

マルチメディアシステム (対話システム)

名古屋大学大学院情報学研究科
知能システム学専攻
東中竜一郎



対話システムの概要

対話システムとは

- 対話

- 情報の授受を繰り返すこと
- それにより外界に作用を及ぼすこと
- ≡ 会話（但し対話システム研究者にとって会話のほうが雑談よりのニュアンス）

- 対話システム：人間と対話するソフトウェア

文脈の存在が前提。 入力は主に話し言葉



- 自然言語インタフェース（文⇒SQL， 一問一答）
- 対話型インタフェース（Unix shell など）

対話システムの類型

- タスクの有無
（タスク指向型，非タスク指向型）
- 人数（一人，二人，多人数）
- モダリティ（テキスト，音声，タッチパネル，ジェスチャ）
- 主導権（システム，ユーザ，混合主導）
- 身体性（バーチャルエージェント型，ロボット型）

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）

タスク指向
よりのメリット

- エンターテインメント・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）

非タスク指向
よりのメリット

工学の他，**認知科学・人間科学**への貢献も期待される

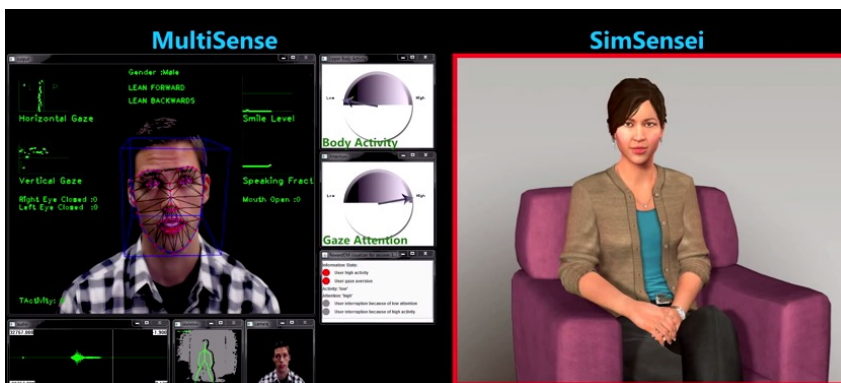


マルチモーダル (マルチメディア) の重要性

マルチモーダル対話

- **モダリティ理解**

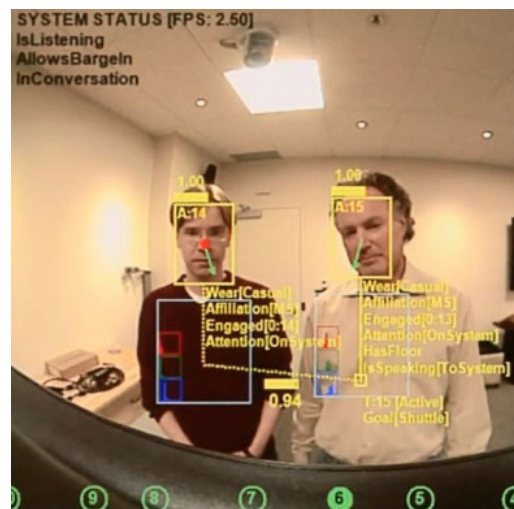
- 顔の向き
- 視線
- ジェスチャー
- 表情



<https://www.youtube.com/watch?v=ejczMs6b1Q4>

- **モダリティ生成**

- ジェスチャー生成
- 視線



<http://research.microsoft.com/en-us/news/features/situatedinteraction-040814.aspx>

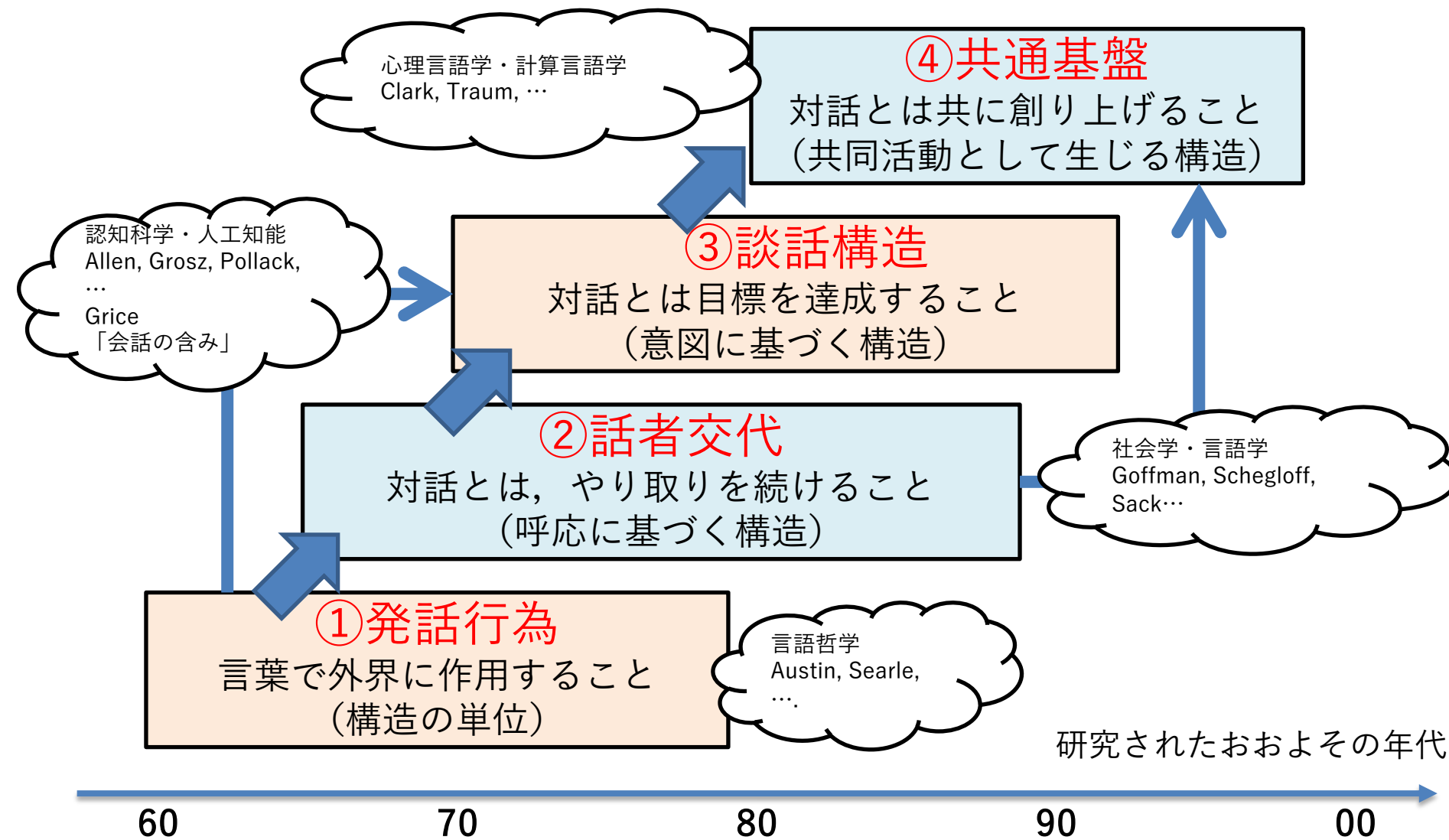
状況依存対話

- センサーを用いた状況理解
 - スマートフォンにおける対話
 - 位置情報
 - 車載対話
 - 位置情報
 - 店舗内対話
 - 人物の位置関係
 - 家具の位置
 - ゲーム
 - ボードの状態



対話という現象

対話理論の構成



① 発話行為

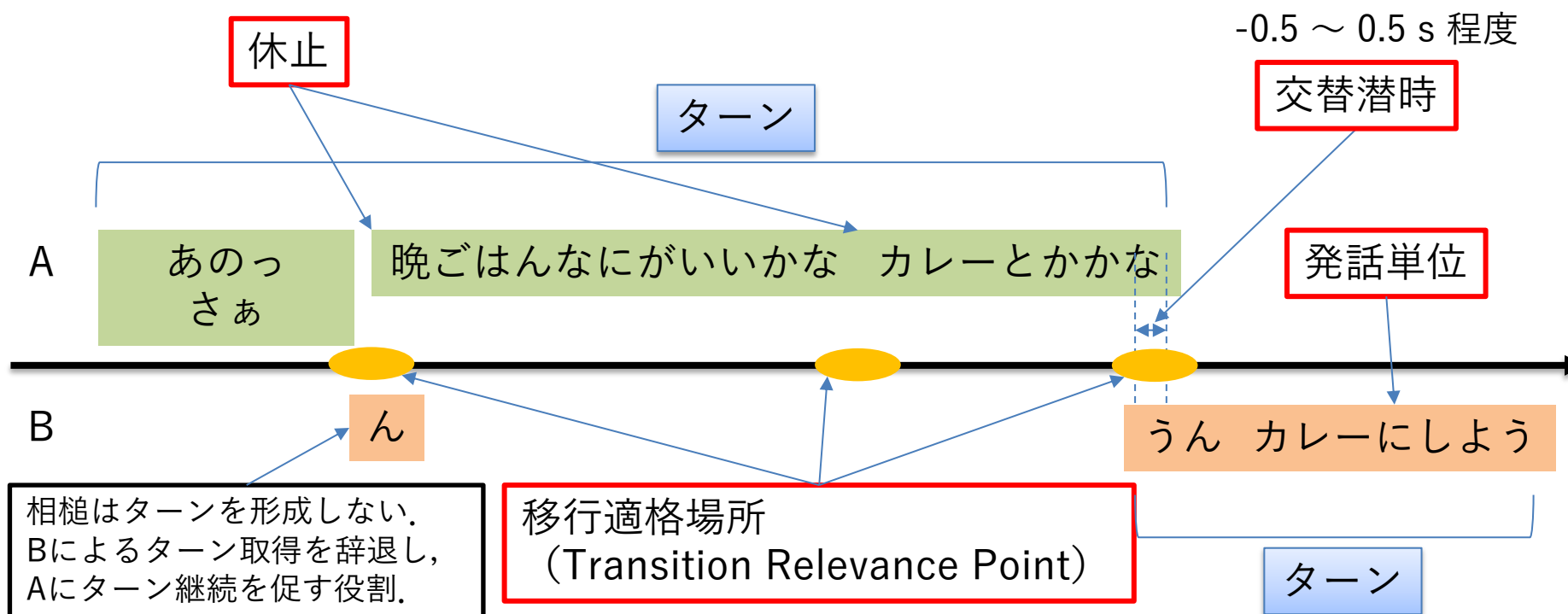
- 「何かを言うことは何かをすることである」

私はこの船をメアリー号と名付ける

発語行為	しゃべる, 板書する
発語内行為	依頼する, 約束する
発語媒介行為	笑わせる, 脅す(相手の捉え方による)
間接発話行為	「とれる?」≒「とって」[Searle 79]

②話者交代

- 対話：ターンの繰り返し（作業仮説）
- ターン：1人の話者が話している区間



話者交代における現象

- **バックチャネル**

- 発話中の話し手へ肯定的・支持的な即時的応答
 - 相槌もバックチャネルの一種
 - ターンを奪取しない

- **バージイン**

- 発話中の話し手への聞き手からの割り込み
 - 相槌はバージインとは捉えない
 - ターンを奪取する
- 音声インタフェースにおける概念（提唱者不明）

③ 談話構造

談話セグメント (対話の入れ子構造)



品川で個室がある沖縄料理店を教えてください

副対話



個室がある沖縄料理店は 1軒あります



残念ながら今日は定休日です



五反田は？

副対話



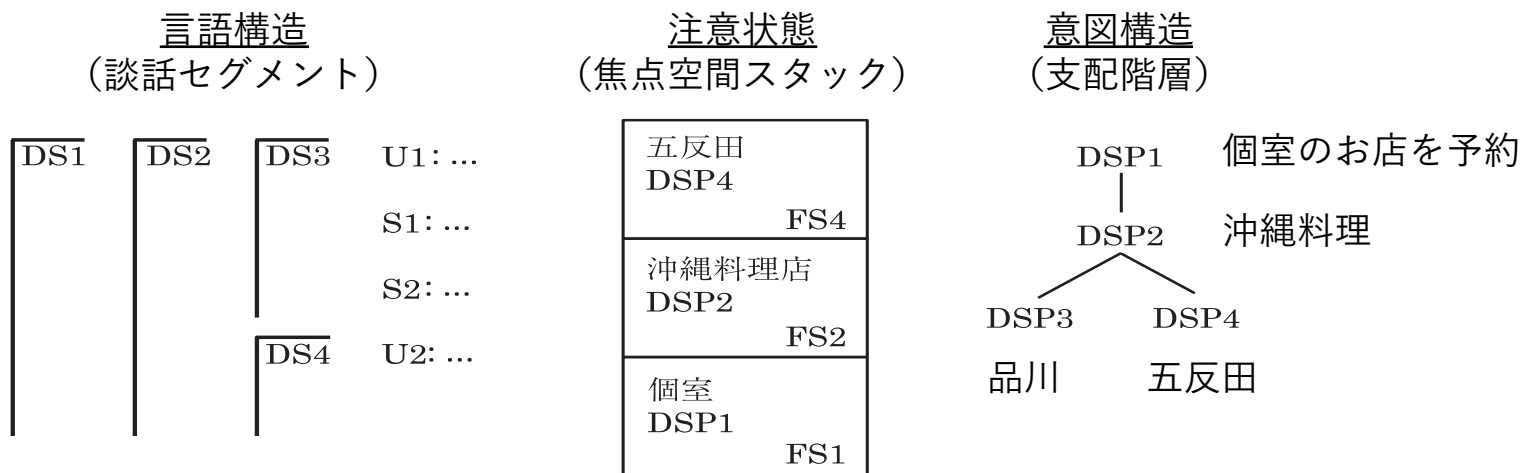
五反田には 2軒あります



お店の情報を読みあげますか？

Groszの談話構造

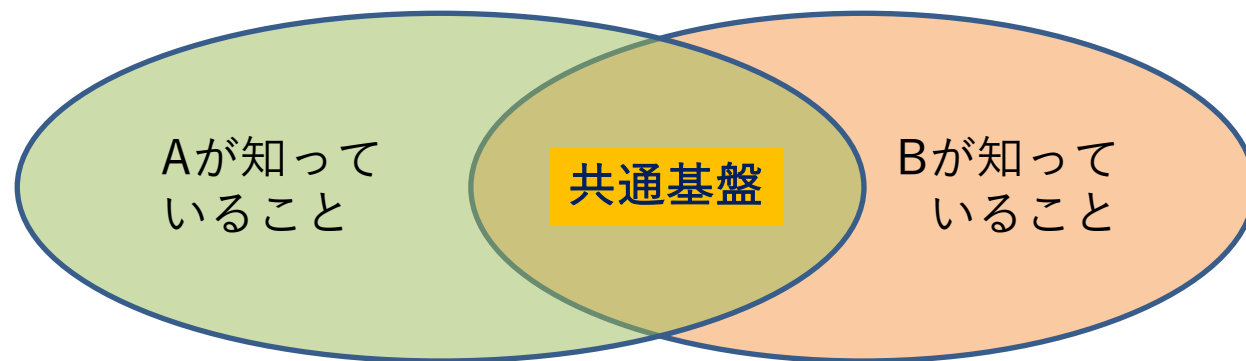
- **< 言語構造, 注意状態, ユーザの意図構造 >**
 - 言語構造はユーザの意図構造に従う
 - 注意状態は現状の話題を管理



※ DSP: Discourse Segment Purpose

④ 共通基盤

- 共通基盤は共同活動により構築される
- 対話は二重構造になっている



メタコミュニケーション（付帯的コミュニケーション）

コミュニケーション
についての対話

コミュニケーション

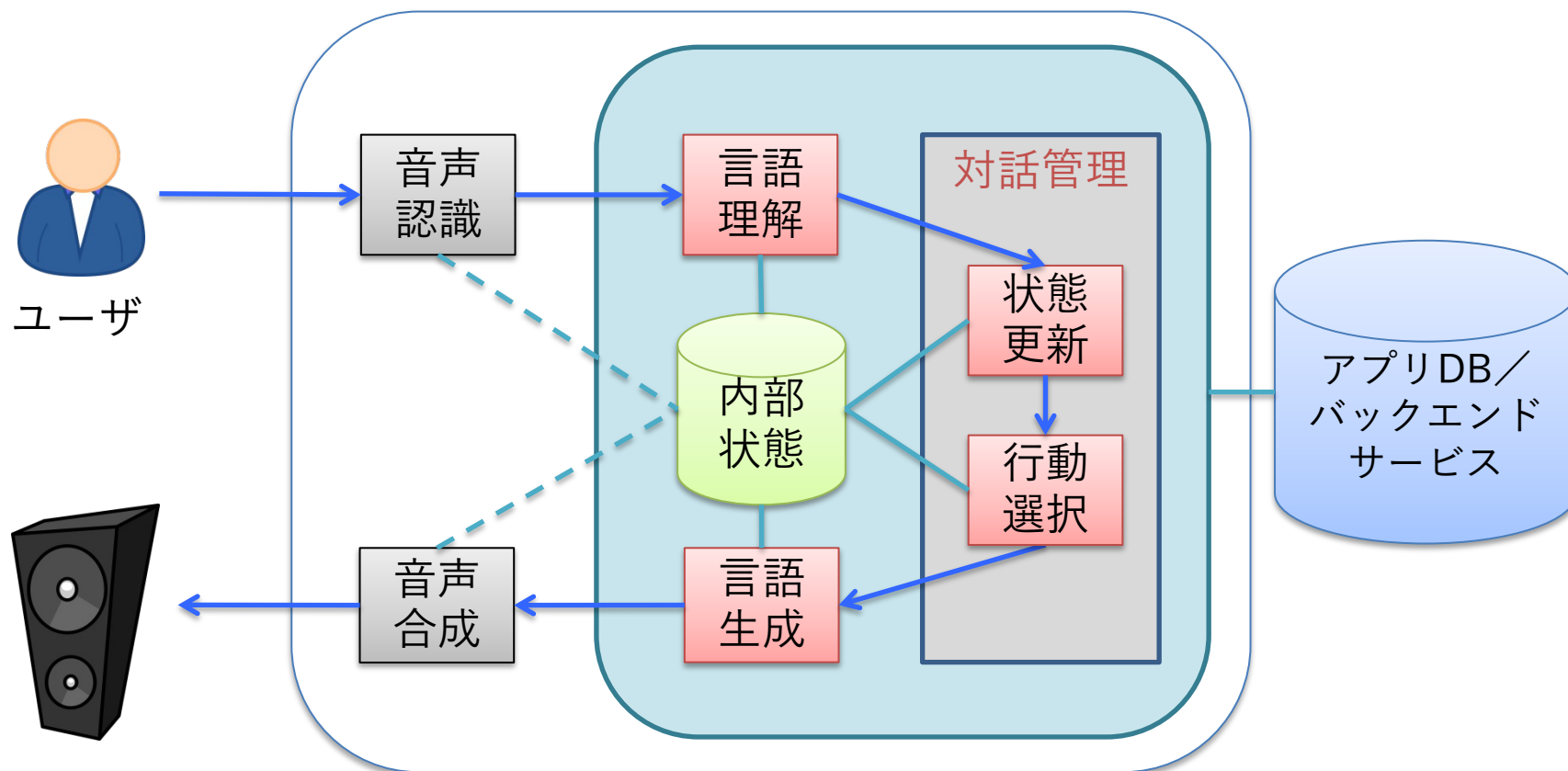
対話の目的についての対話

対話の目的

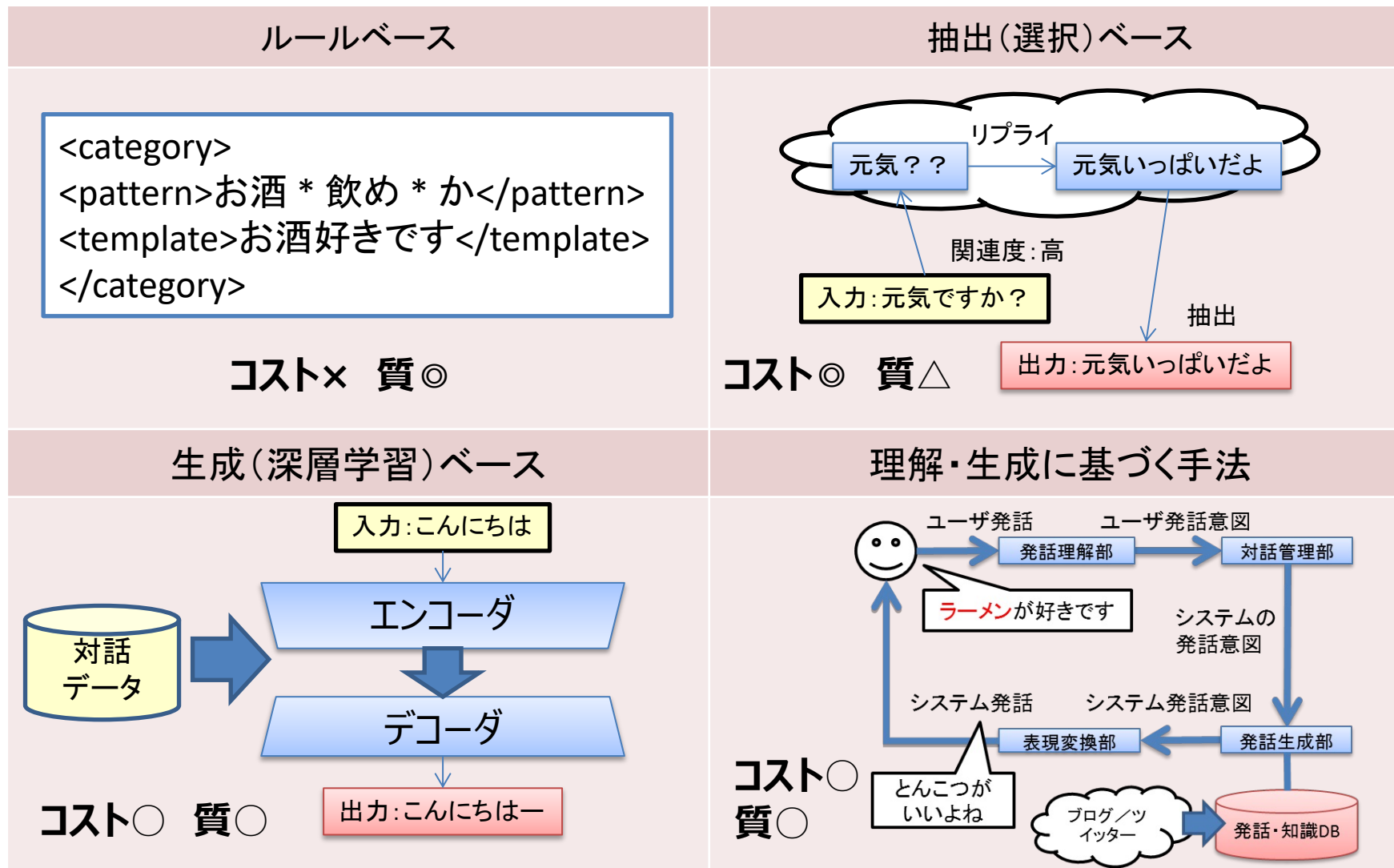


対話システムの構成

タスク指向型対話システムの構成



雑談対話システムの構成





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?

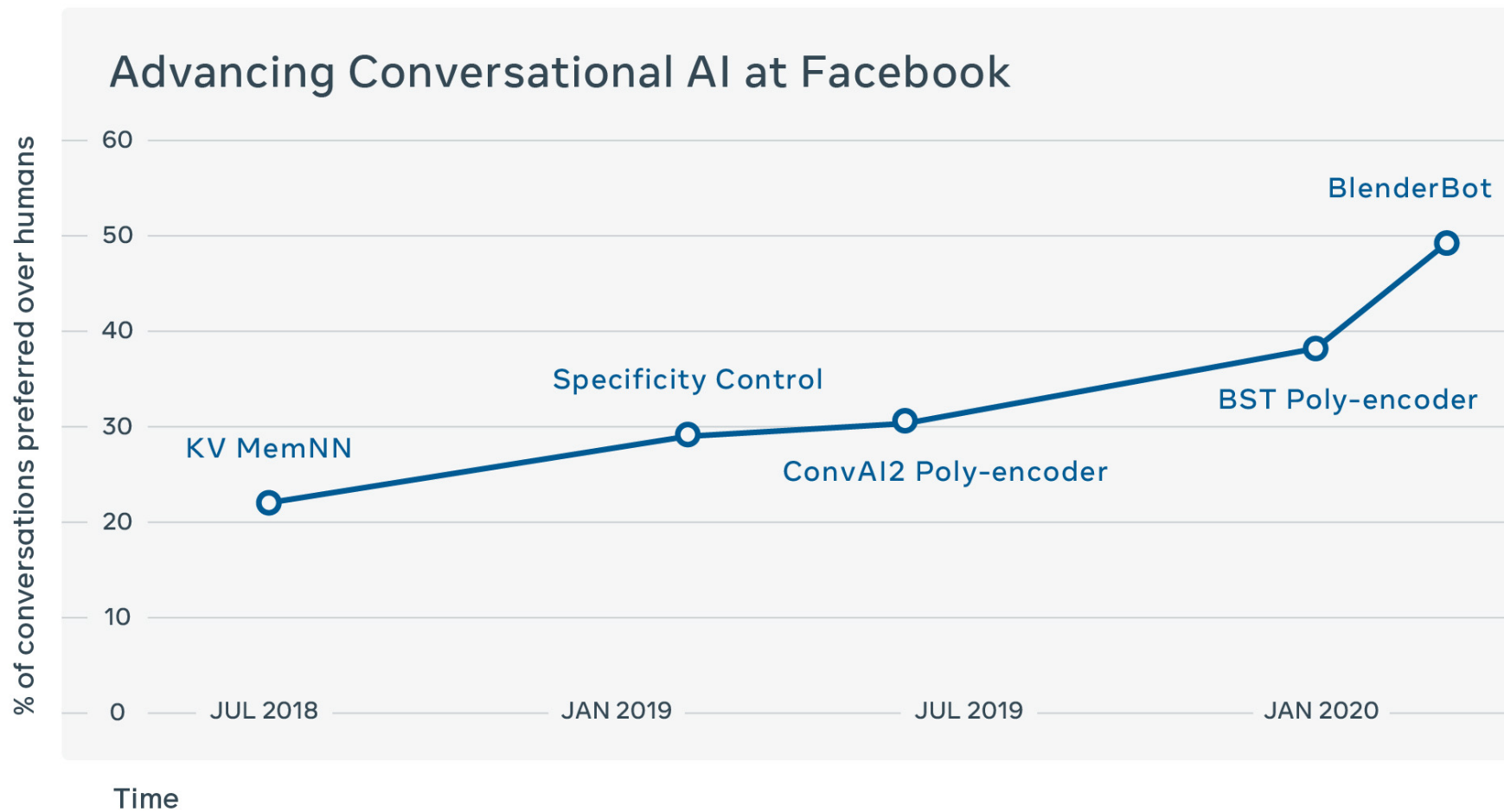
System: I really enjoyed playing as Connor. He was a great protagonist. How about you?

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 (2/2)

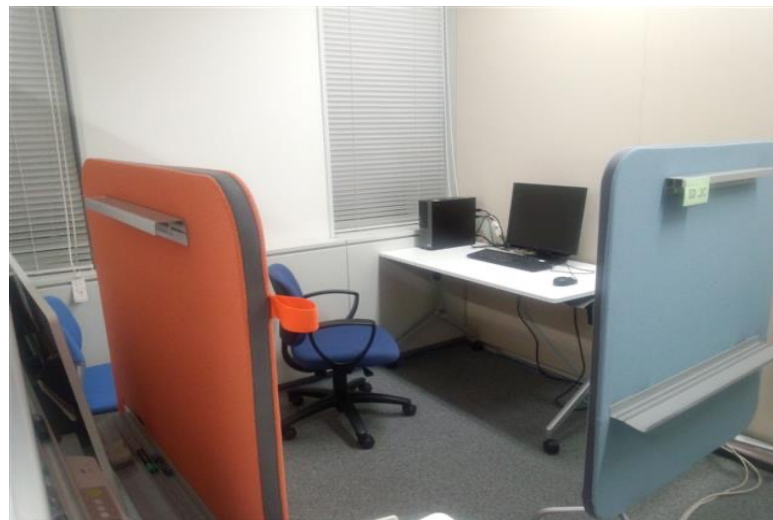


対話システムの評価

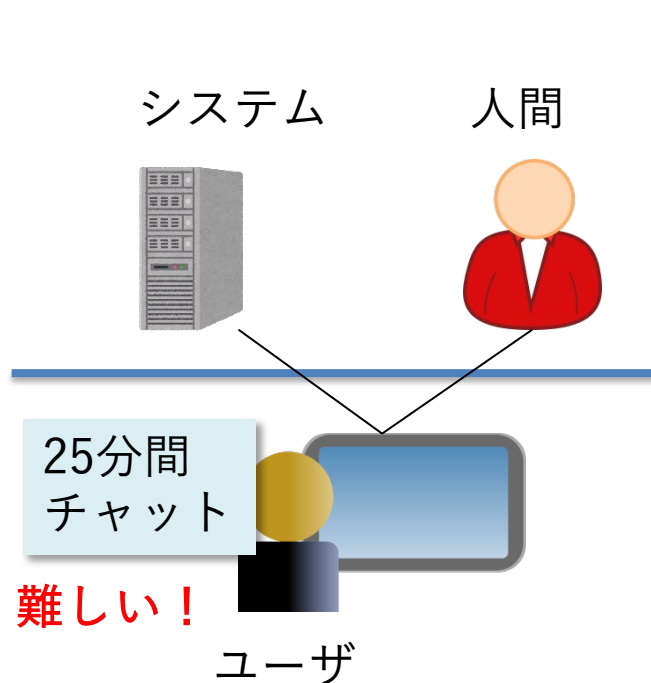
- ユーザ評価
 - ユーザに会話をしてもらいアンケート評価
 - コストは高いが信頼性は高い
- 自動評価
 - 所定のデータセットに対する応答性能を評価
 - コストは低いが信頼性は低い
 - BLEU, distinct など
- シミュレーション評価
 - シミュレーションユーザとの対話
 - 応答アルゴリズムの理論的側面を確認

ユーザ評価

- 実ユーザに対話システムと会話をしてもらいシステムの評価を行う
- 評価の形態
 - ラボ評価
 - フィールド評価
- 評価指標
 - 客観評価指標
 - 主観評価指標
 - 自動評価



チューリングテスト



チューリングテスト

人間・システムの両方と相手
が分からない状態で会話をし
た時，どちらが人間か分から
なければシステムに知性があ
ると判定するテスト [Turing
50]

チューリングテストの一つである
ローブナー賞での結果は人間側の圧勝

倫理的課題

- 対話システムの特徴
 - ユーザに直接的に働きかけることが可能
 - ユーザへの影響が大きく倫理的配慮が重要
- 主な観点
 - 発言の内容
 - 誤情報, 不要な宣伝, ヘイトスピーチ
 - プライバシーの保護
 - 発話・音声・画像データ, 個人情報に基づく発話
 - 利用目的の明示が必要
 - 人体への影響
 - 精神的負担