AUTOMATIC JOKE GENERATION:
LEARNING HUMOUR FROM EXAMPLES

Master thesis by: Thomas Winters

Promotor: Prof. Danny De Schreye

Daily supervisor: Vincent Nys

ABOUT ME: THOMAS WINTERS



Student
CS: HCI

PREPAREE
IMPROTHEATER

Satirical
Cartoons
@
veto





CONTRIBUTIONS

1

GOOFER FRAMEWORK

2

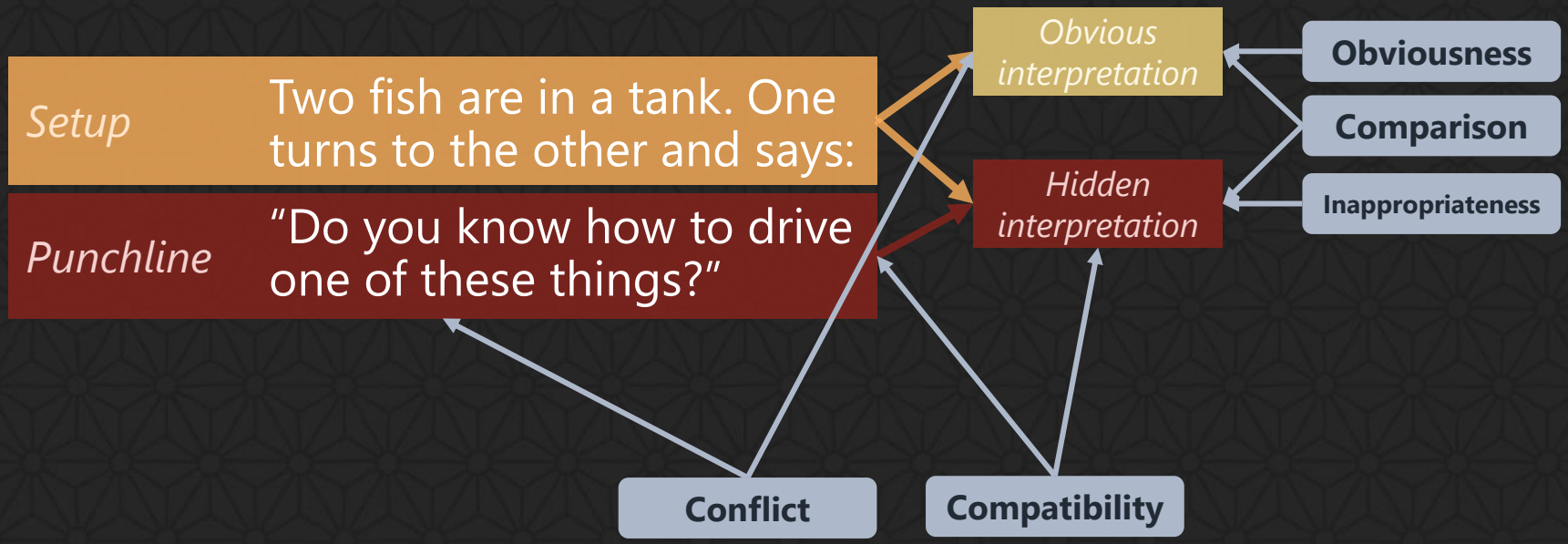
GAG
GENERALISED ANALOGY GENERATOR

3

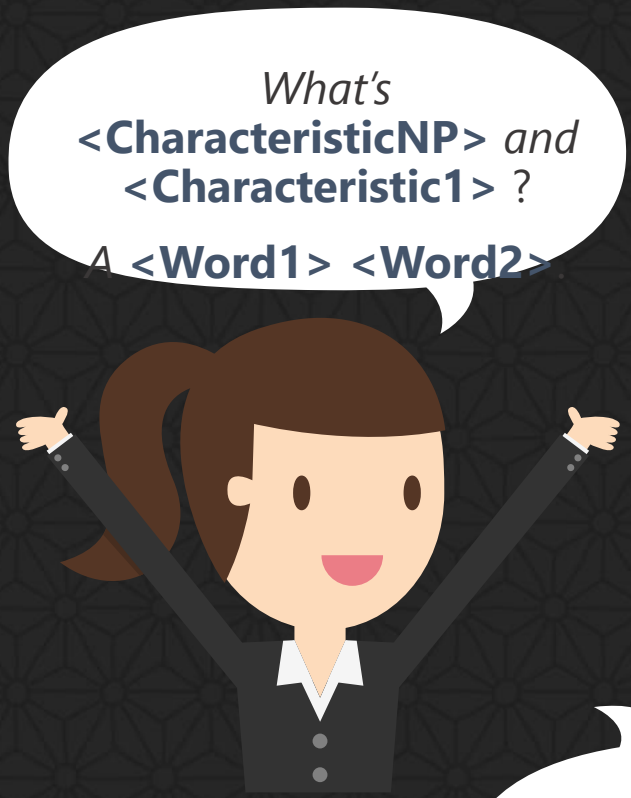
STATE OF THE ART

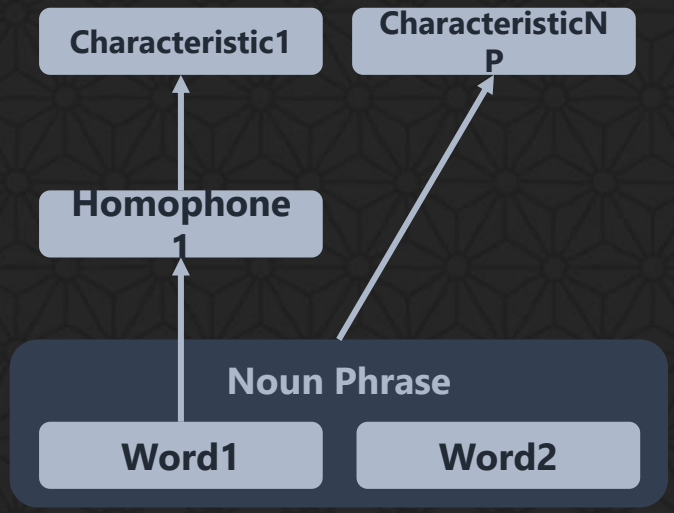
INCONGRUITY RESOLUTION THEORY

(ACCORDING TO GRAEME RITCHIE)



JAPE: TEMPLATES & SCHEMAS

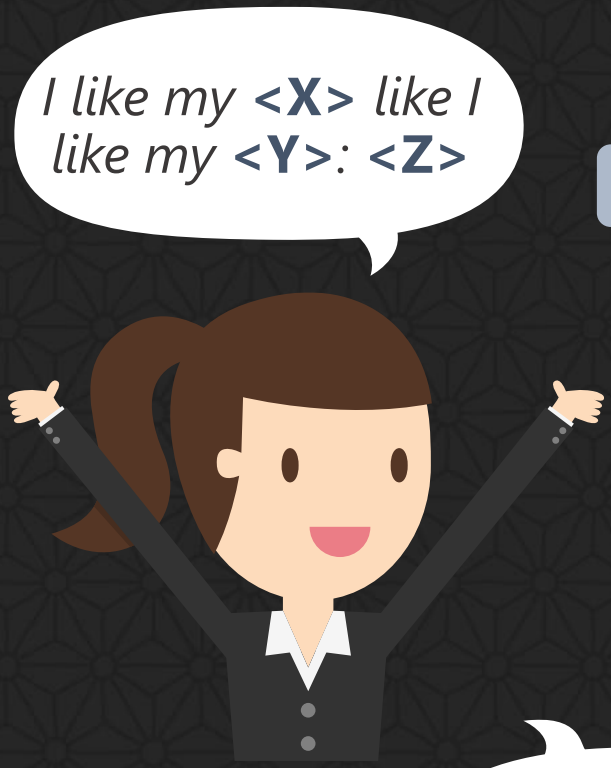


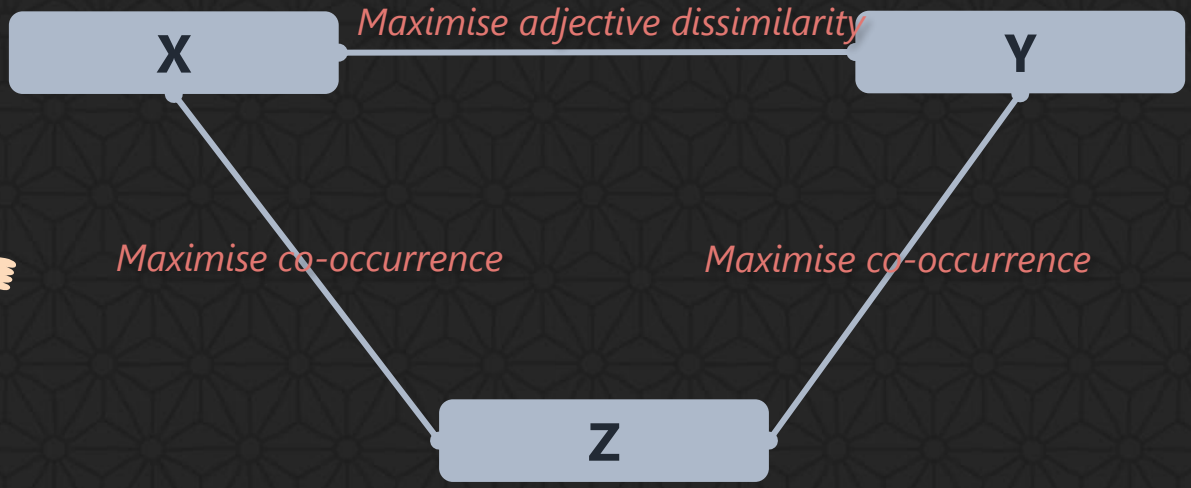


What's **green** *and* **bounces**?

A **spring cabbage**.

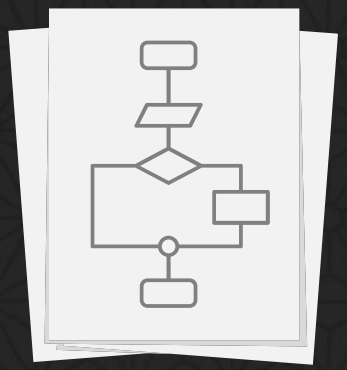
PETROVIC & MATTHEWS ANALOGY GENERATOR



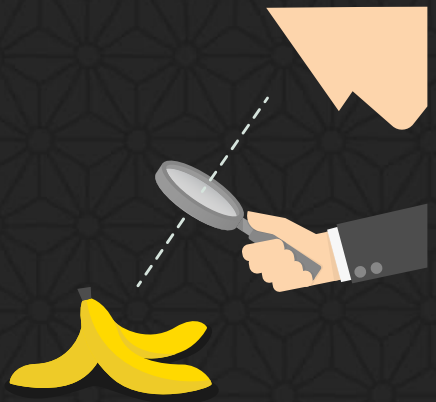


I like my **coffee** *like I like my* **war**: **cold**

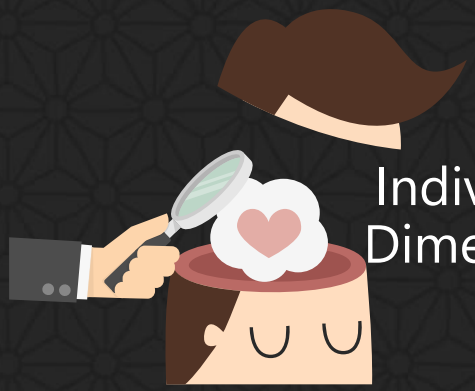
STATE OF THE ART PROBLEMS



Hardcoded Templates & Schemas



No Reinforcement Learning



Individual Dimension



Temporal Dimension

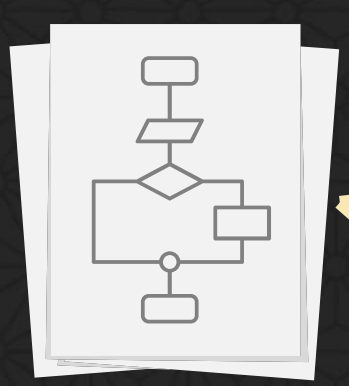
GOOFER FRAMEWORK

(Generator Of One-liners From Examples with Ratings)

SCHEMA GENERALISATION

Constraint-based Schema

Probabilistic Schema



Relations to Boolean functions

All functions return true

Lexicon combinations



Metric functions



Metric aggregator

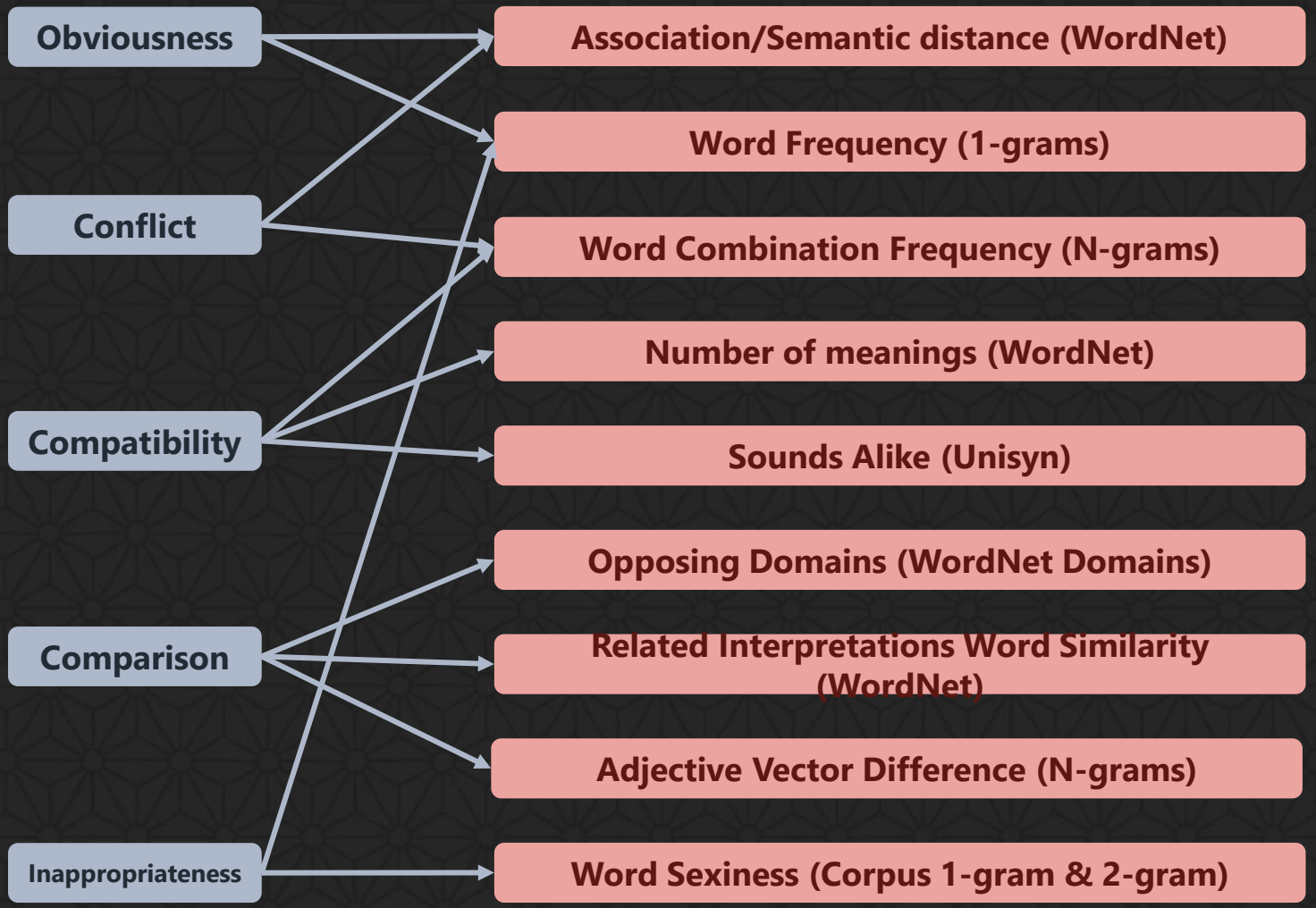
Classification as aggregator

→ ***Learning component!***



Values Generator

HUMOUR METRICS IDENTIFICATION



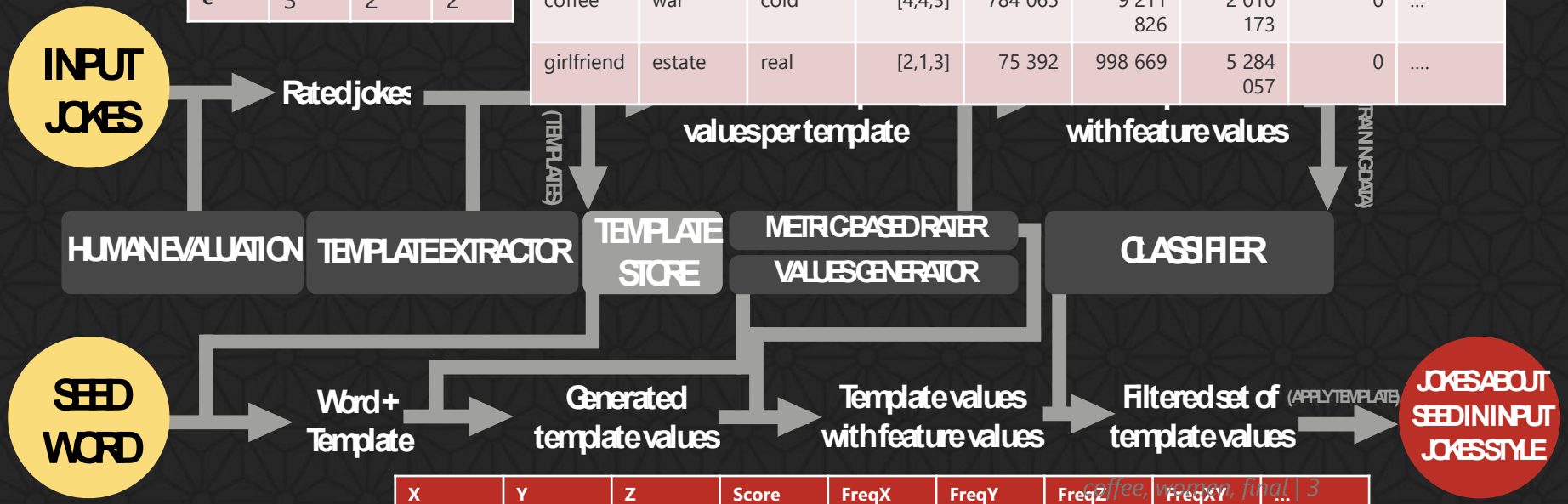
PIPELINE

a) I like my relation
open.
b) I like my coffee li
c) The sailor bears
d) I like my girlfrien
e) The lord cultivate

Joke	Rater 1	Rater 2	Rater 3
a	5	4	5
b	4	4	3
c	5	3	4
d	2	1	3
e	3	2	2

Template 1: "I like my X like I like my Y, Z."
relations, source, open | [5,4,5]
coffee, war, cold | [4,4,3]
girlfriend, estate, real | [2,1,3]

X	Y	Z	Score	FreqX	FreqY	FreqZ	FreqXY	...
relations	source	open	[5,4,5]	3 831 210	4 050 904	4884757	0	...
coffee	war	cold	[4,4,3]	784 065	9 211 826	2 010 173	0	...
girlfriend	estate	real	[2,1,3]	75 392	998 669	5 284 057	0	....



coffee, women

coffee, women | Template 1

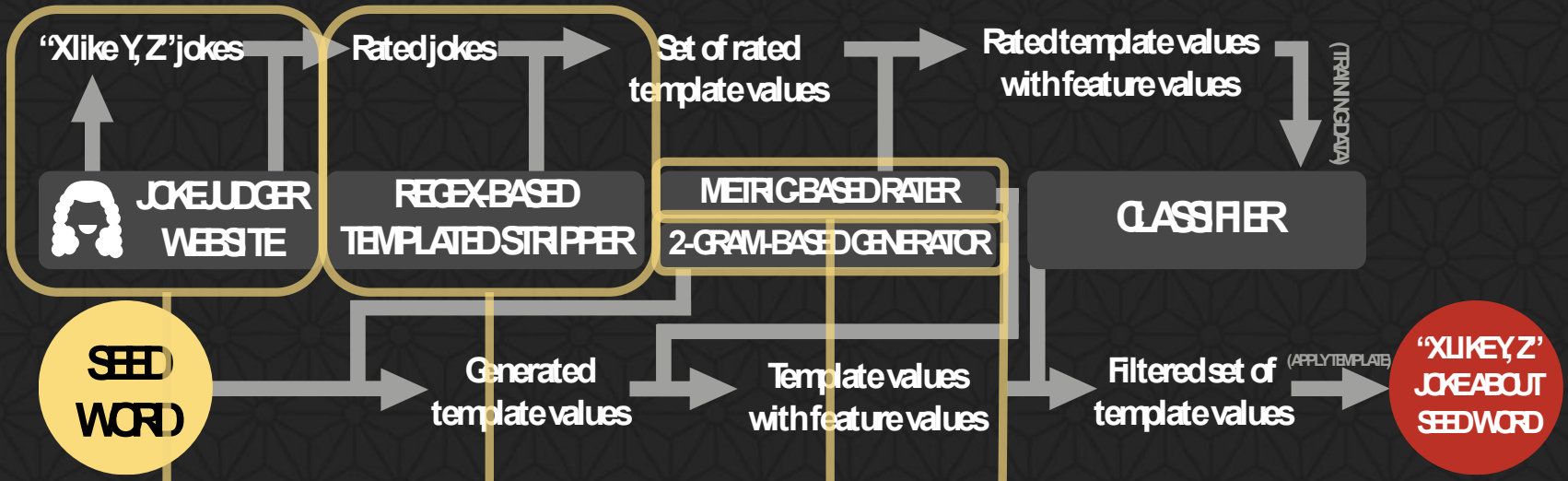
X	Y	Z	Score	FreqX	FreqY	FreqZ	FreqXY	...
coffee	women	final	?	784 065	9 929 965	3 070 183	0	...
coffee	women	black	?	784 065	9 929 965	5 707 260	0	...
coffee	women	popular	?	784 065	9 929 965	2 311 546	0	....
...	...	...	?	...	...	...	...	...

coffee, women, final | 3
coffee, women, black | 5
coffee, women, popular | 2
coffee, women, rich | 2
coffee, women, Irish | 4

I like my women: blac
I like my women: Irish

GENERALISED ANALOGY GENERATOR

GENERALISED ANALOGY GENERATOR



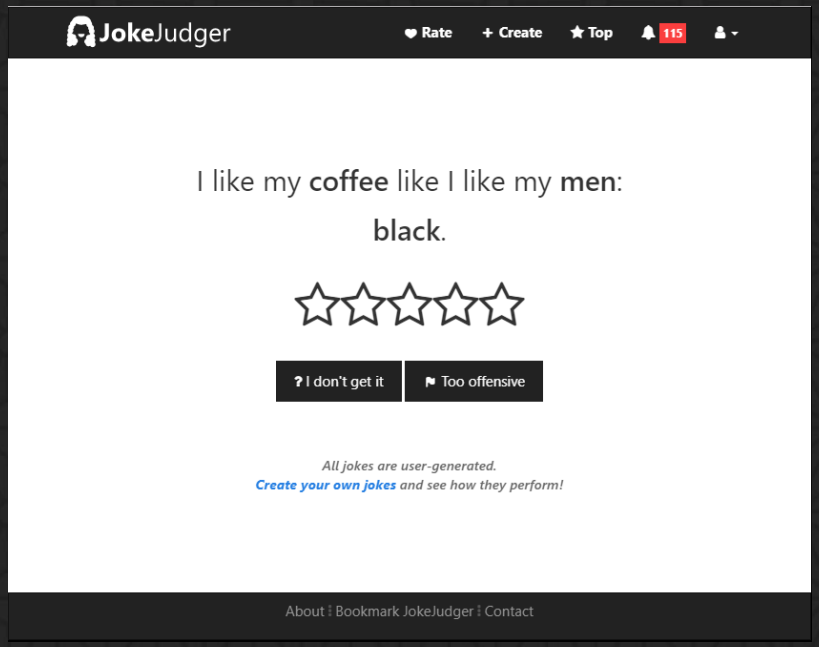
2-gram-based generator

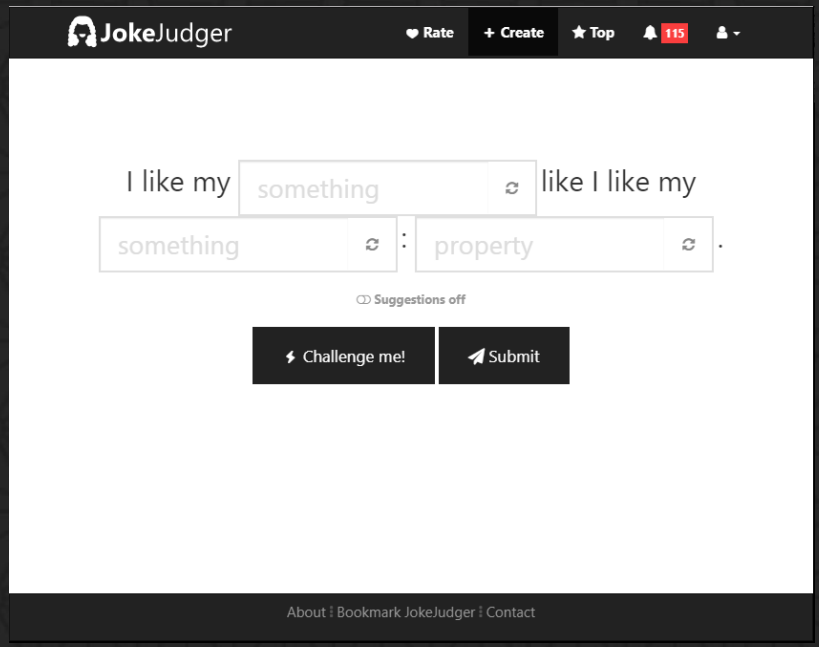
Selected metrics

Simplified Template Extraction: one template

Human Evaluation component also as humour source

JOKEJUDGER


JokeJudger
Rate
+ Create
Top
115
I like my coffee like I like my men:
black.
I don't get it
Too offensive
All jokes are user-generated.
Create your own jokes and see how they perform!
About ⁞ Bookmark JokeJudger ⁞ Contact


JokeJudger
Rate
+ Create
Top
115
I like my something like I like my
something : property .
Suggestions off
Challenge me!
Submit
About ⁞ Bookmark JokeJudger ⁞ Contact

JOKEJUDGER: COLLECTED DATA



203 users



9034 ratings
418 markings



424 human jokes
100 generated jokes



41,9% average agreement on most frequent rating

CLASSIFICATION & REGRESSION



Classification

most frequent rating



Regression

Average rating



5 star jokes

= Small number & High variety

➞ Hard to predict

Humour appreciation varies a lot

➞ Average score close to centre

RANDOM FOREST

Random Trees: **subsets** of features & data

→ **Reduces impact** of noisy features

Classification
61.5% correctly classified

Regression
RRSE: 84%

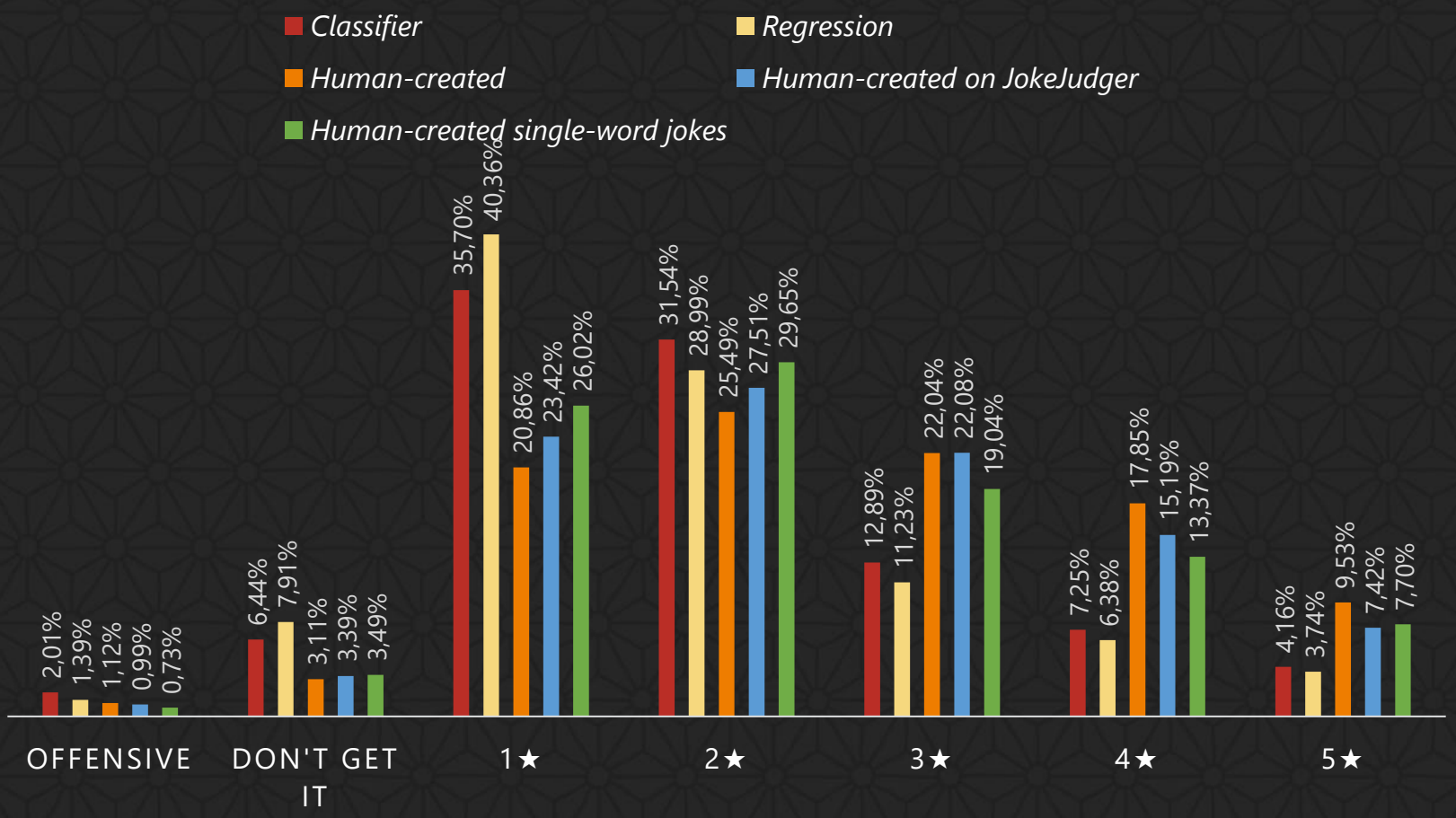




frequency_0 | frequency_1 | frequency_2 | word_senses_0 | word_senses_1 | word_senses_2 | ...

EVALUATION

RELATIVE NUMBER OF RATINGS PER RATING TYPE FOR DIFFERENT SOURCES AND CATEGORIES OF JOKES DURING THE EVALUATION PHASE



CONCLUSION

PROGRAM CONTRIBUTIONS

GOOFER
FRAMEWORK

GAG
GENERALISED
ANALOGY
GENERATOR





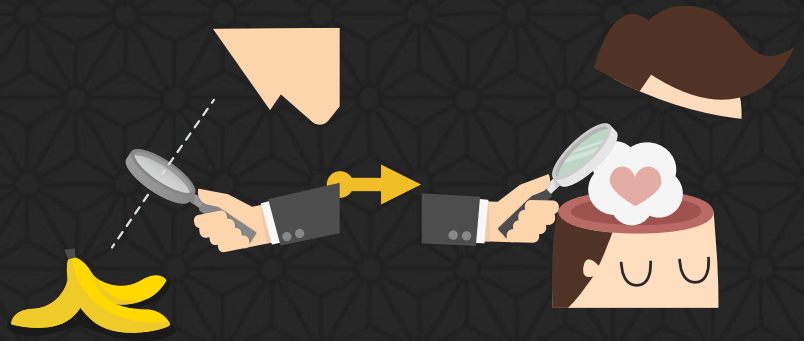
NLP/CH Tools on
github.com/TWinters

CONCLUSION



4+ RATINGS FREQUENCY





DEMO & DISCUSSION