

人工知能学会第34回大会招待講演（2020.6.10）

人間の知 機械の知

Human Intelligence, Machine Intelligence

東京大学・大学院総合文化研究科

植田一博

大会関係者の皆さまへ

貴重な講演の機会を与えていただき
まして、心から感謝申し上げます。

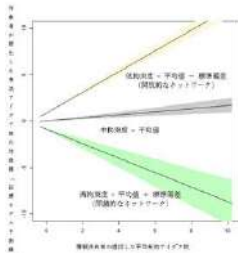
熊本にお伺いできず、とても残念に
思います。

(個人的なこと申し上げます、熊本にあまり縁がない
ようです)

植田研究室の紹介

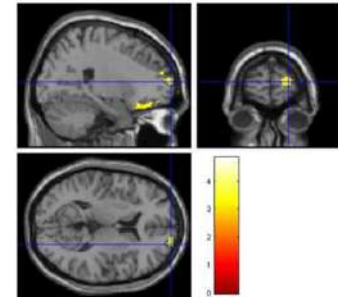
創造性とイノベーション

- 創造的思考に影響を与える諸要因（認知的・社会的）の解明
- 大量の実データの分析



JDM・消費行動研究

- 商品選択と潜在的認知・学習の関係
- **バイアスとナッジの研究**



インタラクション・デザイン

- アニマシー知覚
- **人-人/人工物のインタラクションに関する研究**
- **意図推定に関する研究**
- 集合知に関する研究



- **実世界の人の情報処理**の解明
- 情報学・心理学・脳科学・社会科学を基礎にした学際的アプローチ

エキスパートの認知

- 伝統芸能（文楽・能）の演者・古武術家の熟達化
- マイクロ波による動作計測システムの開発



研究室HP

<http://www.cs.c.u-tokyo.ac.jp/>

人工知能学会との接点

- 学会
 - 人工知能学会（1989年～）
 - 日本認知科学会（1994年～）
- 昔のAI学会大会：100-150名程度の参加者（AI冬の時代）？
 - 深く議論ができて好きだった
 - 認知科学的な議論は今よりもできた
- （しばらくご無沙汰した後）2007年より復帰
 - Deep Learning色，特にその産業応用が主流
 - いずれAIブームが去ったときどうなるのか？
- 科学的な基礎研究 (Scientific Research) にもっと重きをおいても良いのでは？



認知科学研究者との交流を増やすことが
その一助にならないか？

人工知能と認知科学の関係

- 両方に関わる初期のイベント
 - ヒクソン・シンポジウム（1948）
 - ラシュレーの行動主義批判
 - 参加者に心理学者はいない？
 - ダートマス会議（1956）
 - 思考のシミュレーションモデル→AIの創始
 - McCarthy, Minsky, Shannon, Newell, Simonなどが参加
- 人工知能と認知科学は「**双子の学問**」として誕生した
 - 情報科学研究者と後の認知科学研究者が創始
 - 心理学との交流は薄かった？

内村直之・植田一博・今井
むつみ・川合伸幸・嶋田総
太郎・橋田浩一 (著)

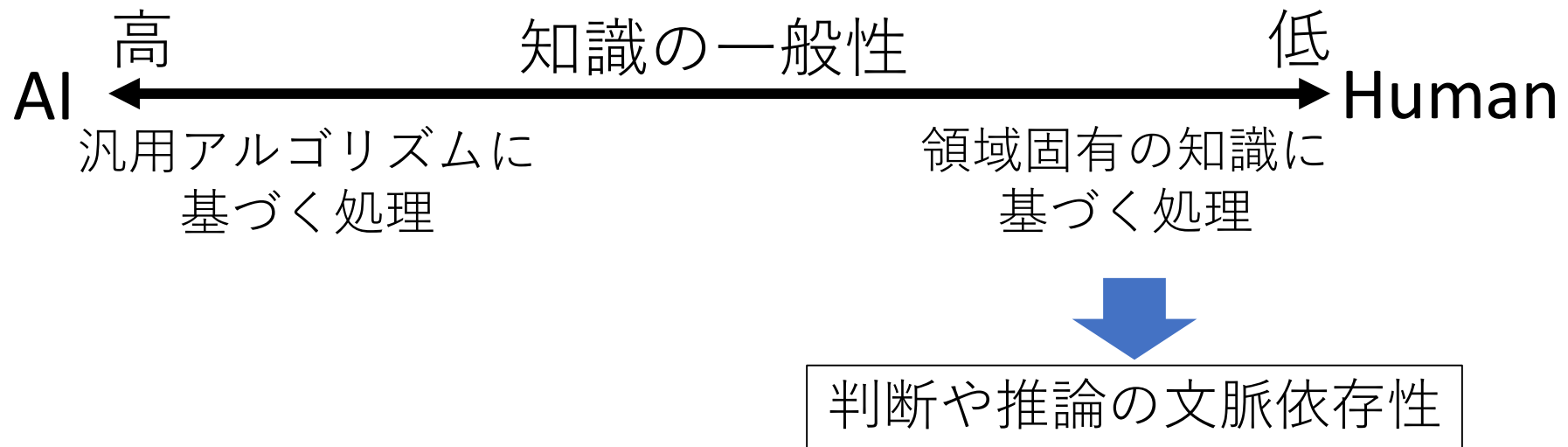
はじめての認知科学

新曜社

(画像消去)

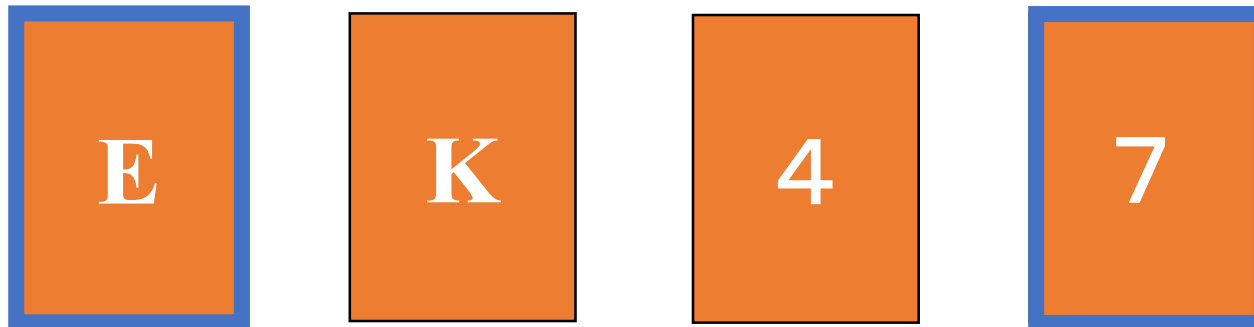
人工知能と認知科学の乖離

- 蜜月状態はどうして崩れたのか？
 - ◀ それぞれの研究が目指す方向性が変わった
- では、いつ頃どのように変わったのか？



4枚カード問題(1)

- オリジナル問題 (Wason, 1966)
 - 「もしあるカードの片面に母音を書いてあるならば、そのカードの裏面には偶数が書いてある」という
ルールの成立しているかどうかを確かめるために、
必ずめくって見なければならぬカードはどれか。
 - 条件文推論 ($p \rightarrow q$, p から q を推論)



- 低い正答率 (4%程度, 追試では10%程度)
 - ➡ 人は一般的な規則を使うのが不得手

4枚カード問題(2)

- 飲酒問題：変形 4 枚カード問題 (Griggs & Cox, 1982)
 - 勤務中の警官
 - 「もしある人がビールを飲んでいるならば、その人の年齢は19歳を超えていなければならない」という規則に違反しているかどうかを決定するために必要なカードを選んでください。

ビール
を飲ん
でいる

コーラ
を飲ん
でいる

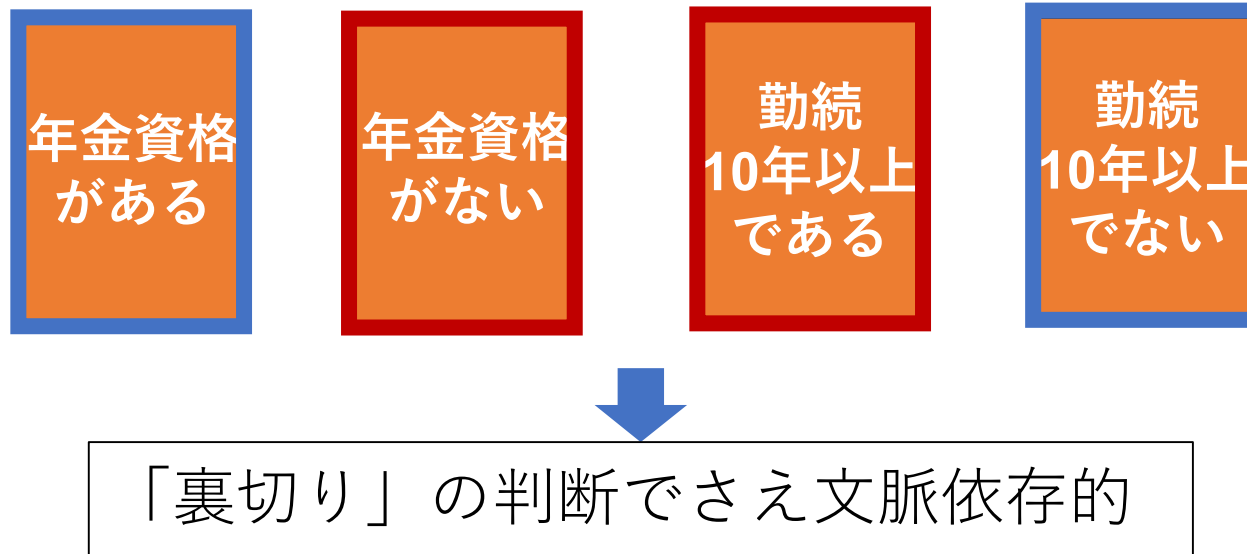
22歳

16歳

- 記憶手がかり説 ➡ 人は領域固有の知識を使う
- 裏切り者の検知(Cheng & Holyoak, 1985)
 - ➡ 中程度に一般的な知識

4枚カード問題(3)

- 年金問題：変形4枚カード問題 (Gigerenzer & Hug, 1992)
 - 「年金の資格を得るには、10年以上勤続していなければならない」という規則の確認
 - 回答者の視点が回答に与える影響を調査
 - 「あなたは経営者の立場にある」という教示を与えた場合
 - 「あなたは労働者の立場にある」という教示を与えた場合



認知科学者からのAI/GOFAIへの批判

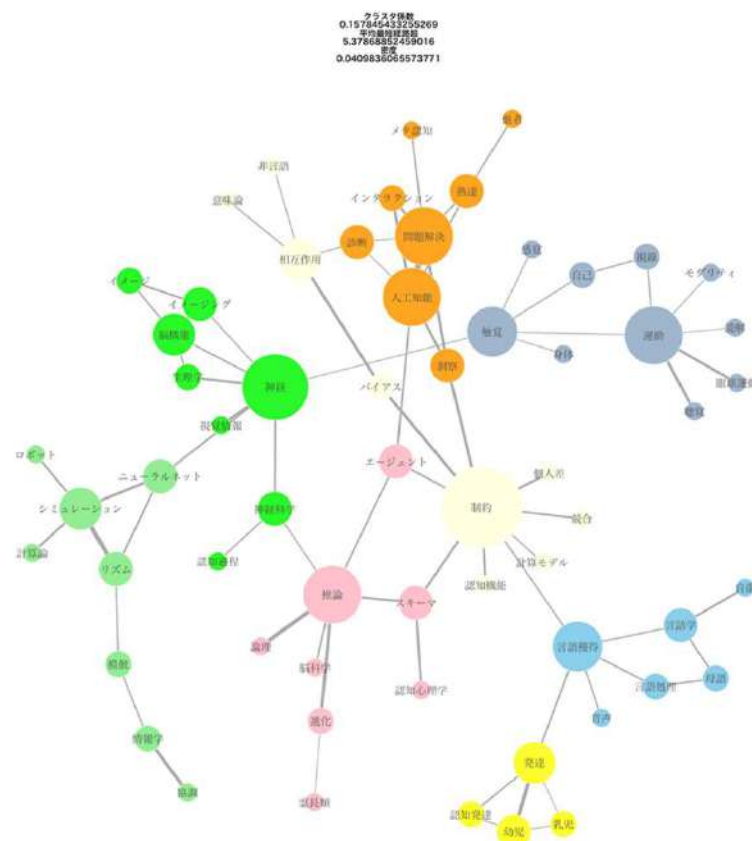
- 人の知能の文脈依存性は計算モデル（認知モデル）では捉えられない
- AIプログラムの学習には大量のデータが必要だが、人ではそうではない
- 将棋のプロに勝てるプログラムも、ノービスができる知的活動を実現できない



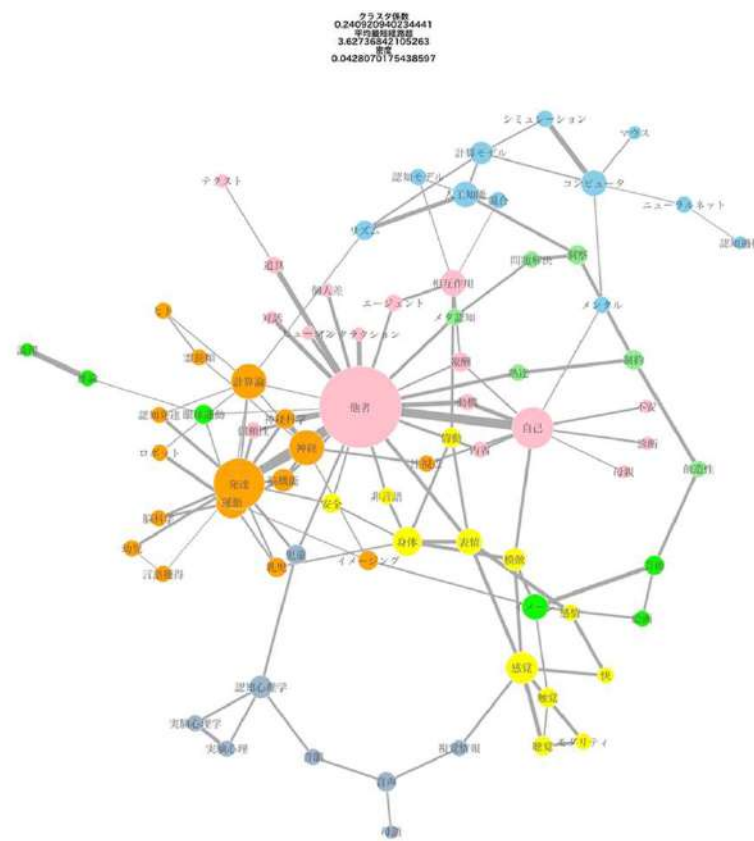
計算論アプローチへの批判をしばしば目にする
（両者の乖離の理由か？）

雑誌『認知科学』のテーマの変遷 (内村ほか, 2016)

2001～2006年



2007～2012年



背景にある事柄と主張したいこと

- 情報が不足している状況で、適切に処理する能力
(昨日の中島先生の基調講演)
 - 人の知能の本質的側面
- 疑問
 - どういう点で適切なのか？
 - 不足している情報をどのように補っているのか？



- 高次認知 = 人の思考（推論，判断・意思決定）の特徴を炙り出すのに，人工知能モデルとの比較や統計的なデータ分析が有効
 - 計算論モデリング：人のデータをモデル（e.g. 強化学習モデル）にフィッティングする

具体的な研究事例

1. ヒューリスティックは合理的な判断の「みすぼらしい代用物」か？
2. 人はどういう状況でヒューリスティックを用いるのか？（文脈依存性）
3. 誤った記憶にも効用はないのか？
4. 人工知能と比較した場合の人の選好推定の特徴とは？

ヒューリスティックは合理的な
判断の「みすぼらしい代用物」
か？

少ない情報から考える

- 友人の旦那さんは、女癖が悪く遊び人で有名。
次のうち起こる可能性が高いのはどちらか？
 1. 夫婦喧嘩による離婚
 2. 浮気が原因の夫婦喧嘩による離婚
- 2.を選ぶのは、論理的には誤り
- 連言錯誤 (conjunction fallacy)
 - ある前提について A という推論と A & B という推論を提示したとき、論理的には A だけの方が可能性が高いにもかかわらず、A & B の方を尤もらしいと感じる
- 少ない情報からありそうなことを考えるのは人間の
特徴かもしれない

リンダ問題 (Tversky & Kahneman, 1982)

- 問題
 - リンダさんは31歳，独身で，率直に意見を言う，頭脳明晰なタイプである。彼女は大学時代に哲学を専攻していた。また学生時代には，差別や社会正義の問題に強い関心を持ち，反核のデモに参加していた。
- 被験者の判断
 - 銀行の窓口係 < 銀行の窓口係かつフェミニスト運動の活動家
- 代表性ヒューリスティック (representativeness heuristic)
 - リンダさんの記述は，「フェミニスト運動を積極的に行う人」の典型例を示しているが，「銀行の窓口で働く人」の典型例を示してはいない
 - 論理ではなく，典型例との類似性によって判断



ヒューリスティックと行動バイアス

- アルゴリズム



- ヒューリスティック

- 必ずしも正しい解が得られるとは限らないが、おおむね正しい解が**短時間で**得られることが経験的にわかっている方法

- 行動バイアス（認知バイアス）

- ヒューリスティックによる判断が客観的で正しい判断と大きく隔たる場合、それを**行動バイアス**と呼ぶ

利用可能性ヒューリスティック (availability heuristic)

- 日常生活で簡単に利用できる情報，想起しやすい情報に頼って判断
 - 想起容易性，検索容易性
- 問題 (Tversky & Kahneman, 1973)
 - a. 小説の4ページ分（約2000語）の中に7文字の単語で末尾がingで終わるものはいくつあると思うか？
 - b. 小説の4ページ分（約2000語）の中に7文字の単語で6番目がnのものはいくつあると思うか？
- 回答の平均：a. 13.4個 vs. b. 4.7個
 - 論理的には，a.の個数 < b.の個数となるべき

あるtwitterでの調査

(@kodai_kusano

https://twitter.com/kodai_kusano/status/790032281020661760)

- 日本にある以下の施設の数で，多い順に表したのは次のうちどれでしょう。
 - 歯科医院 > コンビニ > 美容院 35.3%
 - 美容院 > 歯科医院 > コンビニ 22.4%
 - コンビニ > 歯科医院 > 美容院 **42.4%**
- 次のうちあなたが一番よく頻繁に行く場所はどれですか？
 - 歯科医院 6%
 - コンビニ **89%**
 - 美容院 5%
- 接触頻度の順に多さを判断した可能性（利用可能性ヒューリスティックの影響）

生態学的合理性 (ecological rationality)

- 1990年代以前の意思決定研究では
 - 正答を妨げるような方略
 - 正答を導く推論からすると、みすぼらしい代用物
- 前のスライドの例
 - 人が環境構造を反映した判断を行える可能性
 - 必ずしも行動バイアスを生むとは限らない？
- 生態学的合理性 (Gigerenzer, Todd, and ABC Research Group, 1999; Gigerenzer, 2007)
 - 「生きるために理にかなっている」という視点



- 知らないことがかえって良い判断を生む
 - 再認ヒューリスティック (recognition heuristic)

人口推定課題

(Goldstein & Gigerenzer, 2002; Honda, Matsuka, & Ueda, 2017, *Cognitive Science*)

どちらの都市のほうが人口が多いですか？

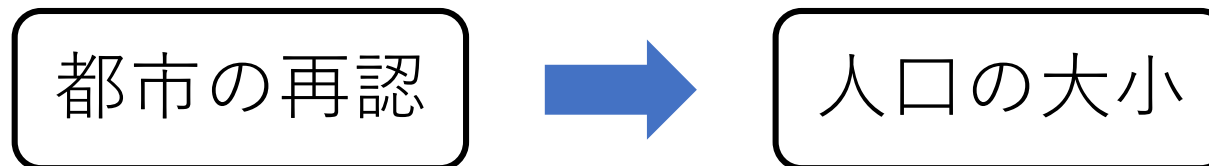
甲府市

北杜市

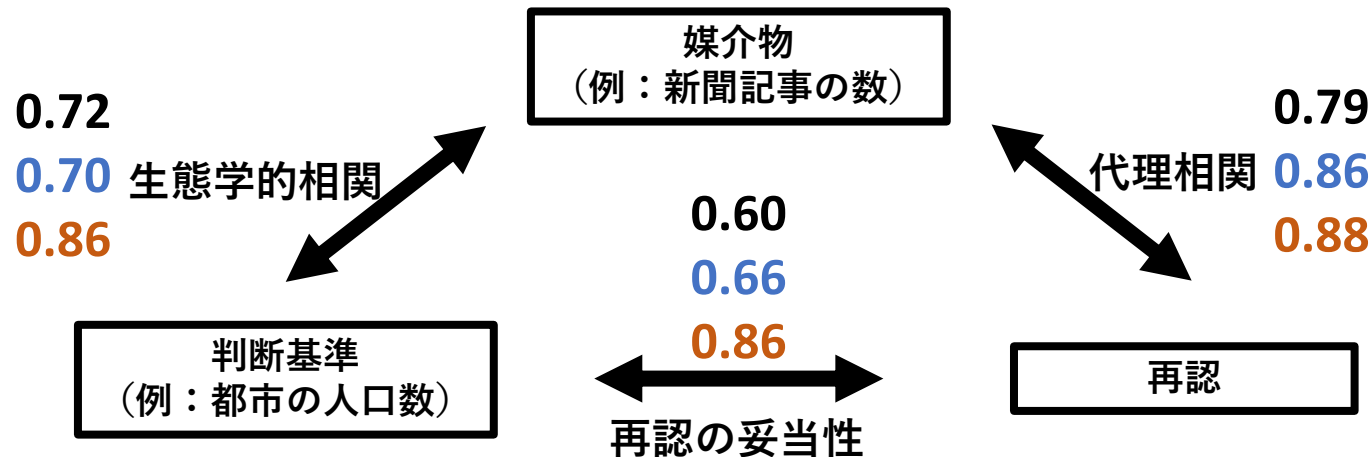
知っている都市の方が多いと考えたのでは？

再認ヒューリスティック

- Recognition heuristic (Goldstein & Gigerenzer, 2002; Gigerenzer & Goldstein, 2011)
 - 1つの都市を知っていて（再認できる），もう一方の都市を知らない（再認できない）場合，知っている都市のほうが人口が多いと考える



どうして再認ヒューリスティックでうまくいくのか？



都市	ドイツの都市	アメリカの都市	日本の都市
新聞誌	Die Zeit	Chicago Tribune	読売新聞 (朝日新聞)
掲載年	1995-1997	1985-1997	1990-2010
実験参加者	ドイツ人学生30名	アメリカ人学生67名	日本人学生107名
都市の人口	10万人以上	10万人以上	上位30都市

(Goldstein & Gigerenzer, 2002; Honda, Matsuka, & Ueda, 2017, *Cognitive Science*)

ヒューリスティックが有効な理由をデータ分析により明らかに

Less is more 効果

- 知らないときに再認ヒューリスティックを使って回答する場合の正答率
≥ すべて知識で回答する場合の正答率
- 「サンディエゴとサンアントニオのどちらが人口が多いか」という質問 (Goldstein & Gigerenzer, 2002)
 - アメリカの学生の正答率：62%（両方とも100%再認可）
 - ドイツの学生の正答率：100%（サンディエゴは再認可だが、サンアントニオの再認率は50%）
- 人口の多い順でアメリカとドイツの各22都市を用いた人口判断課題（実験参加者はアメリカ人学生）
 - アメリカの都市の平均正答率：71.1%
 - ドイツの都市の平均正答率：71.4%

成績の良い選手の選択 (Snook & Cullen, 2006)

- プロのアイスホッケーリーグNHLの選手の中から二者択一で成績が良い選手（**highest career point total**が高い人）を選択させる課題
 - カナダ人の大学生男女が実験に参加
- NHLの知識が少ない女性の成績 $\div$ 男性の成績
- 女性は再認ヒューリスティックを用いて判断



- 再認ヒューリスティックがもつ「無知の知」としての効用
- みすぼらしい代用品ではない

人はどういう状況でヒュー
リスティックを用いるのか？
(文脈依存性)

ヒューリスティックは いつ使われるのか？

- 人はヒューリスティックに頼りがち
 - Tversky, Kahneman, Gigerenzerで一致した見解
- 再認ヒューリスティックの研究では
 - 推論プロセスにおいてヒューリスティックの使用を否定する研究も多い
 - 知識を用いた推論を行っているという証拠もある

解決すべき事から

ヒューリスティック
ベースの判断



使用の境界を
明らかにする必要



知識ベースの判断

人の判断方略を推定するのに計算論モデリングが必要

再認に関わるヒューリスティック

対象を両方とも再認できたとする...

- 思い出しやすさ（どのくらい早く思い出せたか）

➡ 流暢性ヒューリスティック (fluency heuristic)

- 都市の人口数、会社の収益の判断などで有効 (Hertwig, Herzog, Schooler, & Reimer, 2008; Schooler & Hertwig, 2005)
- 対象への馴染み深さ

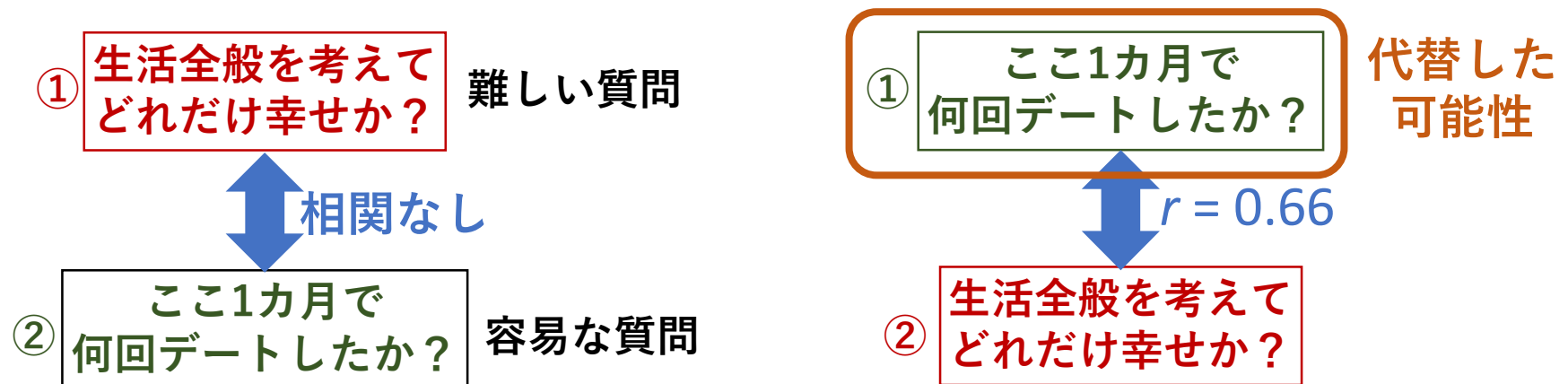
➡ 親近性ヒューリスティック (familiarity heuristic)

- 都市の人口数の判断で有効 (Honda, Abe, Matsuka, & Yamagishi, 2011)

Attribute substitution

(Kahneman & Frederick, 2002)

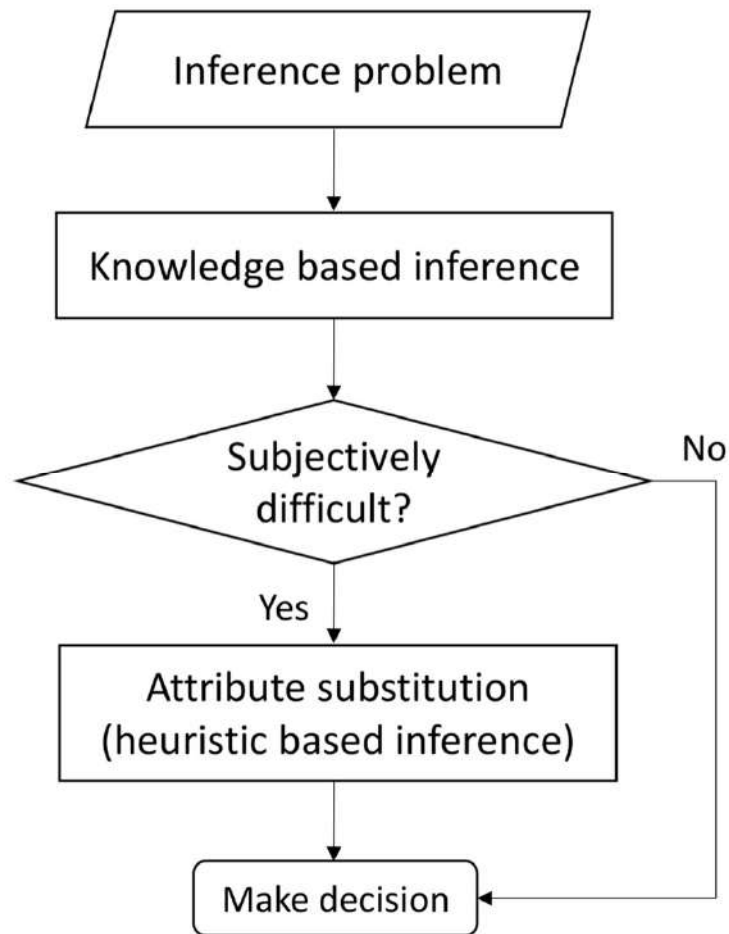
- 大会会場近くのレストランA, Bのいずれに入るか？
 - 知識がないときには，混んでいる店に入る
- 判断の対象の属性が容易に評価できない場合は，もっともらしい回答を出してくれるような属性に“代替”して，その属性からの評価を無意識のうちに行う



(Strack, Martin, & Schwarz, 1988, *Eur. J. Soc. Psychol.*)

仮説

主観的に難しいと思える問題にヒューリスティックを用いる



(Honda, Matsuka, & Ueda, 2017, *Cognitive Science*)

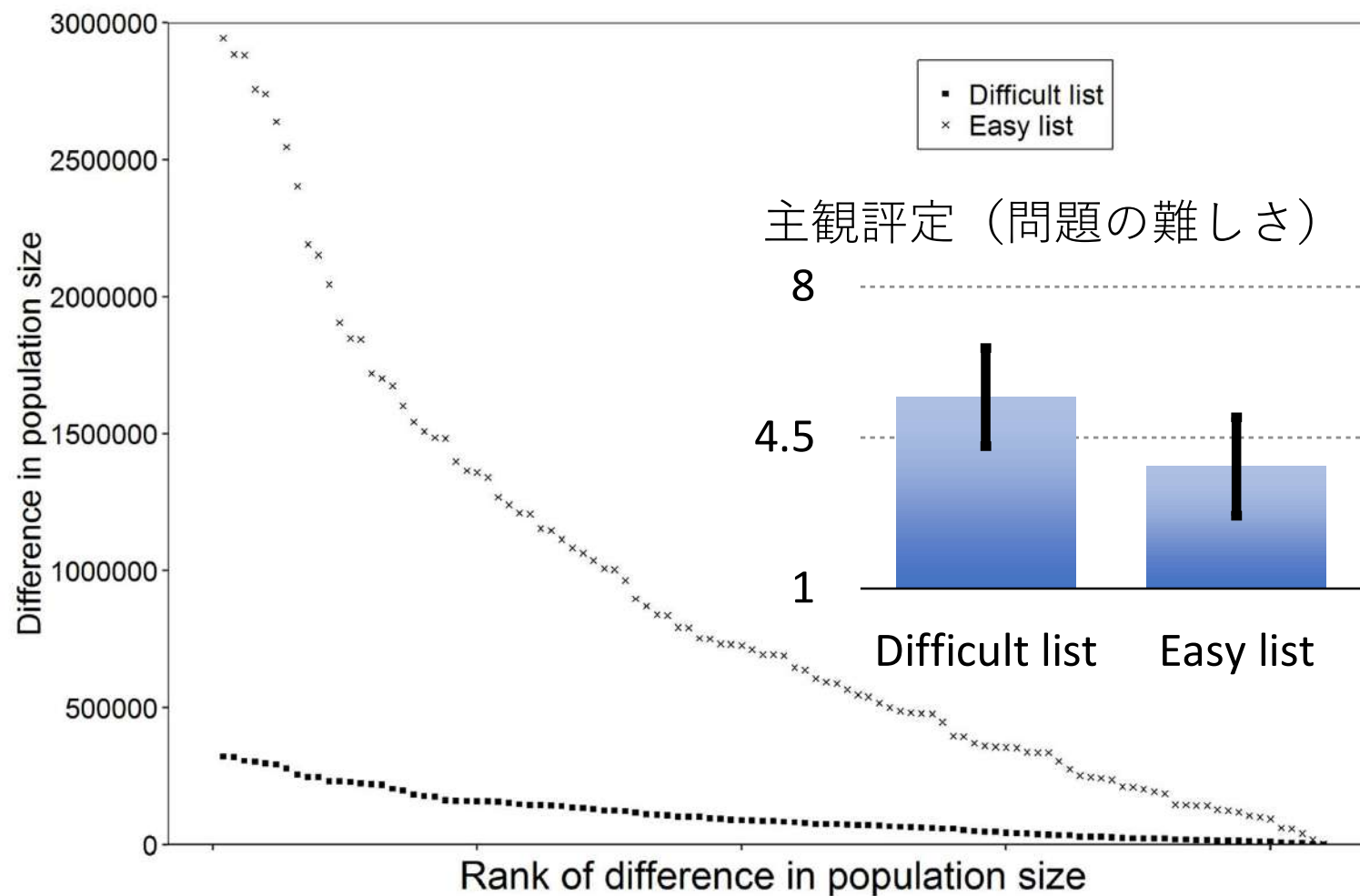
用いた都市のリスト

- 人口推定課題
人口数が近いペア
の問題は主観的に
難しいと感じる？

- 各リスト
全105ペア

順位	Difficult list	Easy list
1	川口市	横浜市
2	町田市	大阪市
3	郡山市	名古屋市
4	高崎市	札幌市
5	津市	神戸市
6	佐世保市	京都市
7	八戸市	福岡市
8	松本市	広島市
9	日立市	仙台市
10	山口市	千葉市
11	高岡市	新潟市
12	今治市	浜松市
13	都城市	熊本市
14	大垣市	岡山市
15	足利市	鹿児島市

各ペアの人口数の差異

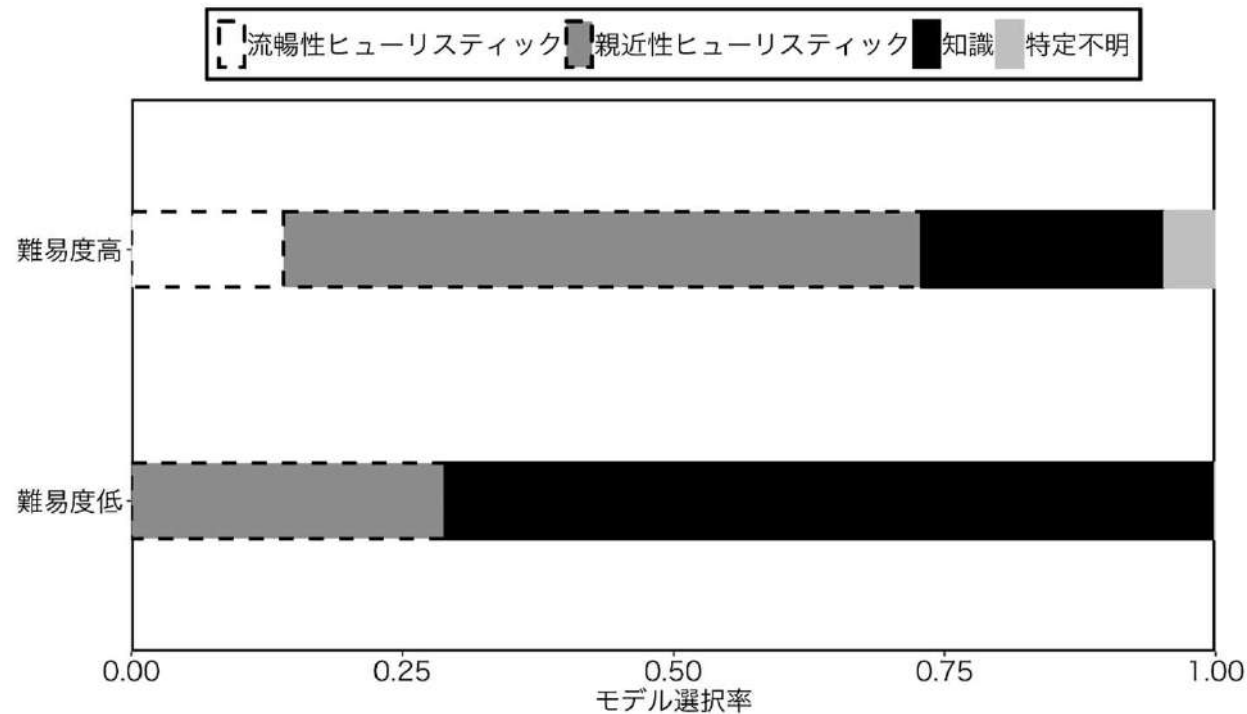


(Honda, Matsuka, & Ueda, 2017, *Cognitive Science*)

分析方法

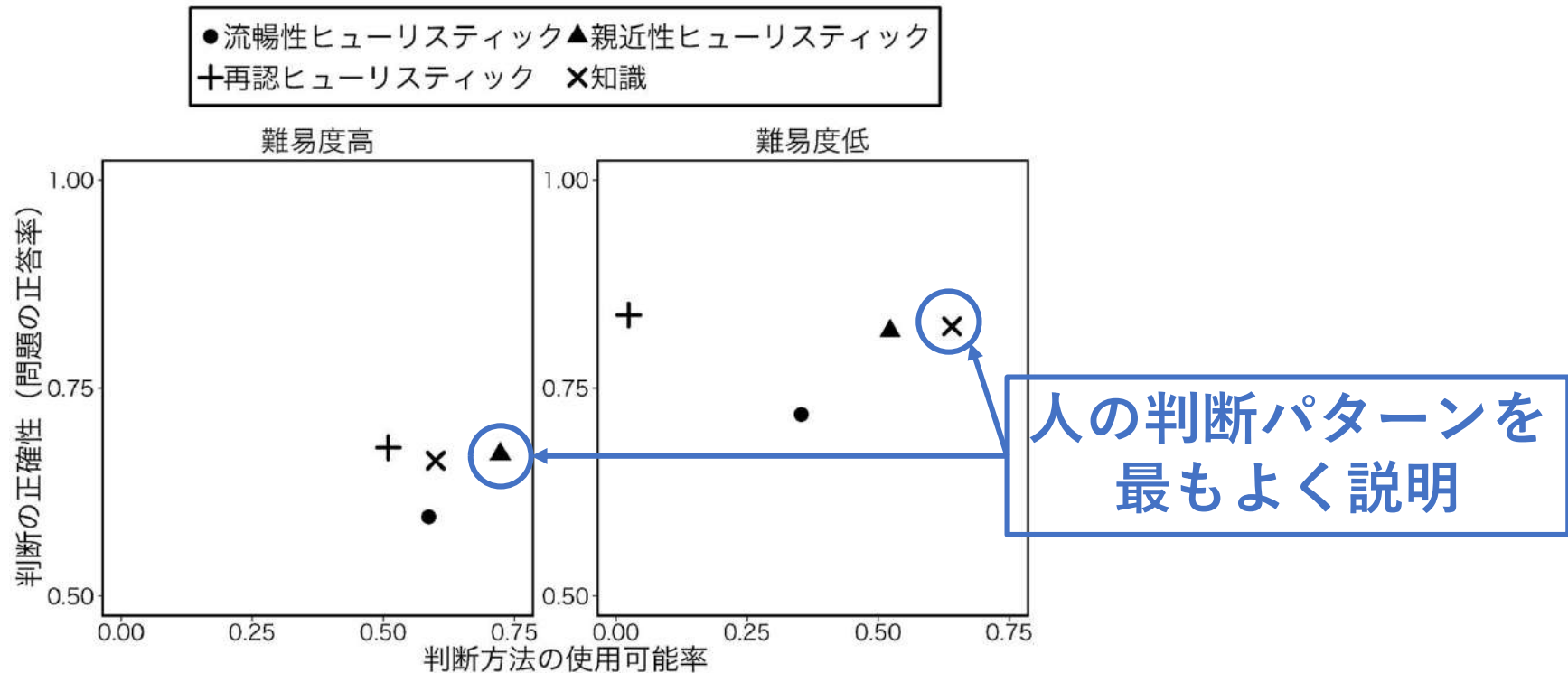
- 個人ごと，都市リストごとに最も説明力が高いモデルがどれであることを統計的に分析
 - 最尤法に基づくアプローチ
 - BICを指標として，最も説明力の高いモデルを選択
 - Bayesian model weightを指標として，選択されたモデルの証拠の強さを求める
- 候補のモデル
 - 流暢性ヒューリスティック (fluency heuristic)
 - 親近性ヒューリスティック (familiarity heuristic)
 - 知識 (knowledge-based)

結果：モデル選択



- 各リストで判断パターンを説明するモデルは異なる
 - 難易度高リスト：親近性ヒューリスティックを用いた判断
 - 難易度低リスト：知識を用いた判断
- Attribute substitutionの考え方と整合的な結果

結果：判断の正確性と使用可能率



- 判断の正確性：いずれのリストでも判断方法による違いはない
 - ヒューリスティックも正確な判断を与え得る
- 使用可能率：難易度高リストでは親近性ヒューリスティックが最も高く、難易度低リストでは知識に基づく判断が最も高い

人は環境構造に適応した形で、かなり正確な判断が可能

まとめ

- ヒューリスティックの使用は，しばしば行動バイアスを生む
- ヒューリスティックは環境構造に適応した形で使用される
 - かなり正確な判断をもたらす得る ➡ みすぼらしい代用品ではない
- ヒューリスティックベースの判断と知識ベースの判断の境界線
 - 人は利用可能性に応じて使い分け，それなりに正しい判断を行っている

Familiarity-matching

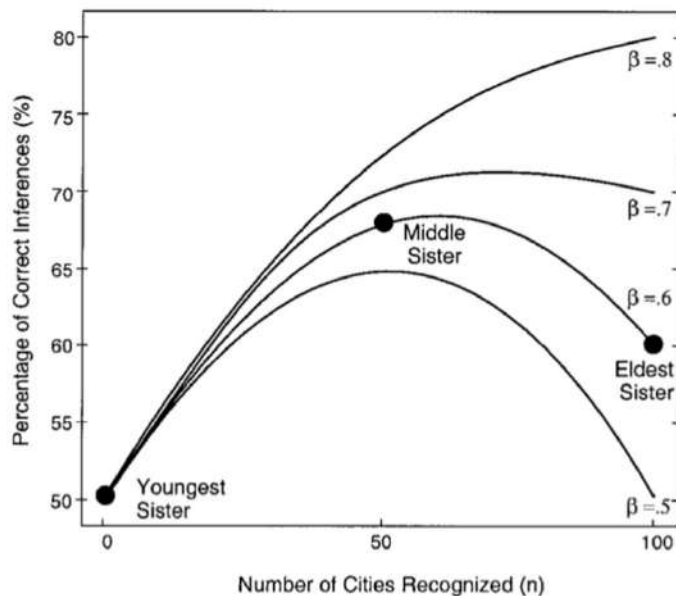
(Shirasuna, Honda, Matsuka, & Ueda, 2020, *Cognitive Science*)

- 新しい問題
 - ダボスという都市がある国はどちらか？：マリ or スイス
 - ダボスは聞いたことがあるので、**馴染みのある**スイス
 - シカソという都市がある国はどちらか？：マリ or スイス
 - シカソは聞いたことがないので、**馴染みのない**マリ
- 新しいヒューリスティックの提案
 - 「問題文で提示された対象(Q)に、より近い familiarity をもつ選択肢(A)を選ぶ」という familiarity-matching
 - $|Fam(Q) - Fam(A)| < |Fam(Q) - Fam(B)| \Rightarrow A$ を選択
- familiarity-matching が ecological rationality を有することの検証
 - 「商品 X を買うなら？ A社製かB社製か」という日常的な商品選択場面でも使われている可能性を示唆

誤った記憶にも効用はないのか？

人の記憶の性質

- 記憶の制約：短期記憶（作動記憶）の容量限界
➡ 忘却
- 最近の意思決定研究は、記憶容量の制約が、必ずしも行動バイアスに繋がらないことを示唆
 - 「知らない」ことが適応的に機能する場面が存在する



↓

Less is more 効果
(Goldstein & Gigerenzer, 2002)

記憶の制約がもつ適応的性質

- これまで量的な側面が注目されてきた
 - 限定的な知識量は適応的判断に繋がる
 - 質的な側面はどうであろうか？
 - 「記憶の誤り」は判断にどのような影響を及ぼすのか？
 - 錯誤記憶 (false memory; Loftus, 1994, *The Myth of Repressed Memory*)
 - 誤記憶：過去の事実が改変，修正されて想起されること
 - 偽記憶：全く経験していないことをあたかも事実のように想起すること
- ➡ negativeな側面に焦点化？

Memory-based false belief

(Honda, Matsuka, & Ueda, 2016, *CogSci2016*)

- 実世界の事象に対する誤った記憶に基づく信念
(誤信念)

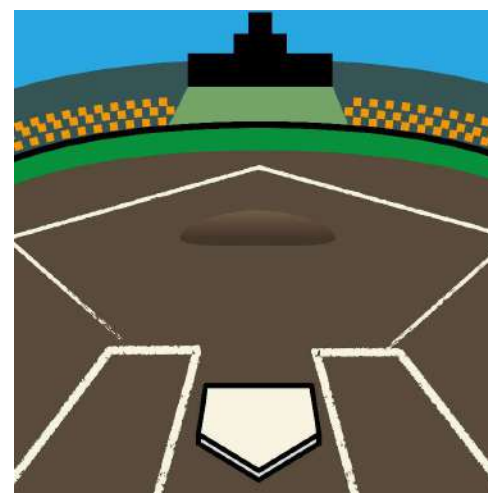
横浜市



高等裁判所がある

False Positive Belief (FPB)

西宮市



プロ野球球団はない

False Negative Belief (FNB)

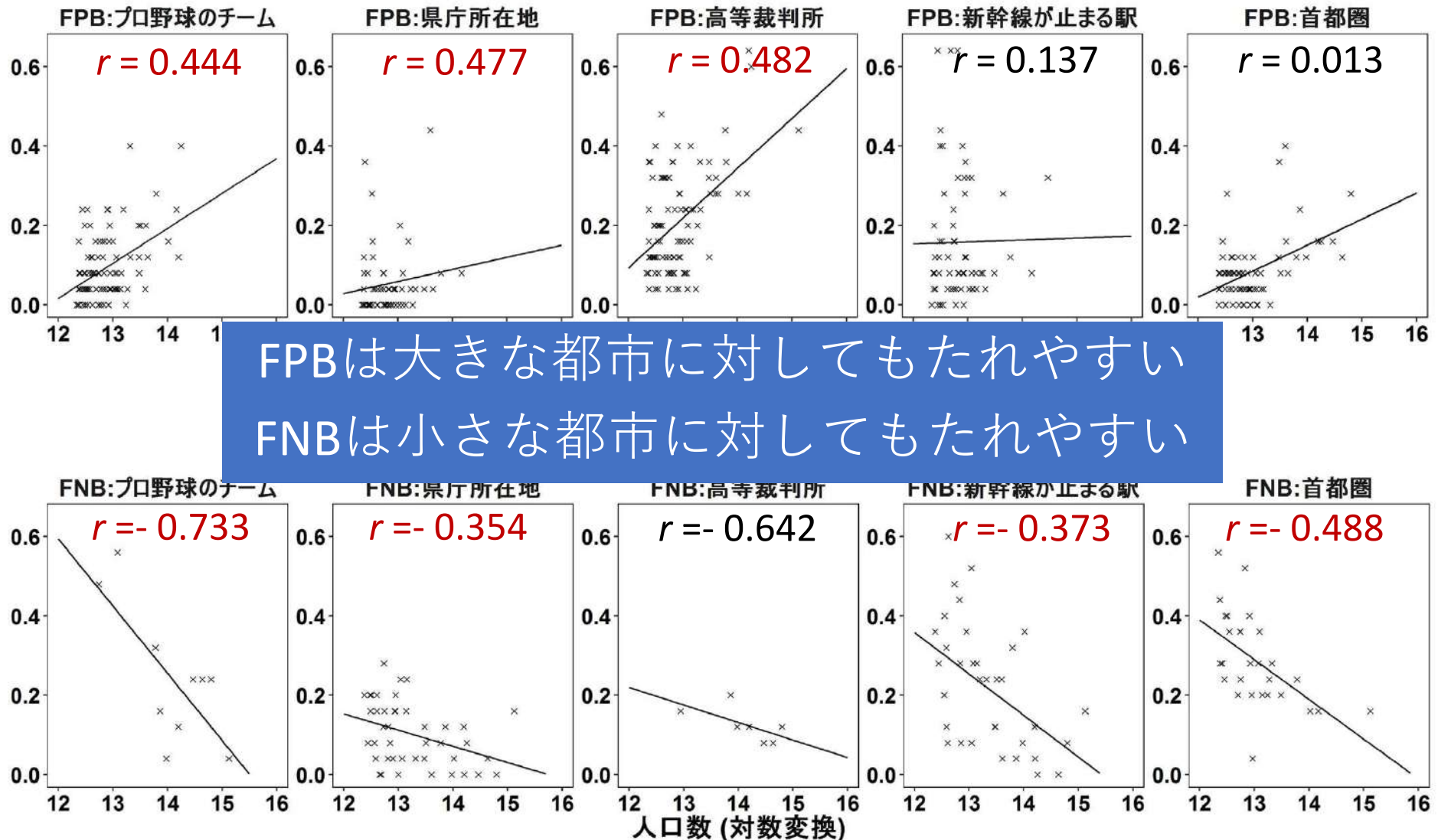
記憶に基づく誤信念

- **FPBとFNB**は系統的にかつ実世界の環境と整合的な形で生じている可能性
 - 誤った記憶は環境構造を反映している
- 「大きな都市」と強い連想がある属性について
 - **FPB**は大きな都市に対してもたれやすい
 - **FNB**は小さな都市に対してもたれやすい
- **仮説**：環境情報を推論する上で、誤った信念は適応的に機能する
 - 適応的に機能する条件を探る

実験方法

- 実験参加者：学部生**25名**
 - 人口上位**100**都市（東京**23**区を除く）の**5**属性を問う
→属性の有無の記憶と人口数に相関関係があるか？
- 質問項目
 - **5属性**
 - 「プロ野球チーム」があるかどうか
 - 「県庁所在地」かどうか
 - 「高等裁判所」があるかどうか
 - 「新幹線の停車駅」があるかどうか
 - 「首都圏」にあるかどうか
 - 馴染み深さ (**familiarity**)  記憶痕跡の指標

結果：人口数と誤信念の関係



とりあえずのまとめと問題提起

- 記憶に基づく誤信念はランダムに発生しているわけではない
- 誤信念が，実世界の構造を反映する形で形成

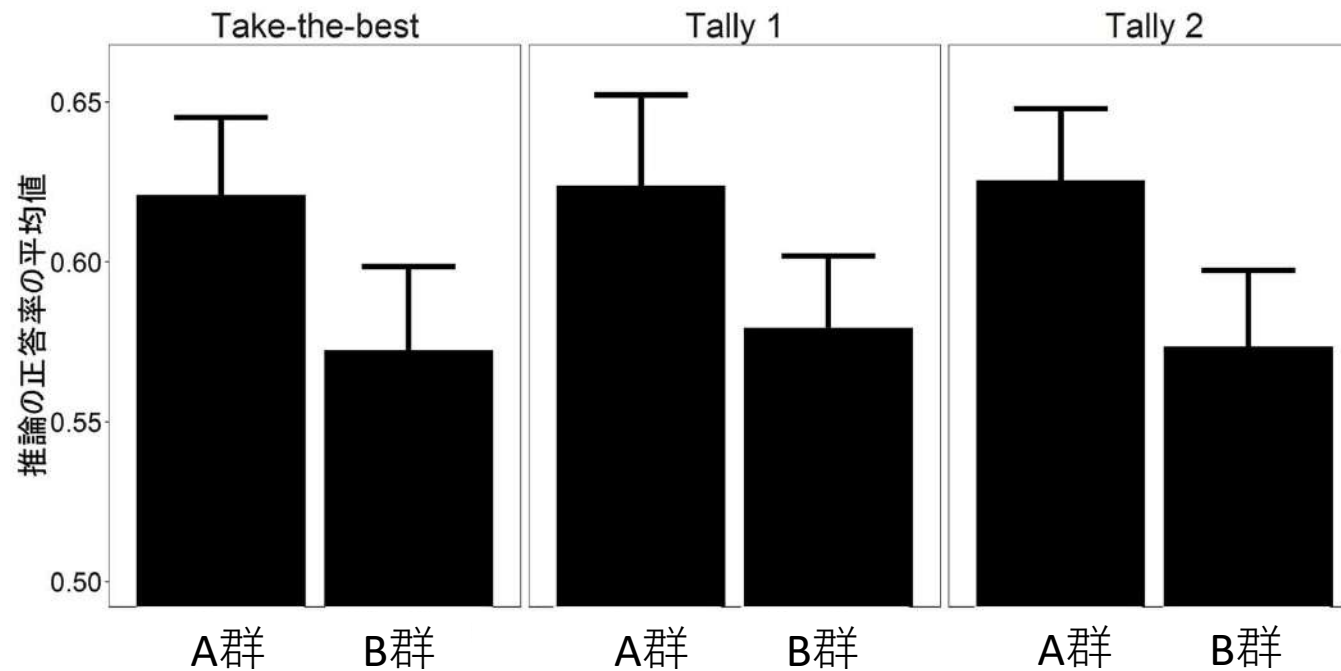


- 記憶に基づく誤信念が実世界に関する正しい理解を促進する可能性
 - 逆説的な関係が成立するかどうかをシミュレーションによって検証

検証方法

- 人口推定課題を用いた計算機シミュレーション
 - 2都市が呈示され、人口数が多いと思う都市を選択する
 - 推定：前述の5属性に関する知識を用いて、TTB (Gigerenzer & Goldstein, 1996), Tally1, Tally2 (Marewski & Schooler, 2011) のアルゴリズムを用いて推論をシミュレート
- 以下の2群を設定
 - A群 (A1, A2, ..., A25, n = 25)：行動実験で得られた知識に関するデータ (100都市, 5属性, 計500属性) を用い, **系統的な誤信念**を含む知識データによる推論をシミュレート
 - B群 (B1, B2, ..., B25, n = 25)：A群と同じ割合の誤信念をもつが, ランダムに再構成された**非系統的な誤信念**を含む知識データによる推論をシミュレート

シミュレーション結果



- いずれのアルゴリズムでもA群 > B群
- **系統的な**誤信念は実世界の環境情報を推論する際に適応的に機能する可能性

まとめ

- 記憶に基づく誤信念はランダムに発生しているわけではない
- 環境構造を我々の認知に反映させていく中で生じた「誤り」である可能性
- 記憶に基づく誤信念が実世界に関する正しい理解を促進する可能性

人工知能と比較した場合の
人の選好推定の特徴とは？

Society5.0の実現にとって必要な事柄

- 内閣府が提唱するSociety5.0で描かれているAI
 - 「AIにより、必要な情報が必要な時に提供されるようになり、ロボットや自動走行車などの技術で、少子高齢化、地方の過疎化、富の格差などの課題が克服される」ことが重要

(http://www8.cao.go.jp/cstp/society5_0/index.html)

- これを実現するには
 - AIが人の心理状態（意図や選好）を読める
 - 「**心の理論**」(Theory of Mind) のモデル化
 - 推定した人の意図に寄り添った出力をAIが行える



人と人のコミュニケーションを対象に
この基本的なメカニズムを明らかにする

科研・新学術領域「認知的インタラクシオンデザイン学」

C01, C02班



人工物デザイン方法論の確立
人の適応性を支える環境知能
システムの構築

人と人工物の協創

自然で持続的な人＝人工物インタラクシオンを実現する
人工物デザインの確立と環境知能の実現

認知的
インタラクシオン
デザイン学

人-人インタラクシオン



- ・ 他者モデルの同定
- ・ 他者モデルの獲得
(学習)機構の解明
- ・ リアルフィールド

人-動物インタラクシオン



A01, A02班

成人同士のインタラクシオン
人工物を介した大人と子供のインタラクシオン

B01班

人＝盲導犬／ウマ／サルの
インタラクシオン



認知的インタラクシオンデザイン学



限られた事例からの他者選好推定

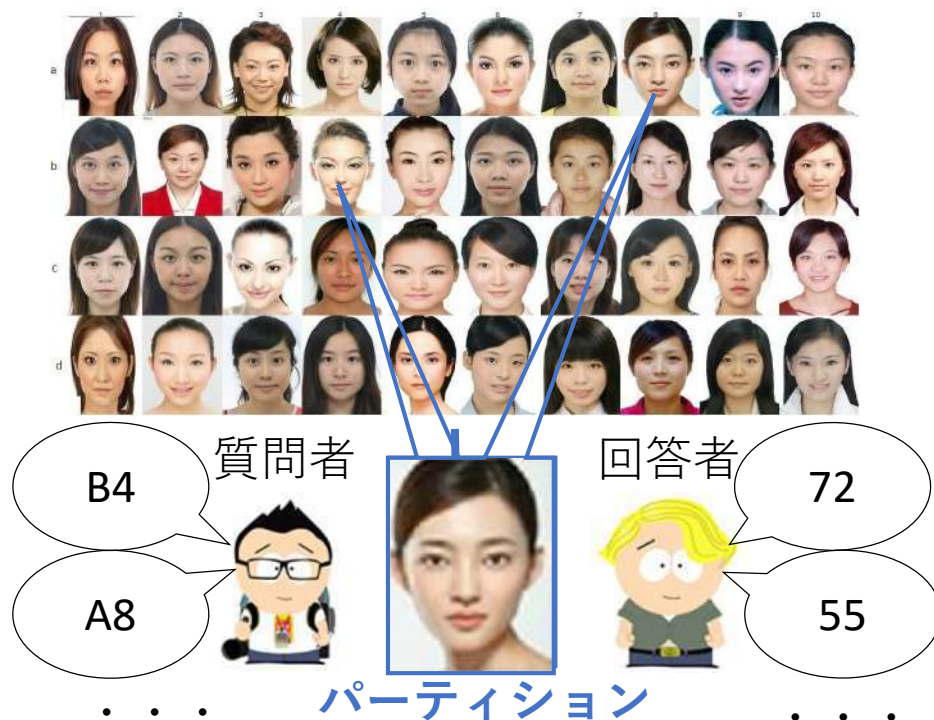
- **人に固有**な他者選好推定の特徴に関する検討
 -  AI (e.g. Sentiment Analysis) と人の違い
 - **限られた事例**から対話相手の選好を推定可能？



- 顔画像に対する相手の選好を推定する実験
 - 計算論モデリングの観点から人とAIの違いを明確化

選好推定実験 (福田・足立原・植田, 2017, HCG2017)

質問フェイズ



テストフェイズ

	質問者 推定点	回答者 回答点
10枚	58	52
	80	84

独立に回答

質問フェイズ

質問者は40枚の顔画像 (Xie et al., 2015) の中から、好きな回数だけ質問をして相手の好みを尋ねる。

テストフェイズ

新たな10枚の顔画像に対して、回答者は回答点、質問者は推定点を答える。

事後回答フェイズ

質問者がいない状況で、回答者はすべての顔画像の好みを回答。

分析のポイント

1. 限られた事例から精度良く推定できていたのか？

- **probably yes**：質問数の平均:5.13, 回答者の回答点と質問者の推定点に差がない（同等性検定）



質問者の推定精度を
ideal observer model（ベイズモデル）の推定精度
と比較

ideal observer model

モデルの概要

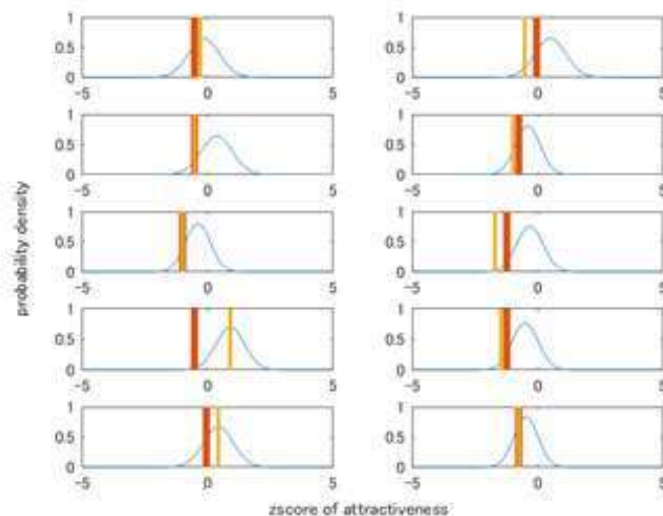
- 一般的な選好の分布の中での、質問フェイズの画像に対する回答者の選好の位置から、テストフェイズの画像に対する回答者の選好を推定するモデル
 - 質問フェイズでの質問画像 $A = \{A_1, A_2, \dots, A_n\}$,
回答者の対応する回答点を $X = \{x_1, x_2, \dots, x_n\}$ としたとき,
テスト画像 T_1 の得点の確率分布は
$$I(T_1 = x | A_1 = x_1, A_2 = x_2, \dots, A_n = x_n)$$



クラウドソーシングサービス
上での529名の選好評価

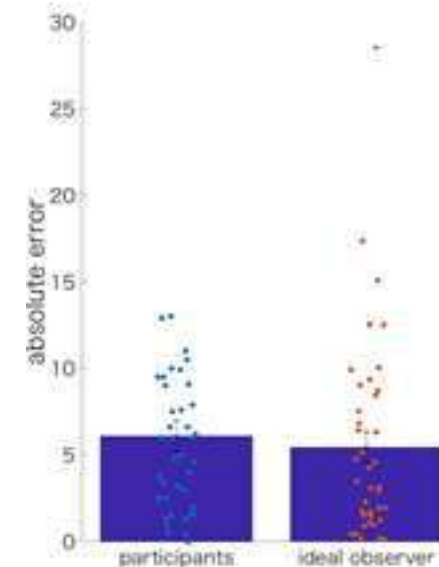
推定精度の比較

回答点と推定点の一例



黄：回答者の回答点
赤：質問者の推定点
青：同時確率モデルの推定分布
（このピークが推定点）

質問者の推定点および
ideal observerの推定点の、
回答者の回答点との絶対値誤差
（エラーバーは標準誤差）



質問者の推定精度は、上限的なモデルと遜色なかった。

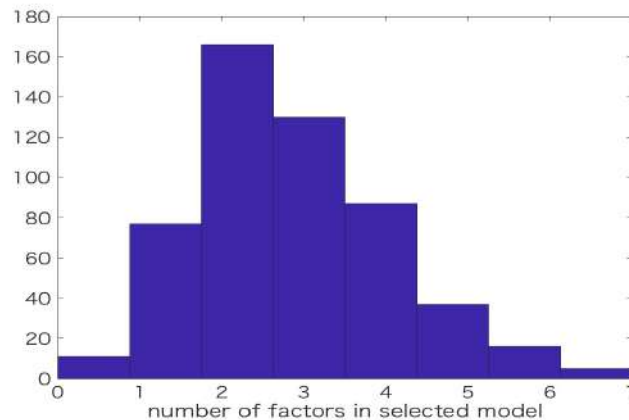
分析のポイント

1. 限られた事例から精度良く推定できていたのか？
 - yes
2. 実際に人が用いていた推定方略とは何か？
 - ベイズ推定？
 - （質問者が）スパース性を仮定した（＝用いる因子を絞った）推定を行っていたのではないか？

好みの評価軸

- 一般的な選好の評価点を因子分析し，選好に関する評価軸（因子）を抽出 ➡ 7つの評価軸（固有値1以上の固有ベクトル）
- 7つの評価軸＋一次の交互作用項の28の特徴量（因子負荷量）で，回答者の回答点を個人ごとにステップワイズ回帰して有意だった特徴量を抽出

選択された評価軸の数と頻度



Ave: 3.12, SD: 1.22

回答者の回答点は**3個**程度の主要な評価軸で説明される

質問者が行っていたこと（モデルによる推定）

- **質問フェイズ**

- n 個（質問回数）の画像について回答者に質問



①抽出

スパース性（少数の因子）
を仮定

回答者の回答点を説明するのに
適した**評価軸（因子）**



②予測

- **テストフェイズ**

- 10個の画像に対する評価（推定）



圧縮センシング（スパースモデリング）で説明

結果の概要

- 圧縮センシングによるテストフェーズ画像**10個**に対する点数の推定
 - 評価軸（因子）の数： $m = 28$
 - 使用する因子数： $k = 3$
- 実際の質問者の推定と推定誤差（絶対誤差）を比較



質問者の推定と
圧縮センシングによる推定
では同程度のエラー

Bayesian rationality との関係

- Baker, C. L., Jara-Ettinger, J., Saxe, R., & Tenenbaum, J. B. (2017, *Nature Human Behavior*)
 - 比較的単純な環境での他者の選好推定をPOMDPでモデル化
- **Bayesian rationality**：人の確率判断はベイズに従う
 - 結果（精度）だけ見ればそう見えても、行っていることは異なっている可能性
 - スパース性のヒューリスティックを利用？
- 人に固有な他者選好推定の特徴に関する検討
 - 計算論モデリングは重要

まとめ

- 人は少ない事例からでも, **ideal observer model** と遜色ない精度で他者の選好を推定可能
- 評価の評価軸を絞っている



スパースモデリング

- 精度だけ見ればベイズ推定で説明可能だが, 実際に行っていることは異なる可能性
- 計算論モデリングを用いて人の行動を理解することの重要性

全体のまとめ

- 人の高次認知（＝思考）の特徴を炙り出すのに，計算論モデリングや統計的な分析が有効
 1. ヒューリスティックは合理的な判断の「みすぼらしい代用物」か？
 2. 人はどういう状況でヒューリスティックを用いるのか？（文脈依存性）
 3. 誤った記憶にも効用はないのか？
 4. 人工知能と比較した場合の人の選好推定の特徴とは？
- AI技術やシステムを構築する際に考えても良いこと
 - 人間にとって「合理的」なこと：生態学的合理性 (ecological rationality)
 - 「機械の知」と比べた場合の「人間の知」の特徴

提案

- 昔のように，認知科学と人工知能とで協力して新しい研究領域を作ることはいできないか
 - 認知モデリング研究会 **by** 大森隆司先生（玉川大学）
 - 興味をもたれた方は，是非入ってください。
 - ついでに日本認知科学会へも入会を！（一応，会長なので...）

共同研究者および協力者

- 鮫島和行（玉川大学）
- 松香敏彦（千葉大学）
- 福田玄明（一橋大学）
- 本田秀仁（追手門学院大学）
- 白砂大（東京大学）
- 足立原功太（元東京大学）

ほかの皆さま

発表タイトルの選択理由

辻井潤一・安西祐一郎 (著)

機械の知 人間の知

東京大学出版会

(画像消去)

「機械の知」の実現と
「人間の知」の解明を
目指す人工知能、認知科学
の成果を吟味し、**二つの**
知の特質とその乖離を
議論する。神経生理学、
計算機科学、言語学など
諸科学との関係、ニュー
ロ計算機、コネクショニ
ズムなど知への多様なア
プローチの可能性を探る。

1988年頃から乖離が
問題になっていた？