

文脈から感情を分析するための 感情判断器の検討

大澤 泰我[†]

マツキン ケネスジェームス[†]
[†]東京情報大学

永井 保夫[†]

研究背景

- 近年コミュニケーションロボットの需要が高まっている
- 対話システムと人との自然なコミュニケーションの実現が求められている
- 人との自然な対話の実現において、感情情報は必要不可欠

対話システムの問題点

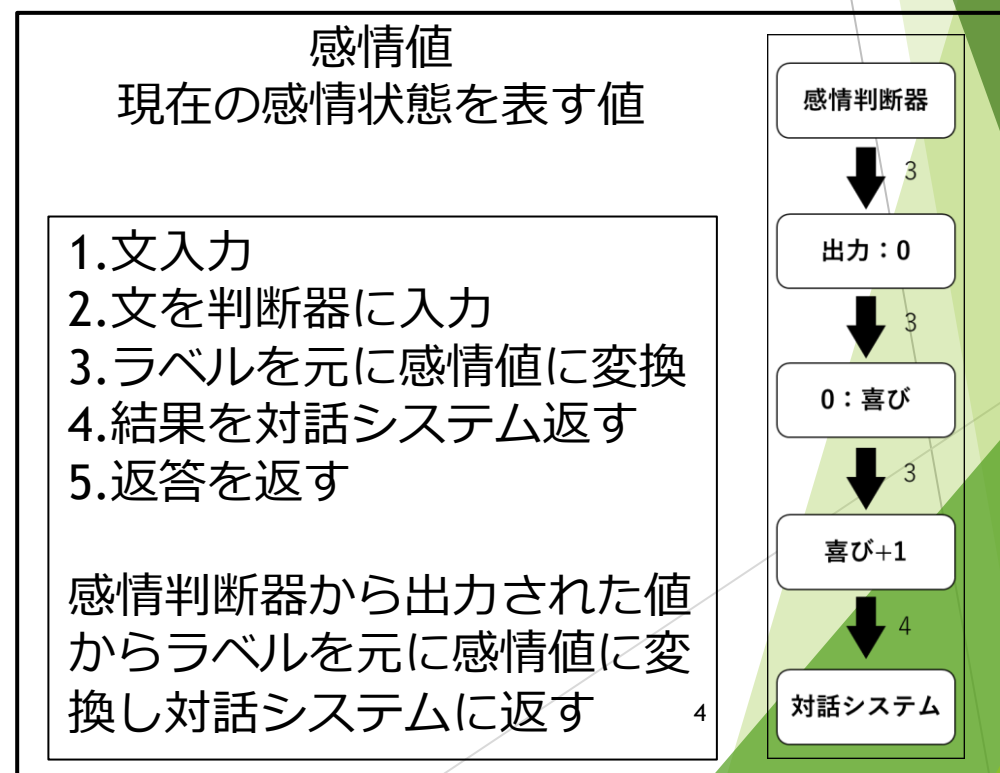
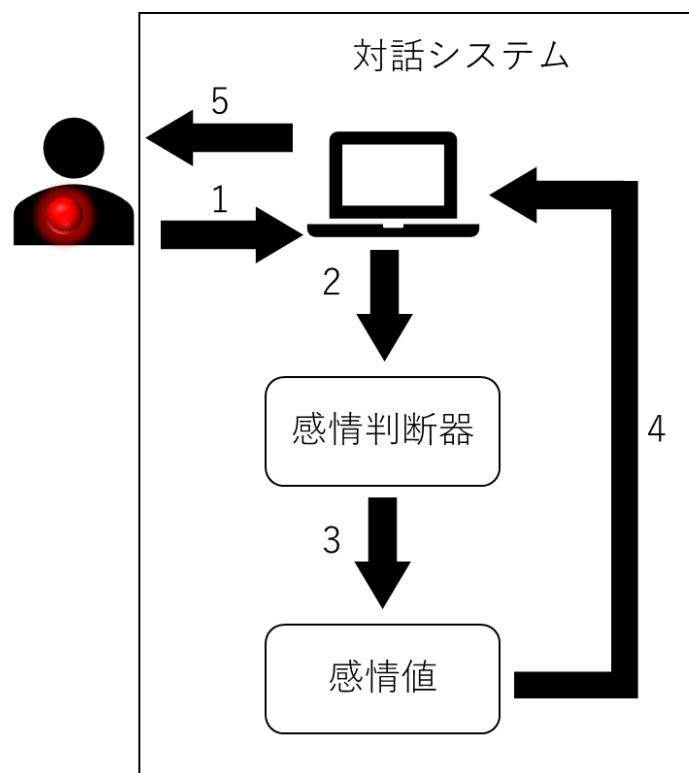
- 対話の際に感じる心理的抵抗感や発話が機械的
- 解決する方法の一つとして、感情を考慮した対話の実現が考えられている[1][2]

[1]河野 泉,池邊 亮志,和氣 早苗,上窪 真一,岩沢 透,西村 健士.感情表現を用いた対話システムEDSの開発(1): システム概要と感情モデル.情報処理学会研究報告, “HI,ヒューマンインタフェース研究会報告 89”, pp.43-48,(2000)

[2]河野 泉,池邊 亮志,和氣 早苗,上窪 真一,岩沢 透,西村 健士.感情表現を用いた対話システムEDSの開発(2): システム概要と感情モデル.情報処理学会研究報告, “HI,ヒューマンインタフェース研究会報告 89”, pp.49-56,(2000)

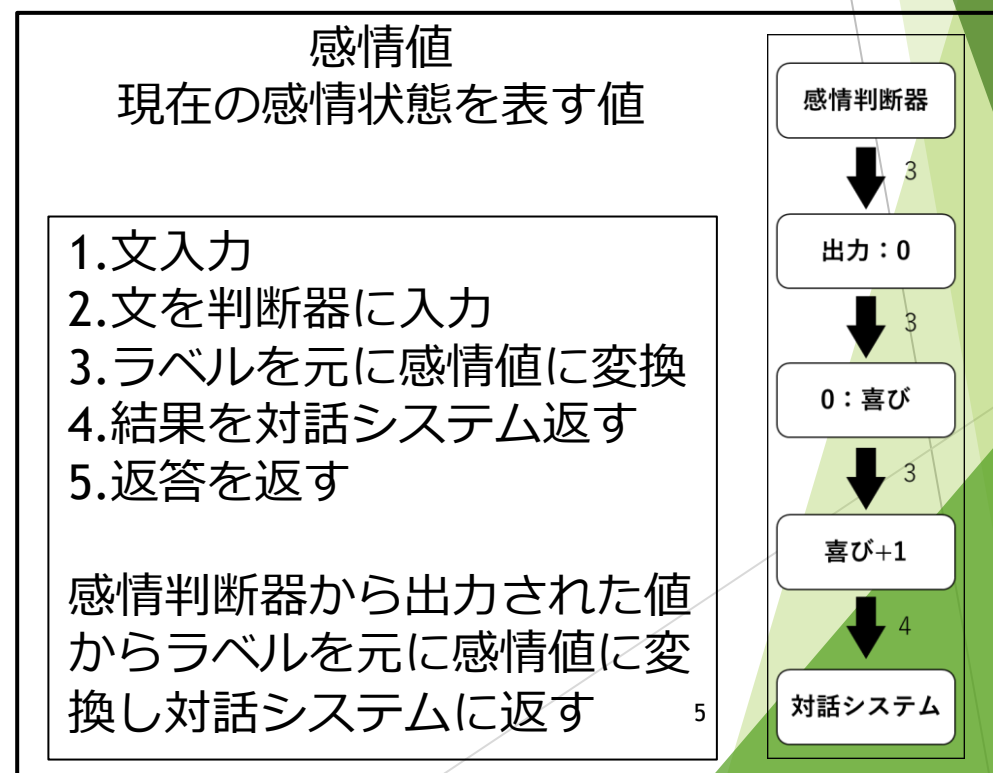
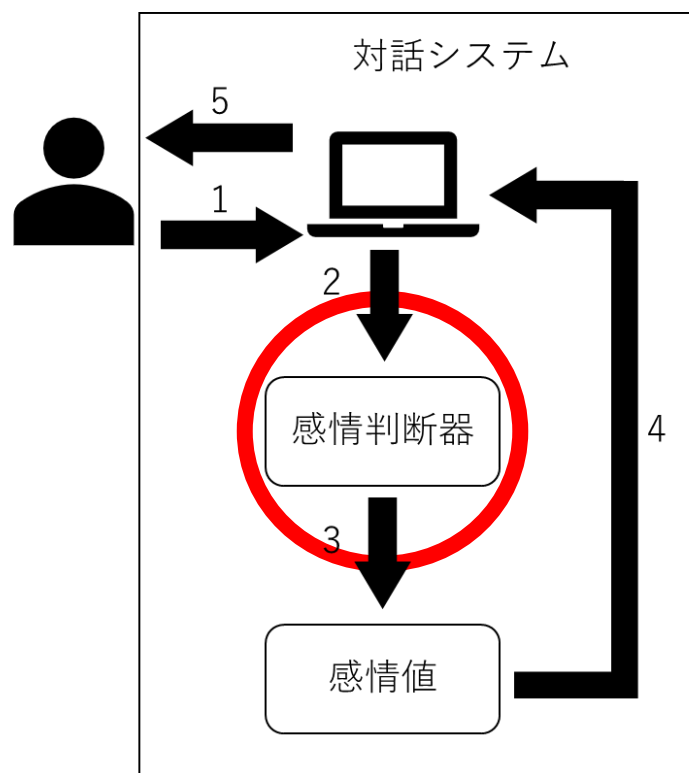
感情を考慮する対話システム

- 感情判断器を組み込むことで、対話システムは感情値の違いによる文章生成を行う



感情を考慮する対話システム

- 感情判断器を組み込むことで、対話システムは感情値の違いによる文章生成を行う



研究目的

- 対話システムから受け取った入力文から感情を判断し、対話システムに感情値を返す
- データセットを作り深層学習（ディープラーニング）を用いて感情判断器を作成する

関連研究1

- ML-Ask[3]（感情的要素/感情表現分析システム）
 - 感情推定システム
- 感情表現辞書
 - 感情を表すとされる単語を定義した辞書
 - 2100語収録されている

関連研究2

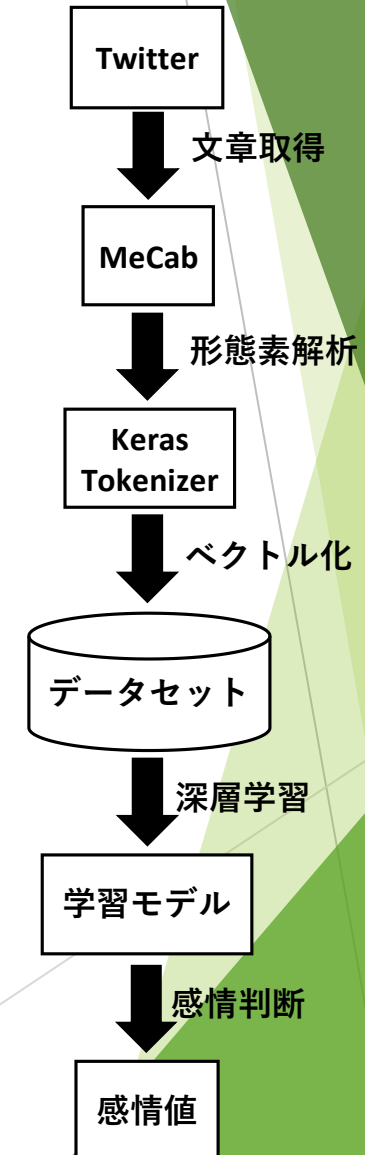
- 「会話文章からの話者の意図・感情判断[4]」
- オリジナル辞書
 - 感情を表すとされる文章を定義した辞書

現状と問題

- 一般的に感情を含んだ単語や文章を、どの感情に当たるのかを定義するのは困難
- 感情語の辞書を作るのには、多くの時間と労力を必要とする
- 辞書は特定の感情モデルに依存する

感情判断手法

- ▶ データセットを作り深層学習を用いて感情判断器を作成する
- ▶ 用意したデータセットを教師データとして教師あり学習を行った



プルチックの感情の輪

喜び joy	希望が達成された時や、やさしさを感じた時の爽やかな気持ち
信頼 trust	心配することなく、信じて安心できる気持ち
恐れ fear	害悪や危険な事柄に対して、逃避したいと感じる気持ち
驚き surprise	予期しない事象を体験した時の瞬間的な感情
悲しみ sadness	物事がうまくいかなかった時や、大切なものを失ったときに感じる残念な気持ち
嫌悪 disgust	憎み嫌い、不快に感じる気持ち
怒り anger	侮辱されたり傷つけられたりした時に起こる不愉快な気持ち
期待 anticipation	事柄が自分の思い道理になることを望む気持ち

感情分析の手法

起こしてくれてありがとう。



喜び

もうこんな時間か、起こしてくれてありがとう。



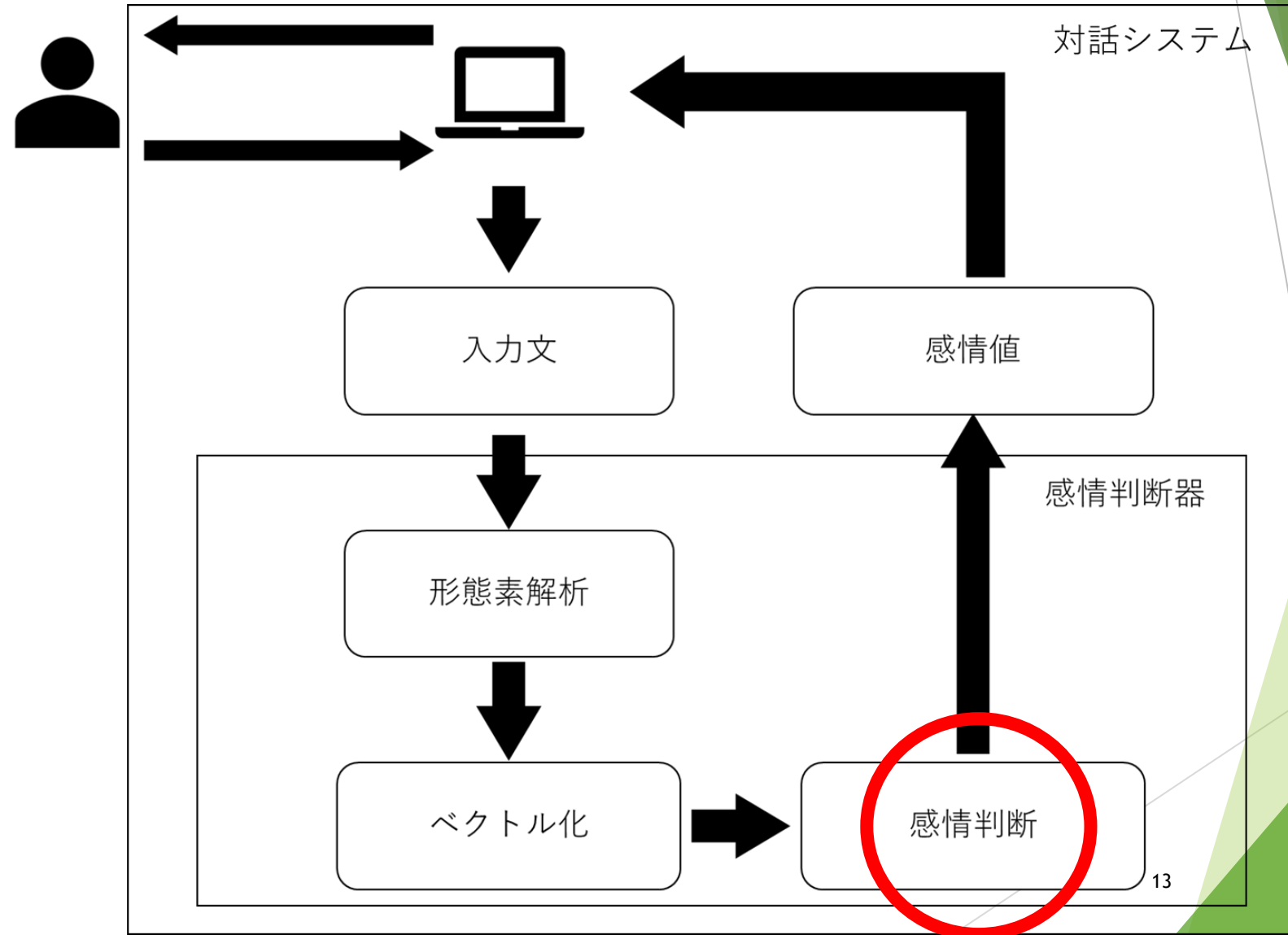
悲しみ

まだこんな時間か、起こしてくれてありがとう。

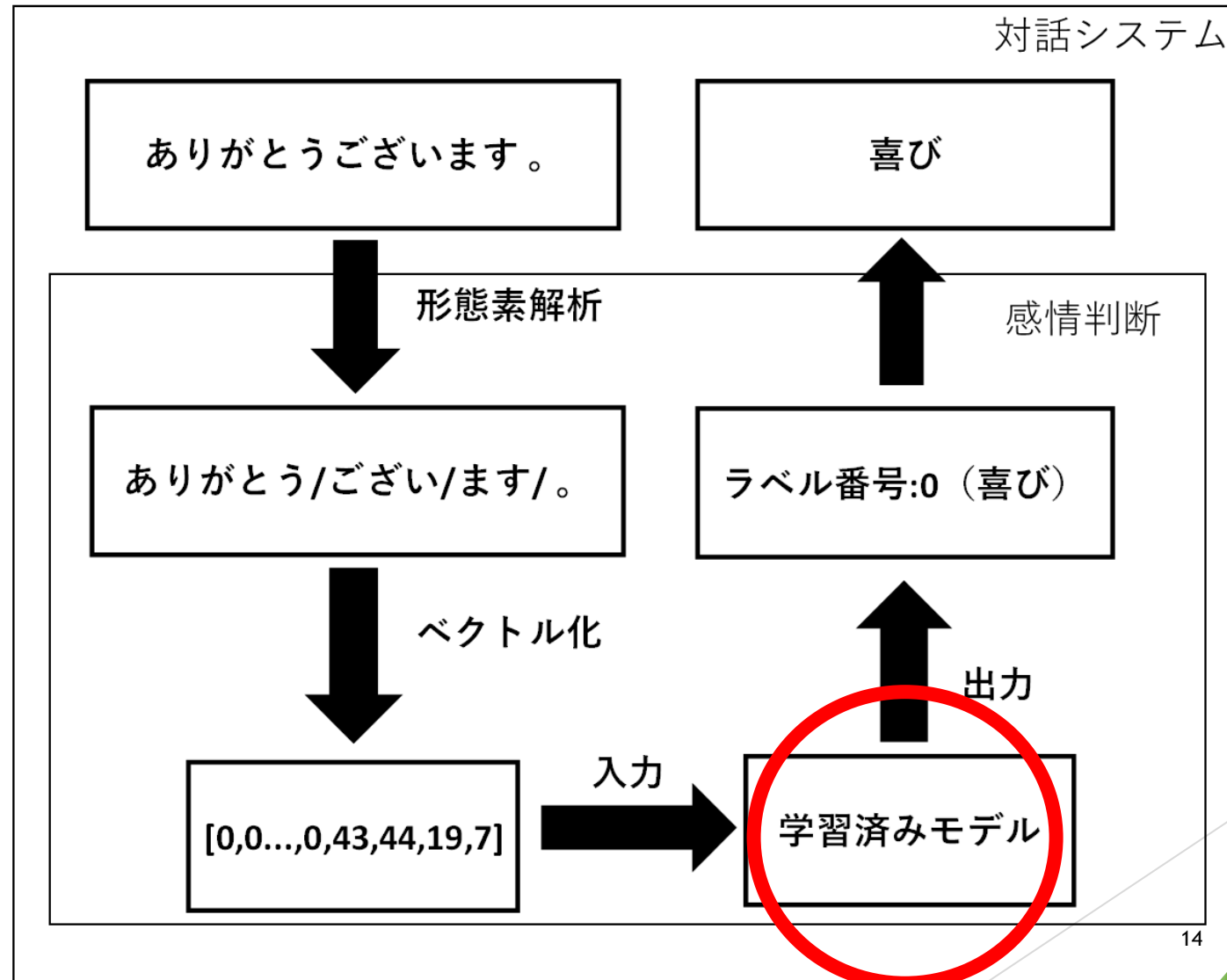


驚き

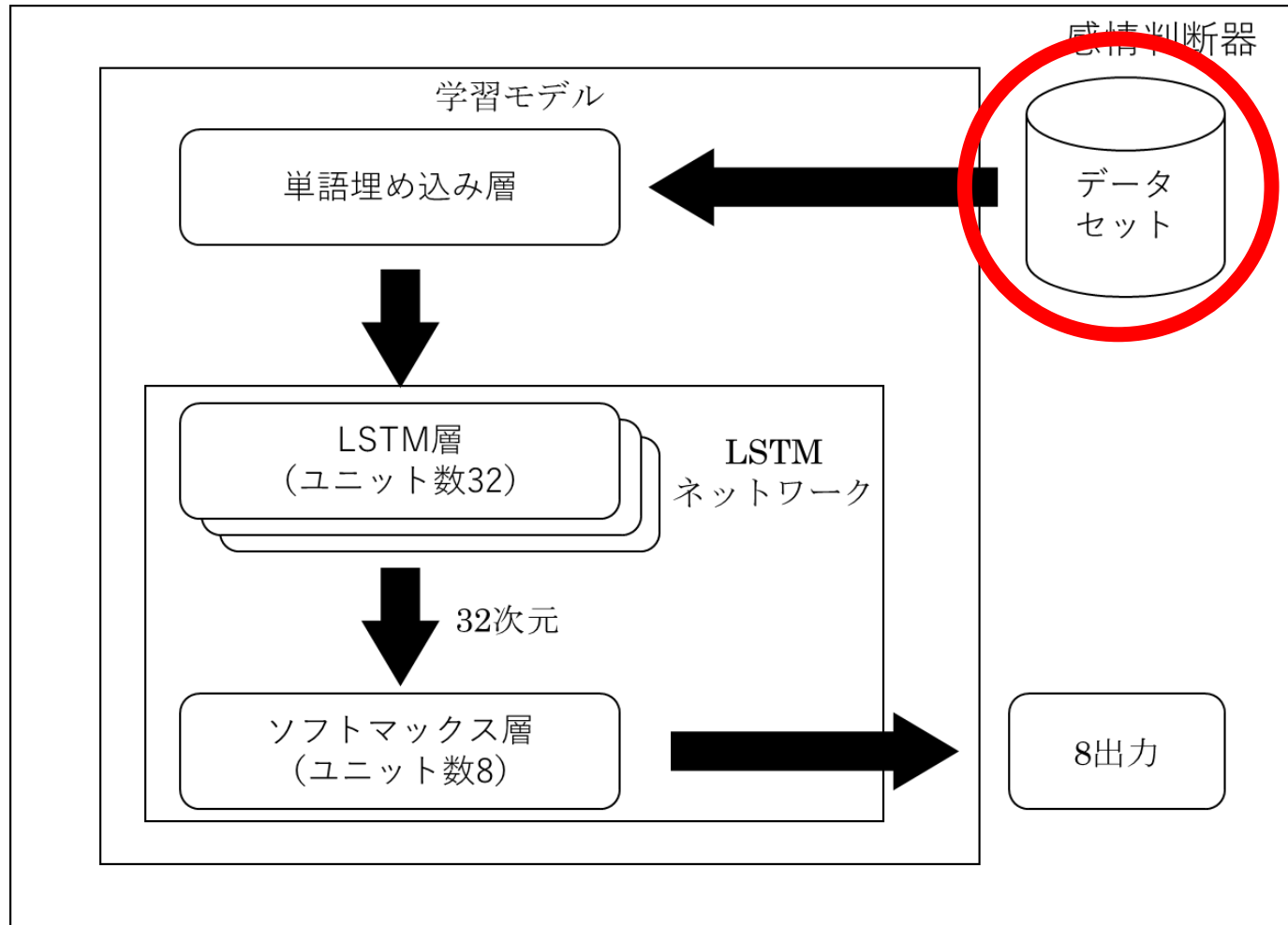
会話から感情判断までの流れ



対話システムと感情判断器



学習



訓練データ件数 : 63,857
テストデータ件数 : 15,964
単語埋め込み層のユニット数 : 100
LSTM層のユニット数 : 32
全結合層のユニット数 : 8
活性化関数 : ソフトマックス関数
学習回数 : 100
最適化手法 : Adam

データセットの作成手法

データセットの例

ありがとうございます！(^^)

(@_@) (T_T) Σ(° A° ;)
＼(^o^)/
(-_-;) (*´▽`*)

Twitterキーワード検索APIで
顔文字のついた文を収集、各パラ
メータ用のファイルにまとめる

Twitter

約7万9千件

喜び

悲しみ

信頼

嫌悪

恐れ

怒り

驚き

予期

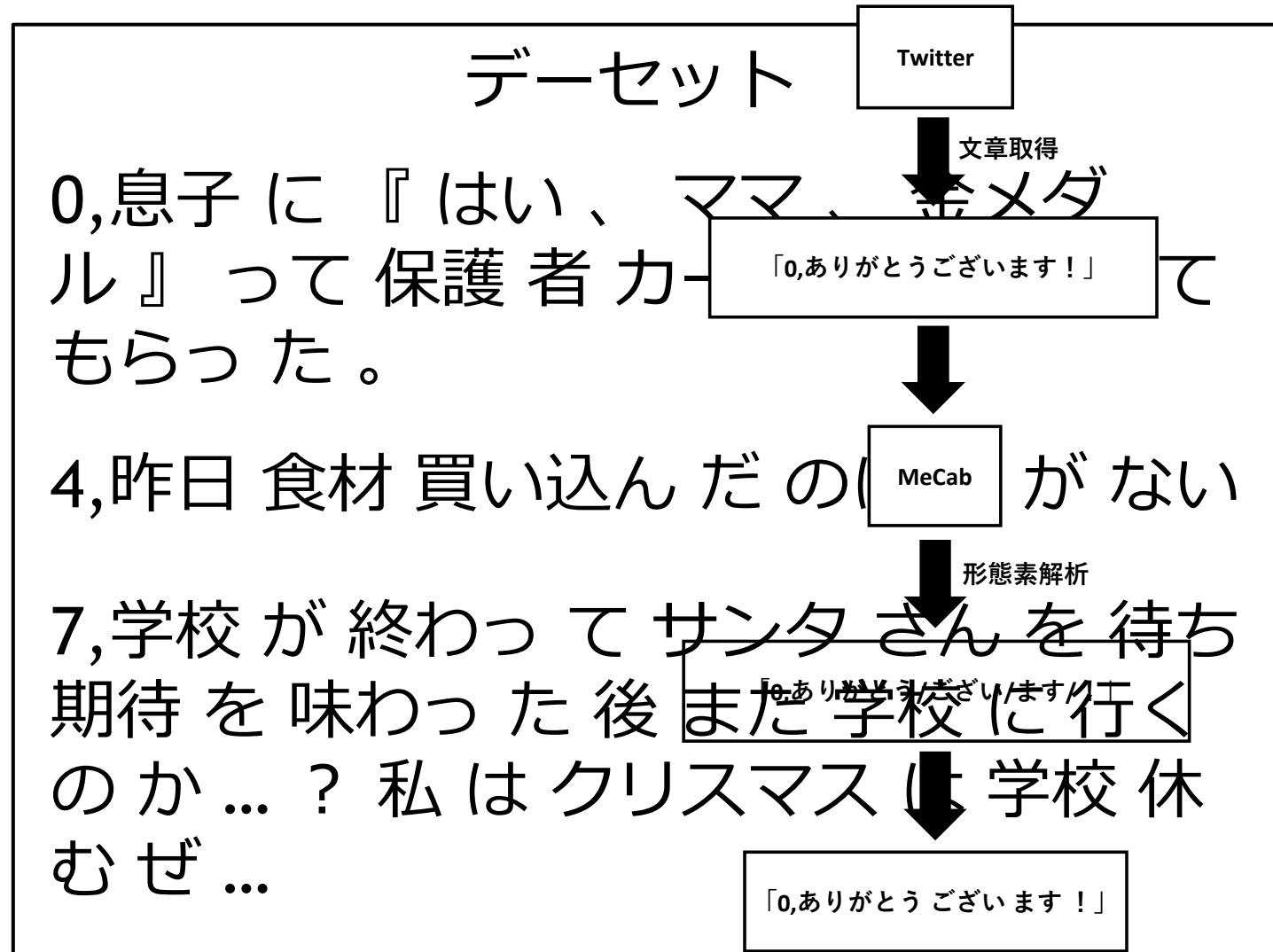
データセットイメージ

0	喜び	$(\cong \nabla \cong)$ $(\wedge \wedge)$
1	信頼	$\varepsilon - (' \vee ` * )$ $( ; ' ? `)$
2	恐れ	$\text{ガ}^* \text{カ}^* \text{ク}(((; ^\circ \text{D}^\circ))) \text{フ}^* \text{ル}^* \text{ル}$ $\Sigma (^\circ \Delta^\circ)$
3	驚き	$\Sigma (=^\circ \omega^\circ =;) \text{マジ} !?$ $\Sigma (^\circ \text{D}^\circ) \text{アア} !?$
4	悲しみ	$(/ _ \cdot \cdot) \text{シクシク}$ $(\rho ^\circ \cap ^\circ) \text{グ ス}$
5	嫌悪	$(' _ \succ `) \text{フーン}$ $(/ _ - ;) \text{ハア} \dots$
6	怒り	$\text{むっ} \dots (\overline{} \cap \overline{} \#$ $\text{ムカムカ} \dots ((\circ (-^\circ - ;)$
7	期待	$((\circ (^\wedge \nabla ^\wedge) \circ)) \text{わくわく}$ $\text{♪} ((\circ (\text{〃} \frown \nabla \text{〃}) \circ)) \text{♪} \text{ド}^* \text{キド}^* \text{キ}$

チューニングイメージ	
原文	# test @you ありがとうございます！(^.^) http://test.com
加工後	ありがとう ございます！

データセットの作成

例：喜びの顔文字「(^_^)」が含まれ文章を取得



0	喜び
1	信頼
2	恐れ
3	驚き
4	悲しみ
5	嫌悪
6	怒り
7	期待

提案

- データセットの前処理手法の提案[5]
 - データセットには重複データが存在する
 - 文に多く存在する接続語は学習時に誤認の原因になる可能性がある

提案

- 文の中には単語間をつなぐための接続語が存在する
- 文に多く存在する接続語は学習時に誤認の原因になる可能性がある

データセット	朝 <u>は</u> きり さめ <u>で</u> あった
加工後	朝 きり さめ あった

提案

接続語	接続詞	自立語（単独で接続語となる）	しかし だから つまり したがって なぜならば
	接続助詞	付属語（単独では接続語にならない）	～が ～に ～の ～は ～も ～を

嫌悪	今日 <u>は</u> 会社 休み ます
嫌悪	今日 <u>も</u> 会社 休み ます
喜び	今日 <u>は</u> 会社 行き ます

提案

接続語	接続詞	自立語（単独で接続語となる）	しかし だから つまり したがって なぜならば
	接続助詞	付属語（単独では接続語にならない）	～が ～に ～の ～は ～も ～を

今日 は ラーメン を 食べ に 行こう
今日 ラーメン 行こう
ラーメン 行こう
ラーメン

提案

- 文の中には単語間をつなぐための接続語が存在する
- 文に多く存在する接続語は学習時に誤認の原因になる可能性がある

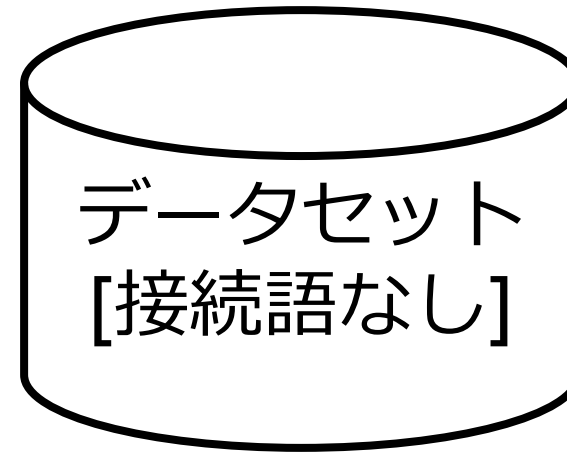
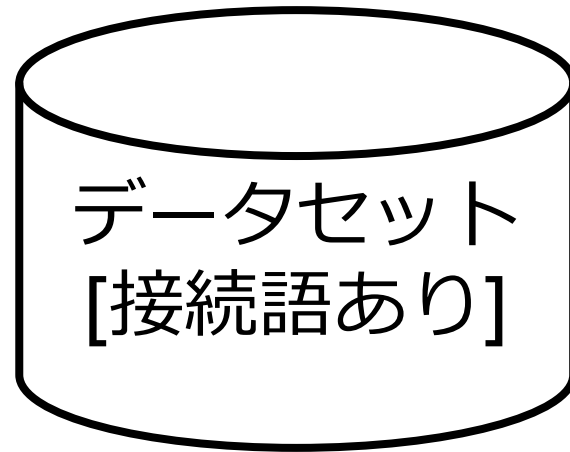
データセット	朝 <u>は</u> きり さめ <u>で</u> あった
加工後	朝 きり さめ あった

朝	は	きり	さめ	で	あっ	た
○	○	○	○	○	○	○

データセット

	件数
総単語数	1,722,475
接続語数	484,214
データセット(接続語あり)	79,821
重複データ	21,202
データセット(接続語なし)	58,619

実験と考察



	接続語あり	接続語なし
正解率	60.5%	62.5%

結論

- 接続語をそのまま分類に利用することは分類精度に悪影響を与える
- 比較的簡単にデータセットを作成することができた
- 会話文から感情を判断し、感情を返すことが可能

精度向上のための手法

- ベクトル化手法そのものを再検討
- 対策
 - 「Keras Tokenizerとword2vec」 2種類のベクトル化手法の比較

参考文献

- [1]河野 泉,池邊 亮志,和氣 早苗,上窪 真一,岩沢 透,西村 健士.感情表現を用いた対話システムEDSの開発(1):システム概要と感情モデル.情報処理学会研究報告,“HI,ヒューマンインタフェース研究会報告 89”, pp.43-48,(2000)
- [2]河野 泉,池邊 亮志,和氣 早苗,上窪 真一,岩沢 透,西村 健士.感情表現を用いた対話システムEDSの開発(2):システム概要と感情モデル.情報処理学会研究報告,“HI,ヒューマンインタフェース研究会報告 89”, pp.49-56,(2000)
- [3] Michal Ptaszynski,Pawel Dybala,Wenhan Shi,Rafal Rzepka,Kenji Araki.A system for affect analysis of utterances in Japanese supported with web mining.日本知能情報ファジイ学会誌,vol.21,No.2,pp.194-213 (2009)
- [4] 山本裕司,吉村枝里子,土屋誠司,渡部広一.会話文章からの話者の意図・感情判断.情報処理学会研究報告,vol.2010-ICS-158,No.2 (2010)
- [5] 大澤泰我,ケネス マッキン,永井保夫.対話システムのための感情判断器の検討.情報処理学会第82回全国大会.1S-06 (2020)