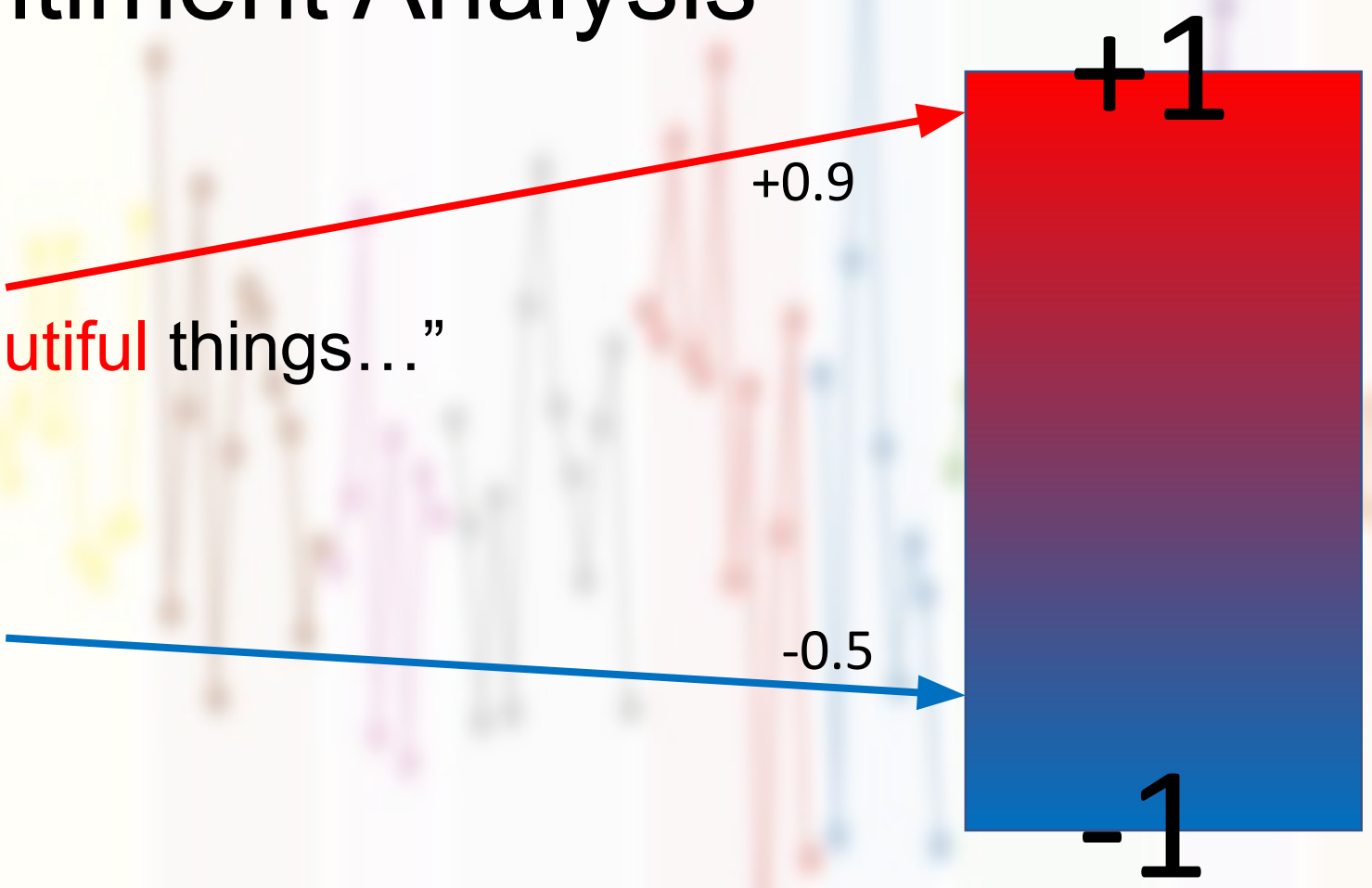
Sentiment analysis

Sentiment Analysis

“This book is fantastic!
It tells wonderful and beautiful things…”

“In those days, sadness
shadowed my heart”



Sentiment Analysis

"[…] is the field of study that analyzes people's opinions, sentiments, appraisals, attitudes, and emotions towards entities and their attributes expressed in written text. The entities can be products, services, organizations, individuals, events, issues, or topics"
(Liu, 2015)

Sentiment Analysis

Matthew L. Jockers

Search... Go

05 Thursday Jun 2014

A Novel Method for Detecting Plot

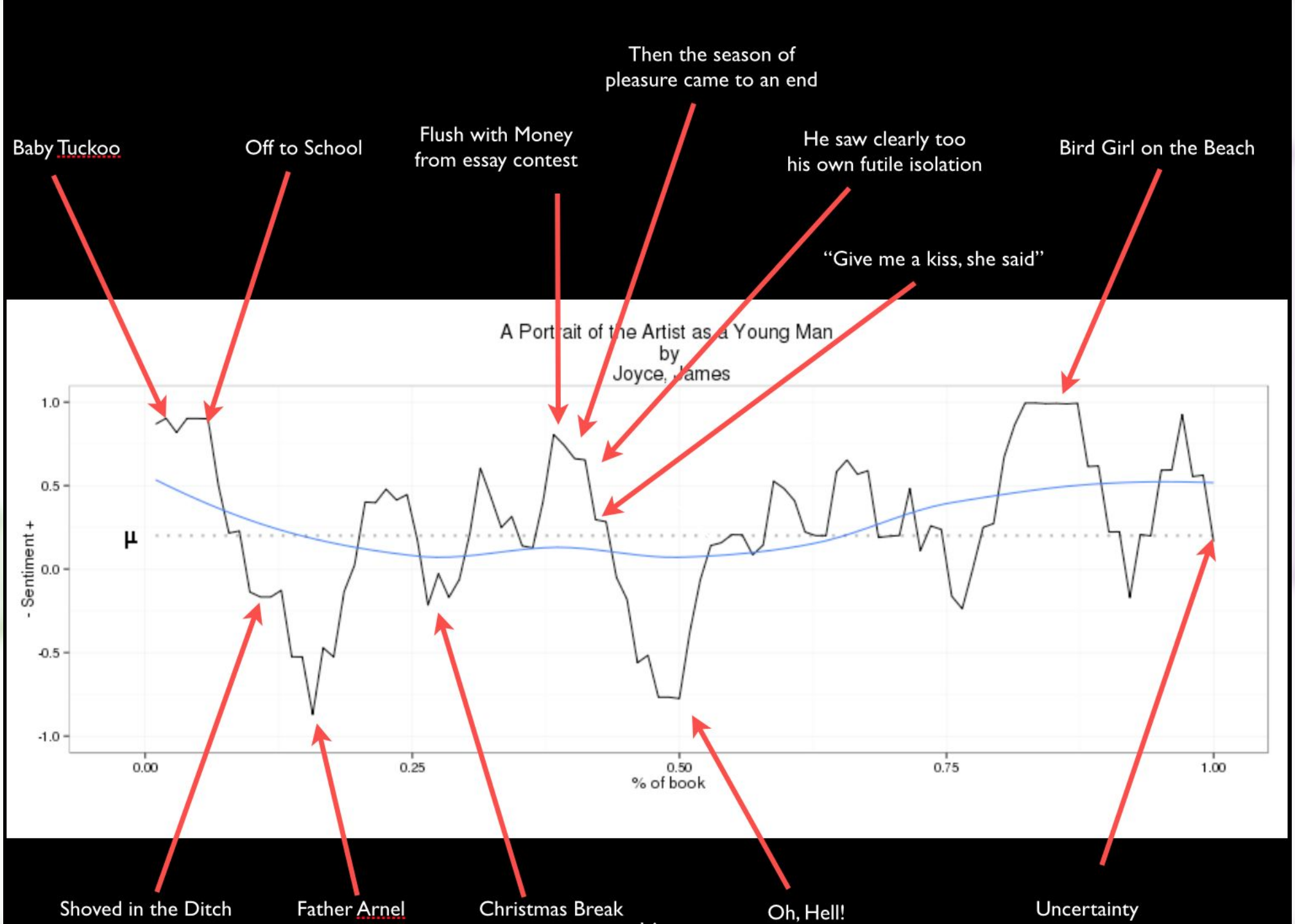
Posted by Matthew Jockers in Commentary, Text-Mining ≈ Comments Off

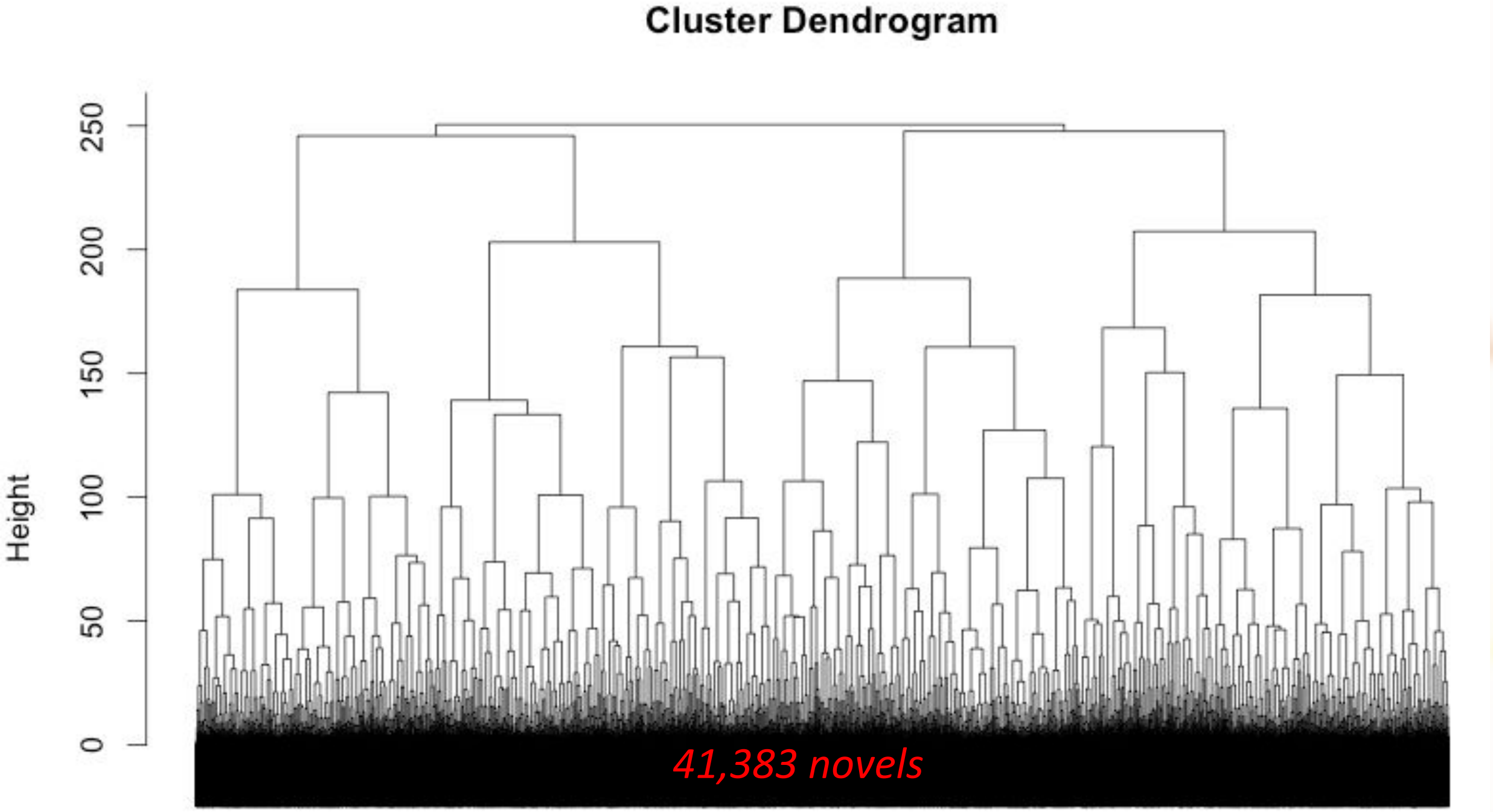
While studying anthropology at the University of Chicago, Kurt Vonnegut proposed writing a master's thesis on the shape of narratives. He argued that "the fundamental idea is that stories have shapes which can be drawn on graph paper, and that the shape of a given society's stories is at least as interesting as the shape of its pots or spearheads." The idea was rejected.

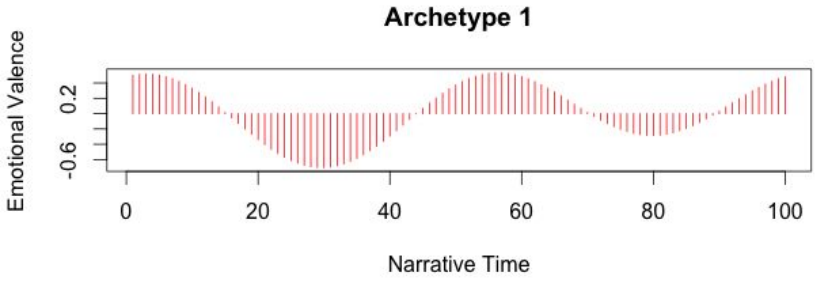
In 2011, Open Culture featured a video in which Vonnegut expanded on this idea and suggested that computers might someday be able to model the shape of stories, that is, the movement of the narratives, the plots. The video is about four minutes long; it's worth watching.

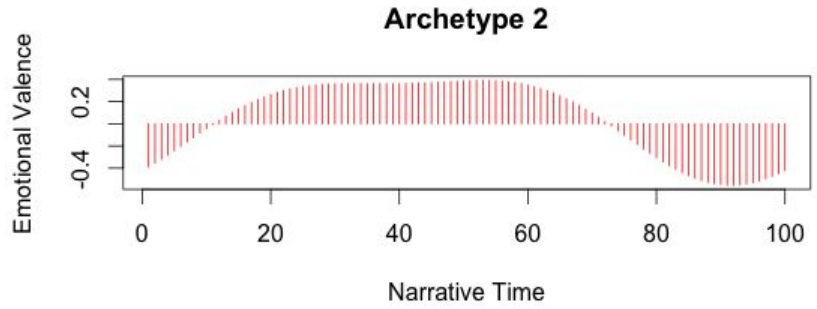
♣ CONTACT

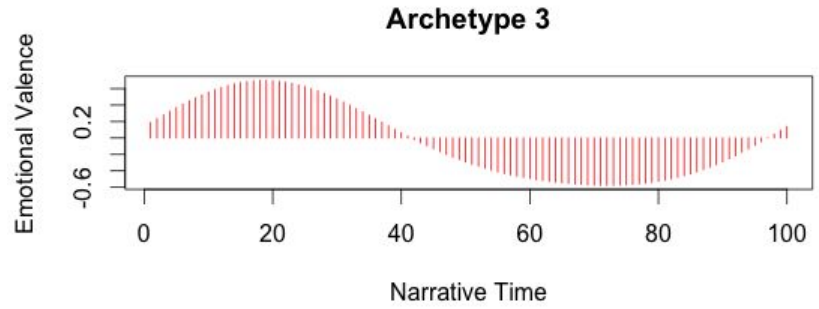
Matthew L. Jockers
Susan J. Rosowski Associate Professor of English
Department of English
University of Nebraska-Lincoln
Lincoln, NE 68588

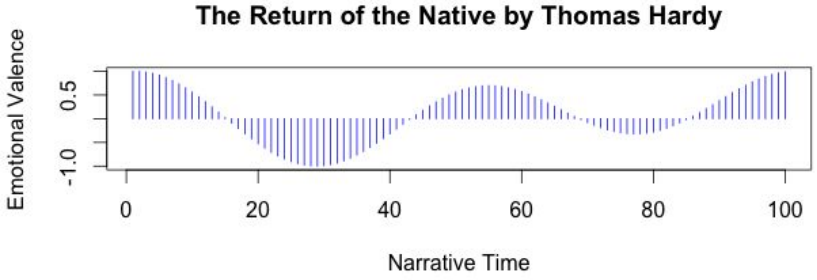

Then the season of pleasure came to an end
Baby Tuckoo
Off to School
Flush with Money from essay contest
He saw clearly too his own futile isolation
Bird Girl on the Beach
"Give me a kiss, she said"
A Portrait of the Artist as a Young Man
by
Joyce, James
1.0
0.5
0.0
-0.5
-1.0
μ
- Sentiment +
0.00
0.25
0.50
0.75
1.00
% of book
Shoved in the Ditch
Father Arnel
Christmas Break
Oh, Hell!
Uncertainty

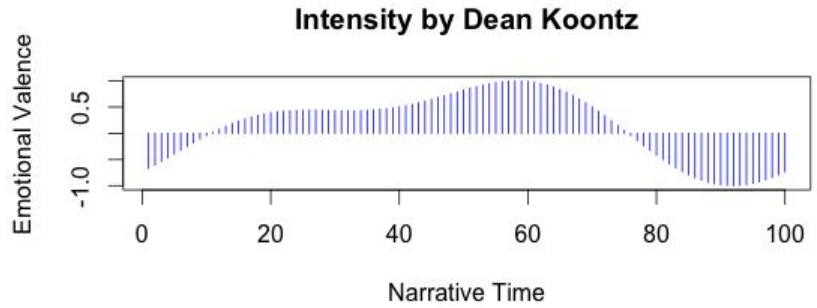

Cluster Dendrogram
Height
0
50
100
150
200
250
41,383 novels

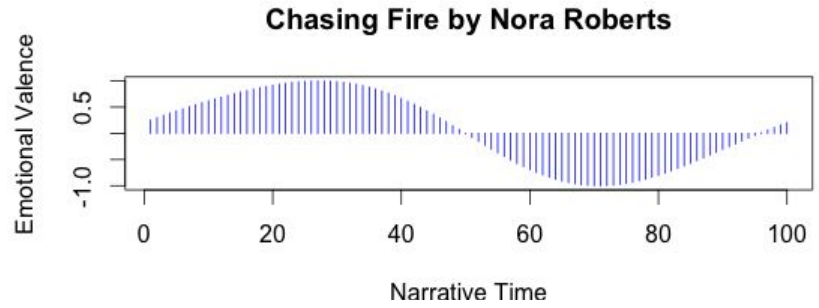

Archetype 1
Emotional Valence
-0.6
0.2
0
20
40
60
80
100
Narrative Time

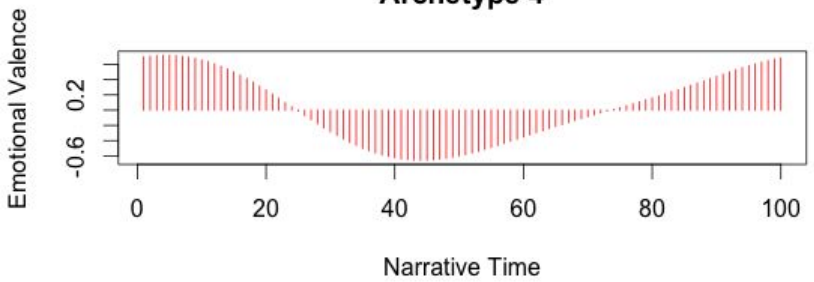

Archetype 2
Emotional Valence
-0.4
0.2
0
20
40
60
80
100
Narrative Time

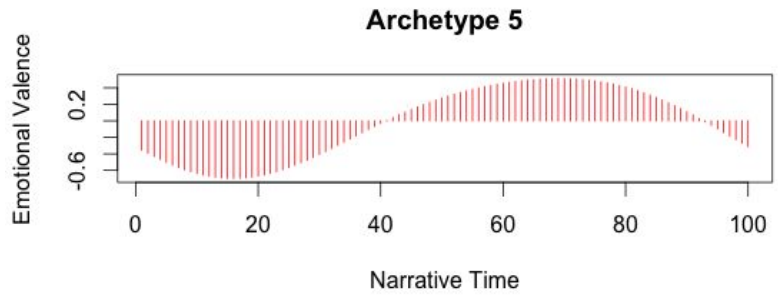

Archetype 3
Emotional Valence
-0.6
0.2
0
20
40
60
80
100
Narrative Time

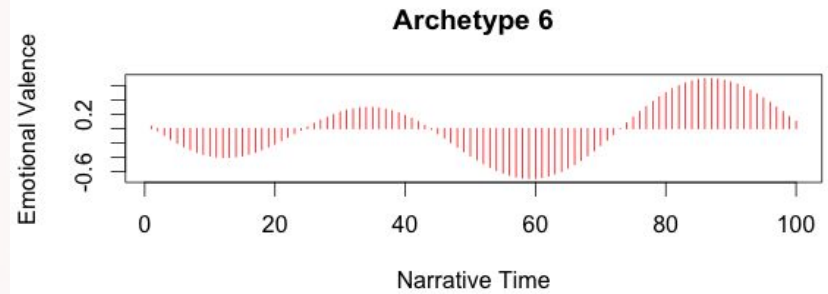

The Return of the Native by Thomas Hardy
Emotional Valence
-1.0
0.5
0
20
40
60
80
100
Narrative Time

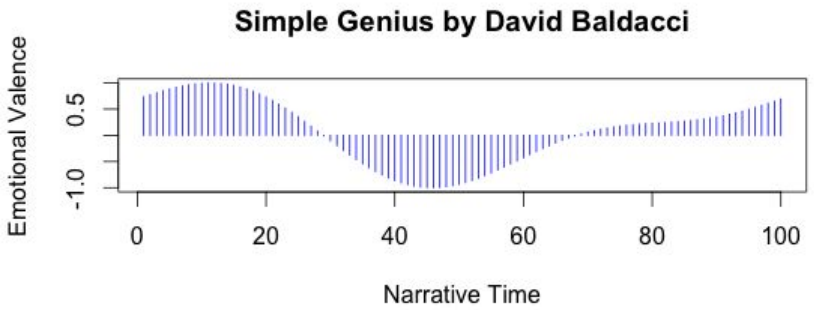

Intensity by Dean Koontz
Emotional Valence
-1.0
0.5
0
20
40
60
80
100
Narrative Time

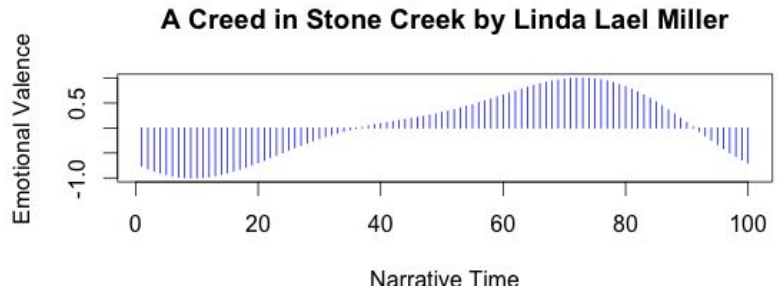

Chasing Fire by Nora Roberts
Emotional Valence
-1.0
0.5
0
20
40
60
80
100
Narrative Time

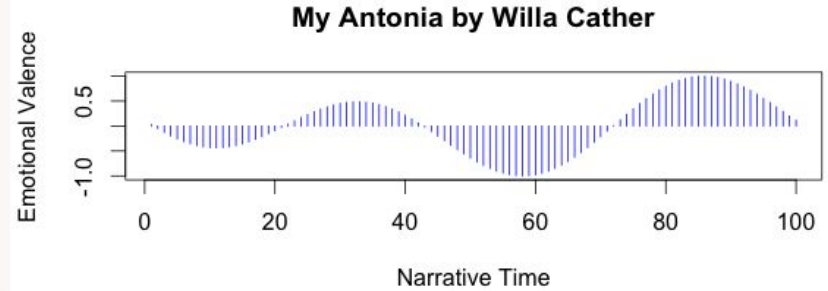

Archetype 4
Emotional Valence
-0.6
0.2
0
20
40
60
80
100
Narrative Time


Archetype 5
Emotional Valence
-0.6
0.2
0
20
40
60
80
100
Narrative Time


Archetype 6
Emotional Valence
-0.6
0.2
0
20
40
60
80
100
Narrative Time


Simple Genius by David Baldacci
Emotional Valence
-1.0
0.5
0
20
40
60
80
100
Narrative Time


A Creed in Stone Creek by Linda Lael Miller
Emotional Valence
-1.0
0.5
0
20
40
60
80
100
Narrative Time


My Antonia by Willa Cather
Emotional Valence
-1.0
0.5
0
20
40
60
80
100
Narrative Time

The Shapes of Stories
by Kurt Vonnegut

Man in Hole

The main character gets into trouble then gets out of it again and ends up better off for the experience.

- Arsenic and Old Lace
- Harold & Kumar Go To White Castle

Boy Meets Girl

The main character comes across something wonderful, gets it, loses it, then gets it back forever.

- Jane Eyre
- Eternal Sunshine of the Spotless Mind

From Bad to Worse

The main character starts off poorly then gets continually worse with no hope for improvement.

- The Metamorphosis
- The Twilight Zone

Which Way Is Up?

The story has a lifelike ambiguity that keeps us from knowing if new developments are good or bad.

- Hamlet
- The Sopranos

Creation Story

In many cultures' creation stories, humankind receives incremental gifts from a deity. First major staples like the earth and sky, then smaller things like sparrows and cell phones. Not a common shape for Western stories, however.

Old Testament

Humankind receives incremental gifts from a deity, but is suddenly ousted from good standing in a fall of enormous proportions.

- Great Expectations with original ending

New Testament

Humankind receives incremental gifts from a deity, is suddenly ousted from good standing, but then receives off-the-charts bliss.

- Great Expectations with revised ending

Cinderella

It was the similarity between the shapes of Cinderella and the New Testament that thrilled Vonnegut for the first time in 1947 and then over the course of his life as he continued to write essays and give lectures on the shapes of stories.

Sentiment Analysis - Critical Aspects

From a **theoretical point of view:**

- Jockers called his software «syuzhet», referring to Russian formalism (see Vladimir Propp) and narratology (the “science of narration”)
- However, traditional narratological studies (see Gérard Genette, Mieke Bal, et al.) do not consider emotions at all
- Only Patrick Colm Hogan wrote a book on *Affective Narratology* (2012), but looking at much more complex phenomena than “plot arcs”

From a **practical point of view:**

- Emotions are subjective (of course!), so their measurement can be unreliable
- SA software are generally unstable: you modify a few parameters, and you get completely different results

Sentiment Analysis

SA and cognitive literary studies

Jacobs, A. M., Schuster, S., Xue, S., and Lüdtke, J. (2017). What's in the brain that ink may character… A quantitative narrative analysis of Shakespeare's 154 sonnets for use in (Neuro-)cognitive poetics. *Scientific Study of Literature*, *7*(1): 4-51.

SA for the study of secondary literature

Mellmann, K. and Du, K. (2018). "Sentimentanalyse in Unstrukturierten Texten (Am Bsp. Literaturgeschichtlicher Rezeptionsanalyse)." In DHd 2018 Konferenzabstracts, 305–8. Cologne: Universität zu Köln.

SA for the study of social reading

Rebora, S. and Pianzola F. (2018). A New Research Programme for Reading Research: Analysing Comments in the Margins on Wattpad. DigitCult - Scientific Journal on Digital Cultures, 3(2): 19–36

SA for Italian language

Sprugnoli, R., Tonelli, S., Marchetti, A., and Moretti, G. (2016). Towards sentiment analysis for historical texts. *Digital Scholarship in the Humanities*, *31*(4): 762-772.

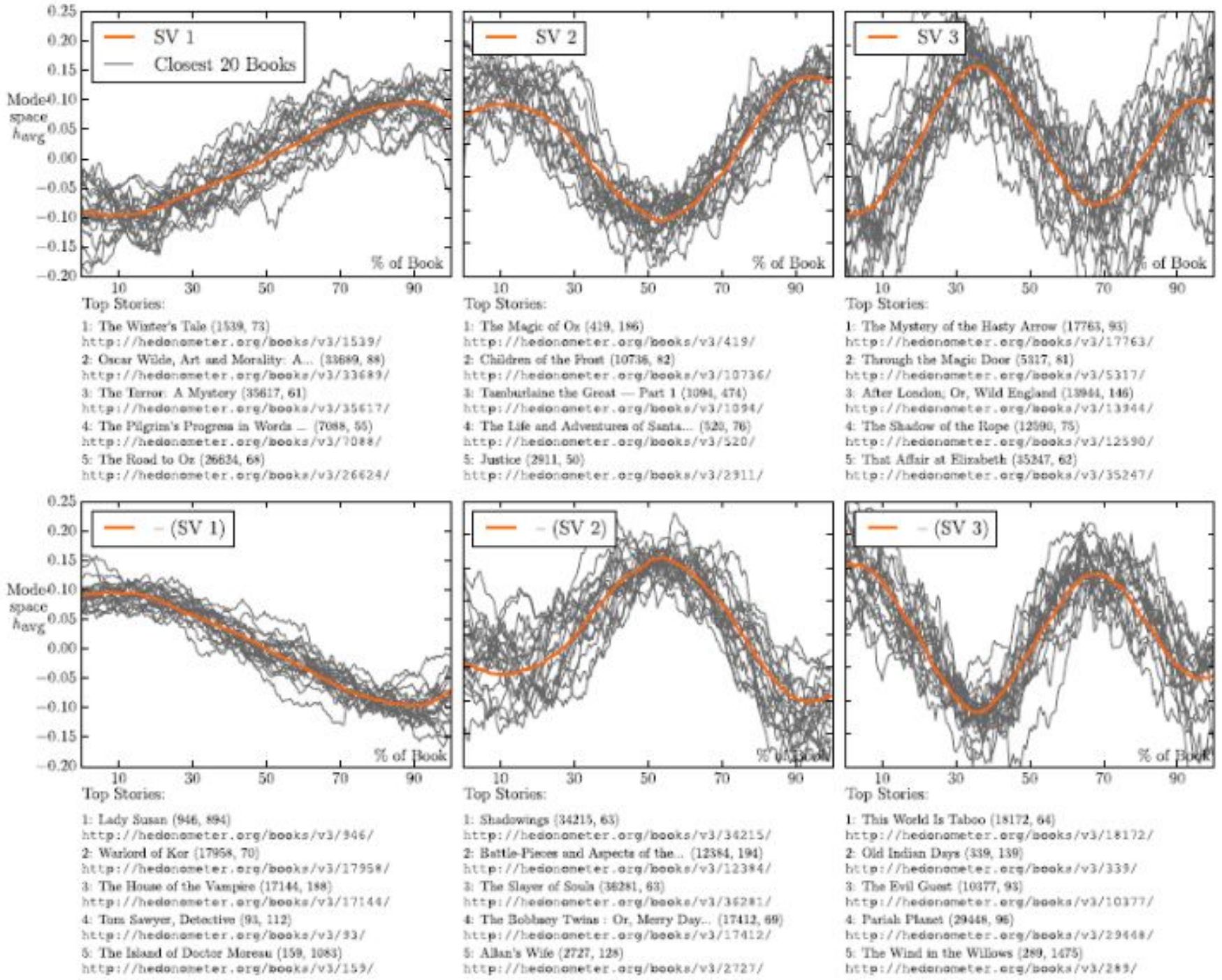
SA for German language

Zehe, A., Becker, M., Jannidis, F., and Hotho, A. (2017). Towards Sentiment Analysis on German Literature. In *Joint German/Austrian Conference on Artificial Intelligence (Künstliche Intelligenz)*. Cham: Springer, pp. 387-394.

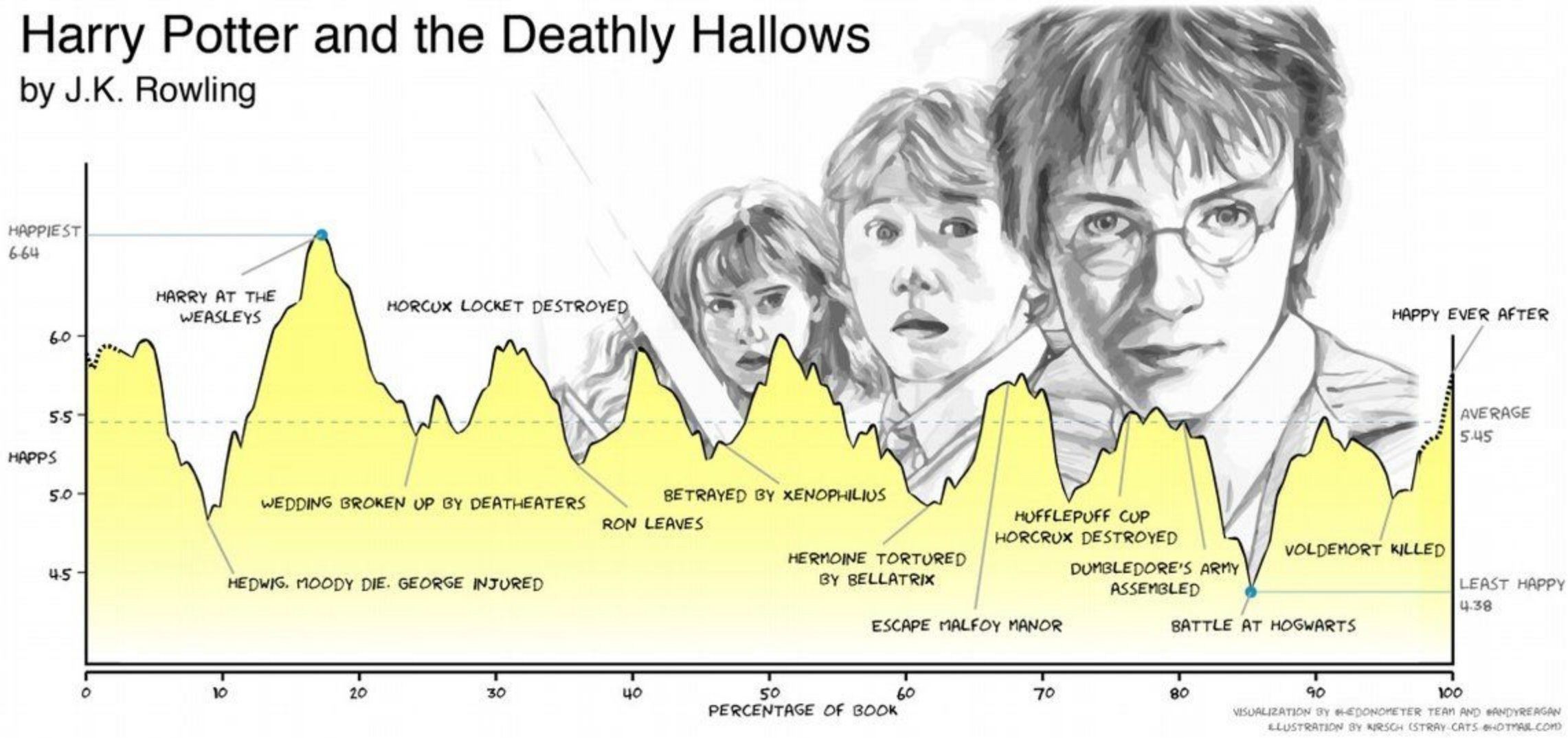
SA for emotional arcs (again)

Reagan, A. J., Mitchell, L., Kiley, D., Danforth, C. M., and Dodds, P. S. (2016). The emotional arcs of stories are dominated by six basic shapes. *EPJ Data Science*, *5*(1): 31.

Andrew J Reagan et al. 2016. “The emotional arcs of stories are dominated by six basic shapes.” *EPJ Data Science*



http://hedonometer.org/



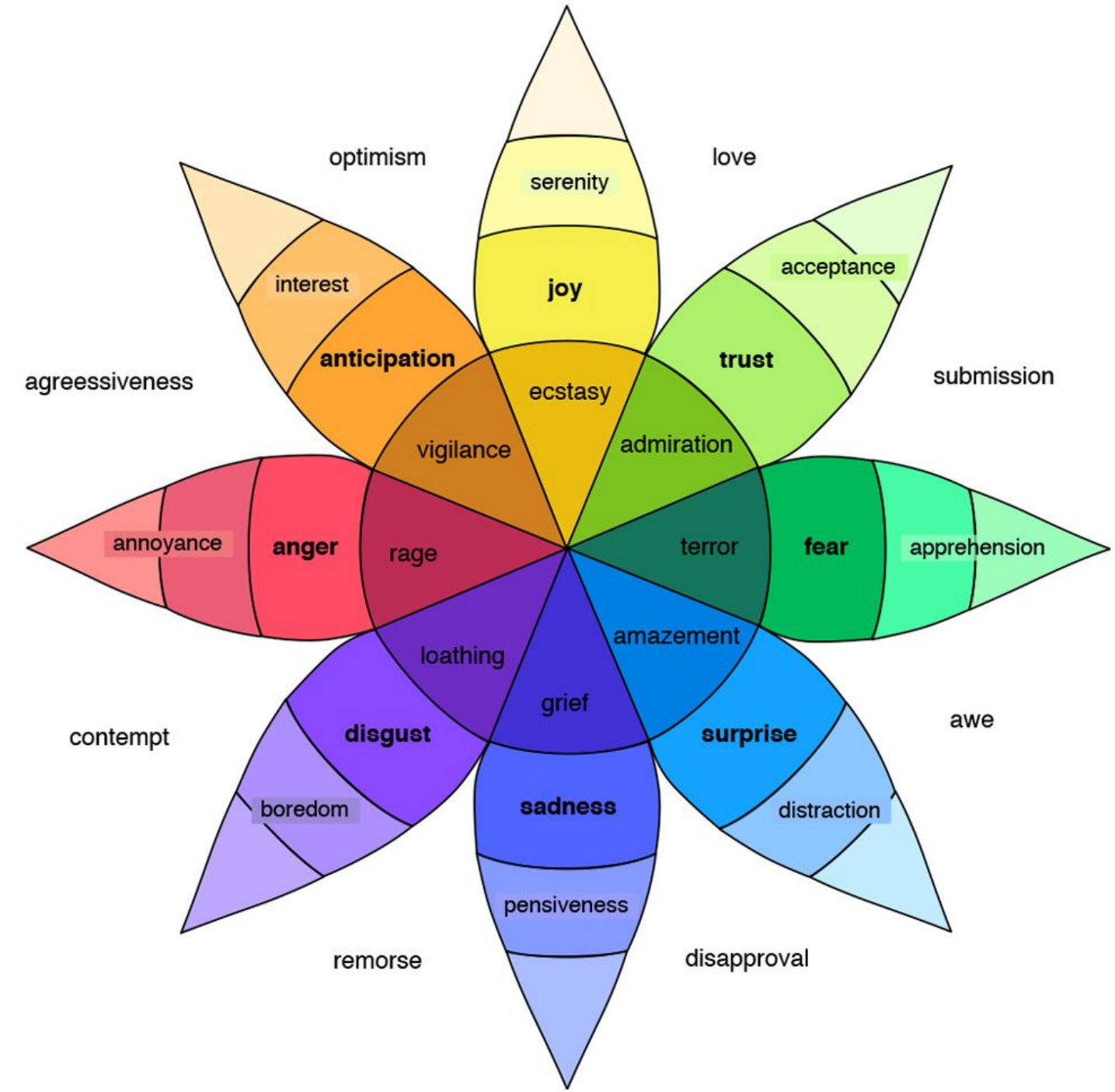
Sentiment Analysis Tools - A (Simple) Taxonomy

Two **defining elements:**

- the emotion theory adopted by the tool
- the technique to accomplish the analysis

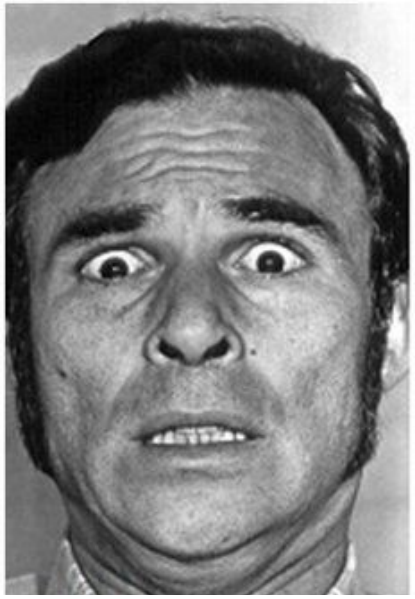
Emotion Theories

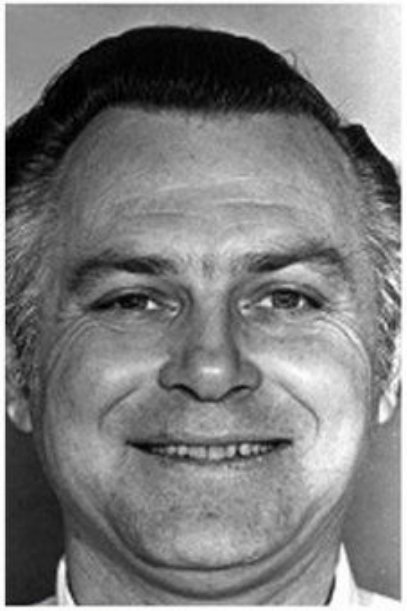
1. Plutchik's eight basic emotions



Emotion Theories

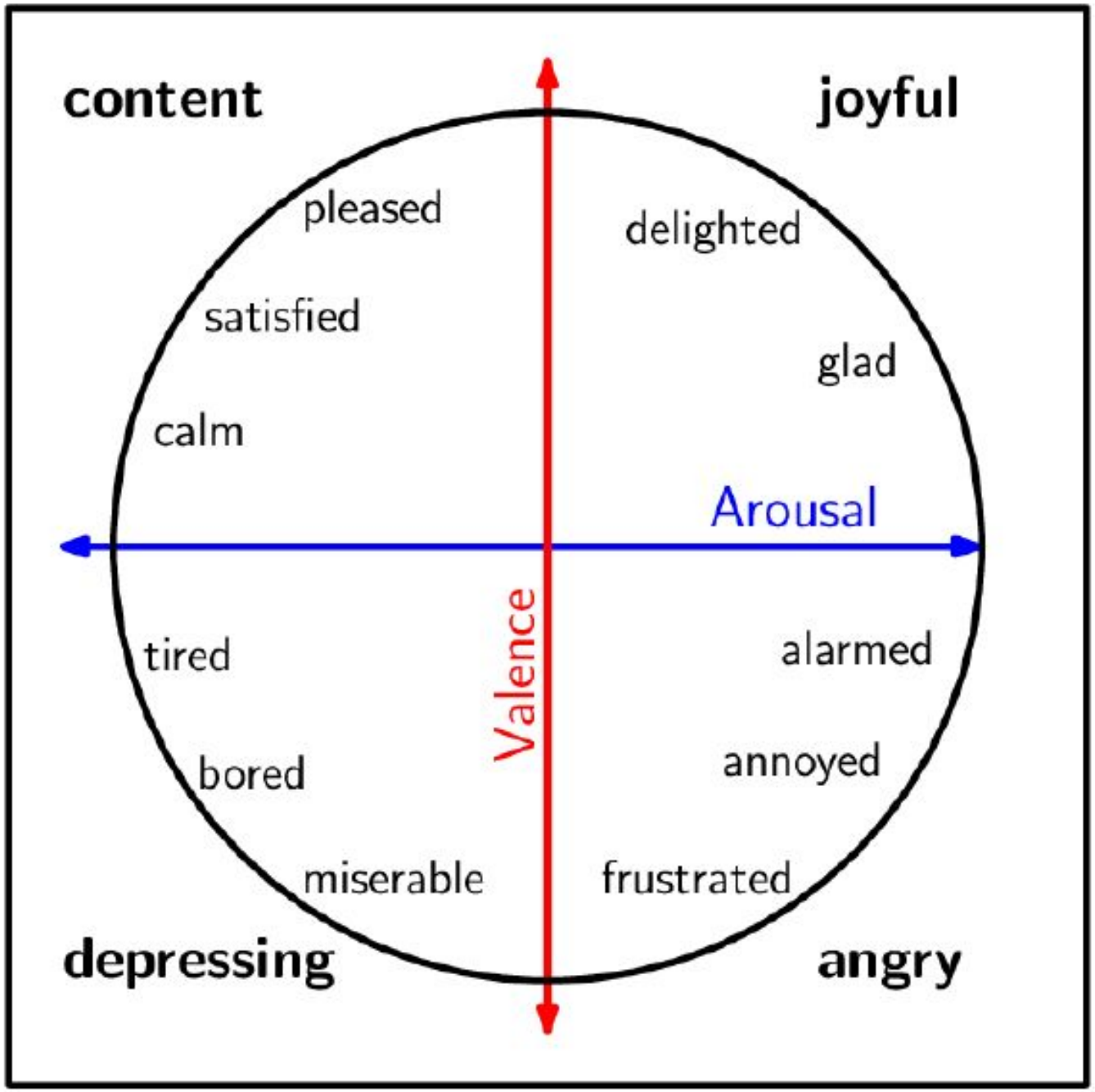
2. Ekman's Theory of (Six) Basic Emotions





Emotion Theories

3. Russel's Circumplex Model

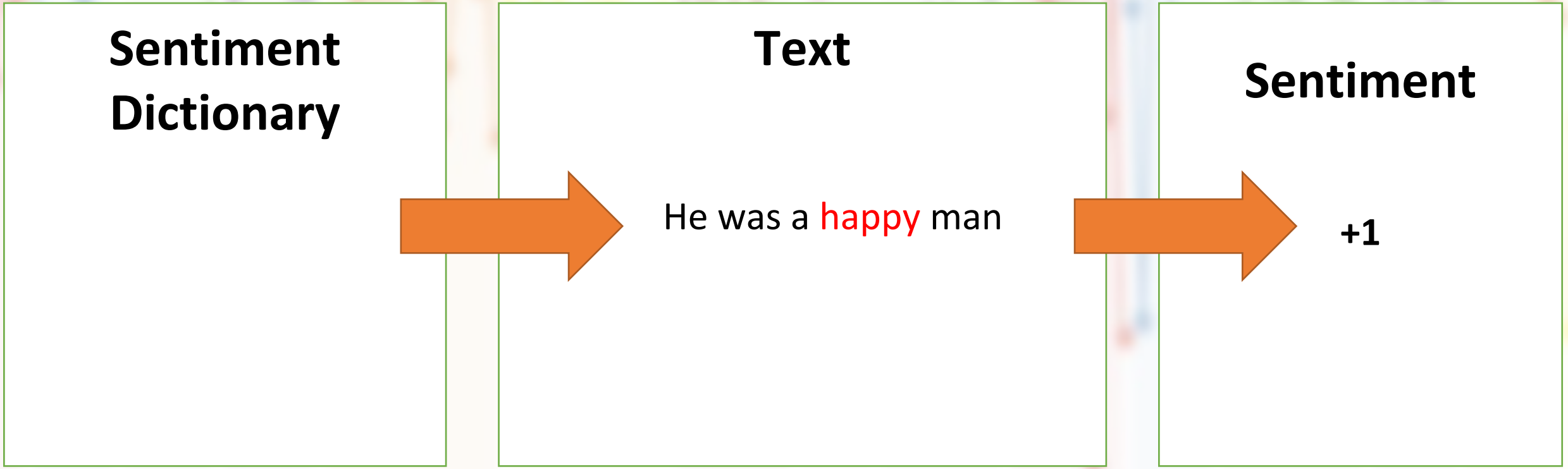


English Word	Arabic Translation (Google Translate)	Basque Translation (Google Translate)	Positive (Valence)	Negative (Valence)
aback	الى الوراء	aback	0	0
abacus	طبلية تاج	abako	0	0
abandon	تخلى	alde batera utzi	0	1
abandoned	مهجور	abandonatu	0	1
abandonment	هجر	abandono	0	1
abate	انحسر	abate	0	0
abatement	انحسار	murrizteko	0	0
abba	أبا	abba	1	0
abbot	رئيس الدير	abade	0	0
abbreviate	اختصر	laburtu	0	0
abbreviation	الاختصار	laburdura	0	0
abdomen	بطن	abdominalak	0	0
abdominal	البطن	sabeleko	0	0
abduction	اختطاف	urrunketa	0	1
aberrant	شاذ	aberranteari	0	1
aberration	انحراف	aberrazio	0	1
abeyance	abeyance	etena	0	0
abhor	مقت	gaitzesten	0	1
abhorrent	مشمئز	nazkagarria	0	1

NRC-Emotion-Lexicon-v0.92
(Mohammad and Turney, 2013)

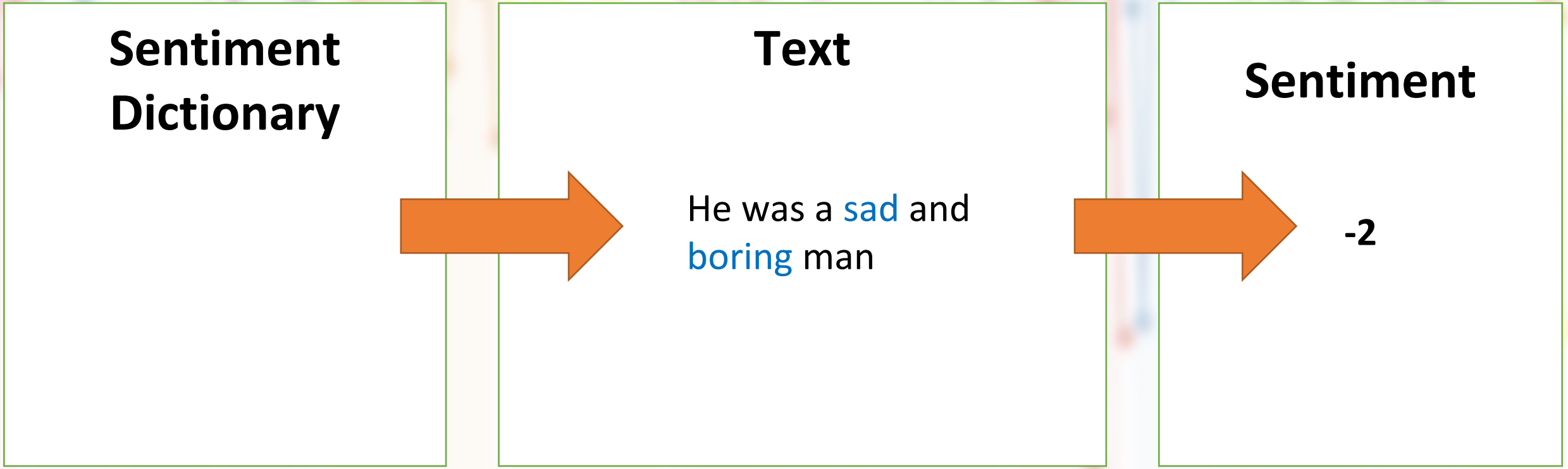
SA – How it works

Approach_1: Syuzhet **(Jockers, 2015)**



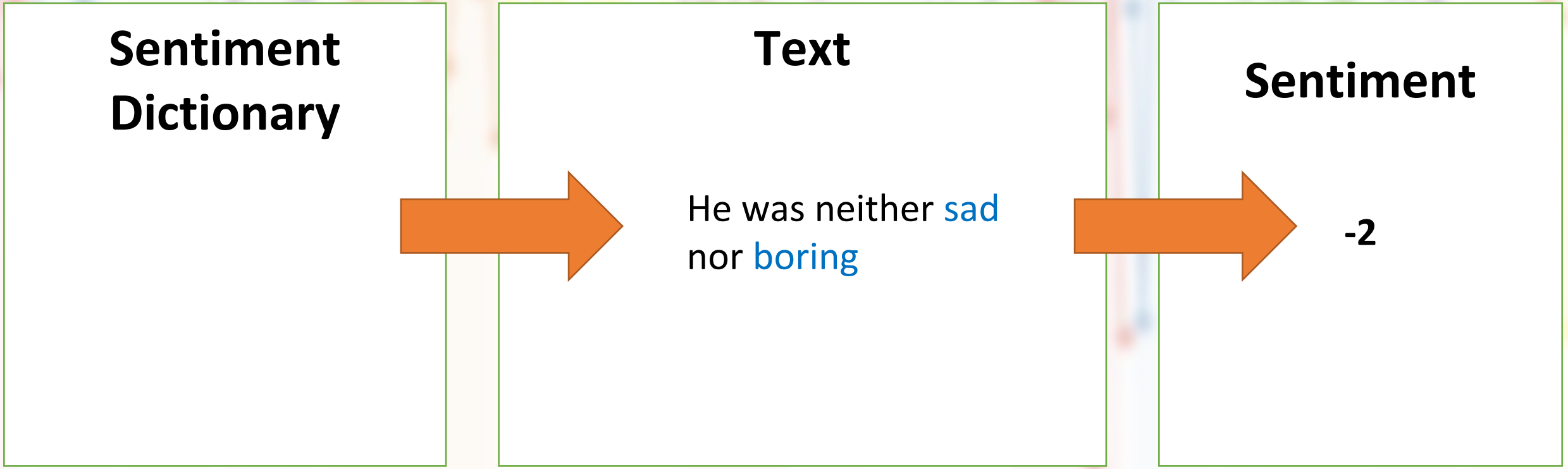
SA – How it works

Approach_1: Syuzhet **(Jockers, 2015)**



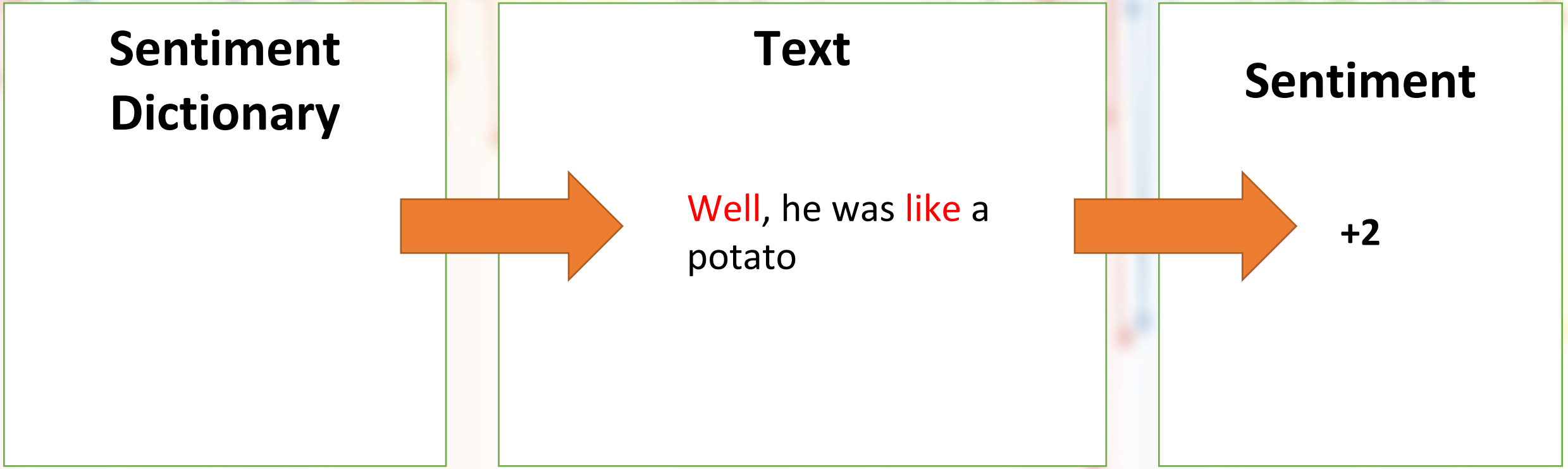
SA – How it works

Approach_1: Syuzhet **(Jockers, 2015)**



SA – How it works

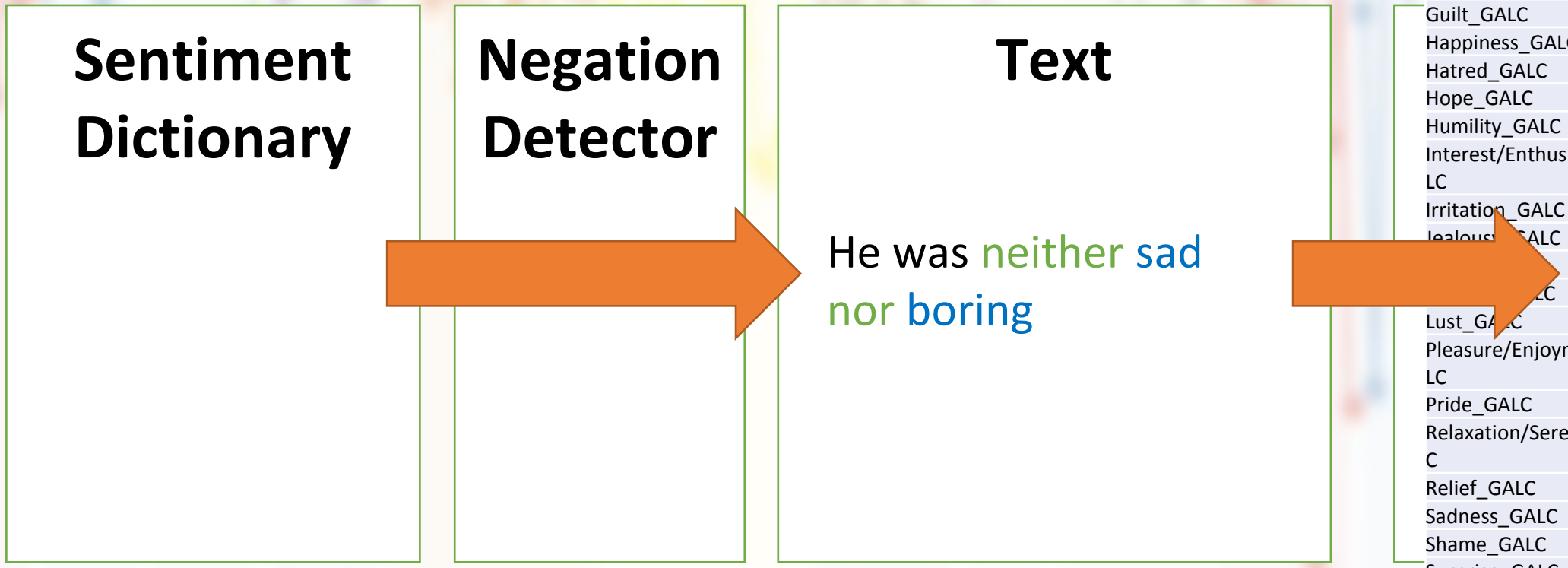
Approach_1: Syuzhet **(Jockers, 2015)**



Number	Index	Variable description	POS	Negation Filtered	examples (not POS specific)
1	filename	name of source file	NA	NA	NA
2	nwords	number of words in source file	NA	NA	NA
3	Admiration/Awe_GALC	Admiration/Awe	All	No	admir*, ador*, awe*, dazed, dazzl*
4	Amusement_GALC	Amusement	All	No	amus*, fun*, humor*, laugh*, play*
5	Anger_GALC	Anger	All	No	anger, angr*, cross*, enrag*, furious
6	Anxiety_GALC	Anxiety	All	No	anguish*, anxi*, apprehens*, diffiden*, jitter*
7	Beingtouched_GALC	Beingtouched	All	No	affect*, mov*, touch*
8	Boredom_GALC	Boredom	All	No	bor*, ennui, indifferen*, languor*, tedi*
9	Compassion_GALC	Compassion	All	No	commiser*, compass*, empath*, pit*,
10	Contempt_GALC	Contempt	All	No	contempt*, denigr*, deprec*, deris*, despi*
11	Contentment_GALC	Contentment	All	No	comfortabl*, content*, satisf*
12	Desperation_GALC	Desperation	All	No	deject*, desolat*, despair*, desperat*, despond*
13	Disappointment_GALC	Disappointment	All	No	comedown, disappoint*, discontent*, disenchant*, disgruntl*
14	Disgust_GALC	Disgust	All	No	abhor*, avers*, detest*, disgust*, dislik*
15	Dissatisfaction_GALC	Dissatisfaction	All	No	dissatisf*, unhapp*
16	Envy_GALC	Envy	All	No	envious*, envy*
17	Fear_GALC	Fear	All	No	afraid*, aghast*, alarm*, dread*, fear*
18	Feelinglove_GALC	Feelinglove	All	No	affection*, fond*, love*, friend*, tender*
19	Gratitude_GALC	Gratitude	All	No	grat*, thank*
20	Guilt_GALC	Guilt	All	No	blame*, contriti*, guilt*, remorse*, repent*
21	Happiness_GALC	Happiness	All	No	cheer*, bliss*, delect*, delight*, enchant*
22	Hatred_GALC	Hatred	All	No	acrimon*, hat*, rancor*

SA – How it works

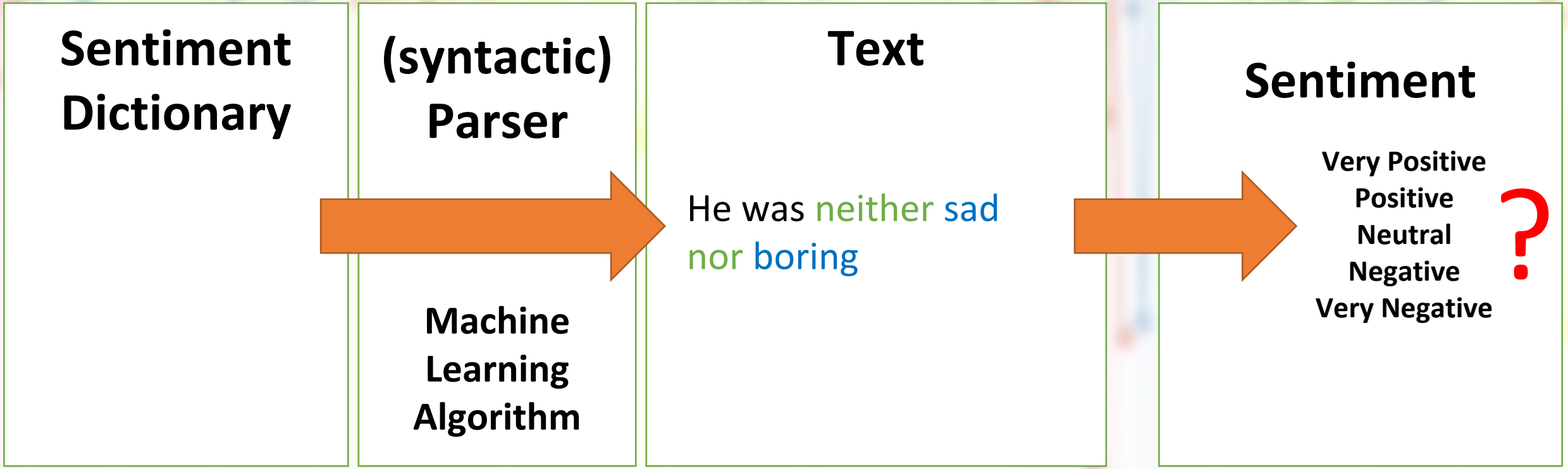
Approach_2: SEANCE (Crossley et al., 2017)



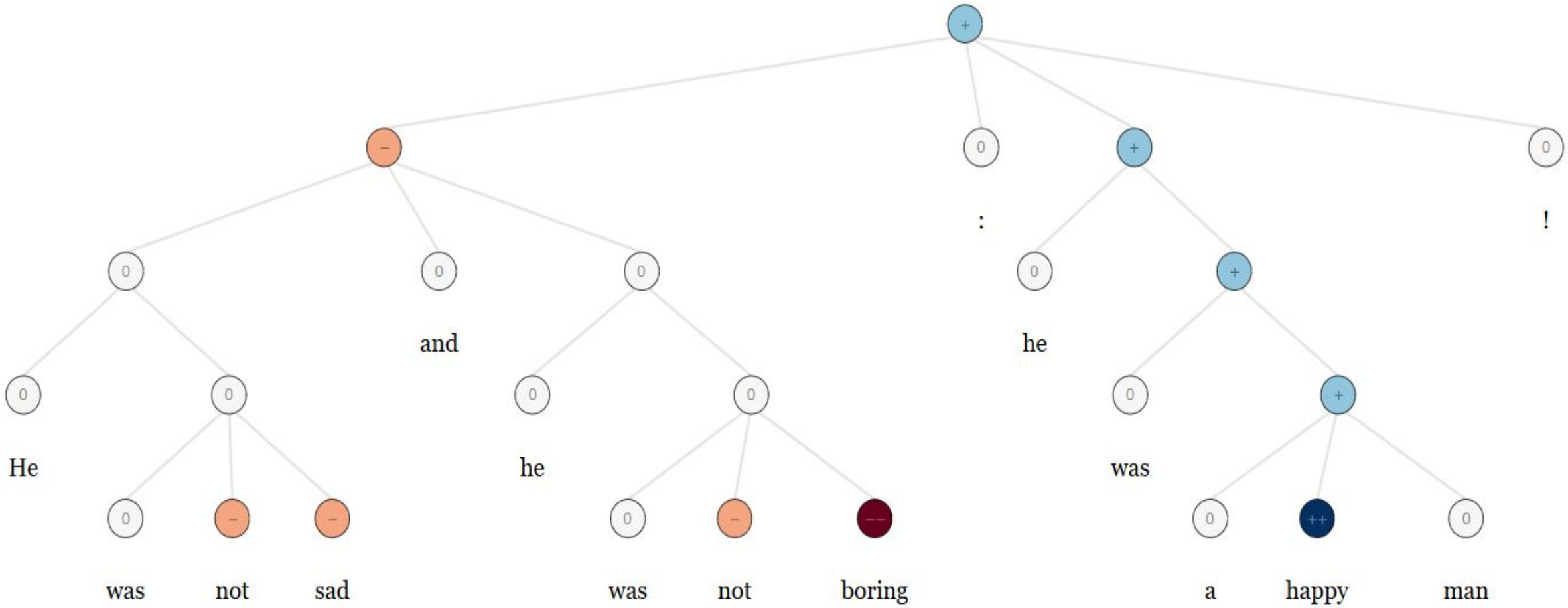
Admiration/Awe_GALC	1
Amusement_GALC	0
Anger_GALC	1
Anxiety_GALC	2
Beingtouched_GALC	-1
Boredom_GALC	0
Compassion_GALC	1
Contempt_GALC	1
Contentment_GALC	1
Desperation_GALC	0
Disappointment_GALC	0
Disgust_GALC	0
Dissatisfaction_GALC	1
Envy_GALC	0
Fear_GALC	1
Feelinglove_GALC	2
Gratitude_GALC	-1
Guilt_GALC	0
Happiness_GALC	1
Hatred_GALC	1
Hope_GALC	1
Humility_GALC	0
Interest/Enthusiasm_GALC	0
Irritation_GALC	0
Jealous[illegible]ALC	1
[illegible]	0
[illegible]LC	1
Lust_GA[illegible]C	2
Pleasure/Enjoyment_GALC	-1
Pride_GALC	0
Relaxation/Serenity_GALC	1
Relief_GALC	1
Sadness_GALC	1
Shame_GALC	0
Surprise_GALC	0
Tension/Stress_GALC	0

SA – How it works

Approach_3: Stanford SA **(Socher et al., 2013)**



He was not sad and he was not boring: he was a happy man!



All labels are now correct

SA Evaluation

TABLE 3 Results on the Sentiment Treebank for binary and fine-grained classification

- Even the most advanced SA tools hardly reach 90% accuracy in distinguishing positive from negative emotions
- Accuracy drops to below 50%, when distinguishing more fine-grained emotions (Ekman, Plutchick, etc.)

(Rojas-Barahona, 2016)