

第Ⅱ章 現地調査 NTT コミュニケーション科学基礎研究所

調査先	: NTT コミュニケーション科学基礎研究所
住 所	: 京都府相楽郡精華町光台 2-4
開催日	: 2015 年 9 月 18 日
参加者	: 6 名
記	: 西原 雅宏 *

1. はじめに

当委員会では、注目技術の技術開発やビジネス展開を行っている研究所・企業等の調査、見学を行い、会員会社に広く紹介していく活動を行っている。今回その一環として、関西の京阪奈地区に立地する NTT コミュニケーション科学基礎研究所を訪問し、そのいくつかの研究室の研究内容を紹介していただいた。

2. NTT コミュニケーション科学基礎研究所について

同研究所は、人間と情報の深い理解に基づくコミュニケーションの実現をミッションとして、「情報」と「人間」を結ぶ新しい技術基盤の構築に向けて、情報科学と人間科学の両面から研究に取り組んでいる。



Fig. 1 NTT コミュニケーション科学基礎研究所全景
(<http://www.kecl.ntt.co.jp/> より)

3. 研究紹介

今回は、協創情報研究部からビッグデータ関連で 2 件、メディア情報研究部から音響・音声処理関連で 1 件の研究内容を紹介していただいた。

3.1. データの潜在構造を反映する確率モデルの自動生成

業務・センサデータ、購買データ、SNS データなど、大量で多様なデータを解析するための潜在変数モデルを自動的に生成する研究である。

これら大量なデータを統計解析することにより、例えばセンサデータから異常、購買履歴データから顧客ニーズや顧客集団、SNS の行動履歴からユーザの目的やユーザの好みといった情報を発見、予測することができる。これまで、この解析に使用される確率モデルを設計するためには、データと機械学習の双方の知識・技術を有する専門家が必要だったが、データの複雑化に伴い、人手で行うことが困難になってきている。

本研究は、データの階層情報を利用した潜在変数モデルの完全自動生成を目的としている。例えばセンサデータは、場所・日・時間・検出量、購買データは、日時・客・店舗・商品、SNS データは、ユーザ・記事・文・単語などの複雑な階層構造を持つ。これらの複雑なデータから適切な階層構造を自動抽出し、多様な潜

* 技術調査専門委員会委員

在変数モデルを自動生成することに取り組んでいる。

蓄積された大量かつ多様なデータから、有用な知見や知識を自動的に抽出し蓄積するシステムが期待される。

3.2. 多種類のデータから共通するパターンを抽出する複合テンソル¹分解技術

センサやウェブを介して得られる実世界に関する多種類のデータを横断的に解析して実世界を理解する汎用技術の構築に取り組んでいる。

その方法としてテンソル分解によりパターンを抽出し、同一の属性に共通の因子行列を仮定し、データ間で共通する横断パターンが抽出し共通の仮想ユーザを仮定する事ができる。これにより例えば人間の行動パターンの可視化することなどが可能となる。

Yelp²のウェブデータを用いた人間の行動パターンの可視化結果を例として、「平日午後の食事に関する因子」、「日本食レストランのランチ・ディナー因子」を紹介していただいた。

3.3. 大量の事例を用いた超高品質音声強調

実環境で音声を収録すると雑音や残響が含まれるが、それらを取り除く従来の技術（音声強調技術）では、処理音声に歪み（残留雑音・残響、不自然な音）が残り、聞き取りにくい音となってしまうことがあった。

本研究は、あらかじめ大量のクリーンな音声データを用意しておき、観測信号（雑音・残響含む）とクリーン音声を賢くマッチングさせ、適切なクリーン音声に差し替える研究である。

ノイズ除去や通常の音声強調とは違い、処理音声に歪みが残らないので、聴き易い「超高品質音声強調」の実現が期待できる。

実際に修正前と修正後の音声を聞かせてもらいながら、研究内容を紹介していただいた。

4. おわりに

本報では、ビジネス機器・情報システムにも参考となる取り組みを探るべく、NTT コミュニケーション科

学基礎研究所の3つの研究室の研究内容を紹介した。当協会の関連業種では何ができるのか、ヒントを掴んでいただければ幸いである。

また、本研究所では毎年オープンハウスというかたちで研究内容を公開している。機会があれば是非参加していただきたいと思います。

最後に、NTT コミュニケーション科学基礎研究所の研究推進／情報戦略の森様、各研究室の皆様方には、見学会の設定にご尽力をいただきこの場をお借りして厚く御礼を申し上げます。

1 行列や複数次元の配列により、要素間の関係を演算可能な数式として示したもの。

2 2004年に米国で立ち上がったWebサービスで、スマートフォンとの優れた親和性を有するクチコミ情報サイト。2014年4月に日本でもサービスが始まった。

<http://www.yelp.co.jp>

データの潜在構造を反映する 確率モデルの自動生成

石畠正和

岩田具治

NTTコミュニケーション科学基礎研究所

研究背景

大量のデータ・多様なデータ



データの統計解析

- 業務効率化・異常検知
- 顧客セグメンテーション
- 友人の推薦・情報推薦
- 興味推定・コンテンツ推薦



統計的機械学習技術が
注目されている

研究背景

ルール、知識、
知見、判断基準

- 統計的機械学習

- 統計解析で**見えてるもの**から**見えないもの**を発見・予測

- **センサデータ** → 異常
- **購買履歴** → 顧客ニーズ、顧客集団
- **SNSの行動履歴** → ユーザの目的、ユーザの好み

- 確率モデル

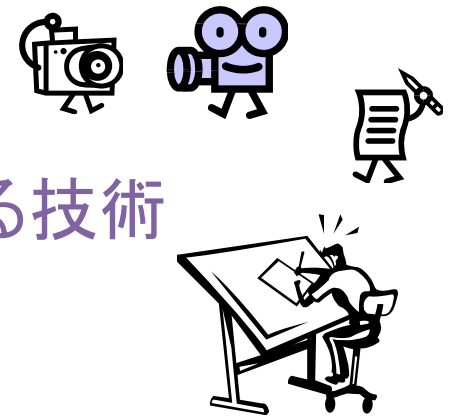
- データの**生成過程**や**依存関係**を表現した確率分布

- **データ**と**目的**に合った確率モデルを**選択・設計**

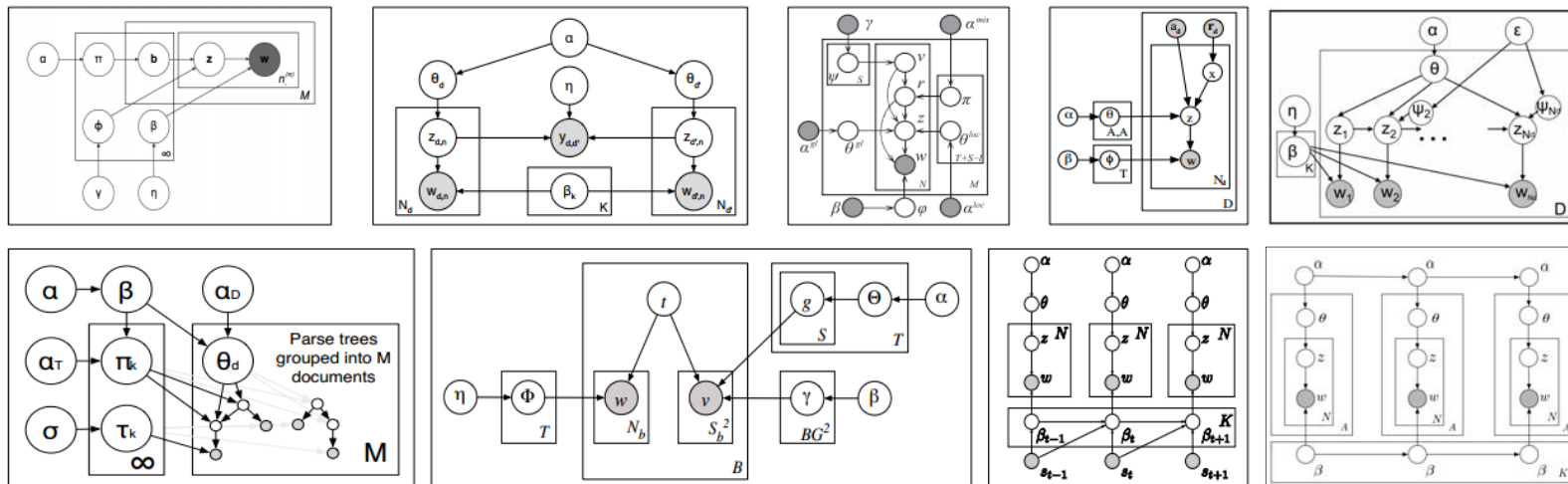
- **データ:** センサデータ、購買履歴、文書、音声...
- **目的:** 分類、推薦、異常値・特異値検出...

研究背景

- 確率モデルでデータ分析をするために必要なもの
 - データに関する知識
 - 解析技術に関する知識
 - 知識を元にモデルを設計、導出、実装する技術



- 複雑化する確率モデル [Blei 11]



研究目的

- 問題

- データがあっても解析ができないジレンマ
 - モデルを設計するための知識・技術がない
 - 人手でモデルを設計するにはデータが複雑すぎる
 - 多様なデータが混ざった複合データ
 - 大量のセンサ群から得られたデータ

- 最終ゴール

- データに適した確率モデルの**完全自動生成**
 - 機械学習の知識に乏しい人でもデータ解析が可能
 - 人手ではモデルが設計できない複雑なデータも解析可能

研究目的

- 最終ゴール
 - データに適した確率モデルの完全自動生成
- 今回のゴール
 - データの階層情報を利用した潜在変数モデルの自動生成

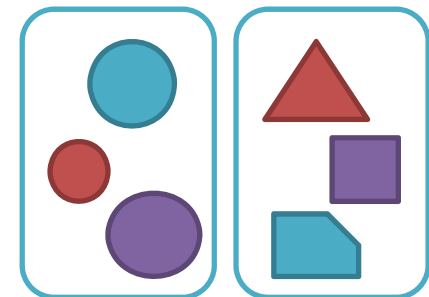
目次

- 研究背景・目的
- 潜在変数モデル
- 関連研究
- データと階層構造
- 提案手法
- 実験
- まとめと今後の計画

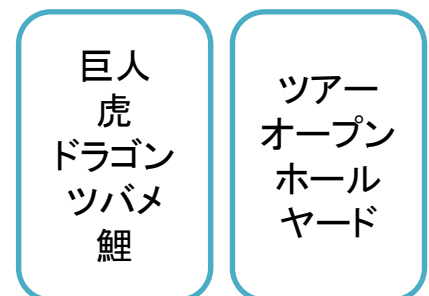
潜在変数モデル

- 潜在変数モデル
 - 観測変数とは別に **潜在変数** を導入した確率モデル
 - データの **潜在的意味構造** に注目
 - データの生成仮定をモデル化するため、**結果を解釈しやすい**
- 潜在変数モデルの例
 - クラスタリング系
 - 解釈: **“似ているデータ”** のグループ
 - 例: 混合多項分布, Naïve Bayes Model
 - トピック系
 - 解釈: **“似た文脈** で現れやすいデータ” のグループ
 - 例: pLSA, LDA [Blei+ 03]
 - トピック + 構造系
 - 解釈: **“文脈”** の **変化**、**依存関係** をモデル化
 - 例: **系列** HMTM [Mark+ 10], HTMM [Boyd-Graber+ 07]
木 STM [Gruber+ 11]

クラス



トピック



関連研究

- モデル選択
 - 複数の確率モデルを比較するための指標 [Akaike+ 73][Rissanen 78][MacKay 92]
 - モデル候補を自動生成する枠組みが必要
- 構造学習
 - Graphical model : 観測変数の **条件付き独立性** を表現 [Cussens+ 13]
 - SRL, PLL : 依存関係を **述語論理** で表現 [deRaedt+ 08]
 - 潜在変数は考慮しない or 限られたモデルクラス上で探索
- 行列分解に基づくモデル生成 [Grosse+ 12]
 - 行列分解を **文脈自由文法** で記述し、モデルクラスを定義
 - 何度も行列分解を行うと、導入された潜在変数を解釈しにくい
- 提案手法
 - データの持つ **階層構造** を利用し、モデルクラスを定義

目次

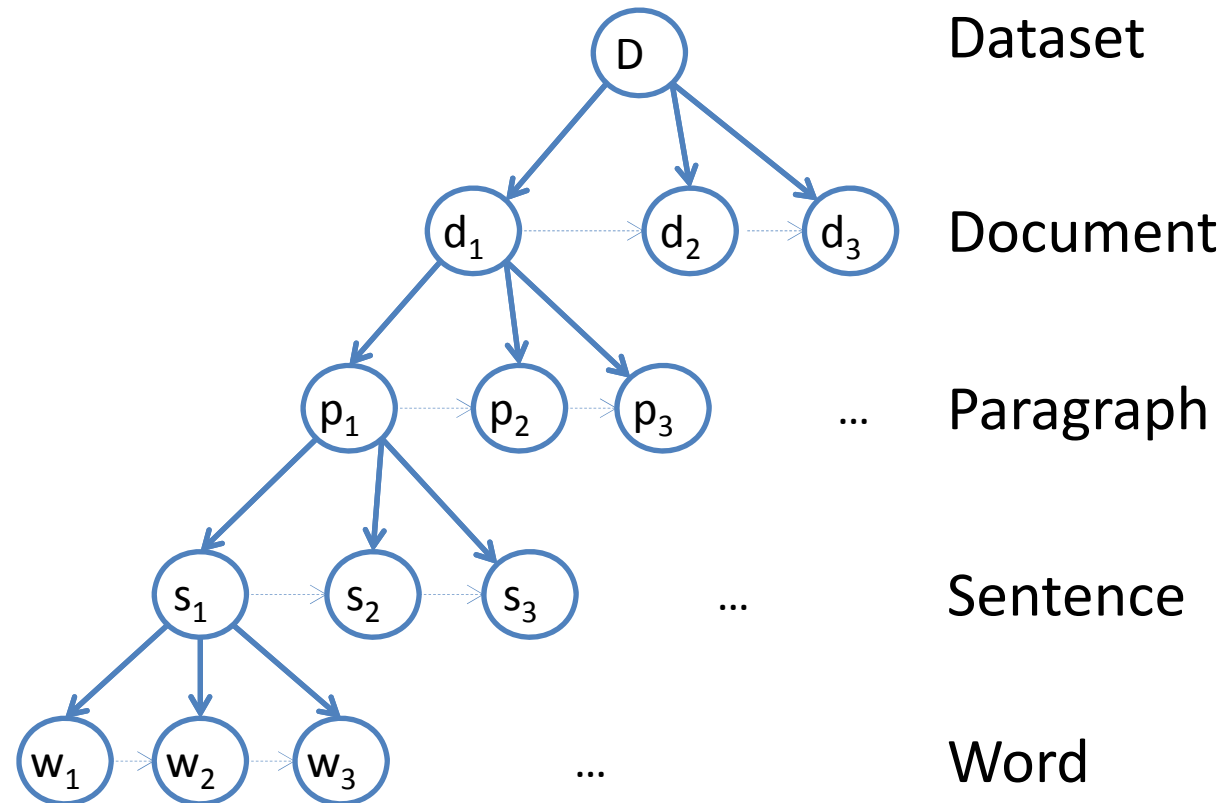
- 研究背景・目的
- 潜在変数モデル
- 関連研究
- データと階層構造
- **提案手法**
- 実験
- まとめと今後の計画

提案手法

- 目的
 - データの階層情報を利用した潜在変数モデルの自動生成
- 提案
 1. データの階層構造を利用したモデルクラス の定義
 2. モデルクラスに対する 汎用な 学習アルゴリズム
 3. モデルクラス上の探索手法

データと階層構造

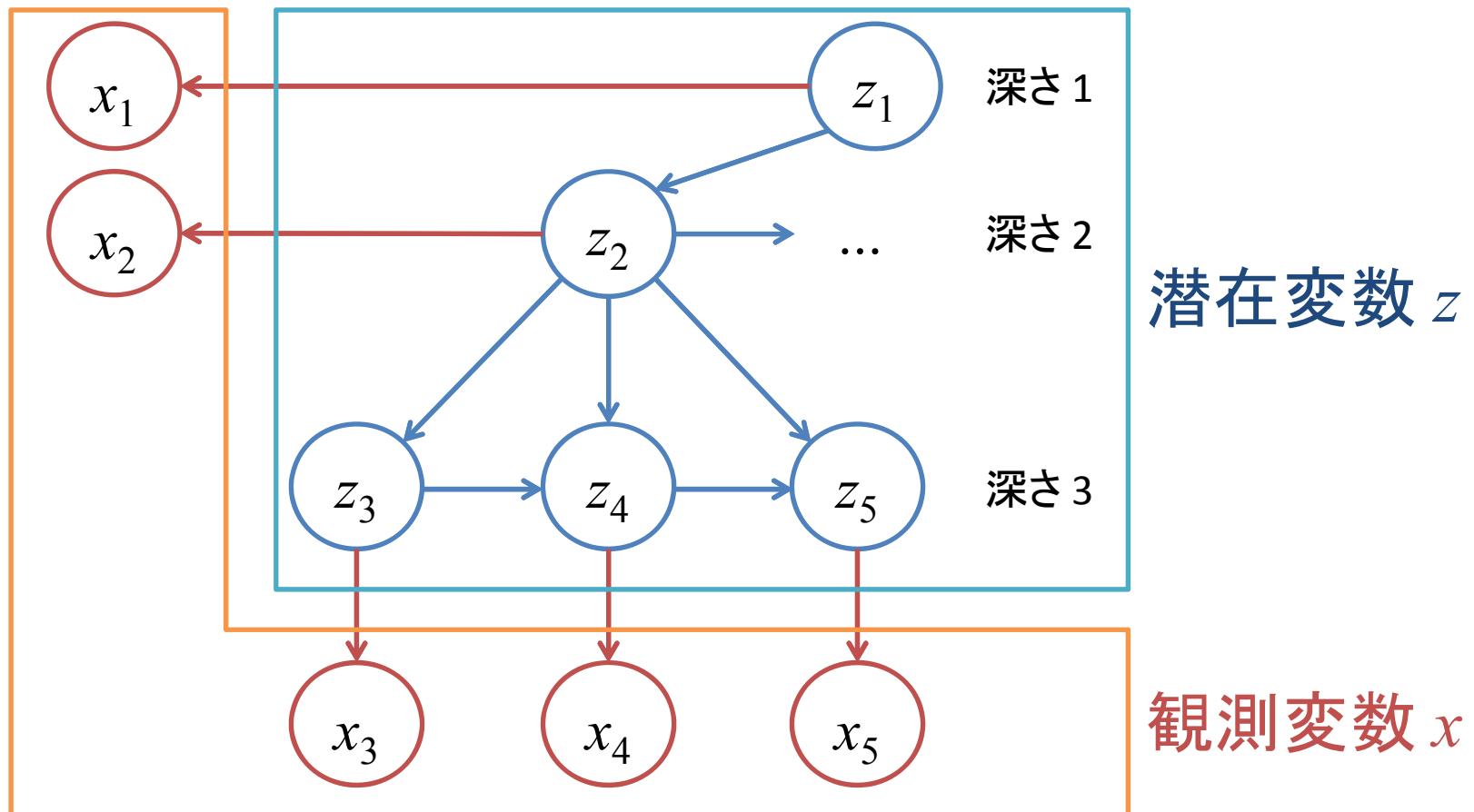
- 階層構造を順序木で表現



Text data の構造を表現する順序木

階層構造に基づく潜在変数モデル ¹⁴

- 潜在変数 z は階層構造と同一の構造を持つ
- 観測変数 x_i は対応する潜在変数 z_i にのみ依存



階層構造に基づく潜在変数モデル ¹⁵

- $p(x, z, \varphi, \theta \mid \beta, \alpha)$
= $p(x \mid z, \varphi) p(z \mid \theta) p(\varphi \mid \beta) p(\theta \mid \alpha)$

- $p(x \mid z, \varphi) = \prod_i p(x_i \mid z_i, \varphi_d)$

- $p(z \mid \theta) = \prod_i p(z_i \mid z_p, z_s, \theta_d)$

- $p(\varphi \mid \beta) = \prod_d p(\varphi_d \mid \beta)$

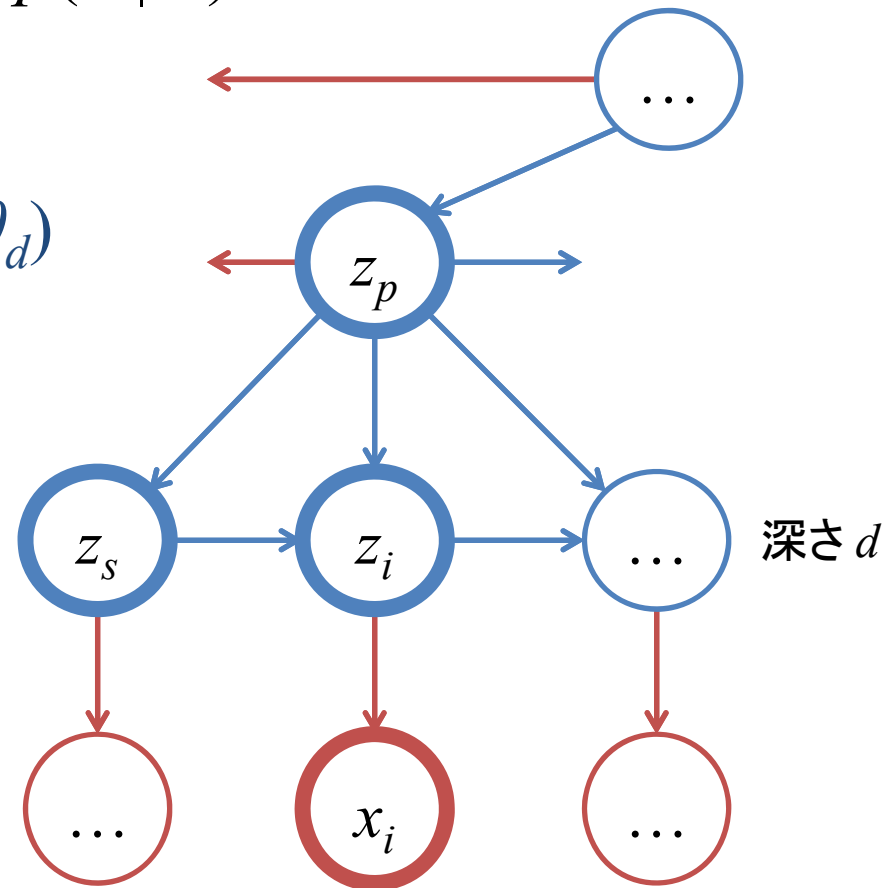
- $p(\theta \mid \alpha) = \prod_d p(\theta_d \mid \alpha)$

- **ただし**

- $d \equiv z_i$ の深さ

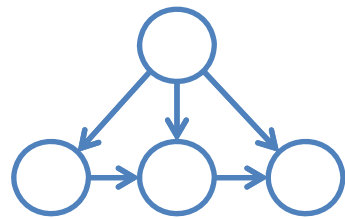
- $z_p \equiv z_i$ の親

- $z_s \equiv z_i$ の兄

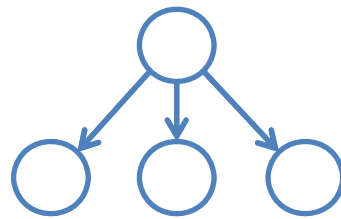


潜在変数の依存関係

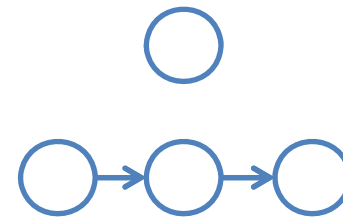
- z_p, z_s, z_i 間の依存関係



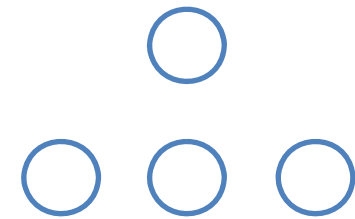
全依存



親に依存



兄に依存



全独立

- “親に依存”の種類

- 親の値に確率的に依存
- 親の値に決定的に依存
- 親のIDに確率的に依存
- 親のIDに決定的に依存
- 親の値, IDに確率的に依存

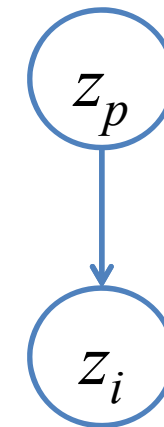
$$p(z_i | z_p, \varphi_d)$$

$$z_i = z_p$$

$$p(z_i | p, \varphi_d)$$

$$z_i = p$$

$$p(z_i | z_p, p, \varphi_d)$$



潜在変数の依存関係

- 深さ毎に以下の10種の依存関係がある
 1. 全独立
 2. 親の値 に決定的
 3. 親のID に決定的
 4. 親の値 に確率的
 5. 親のID に確率的
 6. 兄の値 に確率的
 7. 親の値, 親のID に確率的
 8. 親の値, 兄の値 に確率的
 9. 親のID, 兄の値 に確率的
 10. 全依存 (親の値, 親のID, 兄の値 に確率的)

様々な既存モデルを包含
混合多項分布, LDA, HMM, HTMM, etc

より複雑なモデルを表現可能
人手で扱えない深い階層を持つモデル

潜在変数の依存関係

Ex1) LDA = Word 毎に Topic、Doc. 毎に Topic 分布を持つ

Doc. : 全独立

Para. : 親のIDに決定的 (Doc. ID)

Sent. : 親のIDに決定的 (Doc. ID)

Word : 親の値に確率的 (Topic)

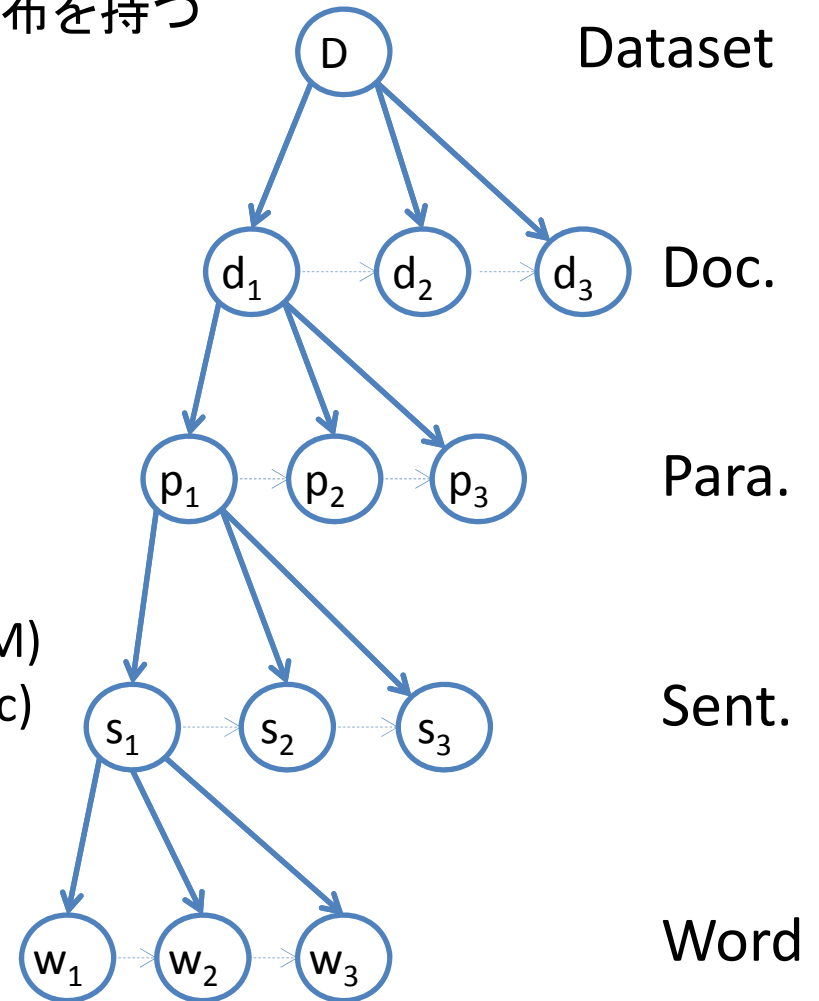
Ex2) Hidden Topic Markov Model (HTMM)

Doc. : 全独立

Para. : 親のIDに決定的 (Doc. ID)

Sent. : 親の値, 兄の値に確率的 (Topic HMM)

Word : 親の値に決定的 (Sent. Topic)



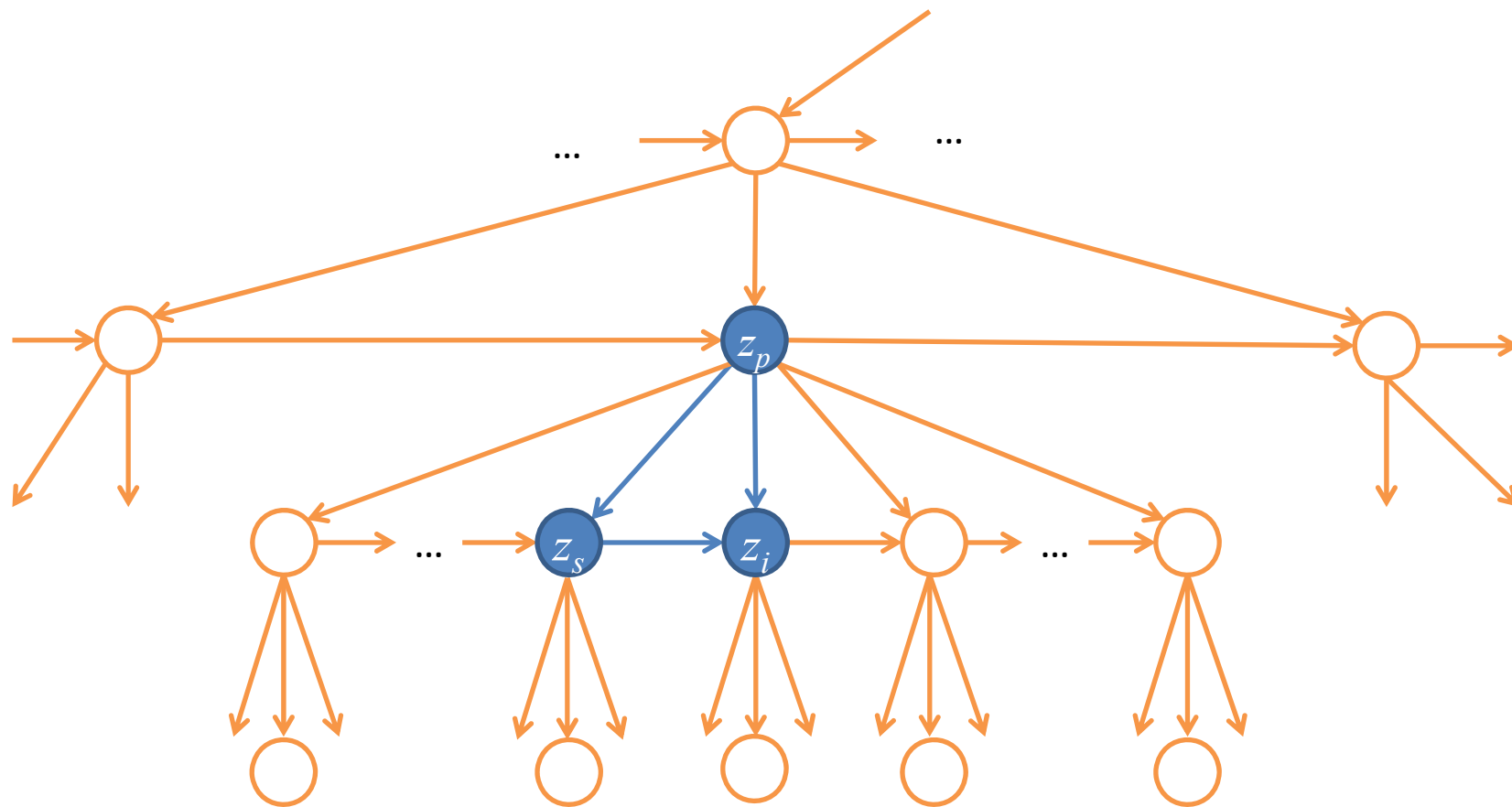
提案手法

- 目的
 - データの階層情報を利用した潜在変数モデルの自動生成
- 提案
 1. データの階層構造を利用したモデルクラス の定義
 2. モデルクラスに対する 汎用な 学習アルゴリズム
 3. モデルクラス上の探索手法

効率的な確率計算

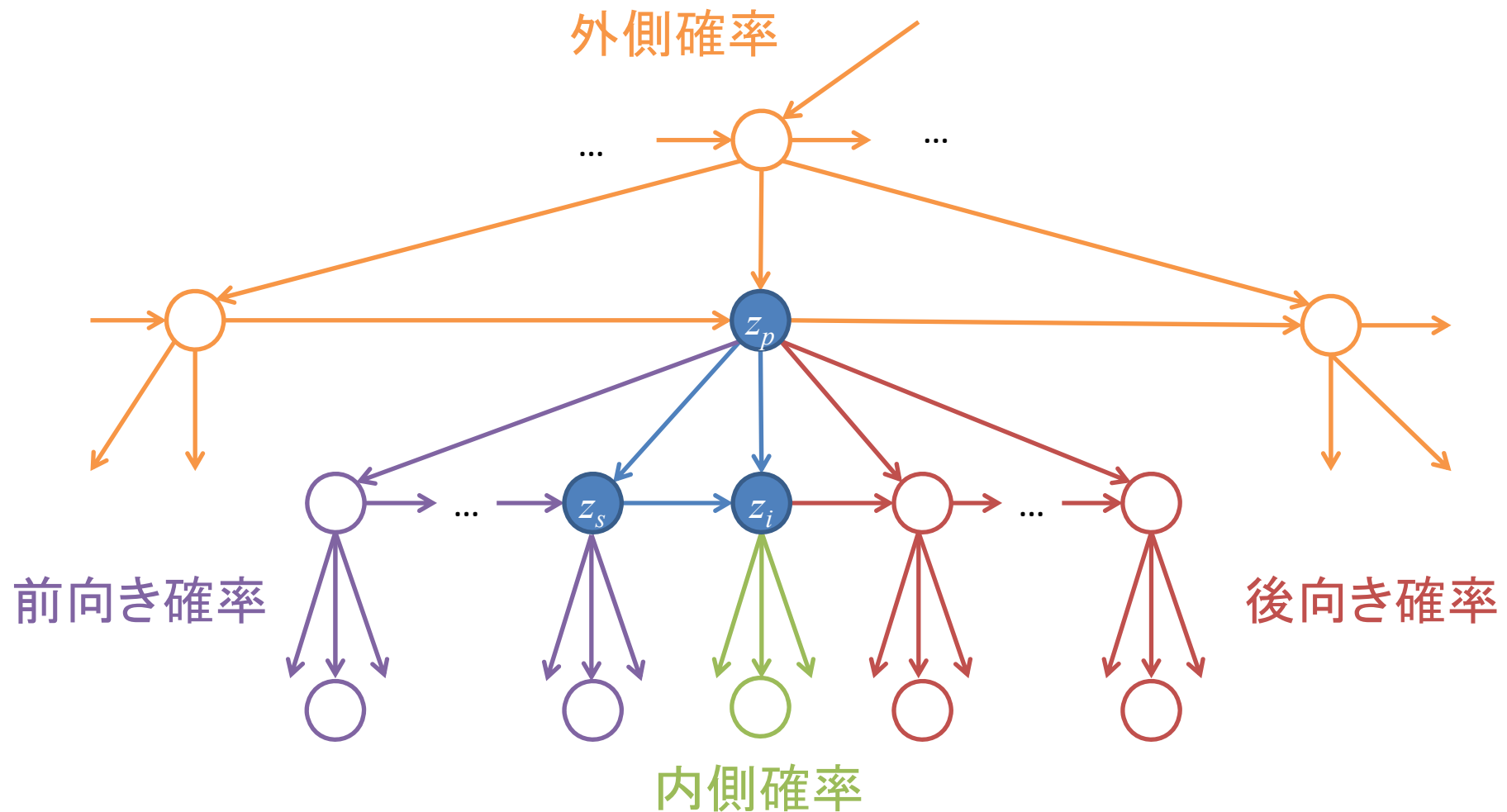
- 学習には $p(z_p, z_s, z_i | x, \theta, \varphi)$ が必要

オレンジをすべて周辺化 → 愚直に計算すると指数時間



効率的な確率計算

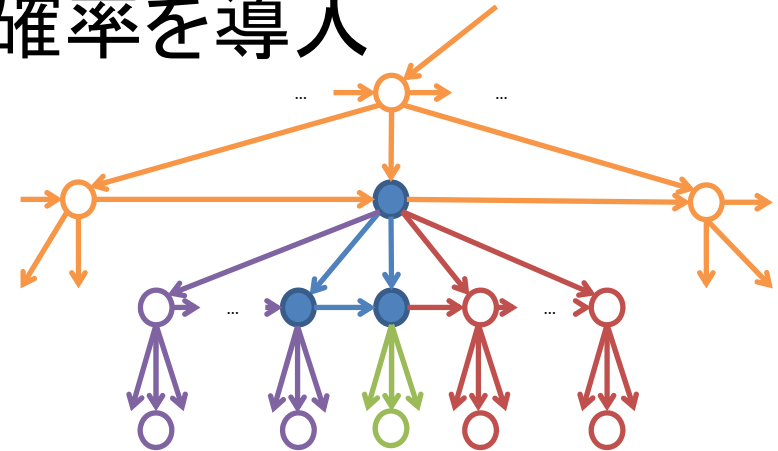
- DP により $p(z_p, z_s, z_i | x, \theta, \varphi)$ を効率的に計算



効率的な確率計算

- 効率的な計算のため4種の確率を導入

- 外側確率
- 内側確率
- 前向き確率
- 後向き確率



- 動的計画法により $O(NK^3)$ で計算可能
 - N = 節点数
 - K = 潜在変数の取りうる値の数 (トピック数)
 - 愚直に確率計算すると $O(K^N)$
- $O(NK^3)$ でモデルの学習が可能

提案手法

- 目的
 - データの階層情報を利用した潜在変数モデルの自動生成
- 提案
 1. データの階層構造を利用したモデルクラス の定義
 2. モデルクラスに対する 汎用な 学習アルゴリズム
 3. モデルクラス上の探索手法

モデルの自動生成

- 各深さ毎の独立性の仮定を探索により自動決定
- 探索法 = 貪欲法
 - 初期仮定 = 全深さにおいて全独立と仮定
 - 遷移 = 最も **選択基準** を向上させるよう、仮定を1つ変更
 - 終了条件 = 選択基準が改善されなければ終了
- 選択基準
 - 変分自由エネルギー = 周辺対数尤度 $\ln p(x, | M)$ の下限値
 - **提案学習アルゴリズムにより効率的に計算可能**
- 目的に応じた拡張
 - タスクに必要な潜在変数があるときは、それを含む空間のみ探索
例) 文書クラスタリング = 文書レベルに潜在変数を必ず入れる

モデル M からデータが得られる確率

提案手法 まとめ

- 目的
 - データの階層構造を利用したモデルの自動生成
- 提案手法
 - 階層構造を元にモデルクラスを定義
 - 潜在変数が **観点** と対応付けれるため解釈しやすい
 - 既存の有用なモデルを包含する
 - 人手では設計困難なモデルも生成可能
 - 高階層なデータからのモデルの自動生成
 - 効率的なモデル生成
 - 動的計画法を用いた学習
 - 貪欲法による探索

目次

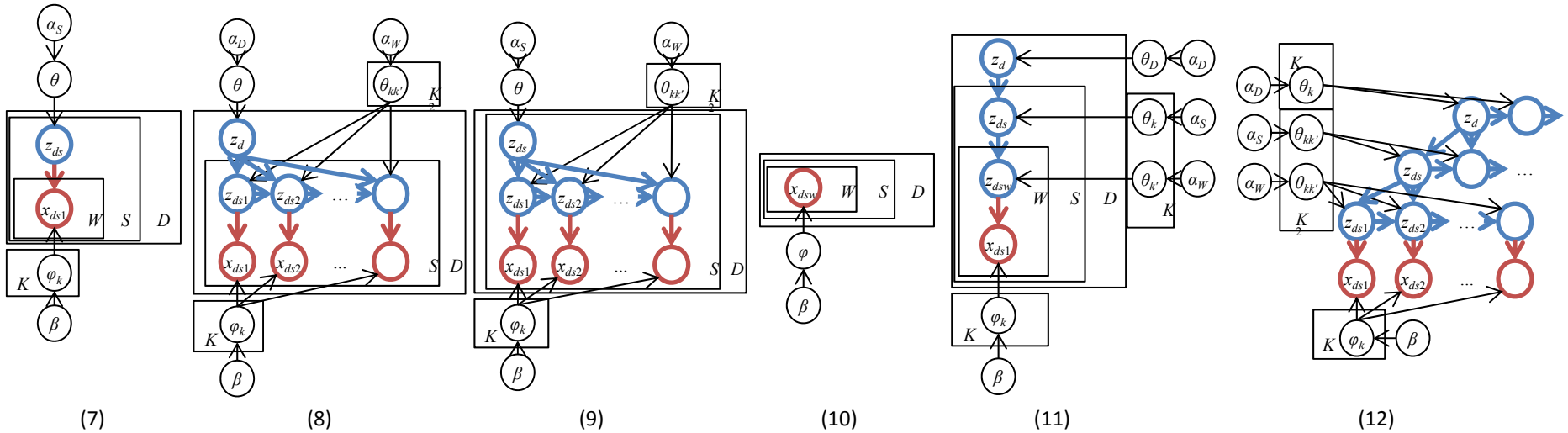
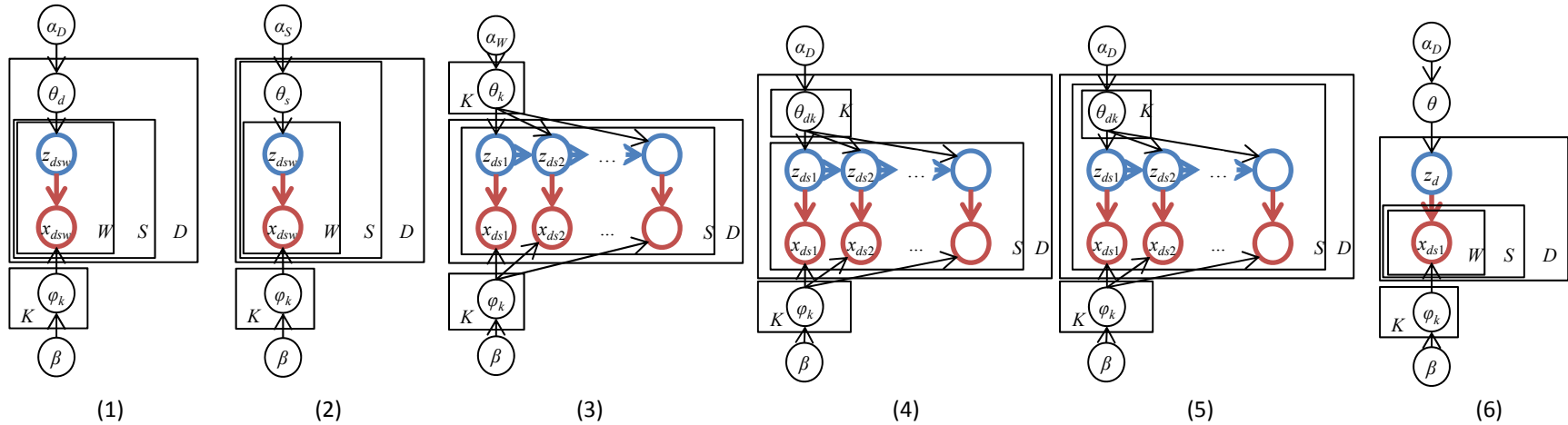
- 研究背景・目的
- 潜在変数モデル
- 関連研究
- データと階層構造
- 提案手法
- **実験**
 - 人工データ
 - 実データ
- **まとめと今後の計画**

実験1

- 人工データ
 - 人手で設計したモデル12個からサンプリング
 - 階層 document – sentence – word
 - 観測 30文書, 30文/文書, 30単語/文
 - データサイズ $30^3 = 27000$ words
 - 語彙数 50
 - クラスタ数 5
- 結果
 - データを生成したモデルを復元できるか確認

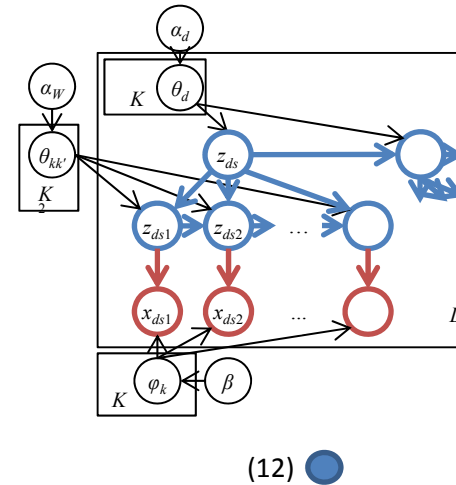
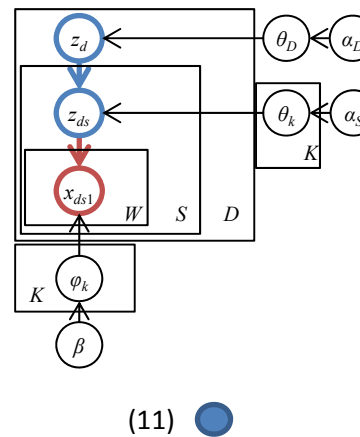
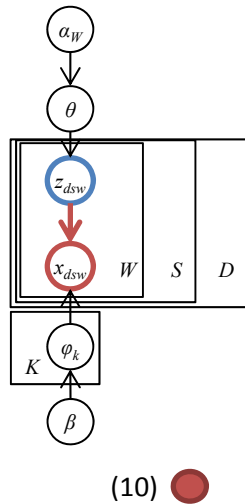
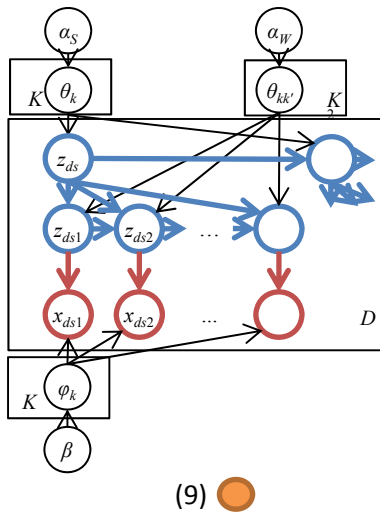
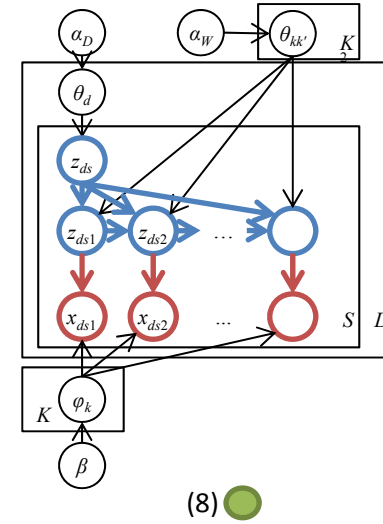
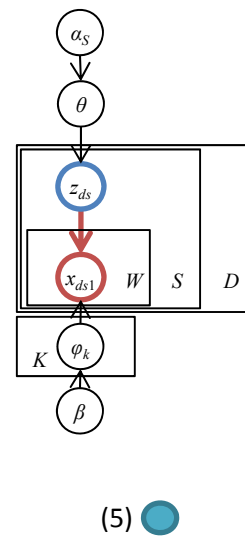
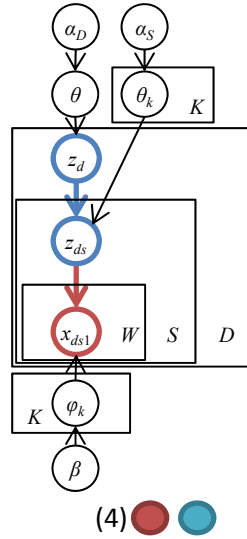
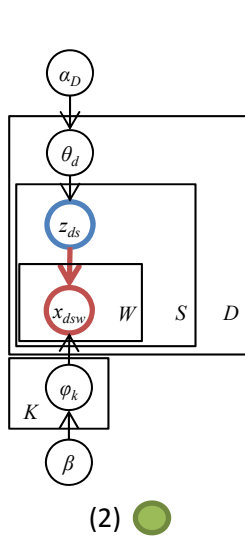
正解モデル

LDA, HMM, HTMM, HMTM, 混合多項分布, それらの組合せなど



正しく予想できなかったモデル

● 階層ズレ ● 階層+ ● 階層- ● 仮定+ ● 仮定-



実験1 まとめ

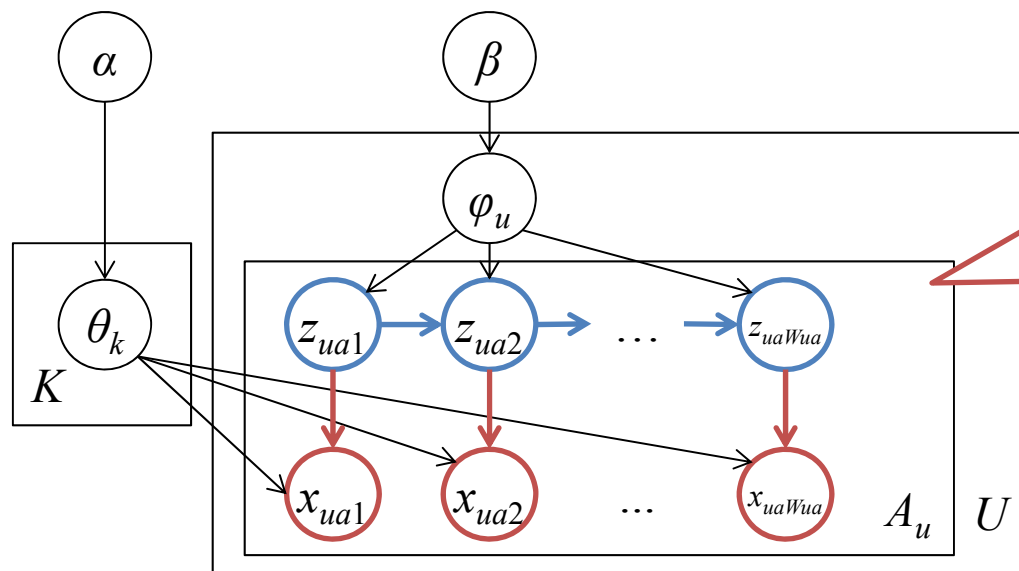
- 各モデルの概形は復元できている
 - 4モデルは正しいモデルを生成
 - 間違った8モデルも高々2のズレ
- なぜ間違えるのか？
 - データ不足
 - 変分自由エネルギーによる近似の影響
 - 貪欲法による局所解への収束
- 今後の課題
 - データを増加させた場合の挙動を確認
 - 変分自由エネルギー以外の近似と比較
 - 貪欲法以外の探索法と比較

実験2

- goo blog データ
 - 50 ユーザの 4～6 月の3ヶ月分のブログ
 - 階層 : user – article – word
- 実験設定
 - データの前処理
 - MeCab で形態素解析
 - 頻出の名詞、動詞を1000語を利用
 - トピック数
 - 全階層で10

得られたモデル

- LDA + HMM (HTMMの一種)
 - User : 親のID に決定的 (不必要)
 - Article : 親のID に決定的 (User ID)
 - Word : 親の値、兄の値に依存 (User 毎のHMM)



各単語は対応する単語トピックに依存
 単語トピックは前の単語トピックに依存
 トピックの遷移はユーザ毎に異なる

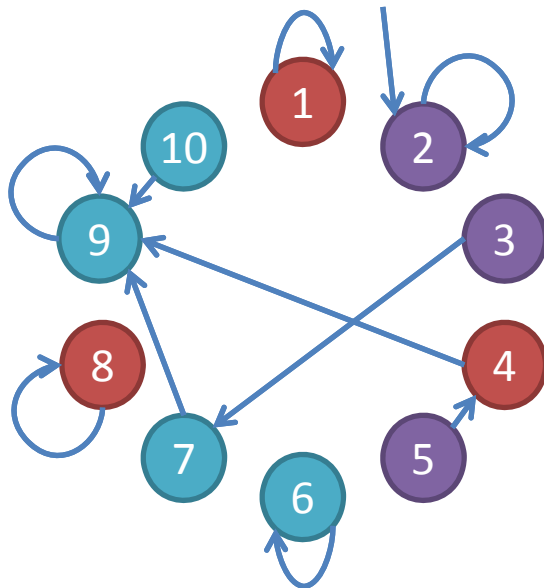
得られたトピック

● 動詞 ● 名詞 ● 混在 33

<p>1</p> <p>居る 有る 言う 成る 見る 物 来る 花 行く 思う 一時 雨 所 度 出る 見える 取る 咲く 天気 中方 今日 気 写真 風 日 分かる</p>	<p>2</p> <p>社会 見る 認知 日本 問題 心 仕舞う 研究 因る 知識 時 自分 情報 表現 大学 分かる 注意 管理 失敗 説明 人間 思考 遣る 集中 効果 評価</p>	<p>3</p> <p>クリック 鮪 ランキング 隊 呉れる 隊長 下さる 頂く 番 願う 応援 会 魚 留守 食べる 止め 今夜 委員 御座る 一 料理 添乗 びっくり</p>	<p>4</p> <p>言う 居る 有る 成る 思う 見る 人 笑い 出来る 物 来る 遣る 分かる 方 行く 話 気 聞く 今 書く 出る 自分 訳</p>	<p>5</p> <p>言う 有る 書く 居る 場合 成る 画面 時 出来る 物 処理 部分 データ つく 行く 作る 設計 考える 遣る 条件 生成 自動 仕様</p>
<p>6</p> <p>年 関連 インターネット 携帯 電話 情報 メディア マルチメディア 事典 用語 テロ ハイコンセプト ローテク 自由 操作 地球</p>	<p>7</p> <p>ページ 月 絵画 日 小説 恐怖 美術 童話 戻る 物語 屋敷 日記 条理 写真 海 カテゴリー</p>	<p>8</p> <p>居る 来る 行く 成る 今日 言う 時 見る 日 有る 思う 明日 時間 遣る 食べる 今 出来る 寝る 呉れる 買う 返る 仕舞う</p>	<p>9</p> <p>年 月 日 居る 新聞 報告 成る 米国 有る 言う 情報 社 因る 広告 為 著作 世界 メディア 日本 公開 テレビ 中国 放送 電話</p>	<p>10</p> <p>回 今日 選手 投手 点 復習 度 民法 勝 日 タイ 明日 ゼム ニュース 囲碁 終わる 試合 タイムリー マリン ラン</p>

得られた解釈：名詞＋動詞

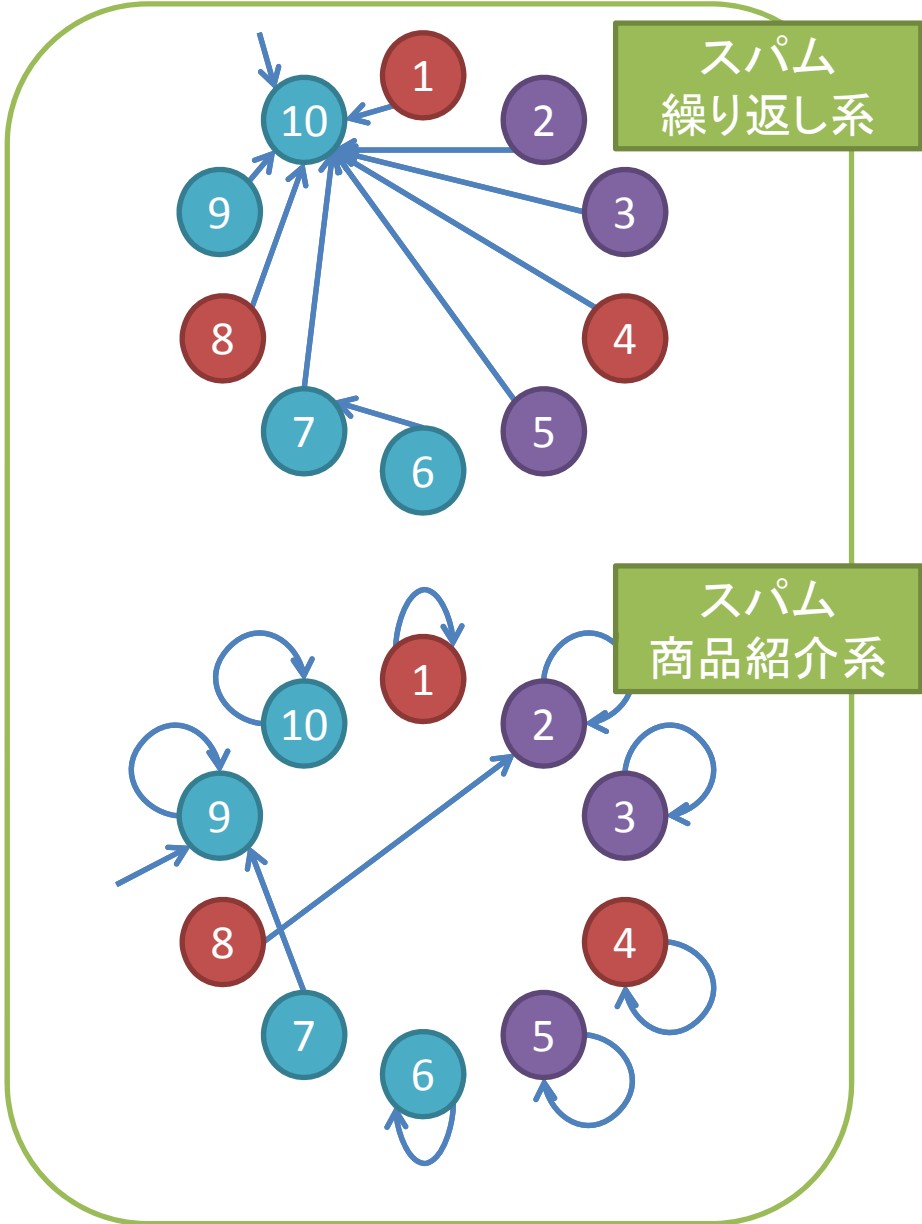
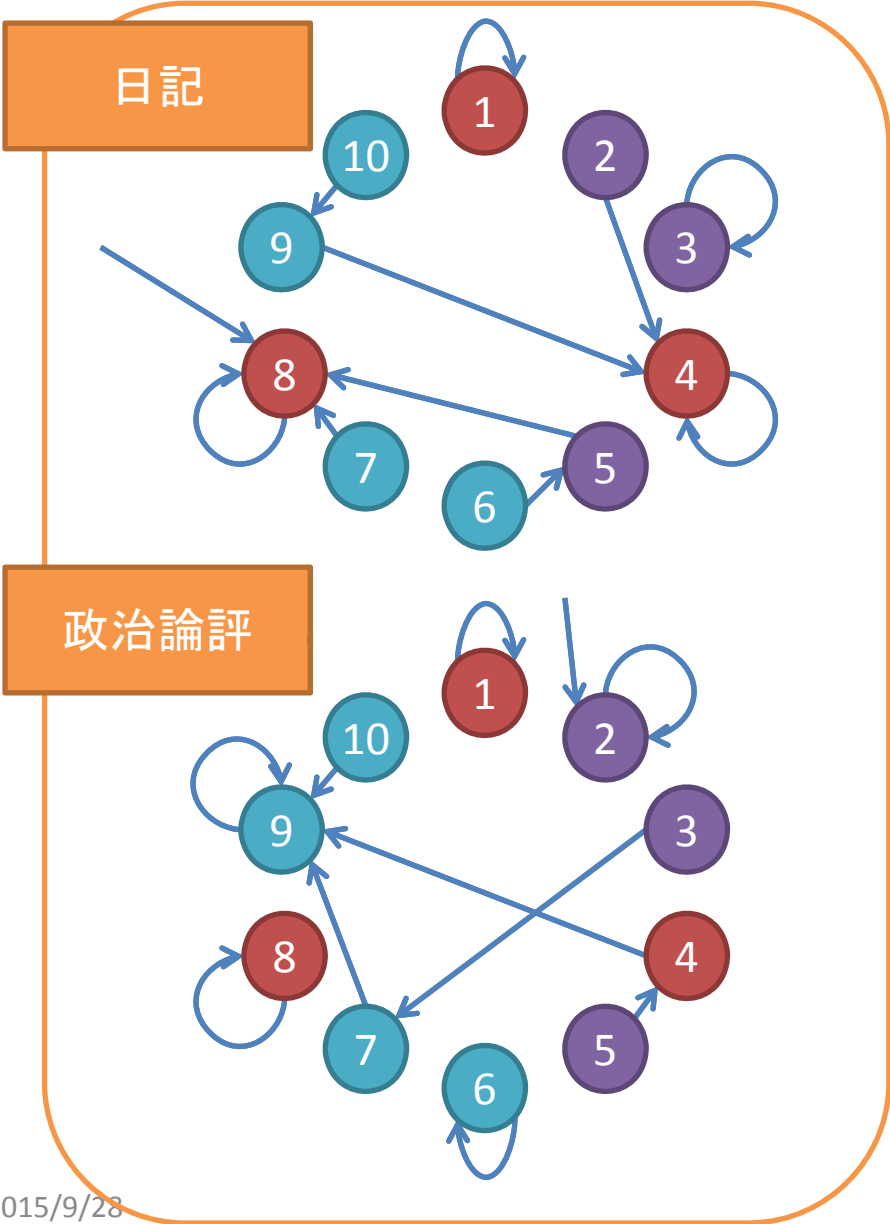
- ① ⑧ ④ : 動詞が多い
- ⑥ : インターネット・モバイル系
- ⑦ : カテゴリー名系
- ⑨ : 政治・経済系
- ⑩ : 野球系



生成されたモデルは
User 毎にトピックの遷移確率を持つ

最も高確率な遷移を可視化

得られた解釈：名詞＋動詞



まとめと今後の課題

- まとめ
 - データの階層構造を利用したモデル自動生成法を提案
 - 階層構造によるモデルクラスの定義
 - 動的計画法による効率的な学習
 - 探索を利用したモデルの自動生成
 - 小規模実データに適用
 - ブログデータから自動的に潜在構造を抽出
- 今後の課題
 - トピック数を自動推定するよう拡張
 - 大規模データに適用
 - 階層構造以外の情報を利用
 - 複数の種類のデータが混在する場合に拡張



Innovative R&D by NTT



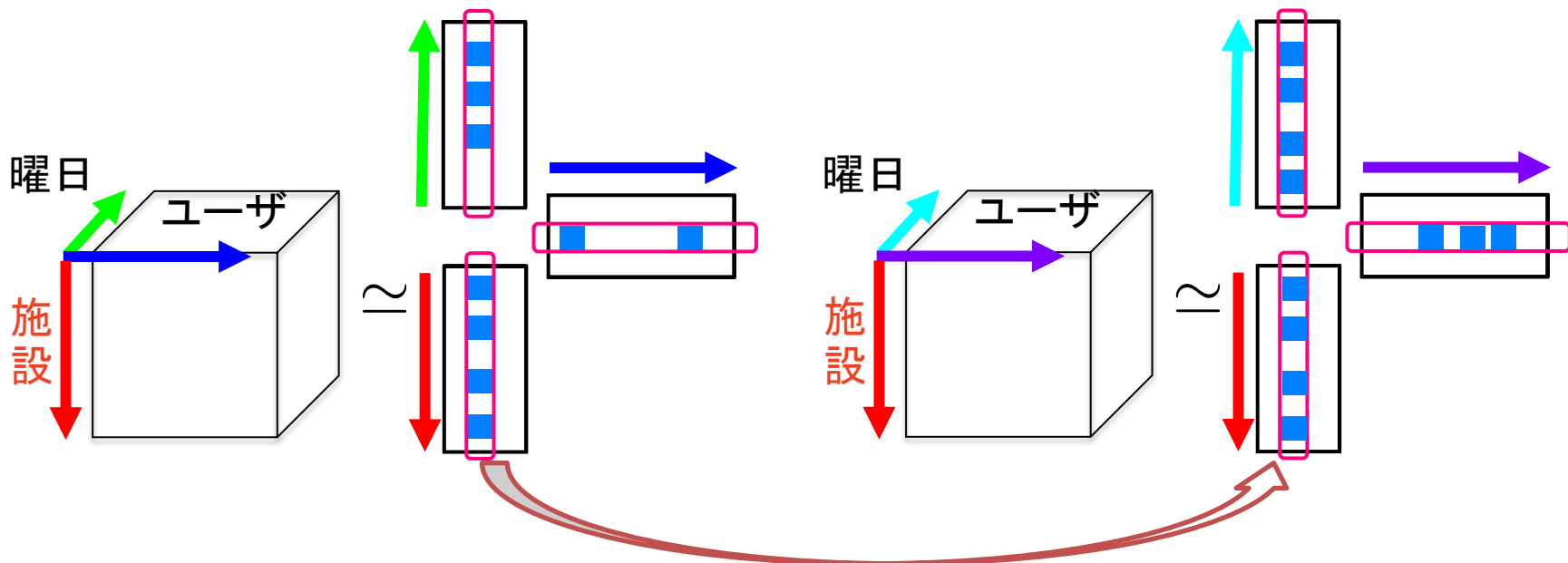
多種類のデータから共通するパターンを 抽出する複合テンソル分解技術

CS研創情部創知G
機械学習データ科学センター

竹内 孝

2015年8月4日

- ✓ 同一の属性に共通の因子行列を仮定することで、データ間で共通する横断パターンが抽出される



$$x_{ijk} \cong \sum_{r=1}^R a_{ir} b_{jr} c_{kr}$$

同一の因子行列を仮定する

- ✓ センサやウェブを介し実世界から多種類のデータが観測されている
- ✓ 多種類のデータから解釈容易性の高いパターンを抽出することで、実世界を理解する手助けが出来ないか？
- ✓ 多種類のデータを横断的に解析する汎用技術を構築する

✓ ある観測が2個の属性と1つの値からなる関係データ

✓ 例: (電信太郎:銀座美術館:利用頻度1)

✓ 3ユーザ4施設の利用頻度データの行列

ユーザ属性

施設属性

	電信太郎	電信花子	電信二郎
銀座美術館	1	5	2
新宿デパート	0	1	1
吉祥寺劇場	2	3	4
三鷹公園	1	5	9

✓ 季節ごとの3ユーザ4施設利用頻度データ

✓ 例: (電信太郎:銀座美術館:春:利用頻度0)

ユーザ属性

冬

秋

夏

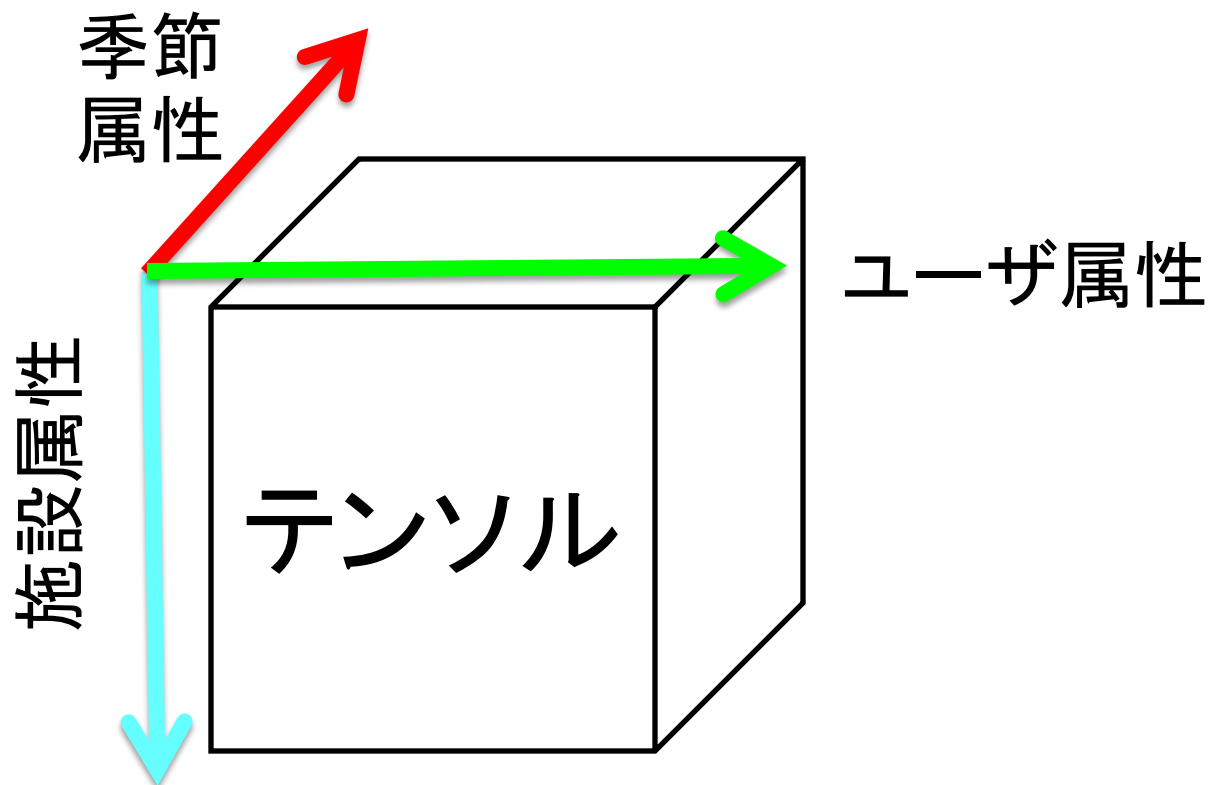
春

施設属性

	電信太郎	電信花子	電信二郎				
					電信太郎	電信花子	電信二郎
		電信太郎	電信花子	電信二郎			
	電信太郎	電信花子	電信二郎				
銀座美術館	1	5	2				
新宿デパート	0	1	1				
吉祥寺劇場	2	3	4				
三鷹公園	1	5	9				

✓ 3ユーザ4施設4季節の利用頻度データ

✓ 例: (電信太郎:銀座美術館:春:利用頻度0)



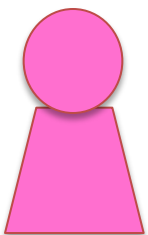
- ✓ 例: 3属性かつ各属性長が1000のテンソル
 - ✓ $1000^3=10$ 億の要素(組合せ)が存在
 - ✓ 頻出組合せによるデータ傾向の把握は困難
- ✓ テンソル分解は10億個の要素を高々 **$K \times 3000$ 個の少要素**で再現する技術 ($K \ll 1000$)
- ✓ キーと成るのは、**低ランク性**という概念

- ✓ データをK人の仮想ユーザ、各ユーザの仮想ユーザらしさ近似できるを学習します

	電信太郎	電信花子	電信二郎
銀座美術館	1	5	2
新宿デパート	0	1	1
吉祥寺劇場	2	3	4
三鷹公園	1	5	9



仮想ユーザ1の
行動パターン



仮想ユーザ2の
行動パターン

行列
分解

電信太郎の重み

仮想ユーザ1 : 仮想ユーザ2 = 4:1

電信花子の重み:

仮想ユーザ1 : 仮想ユーザ2 = 1:5

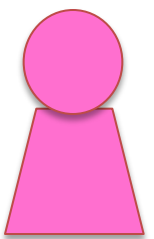
- ✓ データをK人の仮想ユーザ、各ユーザの仮想ユーザらしさ近似できるを学習します

	通信太郎	通信花子	通信二郎
--	------	------	------

仮想ユーザ、元のユーザ毎の重みは、行列分解技術によって、**因子行列**として抽出される



仮想ユーザ1の行動パターン



仮想ユーザ2の行動パターン

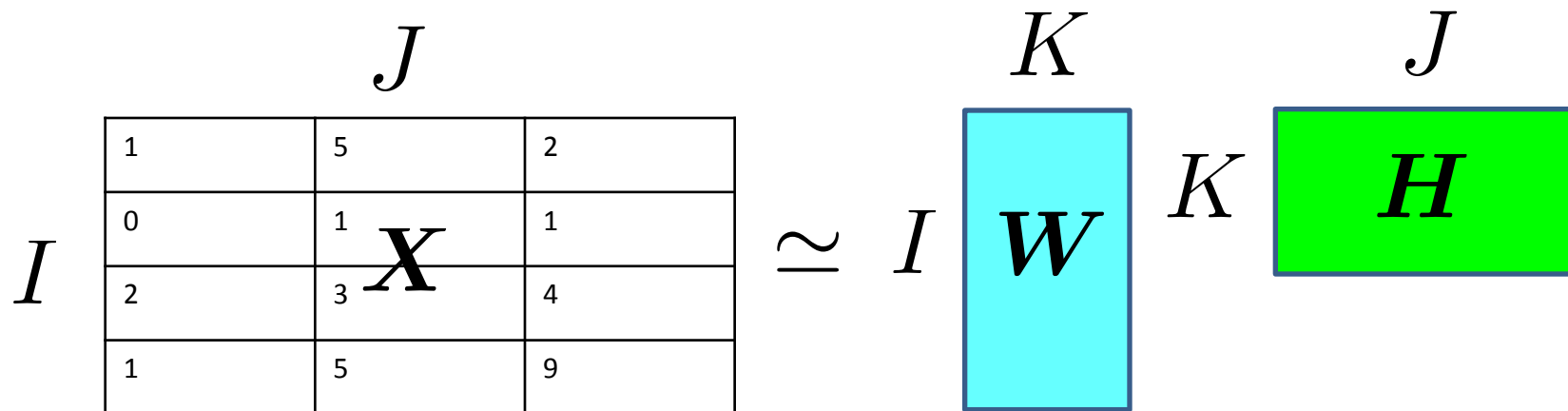
通信太郎の重み
仮想ユーザ1 : 仮想ユーザ2 = 4:1

通信花子の重み:
仮想ユーザ1 : 仮想ユーザ2 = 1:5

定式化と最小化問題

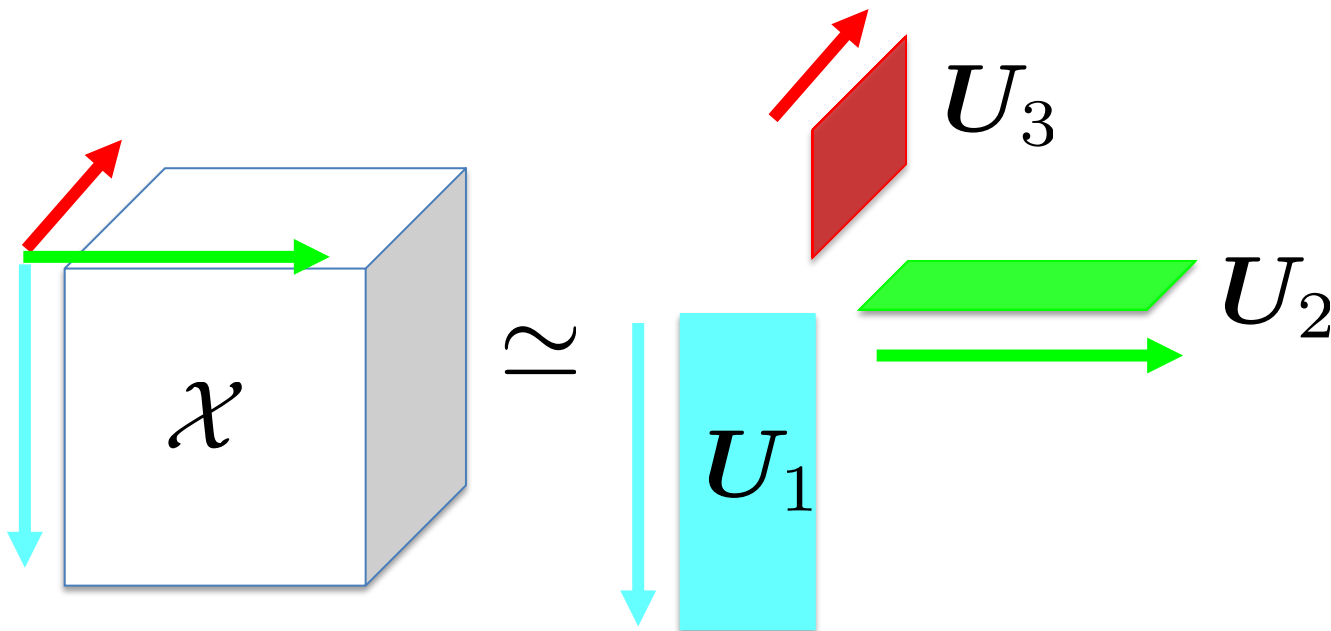
$$\mathbf{X} \in \mathbb{R}^{I \times J}, \mathbf{W} \in \mathbb{R}^{I \times K}, \mathbf{H} \in \mathbb{R}^{J \times K},$$

$$\min_{\mathbf{W}, \mathbf{H}} D(\mathbf{X} \parallel \mathbf{W}\mathbf{H}^T) \text{ subject to } \mathbf{W} \geq 0, \mathbf{H} \geq 0.$$



定式化と最小化問題

$$\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \dots \times I_N}, U = \{U_i \in \mathbb{R}^{I_i \times K} \mid i = 1, \dots, N\},$$
$$\min_U D(\mathcal{X} \| U) \text{ subject to } U_i \geq 0 \ (i = 1, \dots, N)$$



- ✓異なるデータ間でも、**同じ種類の属性**を用いている場合があります
- ✓同じ種類の属性値があるという情報を用いて異なるデータを繋げ合わせます

	電信太郎	電信花子	電信二郎
銀座美術館	1	5	2
新宿デパート	0	1	1
吉祥寺劇場	2	3	4
三鷹公園	1	5	9

	携帯五郎	携帯桃子	携帯史郎
銀座美術館	1	5	2
新宿デパート	0	1	1
吉祥寺劇場	2	3	4
三鷹公園	1	5	9

施設属性が同じ

Yelpデータへの応用例：多種類の関係データを同時テンソル分解



施設名

位置情報

The screenshot shows a Yelp business profile for 'Short Leash Dogs Food Truck'. The page includes a search bar at the top, navigation links, a star rating of 4.5 (236 reviews), categories (Hot Dogs, Food Trucks, American), address (110 E Roosevelt St, Phoenix, AZ 85004), phone number, website, and a menu. It also lists various attributes like 'Price Range: \$\$\$\$' and 'Hours: Tue-Fri 11 am - 10 pm'. A map on the right shows the location. Below the business info, there are user reviews, including one from Luke B. dated 7/23/2013.

時刻

チェックイン

評価点

ユーザ

単語

コメント

✓ 評価点データ

 **施設** × ユーザ × 曜日 → 評価点

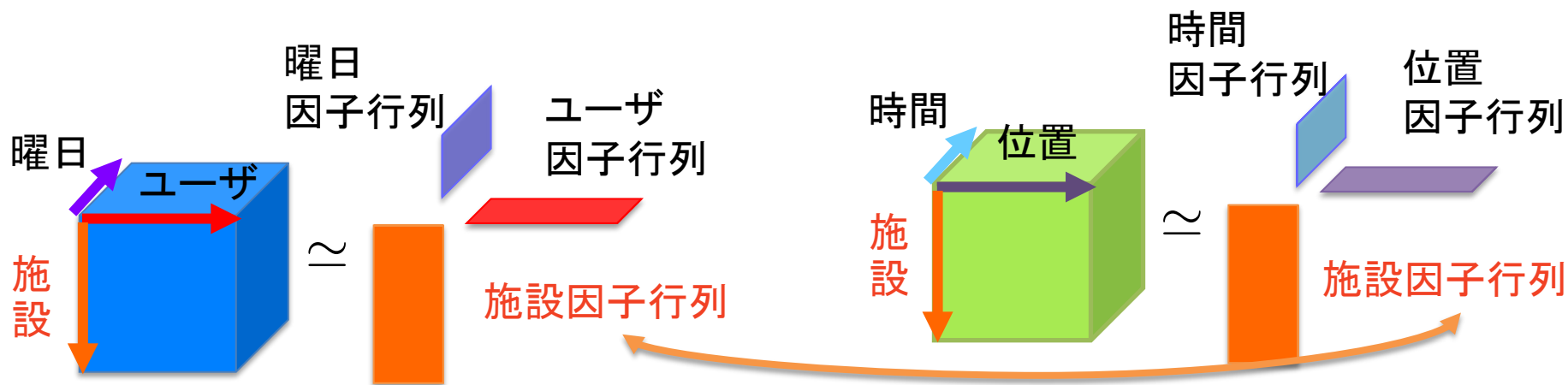
✓ チェックインデータ

 **施設** × 位置 × 時間 → 訪問頻度

✓ レビューコメントデータ

 ユーザ × ジャンル × 単語 → 出現頻度

- ✓ Non-negative Multiple Tensor Factorization
- ✓ 同一の属性に共通の因子行列を仮定することで、データ間で共通する横断パターンが抽出される



定式化と最小化問題

$$\mathcal{X}_m \in \mathbb{R}^{I_{m,1} \times I_{m,2} \cdots I_{m,N_m}} \quad (m = 1, \cdots, M),$$

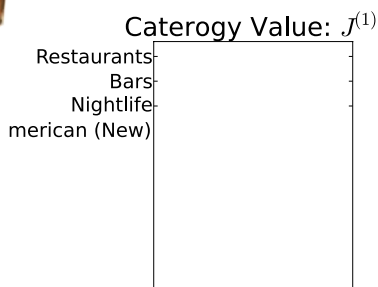
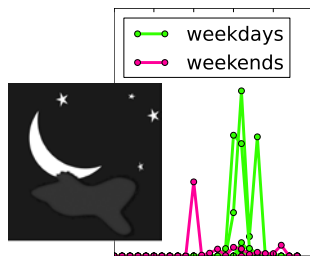
$$U_m = \{U_{m,i} \in \mathbb{R}^{I_{m,i} \times K} \mid i = 1, \cdots, N^{(m)}\},$$

$$U = \{U_m \mid m = 1, \cdots, M\},$$

$$\min_U \sum_m D(\mathcal{X}_m \| U_m) \text{ subject to } U_{m,i} \geq 0 \quad \forall (m, i),$$

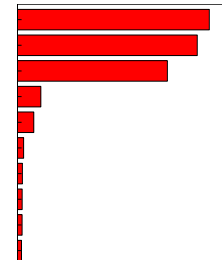
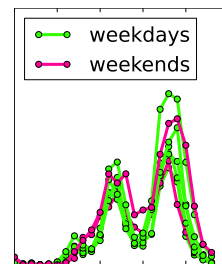
$$\text{and } U_{m,i} = U_{m',i'} \quad \{(m, i), (m', i') \mid I_{m,i} = I_{m',i'}\}$$

平日午後の食事に関する因子



日本食レストランのランチ・ディナー因子

y





大量の事例を用いた超高品質音声強調

メディア情報研究部 信号処理研究グループ

木下慶介

研究の背景 (1/2)

実環境で音声を収録すると**雑音**や**残響**が含まれる

➡音質、音声アプリケーション全般の**品質低下の要因**



(携帯)電話



ICレコーダ

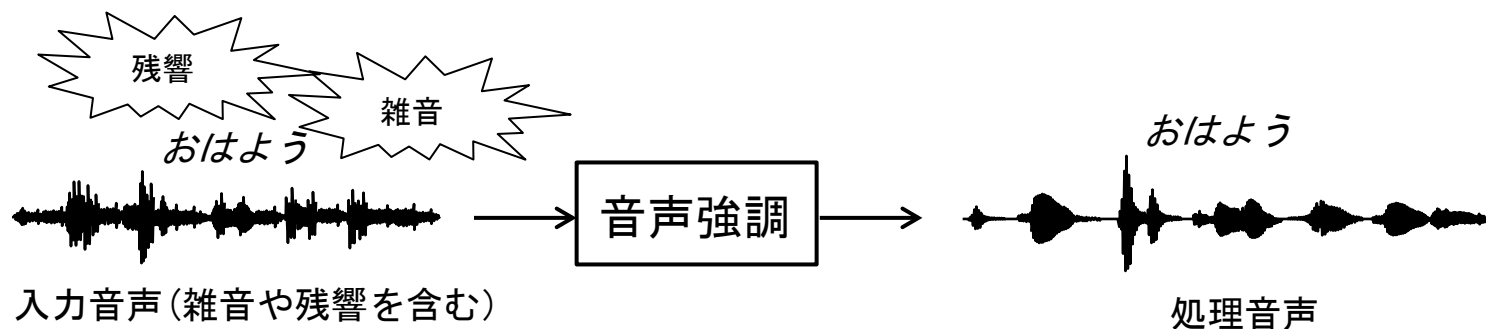


補聴器

その他、TV会議など...

研究の背景 (2/2)

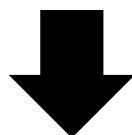
😊 音声強調（雑音除去、残響除去）の発展・普及



- ☹️ 処理音声に歪み (残留雑音・残響、不自然な音) が残ることが多い
→ 処理の結果、むしろ耳障りな・聞きづらい音に…

研究目的と着眼点

目的：処理音声に歪みが残らない、**超高品質音声強調の実現**



着眼点：大量のクリーン音声（キレイな音声）を用意して、
観測信号（雑音・残響含む）を適切な**クリーン音声に**
差し替えてしまおう！（音声強調と音声合成の融合）




学習データ

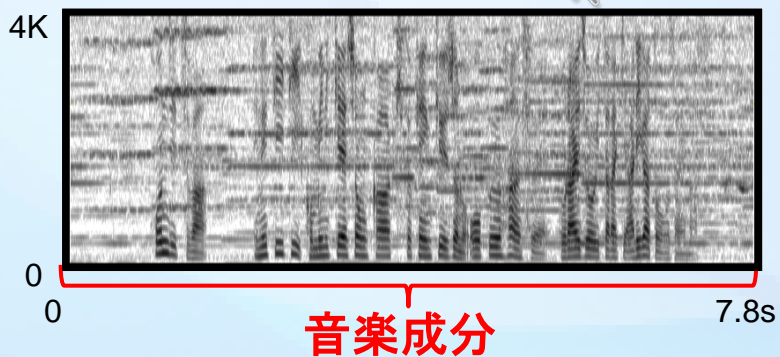



観測データ

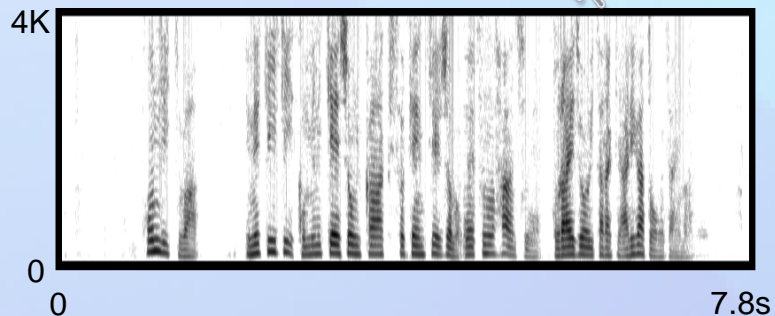
デモンストレーション

効果：非定常雑音除去


a) 観測信号 

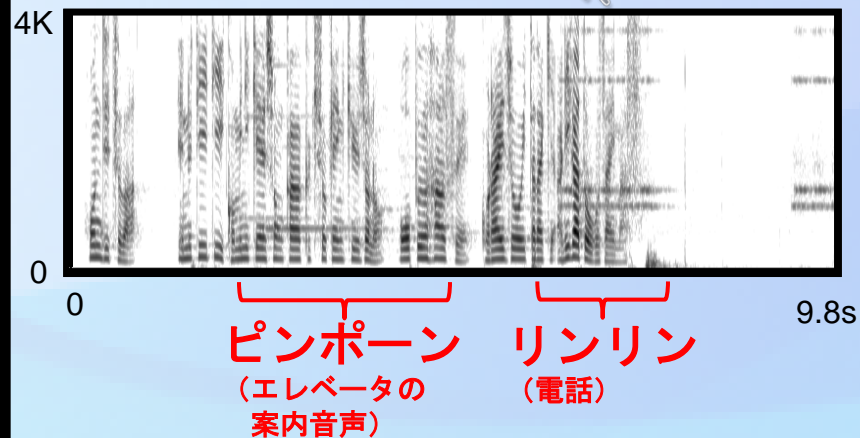



b) 処理音声 

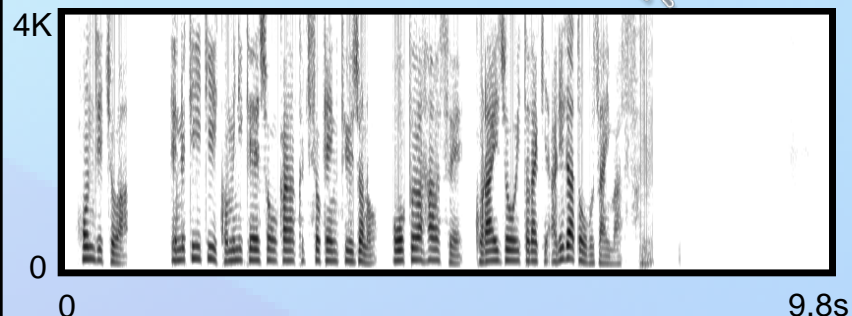


効果：非定常雑音除去+残響除去

a) 観測信号 



b) 処理音声 



処理条件

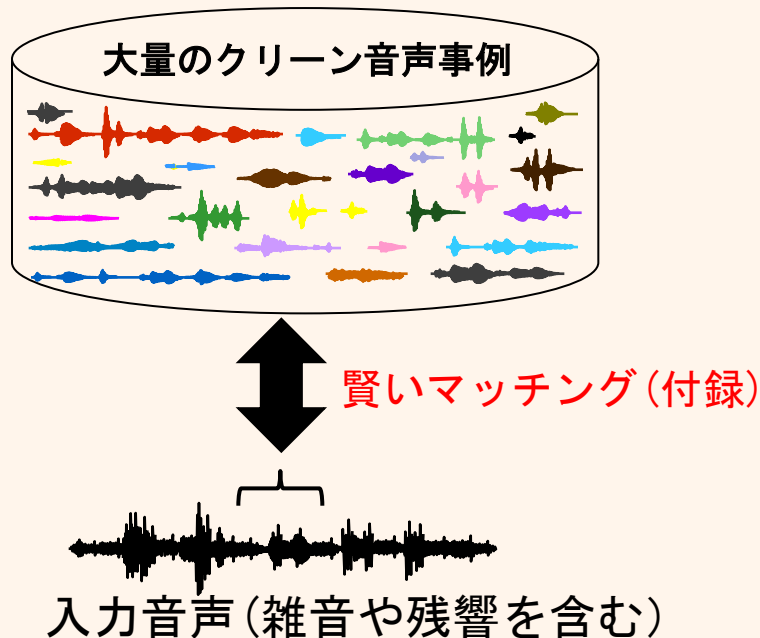
学習用クリーン事例:男性話者1名のクリーン発話(3.6時間)

学習用観測音事例:上記3.6時間の発話に生活雑音を重畳

事例ベース音声強調(1/2)：概念図

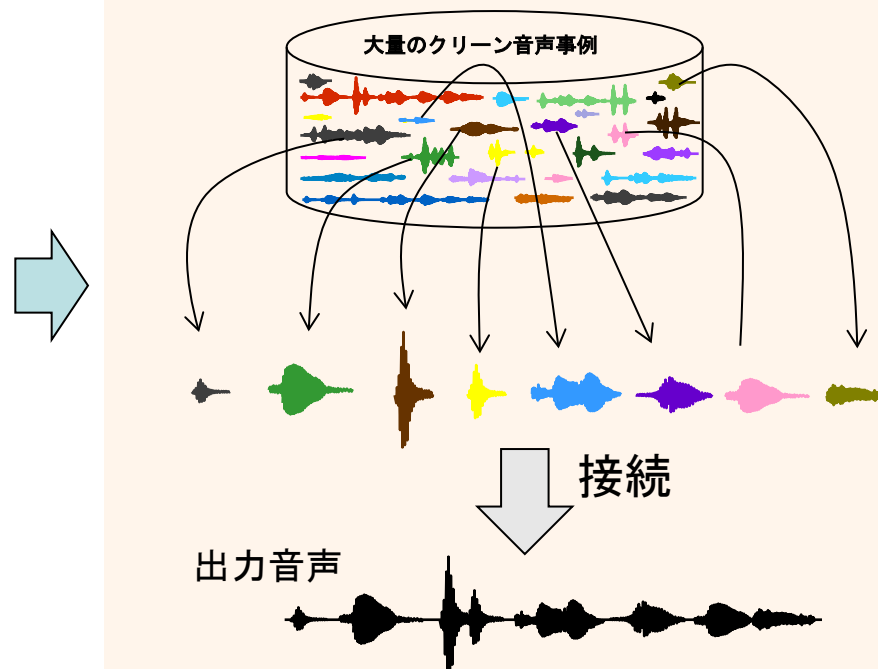
-ステップ1-

雑音や残響を含む入力信号と
大量のクリーン音声事例を賢くマッチング。
入力信号内に含まれる**クリーン音声成分**と
マッチする**最長(※)事例**を探索。



-ステップ2-

マッチした中で**最長のクリーン音声事例**を
選択・接続し、出力音を作成。



クリーン音声事例には雑音が含まれて
いないため、**出力音には雑音のらない**

※より長い事例を探索することで、突発的な雑音・ミスマッチに頑健に、
且つ、ステップ2での接続をより滑らかにできる

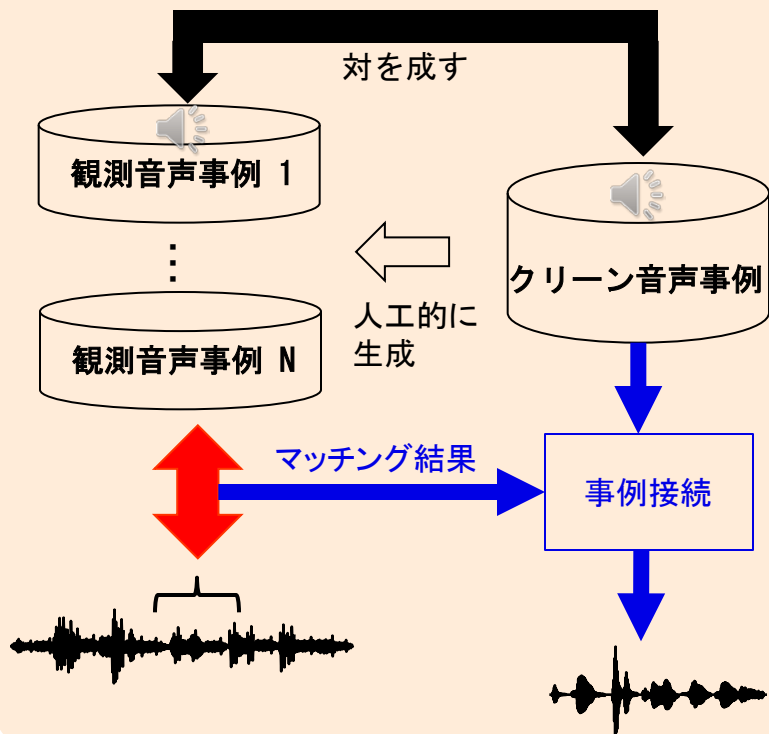
事例ベース音声強調 (2/2) : 賢いマッチ



入力音に含まれるクリーン音声成分にマッチするクリーン音声事例を探すための「賢いマッチング」とは？

-手段1 (基本系)-

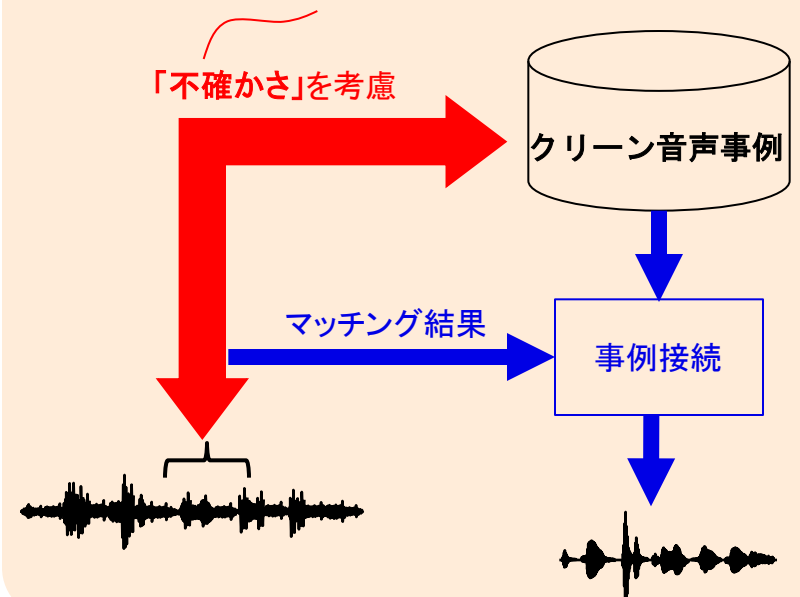
大量の観測音声例を用意し、それとマッチング (音声認識等で使われる高速探索法を導入)



-手段2 (発展系) -

マッチングの不確かさを考慮しながら探索 (手段1との組み合わせも可能)

不確かさの評価: 完全に合致していなくても、どの程度合致する可能性があるか、を評価



Thank you 😊

禁 無 断 転 載

2015年度「ビジネス機器関連技術調査報告書」 “第Ⅱ章”

発行 2016年6月
一般社団法人 ビジネス機械・情報システム産業協会 (JBMIA)
技術委員会 技術調査専門委員会
〒108-0073 東京都港区三田三丁目4番10号 リーラヒジリザカ7階
電話 03-6809-5010(代表) / FAX 03-3451-1770