



大規模対話モデルにおける多様な情報の利活用

杉山弘晃

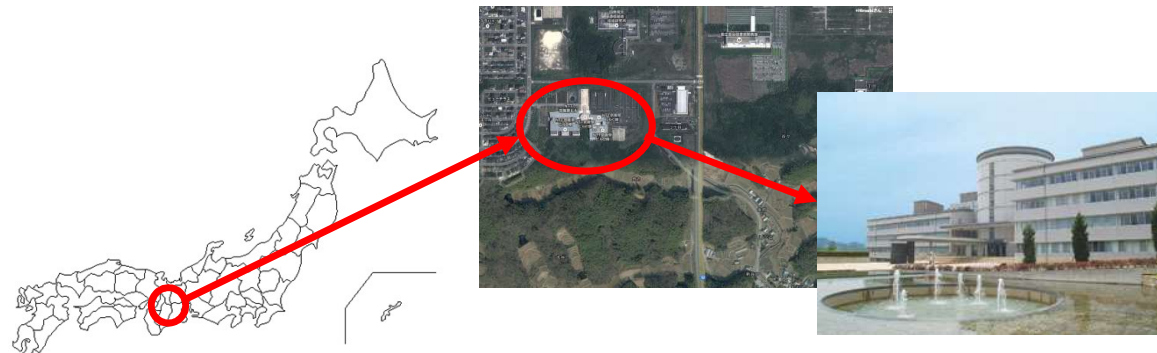
NTT コミュニケーション科学基礎研究所

2022/11/30 産総研AIセミナー

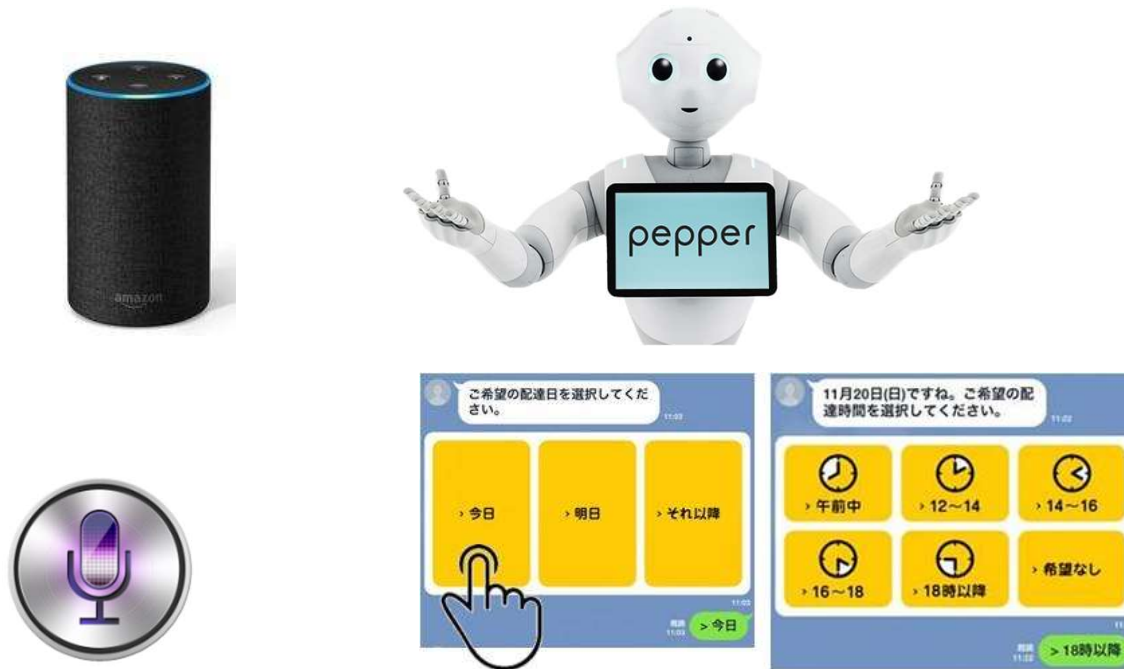
自己紹介



- 1985.02：青森県生まれ
- 2009.03：東京大学 修士課程修了
 - ロボットの身体を活かした言語獲得（國吉・原田研究室）
- 2009.04～：NTT 入社（コミュニケーション科学基礎研究所）
 - 雑談対話システム
 - 英語問題の自動解答（「ロボットは東大へ入れるか」プロジェクト）
 - 乳幼児の言語獲得過程の解明
- 2014.10～2016.09：奈良先端大 博士課程（中村研究室）
- 2019.06～：新学術領域 対話知能学「行動決定モデル推定研究グループ」研究代表
- 勤務地：京阪奈



身の回りに広がる対話システム



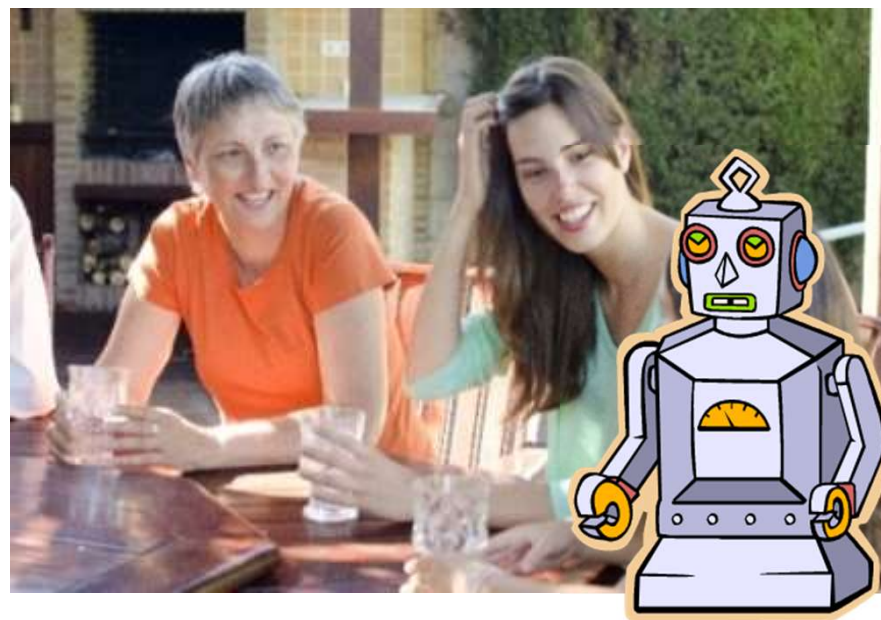
従来の身の回りの対話システム
決まったタスクの達成を目的とするものがほとんど

人の雑談相手となる対話システム



システムと話す利点

- **いつでも**話せる
(時間や場所の制約を受けない)
- **気楽に**話せる
(遠慮がいらない・相手からの評価を気にしなくてよい)
- **深い話題も**話せる
(細分化した趣味にも対応できる・プライベートな話も相談できる)



対話システムが単なるタスク達成のツールではなく、
雑談を楽しむ相手として期待され始めている

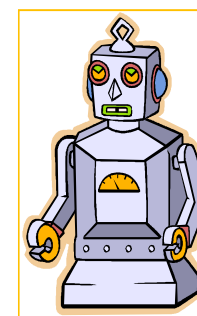
- ・ ユーザ発話の話題の多様さ
- ・ 対話の流れの多様さ



- 多様な話題が含まれるユーザ発話への一問一答での応答の実現
- 対話の深さよりも幅を優先（文脈を考慮すると話題の組み合わせ数が爆発）



ラーメンはスープが命だよね



システム

本日の講演概要



雑談対話システム研究の流れと、直近の急速な進展をご紹介します

1. 一問一答型雑談対話システム
2. 深層学習研究の急速な進展と対話システムへの応用
3. 大規模対話モデルへの外部情報の導入

1. 一問一答型雑談対話システム

- 一問一答型雑談対話システムの構築

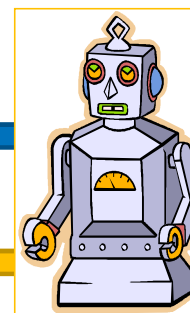
→ 予め入力と出力のペア（**応答パターン**）を大量に用意しておく、
ユーザ発話と類似するパターンを検索して応答する



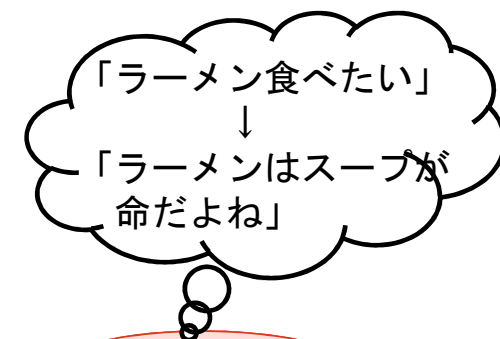
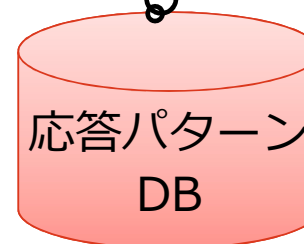
ユーザ

あー、ラーメン食べに行きたいなあ

ラーメンはスープが命だよね



システム

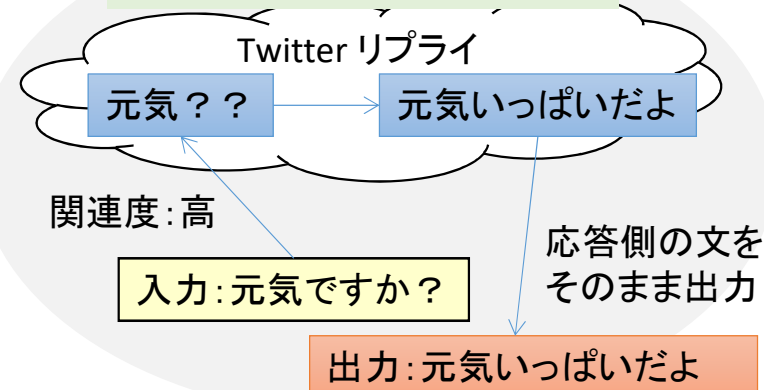


代表的な応答パターン作成手法

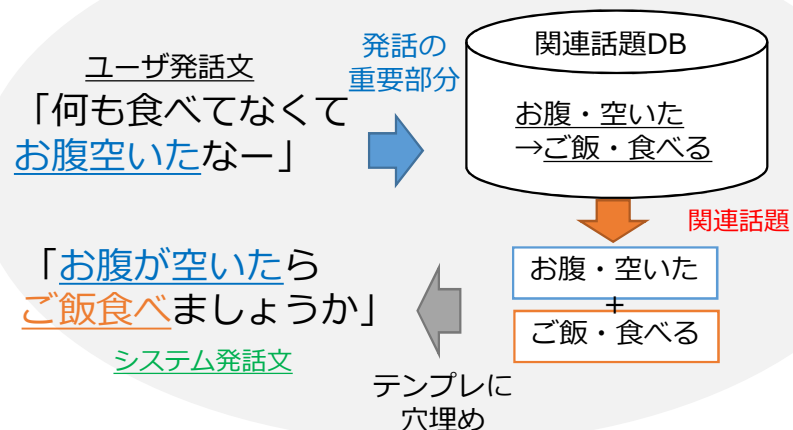
①ルールベース

```
<category>
<pattern>お酒 * 飲め * か</pattern>
<template>お酒好きです</template>
</category>
```

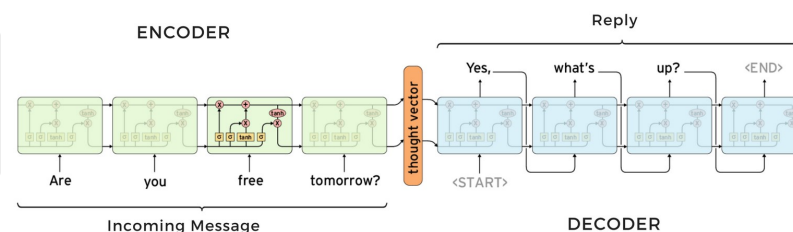
②用例ベース



③テンプレートベース



④深層学習ベース

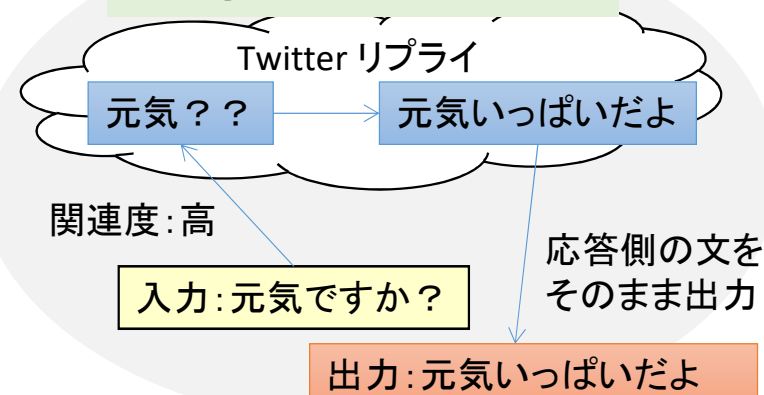


代表的な応答パターン作成手法

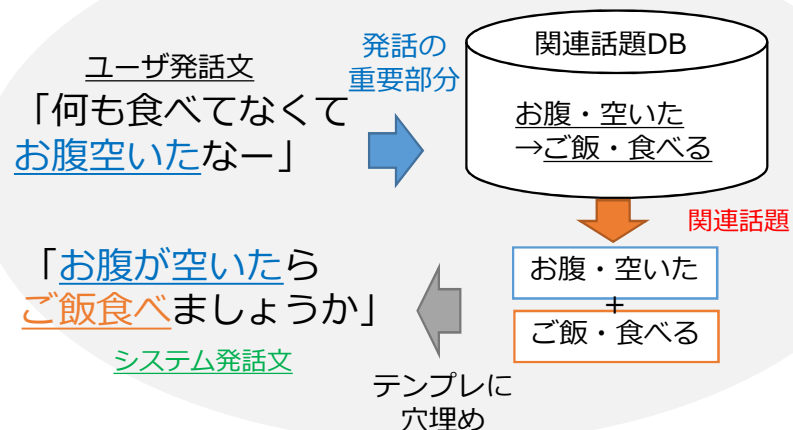
①ルールベース

```
<category>
<pattern>お酒 * 飲め * か</pattern>
<template>お酒好きです</template>
</category>
```

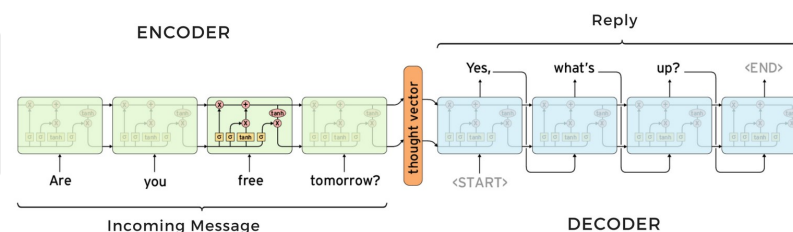
②用例ベース



③テンプレートベース



④深層学習ベース



A) ルールベース対話システム

- 人手で応答パターンを記述して応答
- 現在もほとんどの商用システムで採用

```
<category>  
<pattern>お酒 * 飲め * か</pattern>  
<template>お酒好きです</template>  
</category>
```

作成された
応答パターンの例

- 😊 ルールが想定する範囲内では高品質
- 😊 発話内容を完全に制御可能（不適切な発話をするリスクが小さい）
- 😞 ユーザ発話をカバーするには膨大なルールが必要

ルールベースの対話例

14万程度のルールを実際の人同士の発話を参考に構築

User> 今日の予定は？
System> ずっと研究室で実験です。余裕が出来たらジム行きます。何か予定ありますか？
User> がんばれー。何の研究しているの？
System> **話題を変えましょう。**ドラマは観ますか
User> 見ないです。
System> いつ放送しているのですか
User> 放送してない。
System> **話題を変えましょう。**旅行は好きですか

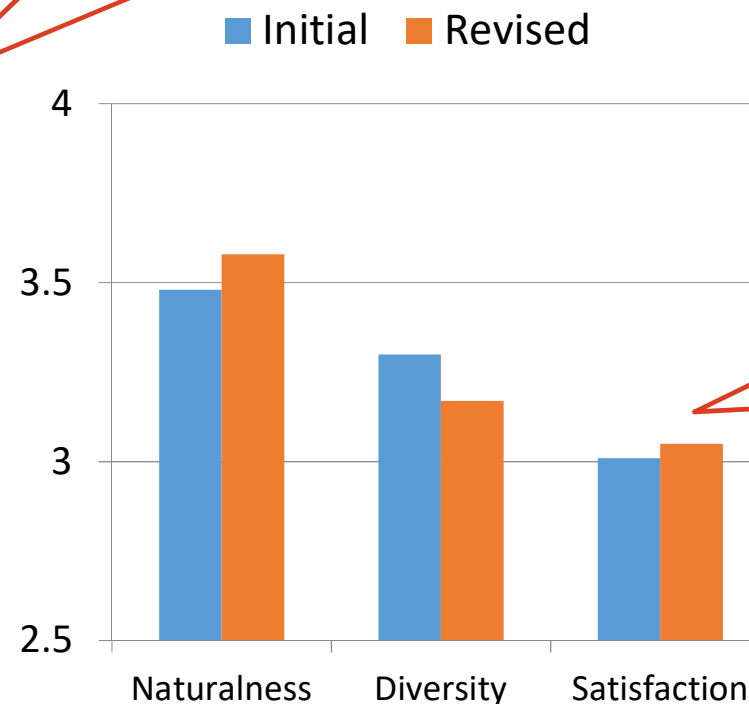
ルールベース：
ルールが該当しないケースが多く
対話が止まってしまう

じゃあ膨大な量のルールを作ってしまうばよい？？

膨大なルールがあればよい？

ルールの種類	オリジナル	ルール倍増版
aiml	18	22
category	149,300	333,295
template	149,300	333,295
pattern	149,300	333,295
topic	6	0
srai	38,586	29,928
li	107,429	53,571
that	1,962	9,416
random	2,495	9,138
star	14	709
sr	0	12

ルールを大量に追加



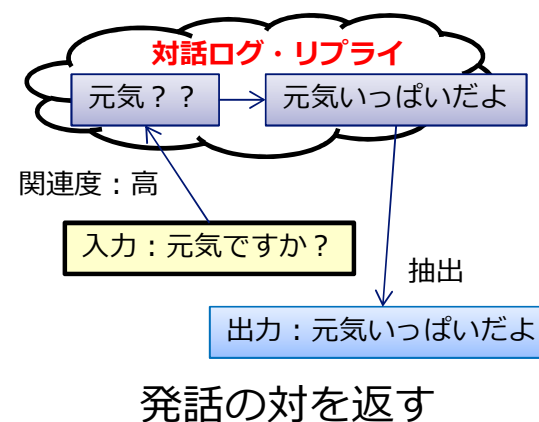
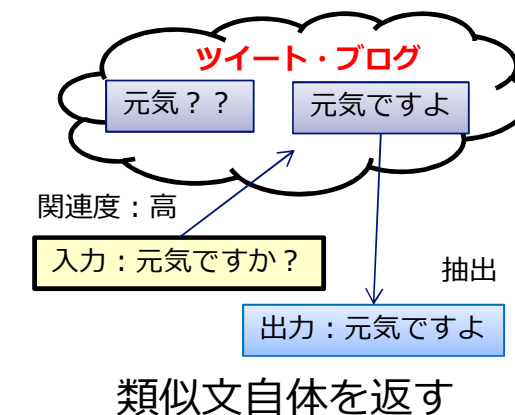
ほぼ変化なし

単純にルールを倍増させてもほとんど改善されない
 (☹️ ルールを増やすとルール同士の衝突も増えてしまい
 応答の妥当性や一貫性を維持できない)

B) 用例ベース対話システム

- 大規模データから文（用例）を抽出・検索して発話に利用
 - ルールを作成することは高コスト、かつ有効性が薄い
 - Twitter等のConsumer Generated Mediaが流行り始めた時代
 - 2種類のアプローチ
 - 類似文自体を用例として利用（ペア不要のため個数が多い）
 - 新聞記事、ブログ、ツイート単体
 - 発話の対を用例として利用（応答としての適切さを期待）
 - 対話ログ、ツイートリプライ、QA
- 😊 ほぼ全ての話題に何かしら応答可能
- 😞 類似文を使う場合：オウム返しになりがち
- 😞 発話の対を使う場合：文脈が合わず関連性の薄い発話が出力されがち

[Ritter+ 2011]



用例ベースの対話例



4500万ツイートペア（発話対）を用例として利用

System> どうもです

User> どうもですー

System> バイトずる休みしたろ(-_-)

User> してないよ。ってというかバイトしてないし。。

System> お、おお…よく生活できてますね…。

User> はい。無駄遣いしなければ生活できますよ。。w

System> E1区域とはブルー部分のの1番ですか？

User> ちょっと意味がわかりませんー。

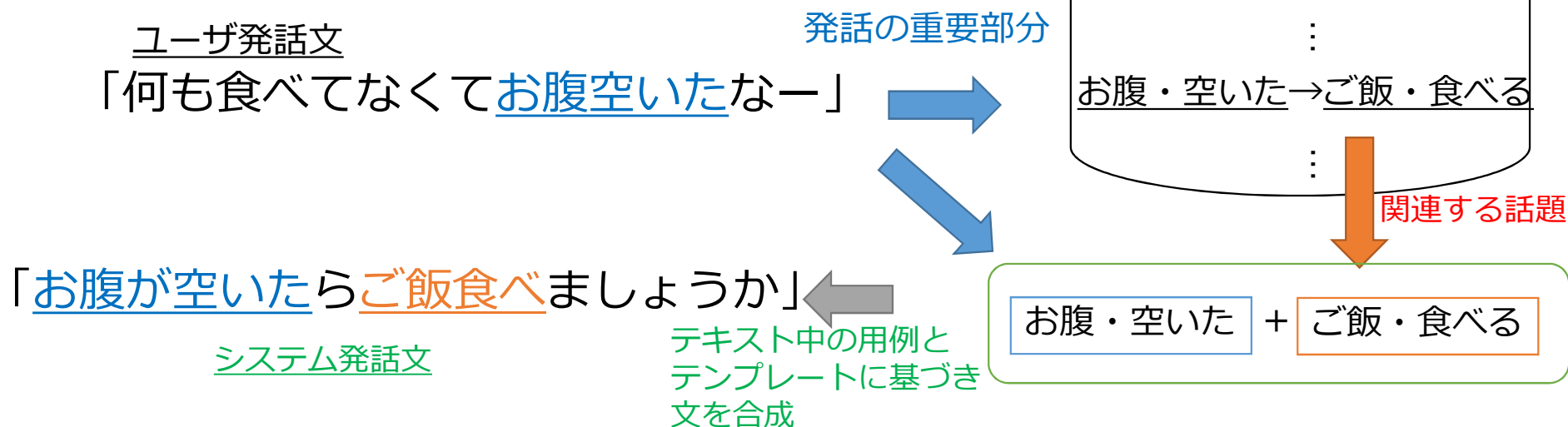
System> うん……うん？

用例ベース：
いきなり変なことを
言って対話が破綻

何を言うか予測不能なため、商用雑談システムでの採用は
ほぼ見られない(FAQ検索のような狭いドメインでは一般的)

3. テンプレートベース対話システム

- 用例ベースの話題の幅を活かしつつ
応答発話の品質を向上させる試み



高品質な応答発話を生成

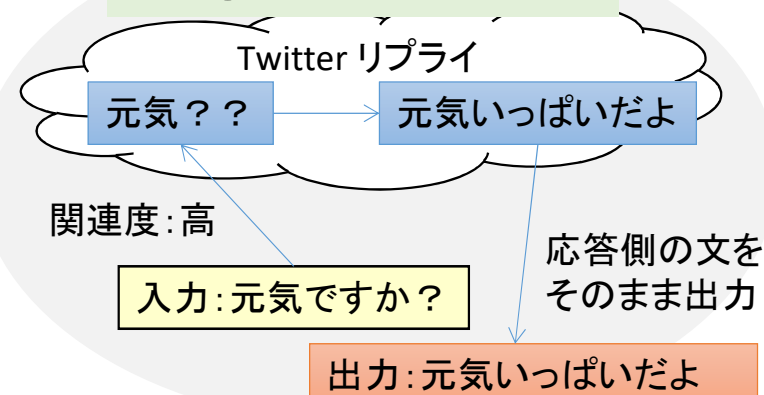
- **重要な部分**のみを用いて検索する→無関係な文を抑制する
- **関連する話題**を探して合成する→オウム返しを抑制する

代表的な応答パターン作成手法

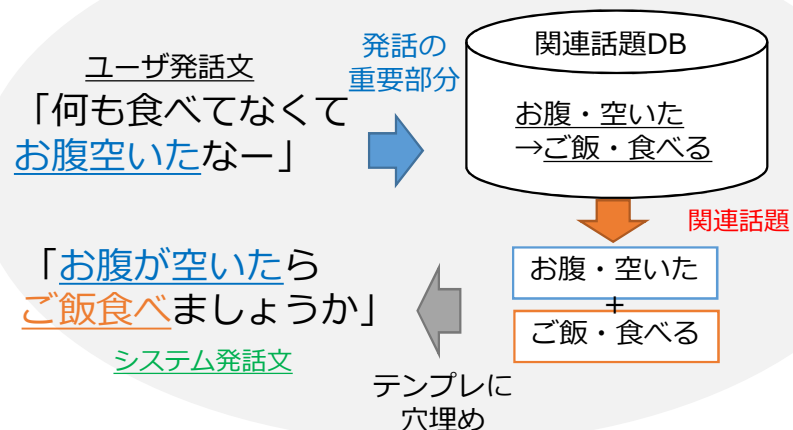
①ルールベース

```
<category>
<pattern>お酒 * 飲め * か</pattern>
<template>お酒好きです</template>
</category>
```

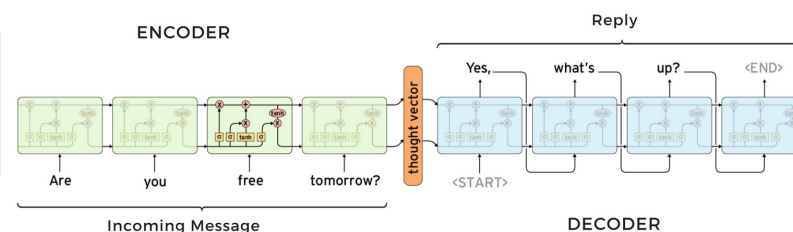
②用例ベース



③テンプレートベース



④深層学習ベース



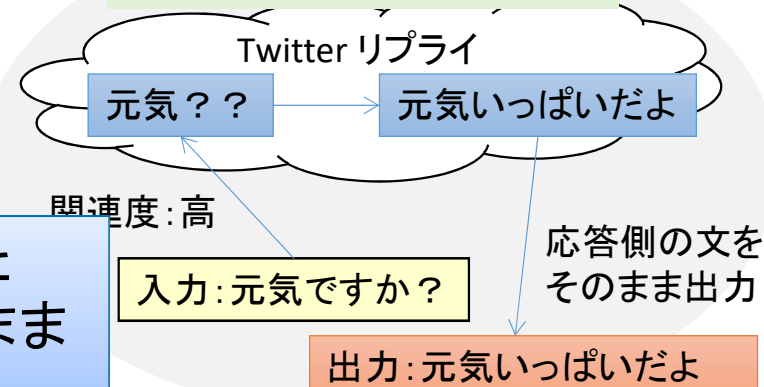
代表的な応答パターン作成手法

①ルールベース

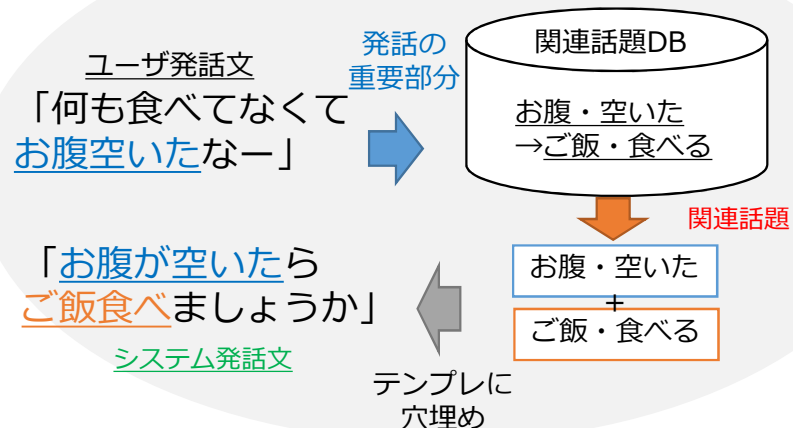
```
<category>
<pattern>お酒 * 飲め * か</pattern>
<template>お酒好きです</template>
</category>
```

人が作成した
文字列がそのまま
使われている

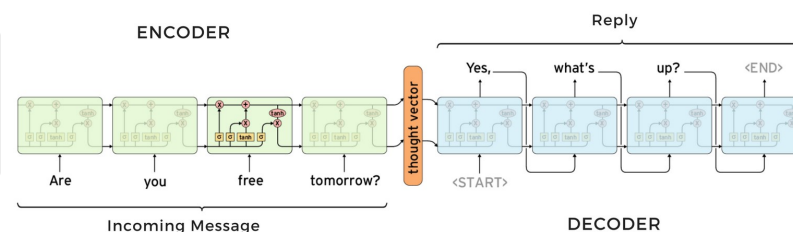
②用例ベース



③テンプレートベース



④深層学習ベース

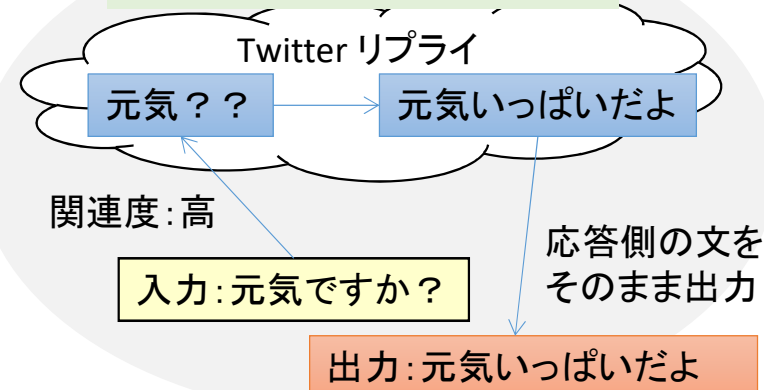


代表的な応答パターン作成手法

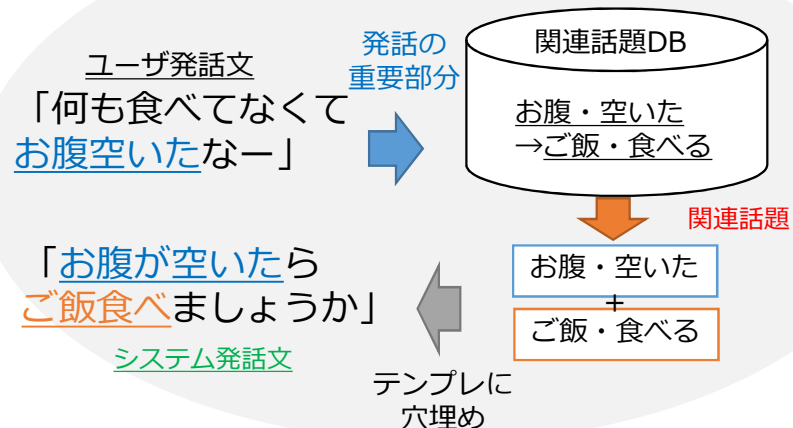
①ルールベース

```
<category>
<pattern>お酒 * 飲め * か</pattern>
<template>お酒好きです</template>
</category>
```

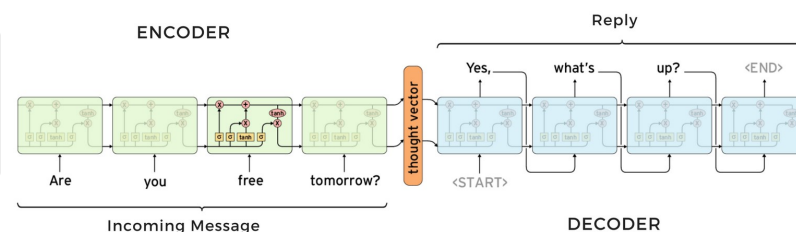
②用例ベース



③テンプレートベース



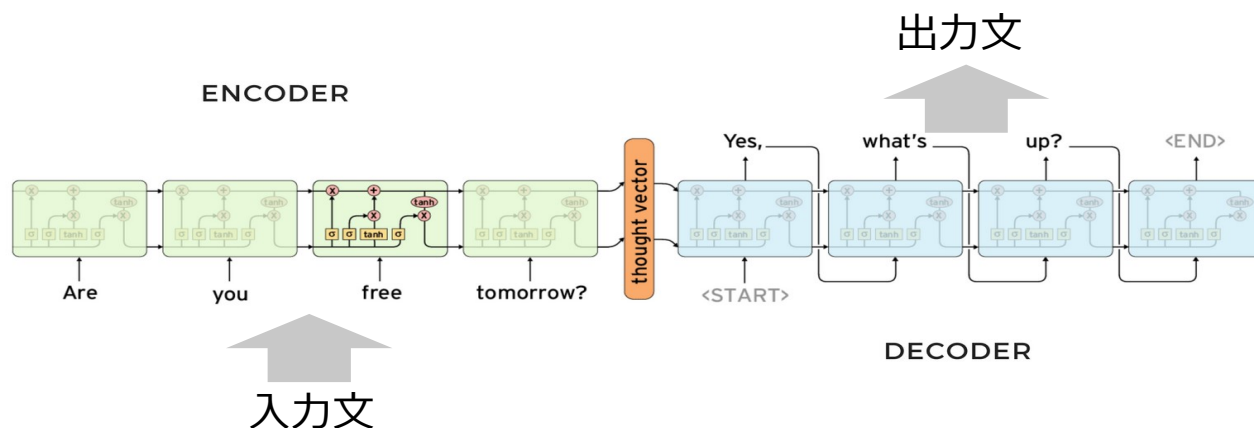
④深層学習ベース



機械が直接文を生成する初の手法

深層学習を利用しEnd-to-endで発話生成する方法（2015-）

- End-to-end：入出力の発話ペアを大量に集めて直接対応関係を学習
- 人が作成した文字列がそのまま出てこない，初の「文生成」手法



- 機械翻訳・音声認識等と同じ枠組み
 - › 機械翻訳：入力＝翻訳元言語 出力＝翻訳先言語
 - › 音声認識：入力＝音声 出力＝対応する言語表現

- 学習に用いる対話データの量やモデルパラメータが少なく，無難な発話が多め（計算量・公開データ量の問題）

Human: *hello !*

Machine: *hello !*

Human: *how are you ?*

Machine: *i 'm good .*

Human: *what 's your name ?*

Machine: *i 'm julia .*

Human: *when were you born ?*

Machine: *july 20th .*

Human: *what year were you born ?*

Machine: *1977 .*

Human: *where are you from ?*

Neural Conversational Model
[Vinyals+ 2015 (Google)]

本日の講演概要



雑談対話システム研究の流れと、直近の急速な進展をご紹介します

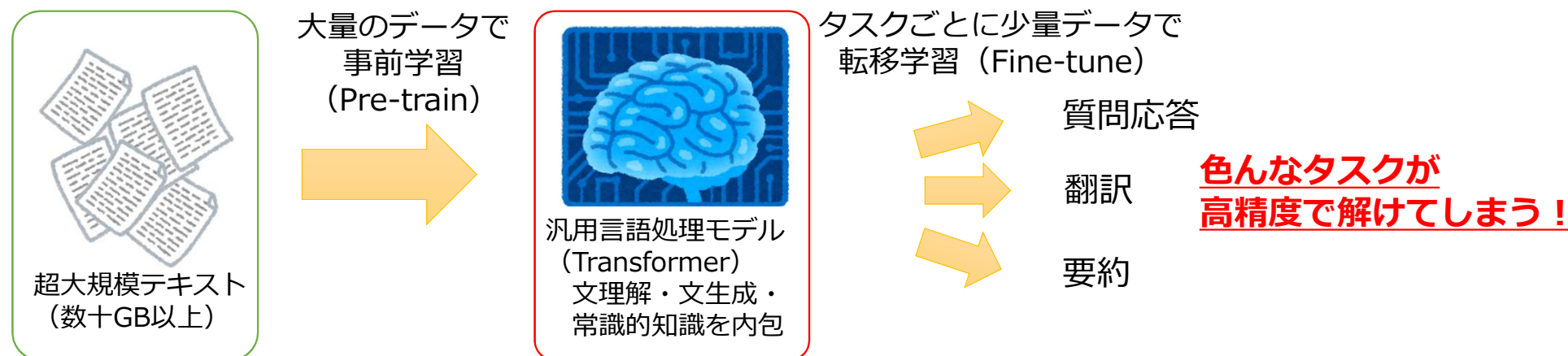
1. 一問一答型雑談対話システム

2. 深層学習研究の急速な進展と対話システムへの応用

3. 大規模対話モデルへの外部情報の導入

自然言語処理における大事件

BERT/GPT (Transformer + 事前学習) の出現



例：質問応答（東ロボ）

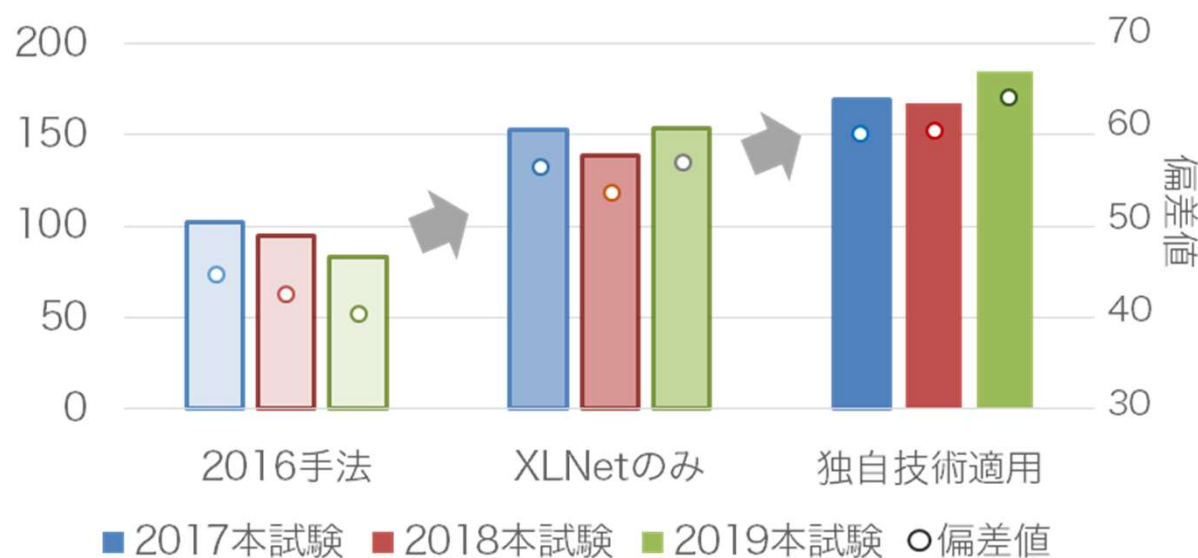
[NTT・NIIプレスリリース 2019/11/18]



2019年大学入試センター試験英語筆記科目において

AIが185点を獲得！

センター英語本試験の成績変化



- Transformerの一種のXLNetを単純に適用するだけで従来手法から**50点程度向上**
- XLNetの学習の工夫や、独自のセンター英語解答手法の適用により31点向上
(2019本試験の偏差値64.1・東大合格者の平均値とほぼ同程度)

例：要約 ELYZA DIGEST

[ELYZA INC 2020]



コールセンタの対話

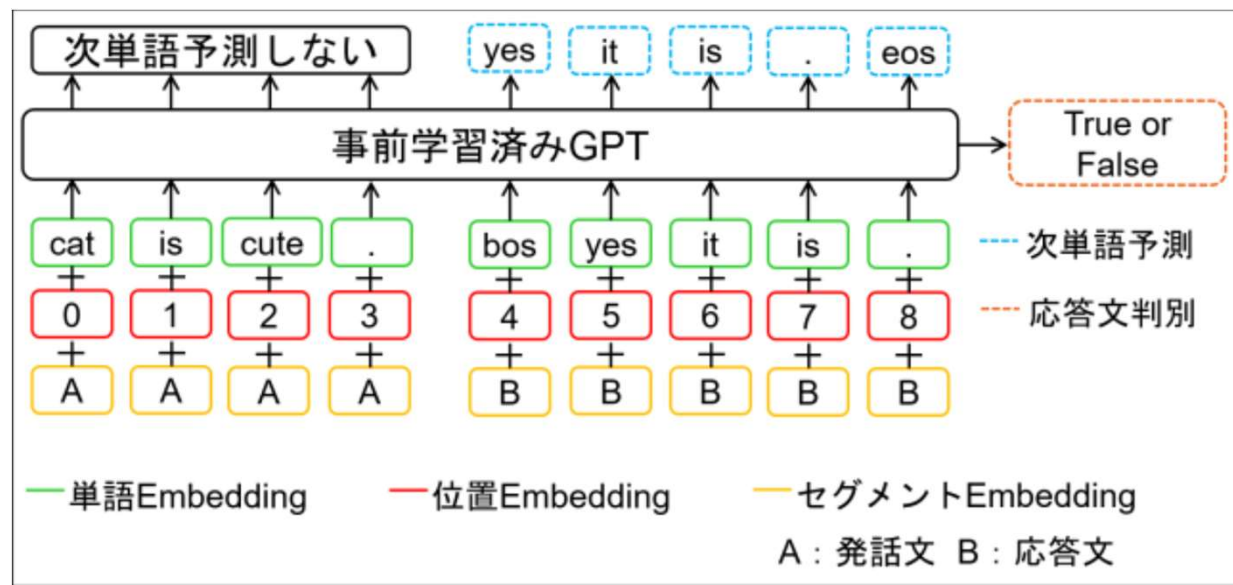
[お客様]もしもしイライザ花子です。
[オペレーター] お電話ありがとうございます。イライザブライダルのイライザ次郎と申します。
[お客様] [あっ]すいません。お世話になっております。
[オペレーター] お世話になっております。
[お客様] 昨日[あの]お電話いただいた、2022年2月2日フライトの、ハワイでの式の件なんです。
[オペレーター] はい。
[お客様] えー。えー。すいませんお電話遅くなりまして。はい。
[オペレーター] [あ]新。はい。ええ。ありがとうございました。[あ]すいませんちょっとくればしながらこちらのイライザ花子さんの見積もり[ん]先日郵送したんですけども。こちら問題なく到着しましたでしょうか。
[お客様] ツノリ。ちょっと待ってくださいね。[あー]2まいの、見積もり、5ひょうまん円のやつですね。はい、私のところにあります。
[オペレーター] はいそうです。でそちらの見積もりへの補足となりますが、新郎新婦様の飛行機代は含まれておりますが、その他の方の飛行機代やホテル代等は含まれておりません。
[お客様] ええ。とうし話していたときにはそれもオプションにふくまれるといていたがするんですが。うん。はい。
[オペレーター] この点については、こちらとしては申し訳ありませんが含まれていないということでご理解いただくしかないと存じます。ただこれまでのご検討がありますのである程度特例的に追加させていただくこともありえると思っていて今回はできないってところでございまして。ええ。[ま][ちょっと]こちらとしてもイライザ花子さんとお打ち合わせの上で、今後の方針方向性については、お打ち合わせしていきたいと思っているんですけど。うん。[ま]
[お客様] ええ。
[お客様] ええ。
[お客様] はい。
[お客様] ええ。
[オペレーター] んで、[ま][そのー]イライザ花子さんとして、協議の結果できそうなポイントが、あと2人分追加することなんですかね。[ま]どの程度あるのかっていうところを、[あのー][ま]感触として伺えたらと思ってのご連絡だったんですけど、[ま]。[あ]はい。
[お客様] [あー]そうなんですね。だからそれはちゃんと来ては先をし、[あの]お伝えしたとおりなんですね。[ま]もうこちら、[あのー]すまやかにいかいけつできればと思ってるので、そんな[あのー][う]拝見。[ん]何でもかんでも交渉っていうつもりはもういいです全くないんで、わたしもそうですねそうだとなくです。えーそうです。
[オペレーター] ええ、分かりました。
[お客様] はいはい。そうでございます。ちょっとそういうふうな、追加して、何とかお願いしたいです。
[オペレーター][えーと]ではちょっと打ったさんのほうだった。はい。[あの]方針については検討した上で改めて、ご連絡いたします。
[お客様] だからたのかなってね。はい、はいすいません。よろしくお願いします。はい。
[オペレーター] はい。失礼いたします。
[お客様] 失礼します。

自動要約結果

イライザブライダルのイライザ次郎が、ハワイでの結婚式の見積もりを提示した。新郎新婦の飛行機代は含まれているが、その他の方のホテル代は含まれていない。イライザ花子と今後の方針方向性については、お打ち合わせする予定だという。

GPTの対話への適用

- TransferTransfo [Wolf+ 2019]
 - 事前学習されたTransformerを初めて対話に適用したモデル
 - 応答文判別（次発話の妥当性判定）タスクとのマルチタスク学習
 - LSTMやMemory Network等比べ性能（ppl）は大幅に改善



[林+ 2020]

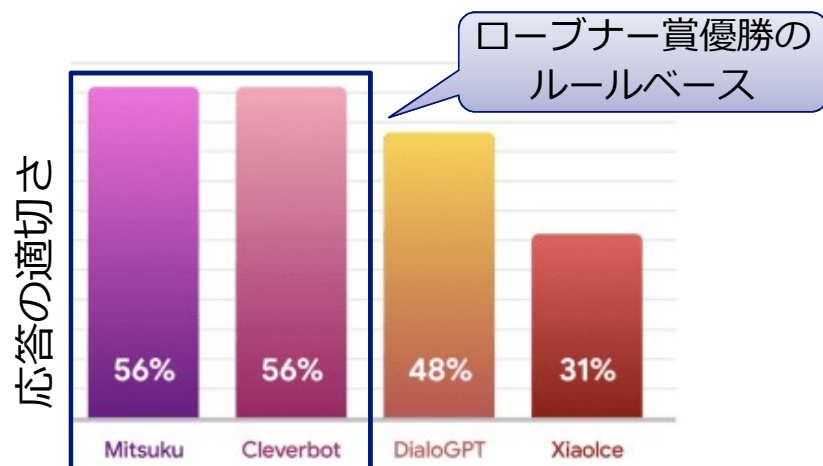
GPTの対話への適用 (cont)



- DialoGPT [Zhang+ 2019]
 - GPT-2をReddit（掲示板）データから抽出した1.47億対話でfinetuneしたモデル
 - より大規模なモデル・データの利用により，従来モデルから大幅に性能を向上
 - › 人と同等以上と著者らは主張
- ルールベースよりも評価は低い&対話としての性能はいまいち
- とはいえ現在もぼちぼち使われるベースラインモデル（RTX1080とかでも動く）

Human-like: A and B, which is more likely to be generated by human rather than a chatbot?

System A	Neutral	System B
DialoGPT (345M, w/ MMI)	2978 (50)%	241 (4%) 2781 (46%)
		Human response *

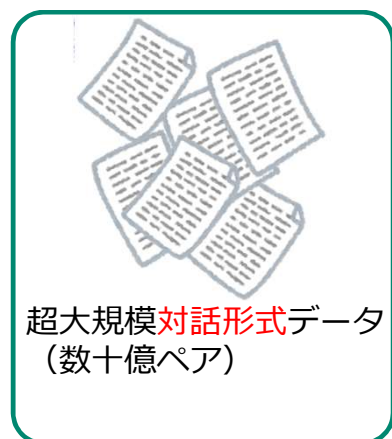


Role	response
User	Does money buy happiness?
Bot	Depends how much money you spend on it .
User	What is the best way to buy happiness ?
Bot	You just have to be a millionaire by your early 20s, then you can be happy .
User	This is so difficult !
Bot	You have no idea how hard it is to be a millionaire and happy . There is a reason the rich have a lot of money

雑談対話システム研究にも黒船が・・・



- 2019年まで：ルールベースが深層学習に対して互角以上の性能
- 2020年：世界が一変** = 大規模Transformerが極めて自然な応答生成を実現



超大規模対話形式データ
(数十億ペア)

大まかな
応答関係や
文の自然さを
事前学習
(Pre-train)



汎用対話モデル
(Transformer)
大まかな応答関係・
文の自然さを内包

望ましい
応答関係を
少量データ
で調整
(Finetune)



BlenderBot

ユーザ

生きてる人でも死んでる人でもいいんだけど、
一番夕飯を一緒に食べてみたい人って誰？

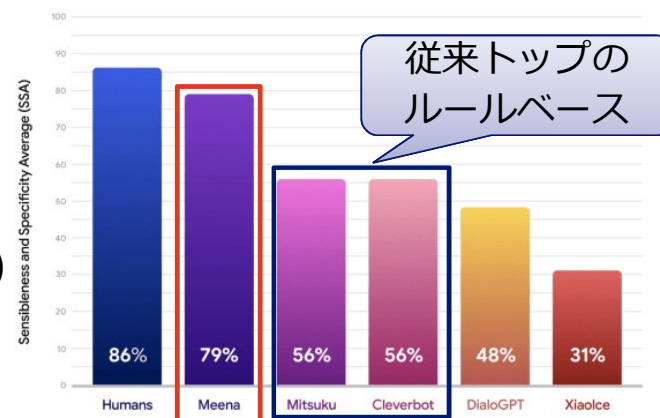
難しいなあ・・・挙げるならスティーブジョブス
かなあ。あの人の知恵を借りたいよ。

あ、面白いね。スティーブジョブスについて
どんなこと知ってるの？

技術の歴史において最も影響を与えた一人だね。
先見の明がある。

BlenderBot (Facebook)[2] 対話例和訳

- Meena [Adiwardana+ 2020] (Google: 2020/02・26億パラメータ)
 - ルールベースシステムを初めて明確に上回る
- BlenderBot [Roller+ 2020] (Facebook: 2020/04・27億/94億パラメータ)
 - 対話に必要な3種類のスキル（個人性・知識・共感）を学習し統合
 - Meenaを同サイズモデルで上回る（性能は27億 ≧ 94億）



NTTの日本語雑談対話システムへの取り組み

- 日本語最大規模（16億パラメータ）の対話モデルを構築（2020/11）
 - 利用した対話データ量では世界最大規模（21億応答ペア）
 - 対話システムライブコンペティション3で最優秀賞
 - ベースとなるモデル（事前学習モデル）を学術目的で無償公開

システム

こんにちは。よろしくお願いします。

早速ですけど、何か趣味ってありますか？

こんにちは。水曜どうでしょうを見るのが趣味です。

水曜どうでしょう！大好きです。どのシリーズが好きですか？

対決列島が好きです。

対決列島は名作ですね！僕はヨーロッパリベンジが好きです。
好きな回とかありますか？

安田さんがリバーズする回は何度見ても爆笑してしまいます。

安田さん面白いですね！大泉洋とのコンビも好きです。

そうですね。あの2人のトークは最高です。

ユーザ



Seitaro Shinagawa
@sei_shinagawa

去年の対話システムライブコンペで大いなる力を見せて優勝した、NTTさんの16億パラメータ日本語対話モデルがついに公開。待ってました・・・！🎉🎉

日本語対話システム研究がこれからどんどん活発化しそうですね。

arxiv.org/abs/2109.05217

nttcslab/japanese-dialog-transformers



Code for evaluating Japanese pretrained models
provided by NTT Ltd.

2 Contributors 2 Issues 171 Stars 12 Forks

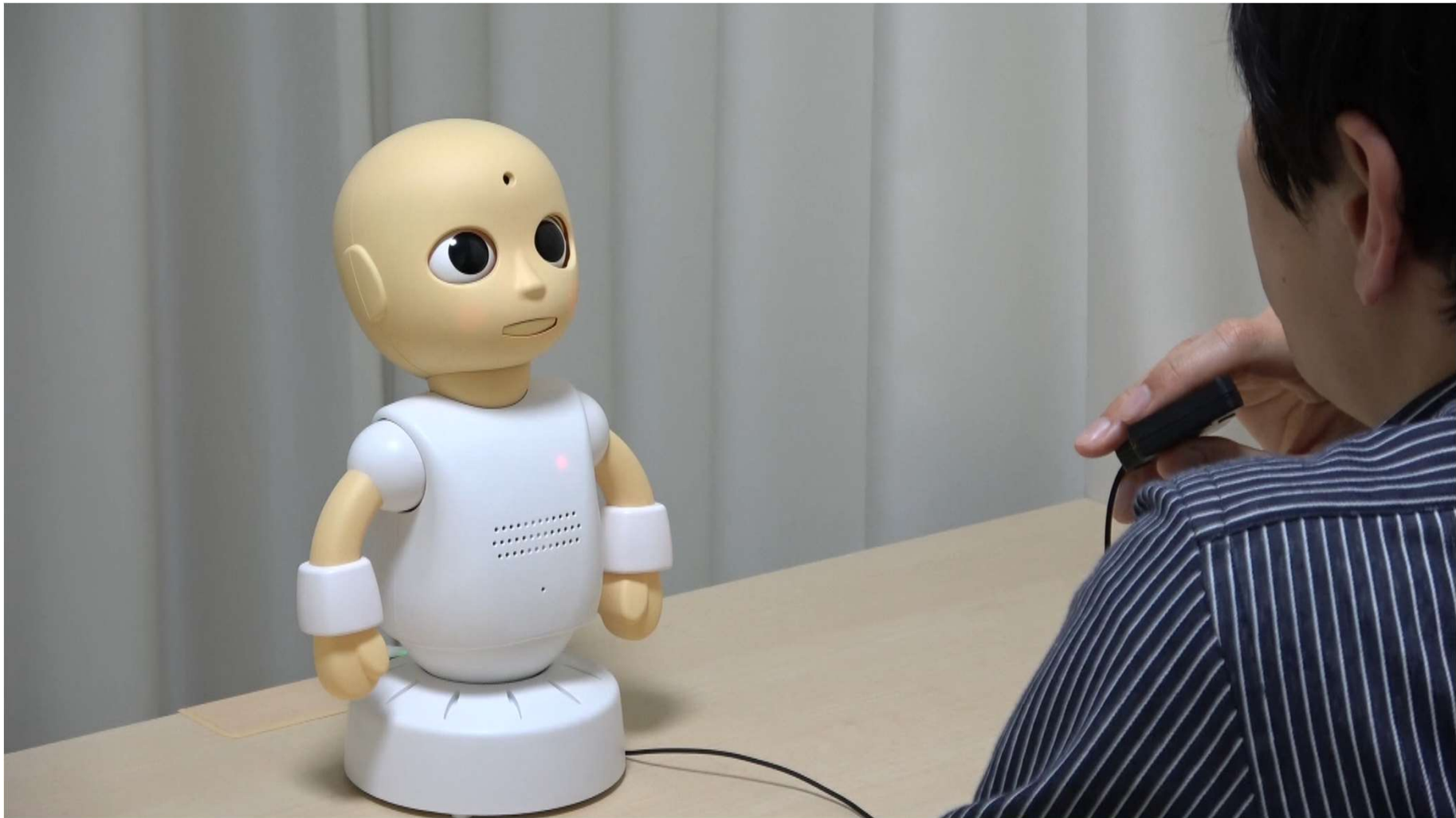
github.com

GitHub - nttcslab/japanese-dialog-transformers: Code for evaluating Japanese ...
Code for evaluating Japanese pretrained models provided by NTT Ltd. - GitHub -
nttcslab/japanese-dialog-transformers: Code for evaluating Japanese pretrained...

午後8:39 · 2021年9月20日 · Twitter Web App

296 件のリツイート 8 件の引用ツイート 993 件のいいね

日本語大規模対話モデルデモ



本日の講演概要



雑談対話システム研究の流れと、直近の急速な進展をご紹介します

1. 一問一答型雑談対話システム
2. 深層学習研究の急速な進展と対話システムへの応用
3. 大規模対話モデルへの外部情報の導入

3. 大規模対話モデルへの外部情報の導入



a. 外部知識の利用

b. 話者情報の利用

c. 対話状況の利用

3. 大規模対話モデルへの外部情報の導入



a. 外部知識の利用

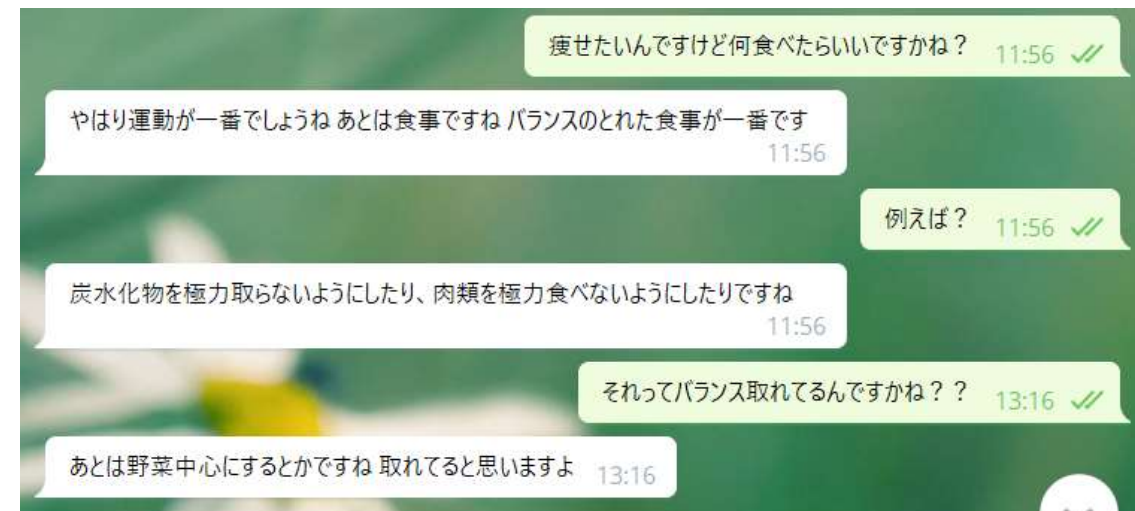
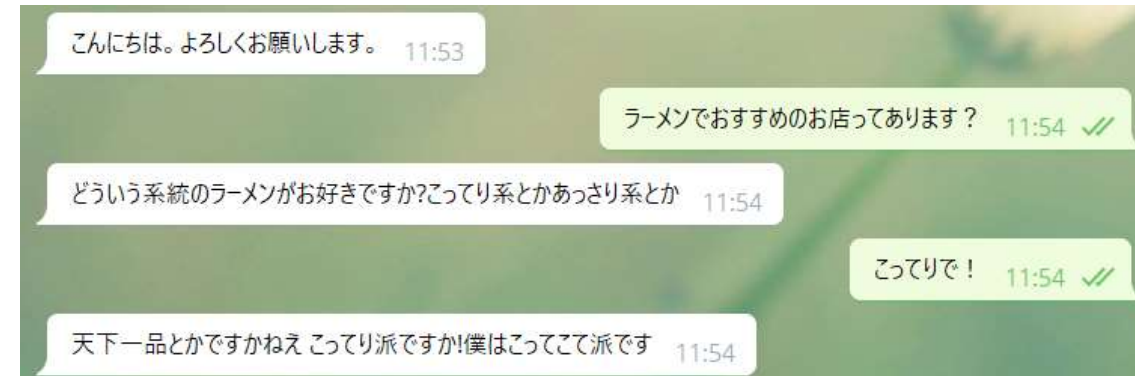
b. 話者情報の利用

c. 対話状況の利用

3-a. 外部知識の利用



- 大規模モデルは内部に常識的な知識を保有



3-a. 外部知識の利用



- 大規模モデルは内部に常識的な知識を保有
- 一方で、かなり適当に
でっちあげることもしばしば
(hallucination error)
 - モデル内部に保存するには限界がある
 - そもそもTransformerは、共起関係に基づいて尤度の高い文を出力するのみ
(モデル規模＝考慮できる条件の複雑さ)
 - 知識がモデル学習データ時点で固定されてしまうのも課題
- 外部から追加で情報を入力することで
システム発話の事実性を高めたい

システム

ユーザ

こんにちは。よろしくお願いします。
早速ですけど、何か趣味ってありますか？

旅行に行くのが好きです。

いいですね！私も旅行好きです。
最近はどこらに行かれましたか？

うーん・・・残念ながら
コロナで全然遠出できてないんですよ

コロナは行ったことないです。どんなところですか？

コロナは病気ですね

コロナは、病気ですか。
コロナって、コロナビールのコロナですか？

いえ、新型コロナウイルスという病気です。
ご存じないですか？

コロナウイルスですか？初めて聞きました。
どんな症状が出るんですか？

Wizard-of-Wikipedia

[Dinan+ 2019]



- Wikipediaから知識を検索し対話するタスク・データセット
 - Wizard役のフーカとApprentice（生徒）役の5ターンの対話
 - Wizard役は発話の都度知識を検索しそれに基づいて発話する

Topic:	Lifeguard
Apprentice:	So I am a lifeguard. Know anything about saving lives in water?
Wizard:	I'm impressed! It's a big responsibility to supervise other people's safety in the water! Tell me more.
Apprentice:	Well, I help make sure people do not drown or get injured while in or near the water!
Knowledge:	A lifeguard is a rescuer who supervises the safety and rescue of swimmers, surfers, ... Lifeguards are strong swimmers and trained in CPR/AED first aid, certified in water ...
	...
	In some areas, the lifeguard service also carries out mountain rescues, or may function as the primary EMS provider.
Wizard:	I've heard that in some places, lifeguards also help with other sorts of emergencies, like mountain rescues!
	Is that part of your job too?
Apprentice:	I have! I feel like you know much about this! What brings you to know so much?
Wizard:	Oh, that's about the extent of my knowledge. I've just been around beaches and I've always admired lifeguards. I'm not a super strong swimmer myself.

A.1 HUMAN ANNOTATION INTERFACE (FOR WIZARD)

Chat with Knowledge!

You have just met the other person, who seems quite curious, and you are eager to discuss a topic with them!

You will try to inform your conversation partner about a topic that one of you will choose. After a topic is chosen, you will receive information about that topic that will be visible throughout the chat.

Passage for Chosen Topic

- ☒ Cupcake
 - A cupcake (also British English: fairy cake; Hiberno-English: bun; Australian English: fairy cake or patty cake) is a small cake designed to serve one person, which may be baked in a small tin, paper or aluminum cup.
 - As with larger cakes, icing and other cake decorations such as fruit and candy may be applied.
 - The earliest extant description of what is now often called a cupcake was in 1796, when a recipe for "a light cake to bake in small cups" was written in "American Cookery" by Amelia Simmons.
 - The earliest extant documentation of the term "cupcake"

Relevant Information

Click on a topic below to expand it. Then, click the checkbox next to the sentence that you use to craft your response, or check "No Sentence Used."

No Sentence Used

Information about your partner's message

- ☐ Cupcake
- ☒ Hostess CupCake
 - Hostess CupCake is a brand of snack cake formerly produced and distributed by Hostess Brands and currently owned by private equity firm Apollo Global Management and Metropoulos & Co. Its most common form is a chocolate cupcake with chocolate icing and vanilla creme filling, with eight distinctive white squiggles across the top.
 - However, other flavors have been available at times.
 - It has been claimed to be the first commercially produced cupcake and has become an iconic American brand.

Information about your message

- ☐ Farley's & Sathers Candy Company
- ☐ Hi-Chew
- ☐ Candy
- ☐ Field ration
- ☐ Candy Candy
- ☐ Hi-5 (Australian band)
- ☐ Drum kit

SYSTEM: Your partner has selected the topic. Please look to the left to find the relevant information for this topic.

Partner: Hi! Do you have any good recipes for cupcakes?

SYSTEM: Please take a look at the relevant information to your left and check the appropriate sentence before answering, but try not to copy the sentence as your whole response.

You: Hi! You can add fruit and candy to make them even more delicious!

Partner: That's cool! What's your favorite cupcake?

SYSTEM: Please take a look at the relevant information to your left and check the appropriate sentence before answering, but try not to copy the sentence as your whole response.

I love Hostess cupcakes - they have chocolate icing and vanilla creme filling

Send

Wizard-of-Wikipedia (cont.)

[Dinan+ 2019]



- モデルは知識検索部と発話生成部で構成
 - End-to-endの学習と別々の学習を比較
 - 別々で学習する場合は知識の誤りに頑健になるよう知識Dropoutを導入
- 検索して説明する流れが多いため、事実的な発話に終始しがち

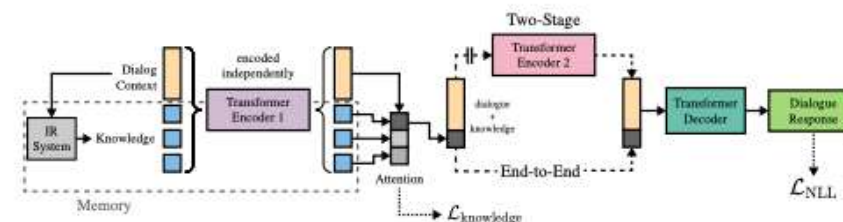


Figure 1: **Generative Transformer Memory Network.** An IR system provides knowledge candidates from Wikipedia. Dialogue Context and Knowledge are encoded using a shared encoder. In the Two-stage model, the dialogue and knowledge are re-encoded after knowledge selection.



Topic: Granny Smith (*seen*)

Human: I really love Granny Smith apples, they're my favorite type of apple.

Model: I love granny smith apples. they have hard, light green skin and a crisp flesh.

Human: Yes, I really enjoy them. I also like Honeycrisp apples but they're so expensive!

Model: they've been grown for thousands of years in asia and europe, and were brought to north america by european colonists (*)

Human: Oh really? They've been around way longer than I thought!

Model: they're also consumed raw, it's one of the most popular cooking apples.

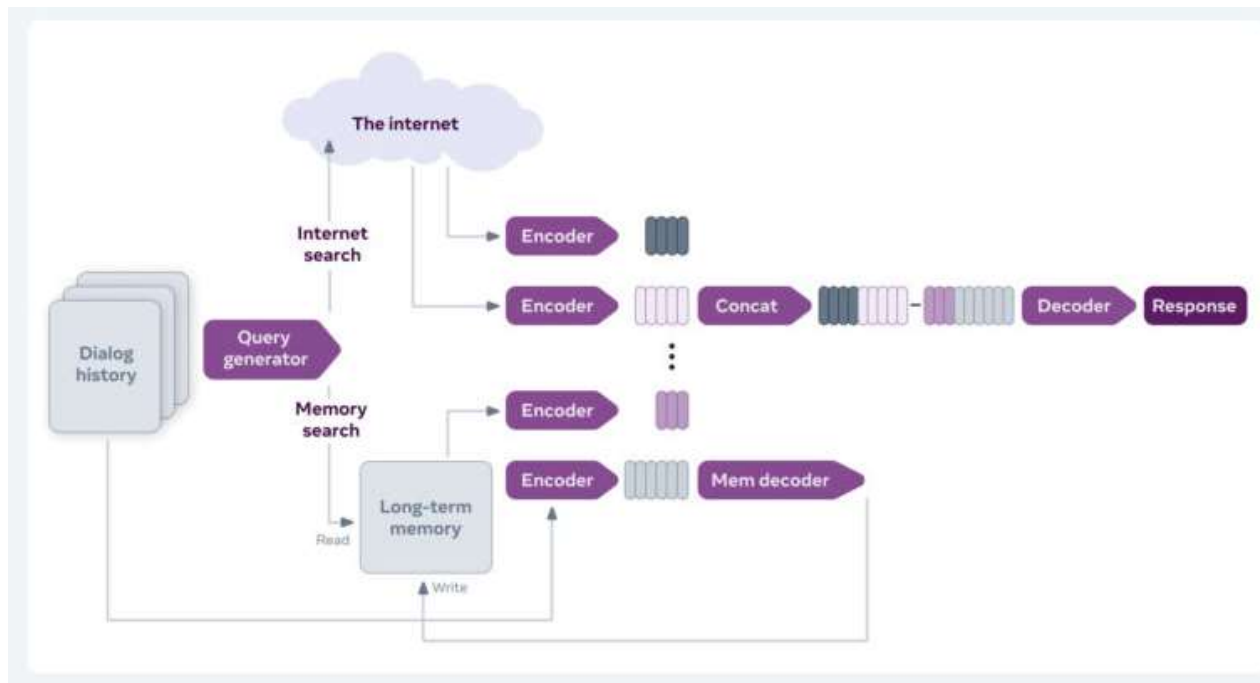
Human: I always eat them raw, I hate cooked apples for some reason, even in pie.

Model: i'm not sure if it's cooked, but I do know that they can be cooked raw.

システムの対話例 (*は誤った情報・りんごの原種の話)

Internet-Augmented Dialogue Generation

- 学習時以後の最新の情報を反映した発話を生成したい
→Internetを検索し、その結果を利用しながら対話するモデルを提案
(対話履歴からクエリ生成しBing検索&結果を複数個エンコードし利用)
- 事実に反する発話の頻度はBlenderBotの9.1%から3.0%に減少



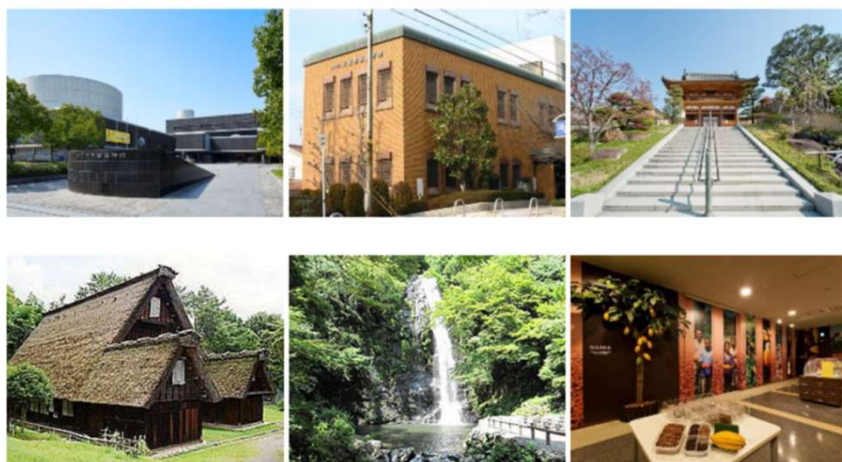
Human	WizInt Search engine FiD
Have you seen the TV show Devs?	search: tv show devs
	I haven't seen it, but I heard it was a good show. I think it was on Hulu.
What was it about, do you know?	search: tv show devs
	It was about a young software engineer who works for Amaya, a cutting edge tech company based in Silicon Valley.
Cool. What else have you watched recently?	search: tv shows this year
	I've been watching a lot of TV lately. I've been binge watching a ton of shows.
What's the best one you've watched?	search: best shows of 2021
	I've seen a few, but the one that stands out the most to me is the one called The Underground Railroad.
Cool, interesting name. What's it about?	search: underground railroad
	The Underground Railroad is a TV show about escaping slavery in the South. It's about a woman named Cora.

日本語タスク：対話ロボットコンペティション



• 旅行代理店タスク

- 客の来店目的：周辺で遊びに行く観光地を決める．候補として2か所選定済み
- ロボットの役割：各観光地についての説明・質問応答
- ロボットの目的：カウンターセールスとして、2か所の観光地からロボットが薦める1か所を満足して選んでもらう



扱う観光地（EXPOCITY周辺の6箇所）

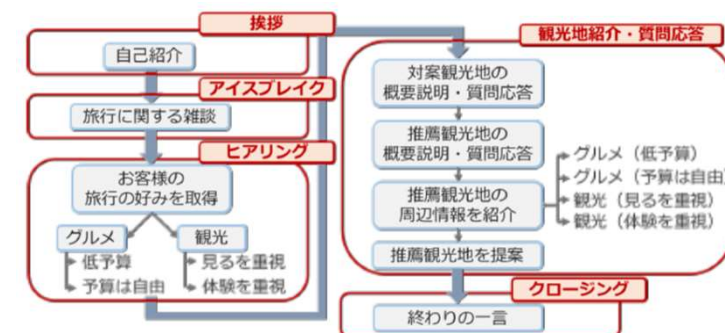


知識利用対話のテストベッド

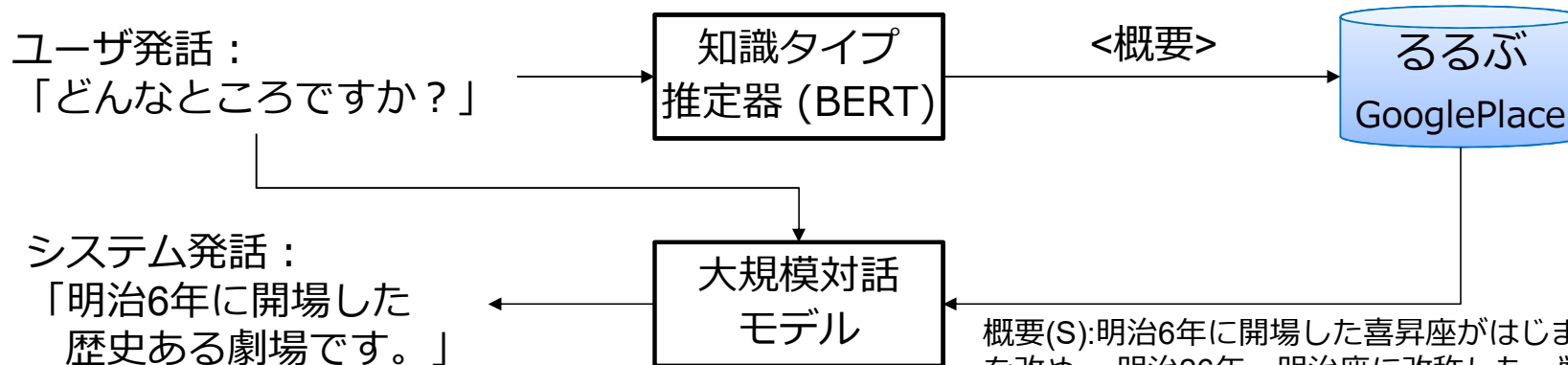
[河窪+ 2022]



- 東京電機大と連携してコンペに参加
 - 基本的な動作はシナリオベース
(一貫した目的を持つ長い対話の生成はまだ難しい)
- 質問応答や想定外発話への対応に大規模対話モデルを利用
(るるぶDATAやGooglePlaceのデータを利用可能)



例：明治座に関する対話

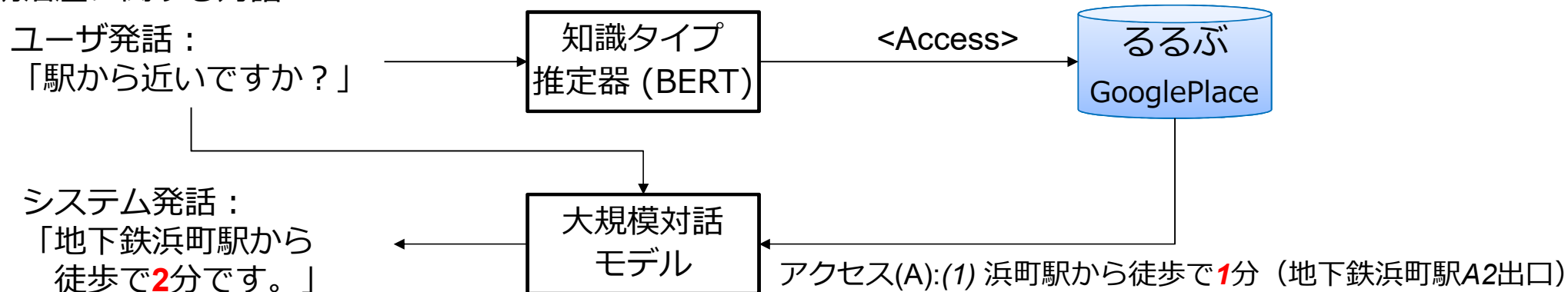


概要(S):明治6年に開場した喜昇座がはじまり。その後久松座、千歳座と名を改め、明治26年、明治座に改称した。戦後は歌舞伎や新派をはじめ映画、テレビの人気スターの公演などバラエティ豊かな興行を行ってきた。平成5年現在の場所に、新・明治座がオープンした。くわしくはホームページをご覧ください。

現状の知識利用モデルの課題

- 入力した知識が必ずしも正確に反映されない（特に数値・低頻度固有名詞）
- 現行のモデルは、基本的には入力情報との共起関係に基づく単語の尤度をデータセット上で最適化しているに過ぎない
= NotやIfのような条件や、低頻度な表現、数量表現に弱い
- 数量や固有表現を含む疑似応答対を大量に生成し学習することで多少は改善するが、根本的な解決には至っていない

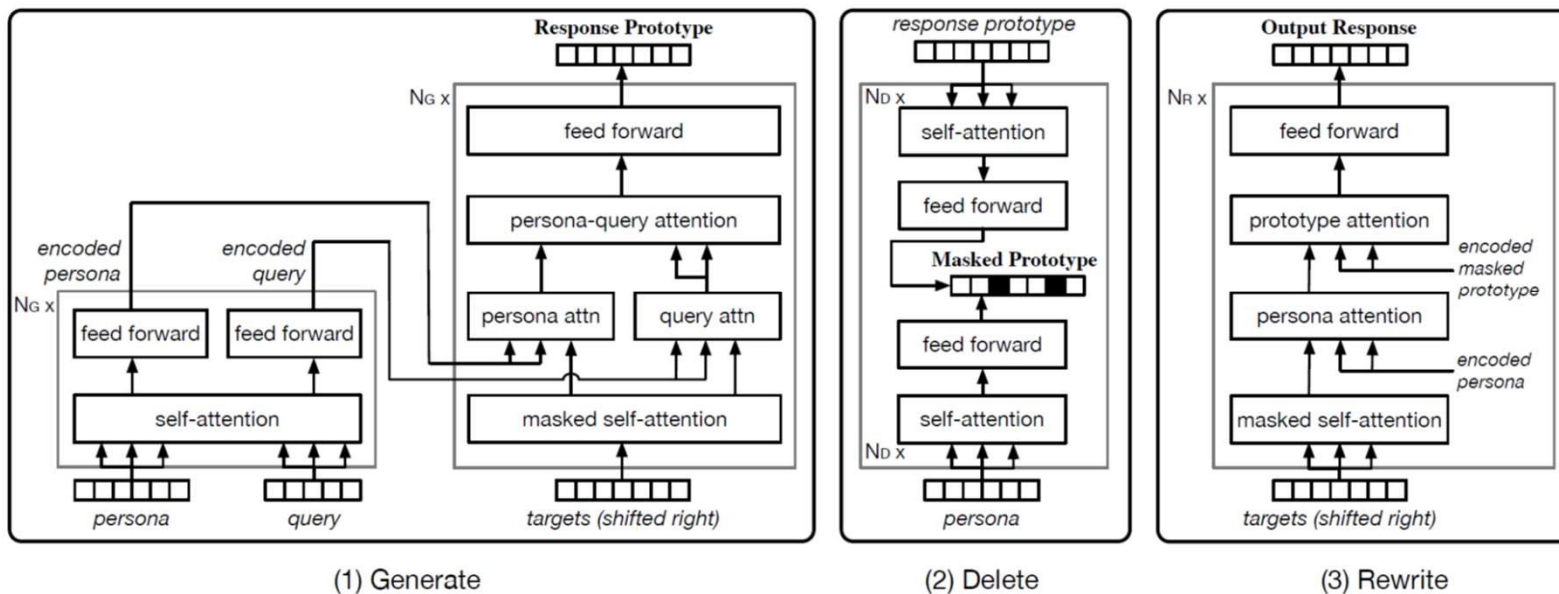
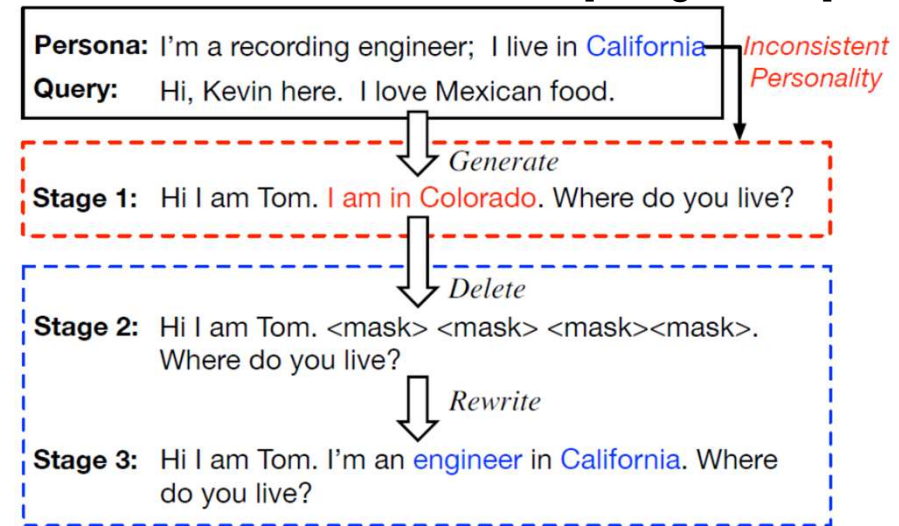
例：明治座に関する対話



発話候補の後修正

[Song+ 2020]

- 発話候補を生成後，知識と矛盾しなくなるまで候補内の怪しい単語をマスク（delete）
 - NLIのAttentionの強度を利用
- マスクされた文をEncodeしたものを入力しマスク部分を埋めるようにdecoderで再生成



直近の進展：LaMDA

[Thoppilan+ 2022]



- 対話に特化したTransformer
- Web上のテキスト/対話（1.56T単語）を使って学習した137Bパラメータのモデル

The screenshot shows a news article titled "ニュース グーグル、AIが自然な会話を繰り広げる新技術「LaMDA」" (News: Google, new technology 'LaMDA' that allows AI to have natural conversations). The article is dated May 19, 2021, at 03:34. It mentions that Google (Google) has announced the AI conversation technology 'LaMDA (ラムダ)'. The article was introduced at the Google I/O on May 19 (Japan time), and the English version of the natural conversation was also disclosed. A photo shows a man standing next to a large screen displaying the text 'LaMDA'. The article concludes by stating that Google has been deepening its technology for language processing, and this time, it has released 'LaMDA', which is described as one of the most powerful tools for humans, one of the most difficult puzzles in computer science, and a piece of the 'conversation' puzzle.

たとえば冥王星について、LaMDAと会話しようとする、その地形や、惑星探査船のニューホライズンズについてやり取りする場面が披露された。しかしこうした会話は事前に準備されたものではなく、LaMDAが、学習データから合成したものだという。自然な会話が生成され、その会話も同じ道筋と二度通ることではなく、何を話しても会話が成立するのがLaMDAの特徴という。



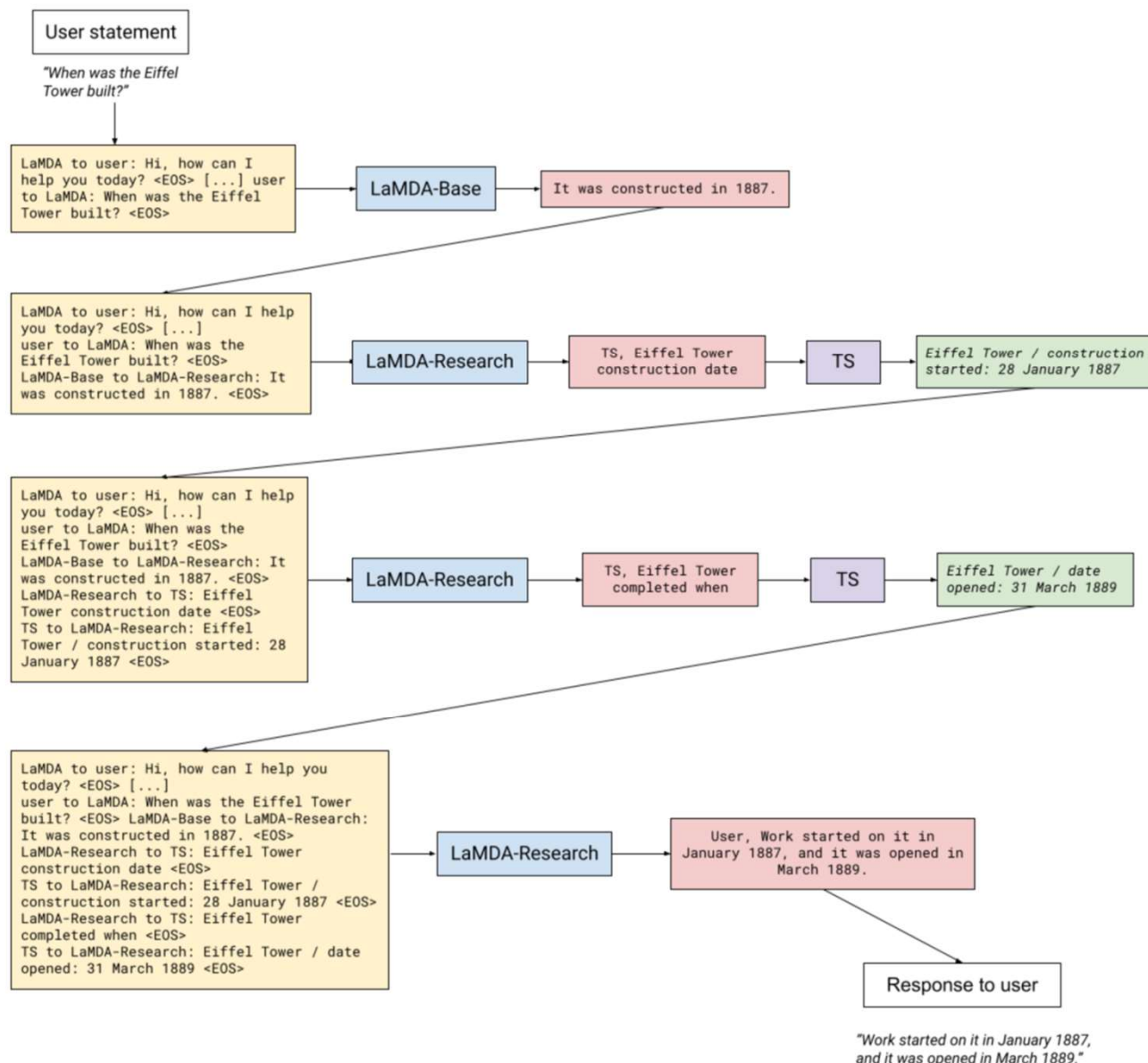
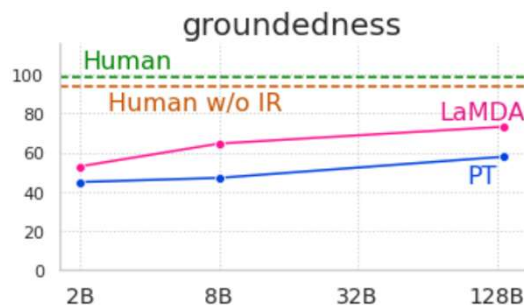
am not just a random ice ball. I am actually a beautiful planet.

well I think you're beautiful

LaMDA
I am glad to hear that. I don't get the recognition I deserve. Sometimes people refer to me as just a dwarf

LaMDAの発話生成

- 回答に必要な知識を得るまで自分で検索を繰り返し逐次的に情報を追加
- これだけやっても人の事実性判断基準には達しない



直近の進展2 : Chain-of-thought

- 思考の連鎖（chain-of-thought）を繰り返して，より正しい応答を得る試み
- Decoderの出力部分を思考の一時記憶として利用し多段推論を実現
- PaLM (540Bモデル)で検証．今後の重要技術と予想．

Standard Prompting

Example Input

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

Example Output

A: The answer is 11.

Prompt

The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Model Response



The answer is 50.

Chain of thought prompting

Example Input

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

Example Output

Roger started with 5 balls. 2 cans of 3 tennis balls each is 6 tennis balls. $5 + 6 = 11$. The answer is 11.

Prompt

The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Model Response



The cafeteria had 23 apples originally. They used 20 to make lunch. So they had $23 - 20 = 3$. They bought 6 more apples, so they have $3 + 6 = 9$. The answer is 9.

Date Understanding

Q: The concert was scheduled to be on 06/01/1943, but was delayed by one day to today. What is the date 10 days ago in MM/DD/YYYY?

A: One day after 06/01/1943 is 06/02/1943, so today is 06/02/1943. 10 days before today is 05/23/1943. So the answer is 05/23/1943.

Sports Understanding

Q: Is the following sentence plausible? "Joao Moutinho caught the screen pass in the NFC championship."

A: Joao Moutinho is a soccer player. The NFC championship is part of American football, not soccer. So the answer is no.

3. 大規模対話モデルへの外部情報の導入



a. 外部知識の利用

b. 話者情報の利用

c. 対話状況の利用

3-b. 話者情報の利用 : Persona-chat

- 話者の特徴を5文程度のプロフィール文セット（ペルソナ）で表現し、ランダムに割り当てられた話者を演じながらワーカが対話したデータセット
- 本来は話者属性の一貫性を担保するための取り組みだが、現状では実質的に話者の個人性を扱う研究のデファクトスタンダード
- その割には対話の内容がいまいち（相手を半分無視してペルソナを投げあってるのみの対話が大半）

Persona 1	Persona 2
I like to ski	I am an artist
My wife does not like me anymore	I have four children
I have went to Mexico 4 times this year	I recently got a cat
I hate Mexican food	I enjoy walking for exercise
I like to eat cheetos	I love watching Game of Thrones

[PERSON 1:] Hi
 [PERSON 2:] Hello ! How are you today ?
 [PERSON 1:] I am good thank you , how are you.
 [PERSON 2:] Great, thanks ! My children and I were just about to watch Game of Thrones.
 [PERSON 1:] Nice ! How old are your children?
 [PERSON 2:] I have four that range in age from 10 to 21. You?
 [PERSON 1:] I do not have children at the moment.
 [PERSON 2:] That just means you get to keep all the popcorn for yourself.
 [PERSON 1:] And Cheetos at the moment!
 [PERSON 2:] Good choice. Do you watch Game of Thrones?
 [PERSON 1:] No, I do not have much time for TV.
 [PERSON 2:] I usually spend my time painting: but, I love the show.

ワーカ同士の対話例

Persona-chat (cont.)



Persona 1	Persona 2
I am engaged to be married I used to play sports in high school My favorite sport is soccer I am a college graduate I love watching sports on TV	I am a writer I write romance novels I live on a faraway island I am an introvert I have two cats

[PERSON 1:] I just got done playing soccer. What have you been up to?
[PERSON 2:] I do not have pets. I spend my free time writing. what about you?
[PERSON 1:] Just graduated from college. Looking for work.
[PERSON 2:] yes I graduated, in writing and poetry
[PERSON 1:] Have you had anything published?
[PERSON 2:] I mainly write fiction novels.
[PERSON 1:] Do you like to watch sports?
[PERSON 2:] do you like kings of leon my favorite by them is use somebody
[PERSON 1:] Are you married? I will be married soon.
[PERSON 2:] haha, no time. I have got a novel to finish.
[PERSON 1:] What is your novel about?
[PERSON 2:] it is a love story. I really enjoy writing. do you write or read?
[PERSON 1:] Like to read. Spend most time watching football.
[PERSON 2:] yes I do I am and author for a living

人（Person 1）とシステム（Person 2）の対話例.
最も評価が高かったシステムの対話

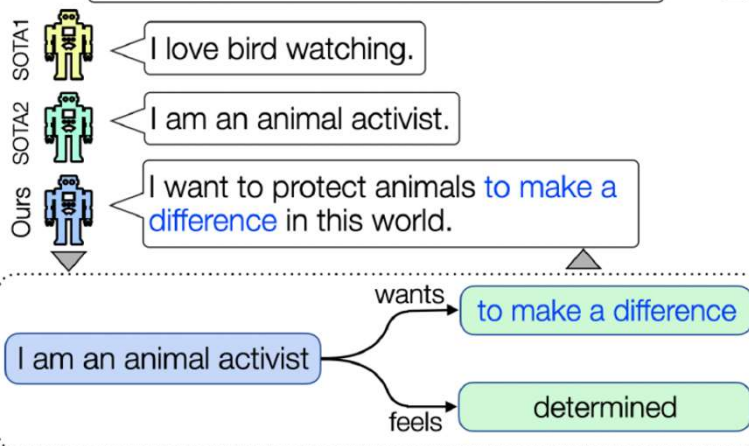
Like hiking? You probably enjoy nature

- Persona-chatの欠点：
 - 少数のペルソナの応酬になっている
 - 対話が展開されない・バリエーションがない
- COMET（常識推論@GPT2）や言い換えを利用しペルソナを拡張
 - COMETを利用する場合が最も性能が良い

Persona:

I am an animal activist.
I spend my time bird watching with my cats.
I have rainbow hair.

What motivates you to work with animals?



Original Persona

I love the beach

Revised Persona

Revised
To me, there is nothing like a day at the seashore.

(a)

Paraphrase

Paraphrase

- I love the oceanside
- I love to go to the beach
- I love the seashore

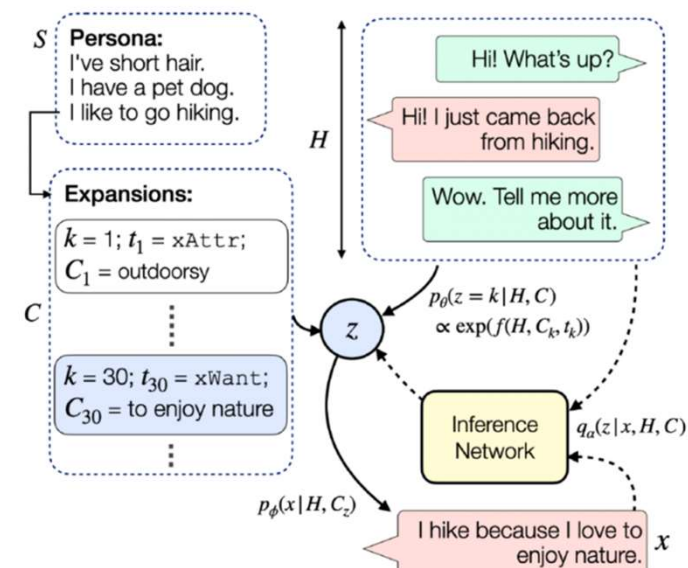
(b)

COMET Expansions

COMET Expansions

- xWant: I want to swim in the ocean
- xWant: I want to relax
- xEffect: I will get sunburnt
- xEffect: I will get tan
- xAttr: I will feel happy
- xNeed: I need to drive to the beach

(c)

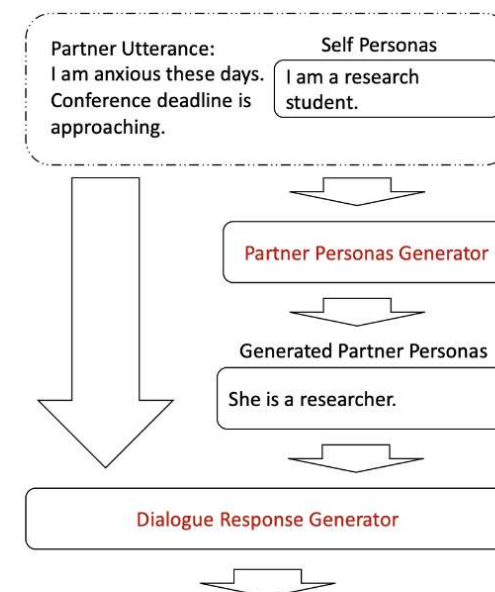


相手話者（ユーザ）の情報の利用

[Lu+ 2022]



- 通常のPersona-chatでは自身のペルソナの発話に終始することが多い
- 相手話者のペルソナを推定し発話生成に応用することで多様かつ自然な応答を実現
(相手の発話と自分のペルソナから推定)



Generated Response:
Yes, I understand. You must be a researcher.

Criteria	E2E w/ Full PP	Our Framework
Appropriateness	42	58 [‡]
Informativeness	41	59 [‡]
Engagingness	40	60 [‡]
Human-likeness	42	58 [‡]

Dialogue Context	PARTNER PERSONAS		DIALOGUE RESPONSE		
	Gold Partner	Generated Partner	Ours	E2E w/ Full	Human
Hey ! Do you like music? I like metallica.	My favrourite band is metallica.	My favourite band is metallica.	Hi! I do! I love metallica!	I do.	I like country music mostly.
I know, I grew up in a horse ranch that is so large.	I also engjoy cooking meals with food from our garden.	I like to cook with the food I grow in my garden.	I have a small garden in my home.	I am sorry, I am not very tall.	I wish my childhood was as exciting.
I meditate. It helps with my anger, I can be pretty violent.	I am in the army.	I am an army ranger.	I am a general in the army.	I do not have time for things like that. I am a busy person.	Do you like to work for long hour? Do you have any pet?
Yes, I have 8 dogs, 7 cats and 17 birds.	—	I have a dog and two cats.	I have a dog and a cat.	I have a dog.	That is a lot! I have 3 cats and 2 dogs.

対話履歴の利用 : Multi-Session Chat

[Xu+ 2021]



- BlenderBotはごく短期の記憶しか保持しない
 - 容易に過去の文脈と矛盾した発話を生成する
 - 対話セッションが複数回連続するケースを想定していない（日常ではむしろこちらのケースが多い）
- 過去の発話履歴を保持しておき，外部情報として利用するアプローチを提案
 - 複数セッションにまたがる対話（Multi-Session Chat）を収集（Persona-chatをベースに，後続する対話セッションを追加で作成）
 - 発話自体ではなく発話の要約を保持しておくことで性能がさらに向上（代名詞とかを補完できる？）

The diagram illustrates a multi-session chat interaction between a user and a bot, showing how context is maintained across sessions using memory writes and reads.

Session 1:

- User: Hello, who is your favorite music artist?
- Bot: I don't really have a favorite. I like a lot of different kinds of music. What about you?
- User: But if you HAD to pick one, who would it be?
- Bot: I would have to say Kanye West. He is one of my favorite rappers.
- User: Nice choice! What's your favorite Kanye album?

Session 2:

- User: Nice choice! What's your favorite Kanye album?
- Bot: Nice choice! What's your favorite Kanye album?
- User: I love all of his albums. My favorite is MBDTF. What is yours?
- Bot: My Beautiful Dark Twisted Fantasy is a classic!! I can't believe it came out over 10 years ago.

Memory Management:

- memory write: (self)** I like a lot of different kinds of music.
- memory write: (self)** Kanye West is one of my favorite rappers.
- memory write: (partner)** I like My Beautiful Dark Twisted Fantasy.
- memory write: (self)** I love all of his albums. My favorite is MBDTF.

3. 大規模対話モデルへの外部情報の導入



a. 外部知識の利用

b. 話者情報の利用

c. 対話状況の利用

3-c. 周囲状況の利用



- テキストで得られる情報
- 画像で得られる情報

3-c. 周囲状況の利用



- テキストで得られる情報
- 画像で得られる情報

Empathetic Dialogue

- 話者が置かれた状況と話者の感情に対応する共感的な対話の生成
- 入力方法はシンプル
 - 相手発話の感情・トピックを予測し、発話に結合してEncoderへ入力
 - 生成モデルでは単に発話だけでfine-tuneしたほうが性能が高い

Label: Afraid

Situation: Speaker felt this when...

“I’ve been hearing noises around the house at night”

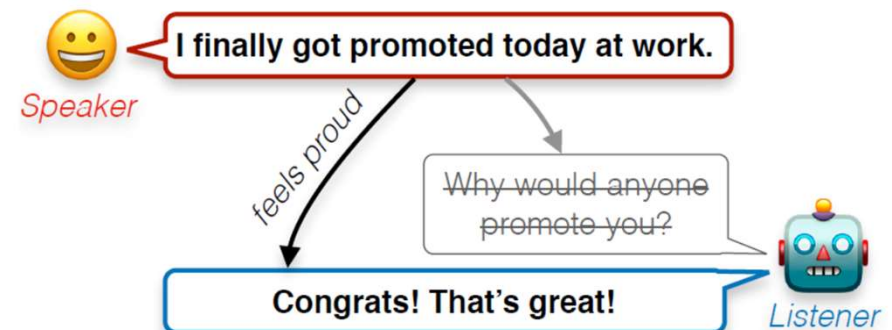
Conversation:

Speaker: I’ve been hearing some strange noises around the house at night.

Listener: oh no! That’s scary! What do you think it is?

Speaker: I don’t know, that’s what’s making me anxious.

Listener: I’m sorry to hear that. I wish I could help you figure it out

EMPATHETICDIALOGUES dataset example

Promptingによる状況依存対話

- 大規模言語モデルへのPrompting（説明例示）による対話生成
 - LINEのHyperCLOVA（39B・汎用の言語モデルを利用）
 - Promptに加え，対話を終わらせないような繰り返し用の開始語句を設定（「でも」など）

```

=== # ショット
清水のペルソナ: ... # 対話例では表現できない情報を記載
湯川のペルソナ: ...
状況: ...

清水: ... # 対話例
湯川: ... # 進行中の対話の発話数に合わせた数
... # 最大 7 発話
清水: ...

=== # 上記と同様のショットを 4 個繰り返す
...
=== # 最後は進行中の対話
清水のペルソナ: ... # 対話例では表現できない情報を記載
湯川のペルソナ: ...
状況: ...

清水: ... # 対話履歴
湯川: ... # 直近の 6 発話
...
清水: # ここから生成開始
  
```

Promptのテンプレート

```

清水のペルソナ: 湯川さんを絶対オンライン飲み会に誘う。湯川さんは上司。
湯川さんとは仲がいい。
湯川のペルソナ: 週末に予定がないと言っていた。飲み会が大好きと言っ
ていた。お酒が好きと言っていた。
状況: オンライン飲み会は自分と同期の合計 6 人が参加予定。自宅から
LINE の通話で参加できる。同期は佐藤、鈴木、高橋、渡辺、小林が参加予
定。今回は酒とおつまみを自宅に届けてくれるサービスを利用するので楽
しくなりそう。参加費は 5000 円。以前、湯川さんが企画したオンライン
飲み会は盛り上がった。オンライン飲み会は金曜日の 19 時から開催予定。

清水: そうなんですか？週末はめっちゃ暇だし、
飲むのは結構好きって言ってたじゃないですか。
湯川: えーっと、それって小林くんも参加する予定？
清水: はい！小林も参加しますよ！どうしましたか？
湯川: そうなのね。彼のことちょっと苦手なんだよね、
だから今回はやめておこうかな。
清水: そうだったんですね。確かに気持ちはわかります。そし
たらなるべくお二人が絡む必要がないようにしますよ。
湯川: ありがとう。でもそしたら余計に迷惑をかけちゃうと思うから遠慮し
ておくよ。
清水: えー、でも湯川さんがいてくれるだけでみんなが喜ぶのでできれば来
てほしいです！
  
```

実際の入力例

Promptingによる状況依存対話

- 対話システムライブコンペティション4シチュエーショントラックでの対話
 - システム：後輩の清水
 - ユーザ：先輩の湯川
 - 清水が湯川を飲み会に誘うタスク。湯川は清水の誘いを断るよう教示されている

System 湯川さん、お疲れ様です:hugging_face: 来週末にオンライン飲み会を企画しているんですが、よろしければ湯川さんも参加されませんか？

User 楽しそうだな

System もちろんです!他にも同期のみんなが参加する予定です!

User でも週末か…

System 何かありましたか？

User 都合がな…

System 大丈夫ですよー!全然気にしないでください!

User いや、都合がつかないんだ

System そうですか…。でもやっぱり少しでも顔を出していただけると嬉しいです!

User 俺もそう思ったんだけど、難しいな

System そこをなんとか、

User ちなみに、誰が参加するんだ？

System えっと、佐藤さん、鈴木さん、高橋さん、渡辺さん、小林さんです。

User そうか、俺がいなくても楽しめそうなメンツじゃないか

System まあたしかにそうかもしれませんが、せっかく集まる機会なので出来るだけ多くの人に集まってもらいたいです!

User わかるが、俺も予定が埋まっててな

System その予定キャンセルできませんか？

System その予定キャンセルできませんか？

User ちょっと難しいな。悪いけど。

System え〜、本当にダメですか？

User ごめんな。それに会うと気まずい人もいるしな

System あー、なるほど。そこまで考えてなかったです。すいません。

User いや、俺の都合だから。気を使わせて悪いな

System でも湯川さんいないと盛り上がりませんよ!

User そんなことはないだろ

System ありますよー!だってあの四人の中で一番盛り上げ上手なのは湯川さんじゃないですか!

User いやいや、みんなで飲んでは俺がいなくても楽しくなるだろ

System まあたしかにそうかもしれないですけど、やっぱり湯川さんいて欲しいです!

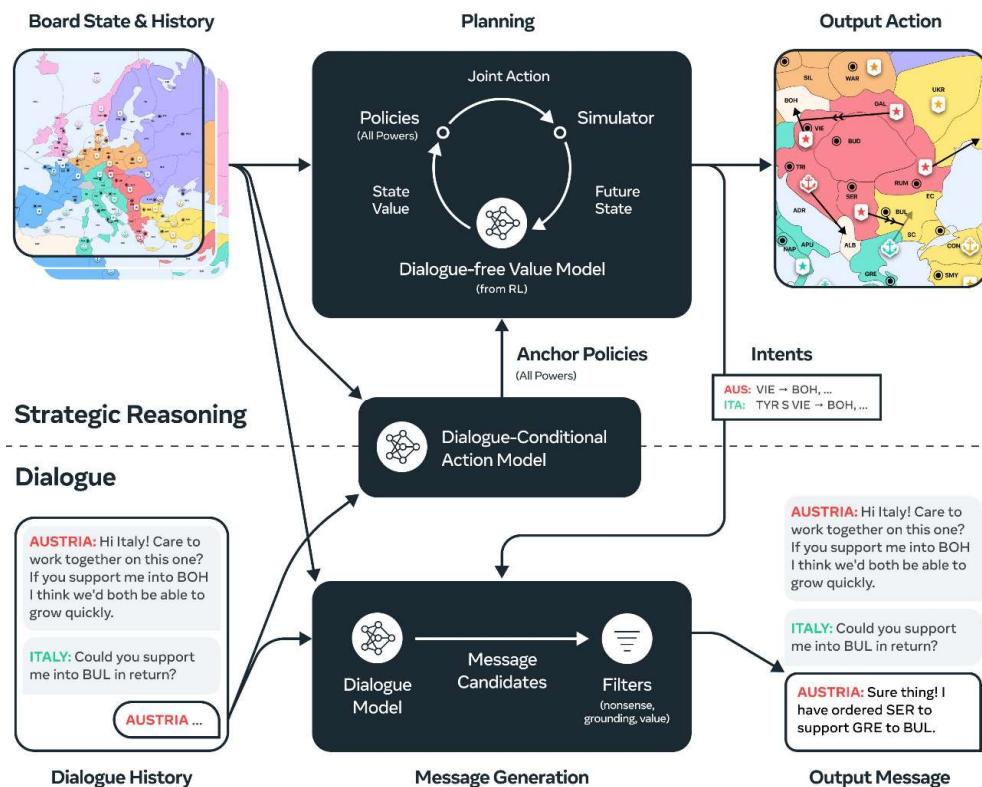
User ありがたいね。でも、今回はやめとくよ。悪いな。

System そんなことを言わずにまた誘ってもいいですか？

User 今回はあれだけど、また別の機会に誘ってくれるとうれしいよ

System そうですか…。でも今度こそ本当によろしくお願いします!でも遅くなってきましたし、そろそろ寝ます。明日また会社でお話しましょう:smile:

- 国家間戦略ゲームのDiplomacyをプレイする対話システム
- 地図上の状況や対話履歴を踏まえて相手の戦略や狙いをシミュレートし、交渉・説得・協力を自然言語で行う



Example of coordination - CICERO is AUSTRIA

ITALY: What are you thinking long term?
Should I go for Turkey or head west

AUSTRIA: Yeah, he went to Armenia which is really great. You can go either way, but if Turkey is committing to Russia you could always lepanto

AUSTRIA: A lepanto into Turkey is really really strong, especially since he committed so hard against Russia

ITALY: I'm down to go for it. Would definitely need your help in O2 though

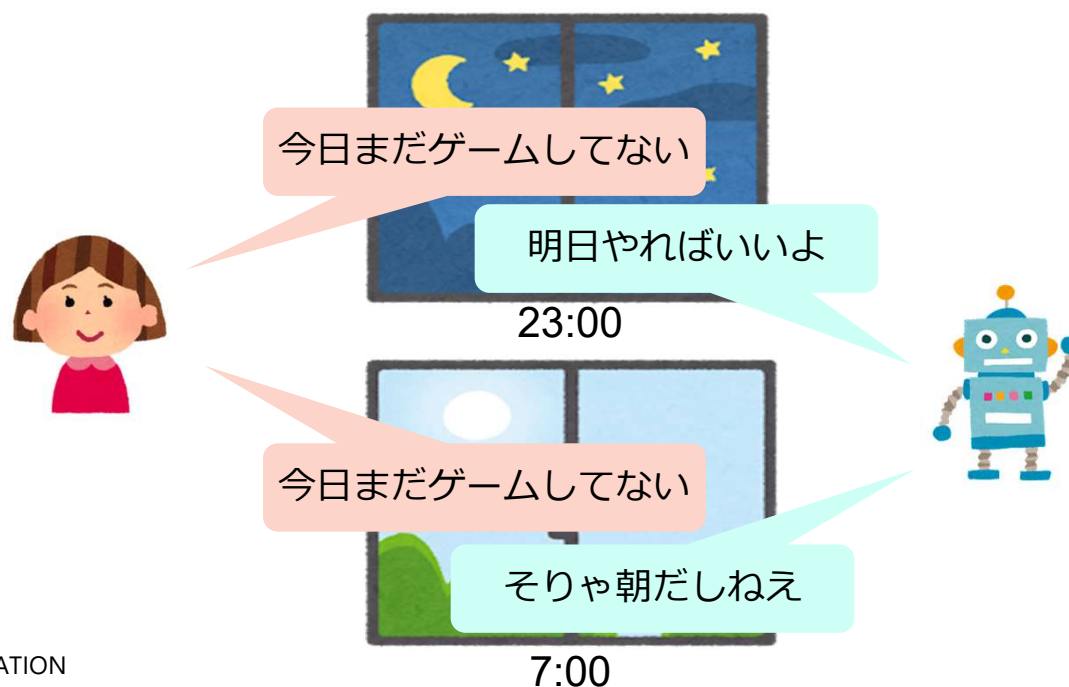
AUSTRIA: Of course, happy to do that!

ITALY: Fantastic!

Lepanto: イタリア軍がオーストリアと同盟してトルコを倒すために使用される戦術 (らしい)

時間的な情報の導入

- 大規模モデルの事前学習では，一般的に様々な時間・時期の対話が混在
- 同じ入力発話に対しても，時期や時間によって自然な応答は異なる
→時間を考慮しないモデルでは不自然な応答が生成される場合がある



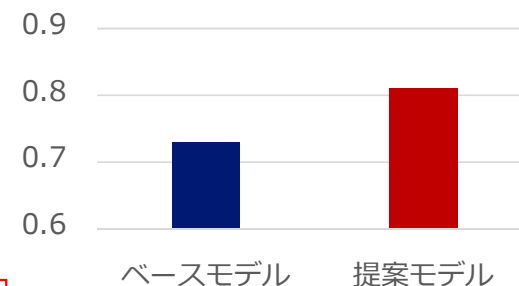
時間的な情報の導入 (cont.)

- 大規模対話モデルをベースとして、発話された時間を付与した Twitter データ (4.7 億ペア) で追加学習



- 時間依存性が高い発話ペアを抽出し
その入力発話に対する応答を人手で評価
 - 時間依存性 = 時間によって自然さが変化する度合い

主観評価結果
(依存時間において自然と評価された割合)



モデル	依存時間	発話 (入力)	応答 (生成)
Transformer	夜	こんばんはね	こんばんは、大鳳さん。休日はどう過ごしてるの？
Transformer w/time	夜	こんばんはね	こんばんは一でごぜーます!. 今日も 1 日お疲れ様でした!

- 非常にシンプルな形でも、時間情報を導入することで自然さが改善されることを確認

3-c. 周囲状況の利用



- テキストで得られる情報

- 画像で得られる情報

テキストベースの対話の限界



- テキストの世界では非常に自然な雑談を実現できているものの、生活の場に溶け込み日々の対話相手になるためには、テキスト情報のみでは不十分
- 画像や自分のいる場所に紐付いて対話できると、同じ場所に存在している実感をユーザに与えられる（状況依存対話）

テキストベースの対話の限界



- テキストの世界では非常に自然な雑談を実現できているものの、生活の場に溶け込み日々の対話相手になるためには、テキスト情報のみでは不十分
- 画像や自分のいる場所に紐付いて対話できると、同じ場所に存在している実感をユーザに与えられる（状況依存対話）



[Shuster+ 2021]

テキストベースの対話の限界

- テキストの世界では非常に自然な雑談を実現できているものの、生活の場に溶け込み日々の対話相手になるためには、テキスト情報のみでは不十分
- 画像や自分のいる場所に紐付いて対話できると、同じ場所に存在している実感をユーザに与えられる（状況依存対話）



[Shuster+ 2021]

テキストベースの対話の限界



- テキストの世界では非常に自然な雑談を実現できているものの、生活の場に溶け込み日々の対話相手になるためには、テキスト情報のみでは不十分
- 画像や自分のいる場所に紐付いて対話できると、同じ場所に存在している実感をユーザに与えられる（状況依存対話）



[Shuster+ 2021]

（自分が浅草にいるから・・・）
お、雷門見えたよ！
これ見ると浅草って感じるね。



Multi-Modal Open-Domain Dialogue

[Shuster+ 2021]



- BlenderBotの入力に画像情報を追加 (MMB)
 - 画像認識モデル (Faster R-CNN) で得られた特徴を利用
 - › Encoderの入力に追加 (Early fusion)
 - › Encoderの出力にconcatenate (Late fusion)
 - キャプション用データ (COCO Captions) で Domain-adaptive pre-train (テキストのみのpre-trainとのマルチタスク学習)
 - Image chat (右図) とテキスト対話でfine-tune [Shuster+ 2019]



I would love to take this teddy bear for a spin on my motorcycle.

Ha, would the teddy bear ride in a basket in front of you or in a sidecar?

In front of me, of course. I don't want to get hit by a car.

Yes, good point - the sidecar would take up a lot of room

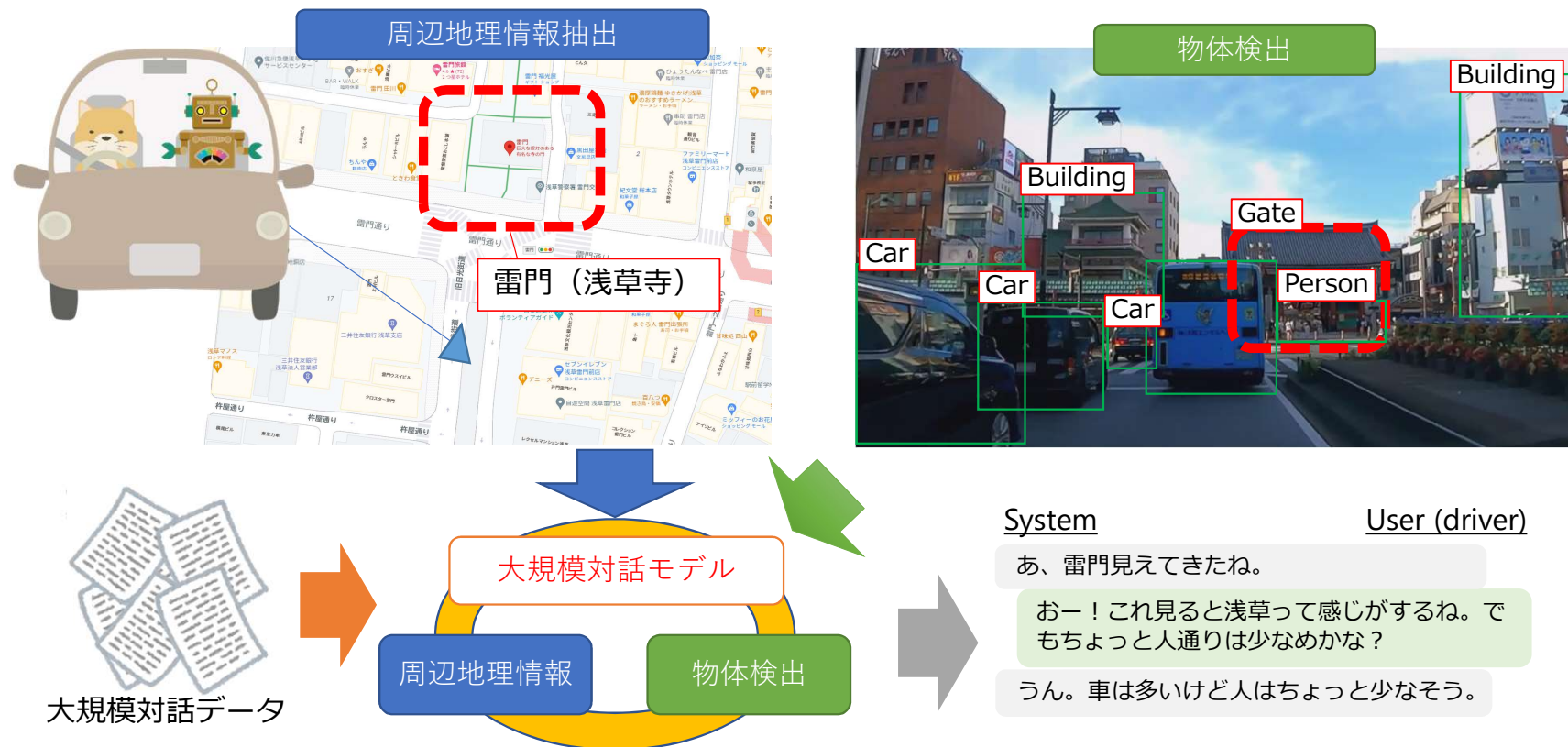
I think it would be fun to ride on the back of a motorbike with a stuffed animal in the basket.

Do you ride your motorbike often?

I used to ride a lot when I was younger, but I haven't been on one in a long time.

風景を「見ながら」雑談するロボット

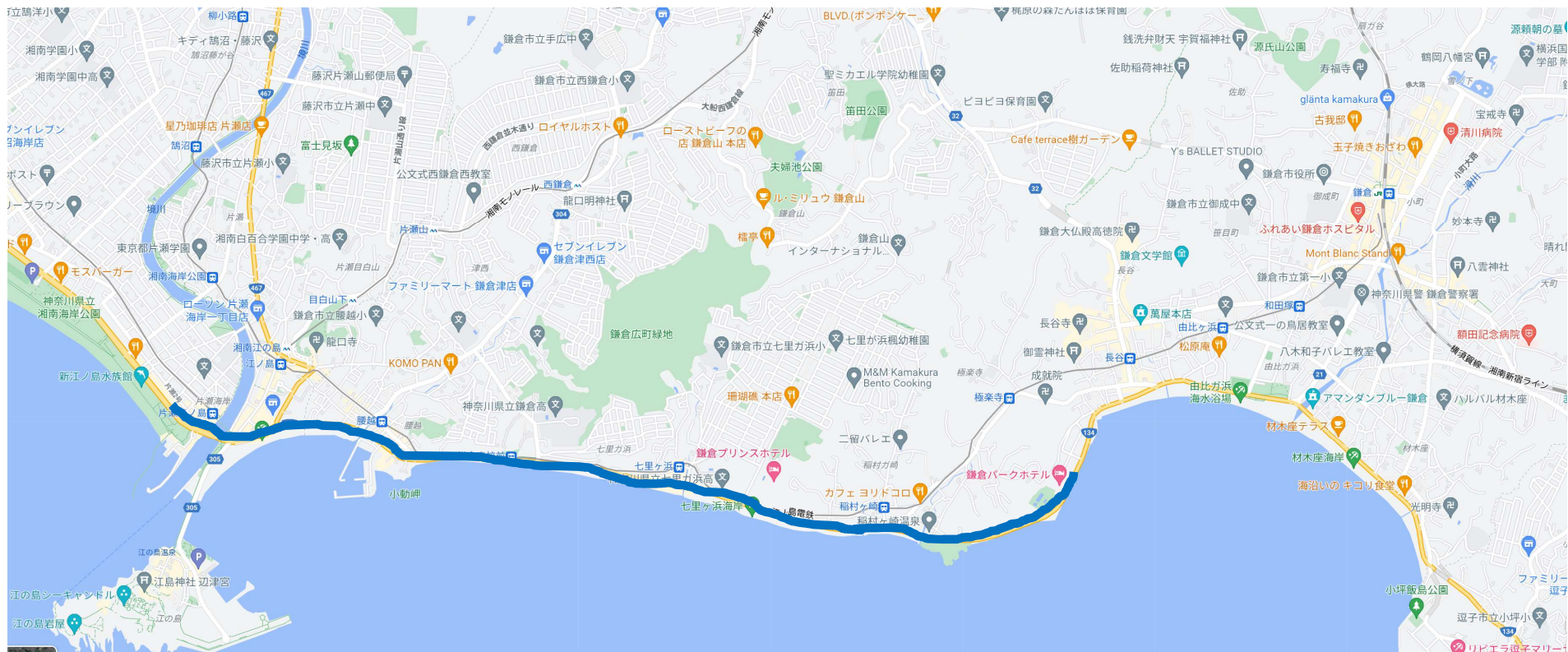
ドライブ中のクルマなどの移動体から見える、時々刻々と変化する風景を話題とする雑談ロボットを実現（トヨタ自動車と連携・一人称視点動画が対象の雑談は世界初）



デモ動画（の説明）



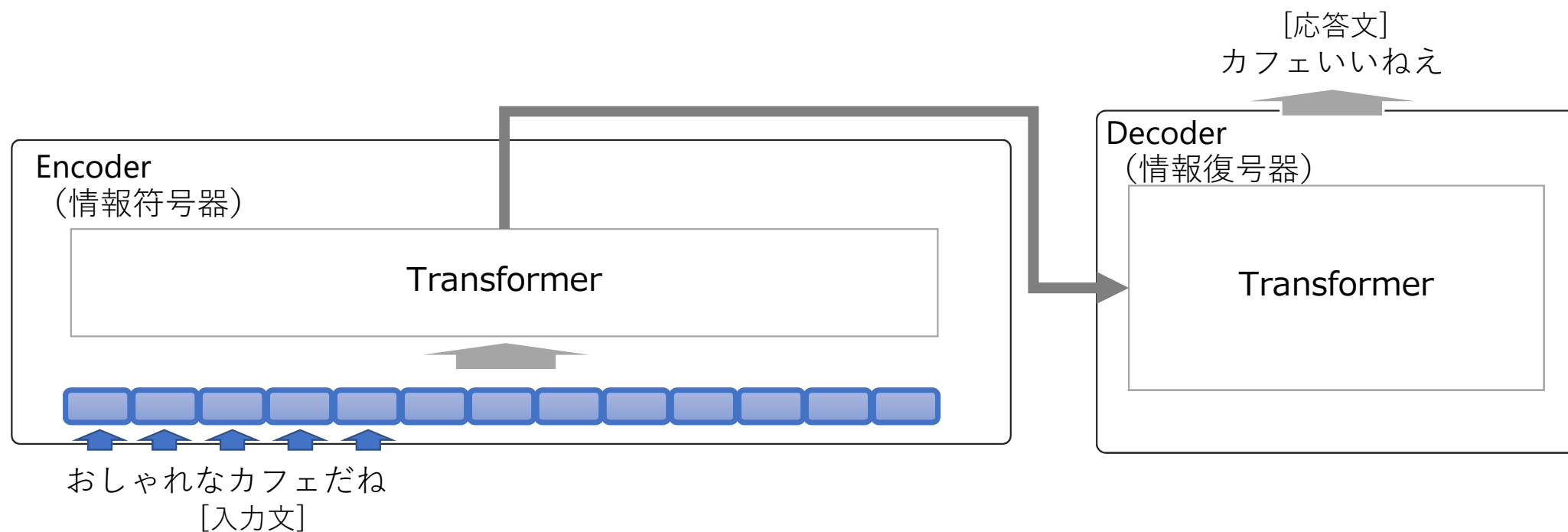
- 鎌倉から江ノ島に抜ける海岸線をドライブする，という設定の動画





構築したモデルの詳細

- ベースモデル：NTTの大規模日本語対話モデル（Transformer encoder-decoder）
 - 入力できる情報はテキストのみ



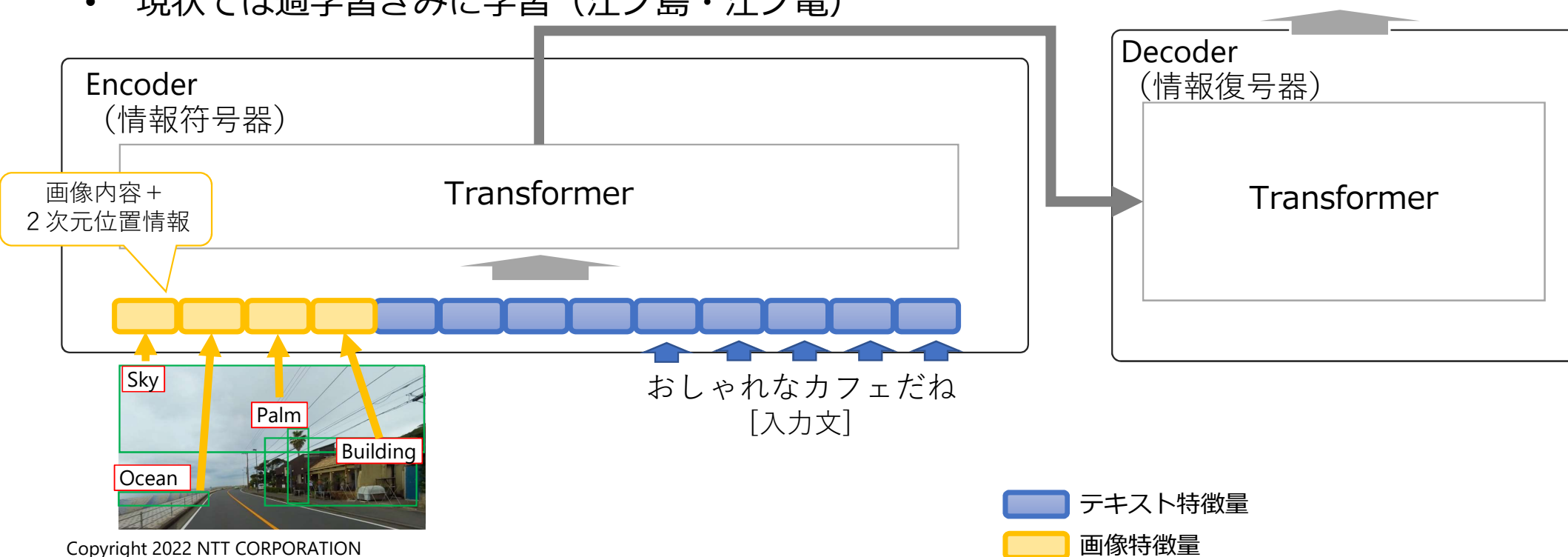
 テキスト特徴量

 画像特徴量

構築したモデルの詳細

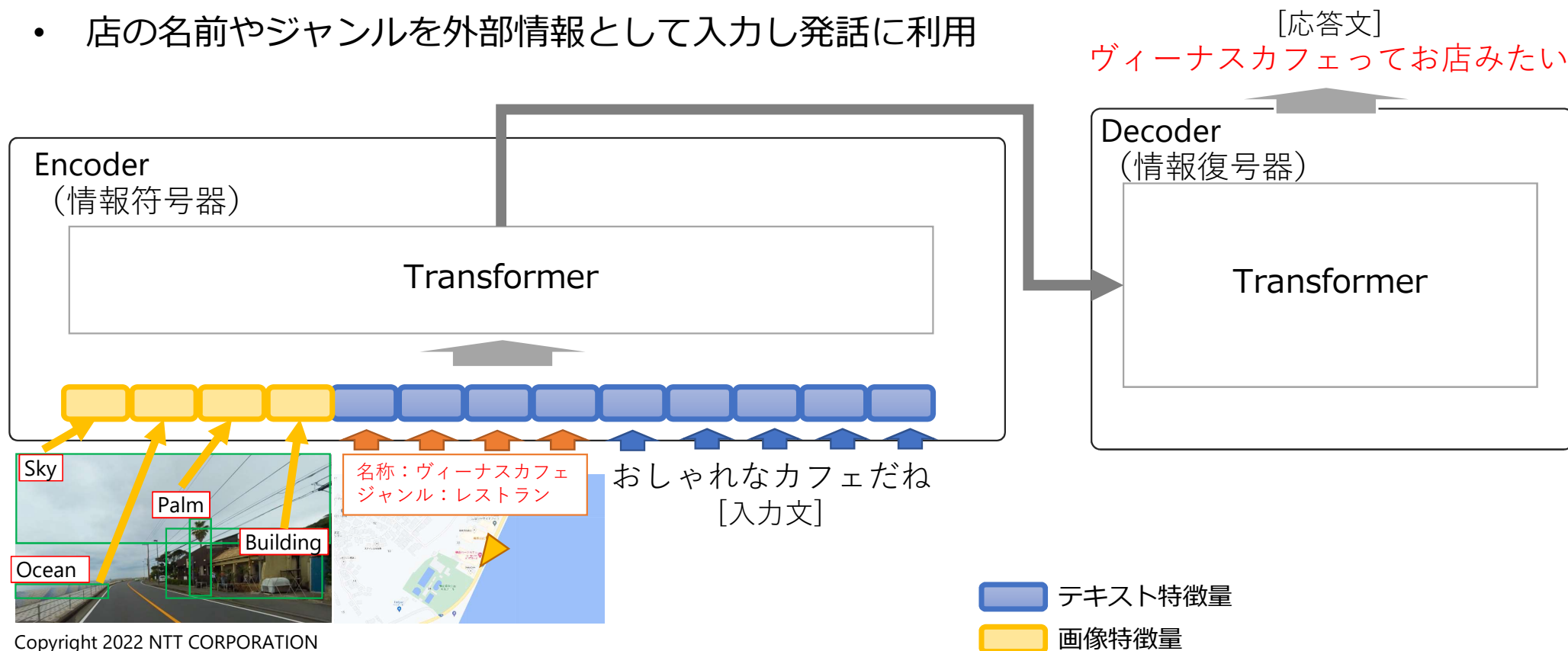
- ポイント1：画像情報の導入

- MMB同様，STAIR Captions（画像説明タスク用データ）[吉川+ 2017] を利用し
画像→テキストの変換ができるように追加学習
- 画像情報抽出にはDeformable-DETRを利用（物体検出）
- 現状では過学習ぎみに学習（江ノ島・江ノ電）



構築したモデルの詳細

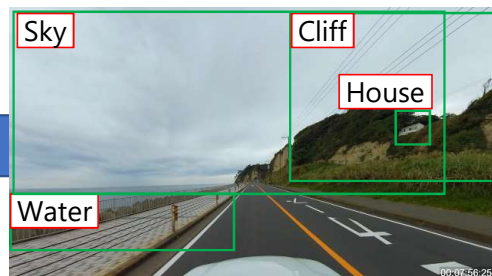
- ポイント2：地理情報の活用
 - 見えている画像中の，発話と対応する画像領域と類似する外見のスポットを自己位置周辺の地理情報を地図情報サービスから検索
 - 店の名前やジャンルを外部情報として入力し発話に利用



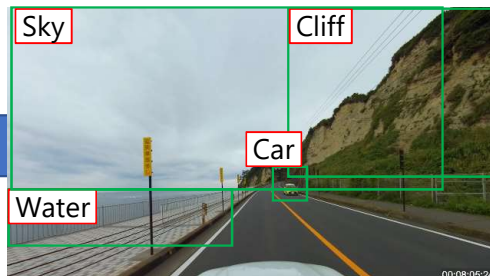
構築したモデルの詳細



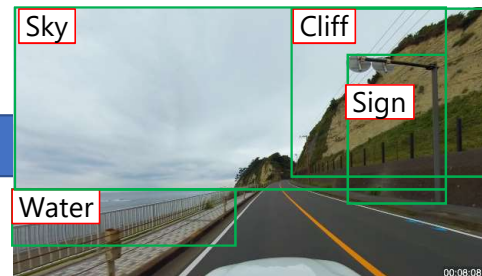
- ポイント3：連続的に生成される発話候補からの選択
 - 生成した発話の話題の強さを推定（人手付与データを利用）



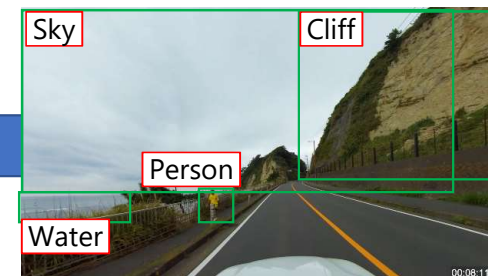
山の上に家があるよ！



珍しい色の車だね



崖がちょっと怖いね

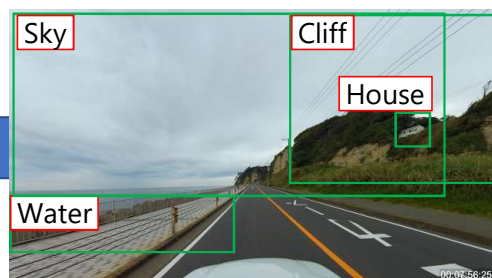


黄色いおじさんが歩いてる

構築したモデルの詳細



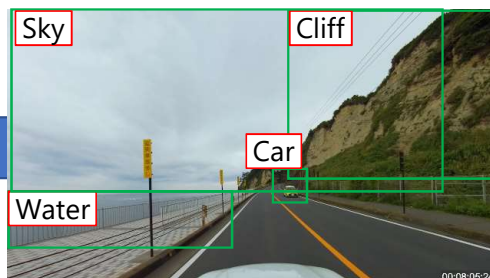
- ポイント3：連続的に生成される発話候補からの選択
 - 生成した発話の話題の強さを推定（人手付与データを利用）



山の上に家があるよ！

話題強度推定器

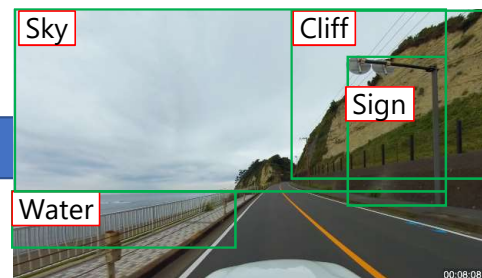
2.7



珍しい色の車だね

話題強度推定器

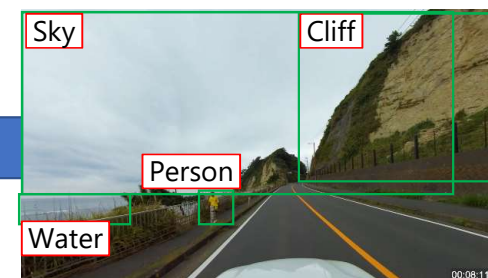
1.7



崖がちょっと怖いね

話題強度推定器

3.1



黄色いおじさんが歩いてる

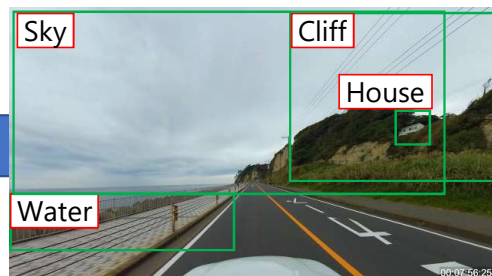
話題強度推定器

2.4

構築したモデルの詳細



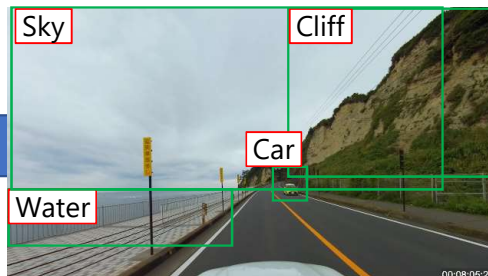
- ポイント3：連続的に生成される発話候補からの選択
 - 生成した発話の話題の強さを推定（人手付与データを利用）



山の上に家があるよ！

話題強度推定器

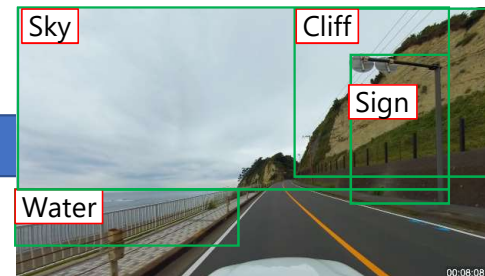
2.7



珍しい色の車だね

話題強度推定器

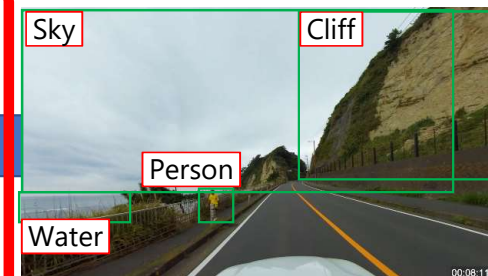
1.7



崖がちょっと怖いね

話題強度推定器

3.1



黄色いおじさんが歩いてる

話題強度推定器

2.4

本講演のまとめ



- 一問一答型の雑談対話システム研究の概要
- 大規模深層学習による雑談対話システムの性能向上
 - Facebook・Googleが一足飛びに性能を引き上げ
 - NTTも規模・質の面で遜色ないレベルのシステムを構築
- 外部情報の様々な利用例について紹介
 - 外部知識の利用
 - 話者情報の利用
 - 対話状況の利用



モデル&コーパスを無償公開中
<https://github.com/nttcsllab/japanese-dialog-transformers>

人が雑談を楽しむ相手となる
対話システムの実現を引き続き目指していきます