

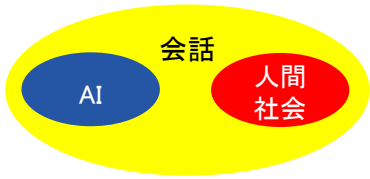
会話からコモングラウンドへ

京都大学情報学研究科
理化学研究所 革新知能統合研究センター

西田豊明

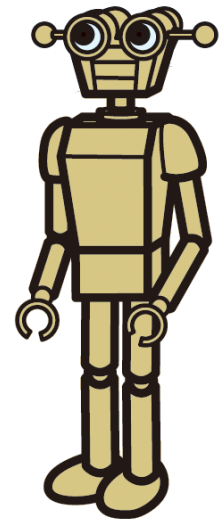
nishida.toyoaki.4e@kyoto-u.ac.jp

会話AI

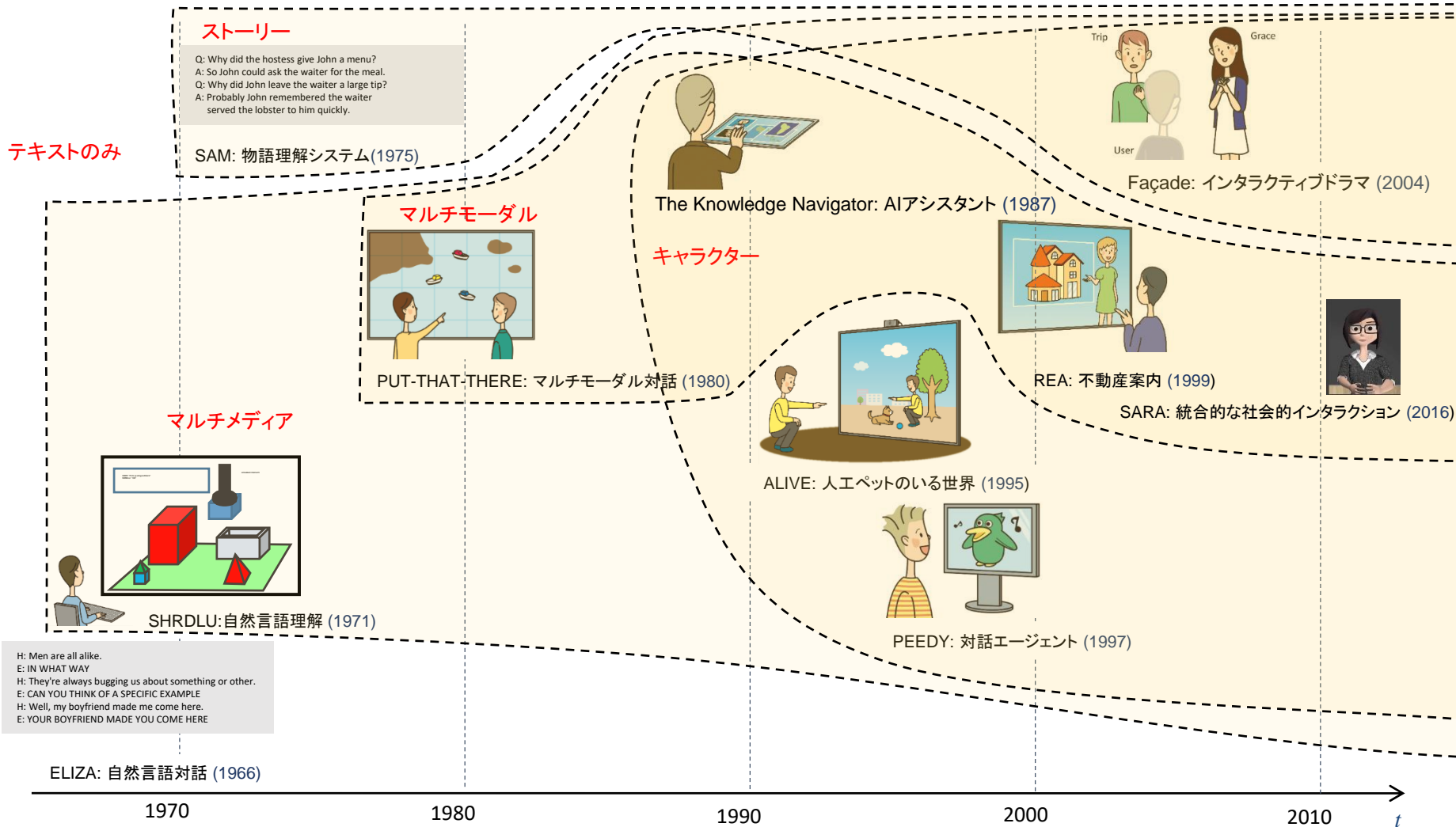


人と会話するエージェント(会話エージェントあるいは会話AI)が広がりつつある

Chatbot, Siri, Cortana, りんな, Pepper, Alexa, Google Home, ...

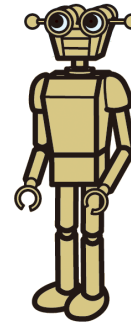


会話システム開発の歴史



私たちの研究グループでも

いろいろなアプローチをとってきた

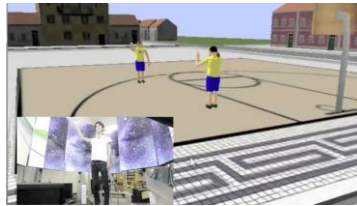


ジェスチャーでやりとり



Vicky (2008)

知的なパス行動



Virtual basketball agent (2013)

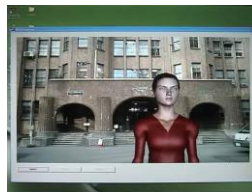
Wizard of Oz(人が遠隔で操る)

キーワードでやりとり



EgoChat (2000)

ストーリーを再生



SPOC (2005)



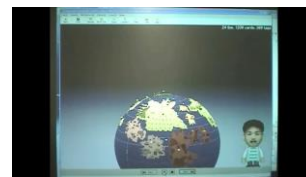
会議に参加するロボット (2013)



CoMeMo Community (1998)



POC TV(2001)



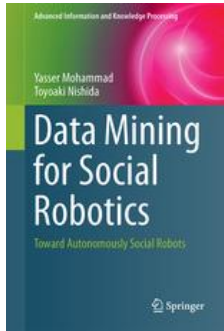
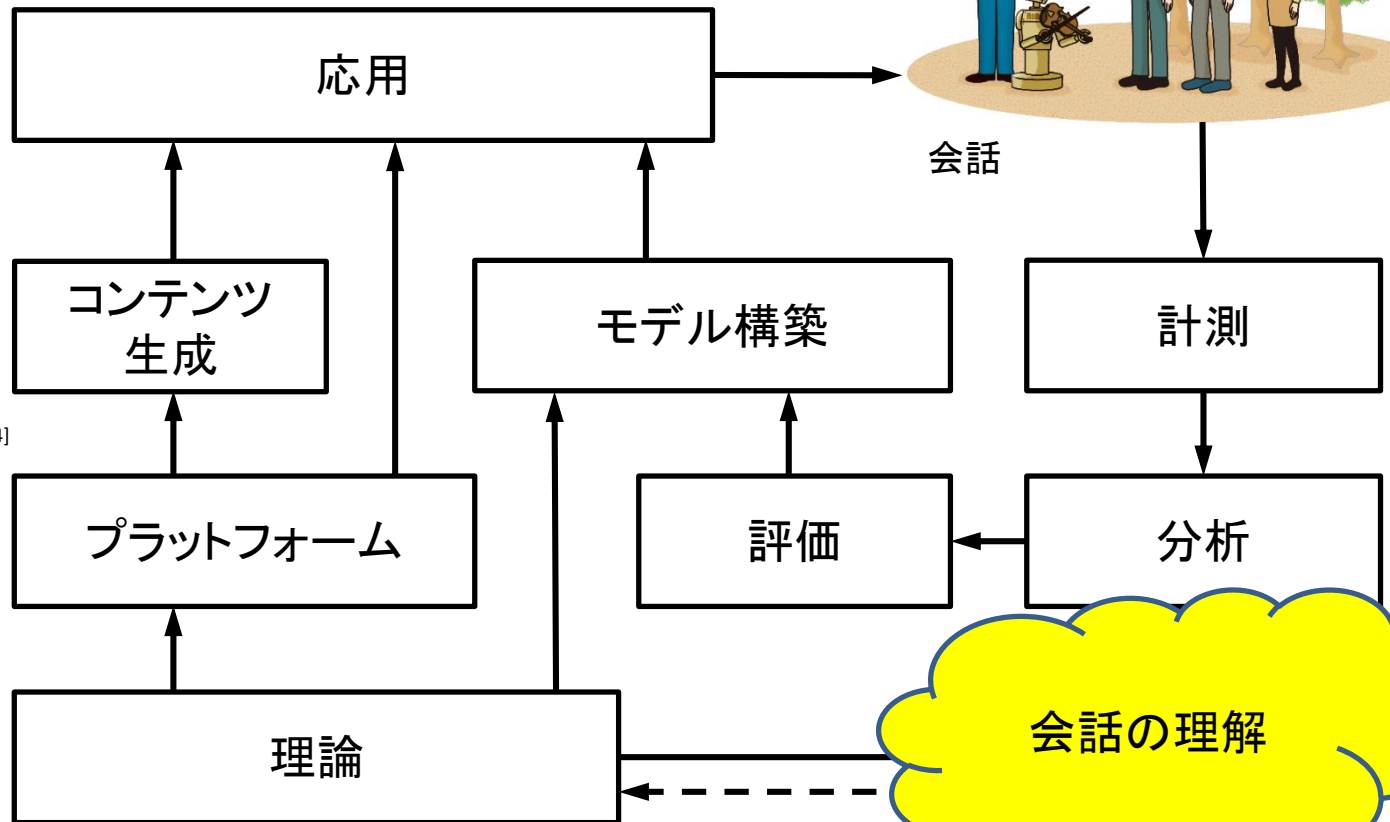
知球(2006)

会話情報学

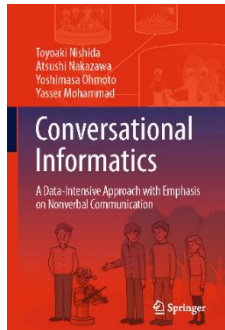
会話システムの構築



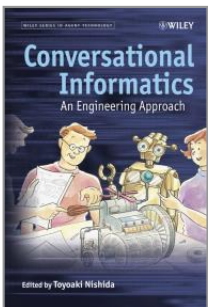
会話



[Yasser, Nishida 2014]



[Nishida, Nakazawa, Ohmoto, Yasser 2014]



[Nishida (ed) 2007]

楽しい会話空間の実現



日常生活空間



MR (VR-AR)による拡張

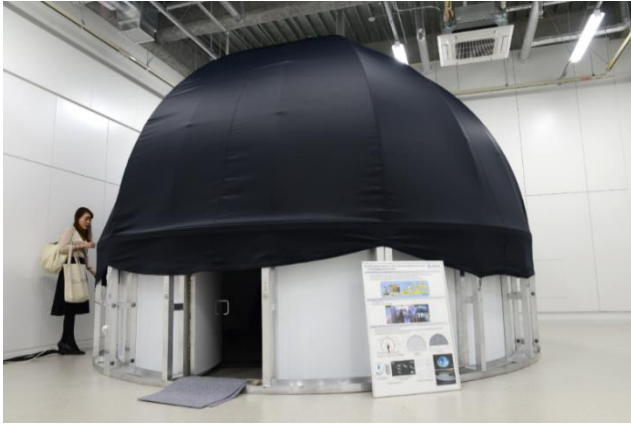
没入型仮想環境ICIE

全周型情報提示十人の動きをセンシングして提示に反映

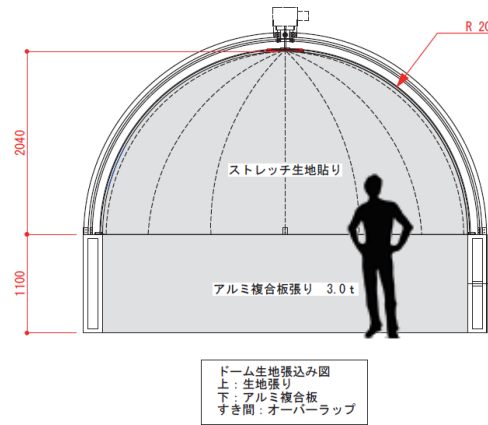


<https://www.youtube.com/watch?v=V-9SKpcMrzk>

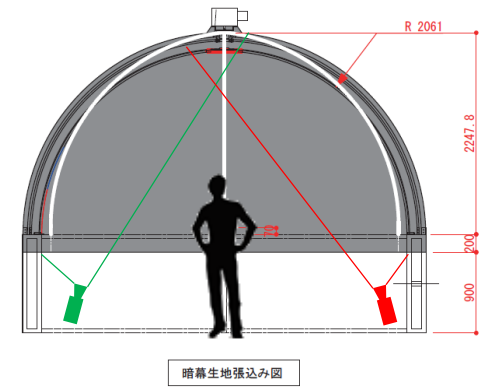
インタラクティブドーム



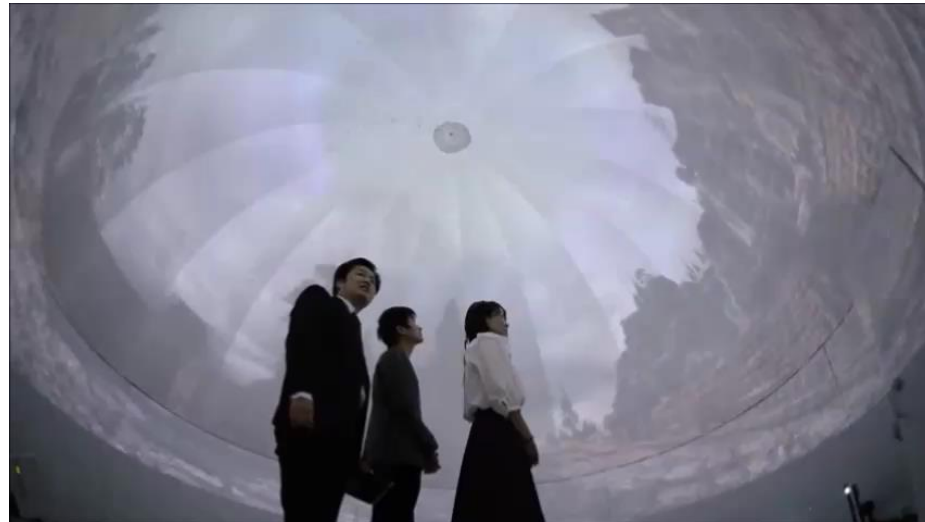
外観



アーキテクチャ

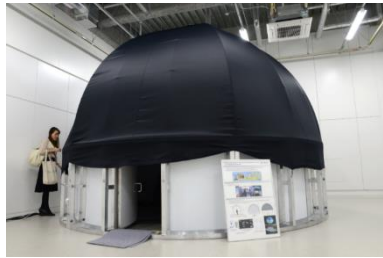


プロジェクション方式



ドーム内部

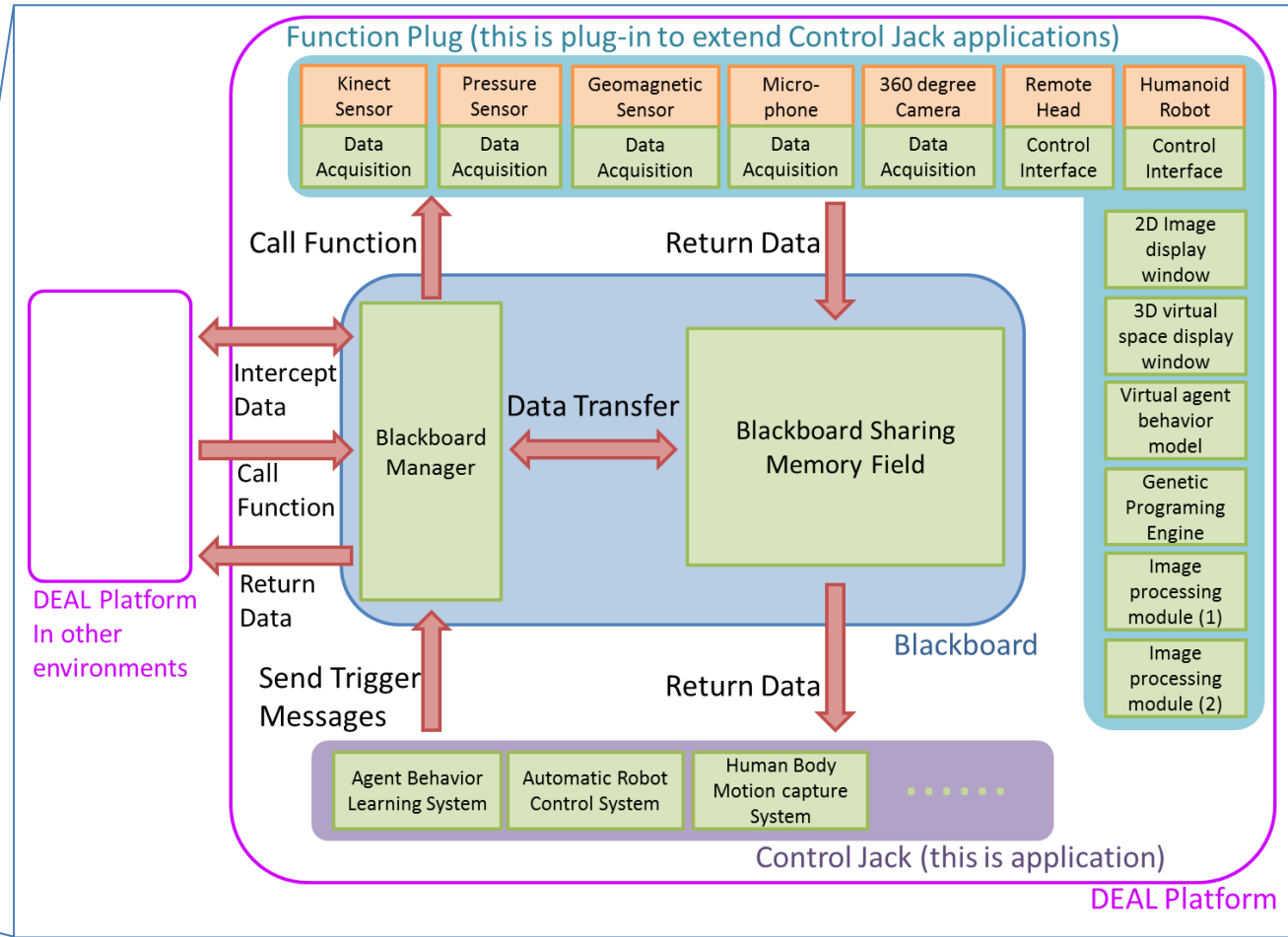
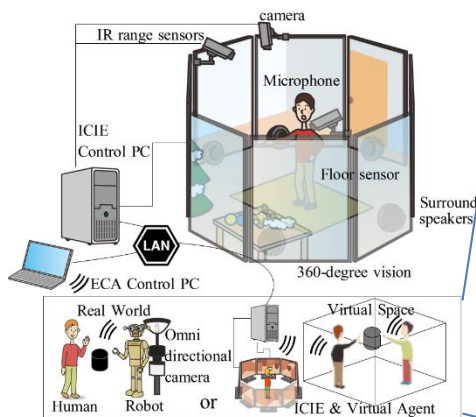
DEAL — 没入型会話環境構築プラットフォーム



インタラクティブドーム

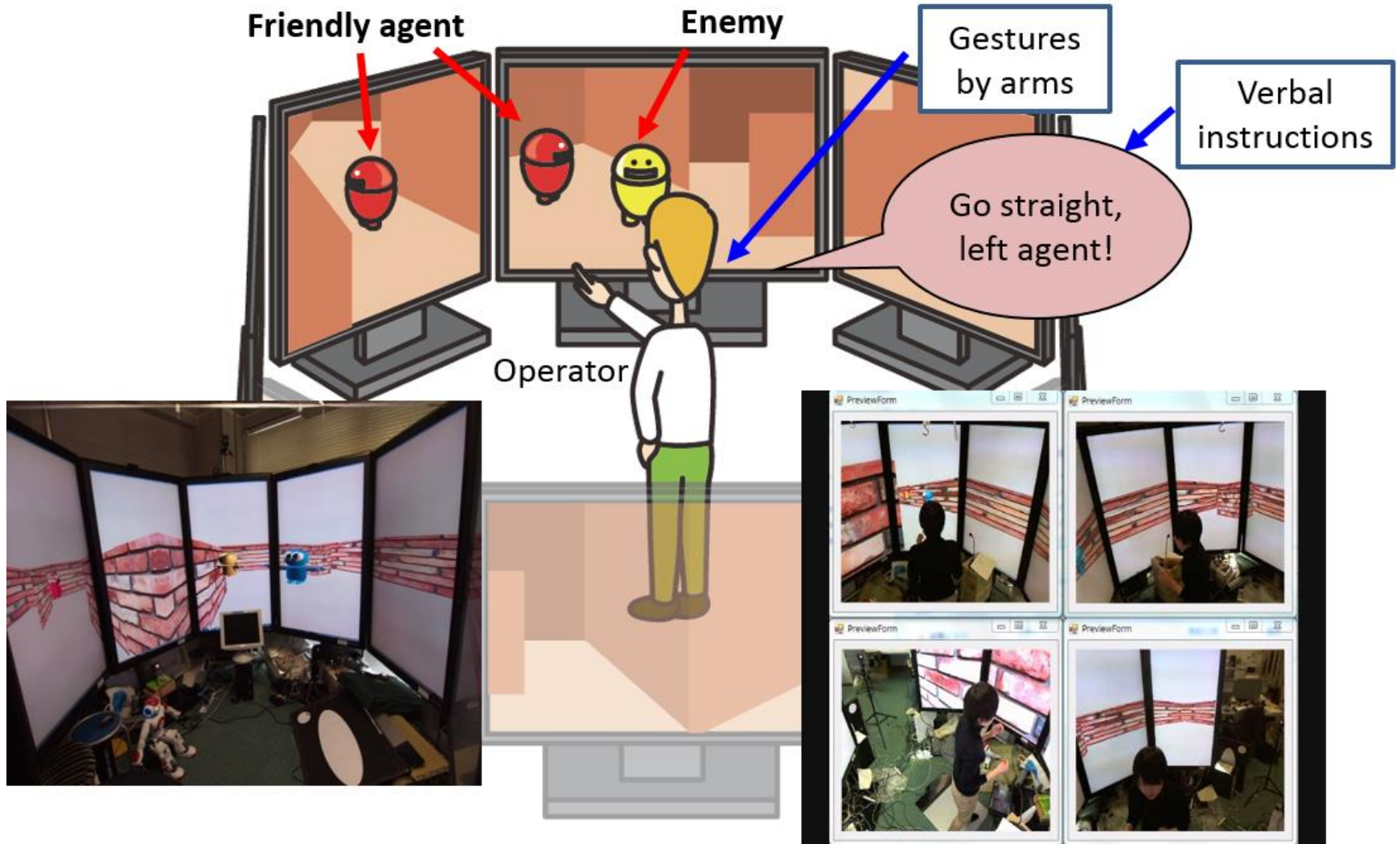


ICIE

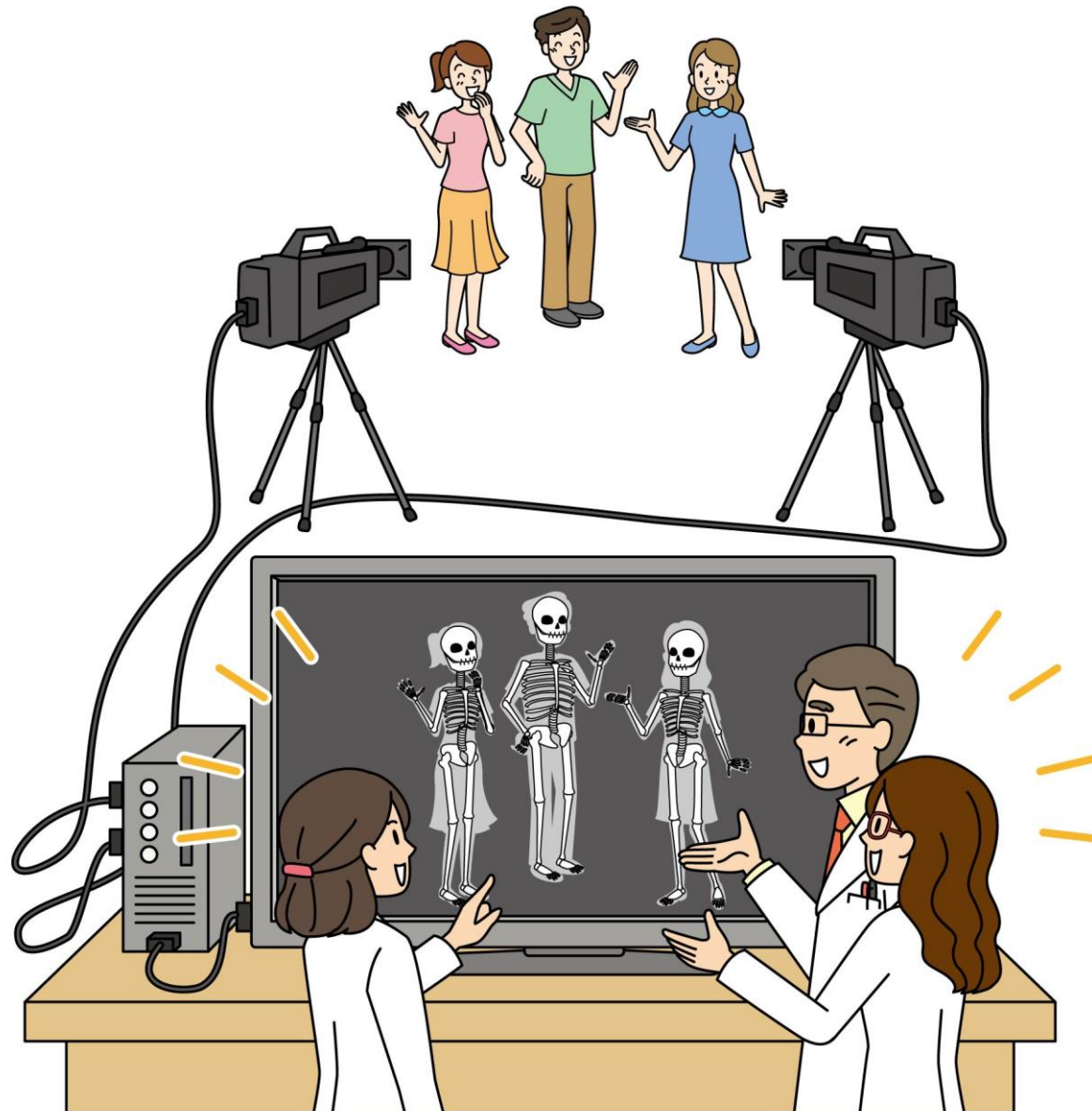


GECAを拡張

没入ゲーム



計測に基づく定量的な理解



会話計測環境



(a) the whole of the group



(b) the right side of the facilitator

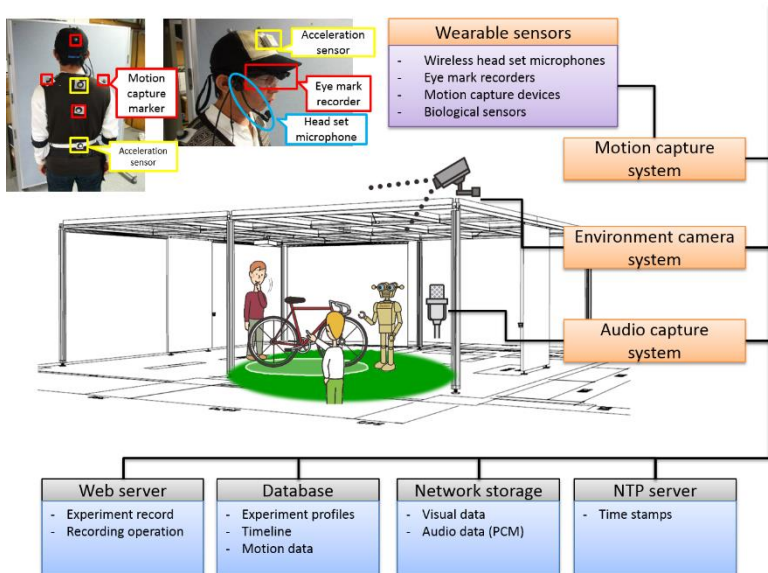


(c) the left side of the facilitator

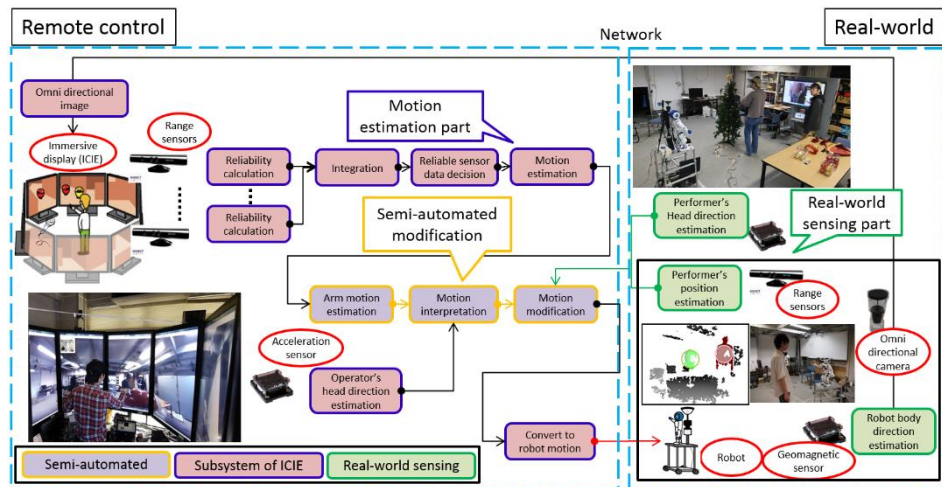


(d) the facilitator face

固定された配置での会話計測



参加者が自由に動き回れる状況での会話計測

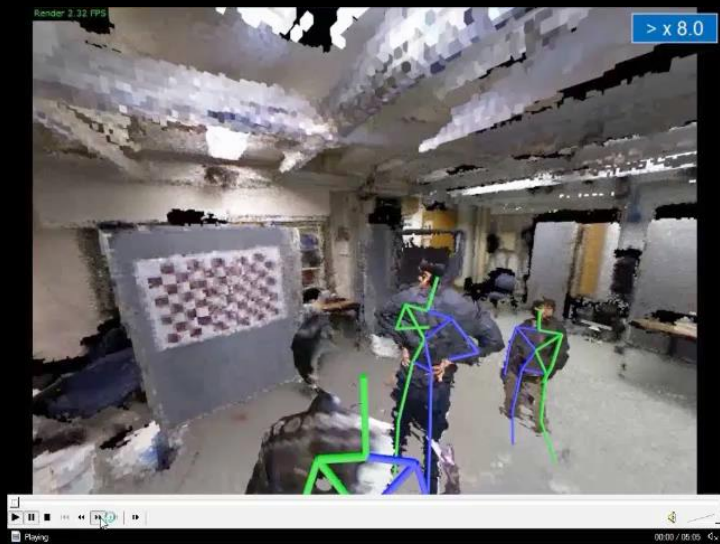


混合空間での会話計測

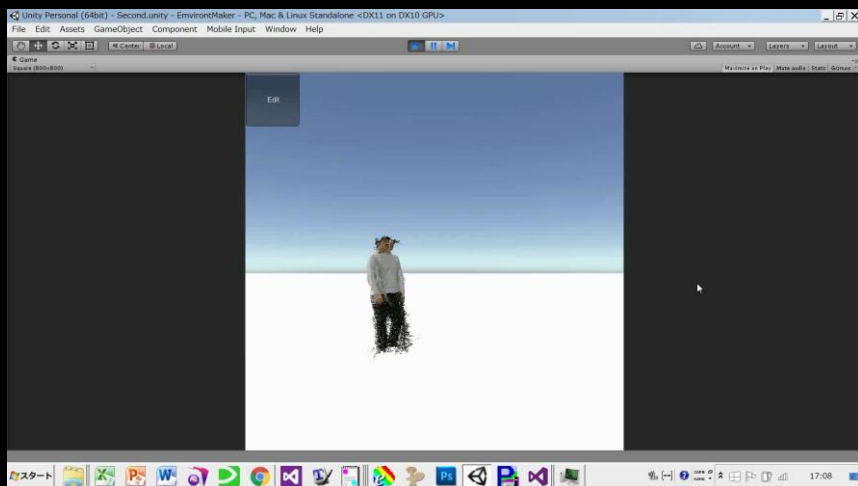


オープンスペースでの計測とキャプチャ

会話情景のキャプチャ

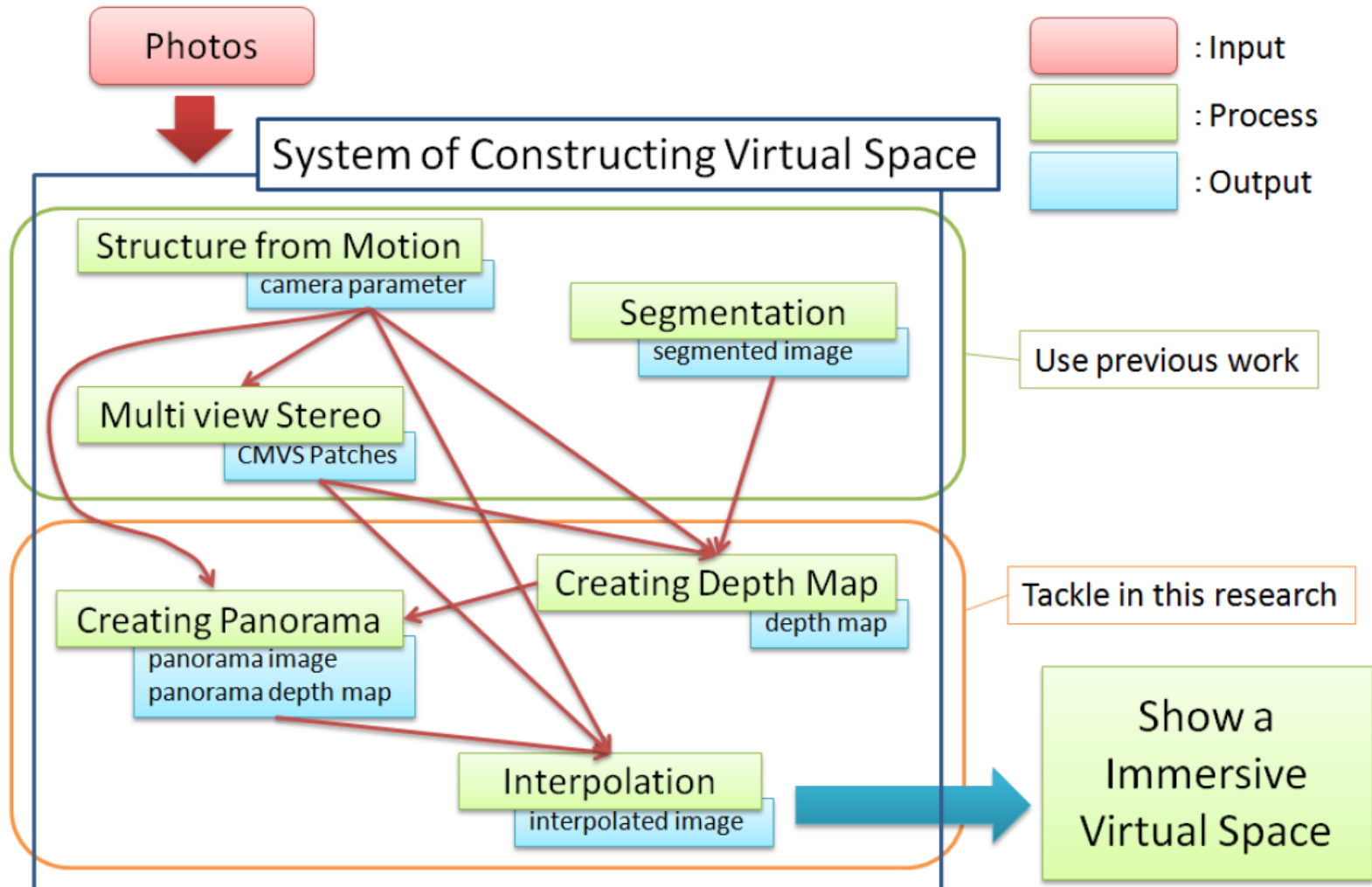


<https://www.youtube.com/watch?v=J08vG8wnrnw>



[Yano 2012; Higuchi 2016, 2018, 2018]

背景のキャプチャ



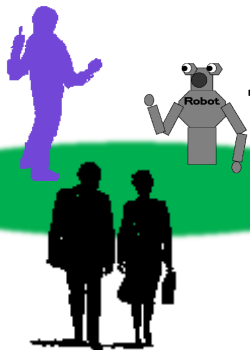
再構成された実世界

x2

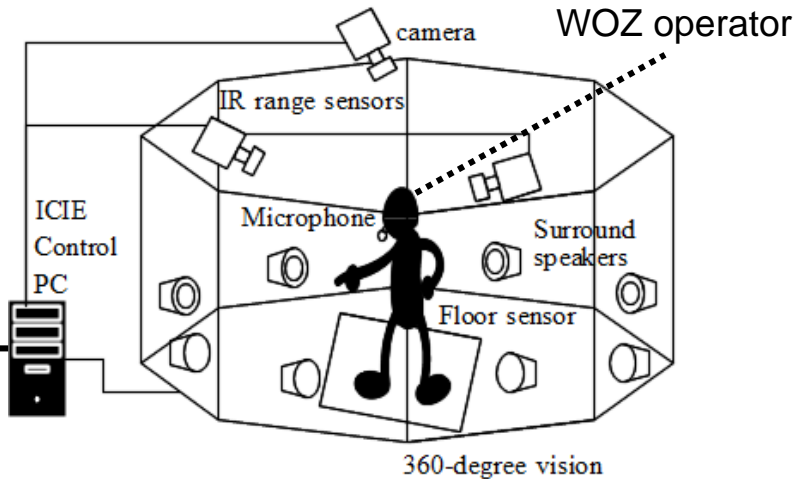


没入型WOZ環境

Tele-operated robot

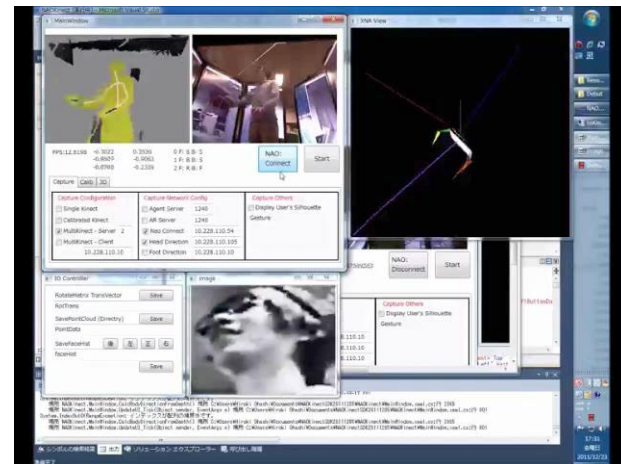


Feedback generation
Motion mapping
User motion sensing
Head recognition
Gesture recognition
Face model
Human body model



The conversation place

WOZ operating environment



[Ohmoto, Ohashi, Saiga]

人の行動のモニタリングー角膜イメージング法

角膜イメージ特徴マッチング

Problem:

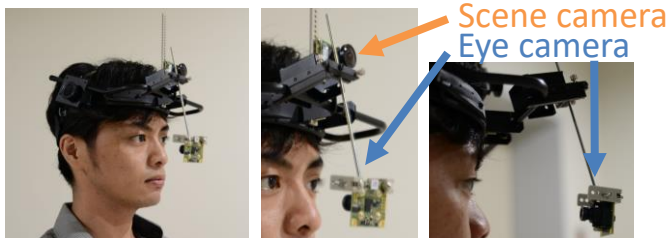
Local feature correspondence + RANSAC does not work due to *large noise in eye images*

Approach:

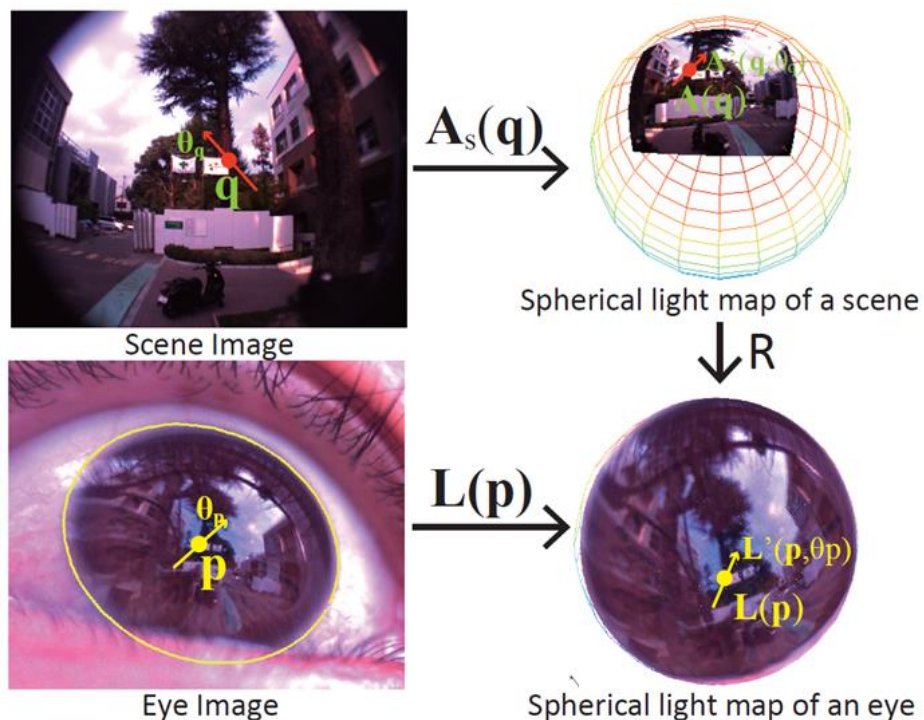
1. Formulate problem as *registration of 3D spherical light maps* of eye and scene image
2. *Single point algorithm* for robust alignment



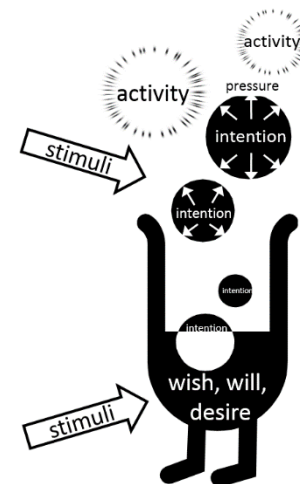
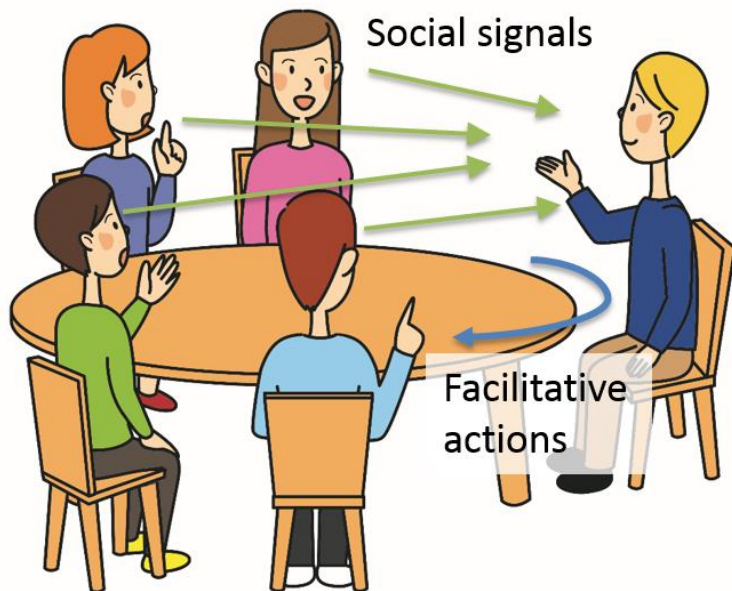
角膜イメージングカメラ



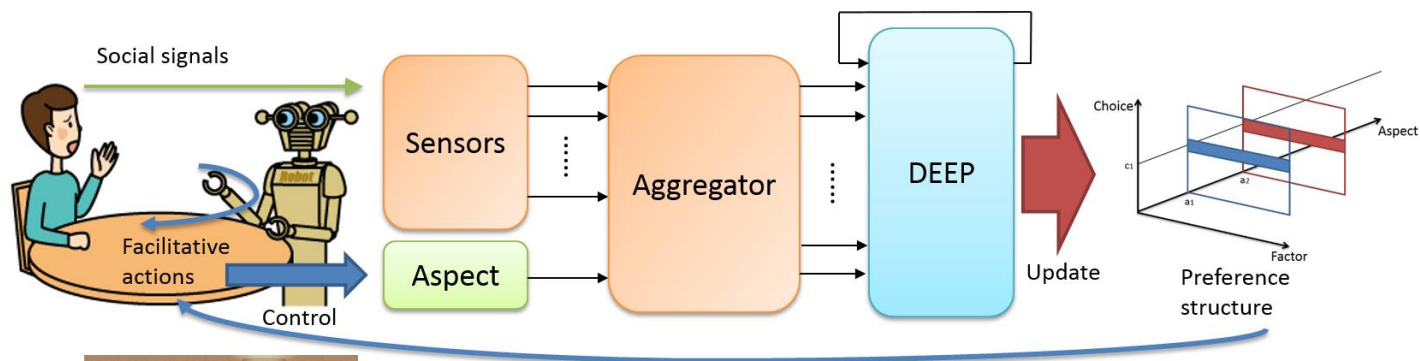
- Lightweight and versatile system
- Appl.: Google Glass like HMD, unconstrained setups



暗黙的な社会シグナル

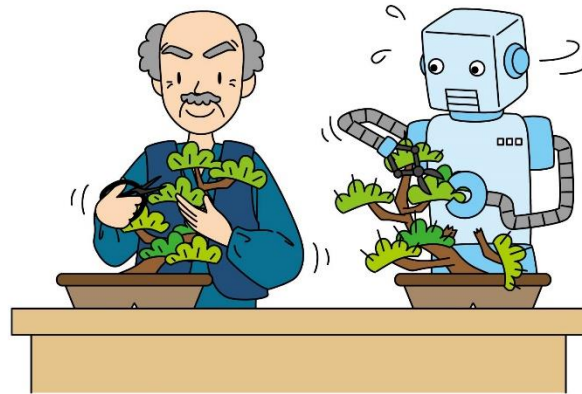


Intentions are dynamic and tacit

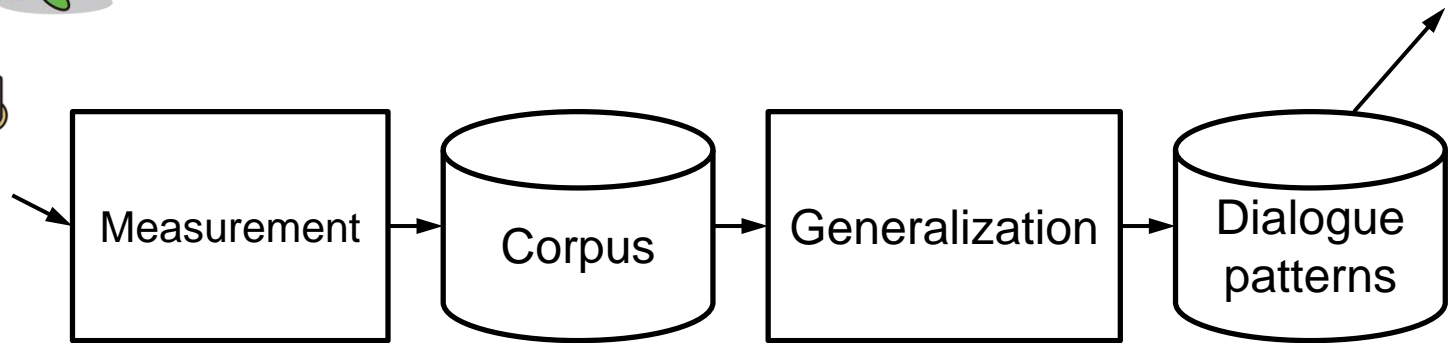
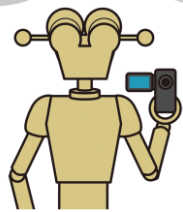
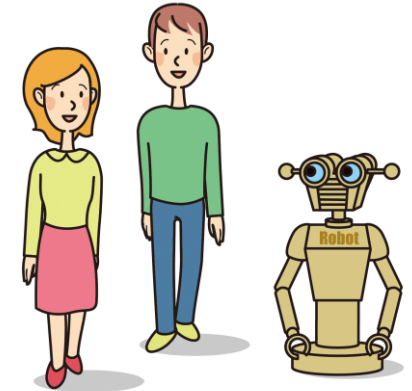


Facilitative agent that can track dynamic and tacit intentions

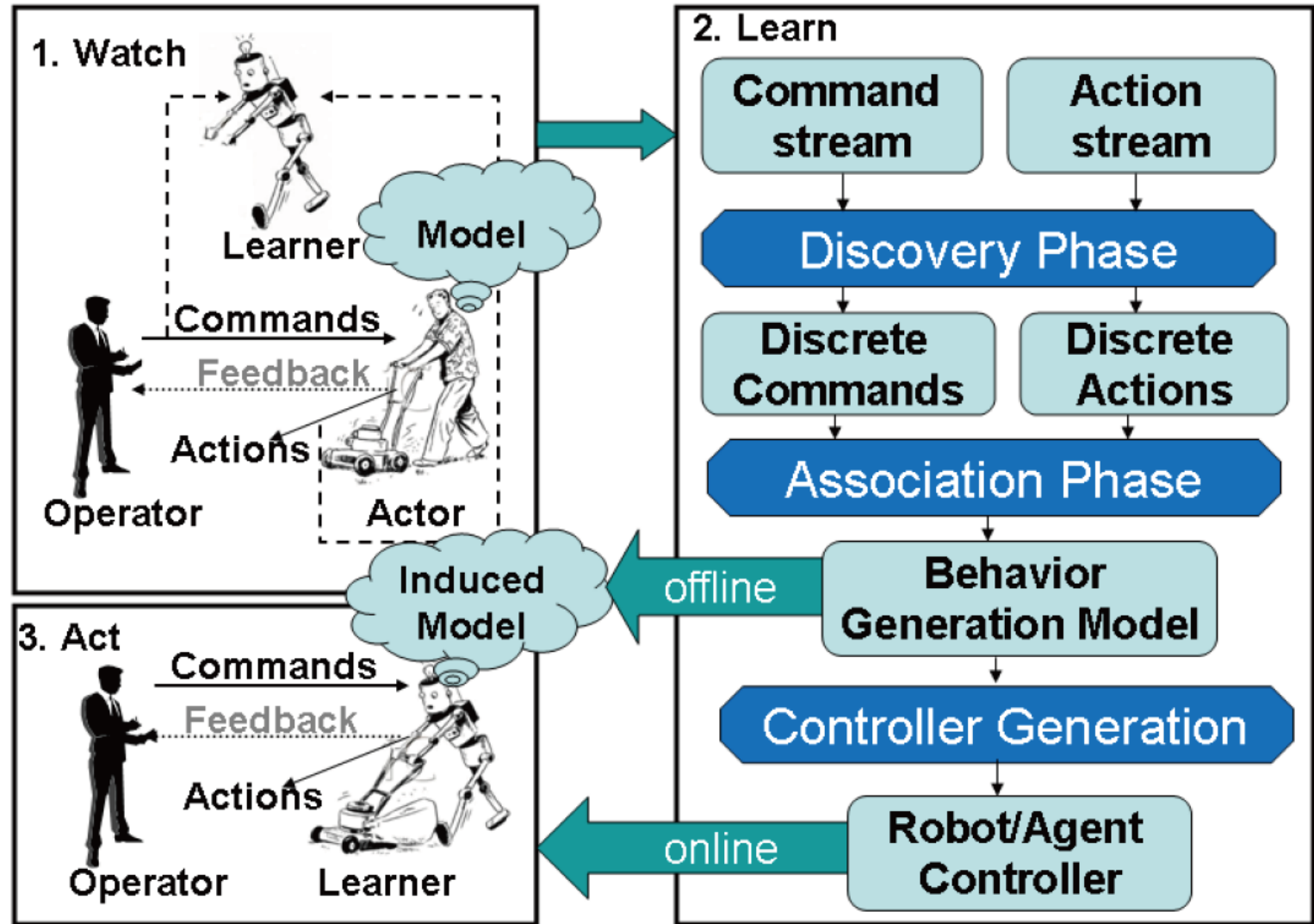
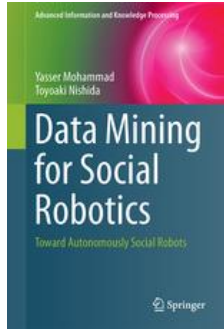
模倣学習：データからの会話行動の生成



Endow robots with an ability of autonomously imitating human behaviors.

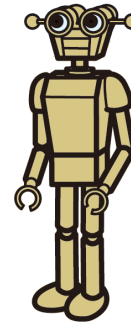


EICAにおける模倣学習



私たちの研究グループでも

いろいろなアプローチをとってきた が限定的



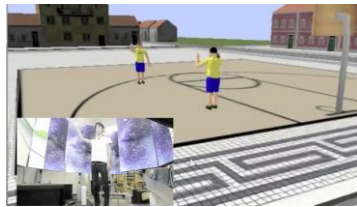
存在感が弱い
すぐ飽きる

ジェスチャーでやりとり



Vicky (2008)

知的なパス行動



Virtual basketball agent (2013)

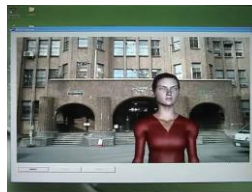
Wizard of Oz(人が遠隔で操る)

キーワードでやりとり



EgoChat (2000)

ストーリーを再生



SPOC (2005)



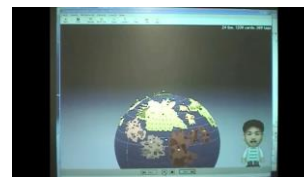
会議に参加するロボット (2013)



CoMeMo Community (1998)



POC TV(2001)



知球(2006)

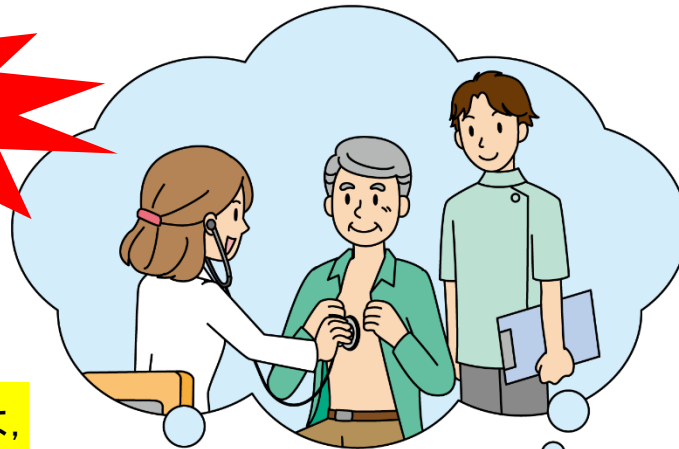
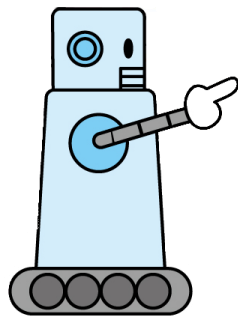
ボトルネックは、人間とAIの間のコモングラウンドの欠如

会話をしている人たちが共有していると信じていることから

チャレンジング

会話をしている当人たちも、傍観者も相互理解できるのは、コモングラウンドのおかげ

AIが会話に参加するためには、コモングラウンドを共有し、発展させるプロセスに参加できなければならない。

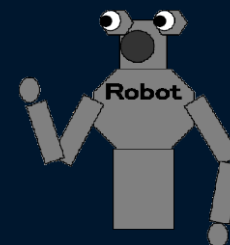


コモングラウンドとは

コモングラウンドを構成するもの



さらに，行動主体の
計算基盤…



	ミクロ	メゾ	マクロ
時間範囲	○秒～○分	○時間～○日	○か月から○年
集団	2人	小グループ	コミュニティ
活動	会話	問題解決	協働，共生
タイプ	認知，言語	概念	文脈
共有対象	参照物と情報	規則，概念	文化

「インタラクションレベル」

「ストーリーレベル」

「背景知識レベル」

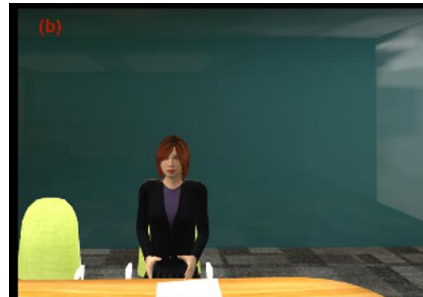
コモングラウンドの恒常的更新としての会話

3人称視点



自分たちはどのような状況の中に置かれているのか？これまで何をしてきたのか？これから何ができるのか？

1人称視点



この人と自分にとってどのような人か？これまでどんないきさつがあったのか？これからどのようなことが起きそうか？

内面



この人は一体何を考えているのか？この状況をどう思っているのか？これから何をしたいのか？

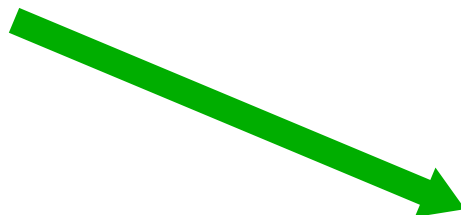
会話の進行



説明可能な会話システム開発には膨大な知見の統合が必要



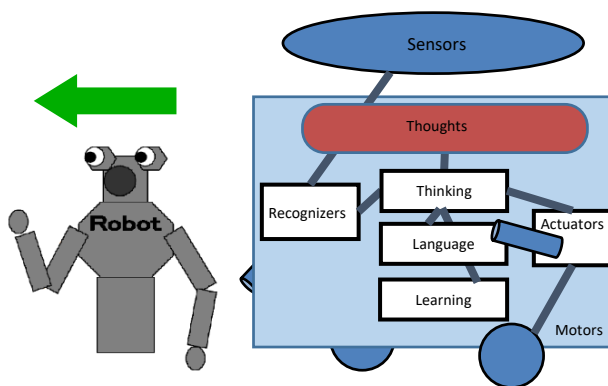
コモングラウンドを会話システムに取り入れるためには膨大な学術的な知見を取り入れる必要.



Reality



[A marketplace in Chiang Mai; courtesy of Sutasinee Thovuttikul]



会話とコモングラウンドに関わる文献の広がり

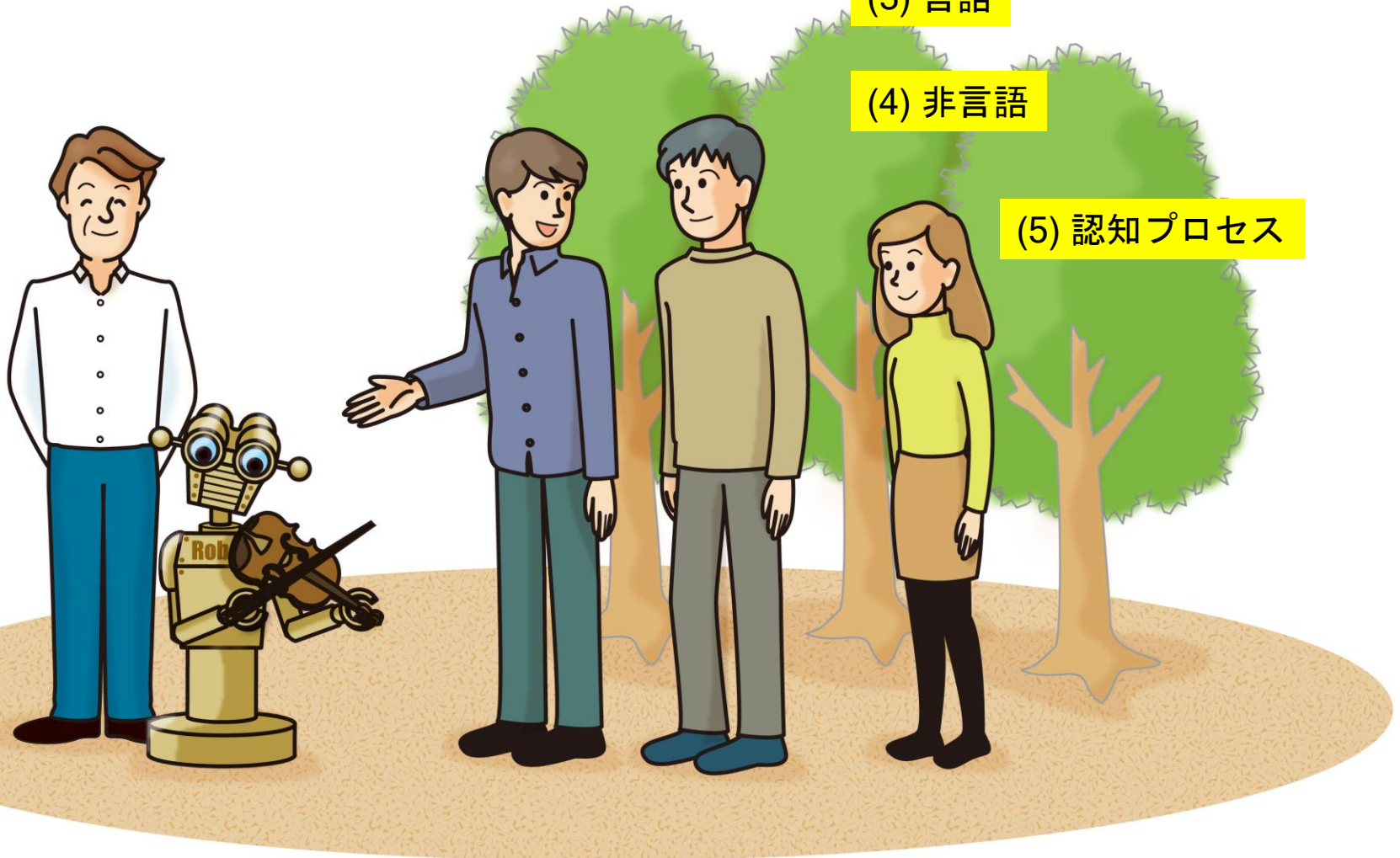
(1) ストーリー

(2) 社会的インタラクション

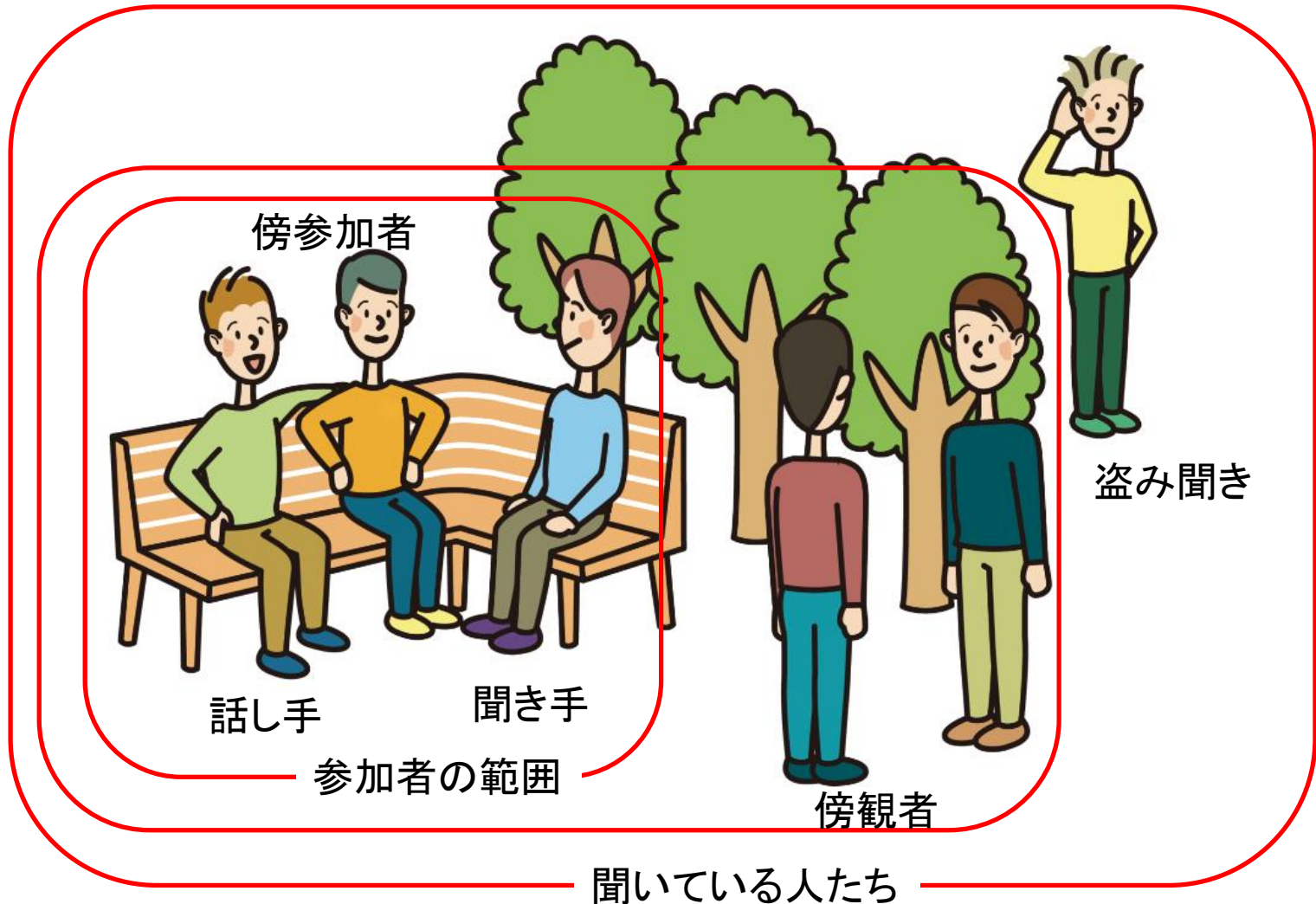
(3) 言語

(4) 非言語

(5) 認知プロセス

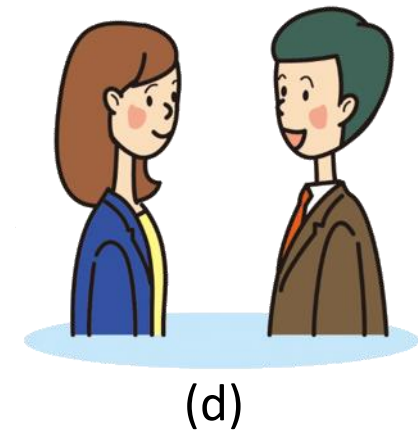
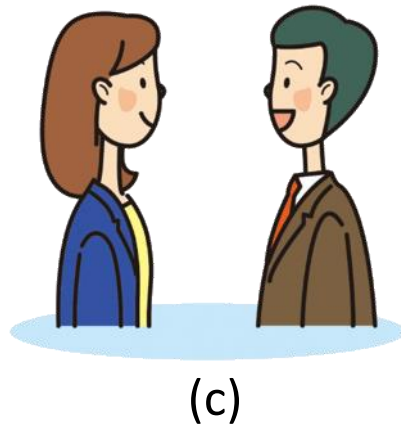
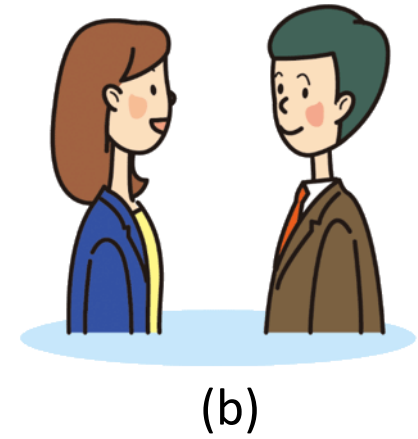
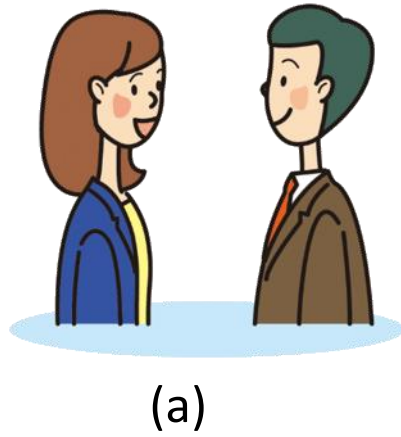


社会的インタラクション: 会話への参加の構造



非言語的なインタラクション

ターン交代



非言語・言語インタラクションの統合

A+B

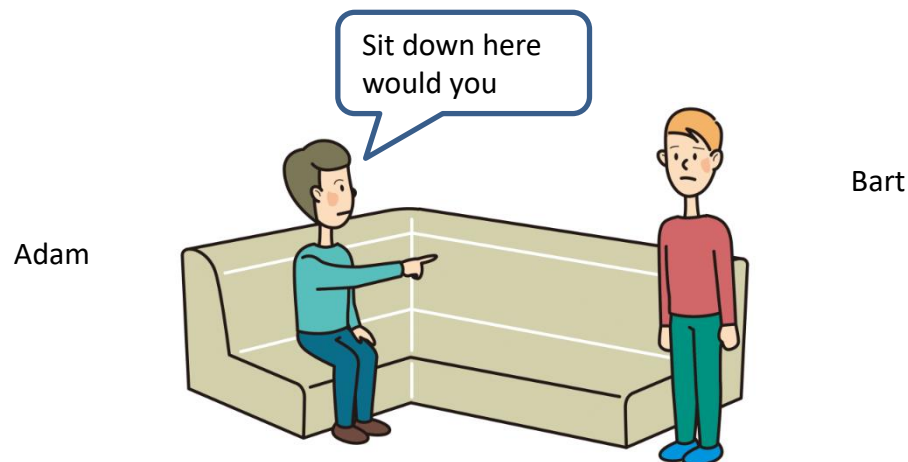
Level	A's actions in progress	B's actions in progress
4	Proposal and consideration	
3	Signaling and recognition / meaning and understanding	
2	Presentation and identification	
1	Execution and attention	

A

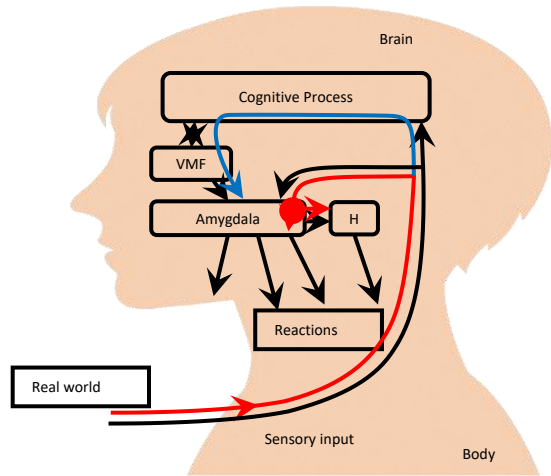
Level	A's actions in progress
4	A is proposing to B
3	A is asking B
2	A is presenting to B a signal
1	A is executing for B's perception

B

Level	B's actions in progress
4	B is considering A's proposal
3	B is recognizing A's request
2	B is identifying A's signal
1	B is attending to A's articulation



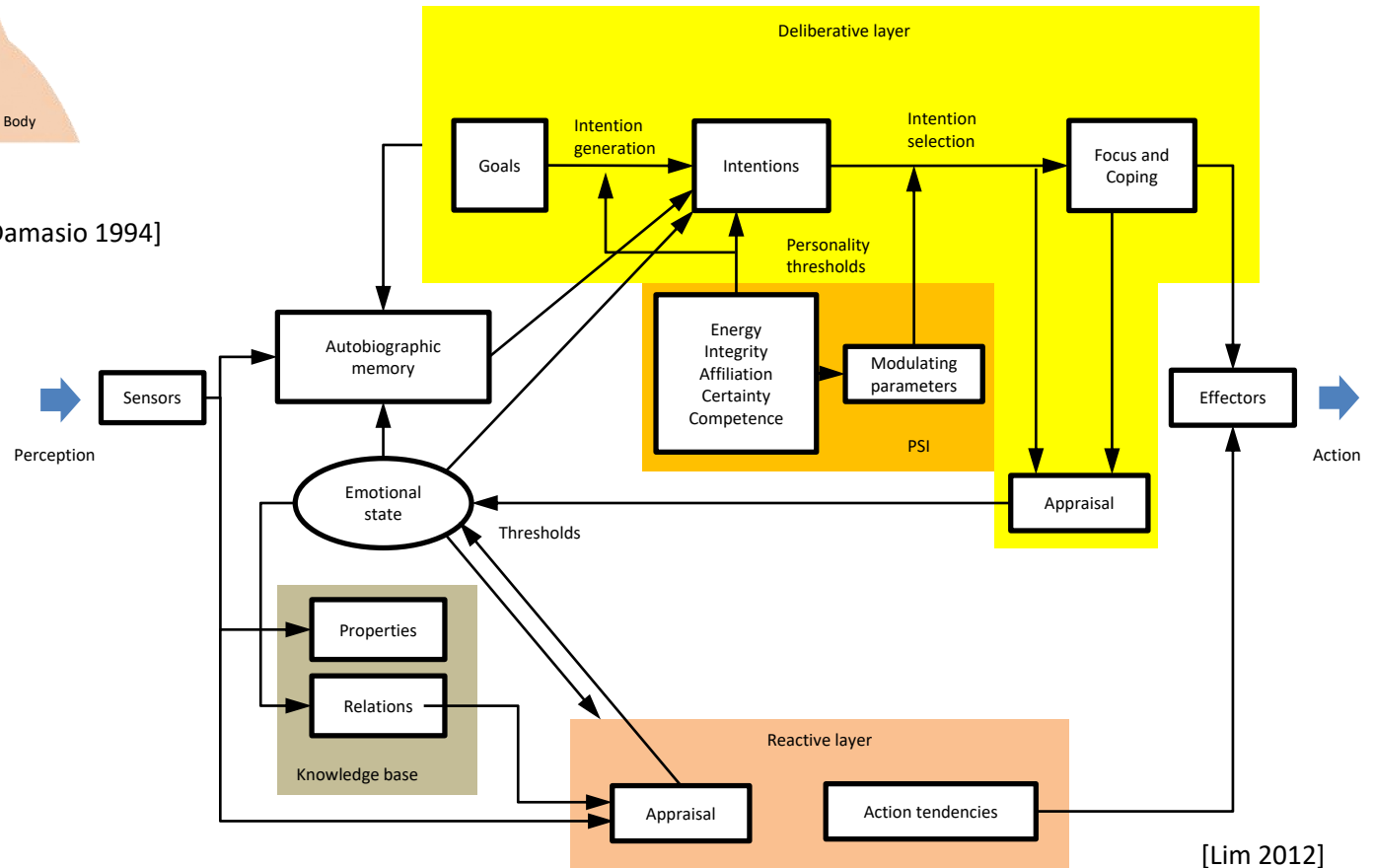
感情のモデル



H: Hypothalamus

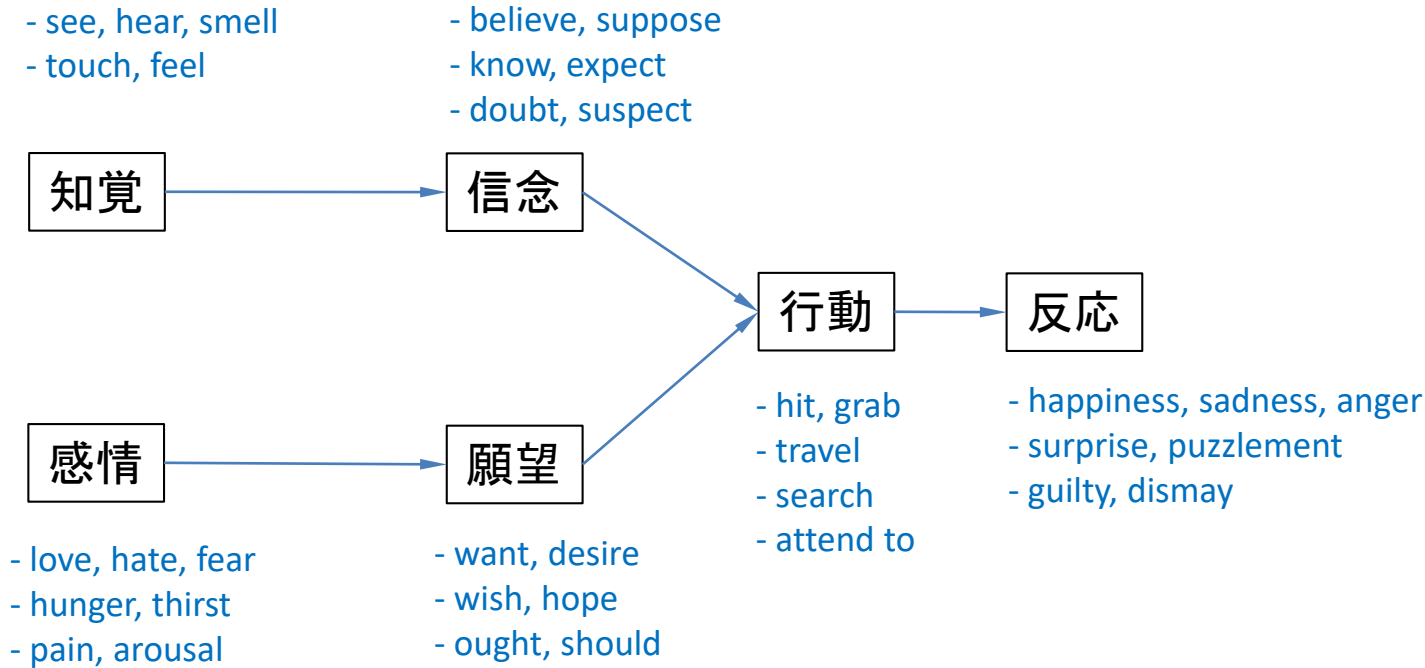
VMF: ventromedial prefrontal cortex

[Damasio 1994]



[Lim 2012]

心の理論—社会的能力



Theory-of-Mind Scale:

Diverse Desire (DD): 願望は人により異なり得る

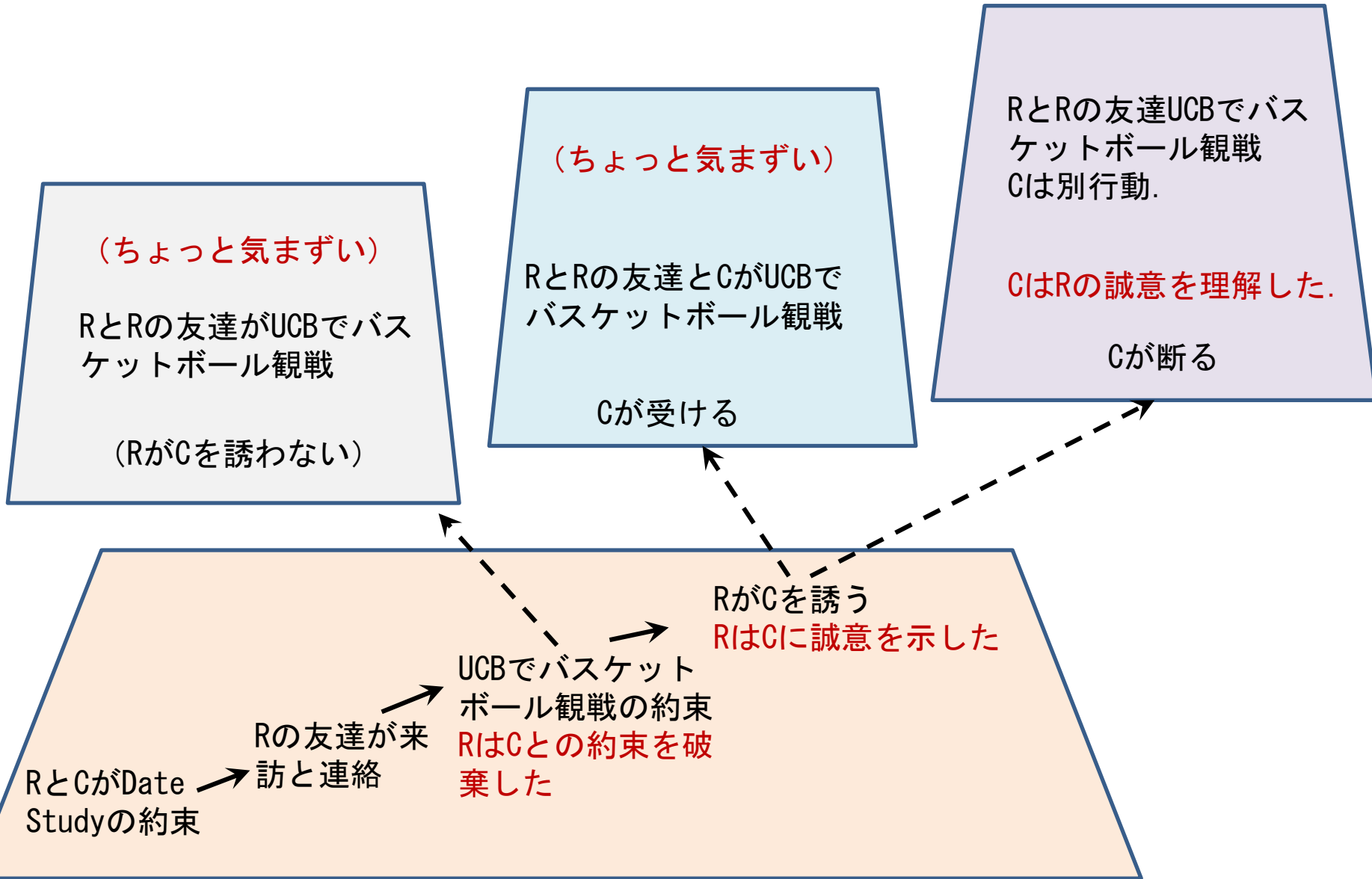
Diverse Belief (DB): 同じ状況でもその解釈は人により異なり得る

Knowledge Access (KA): 状況により知覚できるものが異なる

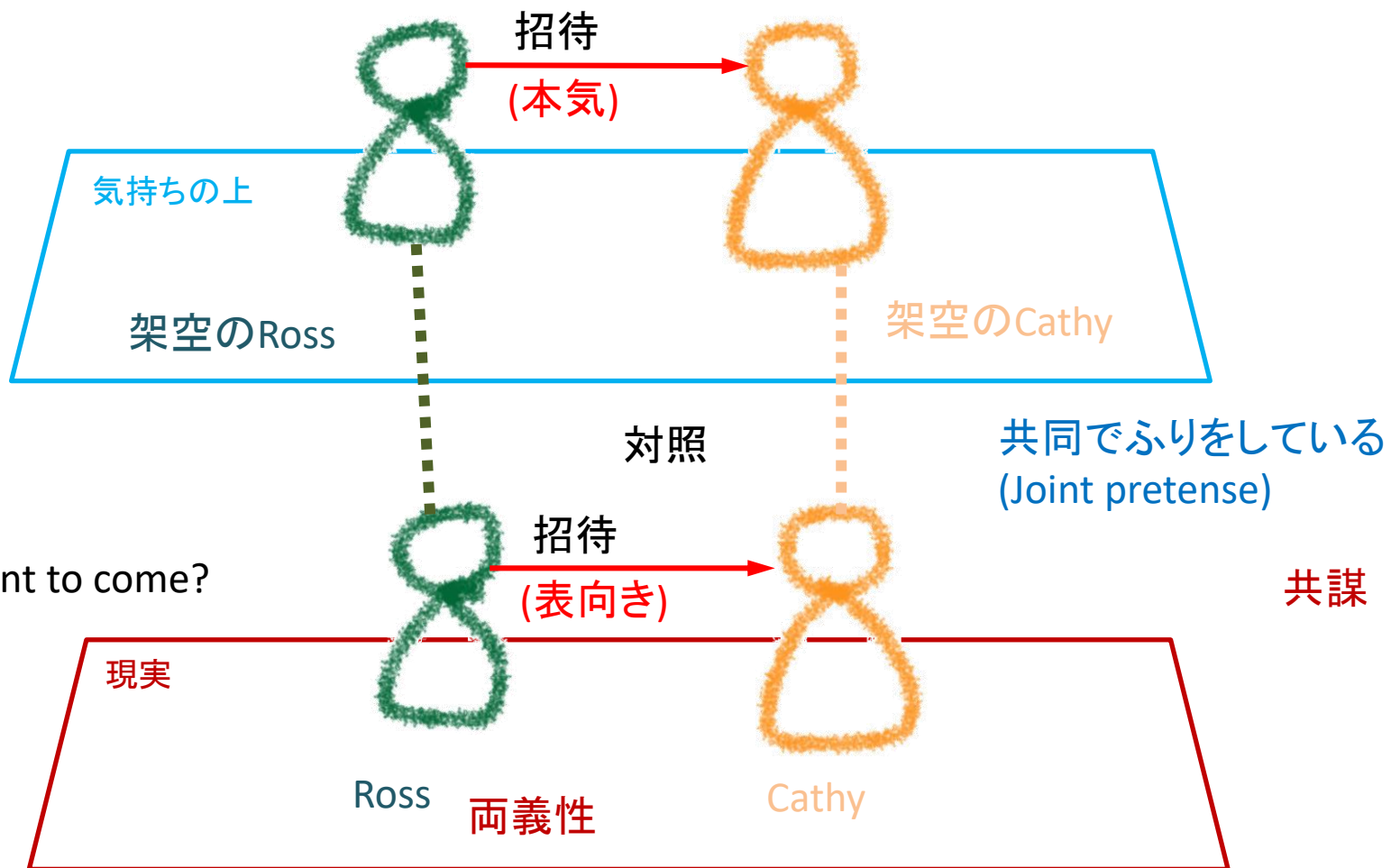
False Belief (FB): 人は誤解するかもしれない

Hidden Emotion (HE): 人は感情を隠すことがある

高度な言語使用—表向きの招待



高度な言語使用—表向きの招待



ペルシャ式価格交渉



Ostensible offer

- C: Hi. How are you?
S: Hi. Welcome to this shop.
C: Thank you. How much are these handicrafts here?
S: Oh, these are called Khaatam .The big ones are 120, [middle sized ones for] 80 tomans and [small ones are] 50 tomans.
C: Um, this is Termeh, right?
S: Yes!
C: What about the Termeh?
S: Actually we have different types of Termeh. We have the red ones there starts from 50 and the brown ones start from 150 and these special ones starts from 300.
C: You know what? I am about to marry, and I am looking for something very nice to decorate my table. Its color is brown. So, what do you suggest?
S: Oh! Congratulations for your marriage!
C: Thank you.
S: We usually go for the red ones for brown tables, but because you are a great customer and you have good taste, probably you would go for this one, this is very beautiful and it is from a very famous designer.

C: Oh, OK. How much is this one?

S: This one? It worth nothing, you can have it [for free].

C: Haha, well, please...

S: The price is 500 tomans.

C: Oh! 500 tomans? It is very expensive? Do you have anything cheaper?

S: This is from a very good brand, we have the no-name ones, you can see up there, and they are cheaper.

C: But the colors of those do not match my table. I would love to buy this if you give me a very good discount. You know, I am about to marry, marriage itself costs a lot, so it would be nice if you would do that.

S: It is your marriage, you should treat us! ... Because it is a happy event I can give it for 450.

C: 450 is still a little bit expensive. Why not we make a deal for 400?

S: 400 is too low, but since it is my first sell of the day, and you are getting married, and everything looks very great I can give it for 430, but you know, I haven't sold it less than 450 so far.

C: You have been so nice so far, why not getting rid of that 30 and make it round for 400. I am sure you are making lots of purchases today with other customers and you can compensate for that. And I'm going to be a frequent customer from now on, I will come to your shop again.

S: We would love to have you again, but the price is somehow at its margin. But, I can guarantee it for one year for you, and I can give it to you for 420, final! This is very good brand, you would never regret it, you would love it, you can have my word that it works for you perfectly and for one year if anything happens to that you can bring it to us and we can fix it for you for free. Ok, I'll write 420 for you.

C: Ok, thank you, you did me a good favor. Thanks a lot.

S: Oh please. We enjoy you being here and please come back soon.

C: Yeah I will. I'm actually going to do more purchases from now on from your shop. I had a good experience.

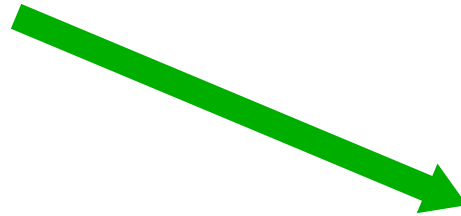
S: Sure! Have a good time.

説明可能な会話システム開発には膨大な知見の統合が必要

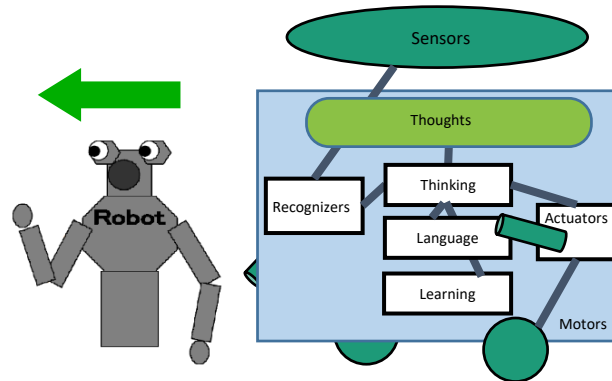


あい路！

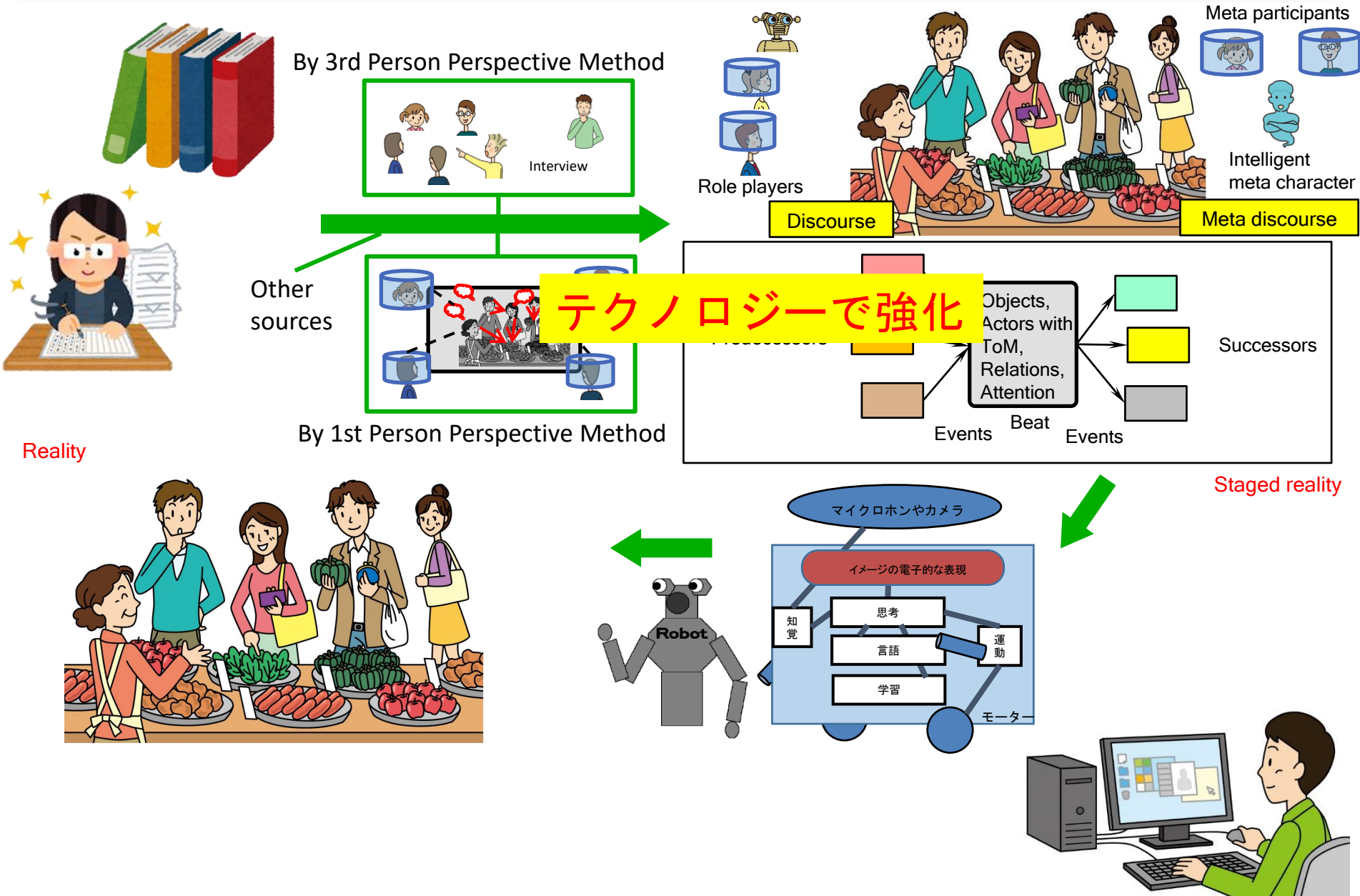
コモングラウンドを会話システムに取り入れるためには膨大な学術的な知見を取り入れる必要.



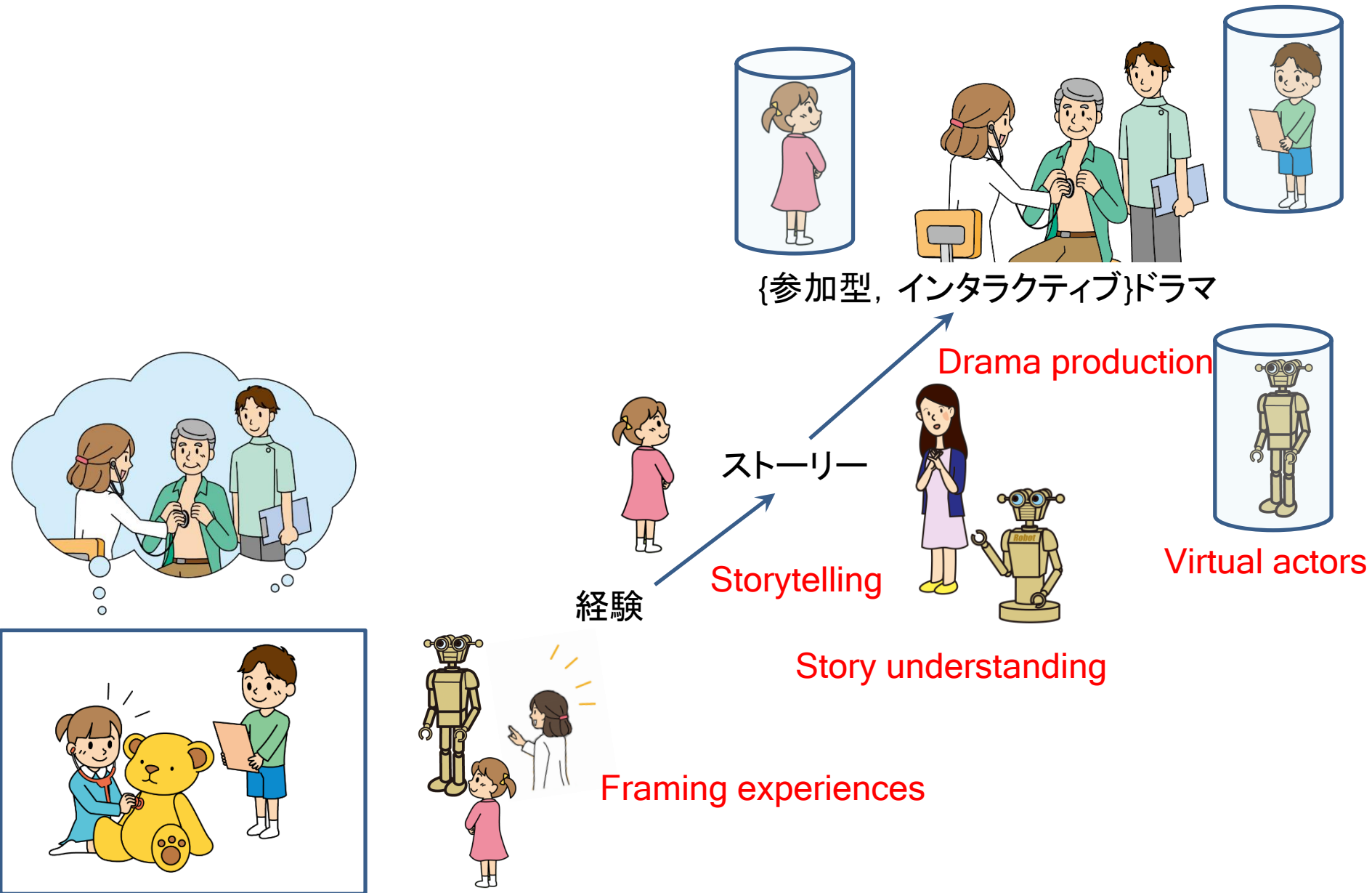
Reality



会話情報学の研究プログラム



インタラクティブドラマ制作支援による構成的手法

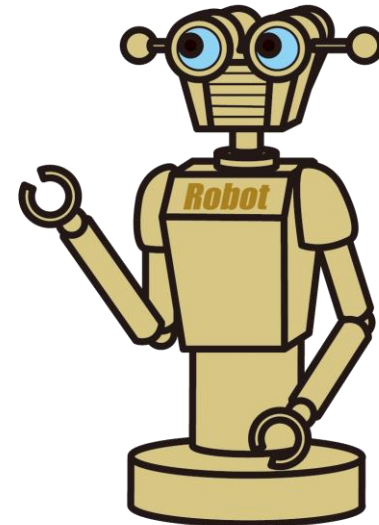


会話のエンビジョニング—構成的考証法

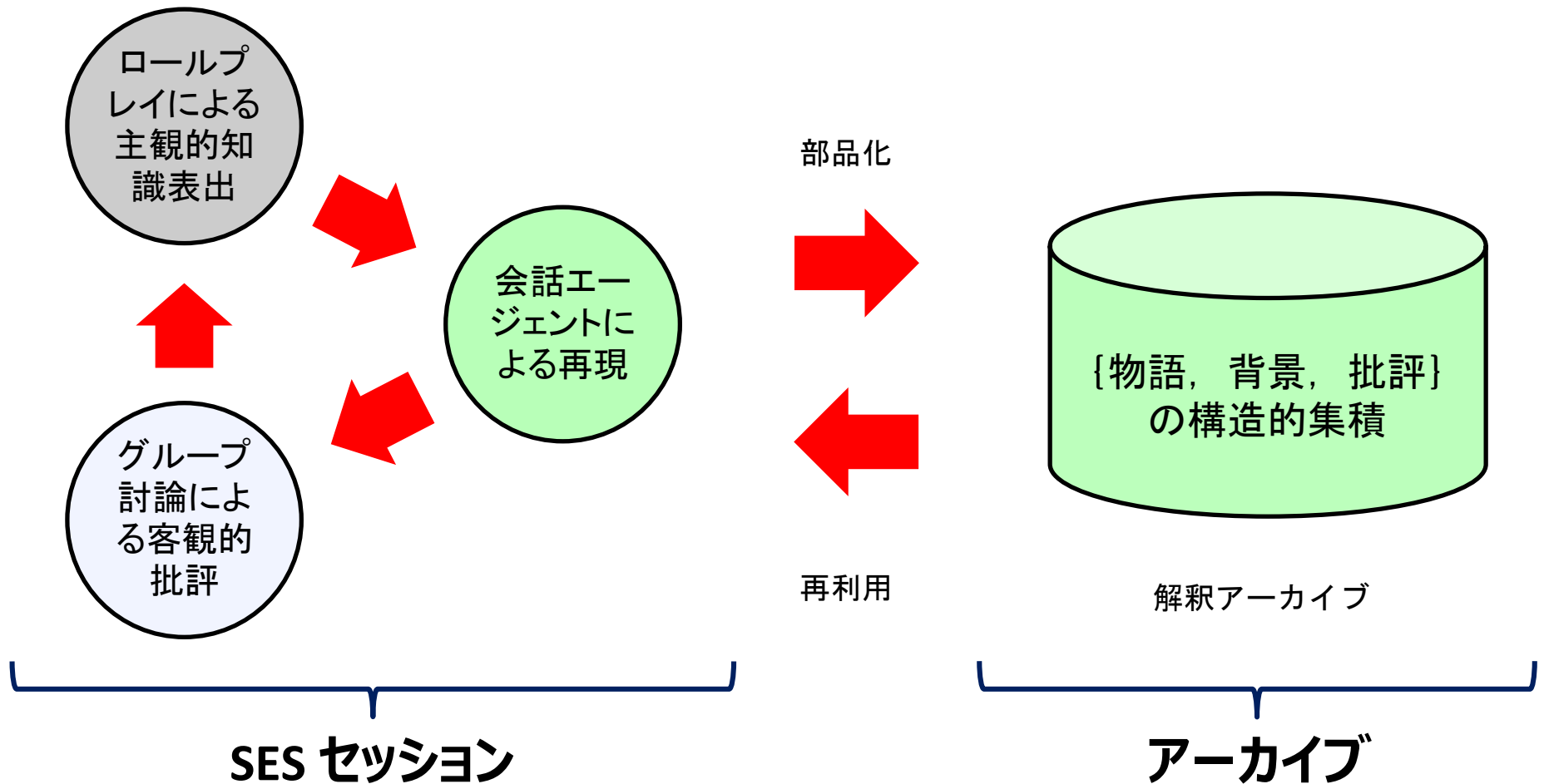
- ストーリーライティング：コストがかかる
- ユーザはストーリーを作れない



構成的考証法：ユーザが全身を使って自分のストーリーを伝える



会話のエンビジョニング—構成的考証法



構成的考証法 — 適用例

例えば、忠臣蔵松の大廊下シーン

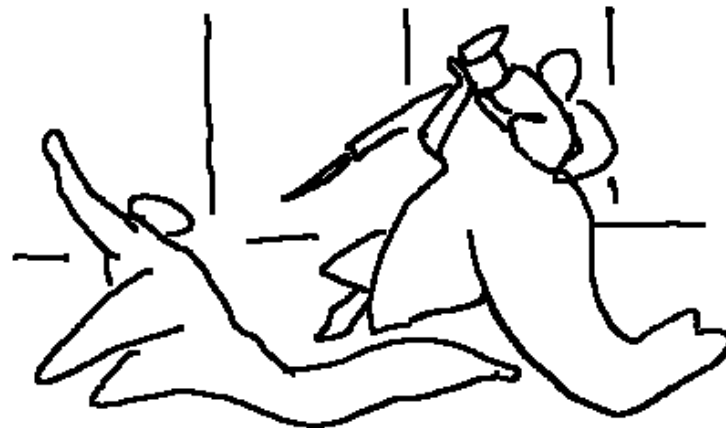
1701年3月14日(4月21日)巳の下一刻(午前11時半過ぎ)、江戸城本丸御殿の大広間から白書院へとつながる松之大廊下で、浅野長矩は「此間の遺恨、覚えたるか」と言いながら背後から吉良義央を小刀で斬りつけた。吉良が振り返ったので小刀は吉良の眉の上を傷つけた。そして吉良が向きかえって逃げるところを追いかけ、また2度斬りつけた。すぐさま、浅野はその場に居合わせた梶川頼照らに取り押さえられ、...

[Wikipedia]

浅野

梶川

吉良



構成的考証法 — 演劇フェーズ



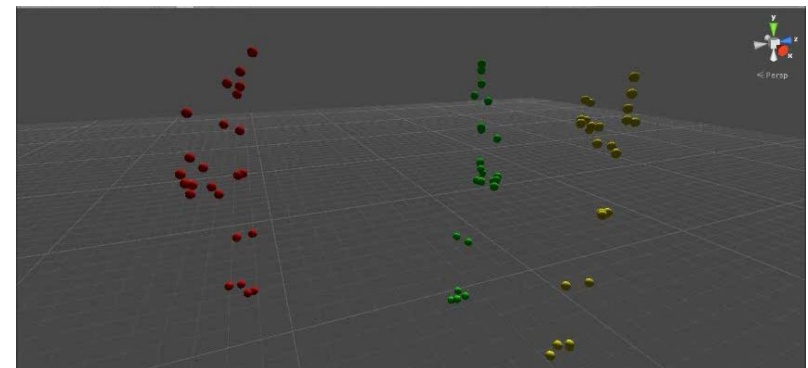
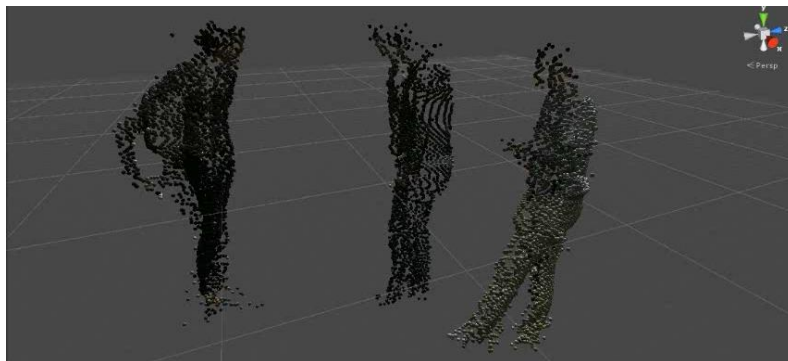
構成的考証法 — 劇をキャプチャ



User Selection: Select 55:011

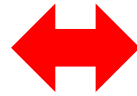


User Selection: Select 55:001



構成的考証法 — 討論フェーズ

3人称視点



1人称視点

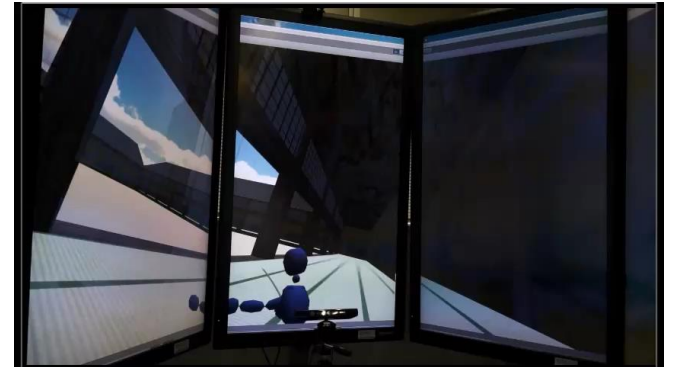
浅野



討論



吉良



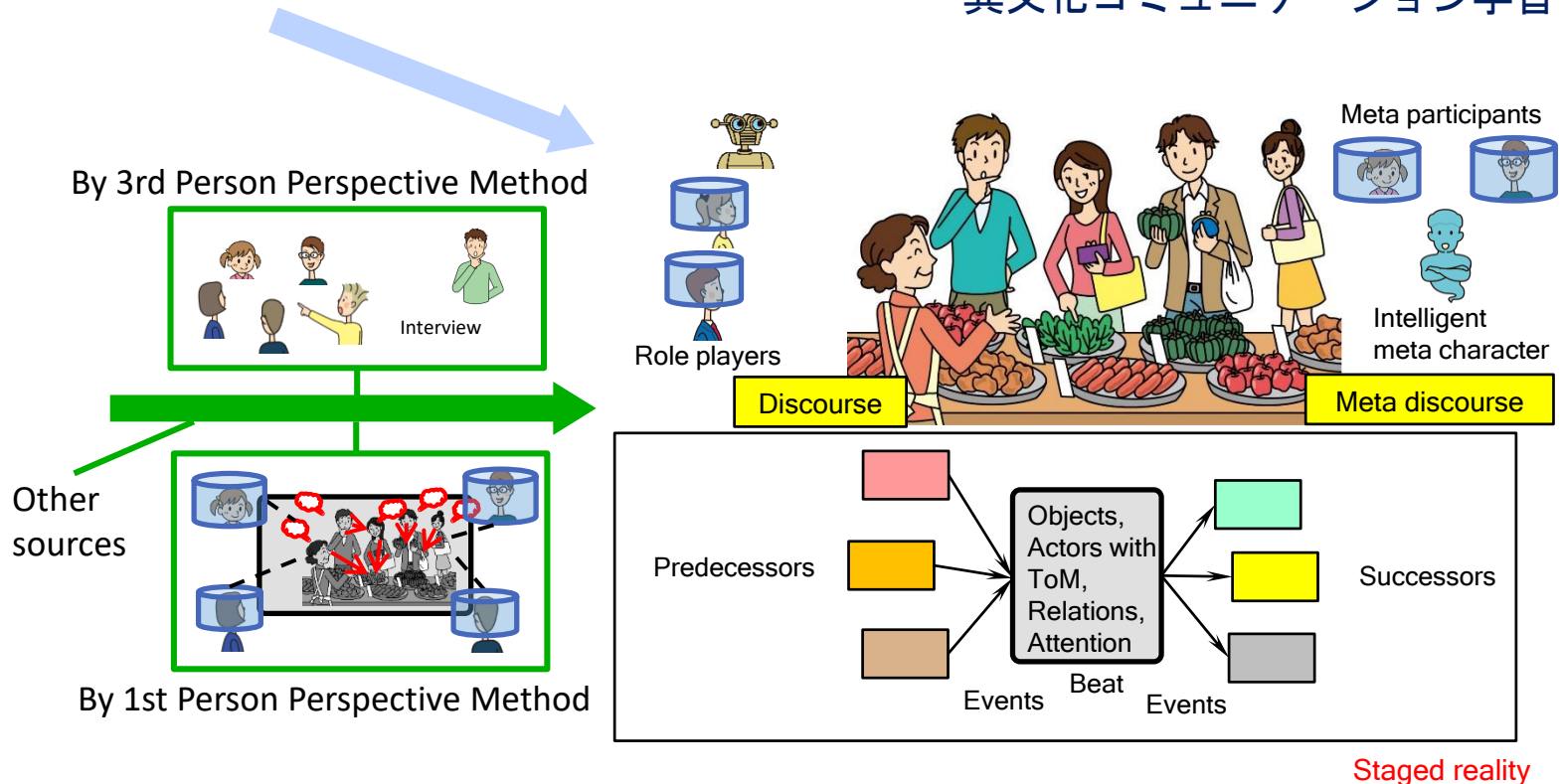
梶川



会話のエンビジョニング

④ クラウドソーシング利用
Movieからのエンビジョニング

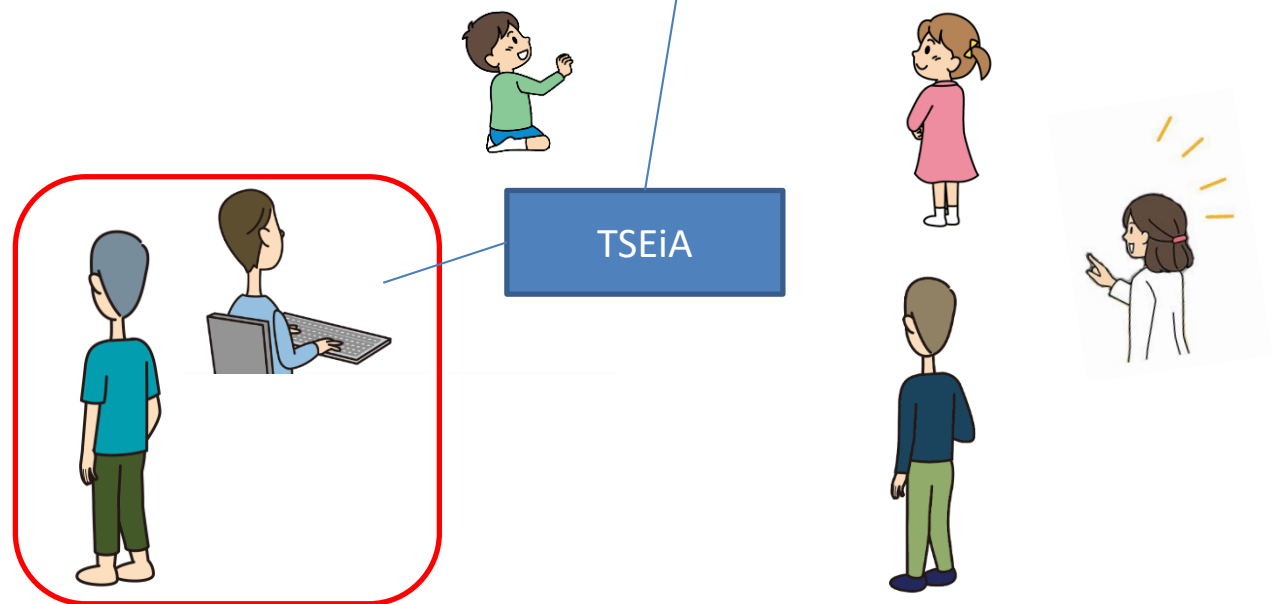
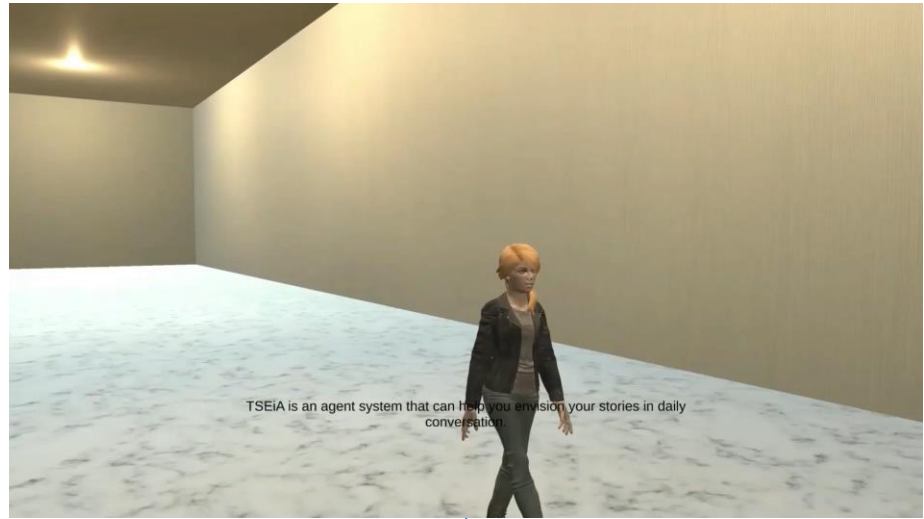
③ 応用：教師トレーニング用仮想クラスルーム
異文化コミュニケーション学習



① エンビジョニングプラットフォームTSEiA

② GECA+FACSVatar→AIアクター

TSEiA (会話エンビジョニング支援エージェント, 開発中)

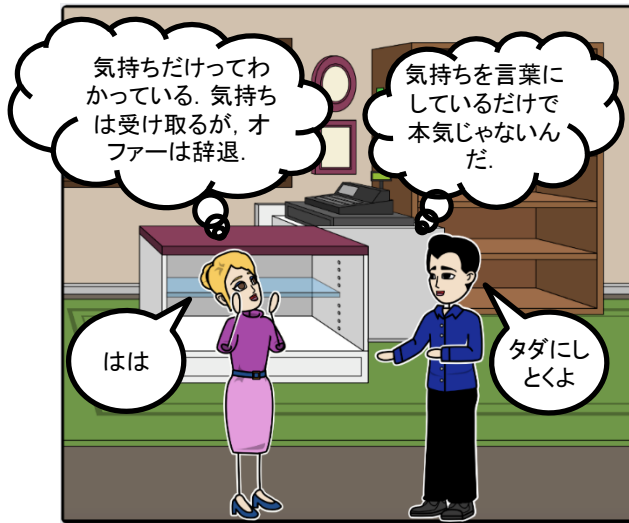


会話エンビジョニングの効果 (1)

たてまえのオファー

- ✓ レイヤ 1: 店主はタダにすると言ったが、客は笑って辞退した。
- ✓ レイヤ 2: 演じられている店主は本気でタダにすると言い、演じられている客はそれを受けた。

いい雰囲気

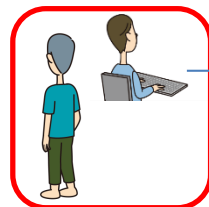


虚構を共演



値段交渉で

Envision it !



TSEiA



会話エンビジョニングの効果 (2)

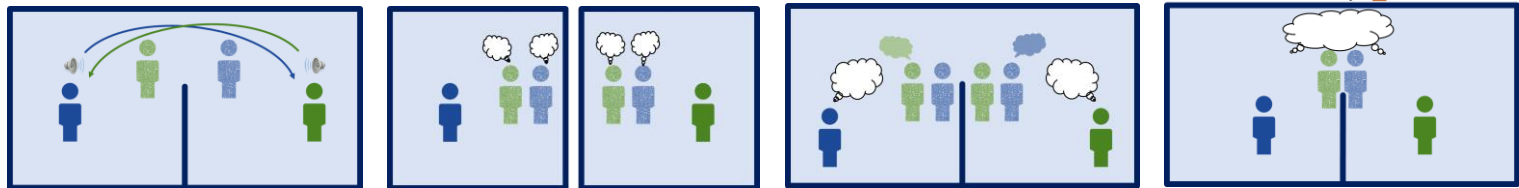
現状認識の食い違いを映像化で明示

ロールプレイ

個人ごと映像化

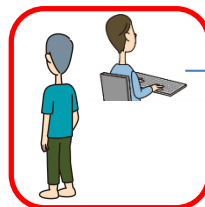
映像化の共有

討論



ジョブインタビューで

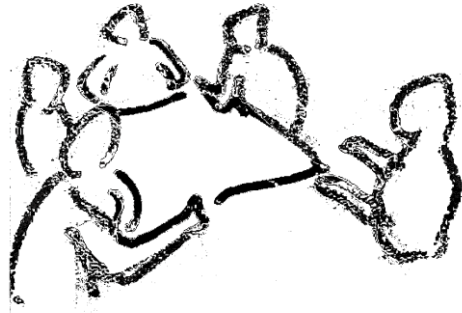
Envision it !



TSEiA



プロモ制作ボランティア



グループ討論



グループ討論

サービスシーン



トモロボ



4SenseVR



水陸両用WKWK



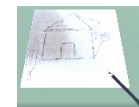
Rimon



ゼログラビティ



スカイカフェ



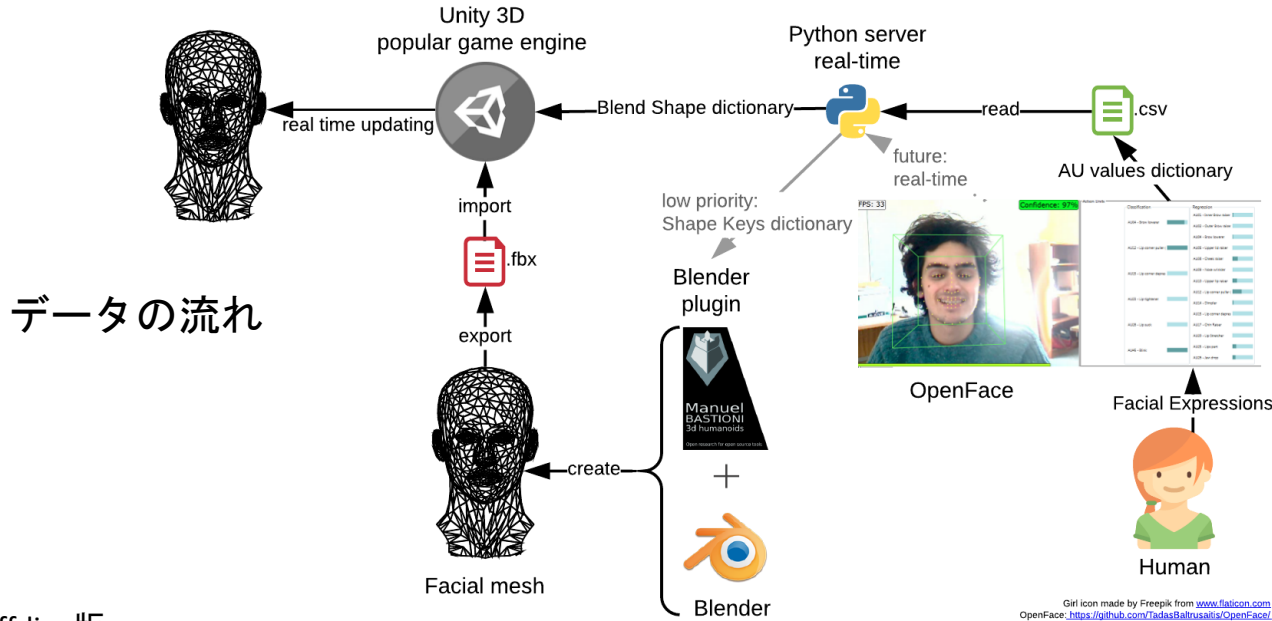
AIE



プロモ制作ブース

FACSVatar

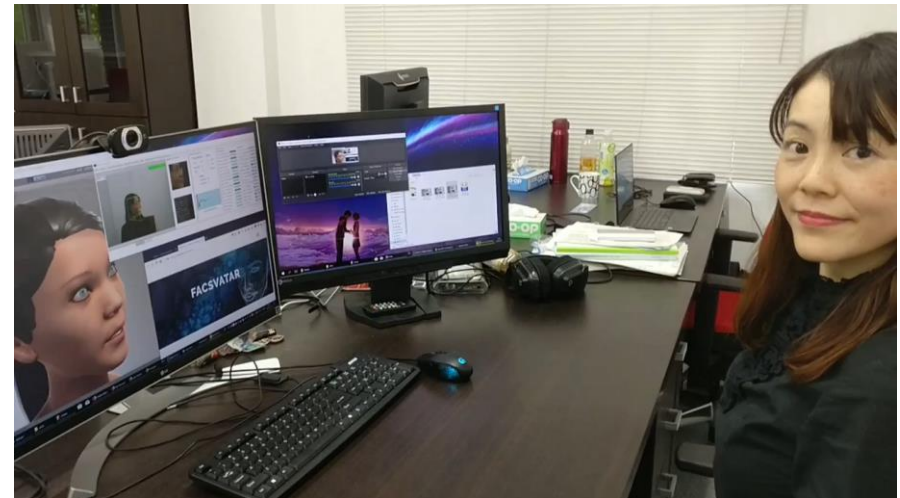
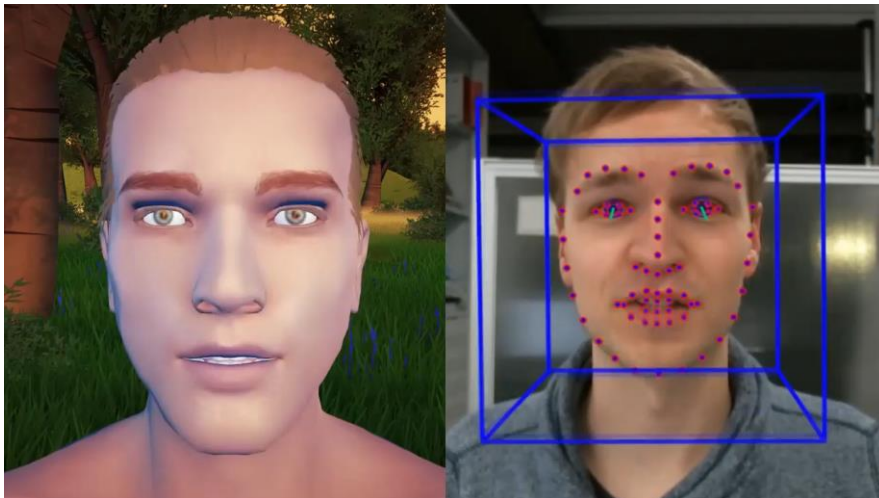
FACSを中間表現とする顔表情認識・生成(オープンソース): 顔表情を用いた研究のための基盤



FACS
 (Facial Action Coding System)
 解剖学的に独立し、視覚的に識別可能な表情動作の最小単位であるAU (Action Unit) の組み合わせで表情を記述する手法

Off-line版

On-line版



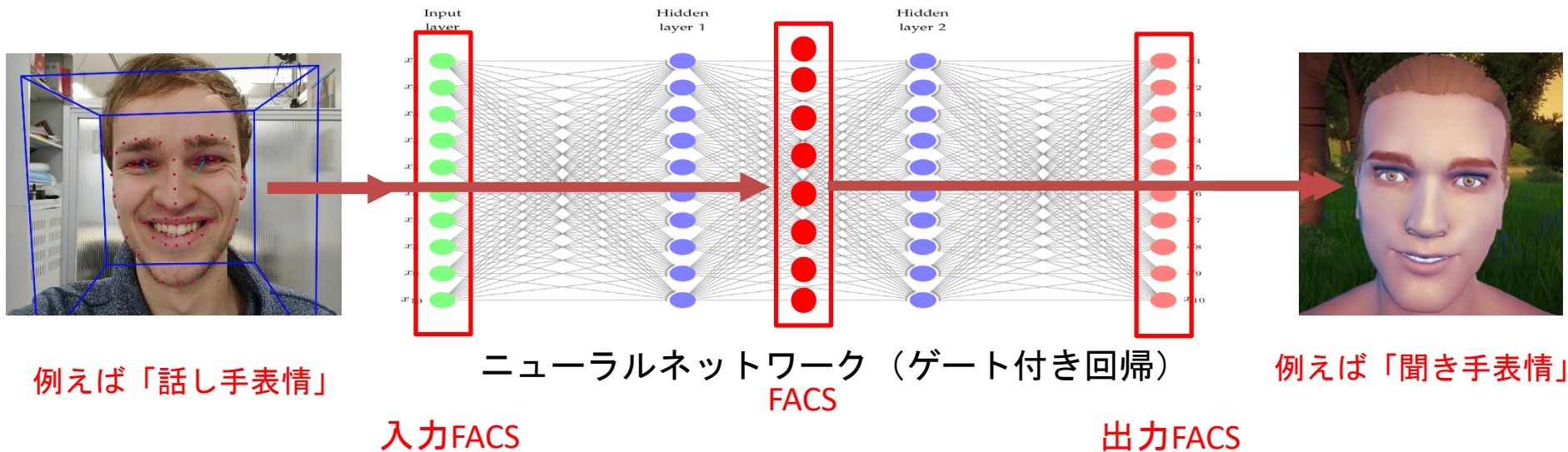
FACSVatar アーキテクチャを使った顔表情変換

顔表情変換

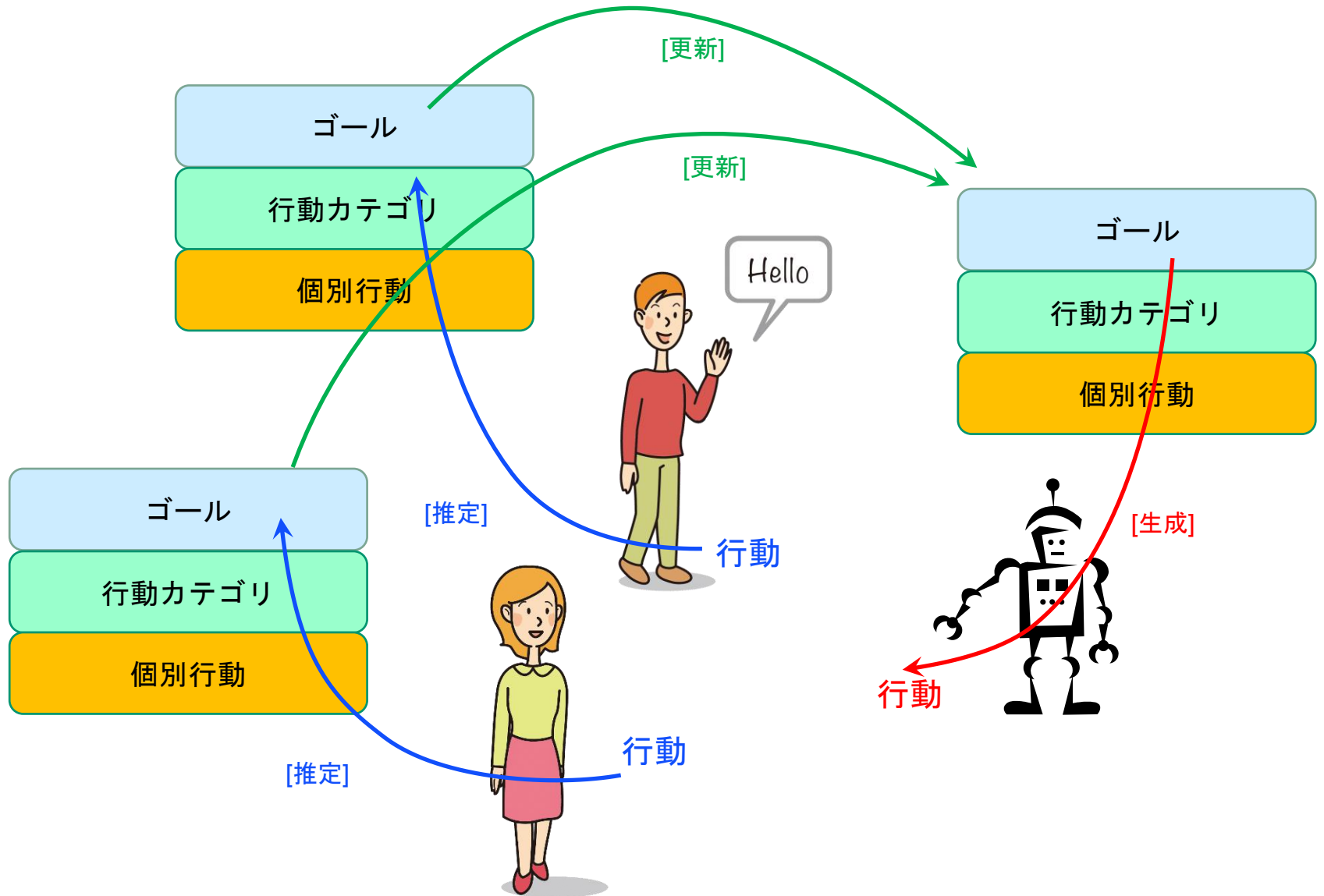
ユーザの表情を分析してFACSデータを生成

ニューラルネットワークを用いてFACSデータを変換

変換されたFACSデータから顔表情生成



循環的意図更新



随伴行動

Self-agentとの一体感を誘発したい

仮説：随伴行動生成規則



・ポイントをゲット



〈条件〉

LF/HF>20.0
SCR>17.0



汗をぬぐう

・相手プレイヤーをヒット



〈条件〉

LF/HF>20.0



こぶしを振り上げる

・ノックアウトされる



〈条件〉

SCR>17.0



地団太を踏む

仮説：エージェントが随伴的な反応（働きかけに関連があると認識できる反応）を返すことで、エージェントがどうしてもそのような反応をするのかに関する関心が高まる

まとめ

1. 会話AIの役割: {人, AI}をつなぐ
2. 会話エージェントの研究開発の歴史
3. 会話情報学の研究プログラム: データに基づく会話理解とシステム化
4. 存在感のあるエージェント ← **しっかりしたコモングラウンド**
5. コモングラウンドは幅広い: いろいろな属性 × {ミクロ~マクロ}
6. コモングラウンドの恒常的な更新としての会話
7. 会話エンビジョニングでインタラクティブドラマによる知の統合を目指す
8. 初期的な枠組み: 構成的考証法(SES)
9. 会話エンビジョニングプラットフォームTSEiA
10. AIアクター
11. 応用: ヴァーチャルクラスルーム, 異文化コミュニケーション学習
12. ムービーからのエンビジョニング
13. 循環的な意図推定, 随伴行動による志向的スタンスの強化

より詳細について (1/2)

- [Mirzaei 2019] Maryam Sadat Mirzaei, Qiang Zhang, Stef van der Struijk, and Toyoaki Nishida. Towards Conversation Envisioning for Cognitive Robots, Chapter 7, in: J. Vallverdú, V. C. Müller (eds.), Blended Cognition, Springer Series in Cognitive and Neural Systems 12, Springer 2019 (in press).
- [Mirzaei 2018] Maryam Sadat Mirzaei, Qiang Zhang, Stef van der Struijk, Toyoaki Nishida. Exploiting Virtual Reality to Advance Sociocultural Knowledge for Language Learners, WorldCALL 2018, Concepcion, Chile, November 13-16, 2018.
- [西田 2017g] 西田豊明(編), 小特集「人工知能とインタラクション」, シミュレーション学会誌, Vol. 36, No. 4, 2017年12月号.
- [西田 2017f] 西田豊明, 選択肢としての人工知能, 情報管理 Vol. 60, No. 9, p. 655-661, 2017.
https://www.jstage.jst.go.jp/article/johokanri/60/9/60_655/article/-char/ja
- [西田 2017e] 西田豊明, スキーマ・マシンとしての人工知能のインパクト, 情報管理 Vol. 60, No. 5 p. 339-344, 2017.
https://www.jstage.jst.go.jp/article/johokanri/60/5/60_339/html/-char/ja/
- [弦間 2017] 【特別対談】人工知能時代の医学教育と日本医科大学の選択, 日本医科大学 弦間 昭彦 学長 × 京都大学大学院情報学研究科 西田 豊明 教授
https://www.jhcs.or.jp/no_category/index.html?cid=28
- [西田 2017d] 西田豊明, 人工知能の歩みー第3次AIブーム, ブリタニカ国際年鑑2017.
<https://www.britannica.co.jp/products/jboy.html>
- [西田 2017c] 西田豊明, 人工知能と技術倫理, 電気評論, 2017年5月号
<http://www.ehyo.sakura.ne.jp/page030.html>
- [Nishida 2017] Nishida, T. (ed.) Human-Harmonized Information Technology, Volume 2, Horizontal Expansion, Springer 2017.
<http://www.springer.com/jp/book/9784431565338>
- [西田 2017b] 西田豊明, 人工知能スキーマ: 人々は人工知能をどうとらえているか, 視点, 情報管理, Vol. 60, No. 1, pp. 50-55, 2017.
https://www.jstage.jst.go.jp/article/johokanri/60/1/60_50/html/-char/ja/
- [西田 2017a] 西田豊明, 人工知能の現状と展望, インターネット白書2017, 2017年2月1日
<http://www.impressrd.jp/news/170201/NP>
- [西田 2016] 西田豊明, 人工知能の光と影(上)ー日米欧, 倫理問題対応急ぐ, 日本経済新聞・経済教室, 2016年9月6日.
有料会員限定リンク: <http://www.nikkei.com/paper/article/?b=20160906&ng=DGKKZ006882280V00C16A9KE8000>
- [Nishida 2016] Nishida, T. (ed.) Human-Harmonized Information Technology, Volume 1, Vertical Impact, Springer 2016.
<http://www.springer.com/in/book/9784431558651>
- [人工知能学会 2016] 人工知能学会監修, 松尾 豊(編), 中島 秀之, 西田 豊明, 溝口 理一郎, 長尾 真, 堀 浩一, 浅田 稔, 松原 仁, 武田 英明, 池上 高志, 山口 高平, 山川 宏, 栗原 聡: 人工知能とは, 近代科学社, 2016年5月30日
<http://www.kindaikagaku.co.jp/information/kd0489.htm>
- [西田 2015] 西田豊明, 人工知能がもたらす未来の「スマート」, CEL (Vol.111), 2015年11月02日
http://www.og-cel.jp/search/1230453_16068.html
- [西田 2014b] 西田豊明. 人間力・社会力を強化する情報通信技術: 人工知能を中心に, 情報管理 Vol. 57, No. 8, pp. 517-530, 2014.
<http://dx.doi.org/10.1241/johokanri.57.517>

より詳細について (2/2)

- [西田 2014a] 西田豊明. 人工知能の社会実装がもたらし得ること. 人工知能, 29巻5号, pp. 524-532, 2014.
<http://ci.nii.ac.jp/naid/110009841788>
- [西田 2013] 西田豊明. 人工知能とは(2), 人工知能学会誌, 28巻2号, pp. 326-335, 2013.
<http://ci.nii.ac.jp/naid/110009594470>
- [西田 2012] 西田豊明. 人工知能研究半世紀の歩みと今後の課題, 情報管理 Vol. 55, No. 7, pp. 461-471, 2012.
<http://dx.doi.org/10.1241/johokanri.55.461>
- [Mohammad 2015] Mohammad, Y., Nishida, T. Data Mining for Social Robotics: Toward Autonomously Social Robots, Springer 2015.
- [Nishida+ 2014] Nishida, T., Nakazawa, A., Ohmoto, Y., Mohammad, Y. Conversational Informatics—A Data-Intensive Approach with Emphasis on Nonverbal Communication, Springer 2014.
[Amazon Kindle] <https://www.amazon.co.jp/Conversational-Informatics-Data-Intensive-Nonverbal-Communication-ebook/dp/B00RZJRQZG>
- [Huang 2008] Huang, H., Cerekovic, A., Pandzic, I., Nakano, Y, and Nishida, T.: The Design of a Generic Framework for Integrating ECA Components, Proceedings of 7th International Conference of Autonomous Agents and Multiagent Systems (AAMAS08), Estoril, Portugal, pp128-135, May, 2008.
- [Nishida 2007] Toyooki Nishida (ed.). Conversational Informatics: an Engineering Approach, John Wiley & Sons Ltd, London, 2007.
<https://www.wiley.com/en-gb/Conversational+Informatics%3A+An+Engineering+Approach-p-9780470026991>
- [Nakano 2005] Masashi Okamoto, Yukiko I. Nakano, Kazunori Okamoto, Ken'ichi Matsumura and Toyooki Nishida. Producing effective shot transitions in CG contents based on a cognitive model of User Involvement, IEICE Transactions of Information and Systems Special Issue of Life-like Agent and Its Communication, IEICE Trans. Inf. & Syst., Vol. E88-D, No. 11, pp. 2623-2532, 2005.
- [久保田 2004]久保田秀和, 角康之, 西田豊明. 「知球」: 持続的に発展可能な時空間記憶の構築, 情報処理学会研究報告「ヒューマンインタフェース」, Vol.2004, No.90, 2004.
- [Nishida 2002] T. Nishida. Social Intelligence Design for Web Intelligence, Special Issue on Web Intelligence, IEEE Computer, Vol. 35, No. 11, pp. 37-41, November, 2002.
- [Terada 2002] Kazunori Terada and Toyooki Nishida. Active Artifacts: for New Embodiment Relation between Human and Artifacts, In Proceedings of The 7th International Conference on Intelligent Autonomous Systems (IAS-7), Marina del Rey, California, USA March 25-27, 2002.
- [Kubota 2000] Hidekazu Kubota, Toyooki Nishida, and Tomoko Koda. Exchanging Tacit Community Knowledge by Talking-virtualized-egos, in: Proceedings of Agent 2000, pp.285-292, 2000.
- [Nishida 1999] Toyooki Nishida, Nobuhiko Fujihara, Shintaro Azechi, Kaoru Sumi, and Hiroyuki Yano. Public Opinion Channel for Communities in the Information Age, New Generation Computing, Vol. 14, No. 4, pp. 417-427,1999.
- [Nishida 1998] Toyooki Nishida, Hideaki Takeda, Michiaki Iwazume, Harumi Maeda, and Motoyuki Takaai. The knowledgeable community: Facilitating human knowledge sharing. In: Toru Ishida (ed.): Community Computing: Collaboration Over Global Information Networks, Chapter 5, pp. 127-164. John Wiley & Sons, 1998.
<https://www.britannica.co.jp/products/jboy.html>
- [西田 1997] 西田豊明. 知識コミュニティプロジェクトからのメッセージ, 招待講演, 第3回知能情報メディアシンポジウム, 電子情報通信学会, 慶応大学三田キャンパス, 1997年12月10日.
https://www.ii.ist.i.kyoto-u.ac.jp/?page_id=5558&lang=ja