MetaMeme: A Dataset for Meme Template and Meta-Category Classification

Benjamin Lambright[*] and **Jordan Youner**[*] and **Constantine Lignos**[**]
Michtom School of Computer Science
Brandeis University
bwlambright@gmail.com, jordan.youner@gmail.com, lignos@brandeis.edu

Abstract

This paper introduces a new dataset for classifying memes by their template and communicative intent. It includes a broad selection of meme templates and examples scraped from Imgflip and a smaller hand-annotated set of memes scraped from Reddit. The Reddit memes have been annotated for meta-category using a novel annotation scheme that classifies memes by the structure of the perspective they are being used to communicate. YOLOv11 and ChatGPT 4o are used to provide baseline modeling results. We find that YOLO struggles with template classification on real-world data, but outperforms ChatGPT in classifying meta-categories.

1 Introduction and Related Work

The majority of the previous research on the automatic classification of memes revolves around specific domains like the detection of politics or hate speech or classifying memes by humor "type" (Courtois and Frissen, 2023). Little work addresses task of identifying which *template*—not necessary a single literal image, but a recognizable reference format for the meme—a meme falls into, but Courtois and Frissen offer one of the most promising recent attempts to do so. Their work uses two datasets: a dataset of randomly selected memes and hand-annotated templates from one of the most widely used meme documentation websites, KnowYourMeme, and a dataset of memes paired with their templates scraped from the aggregator website 9gag. Their method for template identification involved CNNs for detecting features common across examples, calculating accuracy with a geometric mean given that parts of their datasets had templates not seen in training.

Along similar lines, Gleason et al. (2019) attempt simple meme template detection. They use

---

a combination of the the Multi-Scale Structural Similarity (MS-SSIM) index and a color histogram between the input and template image (Wang et al., 2003) to match memes to their templates. Even with their small dataset of 385 memes and 137 templates, they achieved an accuracy of 92.25%.

This paper offers a similar methodology for obtaining the dataset to Courtois and Frissen (2023), but with a larger and more varied set of templates. While Gleason et al. (2019) provides promising results for a small dataset, we expect our template identification task to be much more difficult, with over 2,000 template classes compared to their 137. Additionally, Gleason et al. (2019) likely reached near-ceiling performance because all of their memes were generated using the exact same source images for templates, but the reality of meme usage across the internet is that there is more room for variation, as they are often not made using generators with consistent template images.

Along with examples from Imgflip (a meme generator website), we also include memes sourced from Reddit. In total, our dataset covers 2,059 meme templates with a collection of 274,748 examples scraped from Imgflip and a collection of 242 hand-annotated memes from Reddit. With the inclusion of social media-obtained memes, we expose our model to examples of a given template that are likely much more diverse than those found exclusively on generator websites, which gives a better window into performance on real-world data.

To support the annotation of the Reddit memes, we introduce an ontology of *meta-categories* meant to explore the communicative intent behind memes. Recent work has used LLMs and VL (vision language) models to identify the metaphors represented by a given meme, as well as explaining the entire joke itself (Hwang and Shwartz, 2023). As an alternative to this, we explore a structure-based approach with the assumption that an understanding of the desired intent in using a meme can be

---

[*]Equal contribution. [**]Corresponding author.

identified without requiring any context. Between the meta-category classification task and the more traditional meme template classification task, this dataset contributions a novel resource for the study of meme creation and usage.[1]

2 Dataset

2.1 Data Sources

2.1.1 Imgflip

The large amount of unannotated data that we collect for the purpose of identifying meme templates was collected from imgflip by modifying an open-source codebase.[2] Imgflip was the best website for this dataset for two reasons: it is easy to scrape memes and their corresponding templates, and the website documents thousands of templates.

Our dataset contains 1,992 template classes from the "Top *All Time*" page from imgflip. While there were initially more classes, we found that some of the templates on imgflip were duplicates. In order to handle this, we used an image hashing algorithm to remove duplicates. Some template duplicates were also removed manually, if image hashing could not identify them.

All additional template classes come from our Reddit data. In order to ensure that our model has trained on all templates represented in the Reddit data, we scraped 67 additional templates from Imgflip that appeared in the Reddit data but were not in the original top 2,000 classes. In total, this dataset includes 274,748 meme examples from 2,059 templates.

The dataset is organized in two ways. For image classification tasks, each template is a directory that includes all the image files for the template class. The dataset also includes a json file of all the metadata for each example meme, including the image url, post url, captioning for the meme, and other metadata.

2.1.2 Reddit

In order to have data that reflects more real-world usage than output from meme generator websites, we also scraped images from Reddit's r/memes community (subreddit).[3] We scraped the following from Reddit: 100 memes from the top of all

Meta-category	Count
Reaction	100
Exploitable	89
Image Macro	7
Duality	30
Escalating Progression	9
None	7

Table 1: Counts of the various meta-categories in our annotated Reddit data, including examples that were annotated as having no meta-category.

Figure 1: An example of a meme template. "Drakeposting" or "Drake Hotline Bling" is a meme whose template involves two frames of Drake from the *Hotline Bling* music video that are set up to rate something as bad and something else as good. The user of the template then fills in the whitespace with two images or pieces of text to pair with the Drake ones.

time, 100 from the top of the month scraped on November 13th 2024, and 50 from the top of the month scraped on December 13th 2024. For each example, the annotators identified the corresponding imgflip template (if one could be found) and selected the meta-category that best fits the meme. In total, there were 242 annotated examples,[4] 74 of which did not have a template and 20 that had an identifiable template, but that template was not documented on imgflip.

2.2 Task Definitions

2.2.1 Meme Template Classification

The primary task of this dataset is to identify the template a meme is built from. A meme template is a recognizable joke format that the user fills in the details of when they create a meme. Figure 1 gives an example template. The task involves mapping a human-created meme to the template that was used when creating it.

2.2.2 Meta-category Classification

In order to facilitate capturing the semantic and pragmatic content of memes, we developed an ontology of meta-category categories which define formats and structures that are common across templates. This can potentially augment meme template detection and offer a way to group memes of unseen templates. This set of categories is in part derived from the ontology presented by knowyourmeme.com. Consequently, an alternative task for this dataset is to classify these general categories of memes in order to classify their sentiment.

2.3 Annotation Ontology and Results

The annotators for this project are two male graduate students at Brandeis University between 20-30 years old, receiving course credit for their annotation work. They were asked to annotate 250 images scraped from Reddit and the top 100 most popular template images from Imgflip. The former were annotated for both templates and meta-categories, while the latter were only annotated for the most common meta-category that meme template would have. For the Reddit dataset, the template annotations were almost always consistent between the two annotators. In the few cases they were not, the annotators discussed the difference and performed adjudication jointly.

These meta-categories are derived from some of the higher-level categories presented in the KnowYourMeme ontology. While the website contains hundreds of categories, the ones we chose to use more as they were presented on the website (image macro, exploitable, reaction) are comparably more general and high-level. For duality and escalating progression, we came across the Drake-posting category, which describes memes that resemble the duality/escalating progression structure,

---

⁴Some images were removed due to broken links, deleted posts, or offensive subject matter.

and decided to create two distinct categories that capture similar structural information but mainly differ in terms of semantic content (and enable categorization of such memes without referring to a specific example template). With these selections, we feel that we were able to encompass nearly all memes and enable capturing of some semantic information.

The inter-annotator agreement for our meta-categories was consistently around 70%. This section provides a description of our ontology and a summary of the annotation approach. The full guidelines along with examples of all meta-categories and the guidance for annotation edge cases can be found in the Appendix. Counts for each meta-category annotated are given in Table 1.

Image Macro Image macros can be thought of as memes for which the setup of the joke is given by the image, and the text fills in the details and punchline. For our purposes, we consider image macros as consisting of a single image. The most typical presentation of these memes involves a picture of some kind of entity, often an animal or person, with white text in impact font on the top and bottom of the image.

Reaction Image Reaction images are, in a sense, an inversion of image macros. Instead of the image setting up the joke and the text filling in the punchline, the joke is set up by a text caption that is almost always above or below the image content. The punchline is the image. The images in these memes are usually of people emoting or reacting in some fashion, and the humor often derives in part from the fact that the text caption completely recontextualizes the image from its origin.

Exploitable An exploitable is a meme in which an existing image such as a comic, or still frame(s) from a movie is augmented by adding or replacing text or characters to tell a joke. The idea is that there is some extant structure inside the original image that is "exploited" using additional text or images.

Duality A meme exhibiting duality is one that compares two (or more) situations or contexts that are related across some dimension in opposition to each other. The typical components of a duality meme are a set of discrete contexts and a set of images that visualize the relationship between the contexts. A common format is a "4-panel" layout

in which a pair of contexts and images are stacked vertically.

Escalating Progression Escalating progression memes are those that express a reaction to points sampled along a continuum of context. Unlike in duality where the situations are connected in terms of directly opposing each other (bad vs. good), with escalating progression there is an intensification between each state along an axis (good, better, best). Typically these memes have at least 3 sets of image-context pairs, but it is not required as long as there is an obvious sense of continuum between the contexts.

None While the previous categories are designed to cover as much of the semantic-pragmatic space of memes as is feasible, there are examples that do not align well with any of them. Annotators were always encouraged to mark a category, but it was permissible to label the meta-category as "none" if they felt strongly that none of them fit. There exist other "genres" of memes such as surreal, "deep-fried" or anti-memes which are not addressed by the ontology because they exist in small niches. One could realistically add several more categories to this ontology, but the point of this ontology is for it to be small and representative of most existing memes.

3 Experiments

3.1 Experiment Design

For template classification, we divided our dataset into the images annotated from Reddit and the images scraped from imgflip. From these, we further split the imgflip data into train, dev, and test sets using a 90-5-5 split. We treated the Reddit data as a unique test set, because this would allow us to test on data which the model mostly likely has never seen before, unlike the test set of the imgflip data, which was normally the same images with different text overlayed on it.

Along with classifying the specific templates for each image, we also classified meta-categories (image macro, reaction image, exploitable, duality, escalating progression, and none). In order to do this, we train on two different datasets to see which can best identify the meta-categories. For the imgflip data, each template is annotated with a meta-category and the label is transferred to all images that were generated using that template. For the Reddit data, each image has its meta-category

Task/Dataset	Accuracy
Imgflip Data Template	99.8
Reddit Data Template	40.5
Imgflip Data Meta-category	51.5
Reddit Data Meta-category	68.8

Table 2: Accuracies for the YOLO model on the classification tasks: template identification and meta-category identification.

hand-annotated since the template used to generate is not given automatically through the scraping process (unlike the imgflip data).

The imgflip data was split into train, validation, and test portions using a 90-5-5 split. For the Reddit data, we used an 85-15 train-test split.

3.2 Models

We use two different models to provide baseline results for our dataset: YOLOv11's image classification model and ChatGPT 4o.

3.2.1 YOLOv11

We chose to use YOLOv11 because YOLO has consistently strong performance on image classification tasks (Khanam and Hussain, 2024) and focuses on data examples beyond ImageNet. We fine-tuned the model on scraped data from imgflip. After fine-tuning the model with optimized hyper-parameters,[5] we evaluated the model on a test set from imgflip and Reddit. By evaluating these two test sets separately, we are able to see how the model predicts images that are the same as the training data with different text (imgflip), versus images that are conceptually similar, but often vary significantly from the training data (Reddit).

As shown in Table 2, while the template classification YOLO model reaches near-100% accuracy when testing on the imgflip test data, it struggles with the Reddit data. This is because the memes posted on Reddit do not always use the same images that define the template on imgflip, especially in the cases where they were not directly generated from an imgflip template. As a result, these "in the wild" examples [6] are much harder for the YOLO model to identify. The imgflip meme template classification task is very simple because it is

---

[5]Values: patience=2, image size=640, optimizer="AdamW", learning rate=0.01, momentum=0.937, weight decay=0.0005.

[6]See the appendix for examples.

almost always the same image with different text overlayed on it.

As for the Reddit template identification, 40.5% accuracy indicates that the model is still able to identify the template some of the time. There are 2,059 meme templates for the classifier to select from, so this indicates some level of understanding, even if it does not match the performance of other models trained on fewer classes in prior work (Gleason et al., 2019).

In terms of meta-category classification, YOLO performed reasonably when trained and tested on imgflip data (51.5%), and even better when tested on the Reddit data (68.8%), where a random selection of classes would ne expected to produce 16.7% accuracy.

However, Table 3 indicates that there was a significant amount of overfitting due to dataset imbalance. The less common classes (none, image macro, duality, and escalating progression) were classified correctly significantly less frequently than the common classes of exploitable and reaction Image. Reaction images were correctly classified 82% of the time, and exploitables 89.5% of the time, indicating that at least as a binary classification task these two classes can be correctly understood by the model.

These results indicate that the YOLO model is learning some sort of structural information that allows it to predict at least the most common meta-categories with reasonable accuracy. Furthermore, the fact that the Reddit dataset had a much larger number of templates relative to the number of examples of each may have pushed the model away from relying on direct relationships between templates and meta-categories.

3.3 ChatGPT 4o

In addition to testing our data using the YOLOv11 model, we also used ChatGPT's 4o API. Because of the complexity of identifying the semantic structure of memes, we posited that an LLM might be well suited for the task. In all of these experiments, we always started with the exact same initial and final prompt, which can be found in the Appendix. Additionally, we kept the prompting parameters, temperature and top P, at their default values: 1.00 and 1.00.

We performed three different few-shot experiments with this model: text-based few-shot prompting, image-based few-shot prompting, and a combination of both. In our first experiment, the prompt

describes each of the meta-category categories in a few sentences as text, based on the annotation guidelines. In the second experiment, we prompt with an image example for each meta-category category and explain why that image is an example of the meta-category. In the third experiment, we give the prompts from both the first and second experiments together. All of these prompts are recorded in the Appendix. Because this model is nondeterministic, we ran these experiments three times and calculated the mean score and standard error of the mean.

As shown in Tables 4 and 5, ChatGPT performed significantly better on the Reddit data than the imgflip data, but the specific prompt style used did not have much impact. Given that these two datasets were annotated by the same annotators with the same guidelines, this suggests that the memes in the Reddit dataset were easier for the model to interpret. Having both text and image examples in the prompt likely caused the model to have too much information to properly process it all, but providing text or image prompts provided the same mean score.[7]

While it be reasonable to expect that ChatGPT would be better at reasoning how to classify meta-categories if it necessitates understanding the semantics of the meme, ChatGPT was still not very successful at identifying the meta-categories. Importantly, the model did not always guess the same category, it simply chose a different wrong category most of the time. This suggests that this was still a difficult task for the model, even if the prompts could be improved, and the model was not overfitting.

4 Future Work and Conclusion

While the ChatGPT results were disappointing, the consistency of the poor results could suggest that the model might need something more complex than just prompts. For example, using chain-of-thought or reinforcement learning might help the model improve perform performance by allowing it to reason more. For the same reason, ChatGPT's o1 pro model, which now has vision capabilities, might show stronger results as well (Noda et al., 2025). As for the YOLO models, it suggests that the meta-category ontology can be predicted, with reasonable accuracy, using a pure computer vision

---

[7]The individual scores for text and image prompts generally varied between 40% and 43%; they never produced exactly the same result.

Actual	Predicted					
	None	Exploitable	Image Macro	Duality	Escalating Progression	Reaction
None	0	2	0	0	0	1
Exploitable	0	76	2	0	0	7
Image Macro	0	2	1	0	0	6
Duality	0	18	0	4	0	7
Escalating Progression	0	5	0	1	0	5
Reaction	0	17	1	0	0	82

Table 3: Confusion Matrix for the predictions from the YOLO model on the Reddit data meta-category identification task

Task	Accuracy
Imgflip Data Meta-category	$29.6_{\pm1.8}$
Reddit Data Meta-category	$47.7_{\pm0.2}$

Table 4: Mean accuracies and standard errors for Chat-GPT meta-category classification with image prompts on different datasets.

Prompt Content	Mean Score
Text	$42.2_{\pm1.1}$
Image	$42.2_{\pm0.6}$
Both	$37.8_{\pm1.6}$

Table 5: Mean accuracies and standard errors for Chat-GPT meta-category classification across all test sets using different prompting strategies.

approach. This could help classify the semantic meaning of a meme without requiring the context of an entire joke.

Knowing that annotators considered the text when annotating memes, it would likely also be helpful for the text to be included in the pipeline for meta-category classification. It would be easy to prompt for this with ChatGPT, asking to include this information in its final inference.

The best way to improve the performance of both tasks is likely to be getting more data. Scraping from more sources than imgflip and Reddit would provide a more diverse dataset, as well as being able to handle more meme formats, like GIFs.

The Reddit dataset was also very small due to the limited amount of annotator time available, and more data would likely have improved the performance of our models. On a similar note, in the imgflip dataset, there were meme templates that were extremely similar to each other, as discussed further in the Appendix. While we removed exact duplicates in data preprocessing, there are likely

still meme template classes that could be combined, and would likely improve the performance of the YOLO model on meme template classification.

5 Limitations

Some templates that are functionally duplicates (such as two examples of "Drakeposting" using slightly different stills from the video) may not have been discarded from our image hashing algorithm, potentially causing a dip in accuracy. We recommend using a deep learning model to identify these near-duplicate templates in the future. Also, the total amount of annotated Reddit data is very small, making both learning and inference more difficult. A larger group of annotators combine with a more varied set of data would allow for greater exploration of the machine-learnability of meta-categories, and would also enable further refinement of the ontology. This dataset was even smaller when testing on specifically the meme template classification task, because 74 memes could not be included because the annotators marked them as not having a template.

6 Ethical Considerations and Broader Impact

The authors would like to state that while they made efforts to comb through the data and remove memes with questionable or offensive content, they cannot guarantee that every example in this dataset is free from such content, and furthermore would like to state that they do not endorse anything that may be expressed by memes included in this dataset. Beyond the nature of the memes' content, it is also important to consider that the sources of the scraped memes (imgflip and the r/memes subReddit) may limit perspectives in terms of their respective user-bases, and future work should endeavor to capture memes from a wider variety of creators.

References

Cedric Courtois and Thomas Frissen. 2023. Computer vision and internet meme genealogy: An evaluation of image feature matching as a technique for pattern detection. *Communication Methods and Measures*, 17(1):17–39.

Cole Gleason, Amy Pavel, Xingyu Liu, Patrick Carrington, Lydia B. Chilton, and Jeffrey P. Bigham. 2019. Making memes accessible. In *Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility*, ASSETS '19, page 367–376, New York, NY, USA. Association for Computing Machinery.

EunJeong Hwang and Vered Shwartz. 2023. MemeCap: A dataset for captioning and interpreting memes. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 1433–1445, Singapore. Association for Computational Linguistics.

Rahima Khanam and Muhammad Hussain. 2024. YOLOv11: An Overview of the Key Architectural Enhancements. *Preprint*, arXiv:2410.17725.

Ryunosuke Noda, Chiaki Yuasa, Fumiya Kitano, Daisuke Ichikawa, and Yugo Shibagaki. 2025. Performance of o1 pro and GPT-4 in Self-Assessment Questions for Nephrology Board Renewal. *medRxiv*.

Z. Wang, E.P. Simoncelli, and A.C. Bovik. 2003. Multiscale structural similarity for image quality assessment. In *The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*, volume 2, pages 1398–1402 Vol.2.

A Annotation Guidelines

For each annotation in the project, perform the following steps:

1. Transcribe the text in the meme. Each section of text (either split up as labels, lines of dialogue, or other obvious spacing) should be typed into the box followed by the enter key.

2. Attempt to identify the meme template. If it is immediately recognized, find the corresponding link to the **template** (i.e imgflip.com/memetemplate/... and not imgflip.com/memegenerator/...) if one exists. Otherwise if it is unrecognized, query search engines with a description of the image and the word "meme." If all else fails, obtain the link to the source image (through Label Studio) and input it into Google's search by image feature to obtain a name or description that would lead one to an associated imgflip template. Note that in the case of multiple templates in an image, one should choose a single

template that is ideally the most prominently featured and most easily recognizable in the image. If there are no recognizable templates in the image, leave the link section blank.

3. Choose one of the associated meta-categories according to the ontology. While it is possible to leave this selection blank, it is generally recommended to avoid this unless one feels very strongly that none of the associated meta-categories fit at all.

A.1 Image Macro Example

This example involves an image of a cat in a suit sitting at a table as if it were looking up from reading the newspaper with a shocked expression. With the image as the context, the caption continues the (presumably meta) joke by delivering an announcement about banned meme posts.

A.2 Reaction Image Example

In this example, the joke is derived from the situation set up in the caption "when you donate 1$ to a streamer" and the punchline in the reaction image of Walter White from *Breaking Bad* saying "Now say my name."

A.3 Exploitable Example

This example is using a scene from the cartoon *Scooby Doo, Where Are You!* in which the character Fred is revealing the identity of the villain. The masked villain is labeled as "Being able to cook," and then once unmasked, labeled as "Just following the recipe."

A.4 Duality Example

This example shows the duality between two situations in a video game. The first one is where the player of a game is able to heal their character several times, and the second is where the boss heals a single time. The images of Mordecai from the cartoon *Regular Show* express a neutral-positive reaction to the first situation and a disgusted reaction to the second.

A.5 Escalating Progression Example

The example shown here demonstrates the escalating progression of deciding when to purchase and play a video game. The first context, "pre-order a game" is the neutral ("good") state, associated with the normal image of Winnie the Pooh. The second, "buy the game when it releases," is the next step up ("better"), and the corresponding image is Pooh in fancy dress. The final context, "wait almost a decade so you can get it for free," is the "best" decision, and as a result the corresponding Pooh is dressed in even fancier clothing to communicate this.

A.6 "None" meta-category Example

This example is not much more than a picture with a caption. There does not appear to be any larger

structural relationship between the text and image beyond that.

B Differences Between Imgflip and Reddit Memes

Often, memes that are obtained from sources other than generator websites (i.e "in the wild") look very different from the prototypical template image from which the meme is derived. As an illustrative example, consider the "Sad Hulk" meme. The template image is that of two frames of the Incredible Hulk with tears in his eyes. Typically, examples of this meme will use the exact same image of the Hulk, only having the text changed.
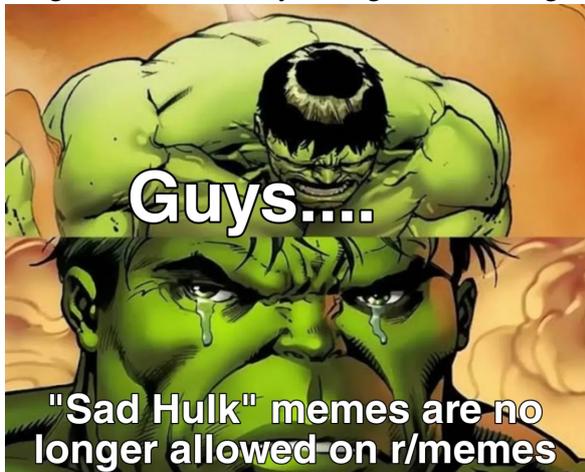
This is the same image as that which is found on the Crying Hulk template on imgflip. However, there are other examples that are clearly referencing the same template without using the exact same picture.

Template	Confidence
odd1sout-vs-computer-chess	0.09
Running-Away-Balloon	0.09
hello-human-resources	0.05
Apu-takes-bullet	0.04
how-i-sleep-homer-simpson	0.04
American-Chopper-Argument	0.10
thanos-what-did-it-cost	0.06
Squidward	0.04
Hide-the-Pain-Harold	0.04
Out-of-line-but-hes-right	0.04

Table 6: Top 5 template labels by confidence from the YOLO model for the Incredible Krunk (top) and Hulk Hogan (bottom) examples.

While both of these examples are using very different images from the template, they are still clearly derived from the original in terms of the joke being made (especially since the first picture is a parody of the Hulk and the second is a picture of Hulk Hogan). Given how different these images are compared to any output from the Crying Hulk meme generator, it stands to reason that these would be especially difficult for a simple CV model to determine that they should be classified as the same template. Accordingly, our YOLO model did not put Crying Hulk in the top 5 for either example.

C Meta-category Edge Cases

Sometimes it is not clear which meta-category label is appropriate for a given meme. This section

provides some examples encountered by the annotators along with discussion of the relevant labels to hopefully elucidate what the ontology is trying to capture.

C.1 Image Macro vs. Reaction

This meme structurally resembles a reaction (the text on top separate from the image) but the actual relationship between the text and image is that of an image macro because the joke is a play on a well-known phrase from the character pictured.

C.2 Reaction vs. Exploitable

Even though the image is from an established piece of media (the Thriller music video), the structure in this meme comes from outside the scene. The image essentially functions as a reaction to the setup in the text.

C.3 Exploitable vs. Reaction

While the image on the bottom of the meme is often deployed as a reaction, in this case the structure of the meme is derived from multiple frames of the actual scene, so it is an exploitable.

C.4 Duality vs. Exploitable
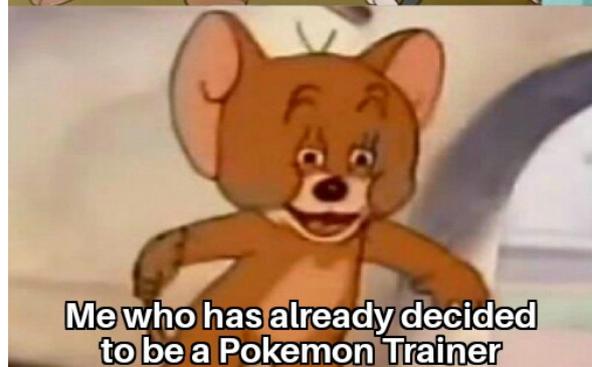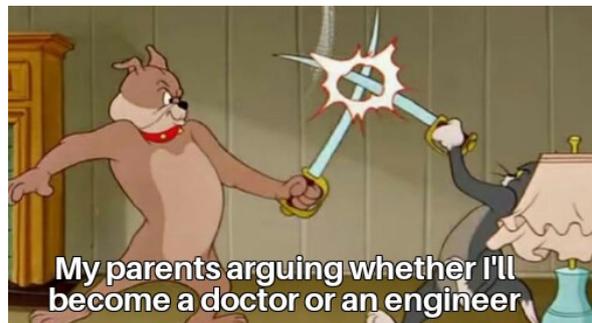
Even though the source for the images in this meme is the TV show Tom and Jerry, it is not considered an exploitable because the two frames are from entirely different parts of the TV show, and the

structure of the meme is as a result not derived from the scene. The duality comes from the comparison between the two image-caption pairs.

C.5 Duality vs. Escalating progression

You have a very specific problem

You search Google

You find a 10 y/o Reddit post with the same question

0 comments

The images in this meme communicate two opposing emotional states occurring across the various stages described by the captions. Rather than being a continuum, they express discrete negative and positive emotions. As a result, despite having more than two parts, this meme is an example of duality and not escalating progression.

D ChatGPT 4o API Prompts

Initial prompt: "You will be given a list of example images for image macro, reaction, escalating progression, duality, and exploitable memes, and afterwards you will have to classify them"

Final prompt: "Given the descriptions of memes from before, how would you describe the following meme? Select a response for the following list and only use the words from this list: image macro, reaction, escalating progression, duality, or exploitable. Only output the template class you decided."

D.1 Experiment 1: Text Prompts

Image Macros: "Image macros can be thought of as memes for which the setup of the joke is given by the image, and the text fills in the details and punchline. For our purposes, we consider image macros as consisting of a single image. The most typical presentation of these memes involves a picture of some kind of entity (often an animal or person) with white text in impact font on the top and bottom of the image."

Reaction Images: "Reaction images are in a sense an inversion of image macros. Instead of the image setting up the joke and the text filling in the punchline, the joke is set up by a text caption (almost always separated from the image portion) and then the punchline is the image. The images in these memes are usually of people emoting or reacting in some fashion, and the humor often derives in part from the fact that the text caption completely recontextualizes the image from its origin."

Escalating Progression: "Escalating progression memes are those that express a reaction to points sampled along a continuum of context. Unlike in duality where the situations are related in terms of opposing each other (e.g good thing vs. bad thing), with escalating progression there is an intensification between each state (e.g good thing vs. better thing vs. best thing). Typically, these memes have at least 3 sets of image-context pairs, but it is not required as long as there is an obvious sense of continuity between the contexts. It does not have to be 3+ images, like it can still be one, the difference is that it represents something escalating, rather than a good or bad. It has a continuum moving in a consistent and intensifyingly humorous direction."

Duality: "A meme exhibiting duality is one that compares two (or more) situations or contexts that are related across some dimension, usually in opposition to each other. The typical components of a duality meme are a set of discrete contexts and a set of images (most often variations on the same image) that visualize the relationship between the contexts. A common format is a '4-panel' layout in which two pairs of contexts and images are stacked vertically."

Exploitable: "An exploitable is a meme in which an existing image (such as a comic, or one or more scenes from a movie) of some sort is augmented by adding and/or replacing some set of things (like dialogue, characters, labels etc.) to tell a joke. The idea is that there is some extant structure inside an image that is 'exploited' using text or additional pictures within the bounds of the original image (unlike a caption in a reaction image, which is typically outside it)."

D.2 Experiment 2: Image Prompts

Image Macros: "Here is an example of an image macro, where you have the classic text on top and

below, where the entire joke could be understood without the image:"

url: "https://imgflip.com/i/3yhgyo"

Reaction Images: "Here is an example of a reaction, where the monkey looks awkwardly in reaction to the text which follows a 'me-when' style:"

url: "https://i.imgflip.com/4zv2v9.jpg"

Escalating Progression: "Here is an example of an escalating progression, where the brain gets bigger and bigger as the text describes something smarter and smarter:"

url: "https://i.imgflip.com/4iyi3q.jpg"

Duality: "Here is an example of an duality, where drake at first thinks it's bad, but then thinks it's good:"

url: "https://i.imgflip.com/4izfsm.jpg"

Exploitable: "Here is an example of an exploitable, where the text is overlayed over all of the people, representing who they are in the context of the joke:"

url: "https://i.imgflip.com/3fys88.jpg"