

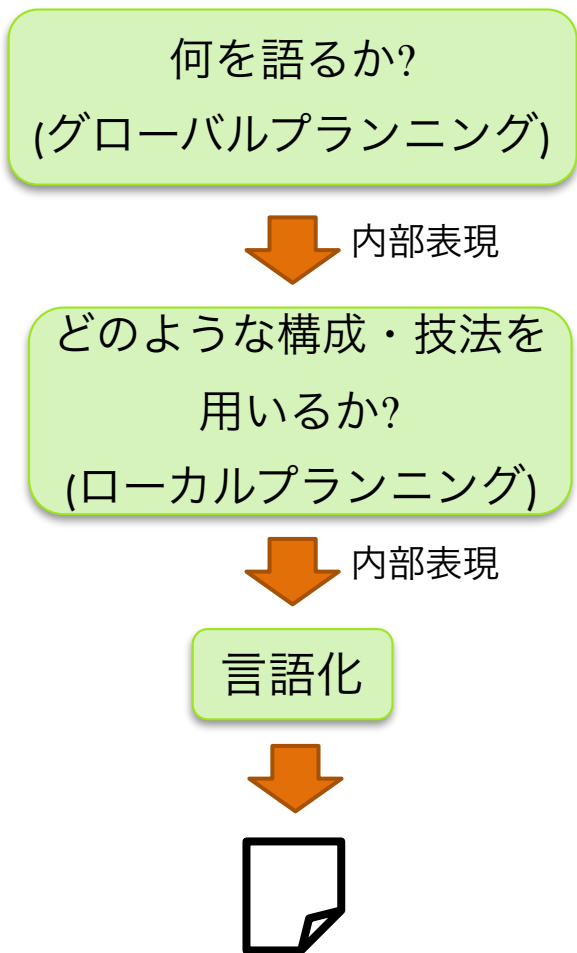
メタファー写像と後編集を利用する 物語文章生成フレームワーク

電気通信大学大学院 情報理工学研究科

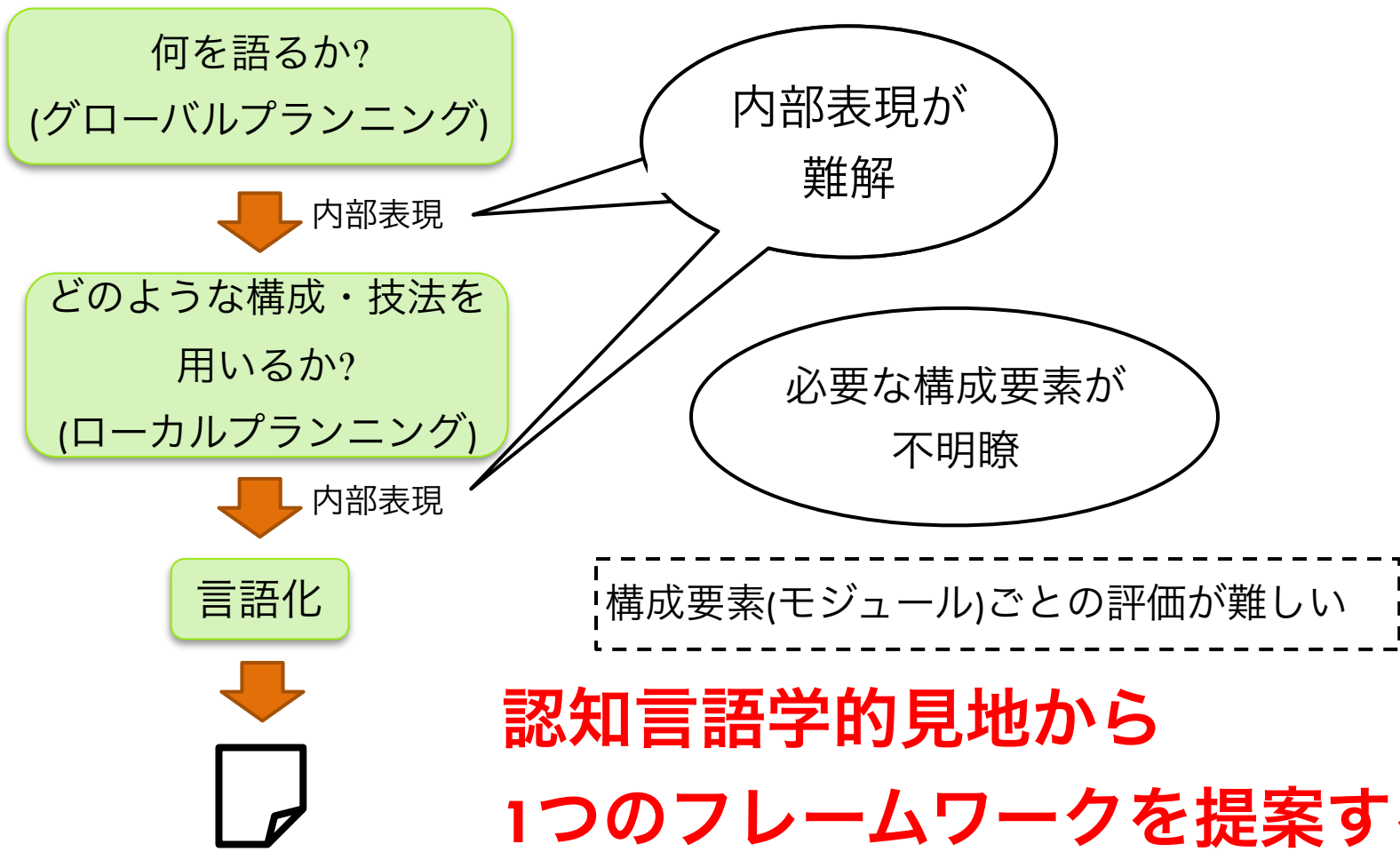
松吉 俊、内海 彰

2018年6月6日(水)

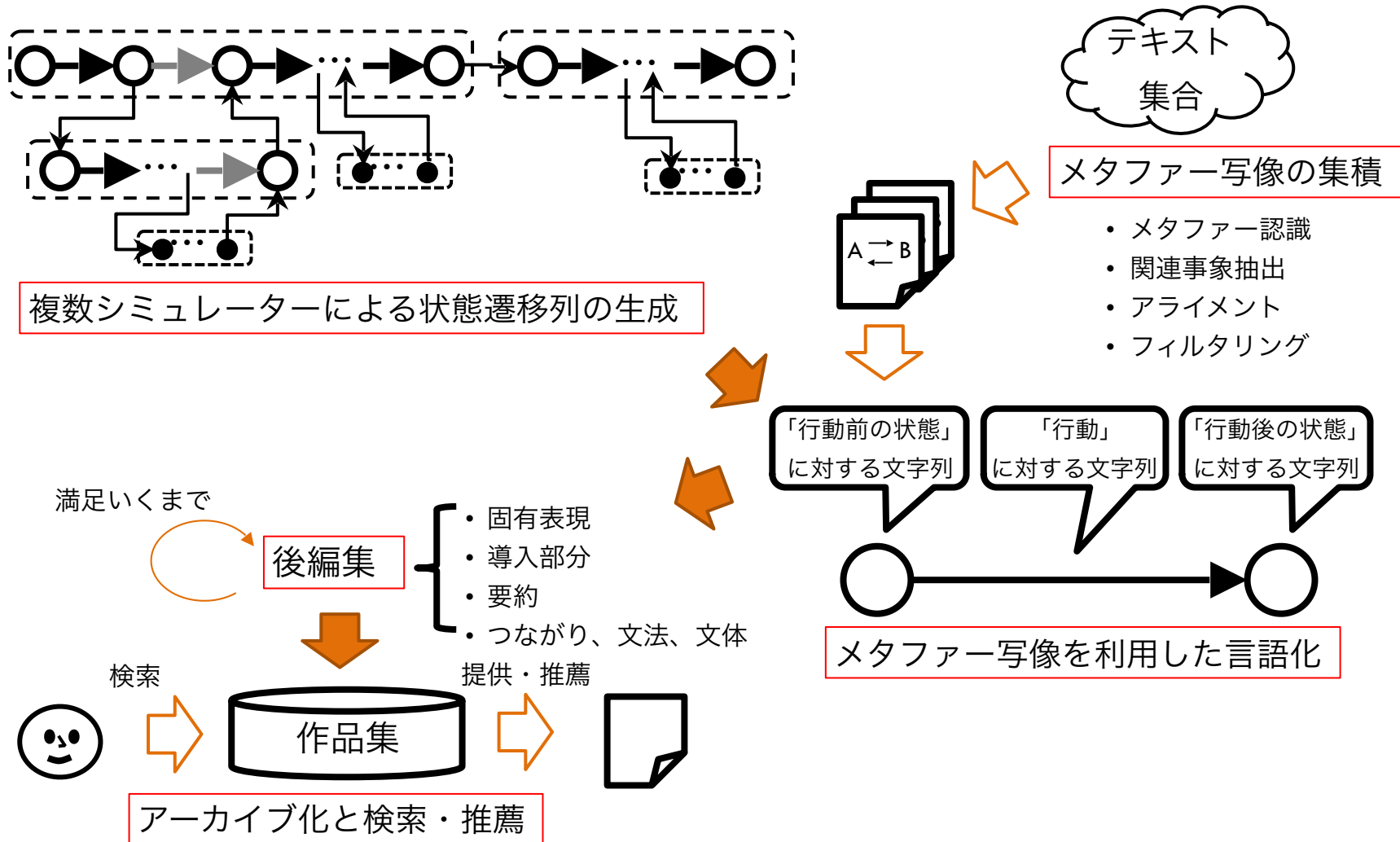
従来手法



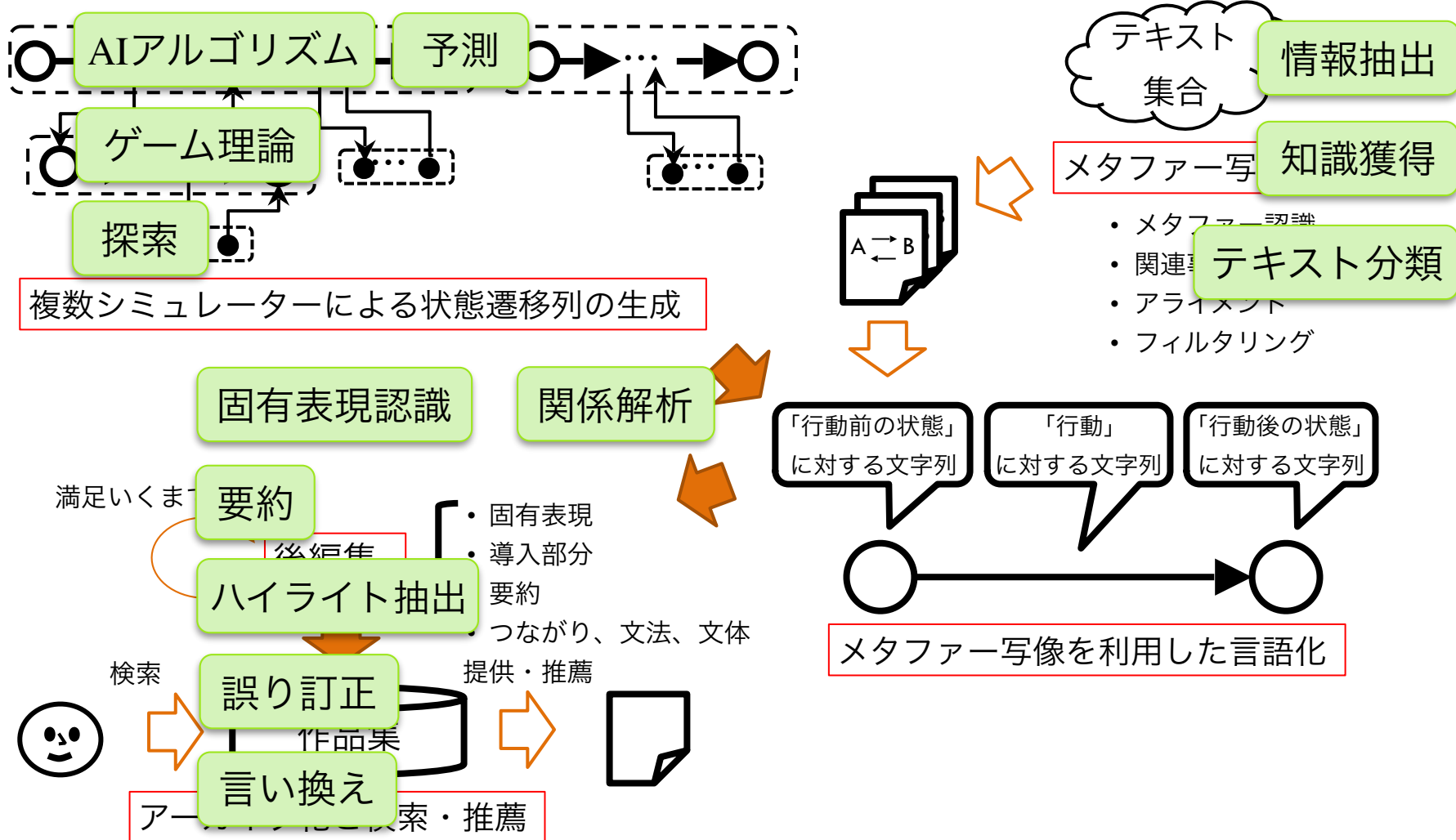
従来手法の(おそらく)短所



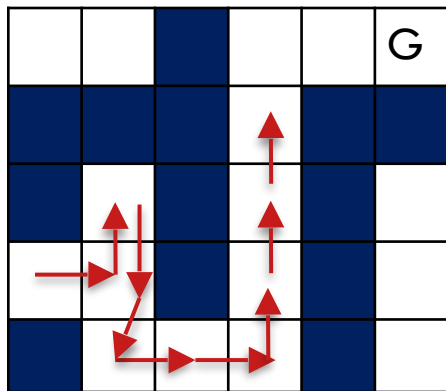
提案手法



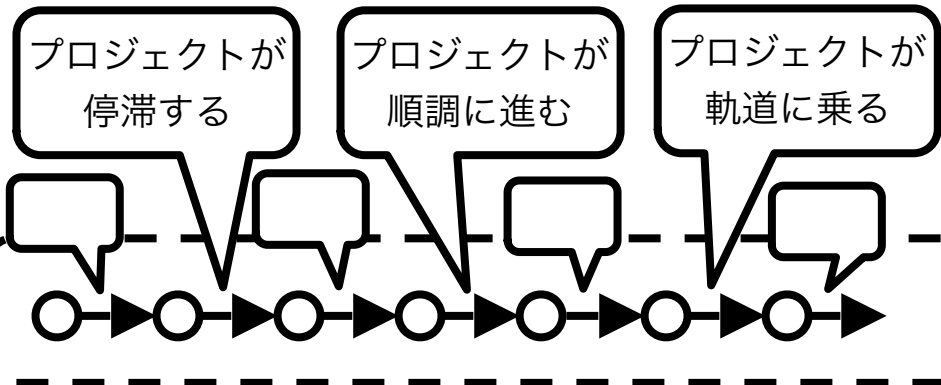
提案手法のモジュール



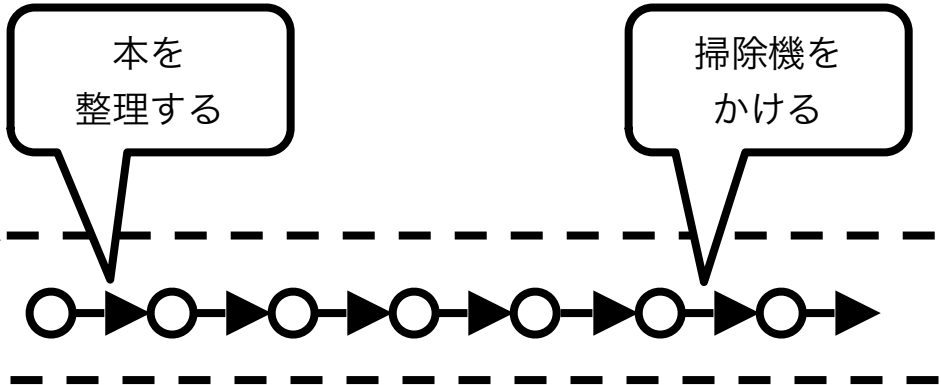
メタファー写像に基づく物語生成



プロジェクト進行



部屋の片付け



物語文章生成の例

```

. . . . .
r . . . . .
P . . . . .
. . . . p . k p
. . . . K . . .
R . . . . P .
. . . . .
. . . . .
8/r7/P7/4p1kp/4K3/R5P1/8/8 b - - 1 67

```

```

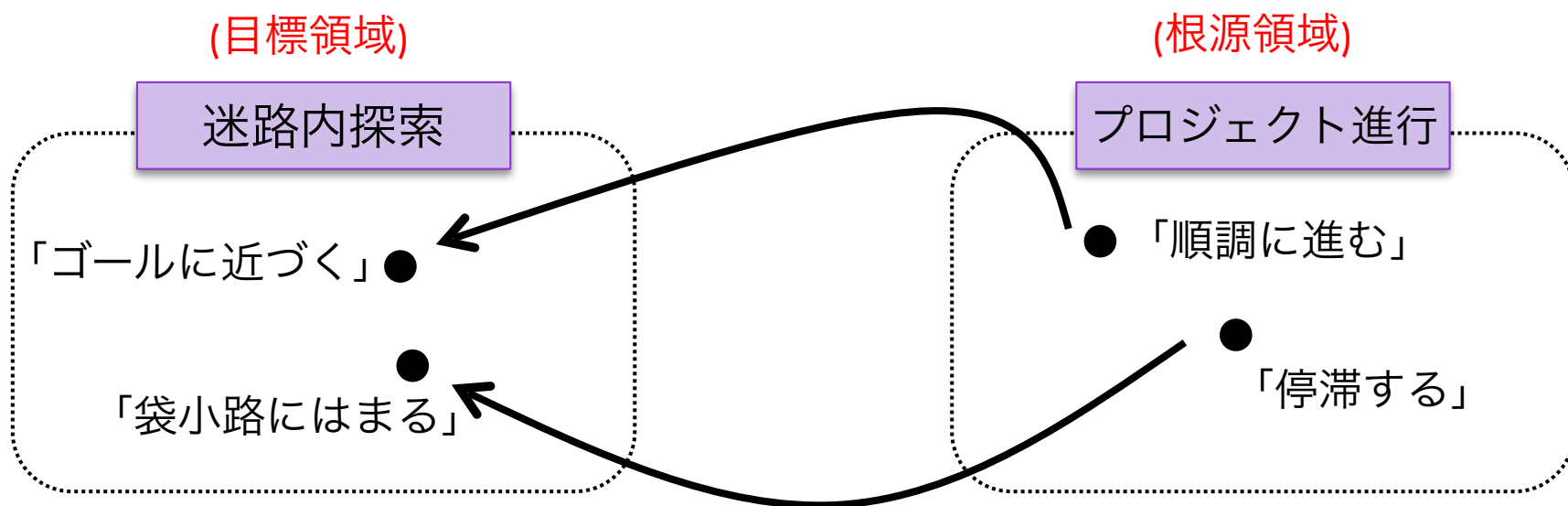
部屋の掃除を開始した
本を整理した
本を整理した
掃除機をかけた
机の引き出しを整理した
本を整理した
掃除機をかけた
本を整理した
本を整理した
机の引き出しを整理した
掛け布団を干した
幼い頃から使っていたブランケットを見つけた
悩んだ結果、ブランケットを捨てることにした
本を整理した
部屋の掃除が無事終わった

```

```
[apricot-(matuyosi):~/work/uec/1jikken/NI G/chess 12468] *
```

メタファー写像

□ 2つの意味領域間の連結



... 迷路で袋小路にはまったかのように、このプロジェクトが停滞している。 ...

... 今月は、プロジェクトがゴールに近づいた。 ...

不変性原理 (Lakoff 1993)

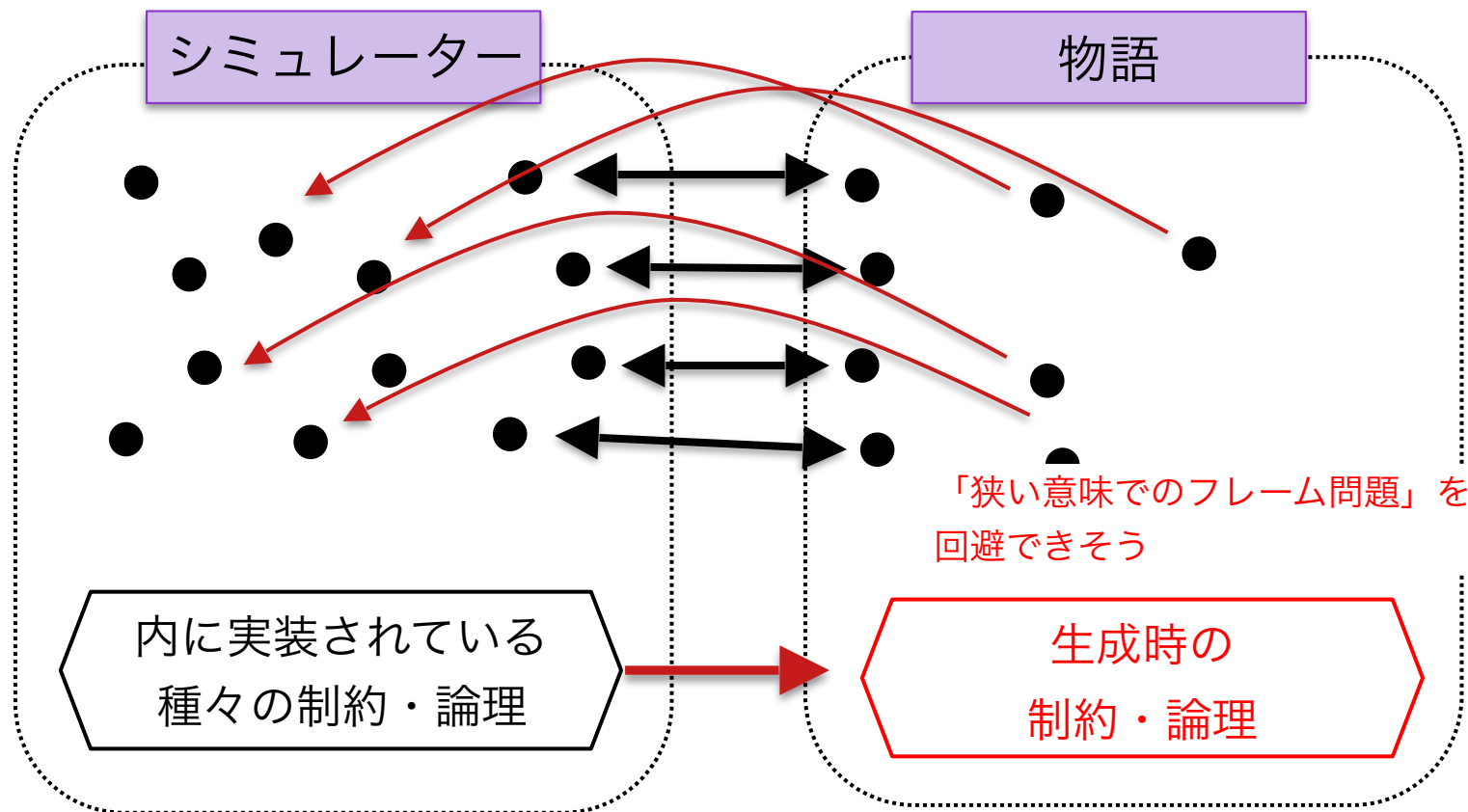
- 「写像において、根源領域の表現間の(位相)構造は目標領域内でも保たれる。」

- この原理に則ると、
 - 任意の2つの意味領域間において
 - 数個のメタファー写像の存在が確認されれば、
 - 根源領域内の他の表現も、位相構造を保ったまま
 - 目標領域内に対応づけることができる
 - 見込みが非常に高い

不変性原理の利用

うまくやれば、
メタファー写像の自動集積が
できそう

テキスト
集合



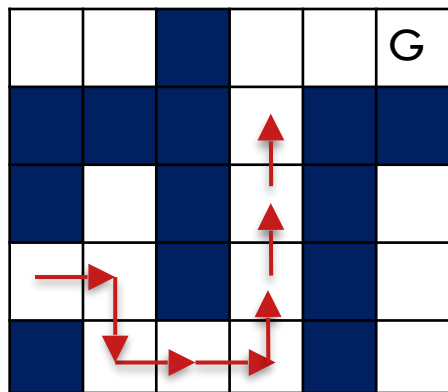
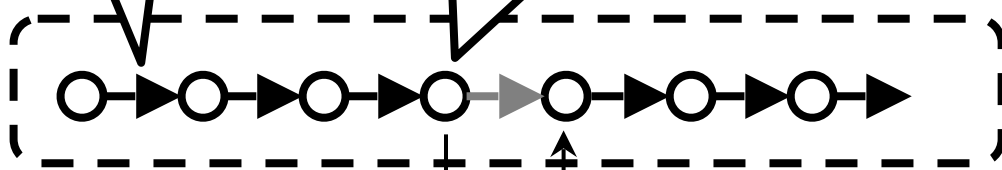
任意の粒度で物語生成



部屋の片付け

本を
整理する

幼い頃から使っていた
ブランケットを見つけた。悩ましい



悩み

捨ててもいいかな
と思う

まだ悩みたい



もう悩まない

発表内容

1. 先行研究
2. 提案手法
3. 物語生成の構造: オーケストラ
4. 評価実験
5. 後編集
6. 提案手法の長所
7. まとめ

物語生成の構造: オーケストラ

オーケストラ

楽器

- 名前
- シミュレーター

演奏者

- 演奏モデル
- 楽器
- パート譜
- 演奏する()

指揮者

- 統制を取る()

パート譜

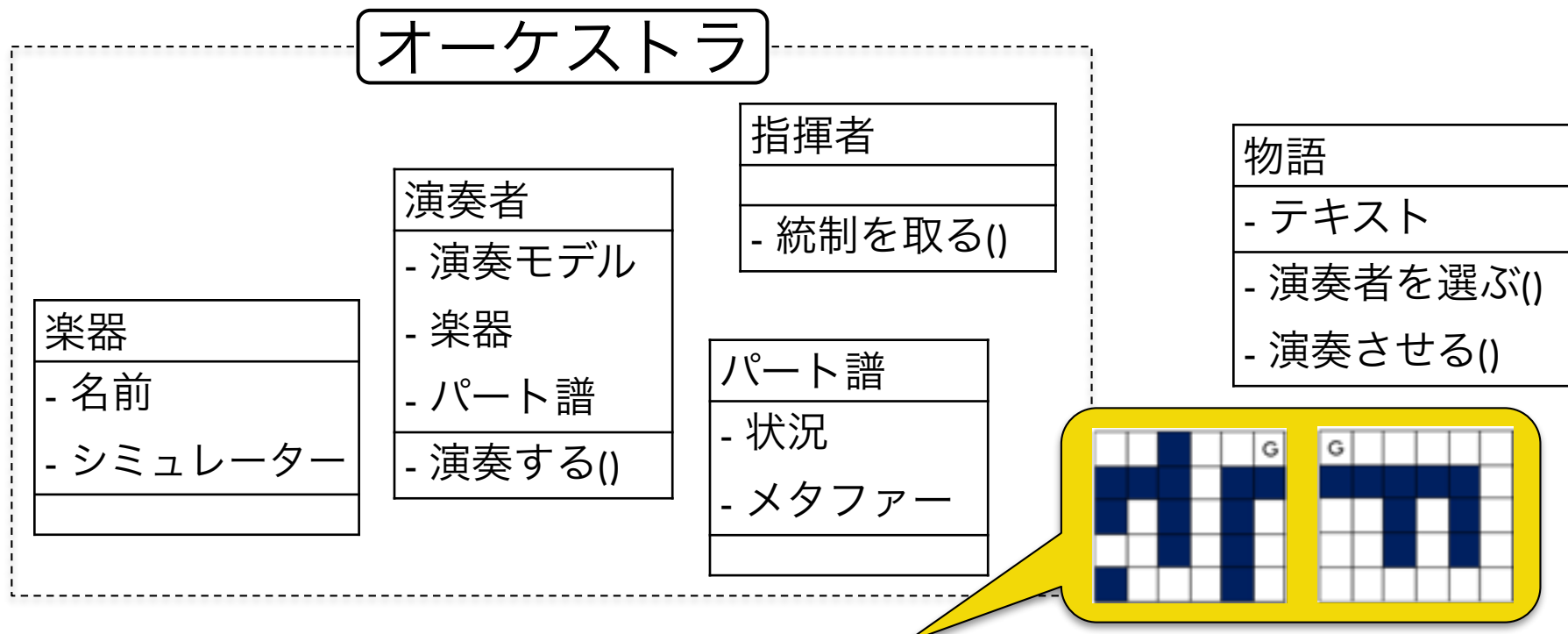
- 状況
- メタファー

物語

- テキスト
- 演奏者を選ぶ()
- 演奏させる()



物語生成の構造: オーケストラ



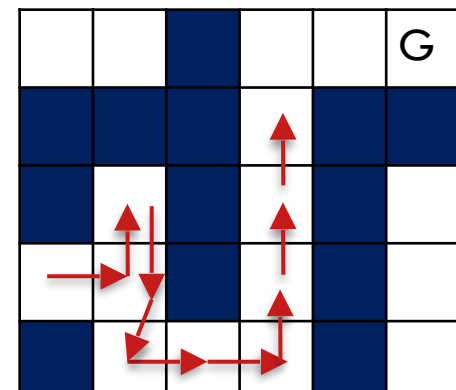
ゲーム数 × AI数 × 局面数 × メタファー数

= 生成できる基本物語の数

みんなが参加できるフレームワーク

バリエーションの評価実験

- ゲーム: 迷路探索
- AI (= エージェントの移動モデル): 強化学習
- 局面とメタファー: 製作を依頼
- 被験者: 電通大 総合情報学科 経営情報学コース 3年生
44名
- 必修のコンピューター演習授業
- 迷路探索に対するメタファー写像
 - ゴールに近づく: 2種類(右移動と上移動)
 - ゴールから離れる: 2種類(左移動と下移動)
- 授業資料を一般公開:
 - <http://www.cl.inf.uec.ac.jp/lec/>



生成された物語の一覧

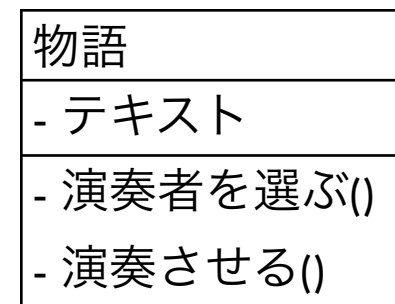
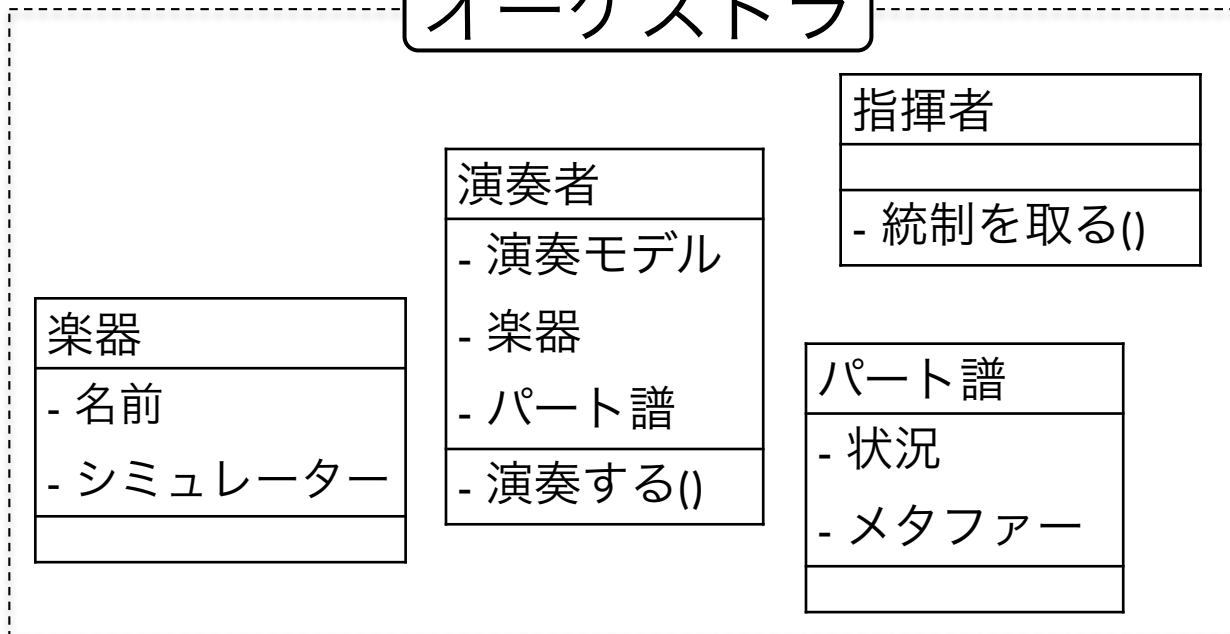
分類	物語	数
構築	植物、料理 (3)、レポート (5)、小説、研究、プロジェクト	12
恋愛	恋愛 (10)、浮気	11
自分向上	試験勉強 (4)、ルックス (3)、リハビリ、転職、解脱の修行	10
競技	サッカー (5)、野球、バスケ、テニス、陸上、ある運動部、将棋界	11
競争・営業	企業競争 (2)、宇宙戦争、バトルロワイヤル、市長業務、村の統治、居酒屋経営、FA宣言、音楽バンド (2)、芸人、ゲーム、パチプロ、賭け事	14
発見	事件捜査 (4)、人参畑を目指すウサギ、伝説の剣の復活	6
退治	魔王討伐、魔物討伐、虫退治、市民革命	4
日常生活	大学生活 (5)、進路の悩み、旅行計画、見覚えのない場所、拾い物、雨宿り、風邪	11
模倣・二次創作	走れメロス、「本能寺の変」の前日、ポケモントレーナー、名探偵コナン、NARUTO、パワプロくん、魔法少女まどか☆マギカ、君の名は。	8
合計		87

生成された物語の一覧

分類	物語	数
構築	植物、料理 (3)、レポート (5)、小説、研究、プロジェクト	12
恋愛	恋愛 (10)、浮気	11
自分向上	試験勉強 (4)、ルックス (3)、リハビリ、転職、解脱の修行	10
競技	サッカー (5)、野球、バスケ、テニス、陸上、ある運動部、将棋界	11
競争・営業	企業競争 (2)、宇宙戦争、バトルロワイヤル、市長業務、村の統治、居酒屋経営、FA宣言、音楽バンド (2)、芸人、ゲーム、パチプロ、賭け事	14
発見	事件捜査 (4)、人参畑を目指すウサギ、伝説の剣の復活	6
退治	魔王討伐、魔物討伐、虫退治、市民革命	4
日常生活	大学生活 (5)、進路の悩み、旅行計画、身 自動収集時は注意が必要 り、風邪	11
模倣・二次 創作	走れメロス、「本能寺の変」の前日、ポケモントレーナー、名探偵コナン、NARUTO、パワプロくん、魔法少女まどか☆マギカ、君の名は。	8
合計		87

現在の実装状況

オーケストラ

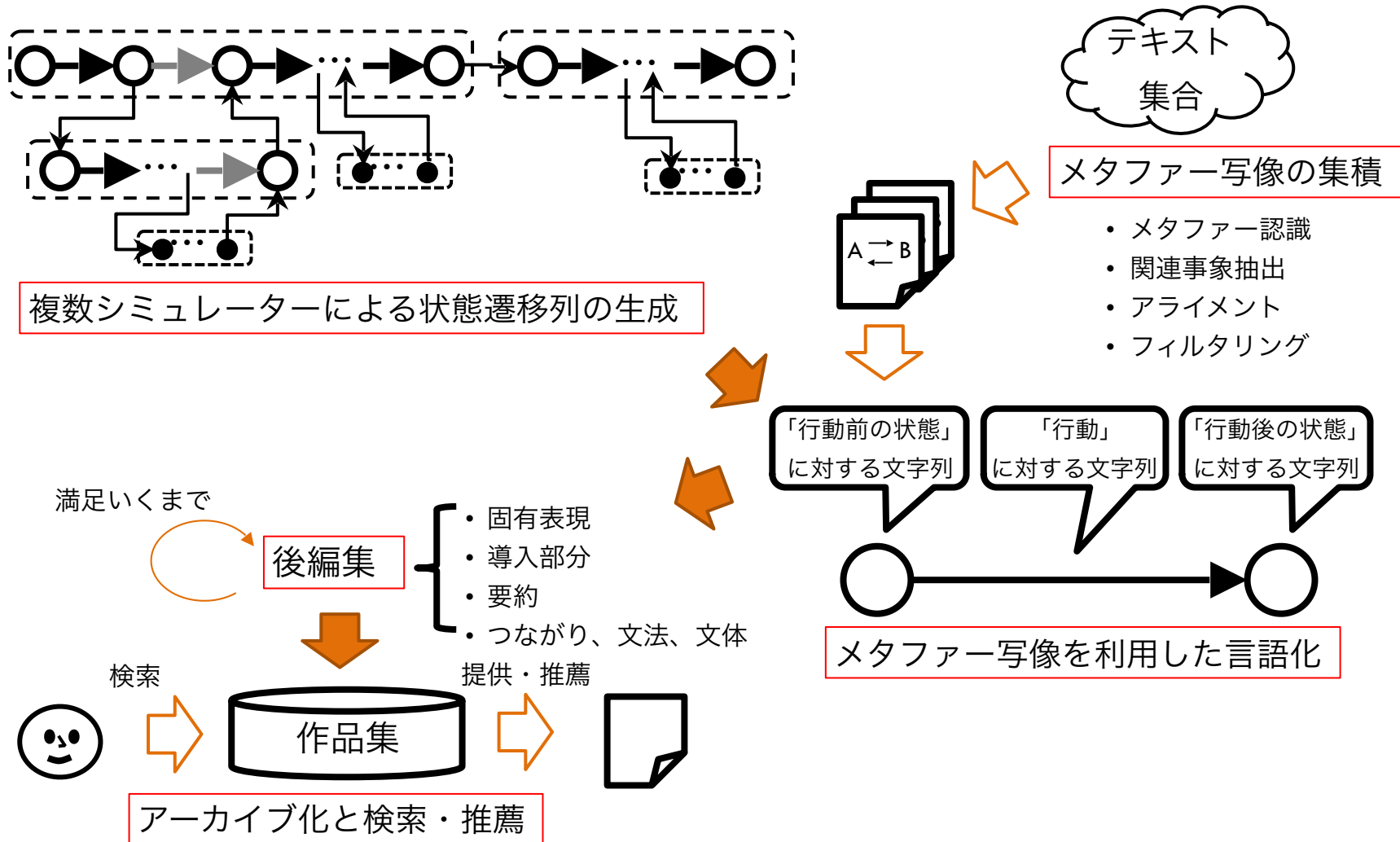


ゲーム	AI	局面	メタファー
チェス	Stockfish	任意	2
迷路探索	強化学習	任意	90

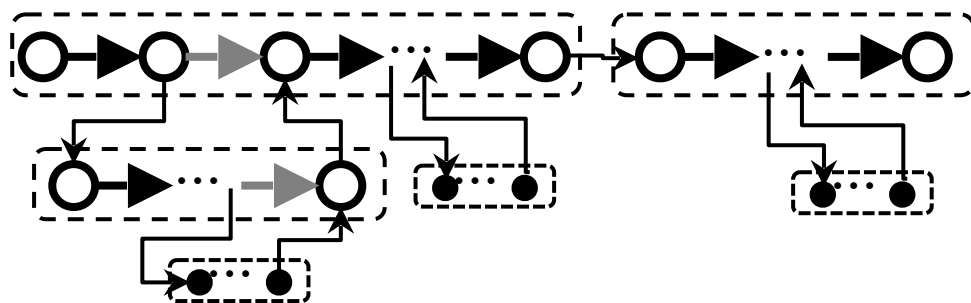
発表内容

1. 先行研究
2. 提案手法
3. 物語生成の構造: オーケストラ
4. 評価実験
5. 後編集
6. 提案手法の長所
7. まとめ

後編集がとても重要



文章 → (変換) → 文章

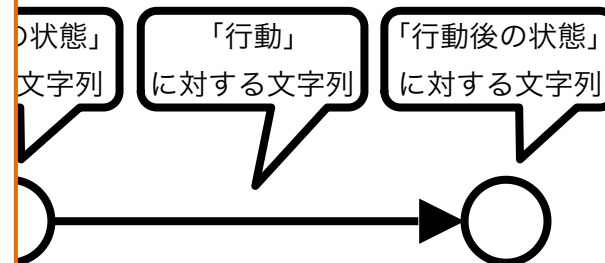
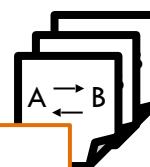
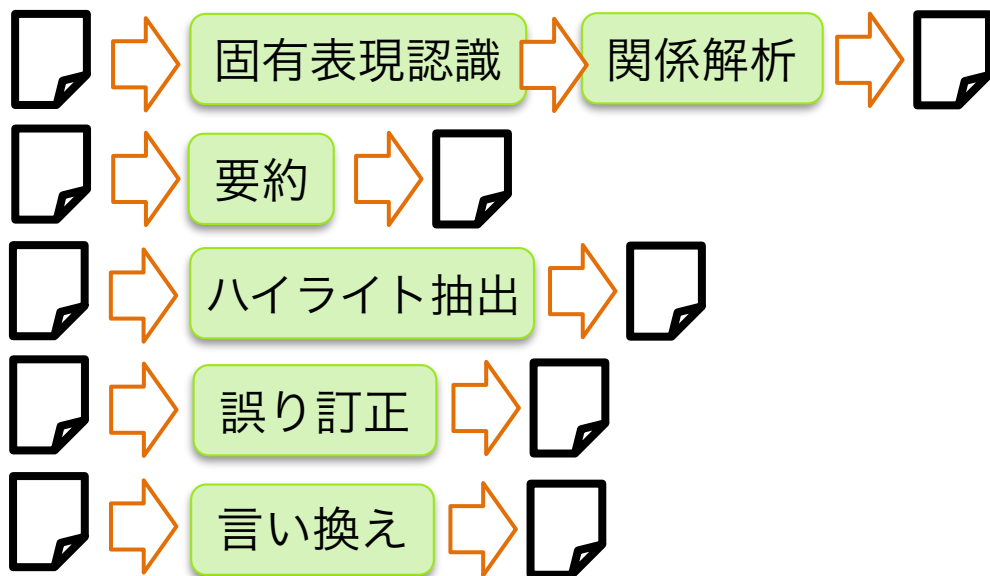


テキスト
集合

メタファー写像の集積

- メタファー認識
- 関連事象抽出
- アライメント
- フィルタリング

満足いくまで順不同で繰り返したい



メタファー写像を利用した言語化

固有表現認識と関係解析

- 固有表現の置き換えが必要

ジョンがイギリス内を旅行

ビル → ジョン

絶対的な置き換え

私の叔父 → ジョン

パリ → ロンドン

相対的な置き換え

山手線 → サークル線

フランス ← イギリス

cf. SemEval-2012 Task 2

東京 ← ロンドン

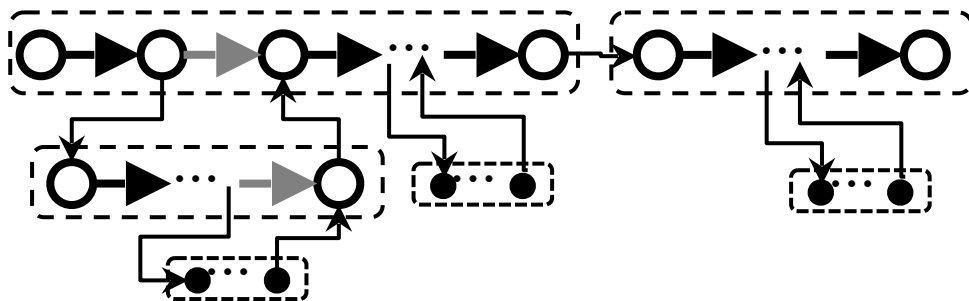
ビルは、その日、パリのホテルで一泊することにした



私の叔父は、山手線に揺られながら、昨日のことを考えていた



次に取り組みたいタスク

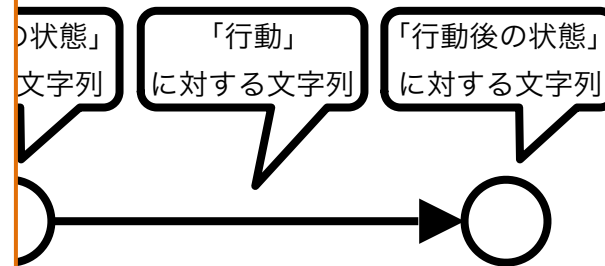
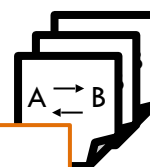
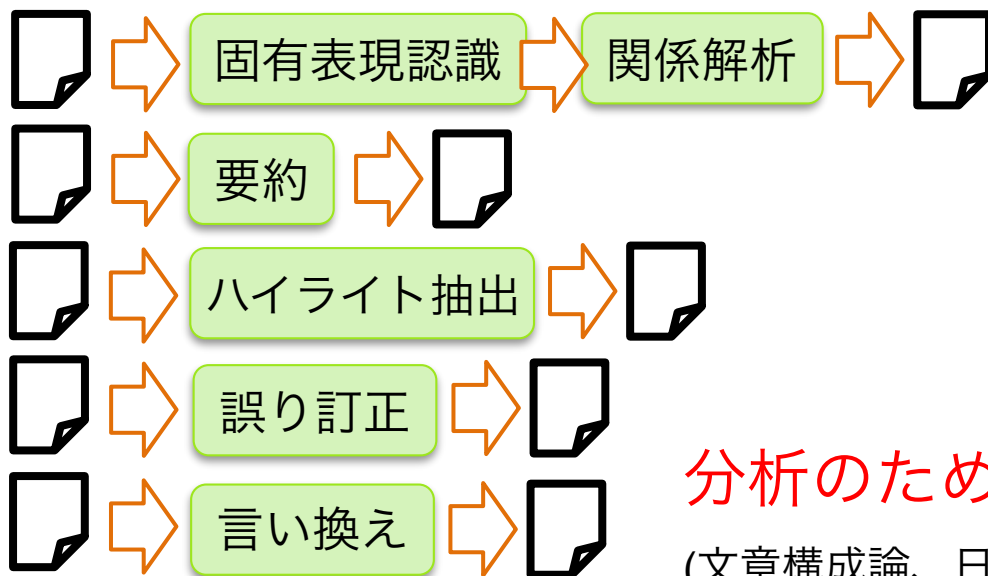


テキスト
集合

メタファー写像の集積

自動集積システムの
試作

満足いくまで順不同で繰り返したい



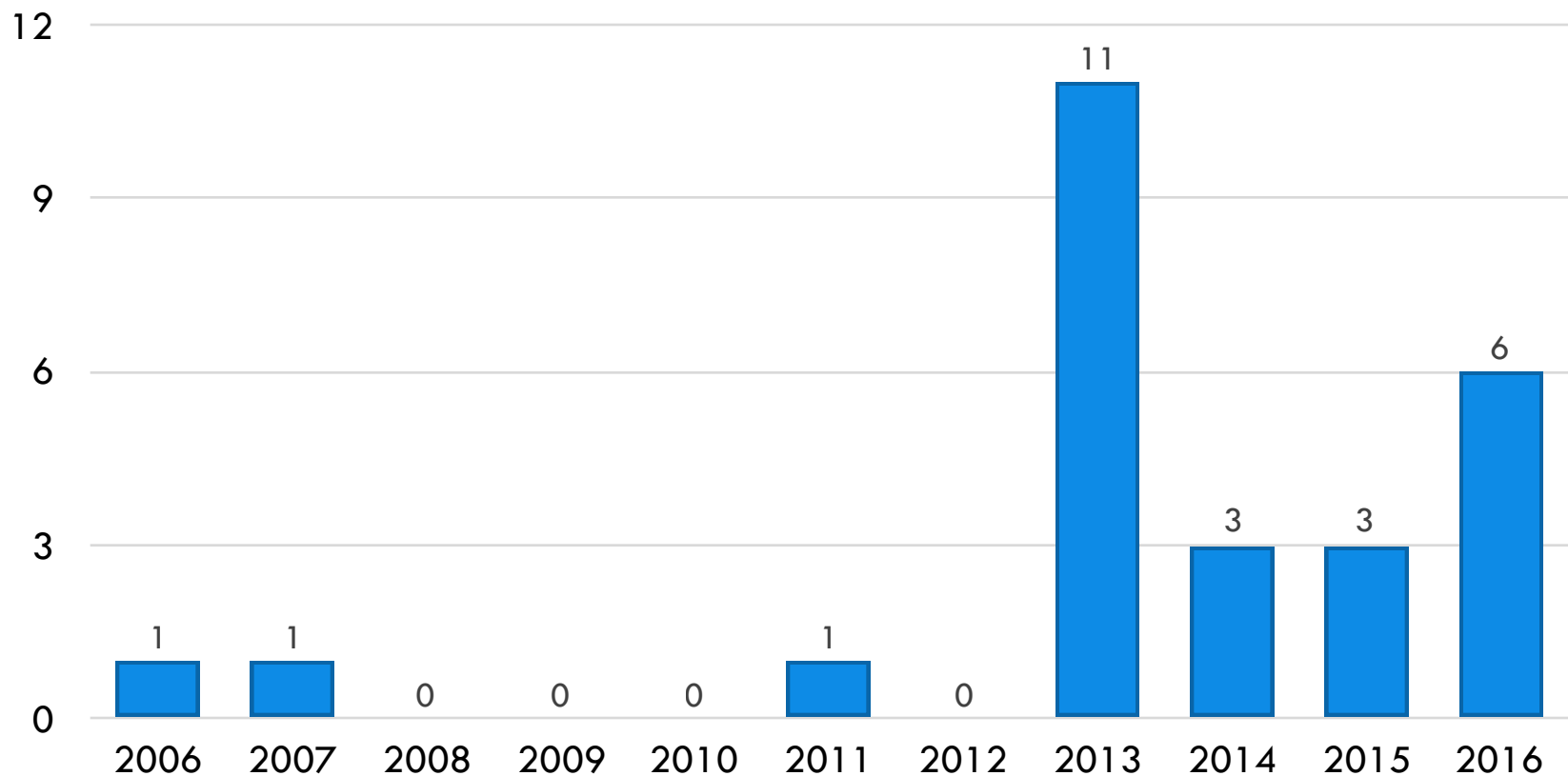
メタファー写像を利用した言語化

分析のためのコーパス構築

(文章構成論、日本語教育、テクニカルライティング)

メタファー自動認識の先行研究

論文投稿数

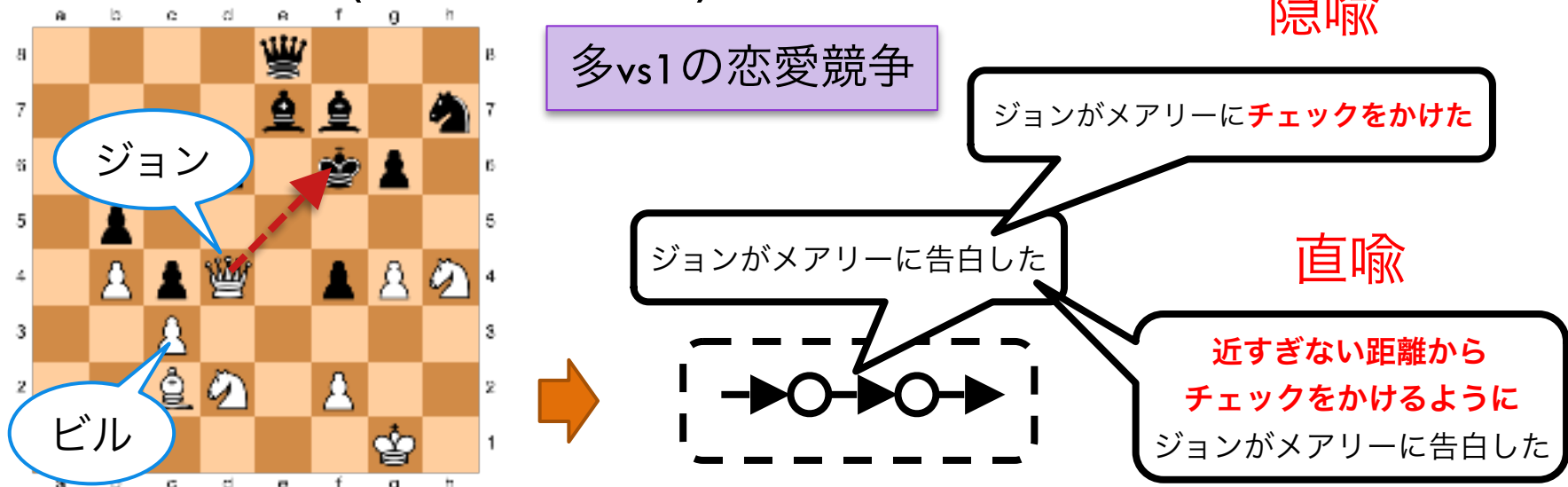


発表内容

1. 先行研究
2. 提案手法
3. 物語生成の構造: オーケストラ
4. 評価実験
5. 後編集
6. 提案手法の長所
7. まとめ

提案手法の長所

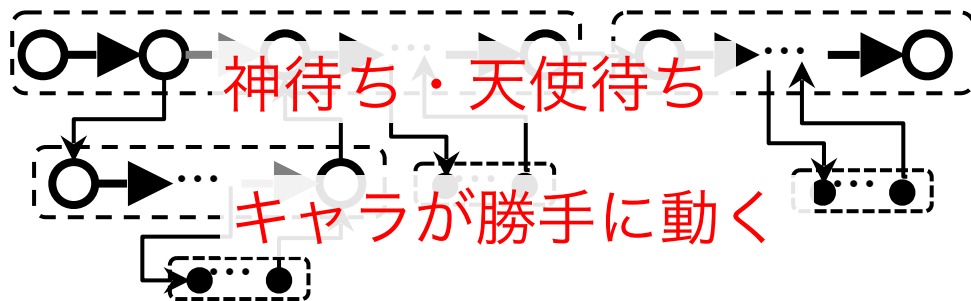
- 比喩表現(直喩・隠喩)が自然に生成可能



- メタ質問への応答可能。要望に応じた途中改変可能

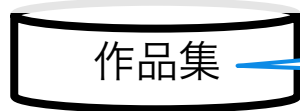


人間の作家との比較



複数シミュレーターによる状態遷移列の生成

「満足いくまで後編集」は、
人間も同様
(独自の内部表現の変換として
実装するのはもったいない)



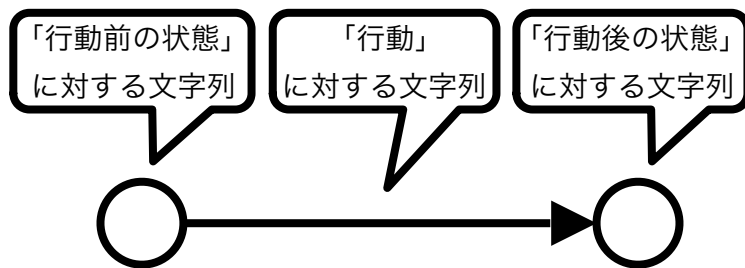
アーカイブ化と検索・推薦



メタファー写像の集積

作家の引き出し
(体験や伝聞)

- メタファー認識
- 関連事象抽出
- アライメント
- フィルタリング



人間が創作しうる物語集合のうち、
ほんの一部

差分を観察したい

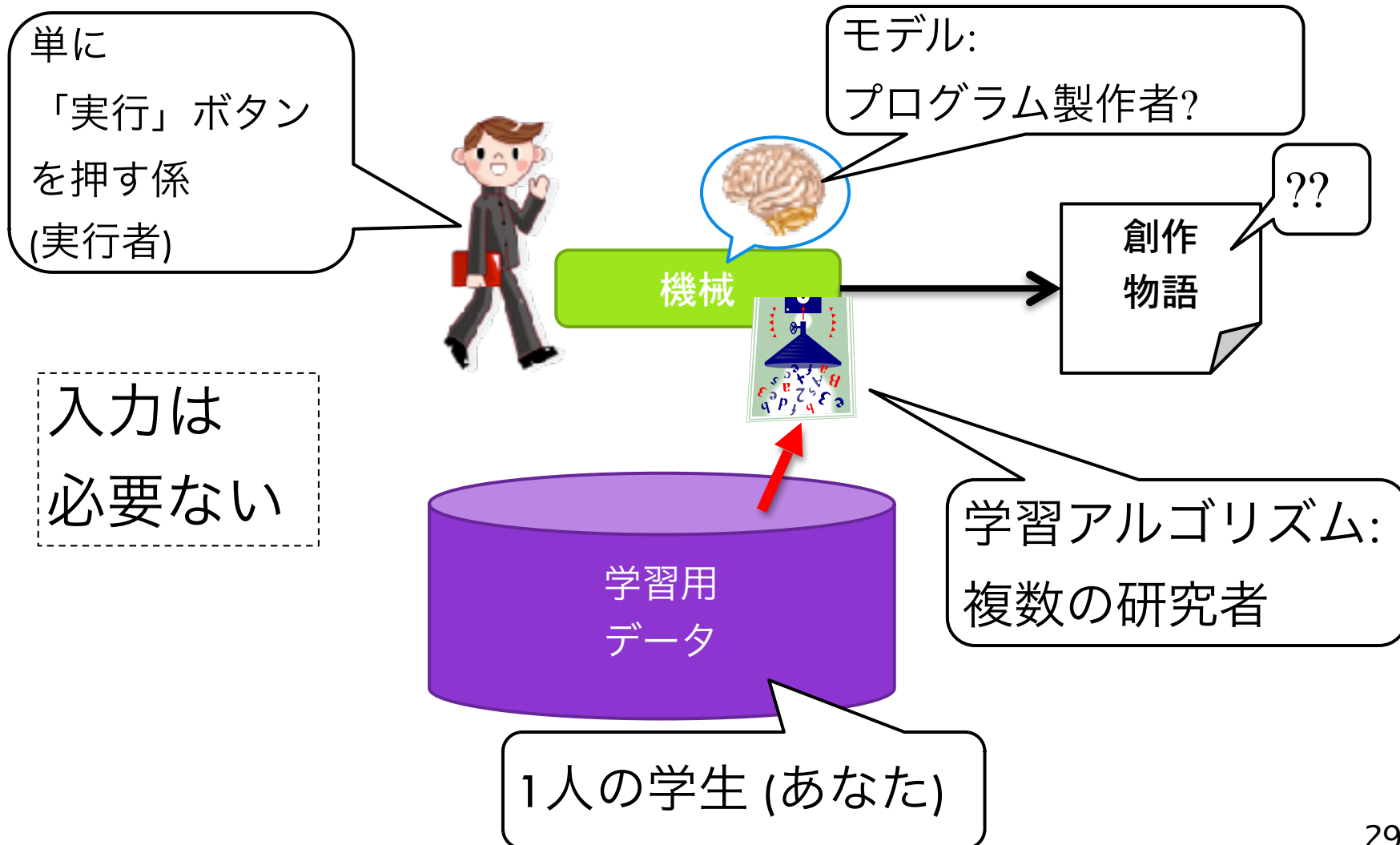
研究倫理

□ 「自律機械と市民をつなぐ責任概念の策定」

研究開発プロジェクト (代表: 東洋大学文学部 松浦和也先生)

- 哲学グループ(哲学者)
- 情報技術グループ(AI研究者)
- 社会制度グループ(法哲学、教育、企業の知財)

物語の自動生成



産物の著作権保有者は？

□ 回答者: 電通大 総合情報学科3年生 45名

保有候補	人数	割合	保有者の完全列挙	人数	割合
あなた	40	89%	あなた	29	65%
プログラム製作者	9	20%	あなた、プログラム製作者	7	16%
実行者	6	13%	あなた、プログラム製作者、 実行者	2	4%
機械	2	4%	あなた、実行者	1	2%
(保留)	1	2%	あなた、機械	1	2%
			実行者	3	7%
			機械	1	2%
			(保留)	1	2%
			合計	45	100%

出版物の責任保有者は？

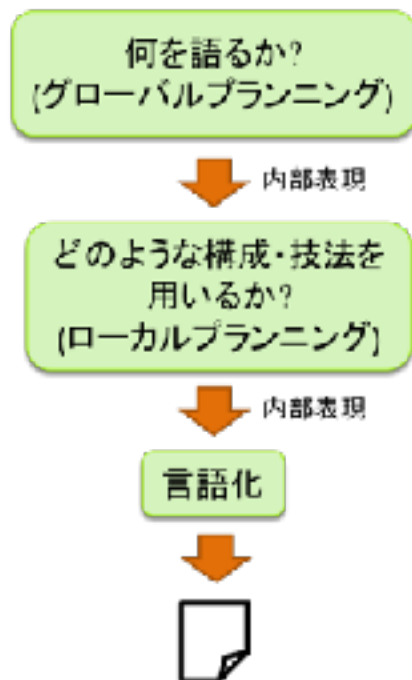
保有候補	人数	割合	保有者の完全列举	人数	割合
あなた	35	78%	あなた	22	49%
出版社	13	29%	あなた、出版社	4	9%
実行者	8	18%	あなた、出版社、プログラム製作者	2	5%
プログラム製作者	5	11%	あなた、実行者	3	7%
いない	2	4%	あなた、プログラム製作者	2	5%
(保留)	1	2%	あなた、実行者、プログラム製作者、	1	2%
機械	0	0%	いない		
			あなた、いない	1	2%
			出版社	5	11%
			出版社、実行者	2	4%
			実行者	2	4%
			(保留)	1	2%
			合計	45	100%

発表内容

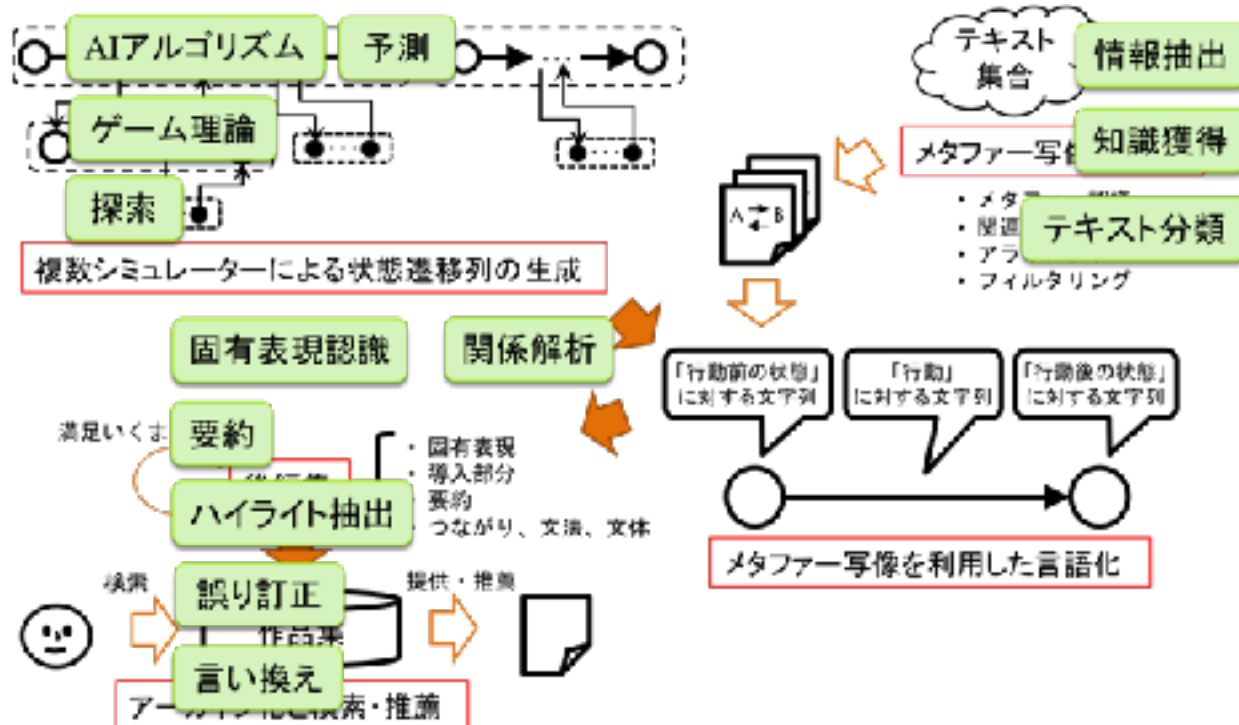
1. 先行研究
2. 提案手法
3. 物語生成の構造: オーケストラ
4. 評価実験
5. 後編集
6. 提案手法の長所
7. まとめ

まとめ

従来手法



提案手法



- ご静聴ありがとうございました

- < 電気通信大学は2018年12月に創立100周年を迎えます >