



コンテキストコンピューティング研究部会 活動報告

2016年9月16日

先端IT活用推進コンソーシアム
コンテキストコンピューティング研究部会
富士ゼロックス株式会社 道村 唯夫

コンテキストコンピューティングとは

活動計画

活動成果

理論、実践、応用

活動の背景と目的

- 背景

- 物事は、人の認知の仕方によって異なった価値や意味を持つ。その価値や意味の違いは、対象とする物事と関連する物事の関係性の違い、つまり"コンテキスト"の違いに現れたりする。IT を使って人がこのようなコンテキストを自由に扱えるようになったら、人々のコミュニケーションが深まったり、人々が協同し活動する成果がもっと豊かになったりするのではないか。そんな思いでコンテキストをコンピュータで研究を行う。この研究活動において、コンテキスト・コンピューティングとは関係性をデータとして記録し、活用する情報活動であり、対象とする関係性は人と人、人と物事、人と環境である。

- 目的

- 近未来の情報社会をビジョンとして描き、コンテキスト・コンピューティングにより個人と社会のインテリジェンス (Social Intelligence) が階層的に連動する情報基盤を提言する。

現在の課題

情報爆発

「情報取得」から「情報活用」へ

検索エンジンなど
機械だけの処理



😊 大量な処理

😞 意味処理に限界
処理が重い

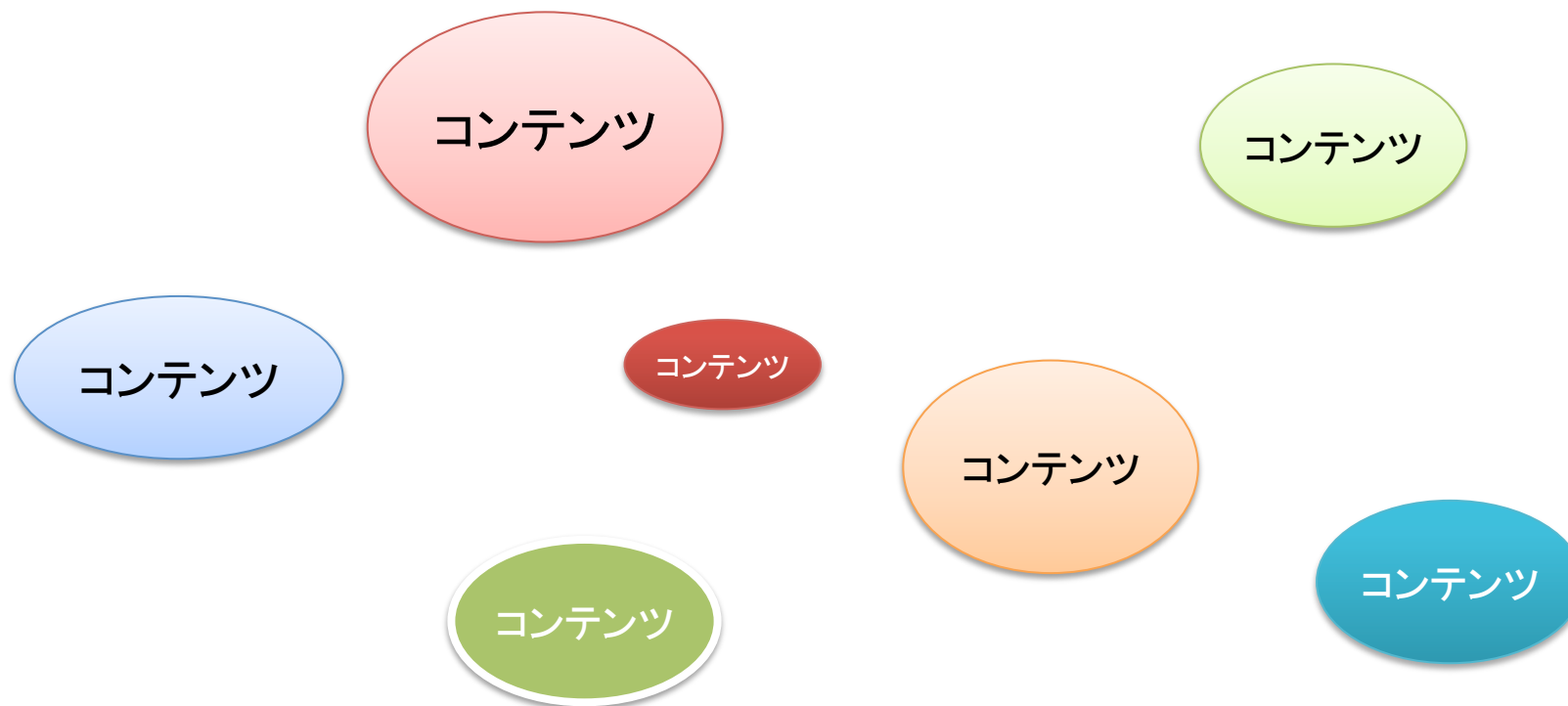
SNSなど
人だけの処理



😊 身近な情報

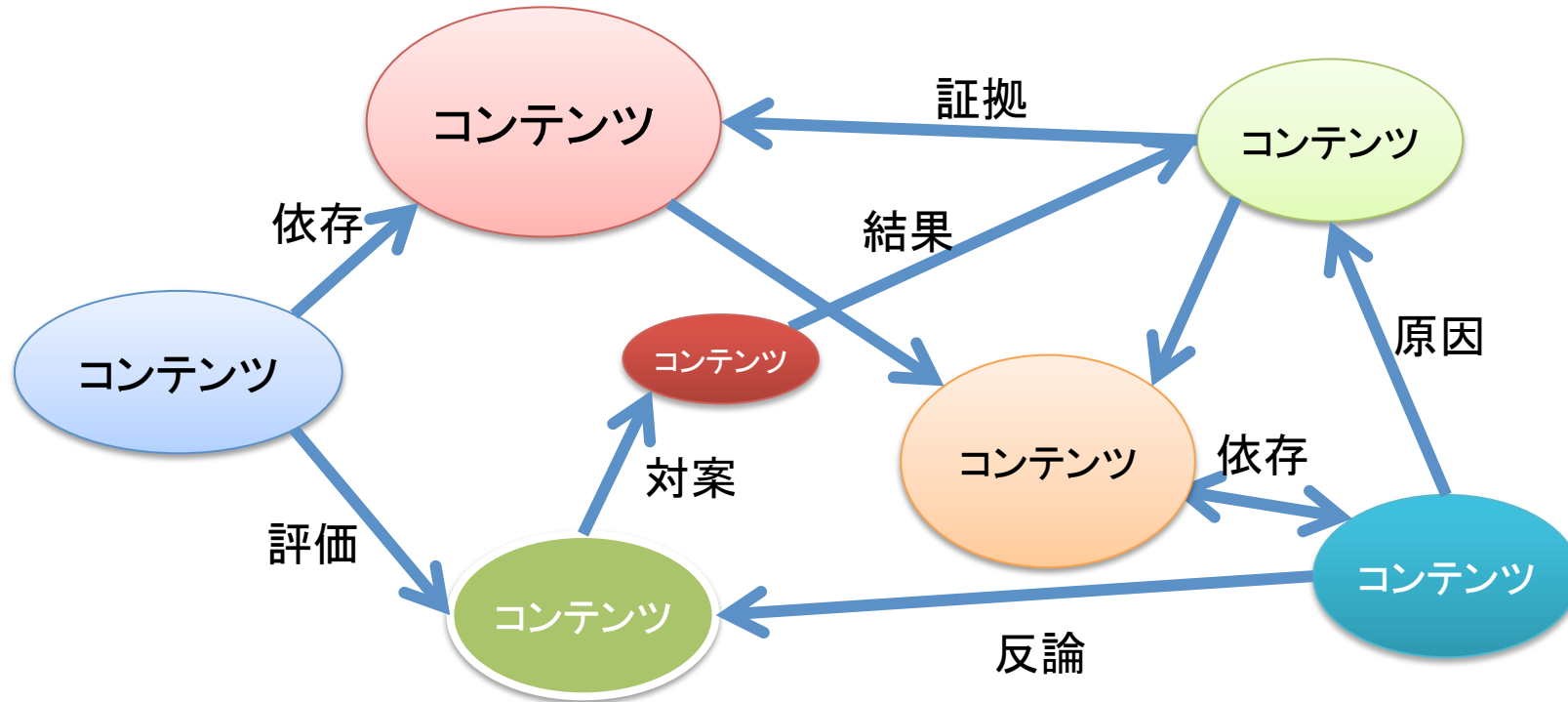
😞 断片的
散在

現在の課題



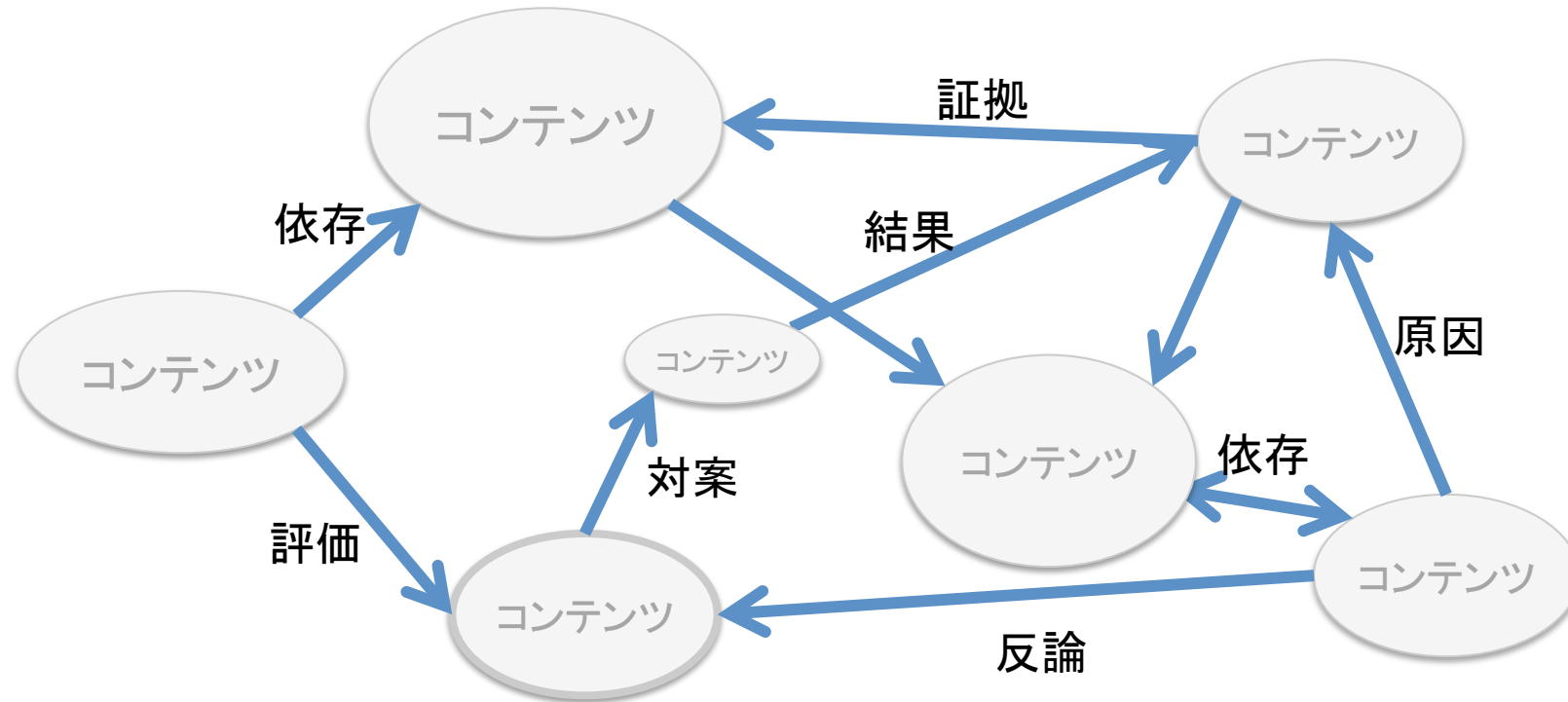
- 多種多様で散在した情報(コンテンツ)をいかに効率的に処理し、意味を抽出するか？

現在の課題



- 多種多様で散在した情報(コンテンツ)をいかに効率的に処理し、意味を抽出するか？

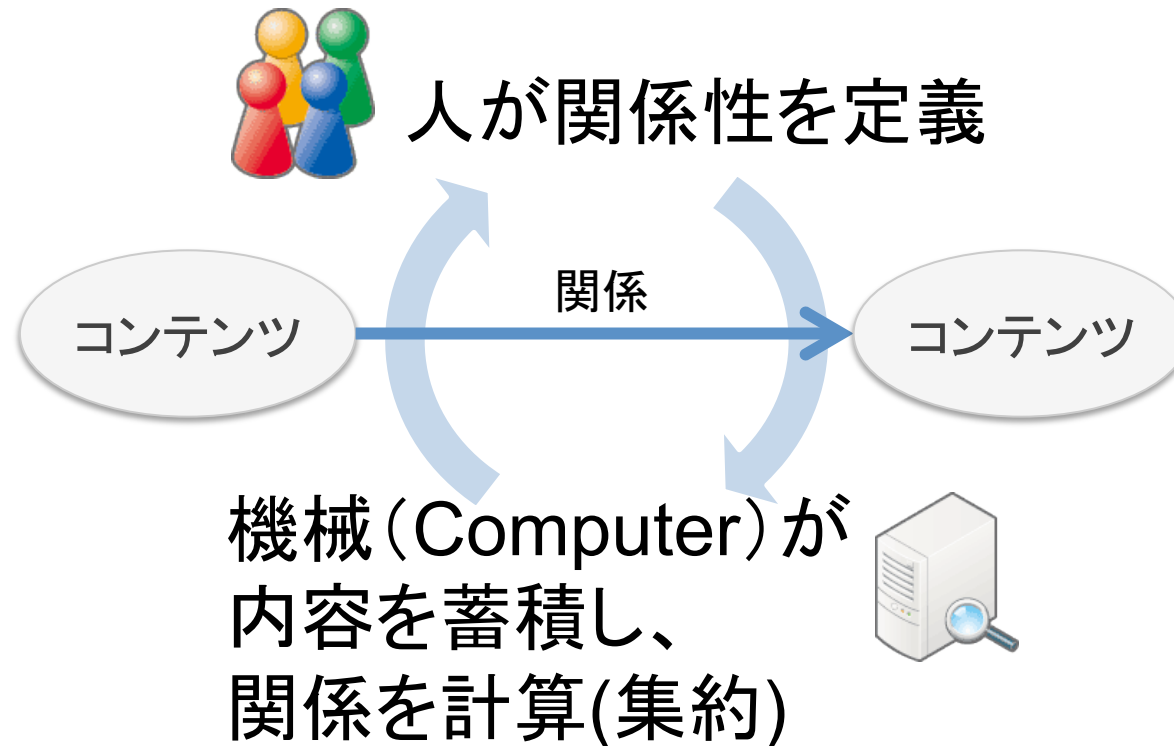
コンテキストコンピューティング



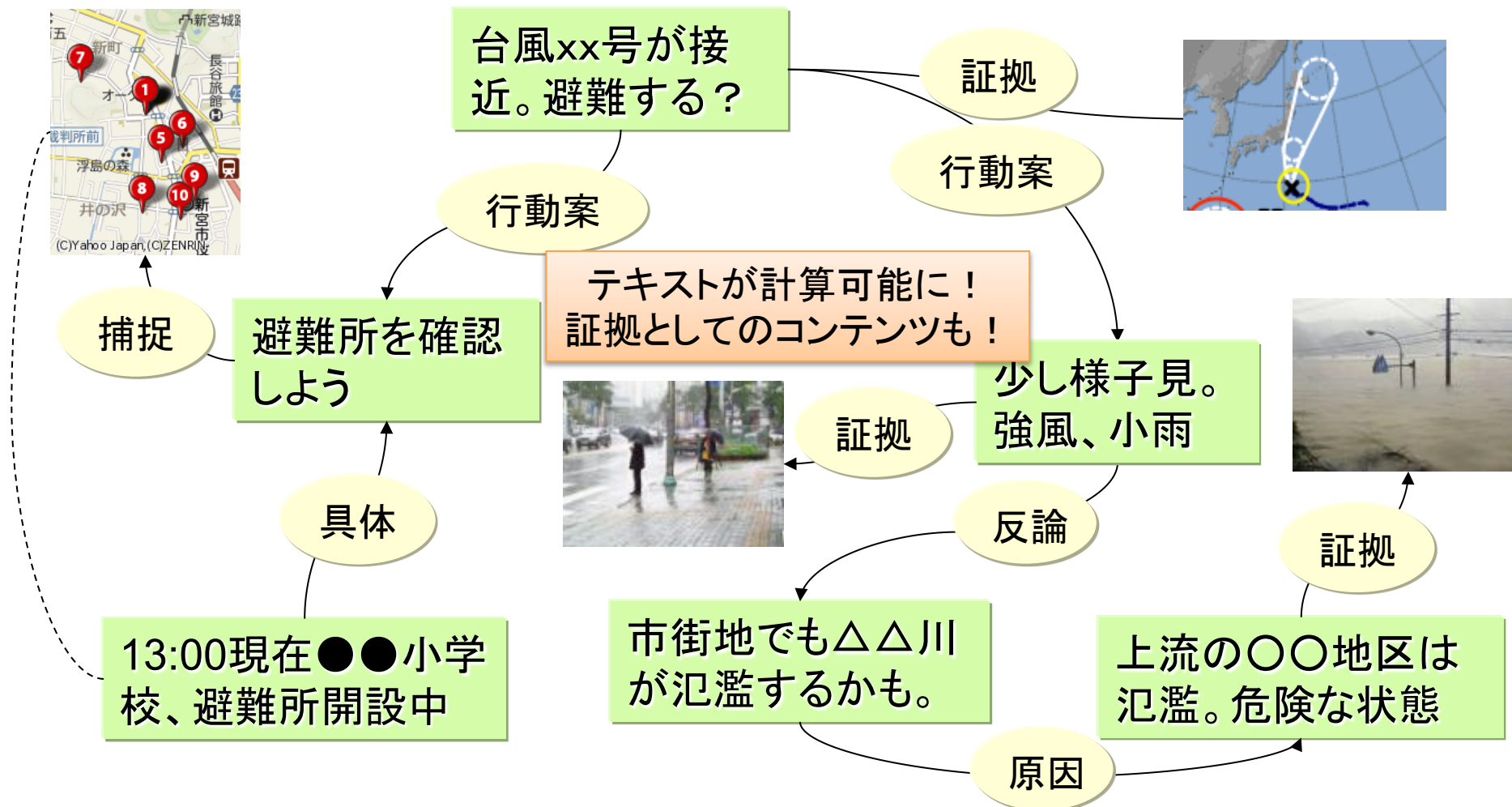
- コンテンツの内容ではなく、**関係性(コンテキスト)**に注目する
 - 依存関係を処理することにより、内容評価と同等の結果が得られるのではないか(仮説)

コンテキストコンピューティング

- Context Computing
 - 人だけでもなく、機械だけでもない
人と機械の協働を目指す



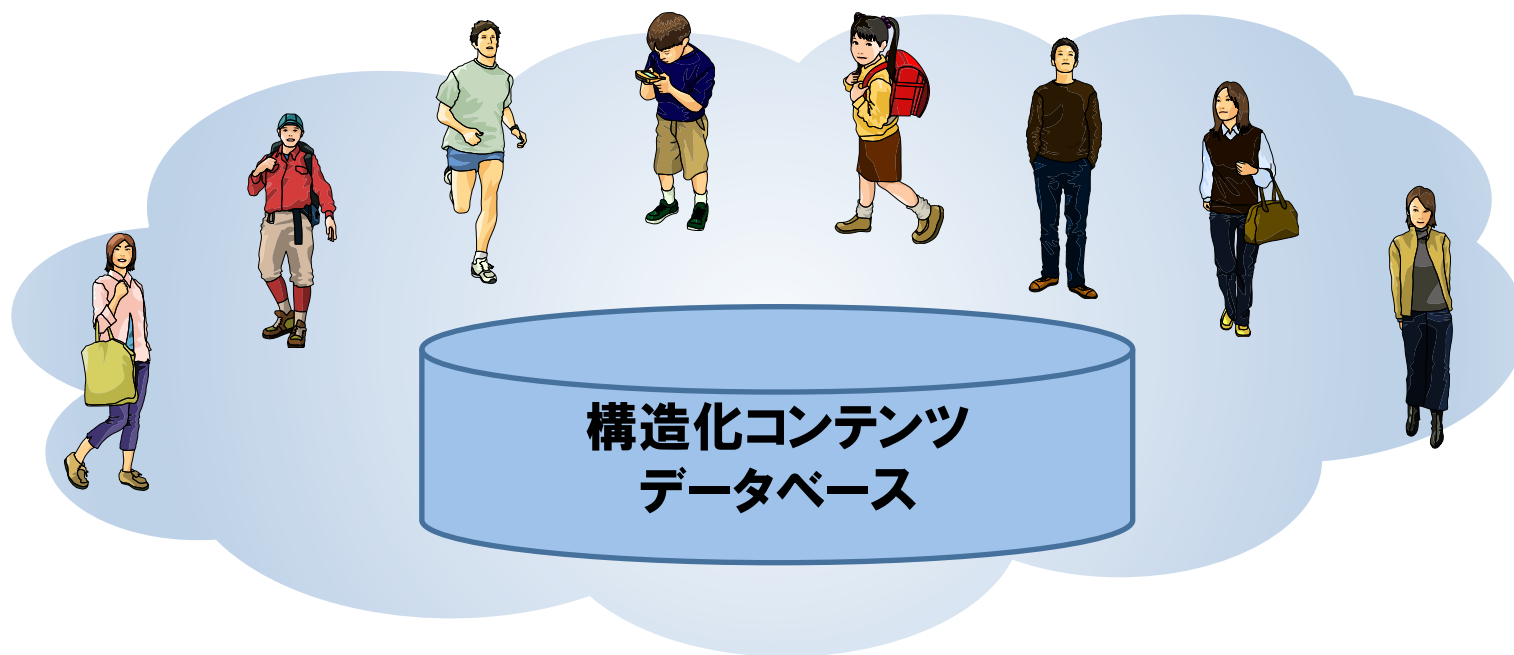
計算可能なコンテンツとは



- コンテンツを構造化することにより、さまざまな技術が適用可能となる

価値共創と知識循環

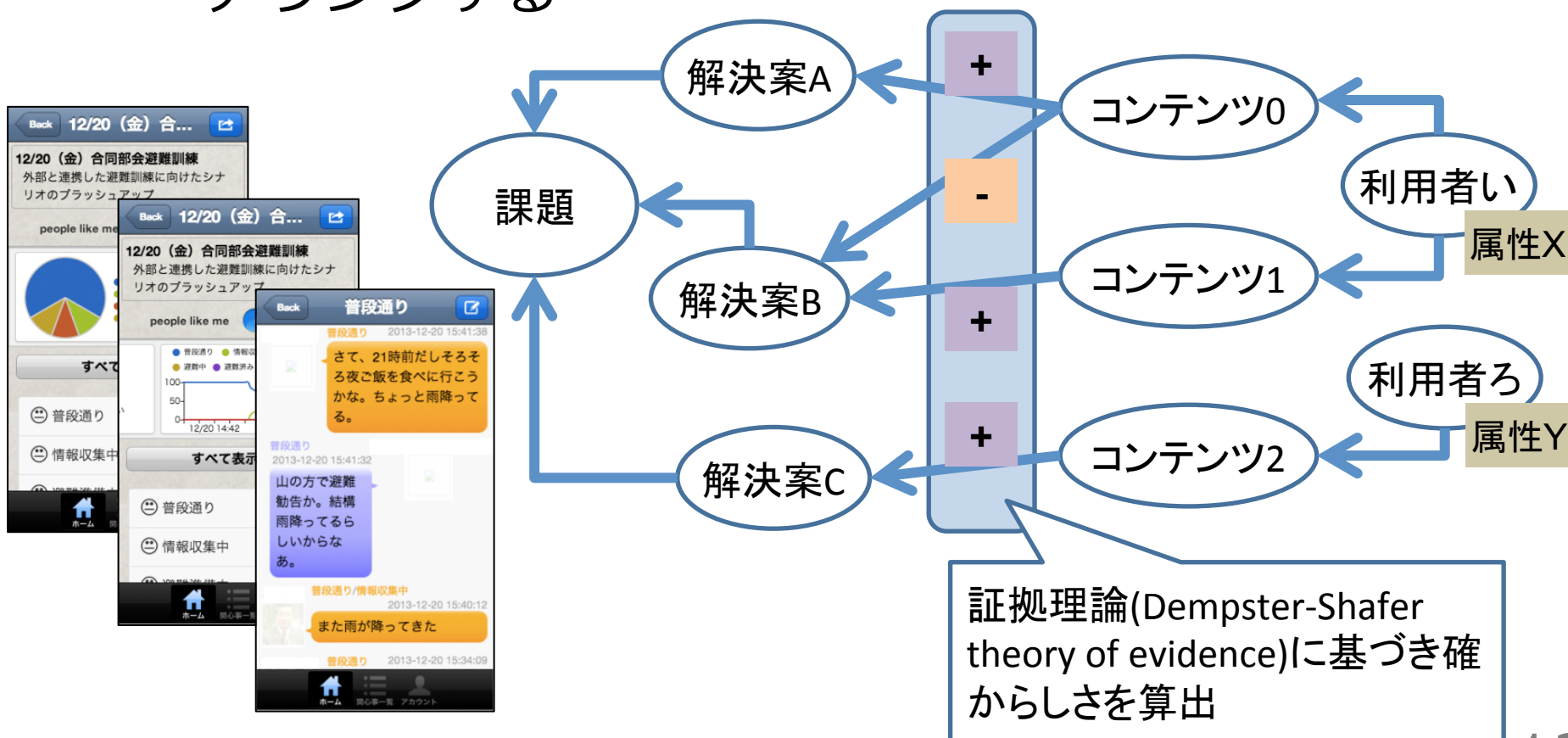
- 構造化されたコンテンツを蓄積し、継続的にコミュニティ全体で共有、共同編集することで、社会との相互作用による価値共創が生まれ、社会全体が知性を持つ



- 社会知としての構造化コンテンツデータベースを構築・活用して、「正解」や「総意」がない課題に対する、「適正解」や「妥協点」を提示するための方法論
 - シナリオの検討 == 仮説の立案
 - 適用技術の調査
 - 検証システムの構築と検証
 - 「空気を読む家」をひとつの応用例として、社会的/個人的合意形成の新たな仕組みを実現する

プロトタイプ

- 関心事にチェックイン
 - 「関心事」という課題に対する解決案を利用者が相互に評価し、機械が解決案に対する評価をスコアリングする



活動計画(進め方)

理論

- コンテキスト・コンピューティング先行研究・関連研究の調査、計算アルゴリズムの検討
- 「合意 (妥協できる合意点)形成」の仕組みの検討

実践

- コンテキスト・コンピューティングの具体化のための技術要素の調査・試用

応用

- 「空気を読む家」、および、社会・ビジネスにおけるコンテキスト・コンピューティングの応用例の検討と提案

➡ これらを検証・実証するためのプラットフォームの検討

活動内容

理論

- 関係を集約するための手法として機械学習を、適正解を見つけるための手法・理論として議論学をスタディ

実践

- 「空気を読む家」におおける、応用例を実現するための技術要素の調査・試用

応用

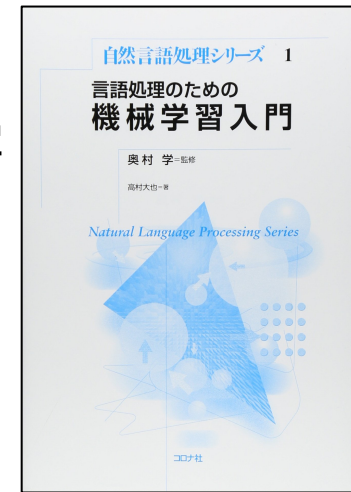
- 応用例として、「早朝会議に遅れない」、「快適な睡眠を得る」、「ルーチンワークの忘れを指摘する」システムの検討

➡ プラットフォームとしては、空間OSを想定

機械学習

- 関係を集約するための手法として、機械学習についてスタディ
 - 「言語処理のための機械学習入門」を輪講
 - 「異種混合機械学習」の紹介
 - 「遺伝的ファジィ決定木」の紹介

- 学んだこと
 - 集約の手法として、機械学習の技法は活用できる
 - 関係の分類やクラスタリングには大いに活用可能
 - コンテンツ間の関係の量や構造から、コンテンツの信頼度や寄与度、異常度などが計算できそう
 - 関係やメタデータのベクトル化、モデル化がキー
 - 対象となる問題領域に応じたロジックが必要となる
 - 高度な意思決定(結果)のロジックがわからないと納得がえられない → 異種混合機械学習



「言語処理のための 機械学習入門
(自然言語処理シリーズ)」
高村大也著、奥村学 監修、
ISBN-13: 978-4339027518

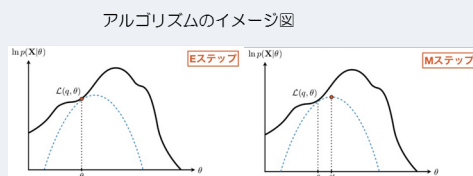
技法

- クラスタリング
 - 似ているもの同士を一つのグループにまとめる
 - 距離空間 (非負性、非退化性、対称性、三角不等式)
 - n次元ベクトルの類似度: 余弦類似度、相関係数

名称	概要	備考
凝集型クラスタリング (単純連結法)	ふたつのクラスタの中で最も近い事例対の類似度をクラスタの類似度とし、近いクラスタを連結する	クラスタが長く伸びてしまう
凝集型クラスタリング (完全連結法)	ふたつのクラスタの中で最も遠い事例対の類似度をクラスタの類似度とし、近いクラスタを連結する	クラスタが長く伸びにくい
凝集型クラスタリング (重心法)	ふたつのクラスタに対し、クラスタの重心間の類似度をクラスタの類似度とする	
k-平均法	無作為に代表ベクトルを決め、代表ベクトルへの近さでクラスタへの帰属を決め、新たにクラスタ内の代表ベクトルを算出し、その近さでクラスタの生成を繰り返す	初期値によって結果が変化する

・ クラスタリング

名称	概要	備考
混合正規分布によるクラスタリング	事例がどれか一つのクラスに帰属させるのではなく、確率的に分け、各事例のクラスに属する確率分の寄与を考慮して代表ベクトルを計算する	
EMアルゴリズム	混合正規分布では平均を最大化するが、最尤推定法を用いて(対数尤度を)最大化する <ul style="list-style-type: none"> アルゴリズムの単純性と収束の安定性、適用モデルの汎用性が利点 多項分布、ポアソン分布、正規分布などの指数分布族の範疇にはいるモデルであれば、算出が容易 イェンゼンの不等式により、EMアルゴリズムは収束する 	確率的潜在意味解析(PLSA) 潜在的ディリクレ配分法(LDA)



－ 問題点と注意点

- ベクトルを作成する際、意味的なクラスが必要なら、意味的な性質を表す組成を使う
- クラス数をいくつにするかについての一般的な解はない
- 繰り返しに基づくクラスタリング手法では結果が初期値に依存するので、乱数を使っていくつかの初期値で計算を行い、平均的な結果を調べる

技法

- 分類
 - あらかじめ決められたグループに分けること

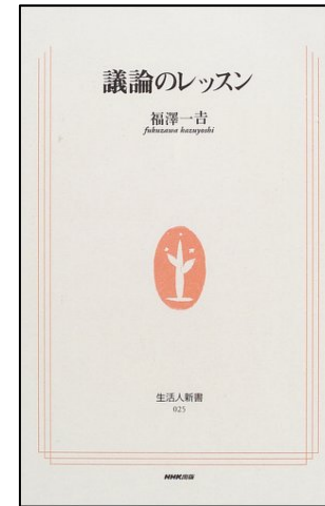
名称	概要	備考
ナイーブベイズ分類器	確率モデルに基づいた分類器で、事例dに対して、 $P(c d)$ が最大となるクラスを出力する 確率 $P(c d)$ を求めるために工夫をする	素性の頻度を数えているだけなので高速
サポートベクトルマシン	事例をよりよく分離する分離平面を、マージン最大化を戦略として算出する	分類性能が高い
カーネル法	カーネル関数(事例の対を引数として、高次元の内積を与える関数)を用いて学習や分類を行う	SVMと組み合わせると非線形で複雑な問題に対応
対数線形モデル	確率的な分類器で、 $P(y d)$ を正則化して最急勾配法などを用いて直接モデル化	
素性選択	ラベル付き訓練データと素性集合が与えられたとき、自己相互情報量や情報利得、 χ^2 乗値から自動的に選択する	

- 系列ラベリング
 - ある系列の各要素に適切なラベル列を付与する

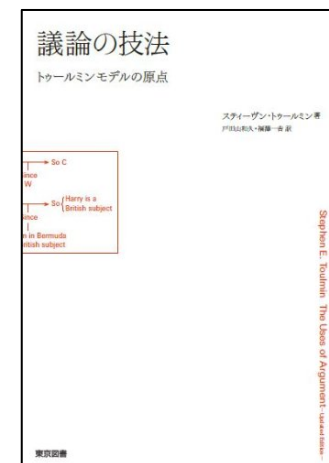
名称	概要	備考
隠れマルコフモデル	各状態はその直前の状態にのみ依存することを仮定し、各状態間の同時確率から、動的計画法を用いて、ラベルを推定する	分類器の逐次適用では、計算時間は長くなるが高精度
条件付確率場	対数線形モデルを系列ラベリング問題に適用したもの	

合意形成

- 「正解」や「総意」が得られない課題の「適正解」を見つけるための手法・理論をスタディ
 - Wikipedia「議論学」を議論
 - 「議論のレッスン」を輪読
 - 「議論の技法」を紹介
- 背景
 - 昨年度の「集合知とは何か」のスタディから、
 - 生命体が生きるための実践活動と切り離せないものが「知」である
 - 「正解」や「総意」はなく、「適正解」と呼べる
 - 社会では、対話と観察を通じて社会としての知を構築する
 - 主観的な知が、コミュニケーションの記憶の蓄積によって、上位の社会的な階層において意味構造＝知を発生させる
 - 近未来のITの姿とは、
 - ローカルな社会集団のコミュニケーションを身体的、暗黙知的な部分から活性化し、集団的な知をまとめあげるマシン



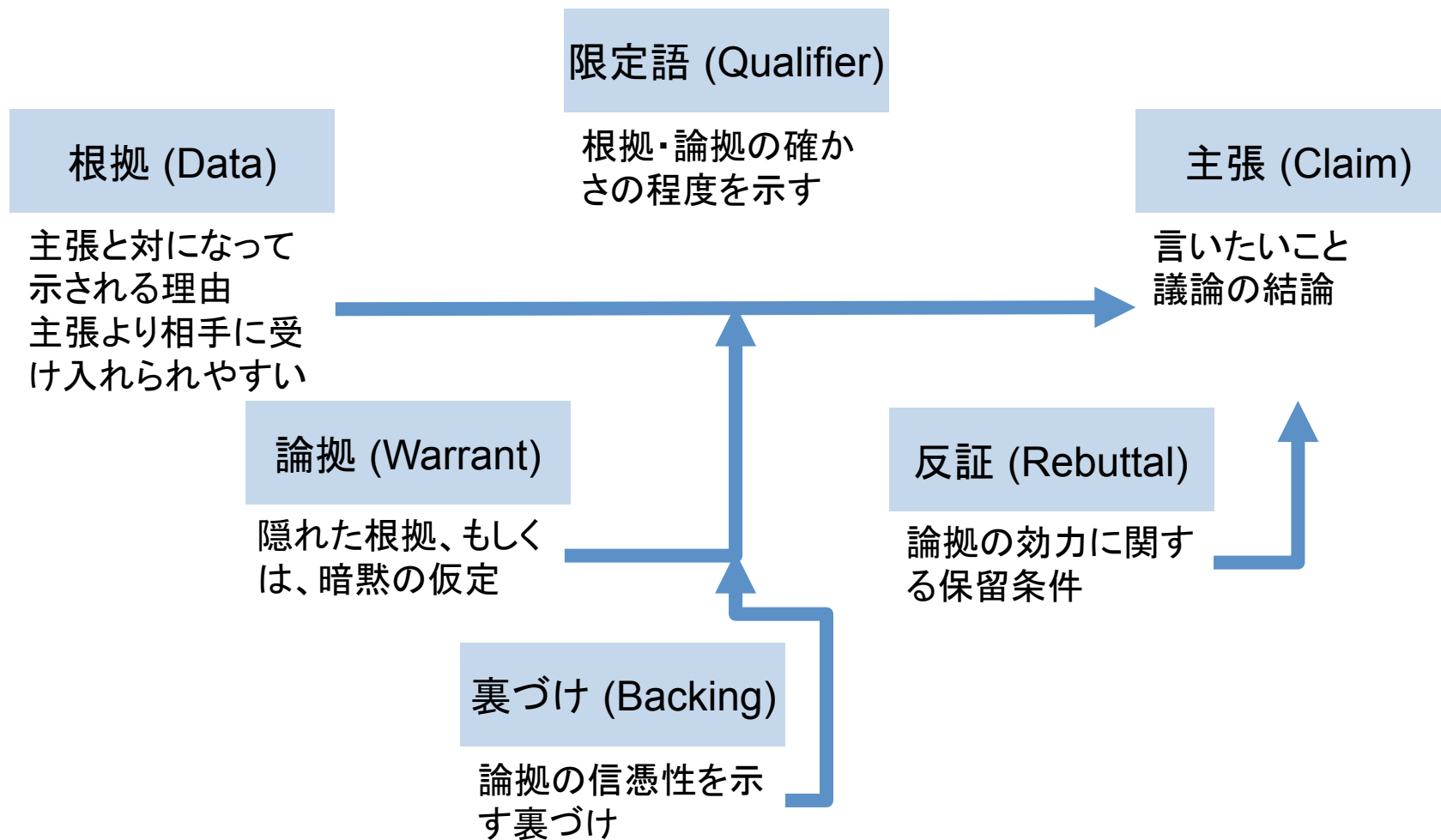
「議論のレッスン」
福澤 一吉著、
ISBN-13: 978-4140880258



「議論の技法」
スティーヴン・トウルミン著、
ISBN-13: 978-4489020940

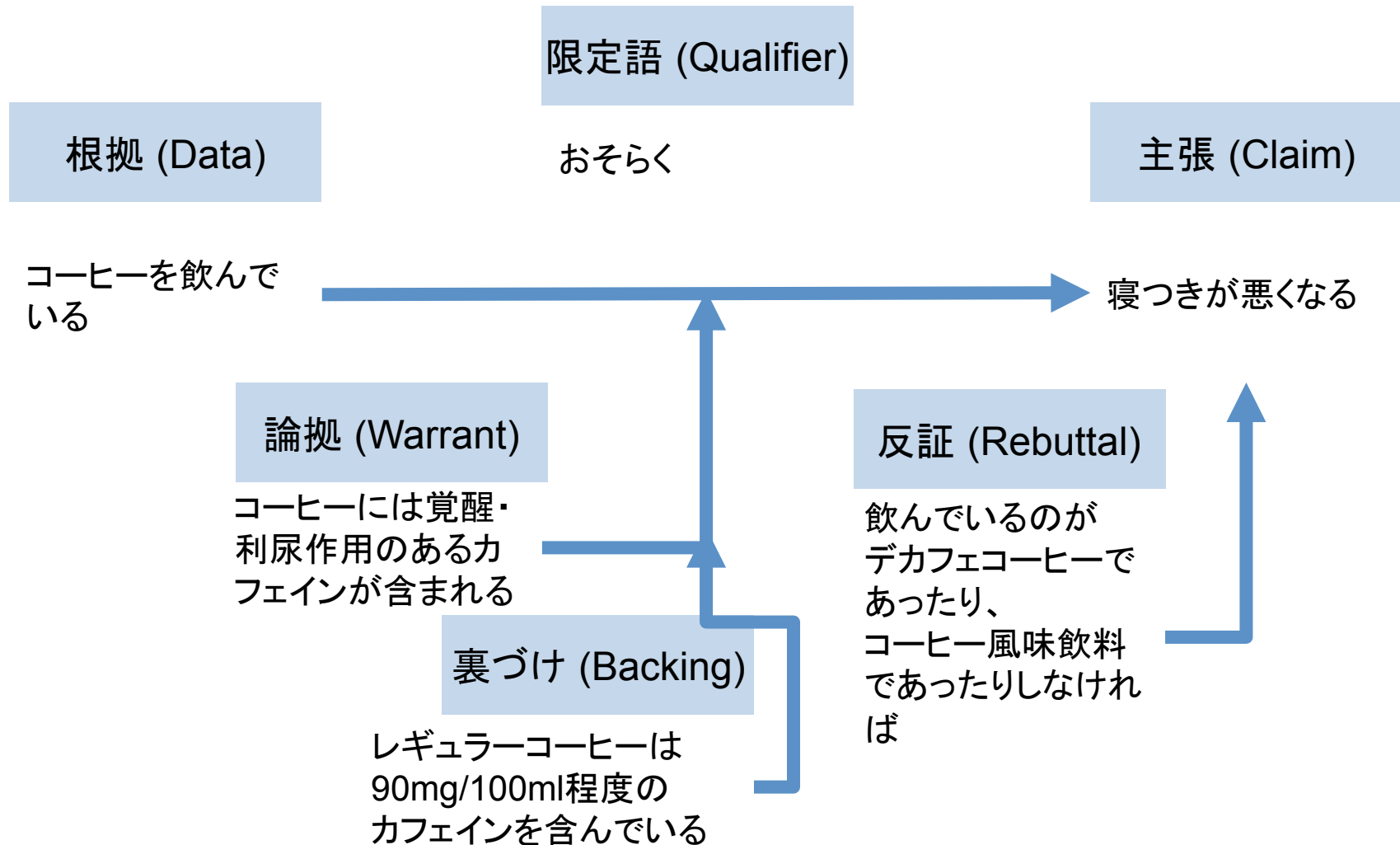
トールミン・モデル

・ トールミンの議論モデル



トールミン・モデル

トールミンの議論モデル



重要な会議に遅れない

- 2月のIT Forumにて



コンテキストコンピューティング

- 「重要な早朝会議に遅れない」を解決する
 - 会議の場所、時間、経路などは知っているとして、..
 - 蓄積された情報（行動）の関係性から、さまざまな経過や結果を予測し、適正解を導き出す
 - つまり、..
 - 一般的な情報間関係性は事前にモデル化が可能
 - 一般常識として、もしくは、機械学習などによって
 - 空間OSに蓄積された情報間の住人に関する特異な関係性は、住人に提示してもらい、この関係性をもとに洗練された「適正解」を導く
 - 感情認識など、自然なユーザインタフェースで
 - 例えば、前夜に気温が低く、酒を飲み過ぎて、重要な会議に遅れてしまったとき、「酒の飲み過ぎ」が「寝過ごし」につながり、「遅れ」となった一連の行動間関係性を提示してもらい、次の重要な会議に遅れないよう予測して、適切な行動を促す

重要な会議に遅れない

- 「家」の妖精が相談しています



太郎さんは、ずいぶんと早く寝たようだね。

明日のスケジュールには7:30から早朝会議と記載されているからかな。「重要」ってタグも付いているし。



なるほど、場所はいつもの会社の会議室だね。それなら、いつもより30分早く起きれば、間に合いそうだね。

でも、天気予報では、早朝は雪が降るとのことだよ。



それだと、電車は混乱しそうだね。それに、風呂に入らずに寝てしまったようなので、その分も早く起きないと、、

この早朝会議では朝食が出るようだから、朝食を家で食べる時間を使ってシャワーを浴びて出る方がよさそう。

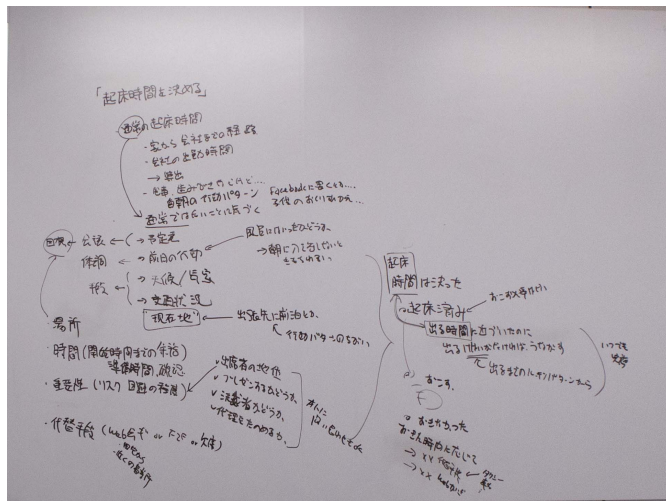
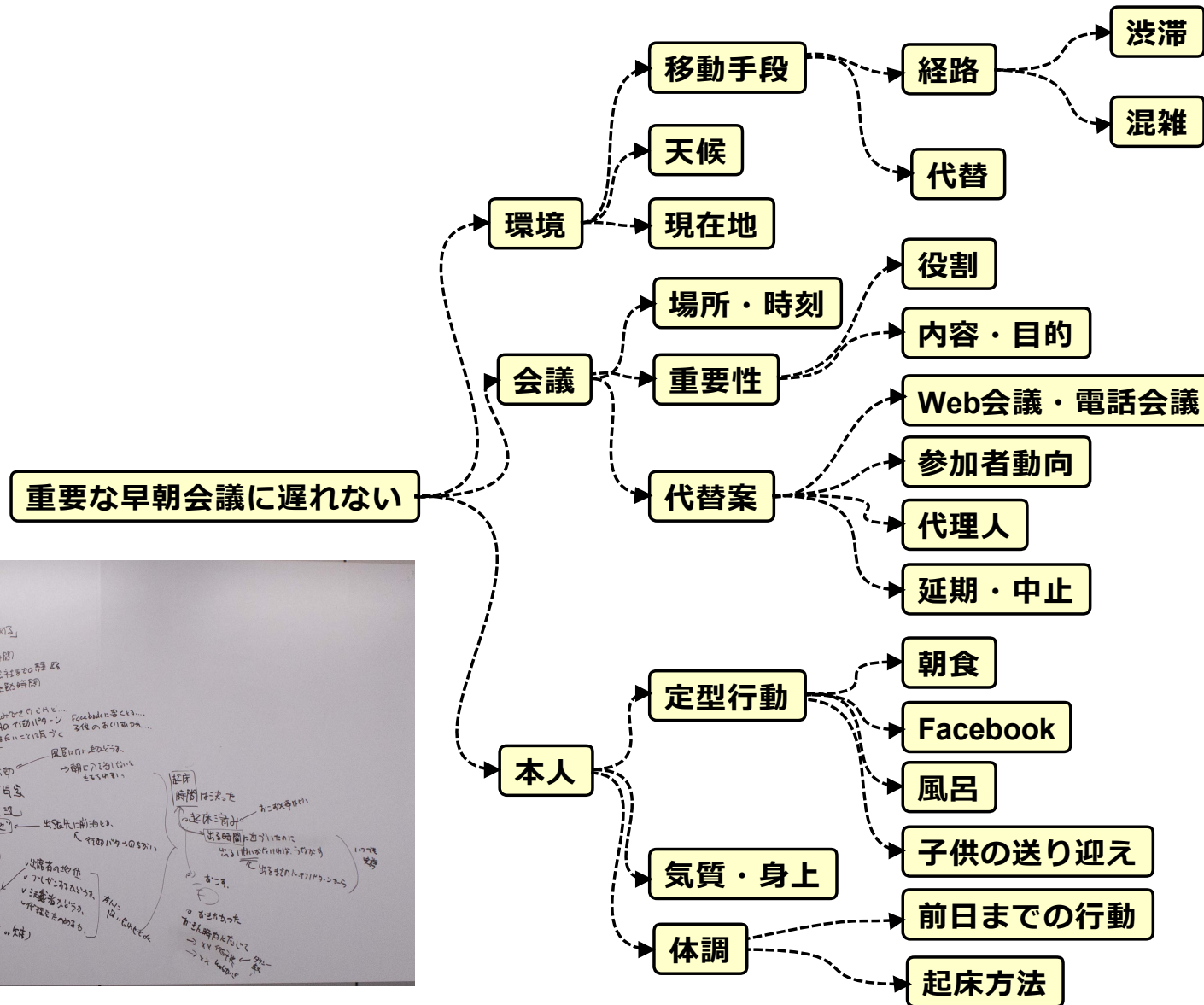


そうだね、前に同じようなことがあったときも、朝食を食べずにシャワーを浴びて出ていったね。

雪が降り始めた。このまま降ると田園都市線は大混雑になるから、小田急線を使うルートで行くことになりそう。



重要な早朝会議に遅れない



重要な会議に遅れない

- 「重要な会議に遅れない」ために「空気を読む」
 - 常識の範囲のルールセットが複雑
 - 考慮すべき要素が多い
 - 情報の構造が複雑
 - 情報元が多様
 - 長期的な時系列情報が必要
 - 「街」レベルの空間OSで解決？
 - 個人化する内容が多様
 - 生活スタイルや身上などが関連する
 - きちんと調べれば類型化できるのかもしれないが
 - 全体としては、かなり複雑系
 - 人間が下した判断を参考にする(扱う)方が現実的
 - 各自の意見を、集約して提示
 - 同意した判断から、集約ルールを強化学習していく

ルーチンワークの「忘れ」防止

- 日常生活における実行すべきルーチンワークの「忘れ」を防止する
 - ゴミの日にゴミを出すのを忘れる
 - 定期的に飲んでいる薬の飲み忘れ
 - 出勤時にSuicaを持っていくのを忘れる
 - 家の中の行動を記録しておき、逸脱(異常)を検知し、指摘する
 - 「トイレでの滞在時間が長くなった」ことも検知
 - 何か病気かも？
 - ユーザとのインタラクションを通じて、行動の内容や、非定型的活動に関するルールを獲得していく

ルーチンワークの「忘れ」防止

- アイデア

- 学習

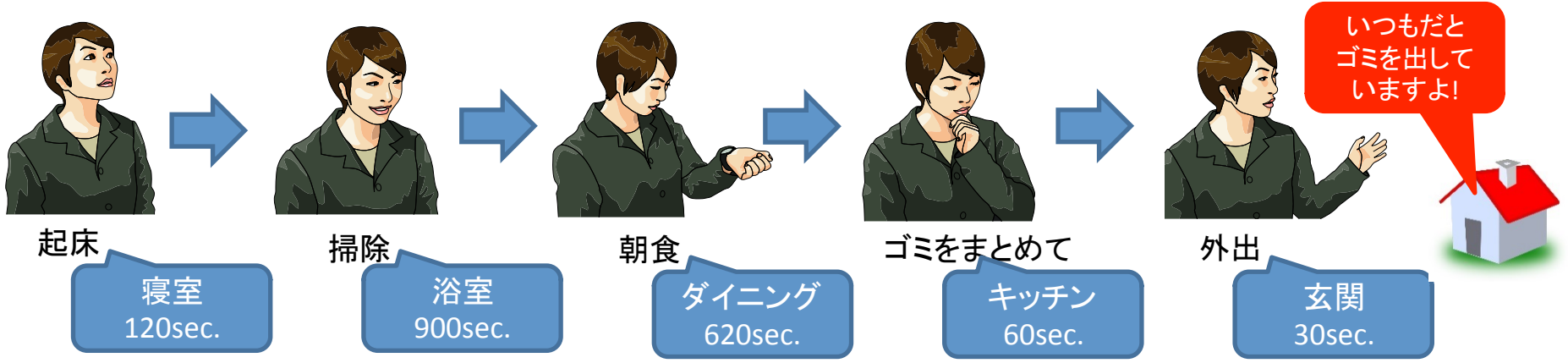
- 人がいつ(時刻)どこ(位置)でどのくらい(時間)とどまっていたかを一つの事例として、時系列に記録・蓄積
 - 事例間での家の中の画像の差分を記録・蓄積
 - 何をしていたのかは気にしない
 - » 場所と時間の流れでやっていることは判別できる(仮説)

- 各事例をストリームとして、クラスタリング

- 起床→外出、帰宅→就寝などで、事例をグルーピングしておく

- 検知

- 頻度の高いストリーム(グループ)の行動パターンに近い行動の中で、逸脱(抜け、過剰など)を検知したら、通知をあげる
 - 通知に対する応答を受け、ルールを強化していく



まとめ

- 活動
 - 理論
 - 機械学習、議論のモデルについて、理解が深まり、コンテキストコンピューティングの中での活用ができると感じられた一方で、課題領域に応じたモデル化が重要であり、困難であることがわかった
 - 実践
 - あまり、調査・検討が進められなかった
 - 応用
 - 「空気を読む家」をテーマに、コンテキストコンピューティングの適用のアイデアを検討した
 - 人間とのかかわりにおいて、ユーザインタラクションが重要であることが、あらためて理解できた
- 今後
 - 個人と社会のインテリジェンス（Social Intelligence）が階層的に連動する情報基盤の提言に向けて、実践に力を入れて活動を進めたい



<http://aitc.jp>



<https://www.facebook.com/aitc.jp>



ハルミン
AITC非公式イメージキャラクター