Visualizing Incongruity in Text: Visualization Strategies for Detection and Modeling of Garden Path Jokes

Category: Research

Abstract—The goal of this project was to investigate the use of visualization as an approach to modeling humor within text. In particular, we developed algorithmic and automated approaches to visualizing and detecting shifts in interpretation as intelligent agents parse meaning from garden path jokes. Garden path jokes can occur when a readers initial interpretation of an ambiguous text turns out to be incorrect, leading them down the wrong path to a semantic dead end. Given new information, semantic incongruities arise that require resolution, often triggering a humorous response. This is a work of visual text mining, that is the use of text visualization as a means of detecting patterns and features related to various text based phenomena such as humor within text that can be useful for modeling and classification. In this paper we describe three successful approaches to text visualization enabling the identification of features relevant to the classification and detection of garden path jokes, a particular joke subtype. These are the use of paired collocated coordinate plotting, heat maps, and two dimensional Boolean plots. The methodology and tools resulting from this project offer a new approach to modeling humor in text, as well as many other incongruity-based linguistic phenomena.

Index Terms—Computational humor, natural language understanding, visual data mining, incongruity modeling.

1 Introduction

Text visualization and discovering patterns in text are growing fields with multiple challenges and relatively unexplored subjects including their integration. This paper focuses on visual text mining which intends to provide deep fusion of pattern discovery and visualization for text exploration. Visual text mining is subfield of visual data mining [12] [4]. In particular we focus on visulaization for modeling and detecting incongruities within garden path jokes, a particular joke subtype. With many joke types semantic incongruities form while reading a text when a reader established two or more interpretations that that oppose or come into conflict with eachother. Some form of resolution, or meaning disambiguation, process is then required. Garden path jokes display a particular pattern of interpretation over time where the reader initially establishes one interpretation given the first part of the joke but given additional information must discard the first interpretation and establish a second [1]. In the field of Computational Humor there is no established theoretical foundation for modeling incongruities within text using computational systems working with language processing capabilities and semantics and many recent attempts at humor detection and generation use different theoretical frameworks that attempt to describe the nature of incongruity and resolution that occurs as we read jokes. In this paper we will describe various visualization based approaches developed that successfully allow for modeling and detection of humor in text by visualizing garden path jokes and non jokes in such a way as to enable the discovery of features relevant humor within text.

The visualization approaches described in this paper represent and work with meaning at a heuristic level through the use of vector space based semantics where the meaning of a word is encoded as a vector describing the relationship with, or in our case the frequency at which the word co-occurs with, other words. Connections between meanings can then be established through calculation of correlation using vector based mathematics. This paper will describe the methods used to establish some meaning using text data resulting from various web searches representing different contexts. Next the data, set simple garden path jokes and non jokes of similar form, will be described. Finally this paper will describe the three visualization approaches used along with the resulting visualizations generated using the data set. The first visualization plots meaning correlation scores as coordinates over time given the two parts of a garden path jokes , which lets us visually see meaning correlation shifts over time with respect different meanings for some ambiguous word. Next we describe an approach based on heat maps based on the differences of meaning correlation scores. Finally this paper describes an approach that allows us to visualize a model space where jokes and non jokes can be plotted on a Boolean space using features generated from the comparison of correlation scores.

Beyond detection and visualization of humor in text the visualization of incongruities can be useful in many text analysis tasks. Incongruities form within Paradigms when there are opposing viewpoints attempting to explain some phenomena, each giving their own evidence. One could visualize the formation and resolution of incongruities within academic document sets as paradigms evolve using the strategies outlined in this paper.. Another example among many is that of hilling, or the leaving of false product reviews. When a bad product has a group of people intentionally leaving good reviews to deceive customers incongruities of sentiment form that can be visualized. Overall our approach offers many useful tools for approaching incongruity in text beyond humor.

2 Related Work

Both Computational Humor and Text Visualization as fields have seen extensive activity lately but tend to work on separate topics. Computational Humor deals a lot with the modeling and detection of incongruities within text and many attempts have recently been made attempting to detect or generate jokes using computers [6] [8] [14] [5] [10] [13] but no attempt focused on visualization has been made. On the other hand people working on Text Visualization tend to focus on other topics such as identifying the central topic within a text. To our knowledge we are the first to visualize incongruities within text.

3 Incongruity Resolution Theory of Humor and Garden Path Jokes

Two fish are in a tank. One looks to the other and asks: How do you drive this thing?

One predominant theory of verbal humor states that humor is triggered by the detection and resolution of semantic incongruities during semantic parsing [11] [9]. The dictionary defines 'incongruous' as lacking harmony of parts, discordant, or inconsistent in nature. Given text that is read, semantic incongruities form when a reader's interpretation of some concept conflicts with other possible interpretations as the remainder of the text is read. The fishtank joke displays an incongruity. When a reader initially reads it in the tank is interpreted as an aquarium, but given additional information indicating that it might actually be a vehicle an incongruity forms and resolution is required.

The fishtank joke is an example of a garden path joke [1]. A garden path joke is a particular subtype of joke that describes a certain pattern of incongruity and resolution. With a garden path joke a reader establishes some interpretation A as they read the first part of a joke, the setup, but given new evidence included in the second part, the punchline, they must discard this interpretation and establish a new interpretation B. It is named after the garden path metaphor where one is lead down a garden path, that is misled, and coming to a semantic dead end has to step back and take some other path. To model this we look at how meaning correlation changes over time given parts of a garden path joke with respect to some ambiguous term. In the next section we will discuss establishing of meaning correlation.

4 Establishing Meanings and Meaning Correlations

We chose to use a vector representation of meaning [7] based on the frequency at which other words co-occur with some given concept. While heuristic in nature as additional information is needed to understand further the nature of the relationship of words that co-occur together often, it works as a good enough heuristic for the task of joke detection. These vectors can be easily extracted from corpuses as opposed to building formal ontologies by hand. In this section we will describe establishing, representing, and calculating correlation between different meanings using a vector representation of meaning.

4.1 Establishing semantic vectors using documents retrieved via web search

Consider some ambiguous word A with a number of possible meanings $AM_1...AM_n$. We establish vectors that contain frequencies at which other words co-occur with $AM_1...AM_x$, that is we establish vectors representing the various opposing meanings of A . To do this we use a web search approach. For each meaning AM_x we established a set of disambiguating keywords $K(AM_x)$ that when used as a search will pull up texts relevant to the meaning AM_x and not other meanings. These keywords can be hand-chosen but can be established using a knowledge resource such as wordnet which includes keywords associated with different meanings of various words, such as hypernyms and hyponyms. Using $K(AM_x)$ as a query for a search engine, we retrieve the top n documents which is designated $D(K(AM_x), j)$. We used j=500 results for the experiments outlined in this paper. Finally we computed frequencies of all words W co-occurring within distance k=3 of A given the document set $D(K(AM_x), j)$. We designate this $F(A, D(K(AM_x), j)))$ where F is a function that returns a vector of word co-occurrence frequencies representing meaning AMx. The frequencies are ordered by the lexicographic order of the words. This results in $F(A, D(K(AM_1), j)))...F(A, D(K(AM_n), j)))$.

In a similar fashion we established semantics for the ambiguous word A given the different parts $P_1...P_m$. That is $AP_1...AP_m$. The search query is in this case is just the part of text itself. This results in $F(A, D(P_1, j))...F(A, D(P_m, j))$.

4.2 Calculating correlation coefficients

We are interested in how the meaning of A associated given a search for some phrase Px correlates with the meaning of A given searches for $AM_1...AM_n$. This allows us to establish which word sense of A is closest to the A as seen in similar contexts given some part of text. When this is calculated for multiple parts of text this allows us to identify any incongruous meaning correlation patterns where two meanings correlate high given two or more section of text. We compute the correlation coefficient: $C_{xy} = C(F(A, D(P_x, j)), F(A, D(K(AM_y), j)))$ where C is a function that returns correlation value given two vectors. In our case C calculates Pearson's R correlation coefficient using the frequencies of words found in one or both vectors.

All of the jokes in our data set involve are two part jokes involving two meanings. Given two meanings of some ambiguous word A and some statement with parts P_1 and P_2 that refer to A, we calculate the following correlation scores:

Meaning correlations given P_1:

$C_{1x} = C(F(A, D(P_1, j)), F(A, D(K(AM_x), j)))$ is a correlation of meaning AP_1 with meaning AM_x,

$C_{1y} = C(F(A, D(P_1, j)), F(A, D(K(AM_y), j)))$ is a correlation of meaning AP_1 with meaning AM_y,

Meaning correlations given P_2:

$C_{2x} = C(F(A, D(P_2, j)), F(A, D(K(AM_x), j)))$ is a correlation of meaning AP_2 with meaning AM_x,

$C_{2y} = C(F(A, D(P_2, j)), F(A, D(K(AM_y), j)))$ is a correlation of meaning AP_2 with meaning AM_y.

4.3 Building Features from correlation coefficient differences

Finally we calculate differences between the correlation coefficients which are useful for joke classification as they can tell us a lot regarding correlation movement patterns. For example The difference between C_{1x} and C_{1y} tells us which meaning has greater correlation given P1. If $C_{1x} - C_{1y} > 0$ then this meaning meaning x is greater than meaning y given the first part of some text. On the other hand the difference between C_{1x} and C_{2x} tells us if a correlation coefficient has increased or decreased given part on or part two of some text. If $C_{1x} - C_{2X} > 0$ then the correlation of meaning x has decreased as the text is read in. We calculate the differences between C_{1x}, C_{1y}, C_{2x}, and C_{2y}.

Finally we define four Boolean variables x1-x4 using the relations by comparing the correlation scores:

$x1 = 1$ If $C_{1x} > C_{1y}$, else $x1 = 0$
1 means Meaning X is greater than meaning Y given P_1

$x2 = 1$ If $C_{1x} > C_{2x}$, else $x2 = 0$
1 means Meaning X decreased going from P_1 to P_2

$x3 = 1$ If $C_{1y} < C_{2y}$, else $x3 = 0$
1 means Meaning Y increased going from P_1 to P_2

$x4 = 1$ If $C_{2x} < C_{2y}$, else $x4 = 0$
1 means Meaning Y is greater than meaning X given P_2.

5 Example

Take a two part garden path joke J with the parts P_1 = fish in tank and P_2 = they drive the tank that contains the ambiguous word tank.. Let $tankM_x$ and $tankM_y$ be the two meanings invoked at different points while reading J, that of an aquarium and that of a vehicle..

P_1=fish in a tank.
P_2=drives the tank
$K(tankM_x)$=[aquarium, tank]
$K(tankM_y)$=[vehicle, panzer, tank]

This is a distilled example of the fishtank joke presented in the introduction. In order to concentrate on the issue at hand, i.e. modeling incongruity, we reduced many jokes to simplified form by cutting out text irrelevant to the actual joke.

We established vectors for the various meanings of tank using data from searches for $P_1, P_2, K(tankM_x), andK(tankM_y)$ and then calculate the correlation coefficients between these meaning vectors. We have potentially four different meanings of tank and want to check which word sense meanings correlate with the meanings associated with different phrase components.

Meaning correlation coefficients given P_1:

$C_{1x} = C(F(tank, D(fish\ in\ a\ tank, j)),$
$\quad F(tank, D([aquarium, tank], j))) = 0.824$

$C_{1y} = C(F(tank, D(fish\ in\ a\ tank, j)),$
$\quad F(;tank, D([vehicle, panzer, tank], j))) = 0.333$

Meaning coefficients given P_2:

$C_{2x} = C(F(tank, D(drives\ the\ tank, j)),$
$\quad F(tank, D([aquarium, tank], j))) = 0.389$

$C_{2y} = C(F(tank, D(drives\ the\ tank, j)),$

$$F(tank, D([vehicle, panzer, tank], j))) = 0.573$$

Hypothesis: Given that when reading a garden path joke there should be a switch in dominant meaning correlation coefficient. Given the first part correlation with meaning X should be greater and given the second part correlation with meaning Y should be greater.

6 Data Set Used in Visualizations

Two part jokes of garden path form containing lexical ambiguities were collected and put into a simple form by hand as we want to focus on modeling incongruities rather than deal with the nuances of natural language processing including coreference resolution, entity relation extraction. Ideally a system would take longer jokes and be able to algorithmically identify the parts of text involved with the incongruity, but in our case material not relevant to the interpretation of the lexical ambiguity was removed. Thus Two fish are in tank because fish in a tank. as the number of fish has little to do with the lexical ambiguity at involved in the incongruity we are attempted to model.

For each joke we created a non joke of similar form. It contains the same first part but a different non-humorous second part. In total we collected and processed 34 statements, 17 pairs of jokes and non jokes for this paper. Overall The following are some examples of jokes and non jokes contained in the data set.

Joke1:
P_1: "A fish in a tank"
P_2 "drives the tank.

NonJoke1:
P_1: "A fish in a tank"
P_2: "swim in the tank.

A = 'tank'

Meaning X search phrase: 'Aquarium tank
Meaning Y search phrase : Panzer tank

Joke2:
P_1: "The computer has a terminal..."
P_2 "A terminal illness.".

NonJoke2:
P_1: "The computer has a terminal..."
P_2: "A terminal and keyboard."

A = 'terminal'

Meaning X search phrase: 'terminal as in monitor'
Meaning Y search phrase: 'terminal as in cancer'

7 Visualization Approach 1. Collocated Paired Coordinates

Our first visualization uses a visualization technique known as collocated paired coordinates [2]. We plot the correlation coefficients given each part of text as coordinates. These coordinates are connected with arrows representing time. Thus meaning correlation patterns over time can be visualized. This allows for visualization of meaning correlation shifts given garden-path jokes.

Visualization Walkthrough:

For P_1 and P_2 we plot the meaning correlation coefficients given two opposing meanings for some ambiguous word with AM_x and AM_y as coordinates:

The X axis represents correlation with some meaning AM_x.
The Y axis represents correlation with some meaning AM_y

1 Plot a coordinate representing the meaning correlations given P_1, that is (C_{1x}, C_{1y}).

2 Plot a coordinate representing the meaning correlations given P_2, that is (C_{2x}, C_{2y}).

3 Connect via an arrow indicating time.

4 Color-code green if humorous, red if not, and black if unknown.

Fig. 1. Our data set of jokes and non jokes plotted as meaning correlation over time using collocated paired coordinates.

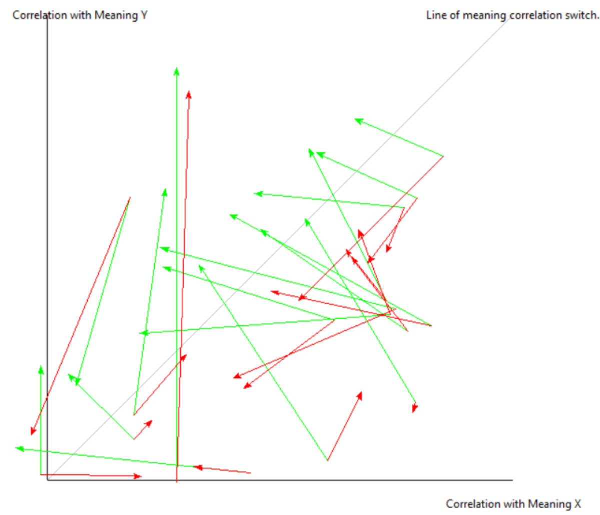
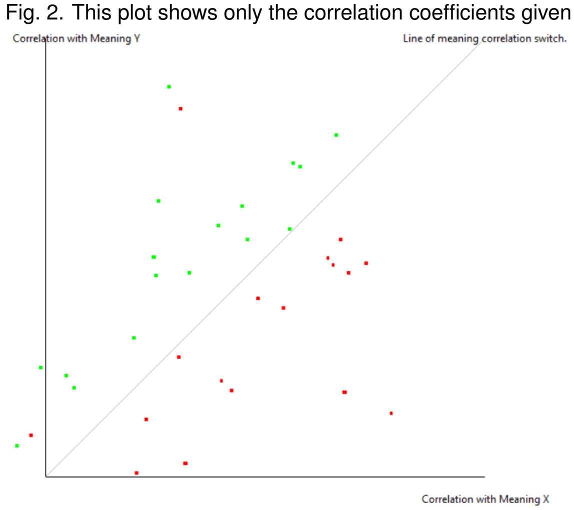


Fig. 2. This plot shows only the correlation coefficients given P_2.



7.1 Discussion

While there are some examples which fail to match the pattern, it is clear that most jokes involve a shift towards the second meaning given part two of the joke. Fig 1 shows this as the green arrows, representing jokes, moves from one axis to another while the red arrows move randomly. An analysis of the handful of cases that do not follow this pattern indicates explainable circumstances such as the web search returning irrelevant documents due to things like a poor choice in keywords. Fig 2 only looks at the meaning correlation coefficients given P_2 which clearly shows that there is correlation with meaning Y which opposes some meaning X that was initially established.

8 Visualization 2: Heat Maps

In the previous visualization we saw that there is a shift from one meaning correlation being higher than another to the opposite. To test this intuition we make use of heat maps based on differences in correlation coefficient values.

Visualization walkthrough:

1. Organize the correlation coefficient differencesy in a data frame along with classification of being a joke or not.

2. Color code the correlation score differences based on value.

3. Sort the rows into groups by classification, that is into two groups of joke and non joke.

4. Identify regions of the heat map where there is a distinguishable difference between the joke and non joke classes. With this approach we can identify potential features that distinguish jokes from non-jokes, assisting in model discovery.

Fig. 3. This heat map color codes the differences in correlation coefficients using three colors.

name	C_{1x}-C_{1y}	C_{1x}-C_{2x}	C_{1x}-C_{2y}	C_{1y}-C_{2x}	C_{1y}-C_{2y}	C_{2x}-C_{2y}	class
webJoke1	0.623	0.239	0.227	-0.384	-0.396	-0.012	joke
soapJoke1	0.453	0.318	0.234	-0.135	-0.219	-0.084	joke
mouseJoke1	0.152	0.191	0.072	0.039	-0.08	-0.119	joke
terminalJoke1	0.187	0.217	0.088	0.03	-0.099	-0.129	joke
framedJoke1	0.42	0.504	0.366	0.084	-0.054	-0.138	joke
balanceJoke1	0.559	0.2766	0.136	-0.2824	-0.423	-0.1406	joke
freeJoke1	-0.432	0.116	-0.026	0.548	0.406	-0.142	joke
dogJoke1	0.18	0.324	0.15	0.144	-0.03	-0.174	joke
fishJoke1	0.491	0.435	0.251	-0.056	-0.24	-0.184	joke
chargeJoke1	0.096	0.1415	-0.044	0.0455	-0.14	-0.1855	joke
chopJoke1	0.273	0.369	0.158	0.096	-0.115	-0.211	joke
potatoJoke1	0.381	0.505	0.249	0.124	-0.132	-0.256	joke
houseJoke1	-0.026	-0.0004	-0.261	0.0256	-0.235	-0.2606	joke
bankJoke1	0.045	-0.069	-0.444	-0.114	-0.489	-0.375	joke
virusJoke1	0.343	0.167	0.014	-0.176	-0.329	-0.153	joke
wavesJoke1	0.379	0.537	0.419	0.158	0.04	-0.118	joke
catJoke1	0.282	-0.001	-0.611	-0.283	-0.893	-0.61	joke
webNonjoke1	0.623	0.008	0.647	-0.615	0.024	0.639	nonjoke
balanceNonJok	0.559	-0.076	0.409	-0.635	-0.15	0.485	nonjoke
framedNonjoke1	0.42	0.121	0.406	-0.299	-0.014	0.285	nonjoke
dogNonjoke1	0.18	0.041	0.281	-0.139	0.101	0.24	nonjoke
chopNonJoke1	0.273	0.197	0.421	-0.076	0.148	0.224	nonjoke
terminalNonjoke	0.187	0.107	0.329	-0.08	0.142	0.222	nonjoke
houseNonjoke1	-0.026	-0.218	-0.023	-0.192	0.003	0.195	nonjoke
potatoNonjoke1	0.381	0.351	0.53	-0.03	0.149	0.179	nonjoke
mouseNonjoke1	0.152	0.311	0.465	0.159	0.313	0.154	nonjoke
soapNonJoke1	0.453	0.135	0.276	-0.318	-0.177	0.141	nonjoke
chargeNonJoke	0.096	-0.041	0.054	-0.137	-0.042	0.095	nonjoke
fishNonJoke1	0.491	0.345	0.418	-0.146	-0.073	0.073	nonjoke
bankNonjoke1	0.045	-0.115	-0.088	-0.16	-0.133	0.027	nonjoke
freeNonjoke1	-0.432	0.213	0.083	0.645	0.515	-0.13	nonjoke
catNonjoke1	0.282	-0.027	-0.561	-0.309	-0.843	-0.534	nonjoke
wavesNonjoke1	0.378	0.067	0.195	-0.311	-0.183	0.128	nonjoke
virusNonjoke1	0.343	0.077	0.247	-0.266	-0.096	0.17	nonJoke

8.1 Discussion

While this heat map only uses three colors when color coding correlation coefficients by value, clearly we can identify areas where the joke data set differs from the non joke data set. Lets look at the column representing the difference between $C_{2x} - C_{2y}$ In Fig 3. This is the column indicating the difference between correlation with meaning X and meaning Y given meaning established in the second part of the joke. If this value is less than 0 then meaning Y is greater given P_2, if it is greater than 1 then meaning X remains dominant. While we already expected this to happen, the heat map would allow us to automatically identify this value as being a distinguishing feature between classes.

9 Visualization 3: Visualizing a model space using monotone Boolean chain visualizations

In he last viusalization we use a two dimensional representation of Boolean space based on the plotting of chains of monotonically increasing Boolean vectors [3] to visualize the difference between garden path jokes and non jokes. This is one way of representing Boolean space in such a way as to preserve some aspects of the topology. Vectors are arranged according to their norm, with the all true Boolean vector at one end of the plot and the all false vector at the other. From the arrangement chains of vectors can where monotonicity is preserved can be observed. Each succeeding vector in the chain contains the same Boolean value as the last but with an additional bit set to true. In our case the vectors represent features and each chain describes an incremental development of a feature set, where the first vector represents no features each succeeding vector establishes a the presence of all feature. This allows the visualization of a model spaceI. In our case these features describe whether correlation values for given meaning move up or down over time, as well as which one is greater given the parts of a statement. In this section we describe the approach and present the results for visualizing a model space based on Boolean features $x_1 \dots x_4$ generated through comparing meaning correlation coefficients as discussed in section 4.3.

Visualization Algorithm:

1. For each joke or non joke establish a vector of Boolean features as described in section 4.3.
2. Establish and visualize a 2D boolean space representation as described in [3]
3. Plot vectors established each joke or non-joke as a dot on the boolean plot.
4. Color code the dot as green if humorous, red if not humorous, and black if unknown.
Resulting Visualization using our data set:

Fig. 4. Our data set of jokes and non jokes plotted as Boolean vectors

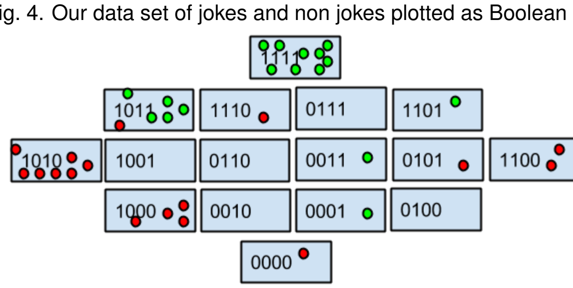
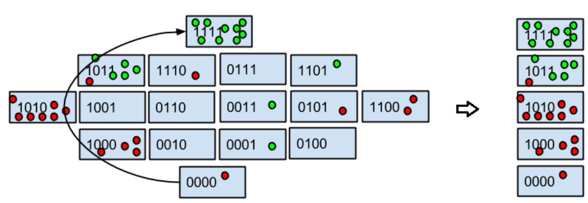


Fig. 5. Here one chain of monotonically increasing Boolean vectors is isolated to establish a border between humorous and nonhumorous examples in terms of features.



9.1 Discussion

This visualization shows we can visualize a feature space representing different models (a model space) in such a way as to identify valid models. Areas of the model space with a large number of jokes but no non jokes ploted represent configurations of features which may form a valid model for that phenomena. By looking at a chains of Boolean vectors where features are added one at a time, we can see a border between the classes that we are trying to model. Looking at one prominent chain, the one with the most examples plotted that as outlined in Fig 5, we can see a border between classes. Let us analyze this chain. First note that other than the all 0 vector (no features) all the examples in the chain indicate that meaning X is higher given part one. This is the meaning of x1 = 1. Now let us look at the border, that is what feature when added turns these statements into jokes. Looking at Fig 5, we can see that the feature is x4. Looking back at section 4.3 we see that x4=1 indicates that meaning Y is larger given part two of the joke. Overall this visualization shows us that given our data set garden path jokes are ones where meaning X is larger given the first part of a joke and meaning Y is larger given the second, matching our intuition and the theory of garden path humor.

10 Conclusion

The results from this study show that visualization can be used as a valid strategy for approaching the modeling and detection of humor within text. This paper presented three approaches that were all successful in enabling a person to identify key features that distinguish humorous and non humorous garden path jokes. One future direction is to use these visualization techniques on other joke types to see what they would look like in terms of patterns of meaning correlation over time. Additionally these techniques can also be used to visualize many other forms of incongruity within texts such as product review sets or academic document sets. They allow for plotting of large data sets at once, letting us visualize big data in terms of natural lanuage texts. Overall the results are promising.

References

[1] M. Dynel. Garden paths, red lights and crossroads. *Israeli Journal of Humor Research*, 1(1):289320, 2012.

[2] B. Kovalerchuk. Visualization of multidimensional data with collocated paired coordinates and general line coordinates. In *IS&T/SPIE Electronic Imaging*, pages 90170I–90170I. International Society for Optics and Photonics, 2013.

[3] B. Kovalerchuk and F. Delizy. Visual data mining using monotone boolean functions. In *Visual and Spatial Analysis*, pages 387–406. Springer, 2004.

[4] B. Kovalerchuk and J. Schwing. Visual and spatial analysis. *Advances In Data Mining, Reasoning, And Problem Solving*, 2005.

[5] I. Labutov and H. Lipson. Humor as circuits in semantic networks. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers-Volume 2*, pages 150–155. Association for Computational Linguistics, 2012.

[6] R. Mihalcea, C. Strapparava, and S. Pulman. Computational models for incongruity detection in humour. In *Computational Linguistics and Intelligent Text Processing*, pages 364–374. Springer, 2010.

[7] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean. Distributed representations of words and phrases and their compositionality. In C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Weinberger, editors, *Advances in Neural Information Processing Systems 26*, pages 3111–3119. Curran Associates, Inc., 2013.

[8] S. Petrovic and D. Matthews. Unsupervised joke generation from big data. In *ACL (2)*, pages 228–232. Citeseer, 2013.

[9] G. Ritchie. Developing the incongruity-resolution theory. 1999.

[10] G. Ritchie. The jape riddle generator: technical specification. *Institute for Communicating and Collaborative Systems*, 2003.

[11] T. R. Schultz and F. Horibe. Development of the appreciation of verbal jokes. *Developmental Psychology*, 10(1):13, 1974.

[12] S. J. Simoff, M. H. Böhlen, and A. Mazeika. Visual data mining: An introduction and overview. In *Visual Data Mining*, pages 1–12. Springer, 2008.

[13] J. M. Taylor and L. J. Mazlack. An investigation into computational recognition of children's jokes. In *PROCEEDINGS OF THE NATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE*, volume 22, page 1904. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2007.

[14] A. Valitutti, H. Toivonen, A. Doucet, and J. M. Toivanen. " let everything turn well in your wife": Generation of adult humor using lexical constraints. In *ACL (2)*, pages 243–248, 2013.