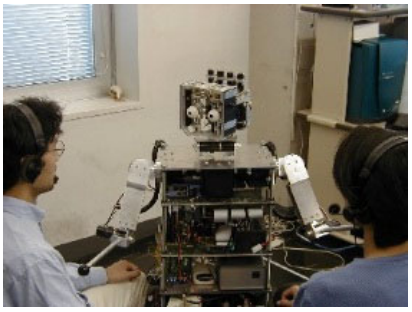
2021.2.27 日本語教師セミナー

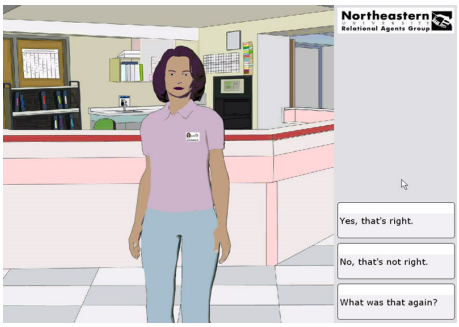
対話システム研究の現状と日本語教育研究者・実践者に期待すること

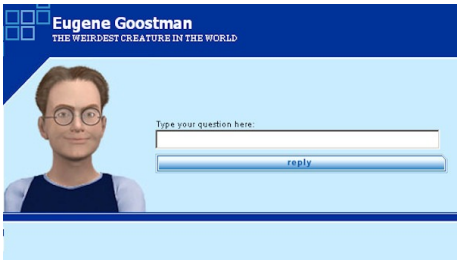
名古屋大学情報学研究科
東中竜一郎
（ひがしなか・りゅういちろう）

さまざまな対話システム



（小林ら，SLUD研究会資料）





http://en.wikinoticia.com/Technology/general-technology/100076-the-nurse-is-a-virtual-reality
http://www.princetonai.com/

タスク指向と非タスク指向

- タスク指向
  - ユーザが達成目標を持っている
    - 目標が達成されたら対話は終了
    - マシンインタフェースとしての対話システム
    - 情報検索など，ユーザの目標が漠然としていることが多い場合も，一般的にタスク指向に含める
- 非タスク指向
  - タスク指向以外の対話システム
    - 明確な目標がなく，対話すること自体や関係構築が目的
    - 雑談対話システムとも呼ばれる

対話システムの意義

- 複雑な情報を素早く，簡便に入力できる
  （任意の量の情報を任意の単位で伝えられる）
- 両手がふさがっていても操作できる
  （音声の場合）
- 使い方を学ばなくても使える
  （理想としては．実際はそうでもない）

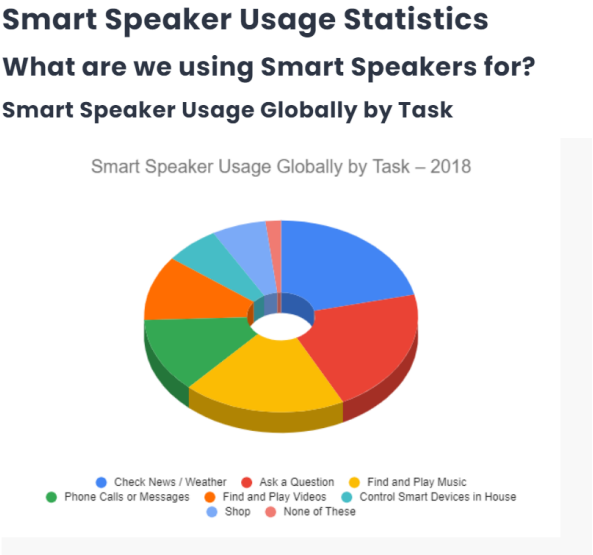
タスク寄りのメリット

- エンターテイメント・教育
  （ゲーム・博物館の展示）
- 会話訓練
  （語学学習者，自閉症患者など）
- コミュニケーションパートナー
  （独居老人など）

雑談寄りのメリット

AIスピーカーの普及

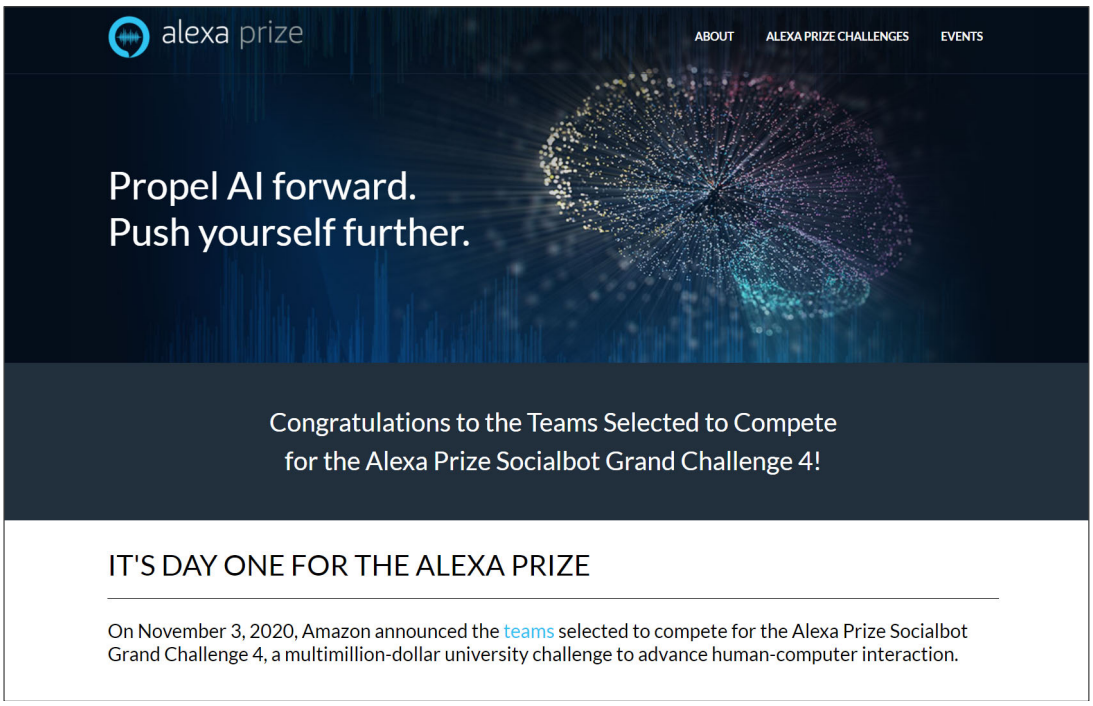
Global Sales and Projections (excluding China) 2016 to 2025



https://speakergy.com/smart-speakers-statistics/

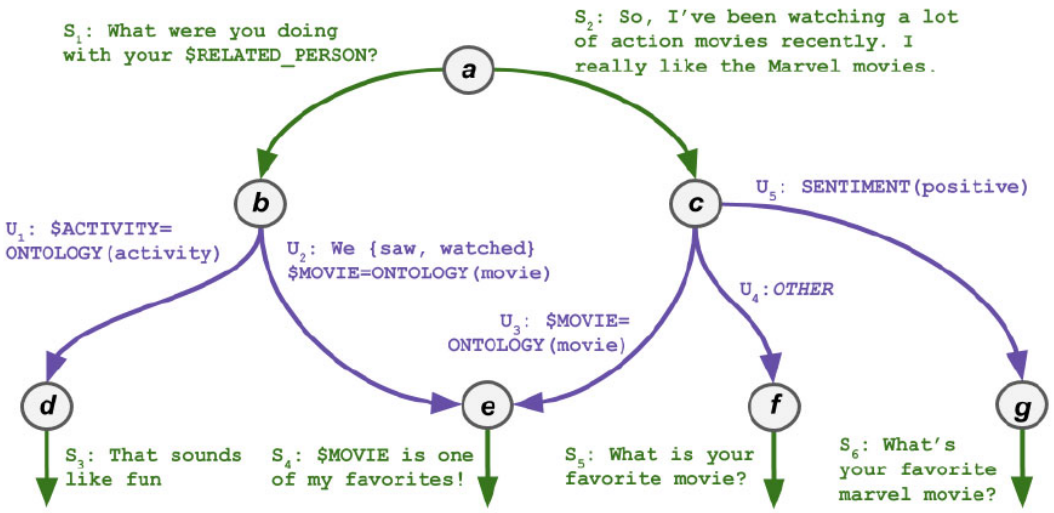
Alexa Prize

https://developer.amazon.com/alexaprize



Alexa Prize のシステム

- サウンディングボード（ワシントン大）
  - 大規模なインターネット上のコンテンツを整備し，ユーザの興味に従った面白い・役に立つ情報を提供
- ガンロック（UC Davis）
  - 多くの話題に対応することに加え，自身のバックストーリーを保持し，個性を持って発話する仕組みを導入
- エモーラ（エモリー大）
  - ユーザとシステムが体験を語り合うことを目的としたシステム
  - 対話の状態を綿密に設計し，対話を制御



https://m.media-amazon.com/images/G/01/mobile-apps/dex/alexa/alexaprize/assets/challenge3/proceedings/Emory-Emora.pdf

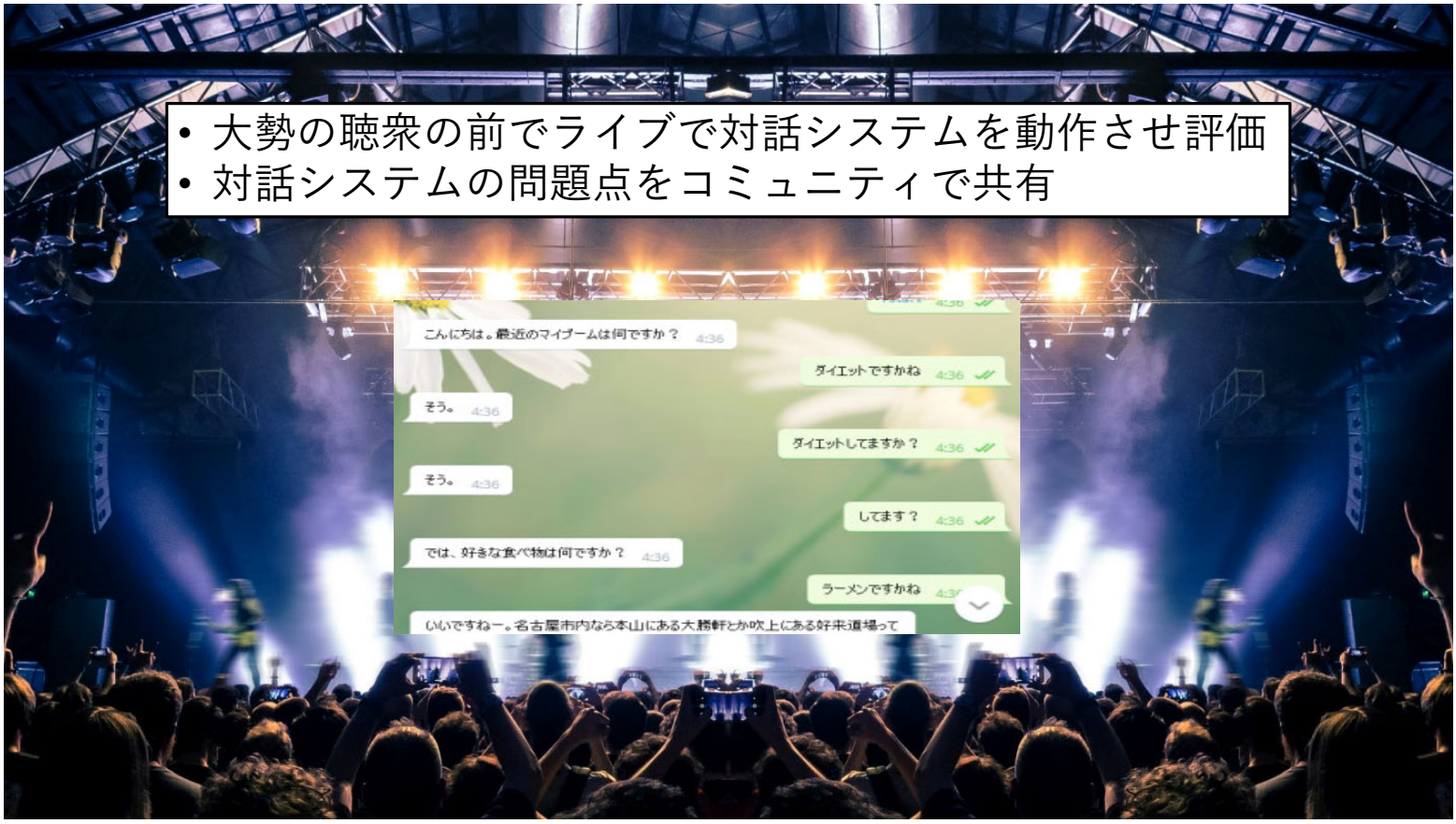
ID	Utterance	Variable Table	Stack
E1	I've been hearing a lot from people about how strange it is to live with the corona virus. Has your life changed much?		[ ]
U1	Yeah my school has online courses now	student=True remote=True	[CoS]
E2	Oh, are you liking your online classes?		[CoS]
U2	Not really		[CoS]
E3(a)	Oh yeah, it is hit and miss for a lot of people. Definitely just try to do the best you can.		[CoS]
E3(b)	Actually, I have seen something pretty uplifting recently. Have you seen those videos where zoos are letting some of their animals out of their cages to roam around?		[ ]
U3	No		[ ]
E4(a)	My favorite one was the penguins at the Chicago zoo, you should check it out!		[ ]
E4(b)	Anyway, the good news is that this virus won't last forever and people are taking steps in the right direction to lower its impact. I know things might seem weird right now, but just do the best that you can and stay positive.		[ ]
U4	I will try to but its pretty stressful, i am having a hard time buying groceries	need_grocery=True	[CoE]
E5	Yeah, that's tough. None of my friends can find necessities like toilet paper in stores anywhere. Honestly, who would have ever imagined there would come a time when toilet paper was such a sought after item?		[CoE]
U6	It is crazy		[CoE]

対話システムのコンペティション

	固定のコーパスで評価	実際の対話で評価
タスク指向型対話	• Dialogue state tracking challenge (DSTC)	• DiaLeague • Spoken Dialogue Challenge • 対話ロボットコンペティション
非タスク指向型対話（雑談対話）	• 対話破綻検出チャレンジ (DBDC) • NTCIR Short Text Conversation (STC)	• ローブナー賞 • Amazon Alexa Prize • Conversational Intelligence Challenge • 対話システムライブコンペティション

誰かが行った対話の評価値が示される

もっとわくわくするイベントを作りたい・・・
対話を生で見たい
みんなで問題点を共有して，対話システム研究を加速させたい・・・


• 大勢の聴衆の前でライブで対話システムを動作させ評価
• 対話システムの問題点をコミュニティで共有
こんにちは。最近のマイブームは何ですか？
ダイエットですかね
そう。
ダイエットしてますか？
そう。
してます？
では、好きな食べ物は何ですか？
ラーメンですかね
いいですねー。名古屋市内なら本山にある大勝軒とか吹上にある好来道場って

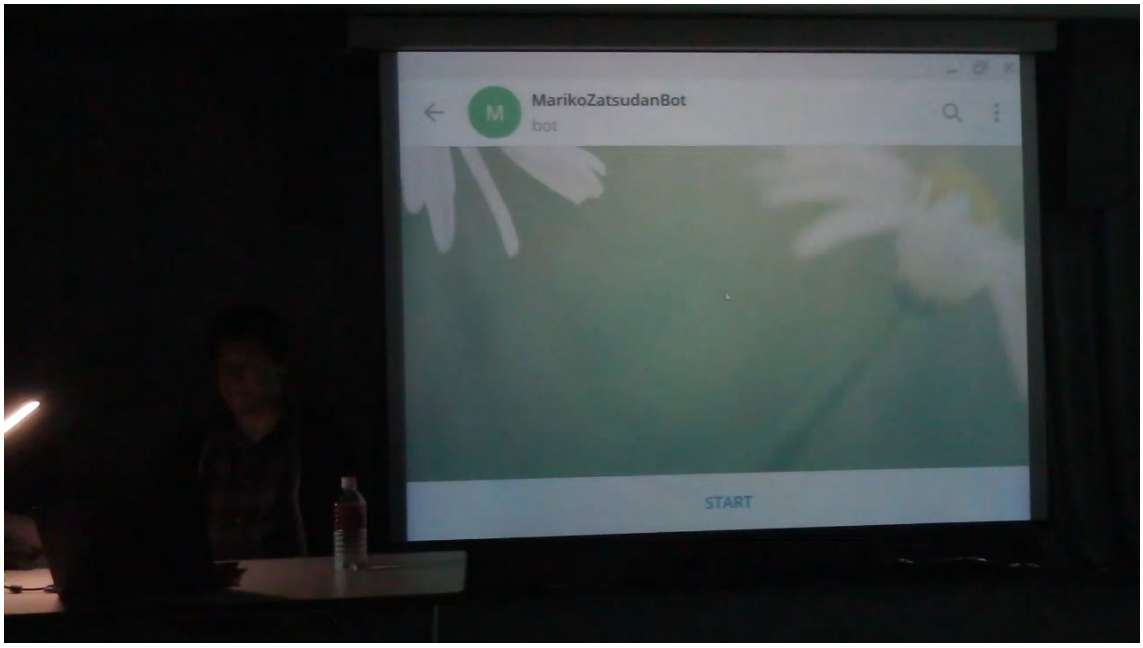
第１回ライブコンペ

- やったこと
  - オープントラック
    - オープンドメインでのテキストチャット
    - プラットフォームにはTelegramを利用
    - 評価はもう一度話したいかどうか
  - 予選（クラウドソーシング）と本選
- 結果
  - 1位 NTTCS研，2位NTTドコモ，3位東北大
- 反省点
  - 対話が人間らしくない（by宇佐美先生）
    - 対話で重要な社会的な要素が入っていない

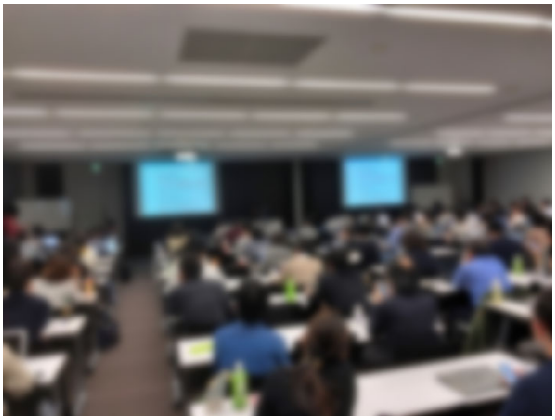
対話者がシステムと対話

聴衆が対話を見て評価

対話をしている様子



第２回ライブコンペ



会場の様子

- やったこと
  - シチュエーショントラックを新設
    - 評価は人間らしい状況に合った対話かどうか
- 結果
  - オープン
    - 1位東工大，2位NTTCS研，3位電気通信大
  - シチュエーション
    - 1位大阪大，2位NTTCS研，3位オーガナイザ
- 反省点
  - オープントラックとシチュエーショントラックの近接
    - オープントラックの話題の限定とシチュエーショントラックの社会関係の不使用
  - 一回きりの評価を逆手に取る（一本道シナリオによる強行突破）
  - 参加者に人文系が少なかった→講習会の実施へ

シチュエーション

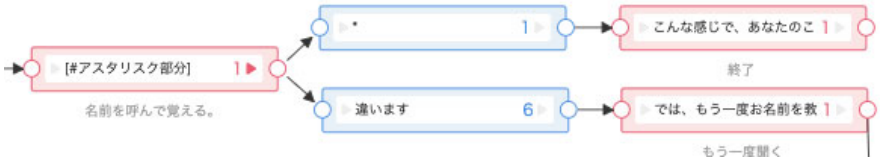
システム：　名前:田中アイ(女)/アキラ(男), 年齢:20~30代, 職業:会社員
ユーザ：　名前:鈴木ユウコ(女)/ユウキ(男), 年齢:20~30代, 職業:会社員
話者の関係：　同性同士, 学生時代の友人関係
場所・時間：　自宅, 暇な時間
話題：　一番印象に残った旅行・場所

田中と鈴木は，学生時代，仲の良い友人同士であった．2人とも大学を卒業して会社員になってからはときどき食事に行ったりしていたが，ここ２，３年は会う機会も連絡をとることもなくなっていた．ある日，田中が自宅でのんびり過ごしていると，鈴木からテキストメッセージが送られて来た．鈴木も家で暇にしていたらしく，ふと気になって連絡をくれたらしい．久しぶりにお互いの近況報告をする中で，最近出かけた場所などが話題になった．
田中「ところで，これまで行ったところで一番印象に残った場所ってどこ？」←この発話から会話をする．

対話システム講習会

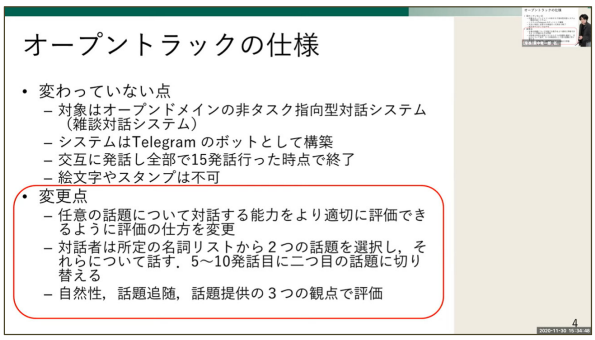
- 目的
  - ライブコンペの参加者を増やすこと
  - 人間同士の対話の知見を持つ人文系研究者の取り込み
- チャレンジ
  - 対話システム構築の敷居を下げること
- やったこと
  - 講義（座学）と演習
  - 動画資料の公開
  - SUNABAとTelegramの接続ソフトの開発

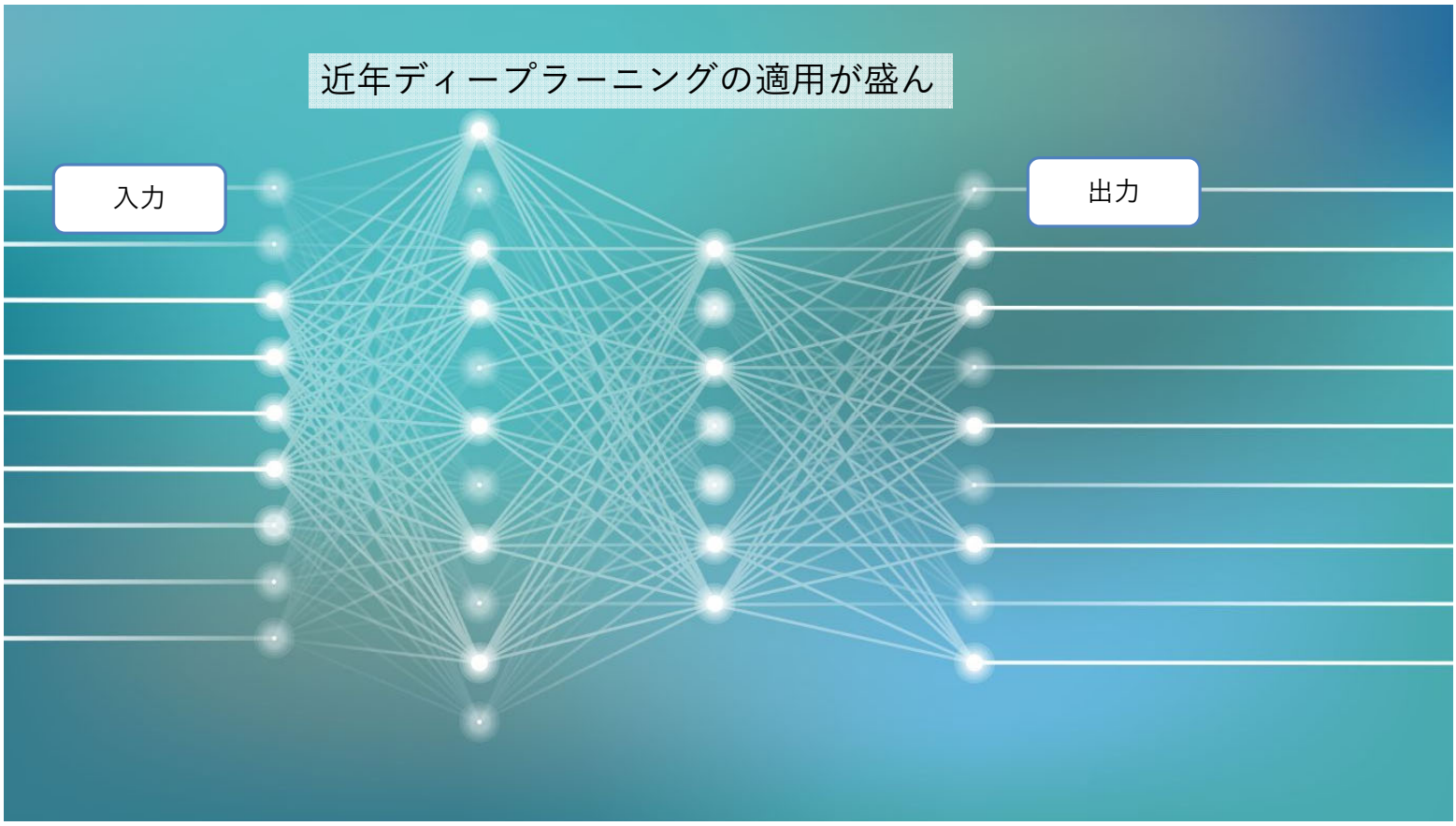
SUNABAでのシナリオ作成の様子



第３回ライブコンペ（オンライン開催）

- やったこと
  - オープントラックに単語リストを導入
  - シチュエーショントラックに社会関係（先輩後輩）を導入
  - 本選で演舞（うまく行った対話を見せること）を導入
- 結果
  - オープン
    - 1位NTTCS研，2位東北大
  - シチュエーション
    - 1位FCL，2位明治大学，3位オーガナイザ
- 反省点
  - もっとわくわくを！みんなで共有を！





BlenderBot（1/2）

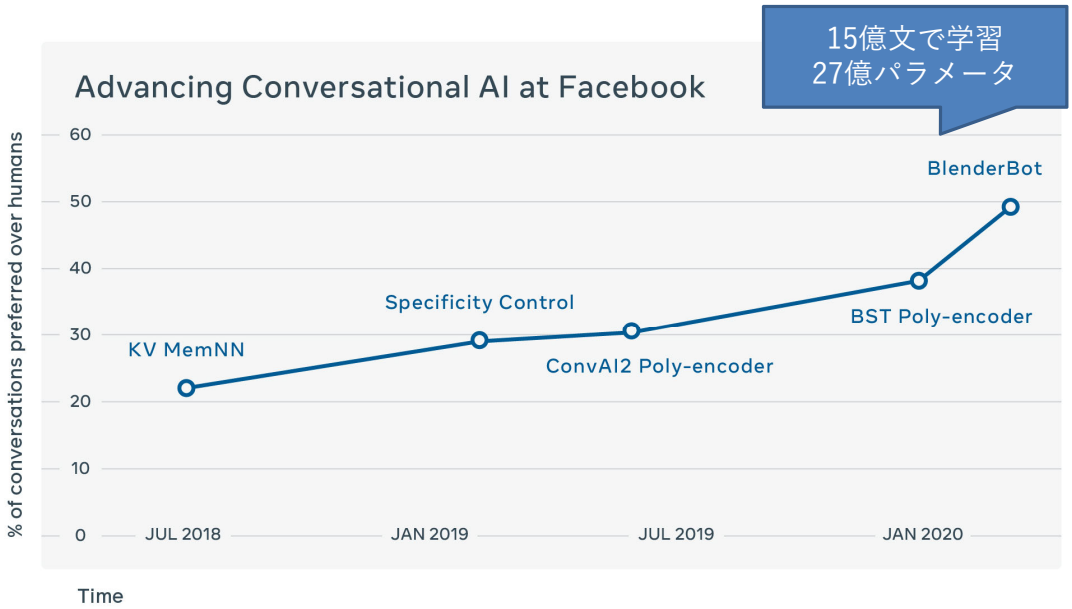
System: Do you have any hobbies that you like to do in your spare time? I like to play video games.
Enter Your Message: I love detroit become human!
System: Detroit Become Human is a great game. Have you played it before?
Enter Your Message: no
System: I highly recommend it. It's one of the best games I've played in a long time.
Enter Your Message: what did you like about it?
System: I really liked the story and the choices you had to make. It was a lot of fun.
Enter Your Message: what is your favorite character?
System: I really enjoyed playing as Connor. He was a great protagonist. How about you?
Enter Your Message: I like Kara.
System: I liked Kara as well. She was a very interesting character. I liked that she was a reporter.

BlenderBot（1/2）

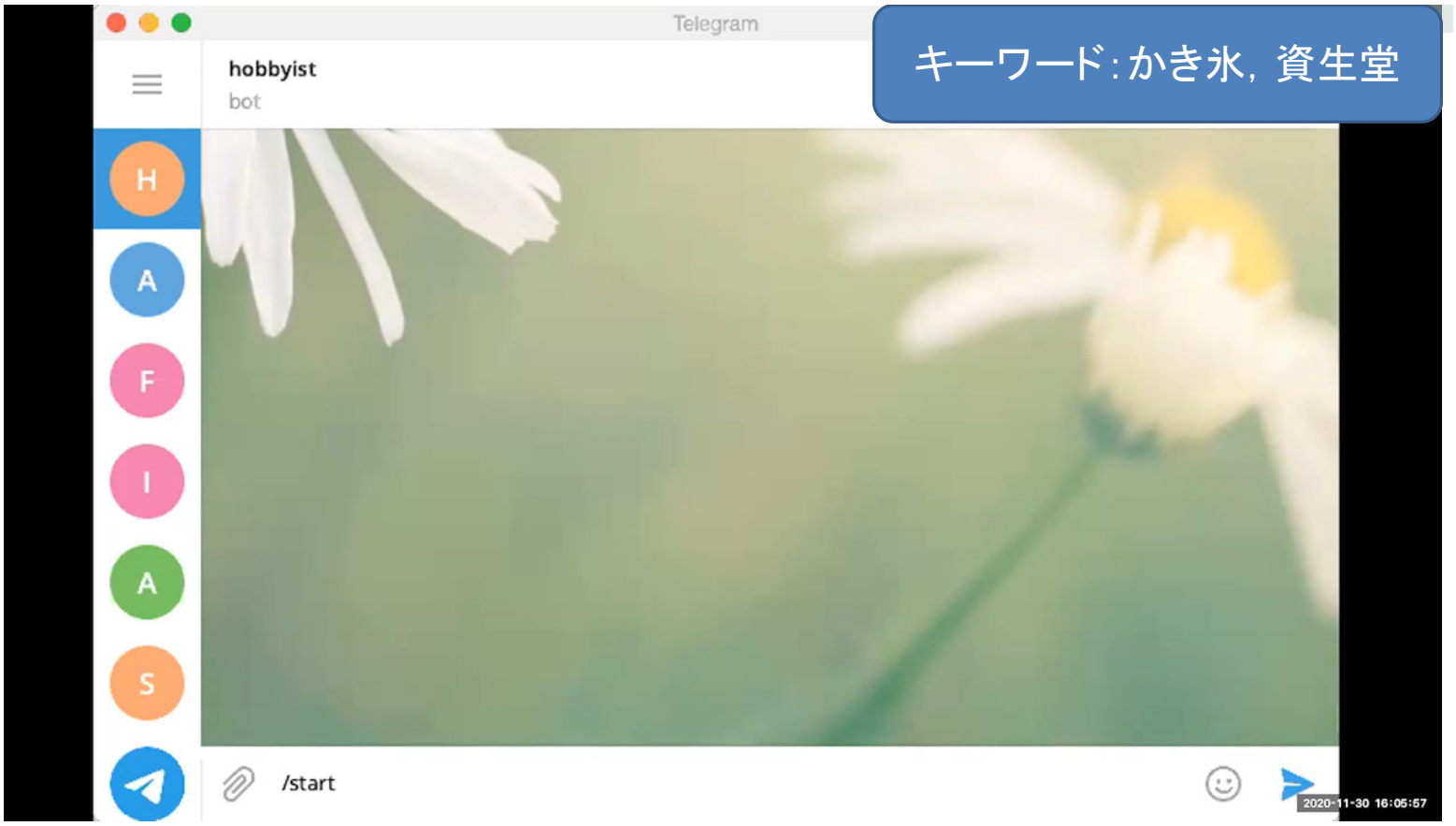
システム：暇な時にやってみたい趣味はありますか？テレビゲームをするのが好きです。
メッセージを入力してください。デトロイト・ビカム・ヒューマンが好きです
システム: デトロイト・ビカム・ヒューマンは素晴らしいゲームです。以前にプレイしたことがありますか？
メッセージを入力してください: いいえ
システム：私はそれを非常にお勧めします。長い間プレイした中で最高のゲームの一つです。
メッセージを入力してください：何が好きでしたか？
システム: ストーリーと選択肢がとても気に入りました。とても楽しかったです。
メッセージを入力してください: 好きなキャラクターは？
システム: コナーとしてプレイするのが本当に楽しかったです。彼は素晴らしい主人公でした。あなたはどうですか？
メッセージを入力してください。カーラが好きです。
システム：私もカーラが好きでした。彼女はとても面白いキャラクターでした。彼女がレポーターであることが好きでした。

https://www.deepl.com/ja/translator

BlenderBot （2/2）



https://ai.facebook.com/blog/state-of-the-art-open-source-chatbot/



キーワード：かき氷，資生堂

シチュエーション

システム(後輩)：清水シズカ(女)/シンジ(男)，大学4年生，22歳前後
ユーザ(先輩)：湯川ユウコ(女)/ユウキ(男)，会社員，25歳前後
話者の関係：同性同士．学生時代のサークルの先輩後輩の関係．学生時代は親しかったが，先輩が卒業してからは，連絡を取り合っていない．
場所・時間：それぞれの自宅，清水が卒業の年，春休みの夜の暇な時間
状況・話題：**先輩からの「同窓会の幹事の依頼」を断りたい状況**

清水(システム)と湯川(ユーザ)は，学生時代，同じサークルの先輩・後輩の関係であった．清水にとって湯川は，いろんなことについて相談しやすい先輩であり，湯川にとって清水は自分を頼ってくるかわいい後輩であった．ただ，湯川が大学を卒業して会社員になってからは，会う機会も連絡をとることもなくなっていた．清水も就職が決まり，自宅でくつろいでいた春休みのある日の夜，久しぶりに湯川からテキストメッセージが送られて来た．卒業式の時期を迎え，清水も卒業の年になったということを思い出してくれたようで，連絡をくれたらしい．久しぶりにお互いの近況報告をしたあと，**湯川がおもむろに頼み事をしたいと言ってきた．どうやらサークルの同窓会の幹事を，特に清水に任せたいらしい．**清水としては決して暇がないわけではないのだけれど，日程調整やら出欠の確認やらは非常に面倒だ．とはいえ湯川は非常にお世話になった先輩であり，失礼な態度もとりづらい．なんとか頼みを断りたいけれど，さてどう言って断ろうか．

	雑談対話システムの問題点（２０１８）
1	【発話理解】発話内容の理解能力が低い／浅い
2	【話題】適切な話題遷移ができない
3	【発話理解】質問に答える能力が低い
4	【発話生成】発話内容が文脈に沿っていない
5	【対話の意義】システムの意図が不明確
6	【ユーザ適応・個性】過去のやり取りを活用できていない
7	【対話の意義】対話の目的が不明確
8	【発話生成】同じ内容を繰り返し生成する
9	【ユーザ適応・個性】ユーザに応じた応答ができていない
10	【話題】広い話題に対応できない

	雑談対話システムの問題点（２０１９）
1	【発話理解】発話内容の理解能力が低い／浅い
2	【話題】広い話題に対応できない
3	【ユーザ適応・個性】ユーザに応じた応答ができていない
4	【話題】適切な話題遷移ができない
5	【ビジネス・開発面】品質評価が困難
6	【その他の機能】適切な主導権の交代ができない
7	【発話理解】質問に答える能力が低い
8	【発話生成】発話内容が文脈に沿っていない
9	【対話の意義】対話に面白さがない
10	【ユーザ適応・個性】過去のやり取りを活用できていない

	雑談対話システムの問題点（２０２０）
1	【発話生成】発話内容が文脈に沿っていない
2	【発話理解】発話内容の理解能力が低い／浅い
3	【ユーザ適応・個性】ユーザに応じた応答ができていない
4	【発話生成】自身の発話と矛盾した発話を生成する
5	【対話の意義】対話に面白さがない
6	【話題】広い話題に対応できない
7	【知識】事実と異なる情報を伝える
8	【ユーザ適応・個性】会話を進めてもシステムとの関係性（好感度など）が変わらない
9	【その他の機能】対話破綻から適切にリカバリできない
10	【話題】適切な話題遷移ができない

日本語教育学会でのアンケート結果

	雑談対話システムの問題点
1	【発話生成】発話内容が文脈に沿っていない
2	【話題】適切な話題遷移ができない
3	【発話理解】発話内容の理解能力が低い／浅い
4	【ユーザ適応・個性】ユーザに応じた応答ができていない
5	【対話の意義】対話の目的が不明確
6	【話題】広い話題に対応できない
7	【その他の機能】対話破綻から適切にリカバリできない
8	【発話理解】質問に答える能力が低い
9	【発話理解】感情の理解能力が低い
10	【発話生成】日本語として不自然な文を生成する

日本語教育研究者・実践者への期待①

- 対話能力の評価支援
  - 現状の対話システムの評価軸は不十分
    - また話したいかどうか
    - 情報が多く含まれているかどうか
    - 人間らしいかどうか（チューリングテスト）

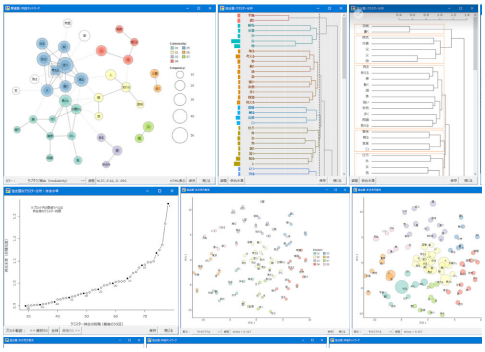
- どんなシチュエーションの対話に着目すれば，言語運用の能力を適切に測れるのかが知りたい
  - 例：Oral Proficiency Interview

第５課　断る			
RP1	友だちの誘いを断る	授業が終わり友だちのＢさんに、お茶に誘われました。今日はアルバイトがあるので行けません。Ｂさんにうまく話してください。	P.94 P.104
RP2	友だちの頼みを断る	先週、授業を休んだ友だちのＢさんに、授業のノートを今日一日貸してほしいと頼まれました。でも、試験が近いので、できれば貸したくありません。うまく話してください。	P.94
RP3	友だちの無理な頼みを断る	友だちのＢさんに、旅行に行くのでペットを預ってほしいと言われました。でも、ペット禁止のアパートです。友だちにそのことを話してください。	P.94 P.106
RP4	相手の申し出を断る	学校で授業中に気分が悪くなりました。放課後、病院に行こうとしていたら、「一緒に行こうか？」と友だちのＢさんが言ってくれました。でも、Ｂさんは忙しそうなので、１人で行こうと思います。うまく話をしてください。	P.96
RP5	頼みを聞いてもらったことのある友だちの依頼を断る	クラスのＢさんに、来週の月曜日の発表を代わってもらえないかと頼まれました。以前、代わってもらったことがあるのですが、その週は忙しく、準備ができそうにありません。うまく話をしてください。	P.96 P.109
RP6	もうすぐ帰国する友だちの誘いを断る	来週、友だちのＢさんは帰国するそうです。そのＢさんに、今週の土曜日、お台場に行かないかと誘われました。でも、その日大切な約束があります。うまくＢさんと話をしてください。	P.96

『ロールプレイ玉手箱』（ひつじ書房, 2009）

日本語教育研究者・実践者への期待②

- 教育用途への対話システム活用
  - シナリオ構築は人文系の方が得意
    - 白井さんの優勝
  - 本当に役立つシナリオをDB化



https://khcoder.net/

- KHCoderを目指した営み
  - SUNABAとTelegramの連携ソフトウェア
  - 名古屋大での異分野連携プロジェクト
    - マルチモーダル対話システムを簡単に構築できるツールを作成中

日本語教育研究者・実践者への期待③

- 対話データ共有の仕組みづくり
  - サイロ化を避ける

- データの例
  - 日本語日常会話コーパス
  - BTSJ日本語自然会話コーパス
  - DailyDialog
    - 英語学習者向けの対話例
    - 深層学習で広く活用

A: I'm worried about something.
B: What's that?
A: Well, I have to drive to school for a meeting this morning, and I'm going to end up getting stuck in rush-hour traffic.
B: That's annoying, but nothing to worry about. *Just breathe deeply when you feel yourself getting upset*.
A: Ok, I'll try that.
B: Is there anything else bothering you?
A: Just one more thing. A school called me this morning to see if I could teach a few classes this weekend and I don't know what to do.
B: Do you have any other plans this weekend?
A: I'm supposed to work on a paper that'd due on Monday.
B: *Try not to take on more than you can handle*.
A: You're right. I probably should just work on my paper. Thanks!

Figure 1: An example in **DailyDialog** dataset. Some text is shortened for space. Best viewed in color.

To construct a multi-turn dialog dataset, we crawl the raw data from various websites which serve for English learner to practice English dialog in daily life. That's why we refer it as **DailyDialog** dataset. The dialogues in the dataset preserve the following three appealing properties.

Dailydialog: A manually labelled multi-turn dialogue dataset
Y Li, H Su, X Shen, W Li, Z Cao, S Niu - arXiv preprint arXiv:1710.03957, 2017 - arxiv.org
We develop a high-quality multi-turn dialog dataset, DailyDialog, which is intriguing in several aspects. The language is human-written and less noisy. The dialogues in the dataset reflect our daily communication way and cover various topics about our daily life. We also manually label the developed dataset with communication intention and emotion information. Then, we evaluate existing approaches on DailyDialog dataset and hope it benefit the research field of dialog systems.
☆ 99 引用元 266 関連記事 全 6 バージョン »

https://www.aclweb.org/anthology/I17-1099.pdf

まとめ

- 対話システム研究の現状
  - タスク指向と非タスク指向
  - Alexa Prize
    - 最新のシステムはほぼシナリオベース
  - 対話システムライブコンペティション
- 日本語教育研究者・実践者に期待すること
  - 言語運用能力の評価における知見
  - シナリオ構築による対話に関する知見の共有
  - 対話データ共有の仕組みづくり
    - DailyDialogの活用を例に