Speech and Language Processing

Chapter 20

Lexicons for Sentiment, Affect, and Connotation

吉永 墾

June 29, 2022

Introduction

- 主題：Affective meaning
  - affective：emotion, sentiment, personality, mood, atitudes
  - affective states の分類 [Scherer, 2000]

Emotion: Relatively brief episode of response to the evaluation of an external or internal event as being of major significance.
(*angry, sad, joyful, fearful, ashamed, proud, elated, desperate*)

Mood: Diffuse affect state, most pronounced as change in subjective feeling, of low intensity but relatively long duration, often without apparent cause.
(*cheerful, gloomy, irritable, listless, depressed, buoyant*)

Interpersonal stance: Affective stance taken toward another person in a specific interaction, coloring the interpersonal exchange in that situation.
(*distant, cold, warm, supportive, contemptuous, friendly*)

Attitude: Relatively enduring, affectively colored beliefs, preferences, and predispositions towards objects or persons.
(*liking, loving, hating, valuing, desiring*)

Personality traits: Emotionally laden, stable personality dispositions and behavior tendencies, typical for a person.
(*nervous, anxious, reckless, morose, hostile, jealous*)

Introduction

- affective states の有用性
  - emotion, mood：学習支援システム，ヘルプライン，twitter・blog，小説
  - interpersonal stance：会議要約（hot spot 検出）
  - personality：ユーザの性格検出
  - attect 生成：対話エージェントの感情付与

用語

- affective lexicons / sentiment lexicons
  - affect や sentiment への手がかりとなる 特定の word のリスト
- connotations
  - 書き手・読み手の emotion, sentiment, opinions, evaluations に関連するもの

20.1 Defining Emotion

- emotion
  - relatively brief episode of response to the evaluation of an external or internal event as being of major significance [Scherer, 2000]
  - *angry*, *sad*, *joyful*, *fearful*, *ashamed*, *proud*, *elated*, *desperate*
  - emotion が判ると何が嬉しいのか：
    - 言語処理タスク改善の可能性（学習支援システム，レビュー，医療）

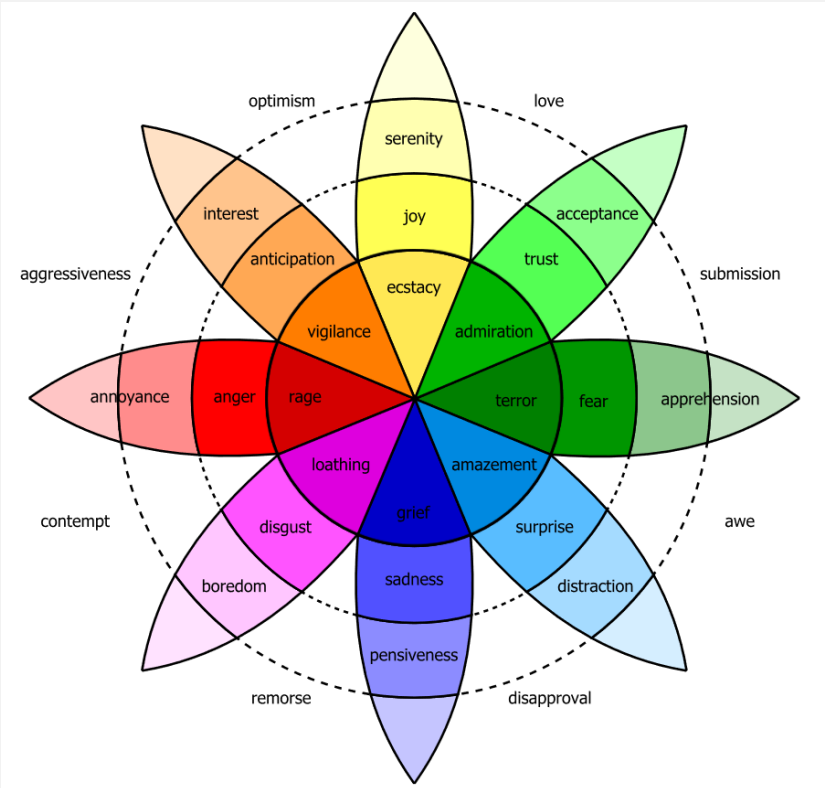
- emotion の計算モデル
  1. basic emotion の組み合わせ
  2. 3 次元 (VAD) 空間

Basic Emotion の組み合わせ

- emotion を 基本単位（basic emotion）の組み合わせとして表現

- 6 basic emotions [Ekman, 1999]
  - surprize, happiness, anger, fear, disgust, sadness

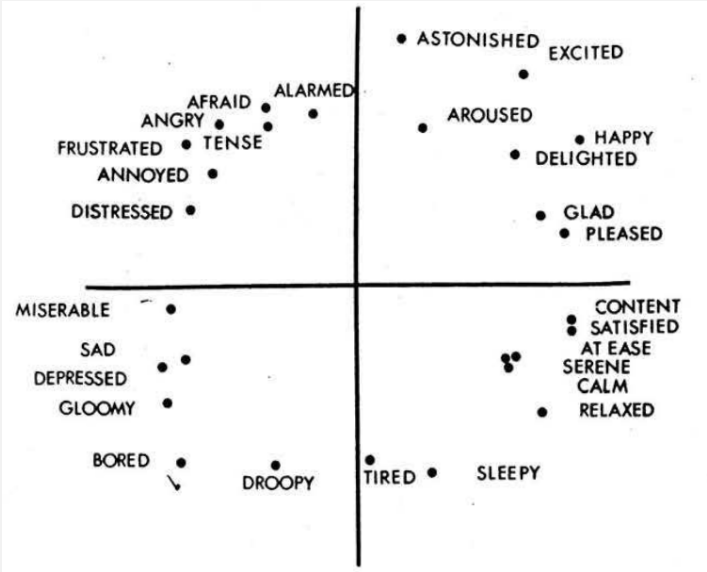
- 8 basic emotions [Plutchik, 1980]
  - joy - sadness
  - anger - fear
  - trust - disgust
  - anticipation - surprize



3 次元 (V A D) 空間

- emotion を V, A, D を軸とする 3 次元空間上の点として表現
  - Valence（感情価）：刺激の心地よさ
  - Arousal（覚醒度）：刺激による感情の強さ
  - Dominance（支配性）：刺激に支配される度合い
    - Dominance を除いて V,A の 2 次元とする場合も

- sentiment $\subset$ emotion
  - valence $=$ sentiment
  - emotion は sentiment の一般化



Russell の円環モデル [Russell, 1980]

20.2 Available Sentiment and Affect Lexicons

- Positive/Negative
  - General Inquire [Stone+, 1966]
  - MPQA Subjectivity Lexicon [Wilson+, 2005]
  - Polarity Lexicon [Hu & Liu, 2004]

Positive	admire, amazing, assure, celebration, charm, eager, enthusiastic, excellent, fancy, fantastic, frolic, graceful, happy, joy, luck, majesty, mercy, nice, patience, perfect, proud, rejoice, relief, respect, satisfactorily, sensational, super, terrific, thank, vivid, wise, wonderful, zest
Negative	abominable, anger, anxious, bad, catastrophe, cheap, complaint, condescending, deceit, defective, disappointment, embarrass, fake, fear, filthy, fool, guilt, hate, idiot, inflict, lazy, miserable, mourn, nervous, objection, pest, plot, reject, scream, silly, terrible, unfriendly, vile, wicked

- Valence, Arousal, Dominance の 3 次元
  - NRC VAD lexicon [Mohammad, 2018]

Valence		Arousal		Dominance	
vacation	.840	enraged	.962	powerful	.991
delightful	.918	party	.840	authority	.935
whistle	.653	organized	.337	saxophone	.482
consolation	.408	effortless	.120	discouraged	.0090
torture	.115	napping	.046	weak	.045

20.2 Available Sentiment and Affect Lexicons

- 8 basic emotions + Sentiments (positive/negative)
  - EmoLex [Mohammad & Turney, 2013]

Word	anger	anticipation	disgust	fear	joy	sadness	surprise	trust	positive	negative
reward	0	1	0	0	1	0	1	1	1	0
worry	0	1	0	1	0	1	0	0	0	1
tenderness	0	0	0	0	1	0	0	0	1	0
sweetheart	0	1	0	0	1	1	0	1	1	0
suddenly	0	0	0	0	0	0	1	0	0	0
thirst	0	1	0	0	0	1	1	0	0	0
garbage	0	0	1	0	0	0	0	0	0	1

- Anger, Fear, Joy, Sadness との関連度
  - NRC Emotion/Affect Intensity Lexicon [Mohammad, 2018]

Anger		Fear		Joy		Sadness	
outraged	0.964	horror	0.923	superb	0.864	sad	0.844
violence	0.742	anguish	0.703	cheered	0.773	guilt	0.750
coup	0.578	pestilence	0.625	rainbow	0.531	unkind	0.547
oust	0.484	stressed	0.531	gesture	0.387	difficulties	0.421
suspicious	0.484	failing	0.531	warms	0.391	beggar	0.422
nurture	0.059	confident	0.094	hardship	.031	sing	0.017

20.2 Available Sentiment and Affect Lexicons

- Positive/Negative + 73 カテゴリ
  - LIWC, Linguistic Inquiry and Word Count [Pennebaker+, 2007]

Positive Emotion	Negative Emotion	Insight	Inhibition	Family	Negate
appreciat*	anger*	aware*	avoid*	brother*	aren't
comfort*	bore*	believe	careful*	cousin*	cannot
great	cry	decid*	hesitat*	daughter*	didn't
happy	despair*	feel	limit*	family	neither
interest	fail*	figur*	oppos*	father*	never
joy*	fear	know	prevent*	grandf*	no
perfect*	griev*	knew	reluctan*	grandm*	nobod*
please*	hate*	means	safe*	husband	none
safe*	panic*	notice*	stop	mom	nor
terrific	suffers	recogni*	stubborn*	mother	nothing
value	terrify	sense	wait	niece*	nowhere
wow*	violent*	think	wary	wife	without

- Concrete/Abstruct words [Brysbaert+, 2014]
  - 単語の具体度（concreteness）
    - Concrete words : *banana* (5), *bathrobe* (5)
    - Abstruct words : *belief* (1.19), *although* (1.07)

20.3 Creating Affect Lexicons by Human Labeling

- lexicon 構成手法：人手でラベリング
  - crowdsourcing：小さなタスクに分割，大人数のアノテータに振り分け

NRC Emotion Lexicon（EmoLex）[Mohammad & Turney, 2013]

1. 同義語選択
   - アノテータの word sense 検証，問題は自動生成

   例 `Which word is closest in meaning (most related) to startle?`
   `automobile / shake / honesty / entertain`

2. 8 basic emotions との関連度の回答
   - 関連度：*not*, *weakly*, *moderately*, *strongly*

- 多数決で関連度を決定
- 関連度よりスコア $\in \{0, 1\}$ を決定
  - *not*, *weakly* $\mapsto 0$
  - *moderately*, *strongly* $\mapsto 1$

Word	anger	anticipation	disgust	fear	joy	sadness	surprise	trust	positive	negative
reward	0	1	0	0	1	0	1	1	1	0
worry	0	1	0	1	0	1	0	0	0	1
tenderness	0	0	0	0	1	0	0	0	1	0
sweetheart	0	1	0	0	1	1	0	1	1	0
suddenly	0	0	0	0	0	0	1	0	0	0
thirst	0	1	0	0	0	1	1	0	0	0
garbage	0	0	1	0	0	0	0	0	0	1

20.3 Creating Affect Lexicons by Human Labeling

- lexicon 構成手法：人手でラベリング
  - crowdsourcing：小さなタスクに分割，大人数のアノテータに振り分け

NRC VAD Lexicon [Mohammad, 2018]

1. N 個（通常 $N = 4$）の word の中から best/worst を選択

   例：Valence
   - Which of the four words below is associated with the MOST happiness / pleasure / positiveness / satisfaction / contentedness / hopefulness OR LEAST unhappiness / annoyance / negativeness / dissatisfaction / melancholy / despair ?
   - Which of the four words below is associated with the LEAST happiness / pleasure / positiveness / satisfaction / contentedness / hopefulness OR MOST unhappiness / annoyance / negativeness / dissatisfaction / melancholy / despair ?

   - best-worst scaling
     - $(\text{word のスコア}) = (\text{best に選ばれた割合}) - (\text{worst に選ばれた割合})$

   - 評価：split-half reliability
     - コーパスを 2 分割
       相関を計算

Valence		Arousal		Dominance	
vacation	.840	enraged	.962	powerful	.991
delightful	.918	party	.840	authority	.935
whistle	.653	organized	.337	saxophone	.482
consolation	.408	effortless	.120	discouraged	.0090
torture	.115	napping	.046	weak	.045

20.4 Semi-supervised Induction of Affect Lexicons

20.4.1 Semantic Axis Method

1. seed words を人手で選択
   (1) single large seed lexicon を使用
   (2) 分野ごとに seed words を使い分ける

Domain	Positive seeds	Negative seeds
General	good, lovely, excellent, fortunate, pleasant, delightful, perfect, loved, love, happy	bad, horrible, poor, unfortunate, unpleasant, disgusting, evil, hated, hate, unhappy
Twitter	love, loved, loves, awesome, nice, amazing, best, fantastic, correct, happy	hate, hated, hates, terrible, nasty, awful, worst, horrible, wrong, sad
Finance	successful, excellent, profit, beneficial, improving, improved, success, gains, positive	negligent, loss, volatile, wrong, losses, damages, bad, litigation, failure, down, negative

分野による seed words 例 [Hamilton+, 2016]

2. 埋め込み空間での計算
   - seed words を埋め込み，重心 $\boldsymbol{V}^+, \boldsymbol{V}^-$，semantic axis $\boldsymbol{V}_{axis}$ を計算

$$\boldsymbol{V}^+ = \frac{1}{n}\sum_{1}^{n} E(w_i^+), \quad \boldsymbol{V}^- = \frac{1}{m}\sum_{1}^{m} E(w_i^-), \quad \boldsymbol{V}_{axis} = \boldsymbol{V}^+ - \boldsymbol{V}^-$$

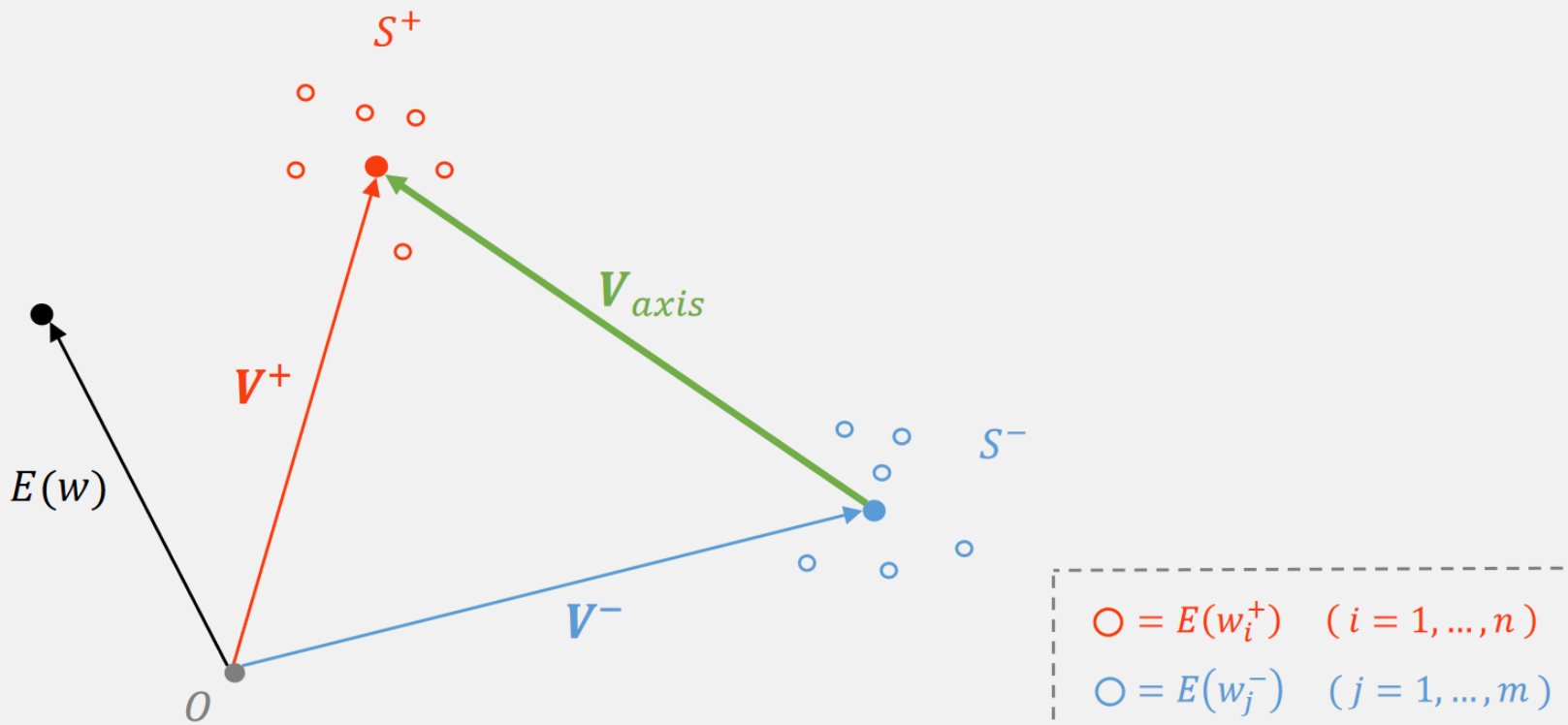
20.4 Semi-supervised Induction of Affect Lexicons

- seed words を埋め込み，重心 $\boldsymbol{V}^+, \boldsymbol{V}^-$，semantic axis $\boldsymbol{V}_{axis}$ を計算

$$\boldsymbol{V}^+ = \frac{1}{n}\sum_{1}^{n} E(w_i^+), \quad \boldsymbol{V}^- = \frac{1}{m}\sum_{1}^{m} E(w_i^-), \quad \boldsymbol{V}_{axis} = \boldsymbol{V}^+ - \boldsymbol{V}^-$$

- word w のスコアを計算

$$\text{score}(w) = \cos(E(w), \boldsymbol{V}_{axis}) = \frac{E(w) \cdot \boldsymbol{V}_{axis}}{\|E(w)\|\|\boldsymbol{V}_{axis}\|}$$

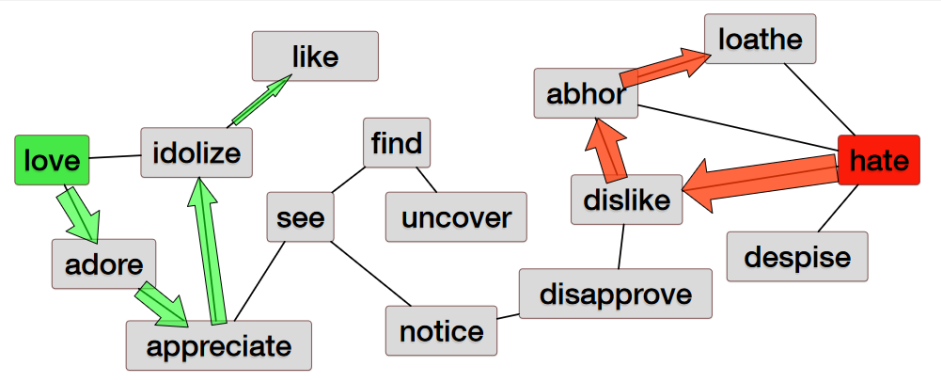


20.4.2 Label Propagation

1. グラフ定義
   - node：word embedding $\boldsymbol{w}$，k 個の近傍 node と接続
   - edge：node $\boldsymbol{w}_i$, $\boldsymbol{w}_j$ 間の重み $E_{i,j}$

$$E_{i,j} = \arccos\left(-\frac{\boldsymbol{w}_i^\top \boldsymbol{w}_j}{\|\boldsymbol{w}_i\|\|\boldsymbol{w}_j\|}\right)$$

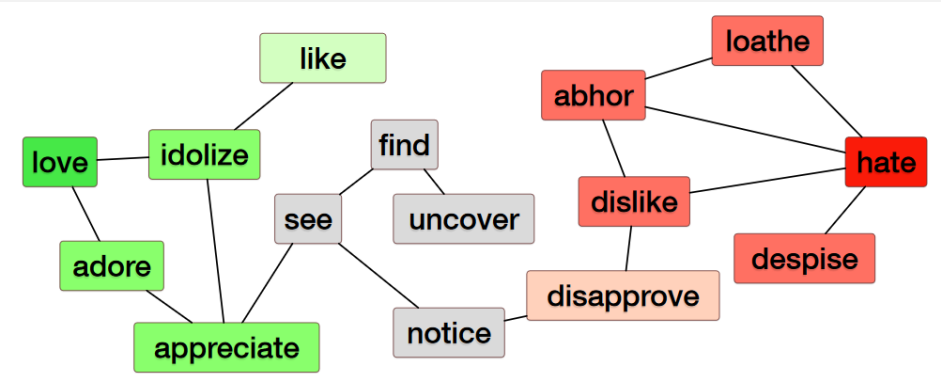
2. seed set 定義
3. seed set から polarity を伝播
   - seed set を起点に edge の重みに従って random walk



20.4 Semi-supervised Induction of Affect Lexicons

4. word score の計算
   - random walk の訪問回数からスコアを計算

$$\text{score}^{+}(w_i) = \frac{\text{rawscore}^{+}(w_i)}{\text{rawscore}^{+}(w_i) + \text{rawscore}^{-}(w_i)}$$



5. スコアの信頼度の計算
   - word score は seed set 依存
   - seed set の部分集合を random に選び propagation (3-4) を複数回実行
   - 各 word のスコアの標準偏差を計算

20.4 Semi-supervised Induction of Affect Lexicons

20.4.3 Other Methods

- seed set との 類似度 の尺度
- and, but による形容詞の並列関係 [Hatzivassiloglou & McKeown, 1997]
  - and：同じ polarity
    - fair and legitimate / corrupt and brutal
  - but：逆の polarity
    - fair but brutal
- 同一の語幹の形容詞の morphological negation
  - *un-*, *im-*, *-less* がつくと逆の polarity
    - adequate/inadequate, thoughtful/thoughtless
- WordNet の synonim/antonym
  - synonim：同じ polarity
  - antonym：逆の polarity

  → SentiWordNet [Baccianella+, 2010]

20.5 Supervised Learning of Word Sentiment

- Online review（スコア，テキストの組）から lexicon を構成
  - 高評価のレビューほど positive な word が現れやすい
  - 多段評価なので分布が得られる

Movie review excerpts (IMDb)	
10	A great movie. This film is just a wonderful experience. It's surreal, zany, witty and slapstick all at the same time. And terrific performances too.
1	This was probably the worst movie I have ever seen. The story went nowhere even though they could have done some interesting stuff with it.

  - word w と評価 c の関連度を計算

$$p(w \mid c) = \frac{count(w,c)}{\sum_{w \in c} count(w,c)} = \frac{\text{Count}(c)}{\text{Total}(c)} = \Pr_{\text{IMDB}}(w \mid c)$$

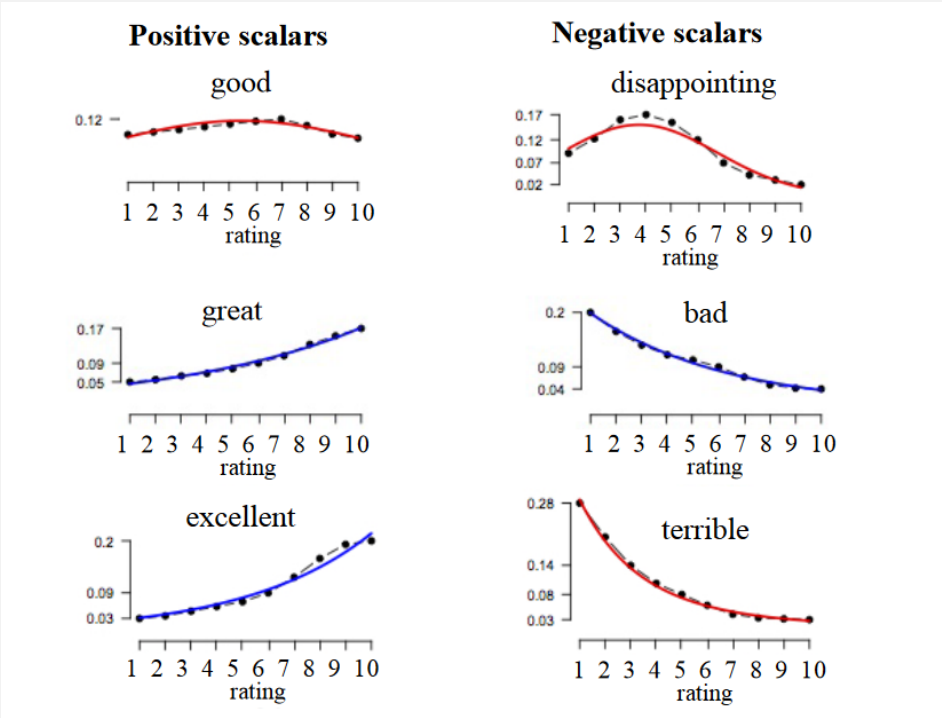
$$PottsScore(w) = \frac{p(x \mid c)}{\sum_c p(w \mid c)} = \Pr_{\text{IMDB}}(c \mid w)$$

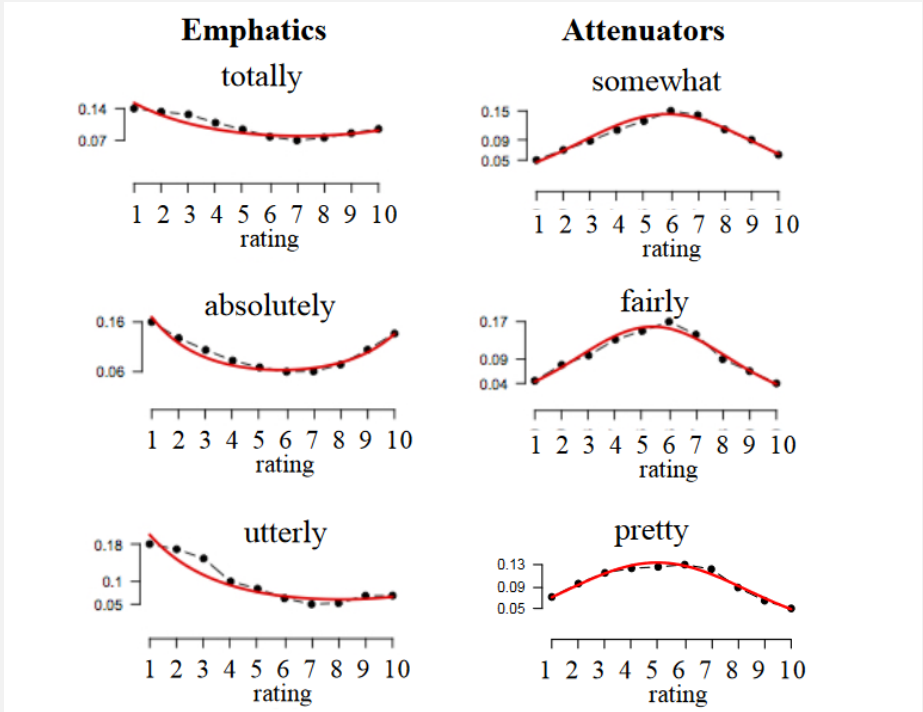
Cat.	Count	Total	$\Pr_{\text{IMDB}}(w\|c)$	$\Pr_{\text{IMDB}}(c\|w)$
1	8,557	25,395,214	0.0003	0.10
2	4,627	11,755,132	0.0004	0.12
3	6,726	13,995,838	0.0005	0.14
4	7,171	14,963,866	0.0008	0.14
5	9,039	20,390,515	0.0004	0.13
6	10,101	27,420,036	0.0004	0.11
7	10,362	40,192,077	0.0003	0.08
8	10,064	48,723,444	0.0002	0.06
9	7,909	40,277,743	0.0002	0.06
10	13,570	73,948,447	0.0002	0.05

$w =$ 'disappointing' の場合 [Potts, 2011]

20.5 Supervised Learning of Word Sentiment

- word$w$ と評価 c の関連度を visualize（Potts diagram）
  - 曲線の形を affective meaning のタイプと見做す





- strongly positive/negative : J-shape
- weakly positive/negative : hump-shape
- empahtics : J/U-shape
- attenuators : hump-shape

20.5.1 Log Odds Ration Informative Dirichlet Prior

- 1 つのカテゴリでのみ使用される 特徴的な word の検出
- word w は コーパス i, j のどちらに より現れるか
  - 対数オッズ比 （ $\mathrm{odds} = p/(1-p)$ ） :

$$\begin{aligned}\mathrm{lor}(w) &= \log\left(\frac{p^i(w)}{1-p^i(w)}\right) - \log\left(\frac{p^j(w)}{1-p^j(w)}\right) \\ &= \log\left(\frac{f_w^i}{n^i - f_w^i}\right) - \log\left(\frac{f_w^j}{n^j - f_w^j}\right)\end{aligned}$$

    where n^i : コーパス i の総 word 数
    f_w^i : コーパス i 内の word w の数

  - 高頻出・低頻出単語への対応
    $\rightarrow$ log odds ratio informative Dirichlet prior [Monroe, 2008]

20.5 Supervised Learning of Word Sentiment

- large background corpus を考慮：

$$\hat{\delta}_w^{(i-j)} = \log\left(\frac{f_w^i + \alpha_w}{n^i + \alpha_0 - (f_w^i + \alpha_w)}\right) - \log\left(\frac{f_w^j + \alpha_w}{n^j + \alpha_0 - (f_w^j + \alpha_w)}\right)$$

where α_0：background corpus の総 word 数

α_w：background corpus 内の word w の数

- 分散を近似：

$$\sigma^2\left(\hat{\delta}_w^{(i-j)}\right) \approx \frac{1}{f_w^i + \alpha_w} + \frac{1}{f_w^j + \alpha_w}$$

- z 変換：

$$\frac{\hat{\delta}_w^{(i-j)}}{\sqrt{\sigma^2\left(\hat{\delta}_w^{(i-j)}\right)}}$$

20.5 Supervised Learning of Word Sentiment

Class	Words in 1-star reviews	Class	Words in 5-star reviews
Negative	*worst, rude, terrible, horrible, bad, awful, disgusting, bland, tasteless, gross, mediocre, overpriced, worse, poor*	**Positive**	*great, best, love(d), delicious, amazing, favorite, perfect, excellent, awesome, friendly, fantastic, fresh, wonderful, incredible, sweet, yum(my)*
Negation	*no, not*	**Emphatics/ universals**	*very, highly, perfectly, definitely, absolutely, everything, every, always*
1Pl pro	*we, us, our*	**2 pro**	*you*
3 pro	*she, he, her, him*	**Articles**	*a, the*
Past verb	*was, were, asked, told, said, did, charged, waited, left, took*	**Advice**	*try, recommend*
Sequencers	*after, then*	**Conjunct**	*also, as, well, with, and*
Nouns	*manager, waitress, waiter, customer, customers, attitude, waste, poisoning, money, bill, minutes*	**Nouns**	*atmosphere, dessert, chocolate, wine, course, menu*
Irrealis modals	*would, should*	**Auxiliaries**	*is/'s, can, 've, are*
Comp	*to, that*	**Prep, other**	*in, of, die, city, mouth*

★1/★5 のレビューと関連する word [Jurafsky+, 2014]

20.6 Using Lexicons for Sentiment Recognition

- lexicon を用いた sentiment analysis の改善
  - 十分な学習データがない場合
  - 文章内の positive/negative word の割合から判定

$$f^+ = \sum_{w \text{ s.t } w \in \text{positive lexicon}} \theta_w^+ \; count(w)$$

$$f^- = \sum_{w \text{ s.t } w \in \text{negative lexicon}} \theta_w^- \; count(w)$$

$$sentiment = \begin{cases} + & \text{if } \dfrac{f^+}{f^-} > \lambda \\ - & \text{if } \dfrac{f^-}{f^+} > \lambda \\ 0 & \text{otherwise} \end{cases}$$

where θ_w^+, θ_w^- : weights

λ : threshold

20.7 Using Lexicons for Affect Recognition

- 大規模データセット
  - 何かしらの修正が必要

- 例：言語情報と個性・性別・年齢の関連 [Schwartz+, 2013]
  - Facebook の投稿（700million words）を n-gram（$n = 1, 2, 3$）の形で使用
  - 1%以上のユーザに使用されている word・phrase のみ使用
  - 2-gram, 3-gram：PMI が充分大きい phrase のみ使用

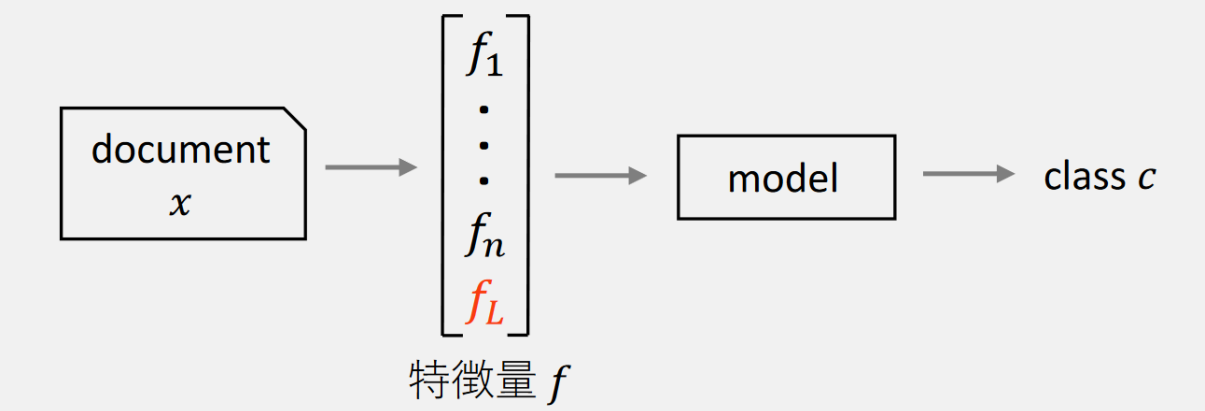
$$\mathrm{pmi}(phrase) = \log \frac{p(phrase)}{\prod_{w \in phrase} p(w)}$$

  - 尤度を計算：

$$p(phrase \mid subject) = \frac{\mathrm{freq}(phrase, subject)}{\sum_{phrase' \in \mathrm{vocab}(subject)} \mathrm{freq}(phrase', subject)}$$

20.7 Using Lexicons for Affect Recognition

- training data が sparser，test set に似ていない
  - lexicon L を利用して document x の特徴量 f_L を作成



  - 特性関数の条件として使用：

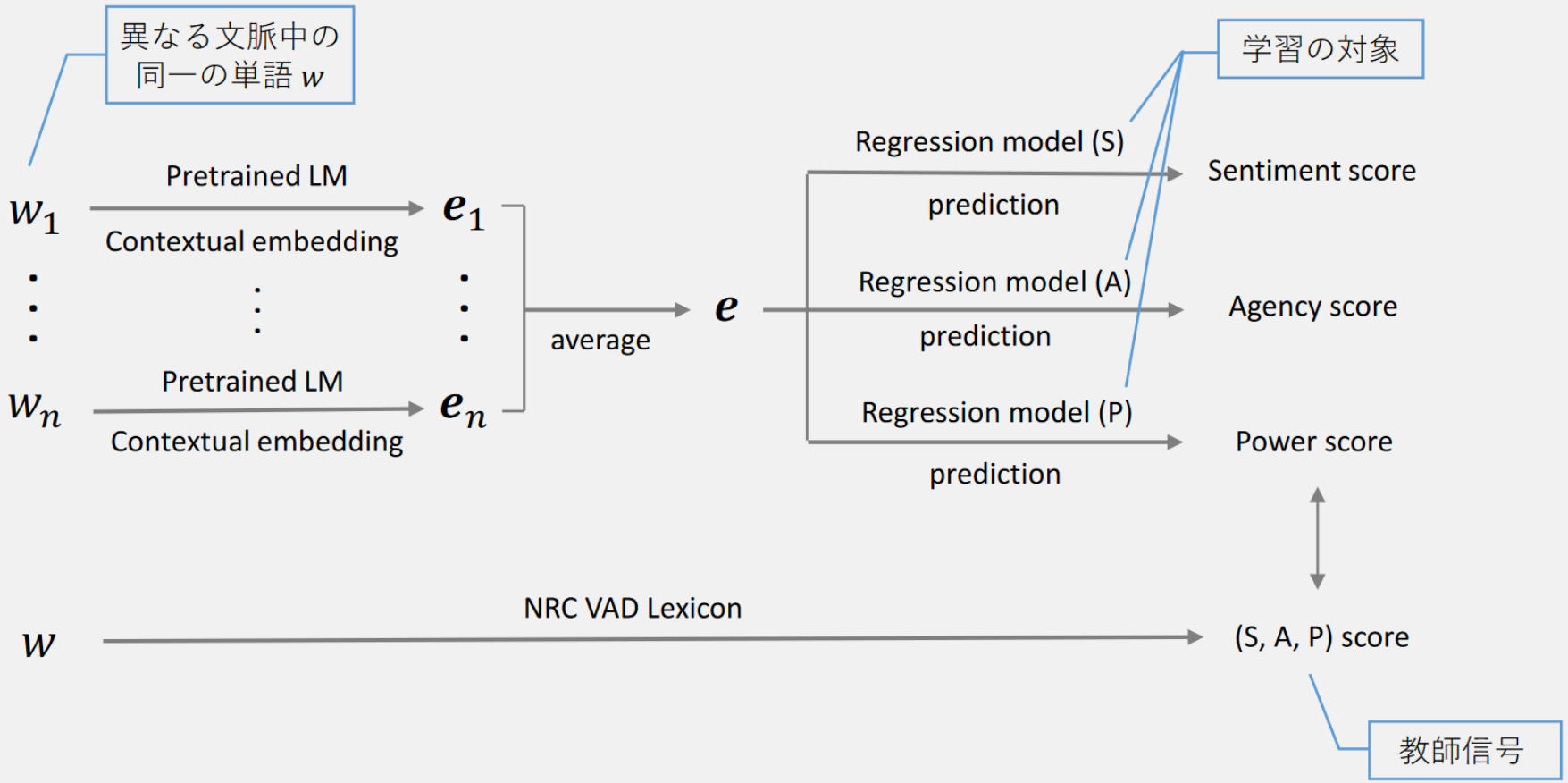
$$f_L(c,x) = \begin{cases} 1 & \text{if } \exists w, w \in L \ \& \ w \in x \ \& \ class = c \\ 0 & \text{otherwise} \end{cases}$$

  - 数え上げの条件として使用：

$$f_L = \sum_{w \in L} \theta_w^L \ count(w)$$

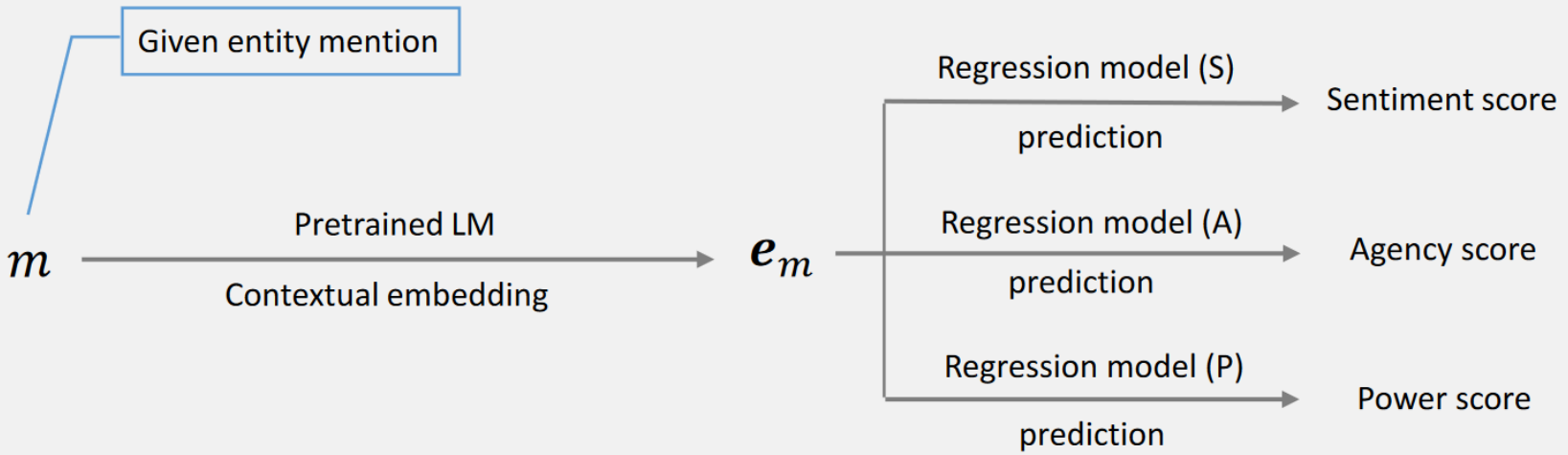
20.8 Lexicon-based Method for Entity-Centric Affect

- 特定の entity の affect score を計算 [Field & Tsvetkov, 2019]
  - Valence / Arousal / Dominance ⇔ Sentiment / Agency / Power



20.8 Lexicon-based Method for Entity-Centric Affect

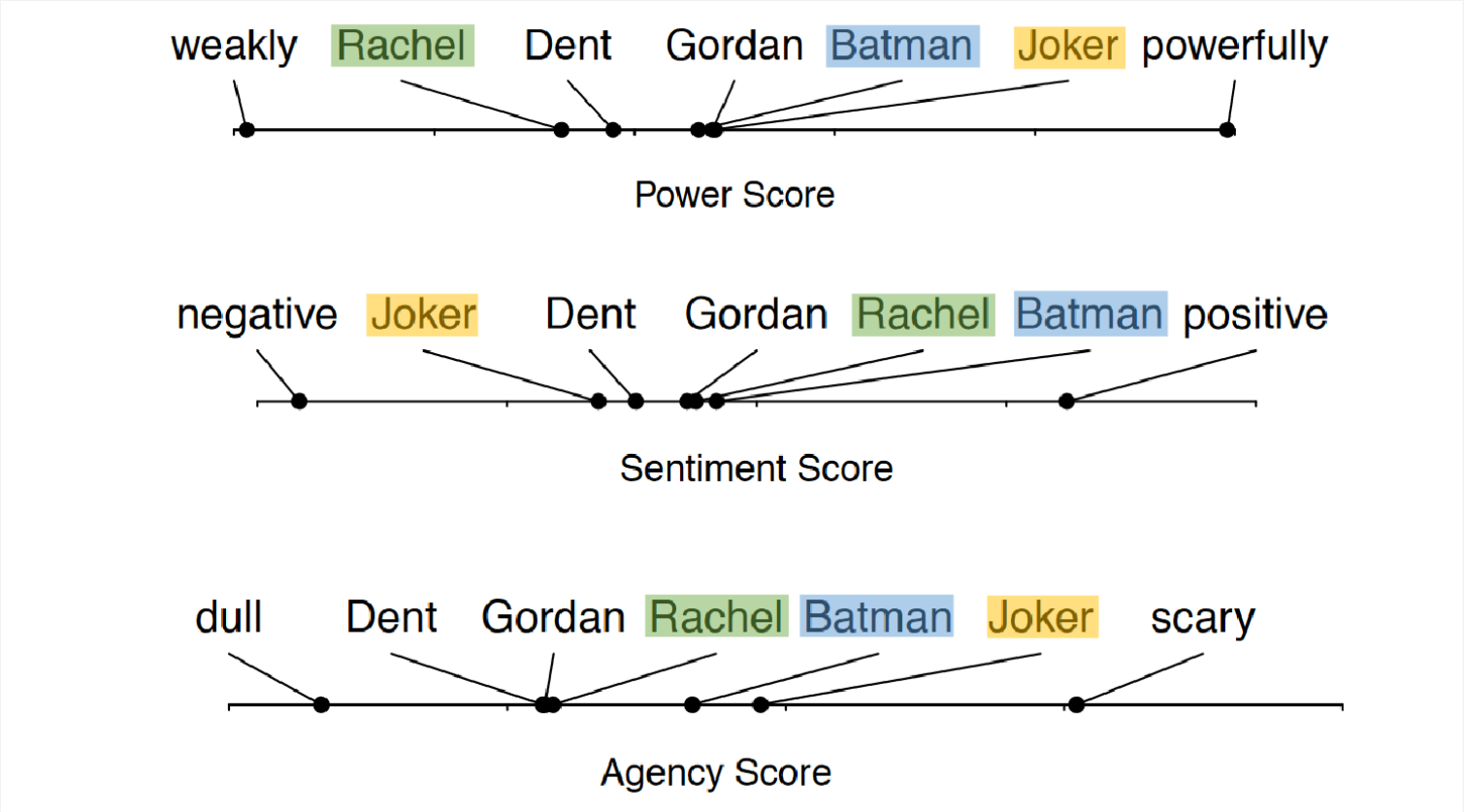
- 特定の entity の affect score を計算 [Field & Tsvetkov, 2019]
  - Valence / Arousal / Dominance ⇔ Sentiment / Agency / Power



  - coreference resolution 実行，得られた全 (S,A,P) score を平均

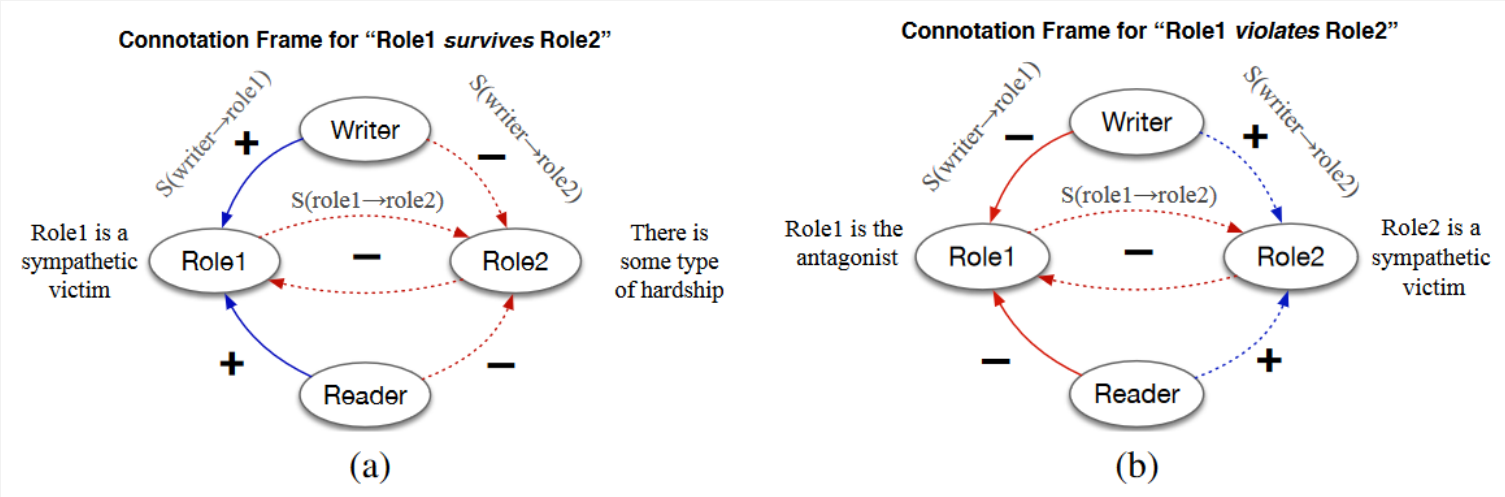
20.8 Lexicon-based Method for Entity-Centric Affect

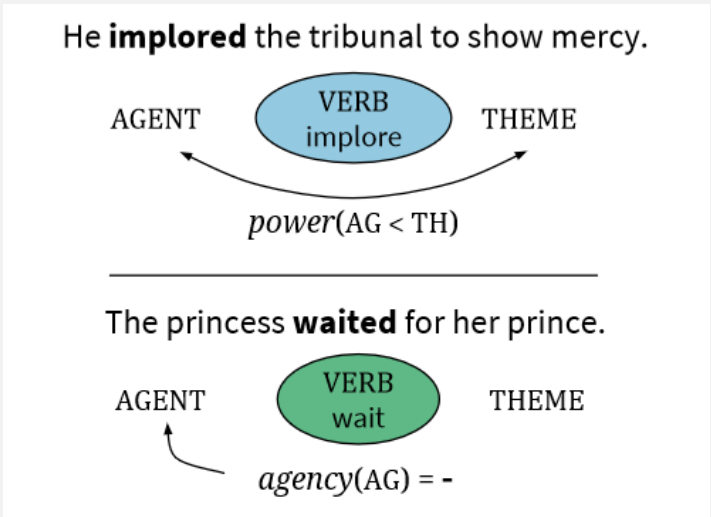
- 特定の entity の affect score を計算 [Field & Tsvetkov, 2019]
  - Valence / Arousal / Dominance ⇔ Sentiment / Agency / Power



20.9 Connotation Frames

- connotation frame：文法構造を組み込んだ lexicon





20.10 Summary

- affective states：
  - emotions, moods, attitudes, interpersonal stance, personality
- emotion の表現：basic emotion, valence-arousal

- lexicon：word と affective states の関連性
- lexicon 構成手法：

  human labeling crowdsourcing
  semi-supervised seed words との類似性
  fully supervised ユーザの review から

- affect の検出
  - lexicon の利用：特徴量作成，学習データ

- connotation frames：文法構造を組み込んだ lexicon