Visualizing the Hollywood Allegations

Trina Sarkar & Basel Al Sharaf
Background
Proposal
We hope to draw parallels between the timelines of individual successes and the known incidents of the allegations to raise the question of Hollywood’s complacency.
Relevant Prior Work
In this era of enforcement, the government has conducted 450 investigations of colleges for possibly mishandling reports of sexual violence.

So far, 96 cases have been resolved and 354 remain open.
Project
Click on a celebrity's face to view their timeline
Selection using multiple filters
Selection with multiple individuals
Our Current Progress
Points for Feedback

Effectiveness of visualization designs

What else would be interesting to learn about from this data?

Sensitivity of topic

Limited data
Thank You
Bitcoin and Cryptocurrencies

Investigative Data Analysis

Danny Diekroeger, CS448B
Google Trends: Bitcoin

Interest over time

Interest by region

1. Nigeria: 100
2. Ghana: 82
3. South Africa: 72
4. Slovenia: 67
5. Estonia: 61
Bitcoin Price and Market Capitalization (2013-present)
Total Market Capitalization (All Cryptocurrencies, 2013-present)
Percentage of Total Market Capitalization (Dominance)

[Graph showing the percentage of total market capitalization over time, with a significant drop in the percentage by 2017.]
<table>
<thead>
<tr>
<th>#</th>
<th>Name</th>
<th>Market Cap</th>
<th>Price</th>
<th>Volume (24h)</th>
<th>Circulating Supply</th>
<th>Change (24h)</th>
<th>Price Graph (7d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bitcoin</td>
<td>$107,317,907,106</td>
<td>$6434.77</td>
<td>$6,763,790,000</td>
<td>16,677,800 BTC</td>
<td>4.40%</td>
<td><img src="https://via.placeholder.com/150" alt="Graph" /></td>
</tr>
<tr>
<td>2</td>
<td>Ethereum</td>
<td>$29,991,447,347</td>
<td>$313.38</td>
<td>$1,220,520,000</td>
<td>95,704,658 ETH</td>
<td>-0.67%</td>
<td><img src="https://via.placeholder.com/150" alt="Graph" /></td>
</tr>
<tr>
<td>3</td>
<td>Bitcoin Cash</td>
<td>$23,689,112,638</td>
<td>$1410.04</td>
<td>$4,999,180,000</td>
<td>16,800,313 BCH</td>
<td>2.02%</td>
<td><img src="https://via.placeholder.com/150" alt="Graph" /></td>
</tr>
<tr>
<td>4</td>
<td>Ripple</td>
<td>$7,773,005,878</td>
<td>$0.201731</td>
<td>$145,955,000</td>
<td>38,531,538,922 XRP*</td>
<td>-0.17%</td>
<td><img src="https://via.placeholder.com/150" alt="Graph" /></td>
</tr>
<tr>
<td>5</td>
<td>Dash</td>
<td>$3,276,606,095</td>
<td>$426.29</td>
<td>$433,957,000</td>
<td>7,686,385 DASH</td>
<td>-0.30%</td>
<td><img src="https://via.placeholder.com/150" alt="Graph" /></td>
</tr>
<tr>
<td>6</td>
<td>Litecoin</td>
<td>$3,266,687,564</td>
<td>$60.70</td>
<td>$282,486,000</td>
<td>53,813,382 LTC</td>
<td>1.48%</td>
<td><img src="https://via.placeholder.com/150" alt="Graph" /></td>
</tr>
<tr>
<td>7</td>
<td>Monero</td>
<td>$1,884,745,240</td>
<td>$122.82</td>
<td>$126,474,000</td>
<td>15,345,839 XMR</td>
<td>-2.44%</td>
<td><img src="https://via.placeholder.com/150" alt="Graph" /></td>
</tr>
<tr>
<td>8</td>
<td>Neo</td>
<td>$1,824,186,000</td>
<td>$28.06</td>
<td>$44,945,600</td>
<td>65,000,000 NEO*</td>
<td>2.05%</td>
<td><img src="https://via.placeholder.com/150" alt="Graph" /></td>
</tr>
<tr>
<td>9</td>
<td>NEM</td>
<td>$1,578,617,000</td>
<td>$0.186513</td>
<td>$8,597,310</td>
<td>8,999,999,999 XEM*</td>
<td>2.04%</td>
<td><img src="https://via.placeholder.com/150" alt="Graph" /></td>
</tr>
<tr>
<td>10</td>
<td>IOTA</td>
<td>$1,664,229,859</td>
<td>$0.598745</td>
<td>$37,671,900</td>
<td>2,779,530,283 MIOTA*</td>
<td>2.45%</td>
<td><img src="https://via.placeholder.com/150" alt="Graph" /></td>
</tr>
<tr>
<td>11</td>
<td>Ethereum Classic</td>
<td>$1,523,053,194</td>
<td>$15.64</td>
<td>$389,522,000</td>
<td>97,407,453 ETC</td>
<td>-6.97%</td>
<td><img src="https://via.placeholder.com/150" alt="Graph" /></td>
</tr>
</tbody>
</table>
Tried predicting price jumps
Old data
Could use some of their novel metrics, like “Whale Index”
- Studies market from ecological lens
- Each cryptocurrency as own species
- Novel approach to studying the market
My Project
Focus: Dynamics of Trading Pairs

- Bitcoin as vehicle for international currency exchange

- KYC compliance causes bottleneck for trading fiat-BTC

- Must first buy Bitcoin (or ETH), then use that to buy Altcoins
First Step: Develop Visualization Tool for Trading Pairs and Flow of Value
Prototype: Visualization for Trading Pairs and Flow of Value

USD → Bitcoin → EUR

USD → USDT → Ether

USD → Other Altcoins

KRW → Bitcoin → JPY

KRW → BCC/BCH → Monero

Other Altcoins
## Seeking: Historical Trading Pairs Data

### Bitcoin Markets

<table>
<thead>
<tr>
<th>#</th>
<th>Source</th>
<th>Pair</th>
<th>Volume (24h)</th>
<th>Price</th>
<th>Volume (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bitfinex</td>
<td>BTC/USD</td>
<td>$754,154,000</td>
<td>$6266.30</td>
<td>11.15%</td>
</tr>
<tr>
<td>2</td>
<td>Bithumb</td>
<td>BTC/KRW</td>
<td>$533,803,000</td>
<td>$6498.12</td>
<td>7.89%</td>
</tr>
<tr>
<td>3</td>
<td>Bittrex</td>
<td>BCC/BTC</td>
<td>$505,980,000</td>
<td>$6569.14</td>
<td>7.48%</td>
</tr>
<tr>
<td>4</td>
<td>Bitfinex</td>
<td>BCH/BTC</td>
<td>$403,621,000</td>
<td>$6549.29</td>
<td>5.97%</td>
</tr>
<tr>
<td>5</td>
<td>GDAX</td>
<td>BTC/USD</td>
<td>$328,305,000</td>
<td>$6401.43</td>
<td>4.85%</td>
</tr>
<tr>
<td>6</td>
<td>Poloniex</td>
<td>BCH/BTC</td>
<td>$316,045,000</td>
<td>$6548.30</td>
<td>4.67%</td>
</tr>
<tr>
<td>7</td>
<td>bitFlyer</td>
<td>BTC/JPY</td>
<td>$308,048,000</td>
<td>$6435.92</td>
<td>4.55%</td>
</tr>
<tr>
<td>8</td>
<td>Bitstamp</td>
<td>BTC/USD</td>
<td>$191,152,000</td>
<td>$6393.54</td>
<td>2.83%</td>
</tr>
<tr>
<td>9</td>
<td>HitBTC</td>
<td>BCH/BTC</td>
<td>$187,283,000</td>
<td>$6585.46</td>
<td>2.77%</td>
</tr>
<tr>
<td>10</td>
<td>Coinone</td>
<td>BTC/KRW</td>
<td>$145,635,000</td>
<td>$6525.82</td>
<td>2.15%</td>
</tr>
<tr>
<td>11</td>
<td>Binance</td>
<td>BCC/BTC</td>
<td>$123,122,000</td>
<td>$6303.65</td>
<td>1.82%</td>
</tr>
<tr>
<td>12</td>
<td>Poloniex</td>
<td>BTC/USDT</td>
<td>$121,417,000</td>
<td>$6285.78</td>
<td>1.79%</td>
</tr>
<tr>
<td>13</td>
<td>OKEx</td>
<td>BCC/BTC</td>
<td>$120,937,000</td>
<td>$6612.11</td>
<td>1.79%</td>
</tr>
<tr>
<td>14</td>
<td>Gemini</td>
<td>BTC/USD</td>
<td>$109,765,000</td>
<td>$6353.45</td>
<td>1.62%</td>
</tr>
<tr>
<td>15</td>
<td>Poloniex</td>
<td>ETH/BTC</td>
<td>$107,534,000</td>
<td>$6270.34</td>
<td>1.59%</td>
</tr>
<tr>
<td>16</td>
<td>Kraken</td>
<td>BTC/EUR</td>
<td>$92,499,200</td>
<td>$6363.98</td>
<td>1.37%</td>
</tr>
<tr>
<td>17</td>
<td>Bittrex</td>
<td>BTC/USDT</td>
<td>$91,107,200</td>
<td>$6306.69</td>
<td>1.35%</td>
</tr>
</tbody>
</table>
Help and suggestions are appreciated!
Cartograms vs. Choropleth Maps

A framework for deciding which to use

Juliette Love
Nov. 15, 2017
The Problem

When visualizing a dataset, which is a better fit?

- Which properties of the data are relevant in this decision?
- How can we create definitive guidelines?
- How do we define better?
Relevant Work

Cleveland & McGill (from class)
- Which marks are perceptually effective
- Includes color and area → area in maps is perceptually different

Dunn 1988, Rittschof & Kulhavy 1998
- Studies choropleths, cartograms compared to other types of maps
- Goal of seeing which is more effective
- Dependent on chosen dataset

Sun & Lee 2010
- Pros and cons of contiguous and non-contiguous cartograms
- Show which cartogram type is best
- Still dependent on dataset type
Relevant Work

Cleveland & McGill (from class)
- Which marks are perceptually effective
- Includes color and area → area in maps is perceptually different

Dunn 1988, Rittschof & Kulhavy 1998
- Studies choropleths, cartograms compared to other types of maps
- Goal of seeing which is more effective
  - Dependent on chosen dataset

Sun & Lee 2010
- Pros and cons of contiguous and non-contiguous cartograms
- Show which cartogram type is best
- Still dependent on dataset type
Relevant Work

Cleveland & McGill (from class)
- Which marks are perceptually effective
- Includes color and area → area in maps is perceptually different

Dunn 1988, Rittschof & Kulhavy 1998
- Studies choropleths, cartograms compared to other types of maps
- Goal of seeing which is more effective
- Dependent on chosen dataset

Sun & Lee 2010
- Pros and cons of contiguous and non-contiguous cartograms
- **Show which cartogram type is best**
- Still dependent on dataset type
Progress

- Discussion of successful and unsuccessful cartograms
- Common problems with unsuccessful cartograms
- Which features of the data cause these problems?
Progress

- Discussion of successful and unsuccessful cartograms
- Common problems with unsuccessful cartograms
- Which features of the data cause these problems?

<table>
<thead>
<tr>
<th>Successful</th>
<th>Unsuccessful</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geographically-based metrics</td>
<td>Non-geographically based metrics</td>
</tr>
<tr>
<td>Comparable units of area</td>
<td>Large differences in areas</td>
</tr>
<tr>
<td>Reasonable(?) data variation</td>
<td>Huge data variation</td>
</tr>
<tr>
<td>Visualizing two metrics</td>
<td>One metric with two normalizations</td>
</tr>
</tbody>
</table>
Feedback

- **How to define success?**
  Accuracy? Proportional accuracy? Recall?
  Different past studies have used different metrics

- **Are there any other types of maps that would make sense to include?**
  Dorling (non-contiguous) cartograms, Gastner-Newman (contiguous) cartograms,
  choropleth maps, dot maps, proportional symbol maps

- **Tips for user studies?**

- **General suggestions for directions/improvement?**
What’s your guess?
Using prediction and feedback to help people remember and trust data.

Sophia Pink
“I would’ve guessed that.”
Hindsight bias
Draw your line on the chart below

Free tip: Your line should go through this point.
Percent of children who attended college

Parents’ income percentile

Poorest  10th  20th  30th  40th  50th  60th  70th  80th  90th  Richest

10%  20%  30%  40%  50%  60%  70%  80%  90%  100%

Your guess

Reality
Data Through Others’ Eyes: The Impact of Visualizing Others’ Expectations on Visualization Interpretation

Yea-Seul Kim, Katharina Reinecke and Jessica Hullman
Results

High, accurate social consensus:

More accurate recollection of data

Prediction and social consensus NOT aligned with data:

Less trust in actual results
Confirmation bias
How can we use prediction and feedback to design graphics that encourage people to trust data that may be misaligned with their and others’ pre-existing views?
Progress

Exploring current solutions

Mock-ups of other ideas

Testing mock-ups with users
Solution ideas

Go in blind: do not show which data is from which category.

Quiz: ask people to answer questions, not draw a line.

Reframe perspective: Emphasize how many people have changed their mind, not how many people got it wrong.

Selective feedback: choose how much feedback to show based on results.
Questions

Do you know of any related visualization strategies that I should look at?

Do you have suggestions for testing out a lot of different visualization ideas?
Creating a Criminal: Visualizing Juvenile Justice Data

James Lyons
Description

- Similar to “Visualize developmental learning data” suggestion
- **Problem:** Individual risk factors increase the chance of juvenile delinquency
- **Motivation:** Identifying trends in risk factors might provide (relatively) simple solutions to reducing the risk of juvenile delinquency/recidivism. Making the data accessible allows more POV for solving that problem
- (General) -> (Specific)
Comparison to Prior Work

- Interactivity (Dynamic vs Static)
- Focus (National, State, International)
- Accessibility
Example A (Interactivity)

Girls' Increasing Share of the Juvenile Justice System, 1992 to 2012/2013


* At the time of this writing, 2012 is the most current available juvenile arrest data and 2013 is the most current juvenile case processing data.
Example B (Focus)
Current Progress

- Compendium of National Juvenile Justice Datasets
- R Shiny
- Cartogram/Bar chart/Line Chart/Scatter plot
- Filterable by student victimization at school, substance use, emotional characteristics
Questions

- Value and interest in a visualization like this
- Formally test for accessibility or focus more on searching for trends in the data using my completed data visualization
CS448B Final Project

Project Progress

The Billionaire Dataset
Jinglin Shan & Kristy Duong
Description of problem

Dataset: Billionaire Characteristics

Problem/Motivation: By understanding the sources and centers of wealth, we can perhaps find trends in wealth inequality and the causes of that inequality. Are certain geographic locations or businesses simply more profitable or is there something more to the story? How does wealth source pattern change through time?
Prior work

Most Of The World's Billionaires Made Their Money In These 5 Industries (Business Insider)
How Billionaires Get Rich: Which Industries Make The Most Mega-Fortunes (Forbes)

<table>
<thead>
<tr>
<th>TOP 5 PRIMARY INDUSTRIES</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Finance, Banking &amp; Investment</td>
<td>19.3%</td>
</tr>
<tr>
<td>Industrial Conglomerates</td>
<td>12.1%</td>
</tr>
<tr>
<td>Real Estate</td>
<td>7.1%</td>
</tr>
<tr>
<td>Non-Profit &amp; Social Organisations</td>
<td>5.0%</td>
</tr>
<tr>
<td>Textiles, Apparel &amp; Luxury Goods</td>
<td>4.9%</td>
</tr>
</tbody>
</table>

Difference

- Find patterns that depends on time
- Analyze data in more detailed category classification
- Explore data by region
Current Visualization

Click to demo
Current Progress

- Got familiar with dataset.
- Initially created a website
- Realized that implementing a data exploration tool is not the right goal to gear toward
- Use Tableau to quickly find the patterns of our interest from the dataset
- Redesign interactive visualization to better communicate our findings to the readers when our conclusion becomes clear
Questions

- How might we explore inequality with such a dataset? Is analyzing geographic location, business, and change over time sufficient or are there other avenues we should explore?
- What other questions would be interesting to pursue with this dataset?
- What might be a better way to visualize the data that provides clearer direction given that there are many variables included in this dataset?
Exploring the Random-Walk: An Interactive Investment Game to Teach People Not to Time the Market

Josh Morris
CS 448B
Its difficult for individuals (and experts!) to consistently outperform the stock market by timing the market (Malkiel 1973; Barber & Odean 2000)

...but when you tell people that, they still don’t believe you (Fernandes et al 2014; Conversations with friends and family)
So instead of telling them... Show them!
Too complicated and doesn’t teach strategy to be passive
Here we go!

Your account has a balance of $10,548.

Sell!

Since 1978, there have only been 126 weeks where a portfolio matching the S&P 500 closed lower 10 years later. Buying and holding the S&P
My Solution

Allow users to:
• Choose market (index, specific stock, bond fund, etc.)
• Choose length of time (1 year? 10 years?)
• Choose speed of application (1 week per sec? 10 weeks per sec?)
• Buy and sell as many times as they want

Make Goal:
• Do your best!

Present:
• Overall performance and performance relative to market for all completed "trials" to show danger of timing in the long-run

Mockup using d3.js
Plan

Use HTML/CSS:
• To allow user input for different parameters

Use Javascript:
• To process stock data from (an API?) when a user selects it

Use d3.js:
• To make line graphs “move”
• To display buys and sells
• To display statistics at end of each “trial”
Feedback

• What parameters should I allow users to choose?

• Advice for getting stock data “on the fly”? Alternatively, I could get data ahead of time and allow for a limited choice set.

• Should I include a dotted line to indicate current performance alongside the market performance?
Analyzing Hollywood’s Attempts to Appeal to an International Audience with Foreign Actors

Sharon Chen & Da Eun Kim
Hollywood & Its International Audience

“Worldwide ticket sales reached a record-breaking $38.3 billion [in 2015]. More than 70% of the film industry’s box office is generated overseas.”

-LA Times
China is acting blatantly protectionist in an industry that’s hugely important to the US

- The Hollywood blackout is implemented during limited periods of the year in China to turn the spotlight on domestically-produced films
- This year’s annual summer blackout coincided with the commemoration of the founding of the People’s Liberation Army on August 1
Hollywood & Its International Audience

Doona Bae at 'Cloud Atlas' premiere

Australian premiere of Chris Hemsworth’s Thor: Ragnarok to fill entire Robina cinema complex
Hollywood & Foreign Actors

Is there a significant relationship between the presence of a foreign actor in a Hollywood movie and its success in the foreign actor’s nation?

Are there thresholds such as actor’s screen time/amount of lines that influence box office success?

How different is the success internationally versus domestically for Hollywood films featuring a foreign actor?
Prior Work: Movies

Shows us relationships between multiple dimensions such as genre, year, ratings, and budget

 Doesn’t delve into actor data
Prior Work: Movies and People in Them

Gender representation with great interactive elements
Prior Work: Hollywood’s Gender Divide

There are interactive visualizations around movies and the effects of different factors on the box office, but those visualizations usually only look at gender.
Hollywood & Foreign Actors: Data

Box office (domestic, international, actor’s homeland)

Actors and their nationalities

Screen time

Lines

Plot influence
Current Progress: General Visualizations

Heat map of what countries Hollywood’s foreign actors and actresses are from
Is there a significant relationship between the presence of a foreign actor in a Hollywood movie and its success in the foreign actor’s nation?
Are there thresholds such as an actor’s screen time/amount of lines that influence box office success?
How different is the success internationally versus domestically for Hollywood films featuring a foreign actor?
Design & Implementation Issues

Collecting all the necessary data may be challenging: finding a tool or website that is able to count screen time and number of lines by actor/actress.

Presenting and visualizing our data in a way that answers our questions.
Seeking Feedback

What are other questions around this topic that we can visualize?

Are there any suggestions as to how we can visualization our current question differently?
Getting More from TimeSearcher: Filtering Stocks by Price, Rate of Change, and Momentum

Amy Chen
CS 448B
The Problem: Timesearcher has Clear Limitations

Prices boxes can be too broad

Prices boxes can be too narrow

Rate of change “angle” queries can be too narrow

And momentum isn’t there at all
What problems does this pose?

PROBLEM:

Rate of Change in TimeSearcher is by angle over a period of time. This relies on precise ROC as well as a precise time to have it over, which provides fewer options while exploring stocks.

PROBLEM:

According to Investopedia, momentum “Indicates Stock Price Strength.”

TimeSearcher has no tool for measuring by momentum.
Prior Work: TimeSearcher 1

DIFFERENCES

TimeSearcher 1 has the aforementioned issues

GOAL

The fluidity of interaction of TimeSearcher 1 with a broader filtering for Rate of Change and Momentum for better exploration of data.
Proposed Solution

3 charts with Timebox Filters

“Flip card” style of switching between

Transparent boundaries to indicate filters on other tabs

Additional ability to compare velocity and momentum of stocks
Questions / Comments

“Flip card” style of switching between

3 charts with Timebox Filters

Additional ability to compare velocity and momentum of stocks

QUESTIONS

Does employing a box instead of an angle take away from the intuitive meaning of the interaction?

Does the layering make sense? (Should I actually just put three connected timebox-filtered charts next to each other?)

Concerns about design? Additional functionality you’d like to have?

Transparent boundaries to indicate filters on other tabs
Visualizing Structured Hypotheses in the Gene Ontology with Focus-and-Context

Junjie (Jason) Zhu and Qian Zhao
CS448B Visualization: Final Project Proposal
Gene Ontology (GO)

- Includes a collection of over 40,000 biological concepts
- Has been used to “annotate” gene functions in over 100,000 studies
- Hierarchically organizes the concepts in a directed acyclic graph (DAG)

[Diagram showing a directed acyclic graph with nodes labeled as cellular_component, biological_process, molecular_function, pigment metabolic process during pigmentation, pigmentation during development, negative regulation of pigmentation during development, etc.]

http://www.geneontology.org
Hypothesis Testing on a DAG

- Suppose a scientist discovers multiple disrupted genes in a cohort
- She would need to investigate every single node in the DAG

![Diagram of a DAG with multiple hypotheses and nodes](image)

Associated genes: GENE1, GENE2, ...
Test if the genes are overrepresented
Calculate a p-value for this hypothesis

[Ramdas et al, 2017]
https://en.wikipedia.org/wiki/P-value
Visualization Challenge

• Multiple biological concepts may be revealed.
• Researchers are interested in understanding:
  • the structures of the multiple discoveries (rejected hypothesis);
  • and contexts of these discoveries (parents, children, depth...).

• There are ~30,000 nodes in this DAG
• We only want to look at a small set of discoveries in detail (- focus),
  • but we also want to see what structures can be revealed (- context)
Prior Art and Existing Work

[Ramdas et al, 2017]
Prior Art and Existing Work

Cytoscape [Shannon et al, 2003]
Prior Art and Existing Work

- Lack of context to display how concepts are discovered
- Balance the layout of edges and nodes
- Rigid display and details that are difficult to interpret
Our Proposed Solution

• Interface with hypothesis testing algorithms

• Render discoveries in hierarchical graphs (and auxiliary information)

• Interactive focus-and-context graph visualization in d3.js

• Display summary statistics/visualizations of the remaining DAG
Our Proposed Solution

**Inspiration: Table Lens**

[Table Lens Image]

**Full Graph Summary Statistics**
(Representation)

**Sub-graph exploration**
(Interaction)
Current Progress

Gene Ontology Database

Information retrieval from database (query, download, local cache)

Graph-based algorithms (path search, hypothesis testing, layouts)

Flask server (data communication, cache, instructions for rendering)

Webpage: d3, jquery (graph and statistics rendering, interaction)
Thank you
Rich Music Playback Slider

CS448B Project Progress
Jianqing Yang
Introduction

The typical music playback slider doesn’t communicate much information to users:
Introduction

Can we get more out of such a ubiquitous widget?

Primary use cases for general music enthusiasts:
  ● Faster navigation to specific points of interest in a song
    ○ For real-life social sharing
    ○ For sampling of new music
  ● Instead of having to randomly scrub through songs

Bonus use cases:
  ● Music browsing and discovery
  ● Arranging playlists visually
Relevant Prior Work

Commercial: Soundcloud

- Adds binned amplitude information and time-tagged social comments
  - Amplitude bins only helpful for navigation for songs with very distinct dynamics
  - Most social comments not informative
Relevant Prior Work

Academic: Timbregrams

Color mapping based on automatic audio feature extraction

- Color band differentiation too fine to determine broad musical structures
- Color mapping is relative to collection of music processed, not independent to each song

Progress

Using Spotify web API:

- Detailed musical features for each track, e.g.
  - Sections, tempo, key
- Overall characteristic scores for each track, e.g.
  - “Danceability”
  - “Instrumentalness”
  - “Liveness”
- Access to enormous music collection
Initial Ideas

Blocks: musical segments
Block lightness / saturation: key (mood), pitch range
Block texture: tempo

Lines: complexity, stability

Overall hues: whole-track values e.g. “liveness”, etc

Loudness / energy

Progress slider

Track time
Your Feedback Please!

1. Do you have other ways of navigating your music tracks?

2. Balance of simplicity versus richness of encodings?

3. Importance of precision of representation (apart from time)?

4. Importance of distinctive overall look for each track?
A tool for visualizing and understanding orchestral scores

Diego Hernandez | CS448B
Figure 3: Examples of visualization of classical compositions.
Questions

- Match by instrument or by note?
- Useful to visualize the physical layout of an orchestra?

http://www.sammamishsymphony.org/Musicians/Seating.php
Interactive Visualizations of Circuit Structure

Ross Daly and Leonard Truong
“Visualization of circuits is an important research area in electronic design automation. Locating errors in a large design may require a high-quality graphical representation of a circuit that allows humans to understand it.”

Problem Description

Hardware designers use textual languages to specify circuit structure

Hardware compiler writers construct, analyze, and manipulate an intermediate representation of a circuit as a graph

Visualizations of the intermediate representation, particularly of the differences across compiler transformations, would be a key productivity tool for both compiler writers and designers.

Prior work [1] has explored the placement and routing of circuits for visualization, we plan to extend this approach to support interactive recompilation of layout when switching between multiple views.

Visualizing the entire design becomes unmanageable as the number of nodes in the graph increases
Extending Prior Work

- Existing graph visualization systems lack support for:
  - Hierarchy
  - Multiple, domain specific views for hardware (undirected/directed graph, DAG, pipeline stages)
Design Goals

• Interactivity using D3
  • Realtime traversal of the levels of hierarchy
  • Switch between different views of the same design
  • Selection/highlighting of graph nodes and edges
  • Should promote fast visual search of the design by applying visualizing techniques (e.g. legends, brushing)
  • Facilitate visualization of graph transformations (compiler passes that manipulate designs)
Ideal Graph Layout Customization

- Statically constrain the placement of certain nodes
- For each node: constrain the locations of each IO port
- Optimization/tuning to explore the tradeoffs of edge routing algorithms (is there a way to formalize this tradeoff, perhaps from navigation?)
  - Number of edges crossing, total path length, distance between nodes
- Edge attributes
  - Path heuristic: align to grid, smooth curve, direct node <-> node
  - Visualizing intersections: ‘X’ vs semi-circle
  - Color, thickness, etc…
Prototype using D3’s Standard Library
Force Graph Issues

• Have to encode layout constraints as physics relationships (charge, etc…)

• Doesn’t encode direction of edges or port locations

• Hard to believe a physics based approach will yield the best results
cola.js

Constraint Based Layout for D3
cola.js
Constraint Based Layout for D3

• Pros:
  • Encode constraints based on domain knowledge of graph qualities
  • Drop in replacement for D3

• Cons:
  • Young project that lacks documentation and examples
  • Has anyone used this before?
Summary

• Placement of circuit elements and routing of connections is the key underlying problem, we hope to leverage existing graph layout and routing techniques.

• Any library suggestions?

• Is D3 the right tool for interactive graph layouts? The standard library seems to be lacking

• Anything applicable from other domains such as maps and navigation?
Ranked-Data Ribbon Visualization

Greg Ramel
Description

- Interested in Billboard Hot 100 data and songs’ paths over time
- Want to create a tool for a general time-series data visualization with stacked ribbons
  - Discrete ranked data
- Possible datasets to explore, very few of which have been visualized:
  - Music charts like Billboard (initial exploration)
  - Box office (prior work)
  - Forbes Billionaire lists
  - Sports: ATP Tennis Rankings, College Football (prior work), major league sports standings, Olympic medal counts
  - New York Times Bestseller list
Description

- Software artifact would seek to create a pipeline to allow people to go from ranked data file to customizable, interactive ribbon visualization
  - Customize look and feel, data subset, filtering
- Surprisingly little prior work in this space, or even to capture ranking over time of any of the datasets I outlined
Prior Work - Box Office
Prior Work - Box Office

- Main point of inspiration is Zach Beane’s Box Office Charts
- Weighted by weekend gross; hue for entry time week-by-week
- Static aside from mouseover providing title and gross; no filtering or customization
- Rank at box office is mostly decreasing - need additional considerations for frequent and more drastic shifts, as well as late reentry
- Draws on Byron and Wattenberg’s streamgraph
Prior Work - College Football Ranking
Prior Work - College Football Ranking

- Same general concept of week-over-week ranked data, presented in a more traditional line plot
- Why not just use line plots?
- Doesn't allow you to capture relative weight and transition between different states
- Harder to track overall path in general, particularly with overlap
- Labeling not possible on path itself
Current Progress

- Have existing, rudimentary visualization of Billboard Top 10 - limited interactivity; slow to render so no filtering possible; bulky data (json)
Current Progress

- Initial time spent on project has been beginning to update old implementation with an eye toward generalized functionality - Billboard data is hardcoded
- Exploring datasets, web scraping to get a sense of how to allow users to plug in custom .csv files and column labels
- Flow:
  - Select file; select columns for axes (timesteps, rank); select columns for attached data (label, mouseover)
  - Customize look and feel (color sequence; custom icons or graphics; weighted or unweighted)
  - Customize filtering options
- Stretch goal: smart scraper to automatically pull rank data in a usable format given an initial URL and limited user pointers
Questions

- What features/filters would you want to see in a tool like this?
- What datasets come to mind that you would want to explore?
- Are there other examples of visualizations in this vein that I missed?
Tracksplore

Na He Jeon, Mathieu Rolfo, Karen Wang
Problem

- **Goal:** compare tracks in CS and SymSys
- Visualize the similarities and differences between tracks
## Current State of Exploring Tracks

| Track Electives: at least three additional courses selected from the Areas and lists above, general CS electives, or the following:  
| CS 238 | Decision Making under Uncertainty  
| CS 275 | Translational Bioinformatics  
| CS 326 | Topics in Advanced Robotic Manipulation  
| CS 334A | Convex Optimization I  
| or EE 364A | Convex Optimization I  
| CS 428 | Computation and cognition: the probabilistic approach  
| EE 278 | Introduction to Statistical Signal Processing  
| EE 364B | Convex Optimization II  
| ECON 285 | Game Theory and Economic Applications  
| M&AE 252 | Decision Analysis I: Foundations of Decision Analysis  
| M&AE 352 | Decision Analysis II: Professional Decision Analysis  
| M&AE 355 | Influence Diagrams and Probabilistic Networks  
| PHIL 152 | Computability and Logic  
| PSYCH 202 | Cognitive Neuroscience  
| PSYCH 204A | Human Neuroimaging Methods  
| PSYCH 204B | Computational Neuroimaging: Methods & Analyses  
| PSYCH 209 | Neural Network Models of Cognition: Principles and Applications  
Current State of Exploring Tracks
Prior Work: Carta
Prior Work: Edusalsa Explore
Prior Work: Edusalsa Journeys
Improvements Needed

- Program sheets and ExploreDegrees list requirements for each track, but lack features to compare them across tracks.
- Third-party software tools don’t provide any information on tracks.
Idea: Forward-Sampling
Track Permutations
Open Questions

Do you think the issue of comparing tracks is a compelling one?

If not, are there related questions students (you) want the answer to?

Do you have any suggestions for computing similarity between tracks?

Do you have any suggestions for visualizing similarity between tracks?
The Effect of Animation and Small Multiples in Dynamic Graphs Surfaced on Mobile Devices

Albert Feng
Pakapark Bhumiwat
Mobile Devices are **small**…

…but Visualizations for Time Dependent Data **Takes Up Space**.
Mobile Visualization Research

Time Dependent Data Visualization Research

Visualization Technique Comparison Research
Mobile Viz Experiments

How does X relate to Y?

A  B  C  D

Next Question
Feedback and Questions

How can we make sure the test visualizations we create as “good visualizations” so that the test results we get from them are actually valid?

What other techniques should we explore in showing time dependent data?

What are some interesting time dependent data sets that we should consider?
Interactive Bayesian Network Visualization with D3 for Non-domain experts

Jesik Min
Bayesian Network
Fig. 2. Causal network in our fictitious example with short node names (greek letters) to be used in the text. Each node has two possible states representing responses ‘yes’ and ‘no’. Direction of causality is from top to bottom.
Why Bayesian Network?

1. Graphical models
2. Causal relationships
3. Handle uncertainty

[Margaritis, 03]
Problem

People without domain knowledge should also be able to construct a reasonable Bayesian network.
Previous Work


Chih-Hung Chiang, Patrick Shaughnessy, Gary Livingston, and Georges Grinstein, Clifford Champion and Charles Elkan, Visualizing the Consequences of Evidence, in Bayesian Networks, arXiv:1707.00791 [cs.AI], 2017


→ Standalone application

→ No integration
Solution

Implement a D3 application that helps non-domain experts construct and visualize Bayesian networks.
Workflow

User enters CSV data

Preprocess the data
1) Provide nodes in pie chart
2) Populate the info panel

Suggest next action
1) Suggest an edge for the selected node
2) Displays conditional probability table for the selected node
Feature Implemented
<table>
<thead>
<tr>
<th>Date</th>
<th>Features to implement</th>
</tr>
</thead>
<tbody>
<tr>
<td>11/7~11/10</td>
<td>Make each node clickable. Enable user to add edges between nodes.</td>
</tr>
<tr>
<td>11/11~11/18</td>
<td>Suggest the conditional probability table for the selected set of nodes.</td>
</tr>
<tr>
<td>11/19~11/22</td>
<td>Suggest an edge from the selected node based on the dataset (e.g. by computing Bayesian score).</td>
</tr>
<tr>
<td>11/23~11/27</td>
<td>Implement information panel on the side.</td>
</tr>
<tr>
<td></td>
<td>Enable user to upload the any CSV file.</td>
</tr>
<tr>
<td>11/27~11/30</td>
<td>Improve encoding schemes (e.g. color scheme, shape, text position).</td>
</tr>
<tr>
<td>12/1~12/5</td>
<td>Finish up code and make poster. Make demo for the live presentation.</td>
</tr>
</tbody>
</table>
Questions?