MVC setup
  - LegoPlot and SurfacePlot
    - » class org.freehep.j3d.plot.**LegoPlot**
    - » class org.freehep.j3d.plot.**SurfacePlot**

Message Flow / Application Architecture

<<Observable>>
RealChanges
NominalChanges

AppModel

UIControls

<<Data>>

Content_Header

Content_Data

<<Observer>>
Agent

<<Observable>>
VizOption

<<Observable>>
VizModel

<<Observer>>
VizController

UIControlsViz

3DView

A → **B**		A sends event to B
A –·–·→ **B**		A updates B
A - - → **B**		B uses A
A•····→ **B**		B is created reading A
A —— **B**		A implements B

ICS 227 - Advanced UI Architectures Presentaion, Spring 2004. UCI
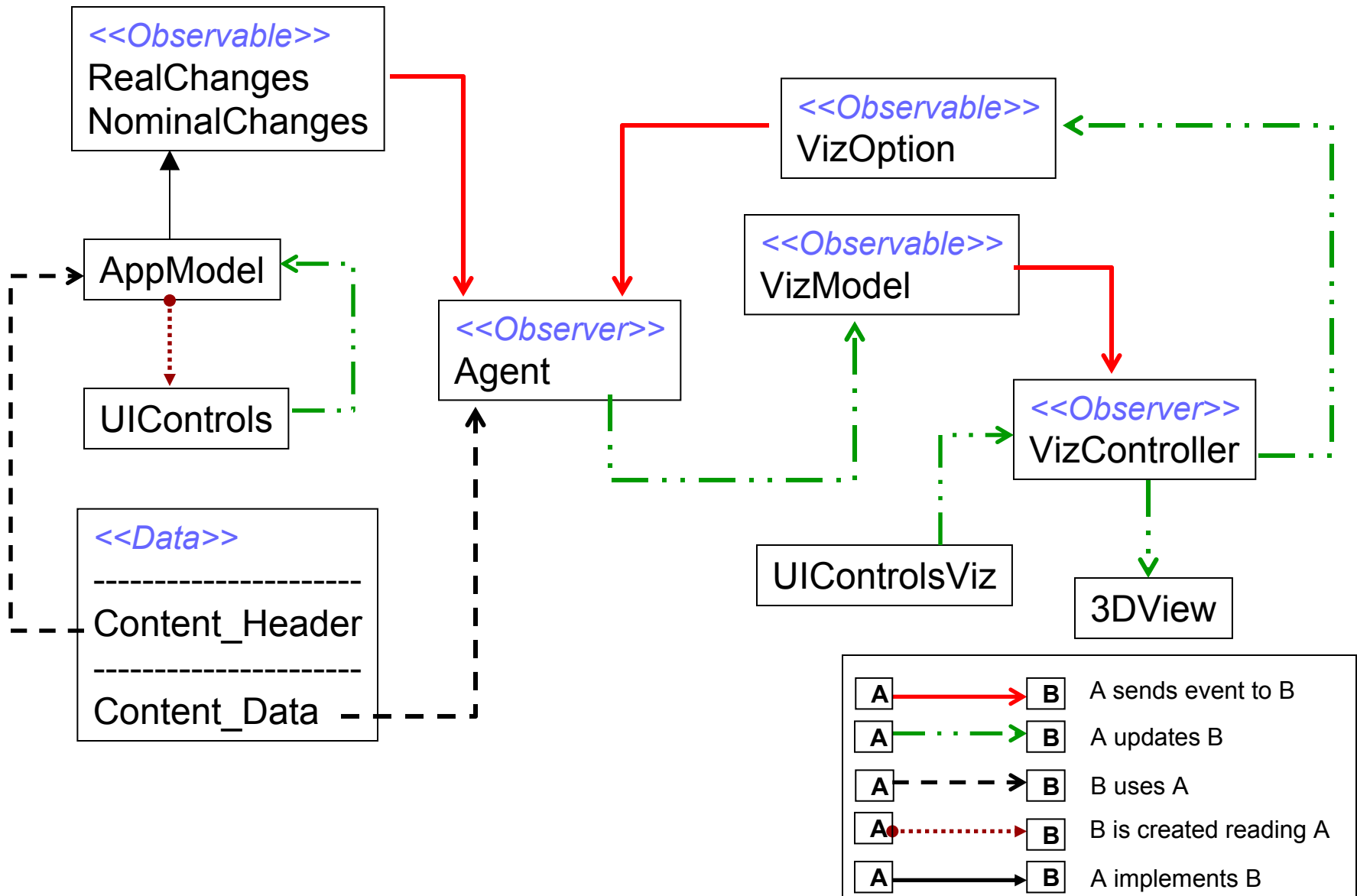
Localizing event triggering method calls using AspectJ

- **public aspect** AspectPredicates {

-

-   **pointcut** predicateChanges():
-    **call** (* RealChanges.* (..)) ||
-    **call** (* NominalChanges.* (..));
-

-   **after**(AppModelImpl o): **target**(o) && predicateChanges()
-   {
-    o.notifyListeners();
-   }
- }

Capture all calls to methods that change Real/Nominal settings

Execute this method after the methods are called

Wrap-up

- Demo
- Questions, and..
- Thank you