

# 共通状態と連結学習を用いたHMM によるコールセンタ対話の要約

東中竜一郎, 南泰浩, 西川仁, 堂坂浩二,  
目黒豊美, 小橋川哲, 政瀧浩和,  
吉岡理, 高橋敏, 菊井玄一郎

# 概要

- コールセンタ対話の要約手法を提案
  - 共通状態と連結学習を用いたHMM
  - 要約の正解データが不要
- 実験的に有効性を検証
- 要約の例(こういうことをしたい)
  - 実例を始めに載せる
  - 特殊なタスクだから課題を分かりやすく提示

# 背景

- 一日数万件(処対に数を確認する;一般論の方がいいかもしれない)のコールセンタ対話
- 自動要約への高いニーズ
  - 1対話が非常に長いので把握が困難
  - オペレータの振り返りの効率化
  - スーパーバイザによる業務内容の確認に有効
- コールセンタ対話の要約研究は少ない
  - ルールによる発話抽出(Byrd et al., 2008)

# 問題

- コールセンタ対話の内容は多様
  - ルールベースの手法には限界
- 要約モデルをデータから学習
  - 要約の正解データを作成しモデルを学習
    - Regina Barzilay and Lillian Lee, 2004
  - われわれのコールセンタ対話（NTTの区分を少し見せる，処対に確認）は多様なドメインを扱うため，ドメインごとに学習データ（要約の正解）を作成する必要があり高コスト

正解を必要としない要約モデルの学習手法が必要

# アプローチ

- コールセンタ対話の特性を用いた要約正解が不要な要約モデルの学習
- コールセンタ対話の特性

- 量が大量にある
- 多様なドメインの対話がある
  - 故障受付, 引越し, 契約, 解約受付, ...
- 各対話にはドメインラベルが付与済み

各ドメインに特徴的な発話系列を抽出し, 要約を作成

# アプローチ(つづき)

ドメイン: **解約** ドメインラベル

オペレータ: お電話ありがとうございます. **共通**

ユーザ: **プロバイダを解約したいのですが**

**解約**

オペレータ: はい. **共通**

ユーザ: **手続きをお願いできますか?** **解約**

**解約**

オペレータ: かしこまりました. それではお客様の  
情報を確認させてください. **共通**

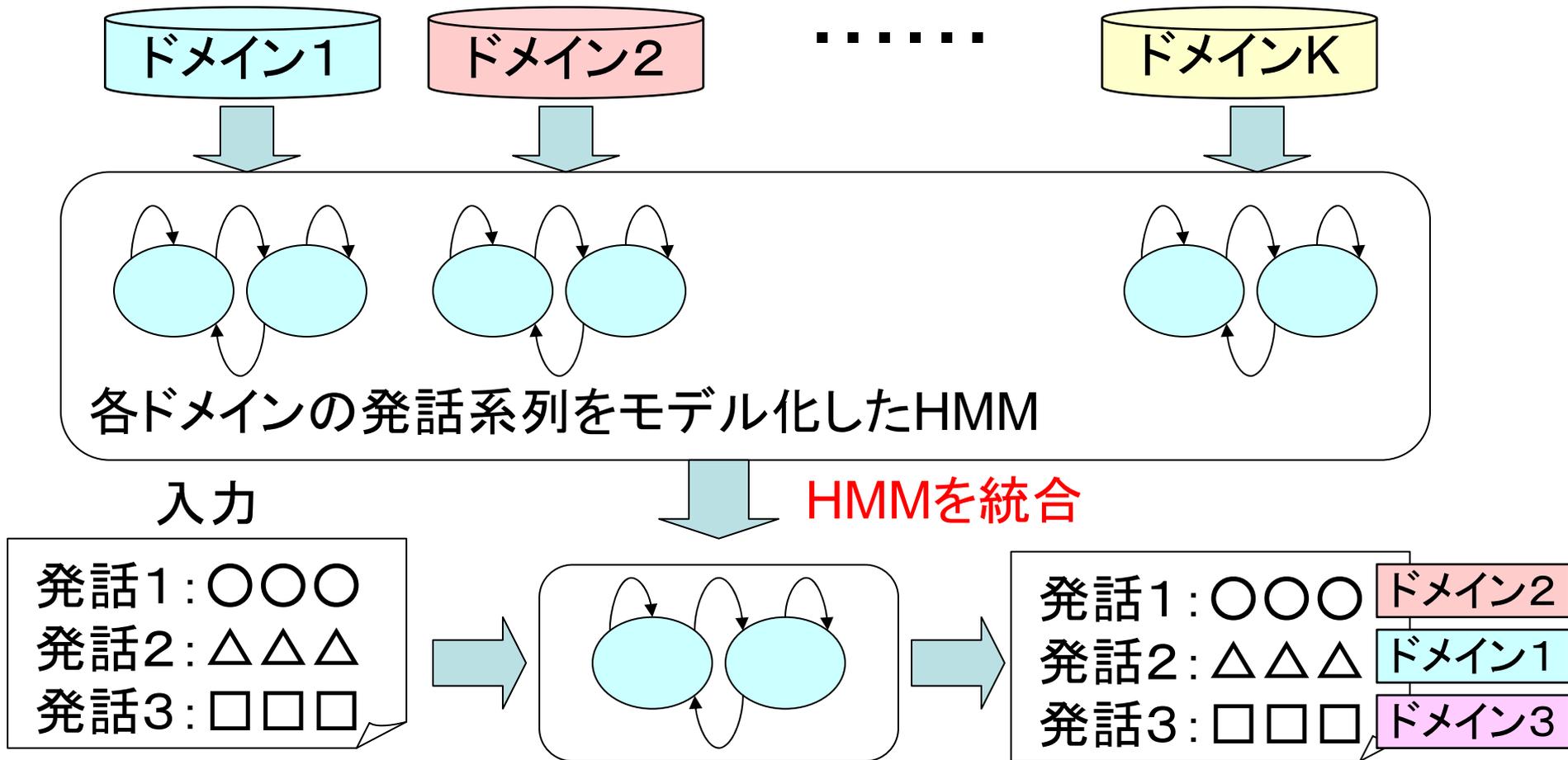
**要約**

ユーザ: はい. **共通**

オペレータ: それでは, お客様番号をいただけますか? **契約**

当該ドメインに特徴的な発話のみを抽出

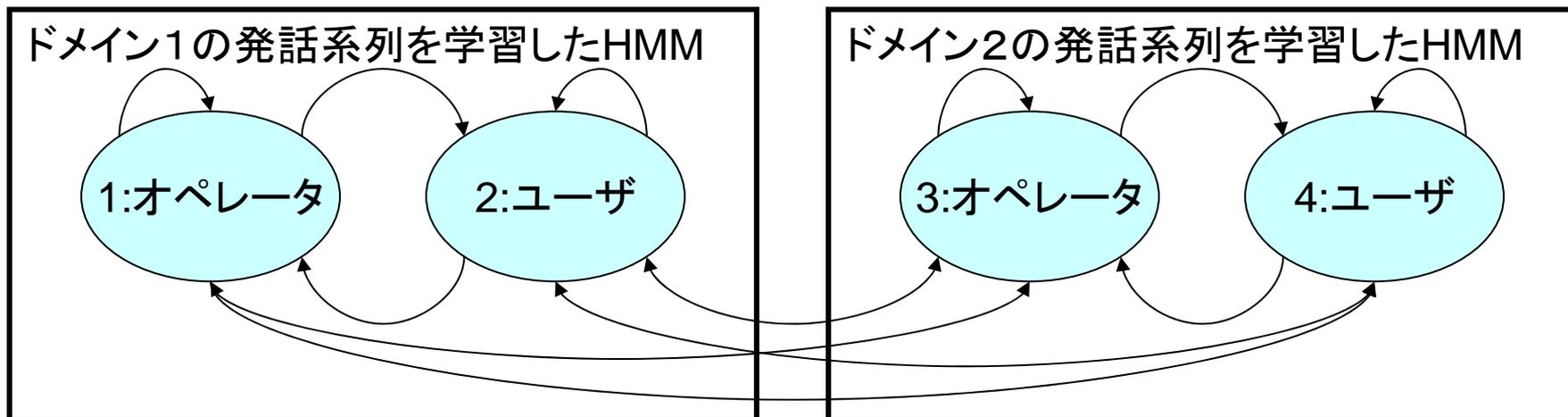
# アプローチ(つづき)



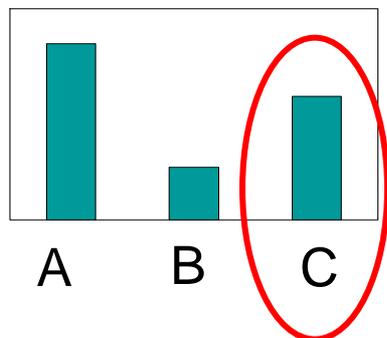
各発話を出力する最尤ドメインを推定しラベルを付与 7

# エルゴディック(単純なモデル)

個々のドメインから学習したHMMを等確率で接続



ドメイン1のオペレータの出力分布

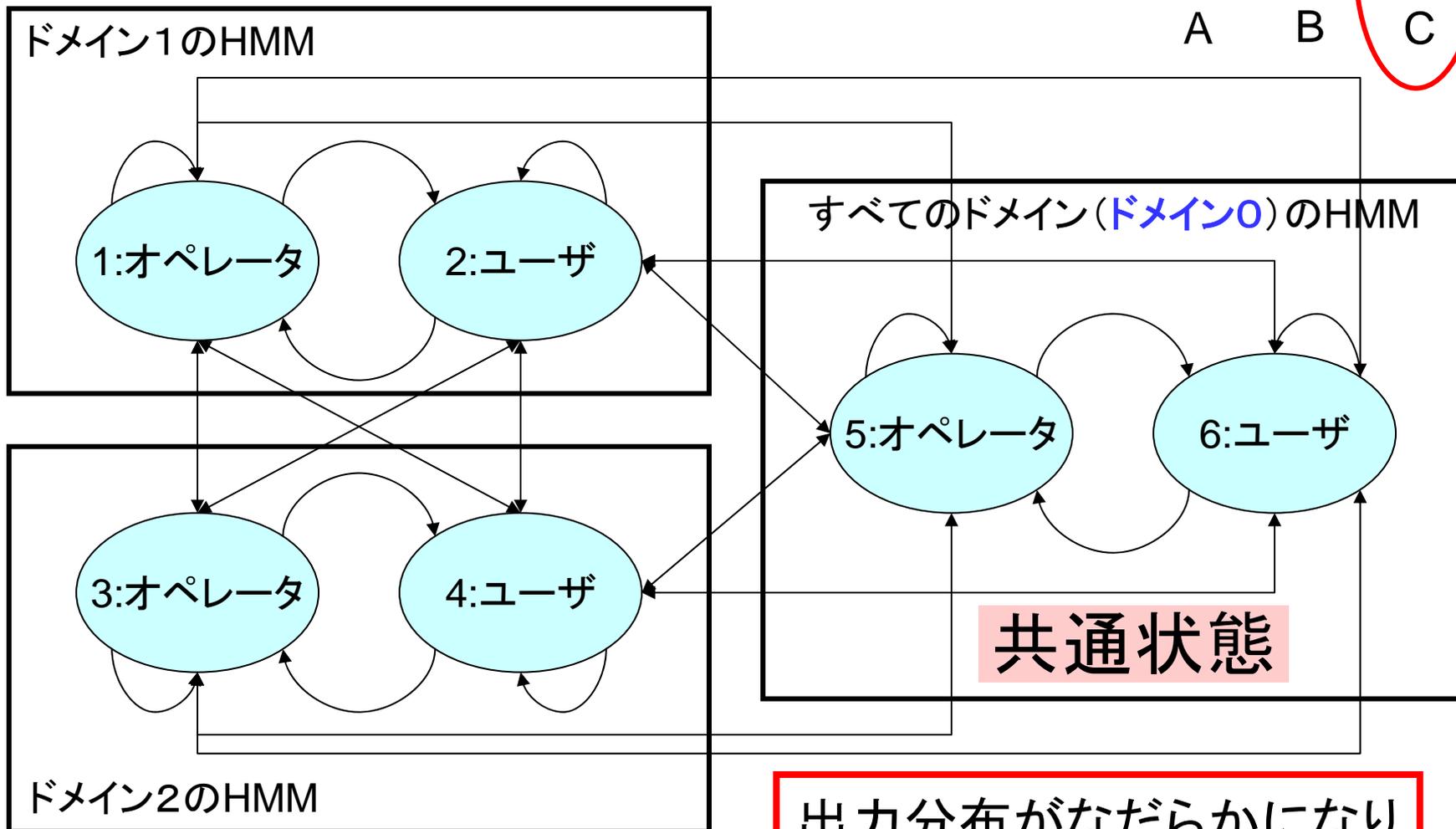


ドメイン2のオペレータの出力分布



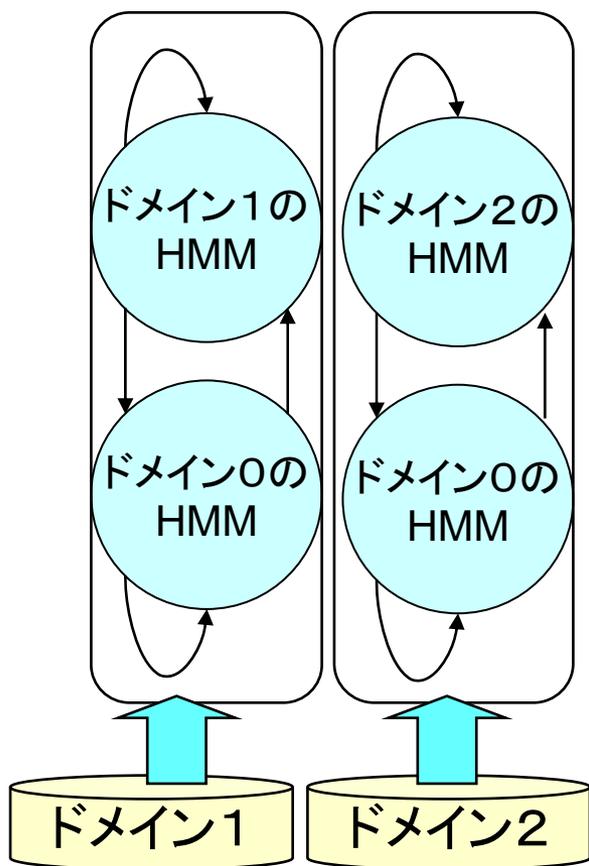
共通的な発話が  
必ず一方のドメインに分類  
されてしまう

# 共通状態の導入



出力分布がなだらかになり  
最尤パスに選ばれない

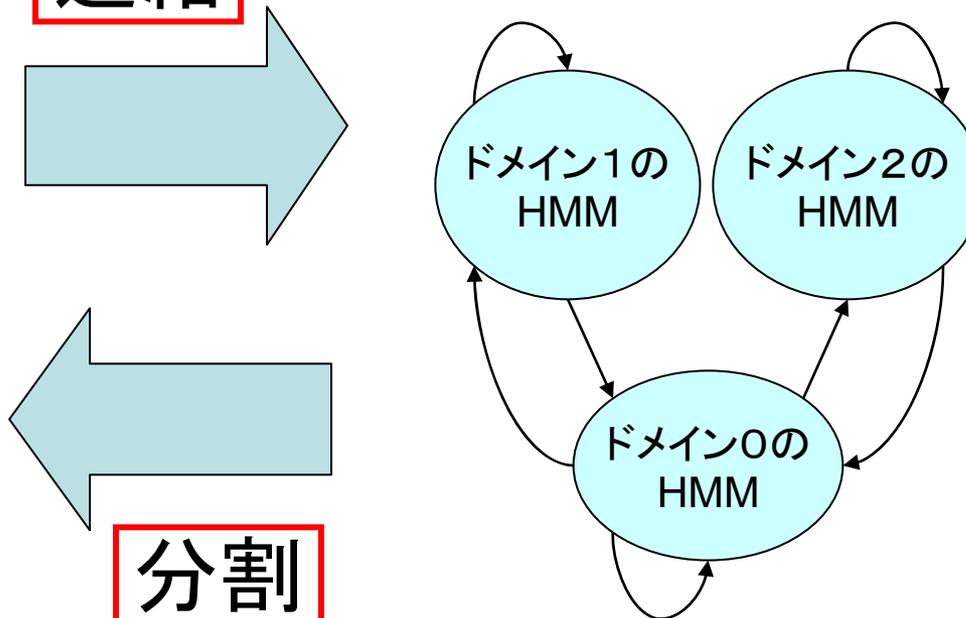
# 共通状態と連結学習の導入



再学習

連結

ドメイン0の確率テーブルを平均化



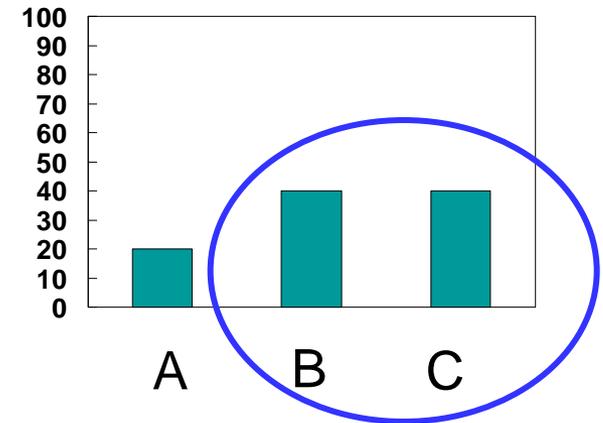
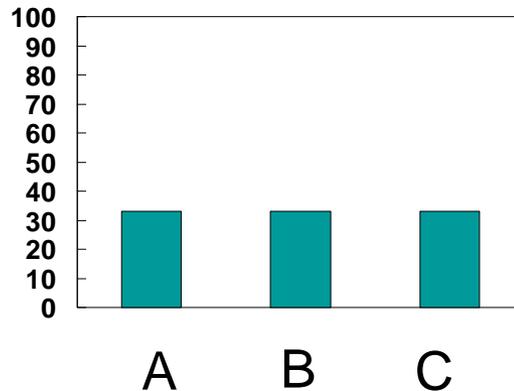
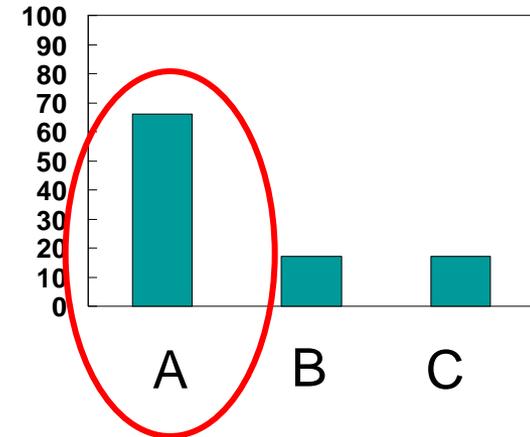
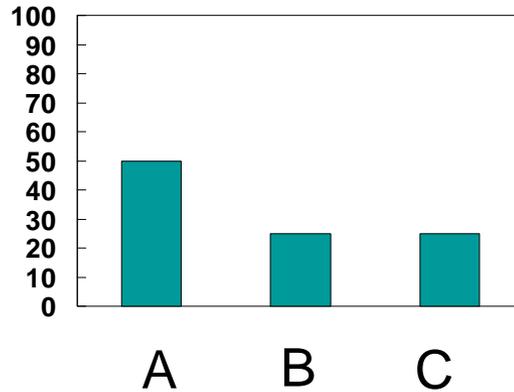
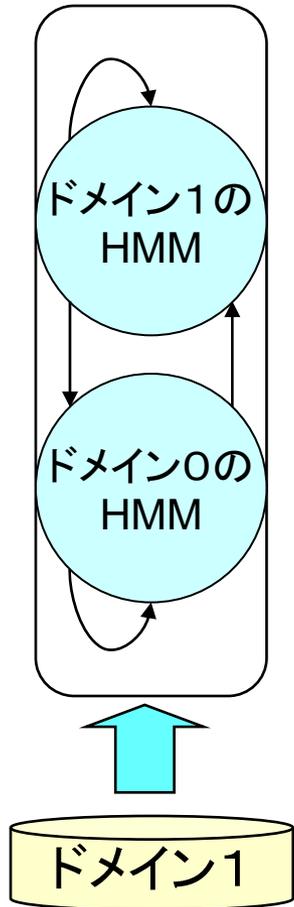
分割

ドメイン0の遷移確率を分配

すべてのドメインについて尤度の改善がなくなるまで連結と分割を繰り返す

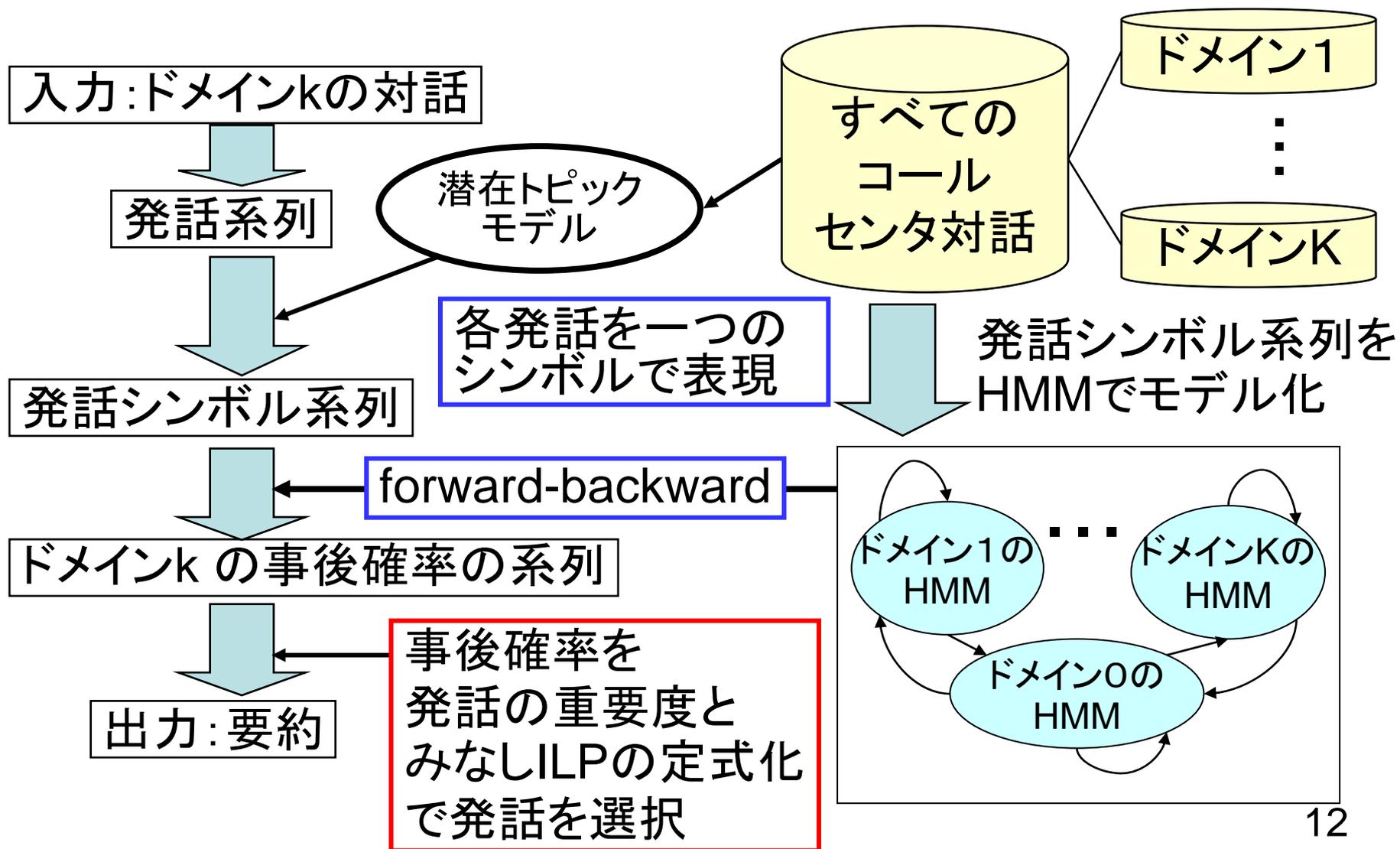
# 連結学習の仕組み

EMによって、よく出現する発話の確率が高まる



ドメイン0では共通的な発話の確率が高まる  
(分布がなだらかでなくなる)

# 要約処理の流れ



# データ収集

- 6ドメインのコールセンタ対話を収集
  - 金融, インターネットサービスプロバイダ, 自治体への問い合わせ, 通信販売, PCサポート, 電話についての問い合わせ(全698対話)
  - オペレータ役は実際のオペレータ業務経験者
  - シナリオに基づき電話で通話
- 書き起こしと要約正解の作成
  - すべての発話を書き起こし
  - 250文字, 500文字要約の正解を発話抽出により作成(アノテーター一致率は0.43, 0.53)

# 実験

- 提案手法
  - 共通状態と連結学習を用いたHMM
- ベースライン
  - エルゴディック
  - 共通状態を付加したもの
- 学習データ
  - 391対話で学習し, 308対話で評価
  - 書き起こしを使用
- 学習データサイズを徐々に増やして実験
- 評価尺度: 発話抽出のF値
- 正解は片方のアノテータのデータ

# 結果（発話抽出のF値）

## 250文字要約

| 学習セット  | (a) エルゴ | (b)+共通状態            | (c)+連結学習                     |
|--------|---------|---------------------|------------------------------|
| set1   | 0.211   | 0.220 <sup>a</sup>  | <b>0.254</b> <sup>aabb</sup> |
| set1-2 | 0.219   | 0.229 <sup>aa</sup> | <b>0.256</b> <sup>aabb</sup> |
| set1-3 | 0.226   | 0.228               | <b>0.248</b> <sup>aabb</sup> |
| set1-4 | 0.225   | 0.235 <sup>a</sup>  | <b>0.268</b> <sup>aabb</sup> |
| set1-5 | 0.226   | 0.237 <sup>a</sup>  | <b>0.263</b> <sup>aabb</sup> |

## 500文字要約

| 学習セット  | (a) エルゴ | (b)+共通状態            | (c)+連結学習                     |
|--------|---------|---------------------|------------------------------|
| set1   | 0.395   | 0.397               | <b>0.432</b> <sup>aabb</sup> |
| set1-2 | 0.403   | 0.406               | <b>0.432</b> <sup>aabb</sup> |
| set1-3 | 0.403   | 0.405               | <b>0.431</b> <sup>aabb</sup> |
| set1-4 | 0.406   | 0.416 <sup>aa</sup> | <b>0.444</b> <sup>aabb</sup> |
| set1-5 | 0.407   | 0.412               | <b>0.431</b> <sup>aabb</sup> |

# まとめ

- コールセンタ対話の要約手法を提案
  - 共通状態と連結学習を用いたHMM
  - 要約の正解データが不要
- 実験的に有効性を検証
  - ベースラインと比較し有意に要約結果が改善
  - 少ないデータからでも効率的に各ドメインの特徴を学習可能
- 今後の課題
  - 音声認識結果への本手法の適用