國立台灣大學電機資訊學院資訊網路與多媒體研究所
碩士論文

Graduate Institute of Networking and Multimedia

College of Electrical Engineering and Computer Science

National Taiwan University

Master Thesis

群眾外包於幽默網路圖片解釋生成技術之研究

Crowd-sourced Explanations for Humorous Internet

Memes

林自均

Chi-Chin Lin

指導教授：許永眞 博士

Advisor: Jane Yung-jen Hsu, Ph.D.

中華民國 103 年 7 月

July, 2014

Acknowledgments

首先要感謝我的指導教授許永真老師，在我對於 "做研究" 這件事情毫無動力時，是老師給我完全的自由，來選擇喜歡的題目，就算這個題目又冷門又奇怪； 而同時也給我適當的提點，讓我不會迷失方向。 不管是在做研究，還是做人處世上，老師都給我非常多的鼓勵， 讓我的熱情可以從碩士開始，到現在要畢業了，都還是熱血沸騰。

然後要感謝 Janet ，她是我研究所的第二位老師， 不管是一開始找方向的時候，還是中間寫 paper 的時候，還是最後修改論文的時候， 她都不斷地提供超級一針見血，血流成河的協助， 如果沒有她專業的指點，我的研究所生涯不可能會這麼順利。

感謝 Dynamic 的組員們，哲瑋、 Jimmy 、 Zoe 、八五、泓其， 每個禮拜我們都一起開會，報告自己的進度， 互相鼓勵，也互相講一些有的沒的鬼話， 這個小組的會議是我加入 iAgent 之後，感到最自在，最快樂的會議。

還有常和我討論研究方向與生活瑣事的大家，怡婷學姊、啟嘉、喬敬大大、宇傑、季恩、恰恰、 昱儒學姊、 Hans 、柏豪、翊瑄、孟穎、佑安， 有你們的陪伴，讓我在實驗室時都好像在家一樣，處處有溫情， 怡婷學姊就像萬能的媽媽，啟嘉就像有活力的表哥， 而其他人就像兄弟姊妹一樣。

還有要感謝我的真正的家人們，爸爸、媽媽、妹妹在台中，給我最有力的支持，而阿姨、姨丈、阿公、阿嬤、表弟妹在台北，給我最體貼的照顧，讓我可以專心做研究，非常感謝他們。

　　最後要感謝我的女朋友阿璞，我論文最重要的靈感，都是和她約會的時候突然想到的。聽起來好像我約會不專心，可是我覺得，也許人要在最開心的時候，想法才會最奔放。

Abstract

Humorous images can be seen in many social media websites. However, newcomers to these websites often have trouble fitting in because the community subculture is usually implicit. Among all the types of humorous images, Internet memes are relatively hard for newcomers to understand. In this work, we develop a system that leverages crowdsourcing techniques to generate explanations for memes. We claim that people who are not familiar with Internet meme subculture can still quickly pick up the gist of the memes by reading the explanations. Our template-based explanations illustrate the incongruity between normal situations and the punchlines in jokes. The explanations can be produced by completing the two proposed human task processes. Experimental results suggest that the explanations produced by our system greatly help newcomers to understand unfamiliar memes. For further research, it is possible to employ our explanation generation system to improve computational humanities.

Keywords: computational humor recognition, crowdsourcing, Internet memes

摘要

在許多的社群網站裡，幽默的圖片是很常見的。 但是，剛到這些網站的新手們通常都很難融入，因爲這種次文化通常都有著一些隱涵的資訊。 在幽默圖片的種類裡，網路模因對於新手們是相對比較難理解的。 在這一篇論文裡，我們開發了一套系統可以藉由群眾外包的技術，來產生網路模因的解釋。 我們宣稱只要看過我們產生的解釋之後，就算對於網路模因不熟悉的人， 也都可以很快地了解其中的笑點。我們模板式的解釋可以突顯出笑話裡正常情況和笑點句子的不和諧。這樣的解釋可以透過兩套人腦計算的流程來產生。 這裡的實驗結果顯示系統所產生的解釋可以很好的幫助新手來了解不熟悉的網路模因。 未來研究的方向可以考慮加入更多電腦計算的幫助，來改善計算人性這個領域的發展。

關鍵字：電腦幽默辨識、群眾外包、幽默網路圖片

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Recently, the subculture of humorous images has been a pervasive phenomenon on the Internet. Social media websites such as 9GAG[1], Reddit, and 4chan provide platforms on which people share their images and opinions. Experienced users relax on the sites by browsing images and leaving their comments. To these users, understanding the gist of the humorous images – that is, what makes it funny – is easy and entertaining.

However, for those who are not familiar with the subculture, it can be very difficult to understand why the images are funny. When newcomers seek explanations in discussion groups, the explanations given are usually very rare or implicit, because the discussion is often dominated by experienced users. Most newcomers have trouble fitting in such communities because of the lack of a desirable explanation system.

---

[1] http://9gag.com/
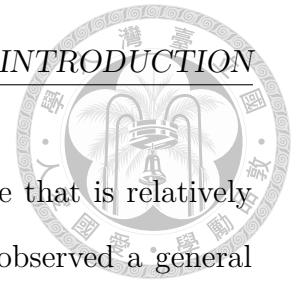
Internet memes are an example of a type of humorous image that is relatively hard for newcomers to understand. Recently, researchers have observed a general fascination with memes [1] [2] [5]. Fig. 1.1 shows a meme background image of the type "Socially Awkward Penguin" and an instance thereof. More implicit background knowledge is embedded in Internet memes than in other images such as LOLCats, face swapping, and pun images. To fully understand Internet memes, one must comprehend the associated shared knowledge and topics. For example, there is always an important characteristic of the background image of a meme that is not explicitly referenced in the meme itself.

Without recognizing this shared knowledge and these topics, one cannot really appreciate Internet memes. A common way to gain this knowledge is to simply read a great deal of memes. Yet this learning process can be very inefficient and frustrating. People who expect to relax on entertainment-oriented websites usually do not want to put so much effort into learning.

In this work, we propose a system that leverages crowdsourcing techniques to generate explanations for memes. The resulting explanations can clearly point out the incongruity between the punchlines and the expectation developed from the setups. We claim that people who are not familiar with Internet meme subculture can still quickly get the gist of the memes by reading the explanations.

There are three main contributions in this paper: (1) applying verbal humor theories to a template-based explanation, (2) developing viable crowdsourcing workflows for generating such explanations, and (3) evaluating our method to show that our workflows help non-native English speakers understand jokes 30% better than with

(a)   (b)
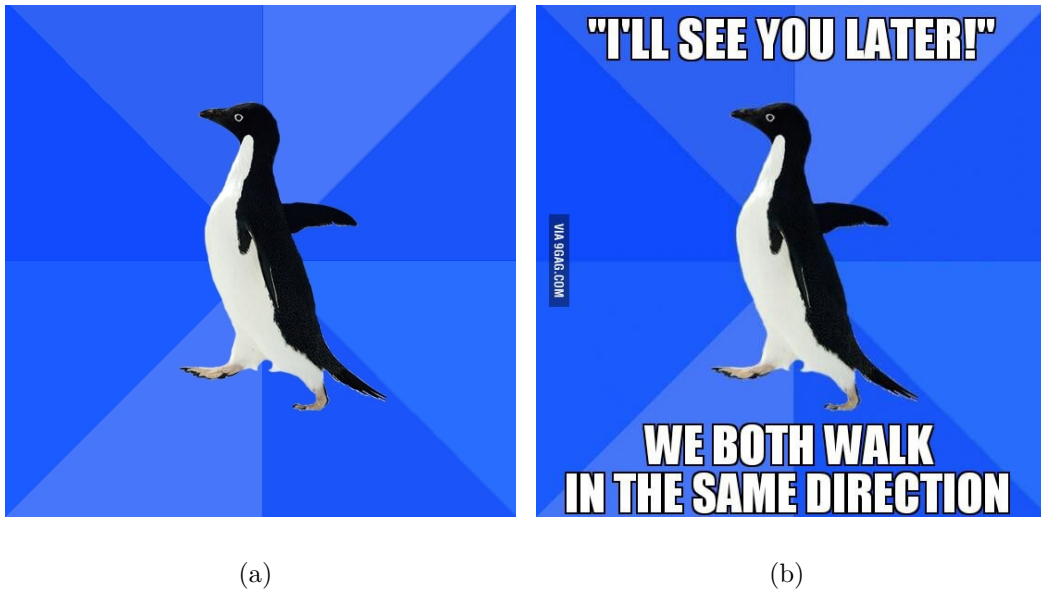
Figure 1.1: An example of humorous Internet memes. (a) A meme template named "Socially Awkward Penguin" and (b) an instance thereof. The text typically narrates uncomfortable life situations, highlighting an exceptionally clumsy or inelegant response.

an expert explanation.

Chapter 2

Related Work

Our work involves several research areas: computational humor recognition, crowdsourcing, and linguistic humor. Based on suitable linguistic humor theories, computational approaches to recognize humor have been proposed. However, we here point out the gap between existing work and the ideal computational humor recognition. We also discuss the innovation of our idea with respect to crowdsourcing.

2.1   Computational Humor Recognition

Computational humor is typically considered to be a hard natural language processing problem. The related work can be divided into two categories: humor recognition and humor generation. Some approaches recognize humorous punchlines by evaluating semantic relatedness and joke-specific features [10]. The result SVM model yields 84% precision when choosing a best punchline from four punchline candidates
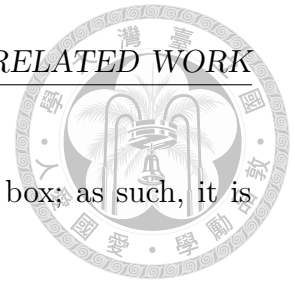
given a set-up sentence. However, the trained model is a black box; as such, it is difficult to determine why a chosen punchline is funny.

Additional methods like finding the keywords and sentence structure in the joke domain have been considered. Some approaches identify "that's what she said" jokes by searching for sexually explicit nouns and sentence structures in erotic domain context [8]. With their proposed method "Double Entendre via Noun Transfer", they achieved 71.4% precision. Even though they are able to find the main elements that made the text erotic, the system still does not understand why the whole text is related to the joke.

Existing approaches all try to recognize humor without dealing with the reasons why people laugh – we fill this gap. In this work, we seek to empower computers to understand the humorous gist by making them discover and comprehend the incongruity behind jokes.

2.2   Crowdsourcing

The wisdom of the crowd has solved many computationally infeasible problems. With crowdsourcing techniques, researchers have successfully achieved goals ranging from generating answers to visual questions for blind people in nearly real-time [4], conducting a variety of remote user studies [9], and enabling people to contribute on microtasks on mobile phones in seconds [13].

In order to explain Internet memes, we choose text explanations as our expression method. In the subfield of text processing, Soylent [3] performs shortening,

proofreading, and other editing tasks on documents. Some approaches attempt to solve the problem of word sense disambiguation by clustering dictionary definitions [11]. Note that the above human tasks were all divided into two or three subtasks for the workers to complete. Our proposed approach is motivated by this find-fix-verify concept, and is divided into explanation generation subtasks.

To our best knowledge, no existing work attempts to generate explanations for humorous materials. With crowdsourcing, humorous text classification can be achieved in an active learning context [6]. In other research, the datasets of joke recommendation have been collected from crowdsourcing [7]. While effort toward joke recognition was made, the algorithms for joke explanation generation are still not clear.

Compared to current work, which only processes general text or classifies humor/non-humor text, we leverage the crowd to generate explanations for humorous Internet memes. It can be said that the idea in our work is relatively innovative. We tackle a problem that cannot be solved by current AI algorithms. However, the generated explanations can be used in further research.

2.3 Linguistic Humor Theories

Incongruity theory is a generally accepted humor theory that was first hinted at in in Aristotle's Rhetoric. The theory states that humor can occur in the resolution of two concepts with incongruity.

The semantic script theory of humor (SSTH) was proposed [12] to implement

incongruity theory. SSTH claims that a text can carry a single joke if:

(a) the text is fully or partially compatible with two different scripts, and

(b) the two scripts are opposite.

The two different scripts are called the "obvious script" and the "alternative script", where the obvious script is the direct meaning in the punchline while the alternative script is the implicit information hidden behind the joke. Humor can be found when the two scripts in a text are opposite. Note that the oppositeness in the two scripts relies on a trait basis. Hence the script can be placed on the two ends of the trait axis.

For example, a one-liner joke goes "I say no to alcohol – it just does not listen." The obvious script is that "the man literally said no to alcohol" and the alternative script is that "the man refused to drink alcohol". In the trait basis of "ability to quit drinking alcohol", the two scripts are at the two ends of the axis. In the fields of both computational humor and linguistic, SSTH is the most suitable theory for analyzing verbal humor.

Chapter 3

Methodology

For every Internet meme, we define three components: the meme template, the set-up, and the punchline. A meme template is the underlying image of a meme and can be repeatedly used in a great number of memes. A set-up and a punchline are the captions superimposed on the upper and lower positions of the meme template, respectively.

To help newcomers understand an Internet meme, a proper explanation is required. Assuming that there are general forms to explain the memes, the question is what is the template of a good explanation for an Internet meme?

3.1   Template-Based Explanation

To find general forms of the explanation, here we take Fig. 1.1(b) as an example. First, we must know that the meme template "Socially Awkward Penguin" describes
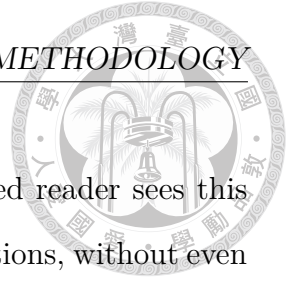
awkward and uncomfortable life situations. When an experienced reader sees this meme, he knows that the captions are a joke about awkward situations, without even actually reading the text. Then we read the upper caption (set-up): "I'll see you later!" For a normal person, saying "see you later" means parting from someone. Therefore, we expect the two to wave goodbye and leave each other in different directions. However, when we read the lower caption (punchline) "we both walk in the same direction", we immediately see that it is an awkward situation because this is contrary to the expectation of parting. The humor lies in the incongruity between the punchline and our expectations.

From the above example, we observe that an explanation can be composed of four lines. The four-line template illustrates the incongruity between a normal situation and the scenario presented in the meme. The template is described below.

1. *Meme template introduction:* A meme template represents the main characteristic in the joke. To newcomers, this is hidden background knowledge. This line helps to build the mental model of the meme template.

2. *Set-up repetition:* In this line, the upper caption in the meme is repeated. The set-up provides the context and premise in the joke.

3. *Anti-punchline:* A normal reaction developed from the set-up. To make the incongruity stand out, the anti-punchline is also an opposite situation to the punchline.

4. *Punchline repetition:* The lower caption in the meme is the punchline which completes the joke.

The anti-punchline is the most important line in the whole explanation. We define an anti-punchline in a meme as the reaction and expectation developed from the set-up, which ends up conflicting with the punchline. Clearly stating an anti-punchline makes the incongruity stand out to readers, making the humorous gist more obvious.

In SSTH, an anti-punchline in a four-line explanation can be treated as the alternative script that is the opposite of the original punchline. The funny gist is revealed when the reader realizes the oppositeness. While finding a trait basis for a general joke can be very challenging, the trait basis for a meme is obvious. The meme template of a meme can determine the trait basis of the oppositeness.

An anti-punchline is not only the most important line in our four-line template: it is also the most difficult to generate. The following subsections describe how we obtain anti-punchlines using crowdsourcing workflows and expert knowledge.

We describe the other lines (meme template introduction, set-up repetition, and punchline repetition) as supporting lines. Supporting lines are also necessary to fully explain the memes. For these lines, existing digital image processing algorithms and natural language processing techniques are sufficient. The generation of supporting lines is described in detail below.

3.2 Three Anti-Punchline Generation Methods

In the following subsections, we propose three methods to generate anti-punchlines. In the first, a crowd is used to collect and then select the best anti-punchline. In
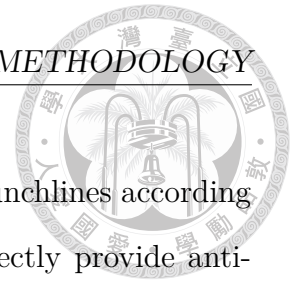
the second, we request crowd workers to directly write up anti-punchlines according to the four-line template. Finally, we ask meme experts to directly provide anti-punchlines.

3.2.1   Collection-Selection Process

In this process, we design two humor subtasks in which to find anti-punchlines. The first subtask is to collect anti-punchline candidates, and the second subtask is to select the best anti-punchline out of the candidates.

Subtask 1: collecting anti-punchline candidates. The workers are asked to imagine themselves in a given situation and fill in their reasonable reaction. Each given situation is the set-up caption that is re-stated in the second person. For each situation, multiple reactions are collected to prepare the selection subtask.

To construct the collection subtask, the meme set-up is extracted and re-stated in the second person. To illustrate the concept of transformation, we take Fig. 1.1(b) as an example. In this subtask, we transform the set-up "I'll see you later" into a situation sentence 'You are a normal person, and you say "I'll see you later!"' When the situation sentence is prepared, the collection subtask can be constructed as Fig. 3.1.

Subtask 2: selecting the best anti-punchline from the candidates from the collection subtask. The workers are asked to choose the best opposite reaction to the punchline with respect to the given trait basis. The trait basis is determined by the meme template and the provided choices are from the collection subtask.

Imagine yourself in this situation:

You are a normal person, and you say "I'll see you later!"

Your reasonable reaction is:

Figure 3.1: An example for Fig. 1.1(b) in the collection subtask.

When enough workers agree with a specific reaction, the reaction is judged to own enough votes, and thus evidences sufficient confidence to be opposite to the punchline. Hence it is selected as the anti-punchline in the explanation. We also provide workers with an "other" option in which they may fill in new reactions when there are no "opposite enough" reactions from which to choose. Once a new reaction is provided, it randomly replaces an existing reaction with the least votes.

After receiving the responses from the collection subtask (Fig. 3.1), we continue publishing the selection subtask shown in Fig. 3.2. The labels on the axis are the clues for the workers to choose a proper opposite reaction. The label on the anti-punchline side is always "Normal situation" while the one on the punchline side is determined by the trait of the given meme template.

After a sufficient number of workers agree that "wave and leave" is the best opposite reaction, we have the anti-punchline sentence against the original punchline. Therefore, a template-based explanation is completed and shown below.

1. *The background image is called "Socially Awkward Penguin". The text typically*

You say "I'll see you later!"

Normal situation   Awkward situation

- Add "alligator" as a lame joke.

- Wave and leave.

- To walk off.

- other: _____

- We both walk in the same direction.
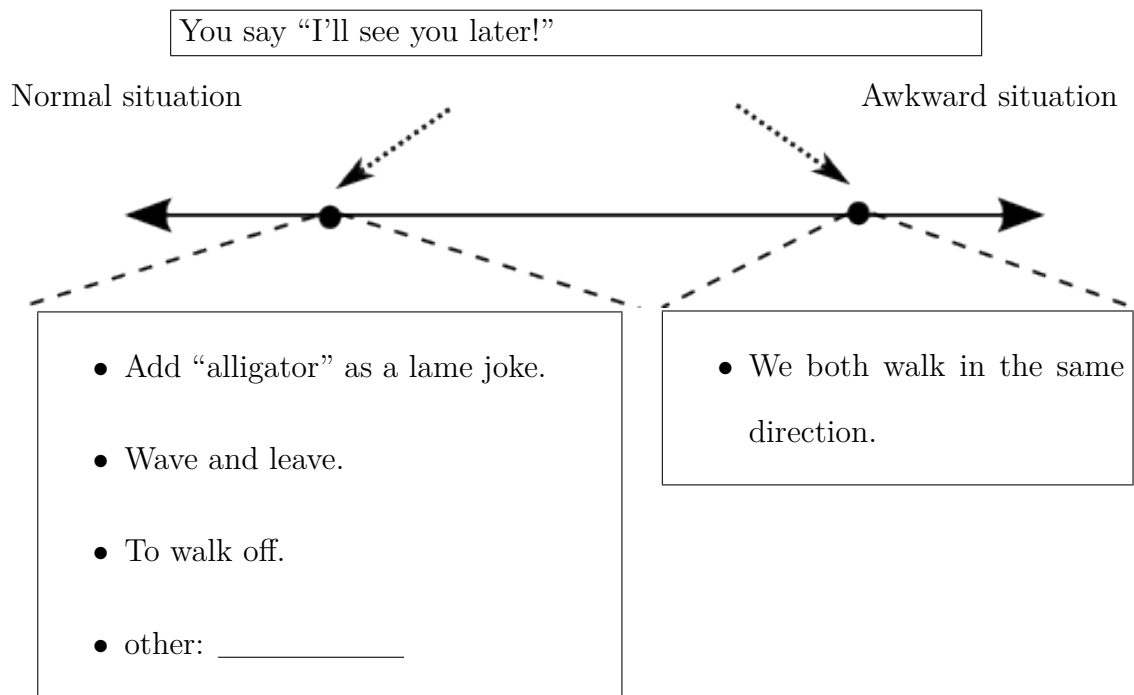
Figure 3.2: An example selection subtask for Fig. 1.1(b).

narrates uncomfortable life situations, highlighting an exceptionally clumsy or inelegant response.

2. *The upper caption says that he says "I'll see you later!"*

3. *We expect him to wave and leave.*

4. *However, in the lower caption, it turns out that they both walk in the same direction.*

The collection subtask is to collect anti-punchline candidates from the workers. After the candidates are collected, the best anti-punchline is selected in the selection subtask. Note that during the process, the workers are not aware that they are explaining memes. As a result, they need not be familiar with memes to complete the subtasks.

3.2.2 Direct Filling Process

While the collection-selection process is divided into two subtasks, the direct filling process consists of a single subtask. An anti-punchline is directly obtained from the worker in the subtask.

In this subtask, the meme image is completely visible to the worker. In addition, the worker can read the supporting lines (meme template introduction, set-up repetition, and punchline repetition) of the template-based explanation. In other words, the workers have access to the full information about the meme and explanation structure, except for the anti-punchline. The direct filling process is to request the

> *The background image is called "Socially Awkward Penguin". The text typically narrates uncomfortable life situations, highlighting an exceptionally clumsy or inelegant response.*
>
> *The upper caption says that he says "I'll see you later!"*
>
> **We expect that _____.**
>
> *However, in the lower caption, it turns out that they both walk in the same direction.*

Figure 3.3: An example subtask for Fig. 1.1(b) in the direct filling process.

workers to enter a good anti-punchline into the blank. Fig. 3.3 demonstrates an example subtask for Fig. 1.1(b).

The following text is one of the output explanations:

1. *The background image is called "Socially Awkward Penguin". The text typically*

narrates uncomfortable life situations, highlighting an exceptionally clumsy or

inelegant response.

2. *The upper caption says that he says "I'll see you later!"*

3. *We expect that they will walk in opposite directions.*

4. *However, in the lower caption, it turns out that they both walk in the same direction.*

Because the workers have to understand the funny gist before entering their anti-punchline, this process has an higher entry barrier. However, they may use the four-line template structure as an aid for comprehension. The explanation breakdown can help them organize what information the meme is trying to communicate.

3.2.3 Expert Process

In order to understand the advantage of the above proposed methods over experience readers' help, we designed a process to obtain anti-punchlines from meme experts.

The expert process is identical to the direct filling process, except that the workers are selected meme experts. The memes and supporting lines are shown to the experts. The experts are asked to come up with a proper anti-punchline to fill in the blank line. Here is one of the explanations produced by an expert:

1. *The background image is called "Socially Awkward Penguin". The text typically narrates uncomfortable life situations, highlighting an exceptionally clumsy or inelegant response.*

2. *The upper caption said that he says "I'll see you later!"*

3. *We expect that he and his friend will part ways.*

4. *However, in the lower caption, it turned out that they both walk in the same direction.*

3.3   Supporting Lines

The supporting lines (meme template introduction, set-up repetition, and punchline repetition) can be extracted automatically. Basic digital image processing and natural language processing techniques are sufficient to produce these lines.

Lists of the most popular meme templates can be found online (for example Quick Meme[1]). When the meme templates are given, the template of a meme can be easily recognized by calculating the similarity between the target meme and all of the templates. After the template is determined, additional information can be retrieved from Know Your Meme[2].

The set-up and punchline repetition requires optical character recognition (OCR) techniques. Since the state-of-the-art OCR algorithms are able to recognize computer-printed characters with very high precision, we use it as a tool to extract the set-up and punchline captions from memes. In the explanations, these two lines are usually stated in the third person; this necessitates basic substitution of the subjects and grammatical verb tenses.

---

[1]`http://www.quickmeme.com/`
[2]`http://knowyourmeme.com/`

Chapter 4

Experiments

4.1 Datasets

Since 9GAG is one of the most popular websites for funny images, we chose it as our Internet meme source and crawled 39447 posts from it. From these posts we recognized 3051 memes in 38 unique meme templates. Table. 4.1 shows the ten most popular meme templates in our datasets.

For each meme template, we randomly selected two memes to be used in our experiment, resulting in a 76-meme testing dataset. After completing the human task processes, the output explanations for each memes were generated.

As Fig. 1.1(b) is one of the randomly selected memes in the testing datasets, in the following sections we continue to use it as an example describing how we execute the collection-selection and direct filling processes.

Rank	Template name	Proportion
1	Confession Bear	16.91%
2	Actual Advice Mallard	16.65%
3	Success Kid	9.40%
4	Socially Awesome Awkward Penguin	5.24%
5	Scumbag Steve	5.04%
6	Good Guy Greg	4.49%
7	First World Problems	4.29%
8	The Most Interesting Man in the World	3.90%
9	Futurama Fry	3.73%
10	Socially Awkward Penguin	3.31%

Table 4.1: Proportion of top ten meme templates in our datasets.

4.2 Experiment Settings

To conduct the experiment, we used Amazon Mechanical Turk (MTurk). However, the human tasks are not limited to MTurk and can be ported to any crowdsourcing platform.

In the collection-selection process, we collected three responses in the collection subtask. After collecting the three anti-punchlines candidates, we selected the best anti-punchline in the selection subtask. In the direct filling process, we also gathered three anti-punchlines. For the expert process, we invite an American subculture lover who has been reading memes for over two years. This expert is a 24-year-old male graduate student in computer science.

4.3 Evaluation

In the evaluation, ten memes were randomly chosen from the testing datasets. We designed a questionnaire to ask each participant whether the explanations were meaningful and helpful to them. For each of the ten memes, five explanations were attached. The explanations consisted of three sources: (a) one explanation generated from the collection-selection process, (b) three explanations generated from the direct filling process, and (c) one explanation provided by an experienced meme reader. The motivation behind including a meme expert's explanation was to use it as a baseline. The order of the five explanations was randomized. When the participants read the explanations, they were not aware of the respective sources.

After recruiting, nine volunteers participated in the evaluation and provided

their opinions for the memes and explanations. The participants were all non-native English speakers. They consist of five females and four males, all graduate students. For each meme in the questionnaire, a participant was required to answer whether they had seen the meme template before in a yes/no question. Immediately, without any explanation, the questionnaire asked the participant if he or she understood the funny gist right after reading the meme. The participant was permitted answers of "yes", "no", and "not sure" for the prior understanding question. This was followed by five explanations. For each explanation, two questions were asked. The first question was if the participant considered the given explanation reasonable, and the second was if the participant believed the explanation to be helpful. The options for the two questions were also in the form of "yes", "no", and "not sure".

From the results of the questionnaire, we found that none of the participants were familiar with the meme subculture. Four of them had seen one meme template out of ten memes, and the other six participants had never seen any meme template. As mentioned in previous sections, they can be considered newcomers to the subculture.

To compare the helpfulness of the different approaches to generating anti-punchlines, we show the statistical results of the participants' feelings that an explanation was helpful. In order to measure the helpfulness of an explanation, we simply counted the number of "yes" answers to the helpfulness question. The answers of "no" and "not sure" did not count. The overall statistical results are presented in Fig. 4.1. In the chart, we observe that both crowd processes produced anti-punchlines comparable with those of the expert.

We compared not only the individual processes, but also calculated the union
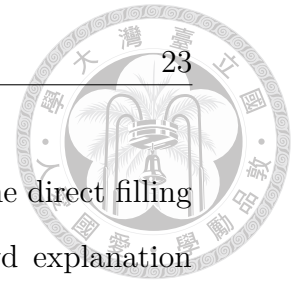
of the crowd processes, that is, the selection-collection process and the direct filling process. In our evaluation settings, a participant considers a crowd explanation helpful if at least one of the explanations from the selection-collection and direct filling processes helps him understand the meme. In Fig. 4.1, we see that the crowd union yields a 30% improvement over the expert.

Also note that the crowd union is substantially improved over the individual collection-selection process and direct filling process. It is interesting that the two processes are mutually complementary. When one crowd process cannot explain a meme well, the other one is very likely to be able to explain it.

The helpfulness ratios with respect to individual memes are shown in Fig. 4.2. The blue bars stand for the ratios of participants that understood a meme before reading any explanation. The red and yellow bars are the helpfulness of the explanations from the expert and our crowd processes, respectively.

Meme #3 is a meme nicely explained by the crowd explanations; the explanations have a helpfulness ratio of 77%. The meme is shown in Fig. 4.3 and feature a meme template named "Scumbag Steve". The five explanations shown to the participants are:

1. *The background image is called "Scumbag Steve". The overlaid text generally centers around unethical behavior regarding drugs, partying, and other hedonistic behavior.*

2. *The upper caption says that he borrows a bunch of your stuff.*

3.a *(collection-selection process) You expect that he will be careful to return it in*
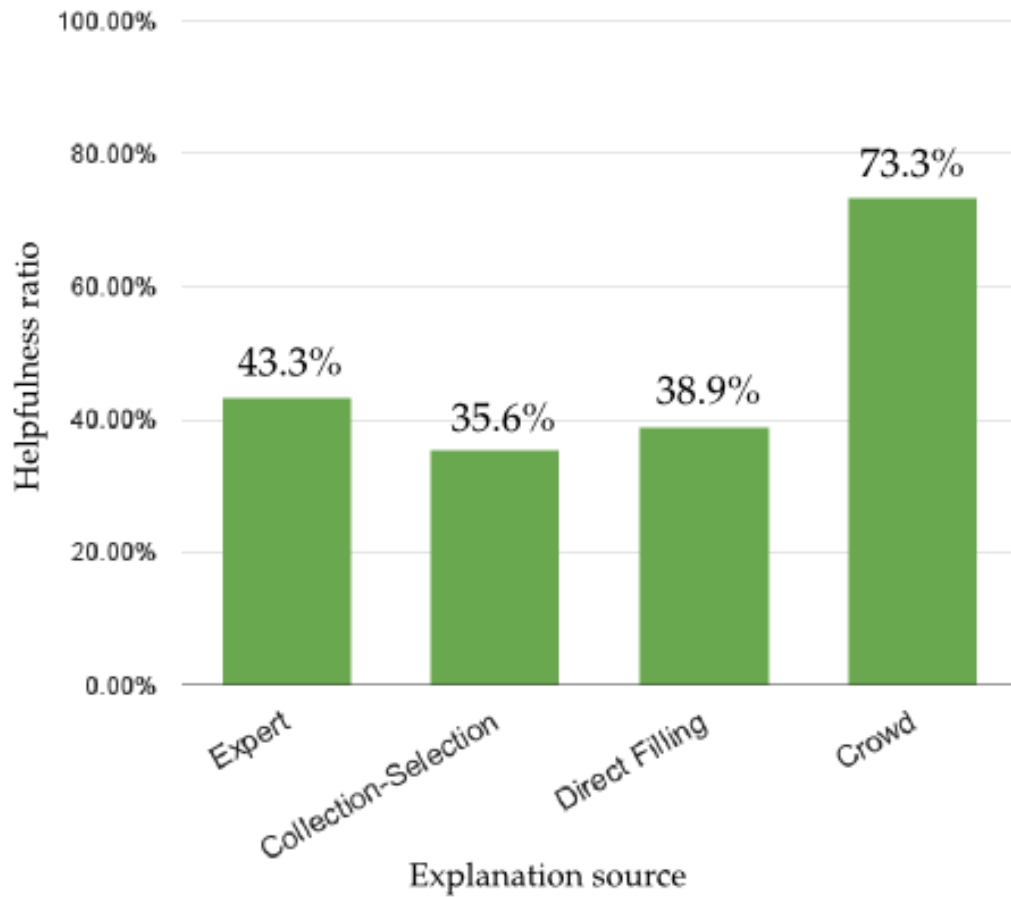
Figure 4.1: The helpfulness ratio for the explanations from experts, the selection-collection process, the direct filling process, and the crowd processes. The crowd processes are the union results of the selection-collection and direct filling processes. In other words, a newcomer indicates that an explanation from the crowd processes is helpful if he feels at least one explanation from the selection-collection and direct filling processes is helpful.

good condition.

3.b (direct filling process) You expect that he will return your stuff later.

3.c (direct filling process) You expect that he will not return it in a timely manner.

3.d (direct filling process) You expect that he will return it to you.

3.e (expert process) You expect that he will return your stuff on time.

4. However, in the lower caption, it turns out that he lends it out to people you don't know.

The above explanation contains five unique anti-punchlines, each preceded by its source in parentheses. As we can see, the majority of the anti-punchlines follow the logic and flow in a four-line template. Among the five anti-punchlines, only one of them from the direct filling process (3.c) expresses a different idea, that which is not an appropriate anti-punchline.

One possibility why the expert's anti-punchline is not helping the participants is that the set-up and punchline do not mention anything about when the stuff will be returned; referring to this concept thus only confuses newcomers. It may be that experts are not able to produce proper anti-punchlines for every meme because experts have already forgotten what is confusing to a newcomer.

To evaluate the cost efficiency of the processes, we discuss the payment strategy as follows. In the collection-selection process, we collected three anti-punchline candidates and selected the best one as the average three votes. For each collection and selection subtask, we paid \$0.05 to the workers. We thus spent \$0.05 \times (3 + 3)

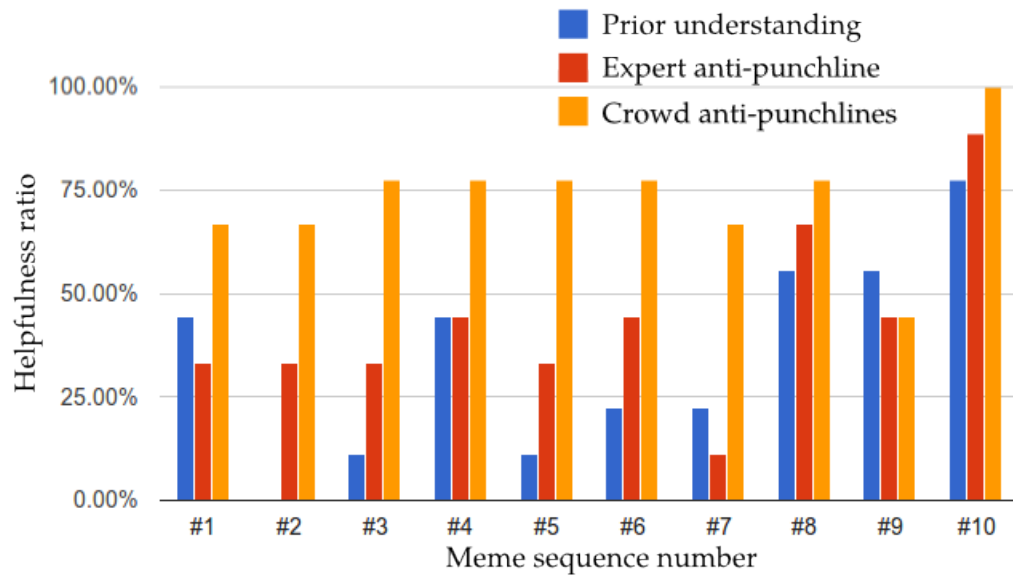Figure 4.2: Comparison for the helpfulness of the expert's explanation (red) and the proposed crowdsourcing workflows (yellow) to the participants. The ratios of participants that get the funny gist before any explanation (blue) are shown to indicate the readers' knowledge gain. The horizontal axis stands for the sequence number of the ten memes and the vertical axis is the ratio of participants who agree with the offered opinions.

Figure 4.3: Meme #3 in the evaluation questionnaire. It is one of the memes well-explained by the crowd processes.

= $0.30 to find an anti-punchline. In the direct filling process, we also offered $0.05 for the workers to fill in an anti-punchline.

Chapter 5

Discussion

Here we review the possibilities and limitations in generating sufficient explanations for Internet memes by crowdsourcing techniques.

The evaluation results suggest that our collection-selection process and direct filling process produce helpful explanations. However, the helpfulness is obvious only after the readers have read all four explanations. Reading more than one explanation requires extra patience and is unlikely for most Internet users. A good mechanism for selecting a best explanation from those generated would be significant.

Considering the aspects of time efficiency and monetary cost, the two proposed processes have different strengths. The collection-selection process is much faster than the direct filling process. As dynamic changing of the content and topics is the nature of memes, speed is a decisive factor. If we emphasize the monetary cost, the direct filling process is more prominent. When dealing with millions or even billions of memes on the Internet, a process with high cost is impractical.

As far as limitations, we are aware that our four-line explanations are not suitable for explaining all Internet memes. For example, some memes contain only one caption (Fig. 5.1(a) for example) and cannot be used in our template-based explanation. Another condition that we cannot handle is meme templates that have one conventionally fixed caption (e.g. Figures 5.1(c) and 5.1(d)).

Fortunately, less than 1% of memes contain only one caption, and only 8.4% of meme templates have a conventional fixed caption: these are sufficiently low ratios. To sum up, we leverage crowdsourcing to produce explanations for humorous Internet memes. Our experiments demonstrate that acceptable explanations are successfully generated by completing the humor tasks.

5.1 Online Product

On Google Chrome Web Store[1], an online explainer called "9GAG Lookup" is available in a form of browser extension. As the previous introduction describes, 9GAG is one of the most famous social website for Internet meme sharing. We build 9GAG Lookup to assist readers to understand the funny gist of the posts on 9GAG.

9GAG Lookup provide mainly two services: phrase lookup and meme understanding. The phrase lookup service is for clarification of input phrases. Not all of the English vocabularies on the website can be mapped to a text explanation. Some of slang or buzz word can only be explained by a short video or a series of images. When a user query a phrase, 9GAG Lookup will search in possible resources
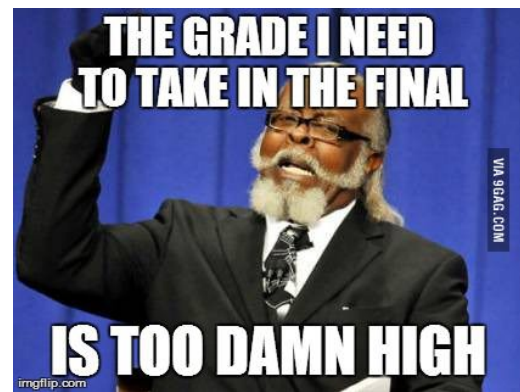
---

[1]`http://chrome.google.com/webstore/`

(a)   (b)

(c)   (d)

Figure 5.1: (a) An meme with only one caption with a meme template named "Actual Advice Mallard". The intention to use only one caption "take a walk" is to express that taking a walk is a good advice in any situation. (b) A normal meme uses "Actual Advice Mallard" template with two captions. (c) An meme with the meme template named "One Does Not Simply". The meme template usually begins with the phrase "one does not simply". (d) An meme with the meme template named "The Rent is Too Damn High". Memes using this meme template usually end with "is too damn high".

(e.g. Google Image, Youtube, Urban Dictionary) and return a list of explanation candidates.

The meme understanding service is the implementation of the proposed method. Through the human task processes, explanations of memes can be produced. First of all, the system will determine the meme template of the given image by digital image processing techniques and reconfirm by users' suggestions. After this, the captions on the meme can be extracted by optical character recognition algorithms and corrected by the users. With the meme template and captions set, the only missing component in the four step explanation is the anti-punchline. With the proposed collection-selection and direct-filling processes, the full explanation can be produced.

Chapter 6

Conclusion and Future Work

Humorous Internet memes can be seen in many image-based online forums. However, steep learning curves are observed before newcomers can fit in to this subculture. In this paper, we propose a template-based explanation approach to extract the incongruity in humorous Internet memes. With crowdsourcing techniques, we are able to generate such explanations for most of the cases.

In this paper, there are mainly three contributions: template-based explanation applying humor theories, workflows of anti-punchline generation, and an evaluation of our crowd explanations. The four-line explanations are based on SSTH, and illustrate the incongruity and oppositeness between the anti-punchline and the punchline. The proposed two anti-punchline generation processes demonstrate how crowdsourcing may be leveraged to generate explanations for humor materials.

The experimental results and evaluation show that the explanations produced by our processes indeed do help newcomers to pick up the funny gist. An additional

comparison between explanations from an expert and our system is made to highlight the strength of our proposed processes.

For further study, is it possible to leverage the proposed system to build an explainer system for other forms of humor. For example, an explainer system for comic strips could be built if the last dialogue can be extracted as the punchline and the previous ones as the set-up. Pure computational humor recognition and generation can also consider our system as a source of explained humorous data for better analysis.

Bibliography

[1] C. Bauckhage. Insights into internet memes. In L. A. Adamic, R. A. Baeza-Yates, and S. Counts, editors, *Proceedings of the Fifth International Conference on Weblogs and Social Media.* The AAAI Press, 2011.

[2] C. Bauckhage, K. Kersting, and F. Hadiji. Mathematical models of fads explain the temporal dynamics of internet memes. In *Proceedings of the Seventh International Conference on Weblogs and Social Media*, 2013.

[3] M. S. Bernstein, G. Little, R. C. Miller, B. Hartmann, M. S. Ackerman, D. R. Karger, D. Crowell, and K. Panovich. Soylent: A word processor with a crowd inside. In *Proceedings of the 23rd Annual ACM Symposium on User Interface Software and Technology*, pages 313–322, New York, NY, USA, 2010. ACM.

[4] J. P. Bigham, C. Jayant, H. Ji, G. Little, A. Miller, R. C. Miller, R. Miller, A. Tatarowicz, B. White, S. White, and T. Yeh. Vizwiz: Nearly real-time answers to visual questions. In *Proceedings of the 23rd Annual ACM Symposium on User Interface Software and Technology*, pages 333–342, New York, NY, USA, 2010. ACM.

[5] M. Coscia. Competition and success in the meme pool: a case study on quickmeme.com. *International Conference of Weblogs and Social Media*, abs/1304.1712, 2013.

[6] J. Costa, C. Silva, M. Antunes, and B. Ribeiro. On using crowdsourcing and active

learning to improve classification performance. In *International Conference onIntelligent Systems Design and Applications*, pages 469–474, Nov 2011.

[7] D. Gupta, M. Digiovanni, H. Narita, and K. Goldberg. Jester 2.0 (poster abstract): Evaluation of an new linear time collaborative filtering algorithm. In *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 291–292, New York, NY, USA, 1999. ACM.

[8] C. Kiddon and Y. Brun. That's what she said: Double entendre identification. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, pages 89–94, Portland, OR, USA, Jun 2011.

[9] A. Kittur, E. H. Chi, and B. Suh. Crowdsourcing user studies with mechanical turk. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 453–456, New York, NY, USA, 2008. ACM.

[10] R. Mihalcea, C. Strapparava, and S. Pulman. Computational models for incongruity detection in humour. In A. Gelbukh, editor, *Computational Linguistics and Intelligent Text Processing*, volume 6008, pages 364–374. Springer Berlin Heidelberg, 2010.

[11] G. Parent and M. Eskenazi. Clustering dictionary definitions using amazon mechanical turk. In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk*, pages 21–29, Stroudsburg, PA, USA, 2010. Association for Computational Linguistics.

[12] V. Raskin. *Semantic Mechanisms of Humor*. D. Reidel, 1 edition, Dec 1985.

[13] R. Vaish, K. Wyngarden, J. Chen, B. Cheung, and M. Bernstein. Twitch crowdsourcing: Crowd contributions in short bursts of time. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2014.