

概要

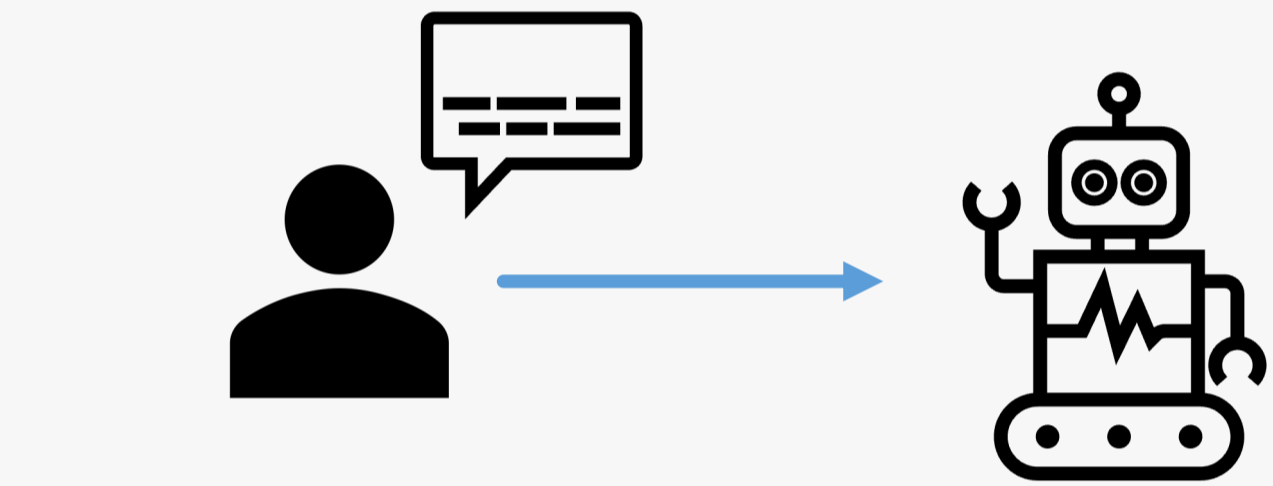
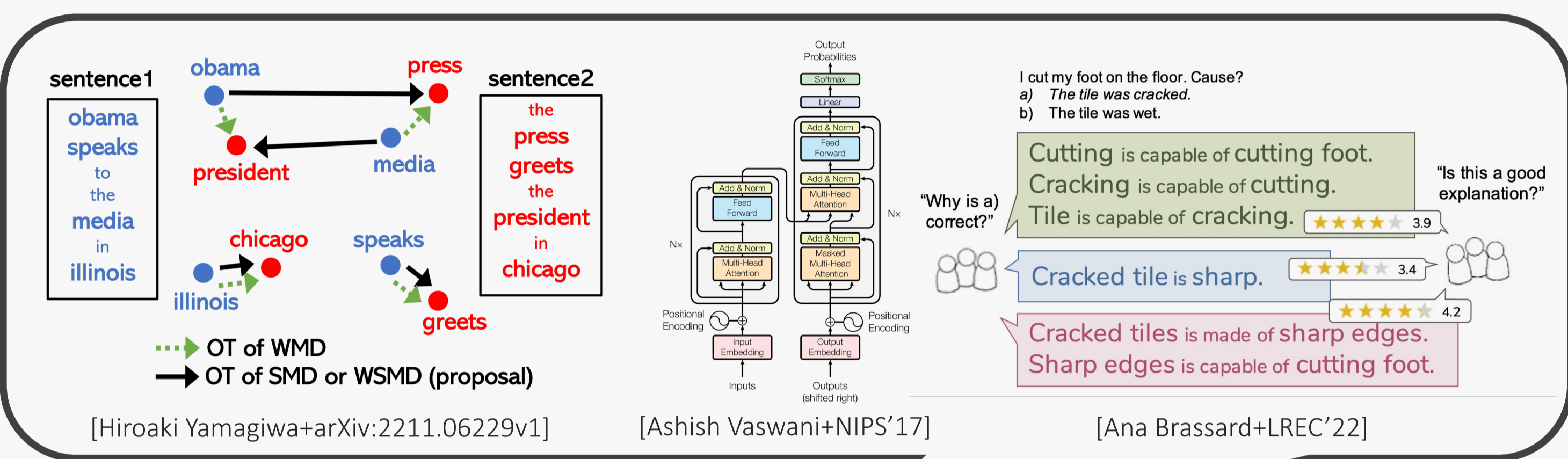
ダイアグラムの自動生成を行うための準備として**評価の軸**を検討し、**既存のモデルの性能**を調査した

背景

ダイアグラム：情報を**視覚的に**表現した図

- 伝えたい情報をわかりやすく表現できる
- もっと簡単に作れるようになると嬉しい

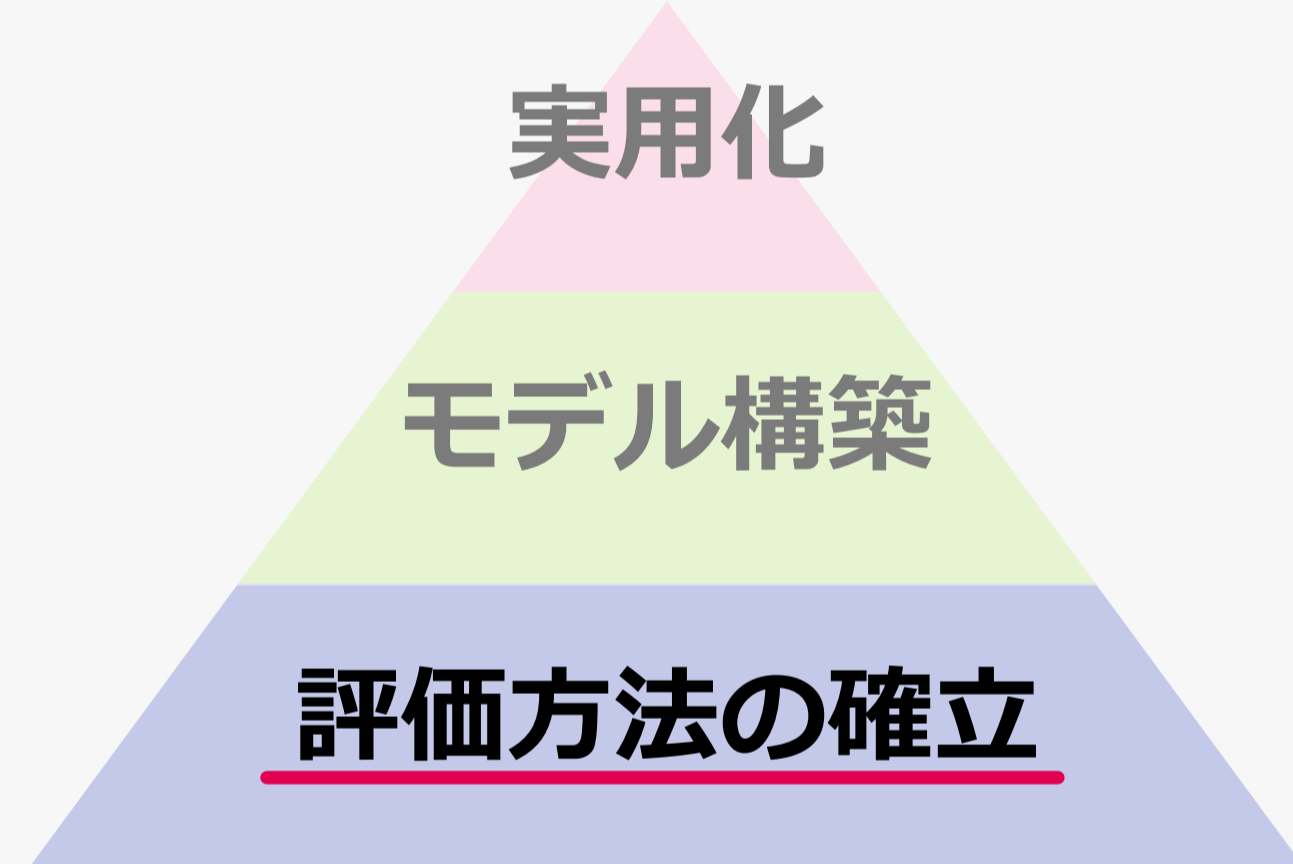
自然言語から**自動生成**したい！



課題と目的

- ✗ 評価の軸が定まっていない
- ✗ 有効な生成手法がない

- ✓ 評価の軸を検討
- ✓ ベースラインの提示

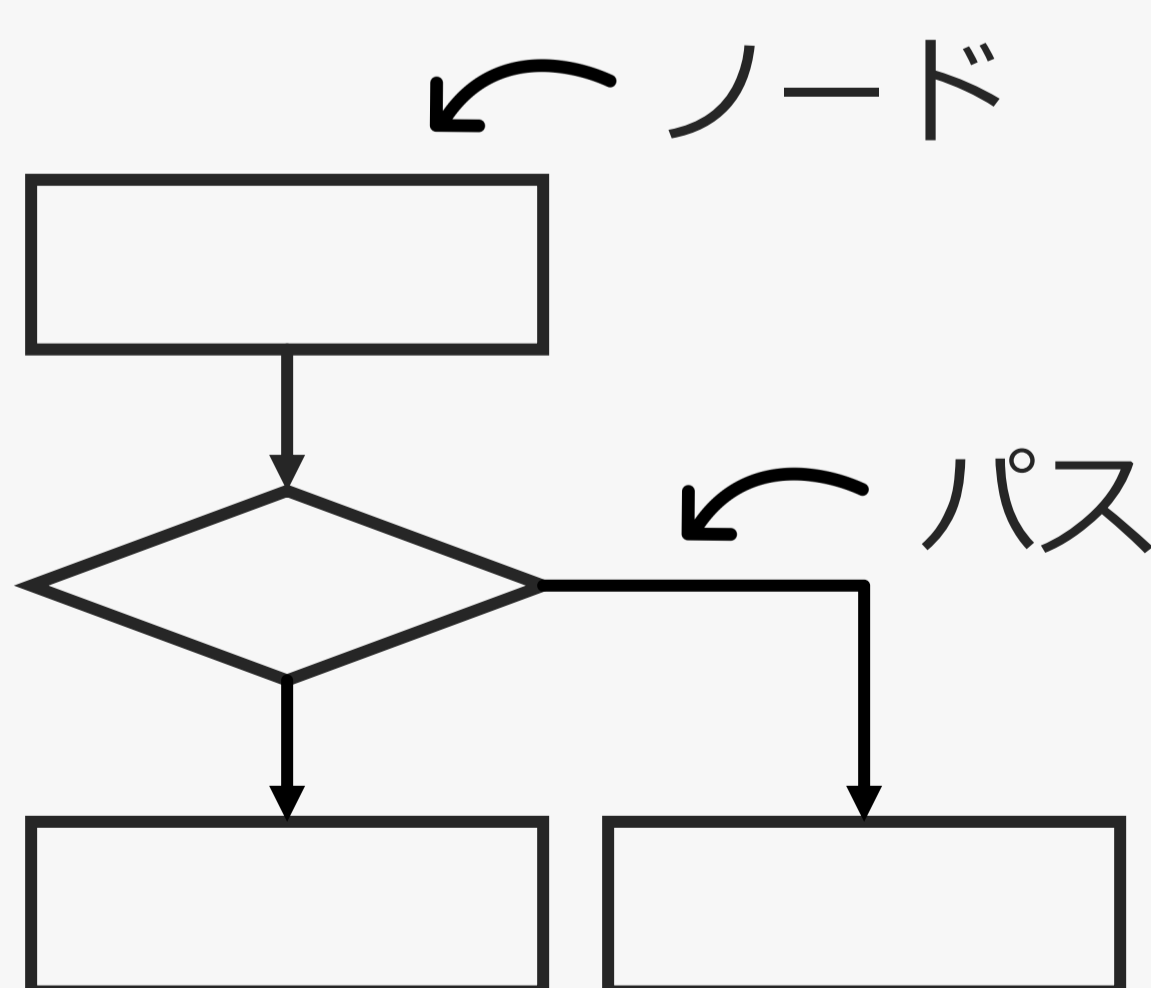


評価軸の検討

意味の評価軸

- ノードの数が必要十分か
- パスの接続が正しいか
- パスのラベルが正しいか

フローチャートに限定



デザインの評価軸

- 適度なスペースがある
- 色使いに意味がある
- 雰囲気が一貫している etc...

まとめ

意味

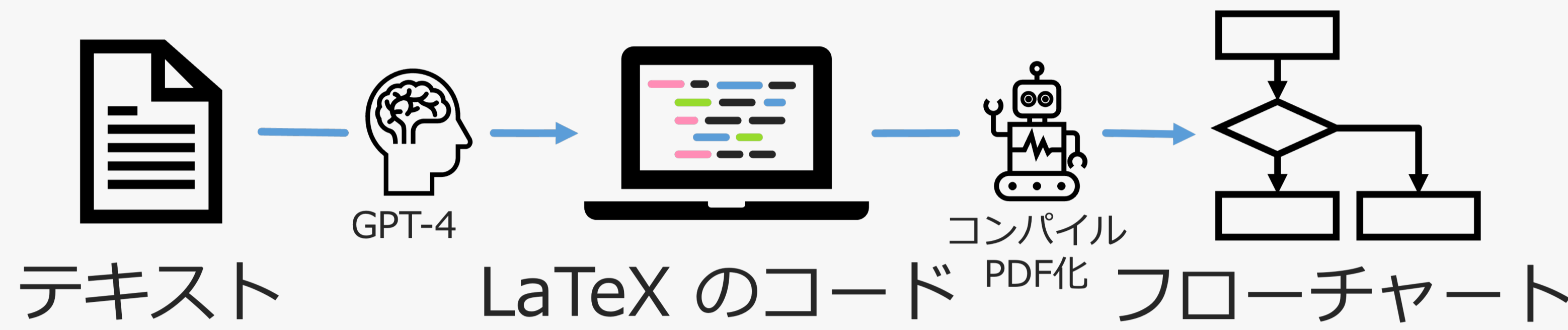
- 高い精度で正しい出力を得た
- プロンプトを正しく理解
 - コード生成能力も十分

デザイン

- 出力結果にばらつきがあった
- ノードサイズの考慮が不十分
 - 色使いがワンパターン

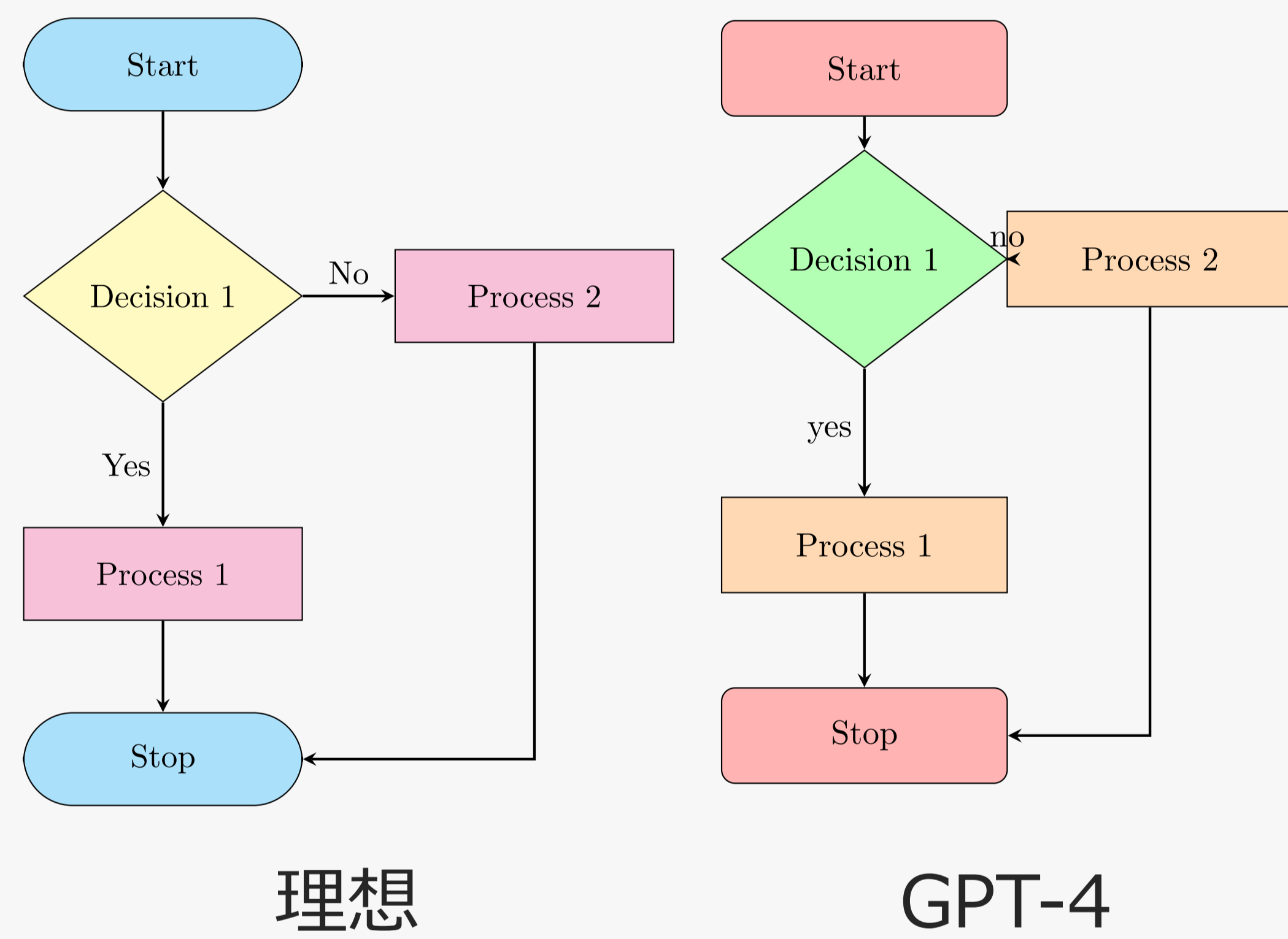
実験概要

- **text-to-code モデル** ← 本研究ではこちらを採用
- text-to-image モデル



実験結果と分析

A flowchart that starts at the start node and goes through process 1 if yes, or through process 2 if no, to reach the end node

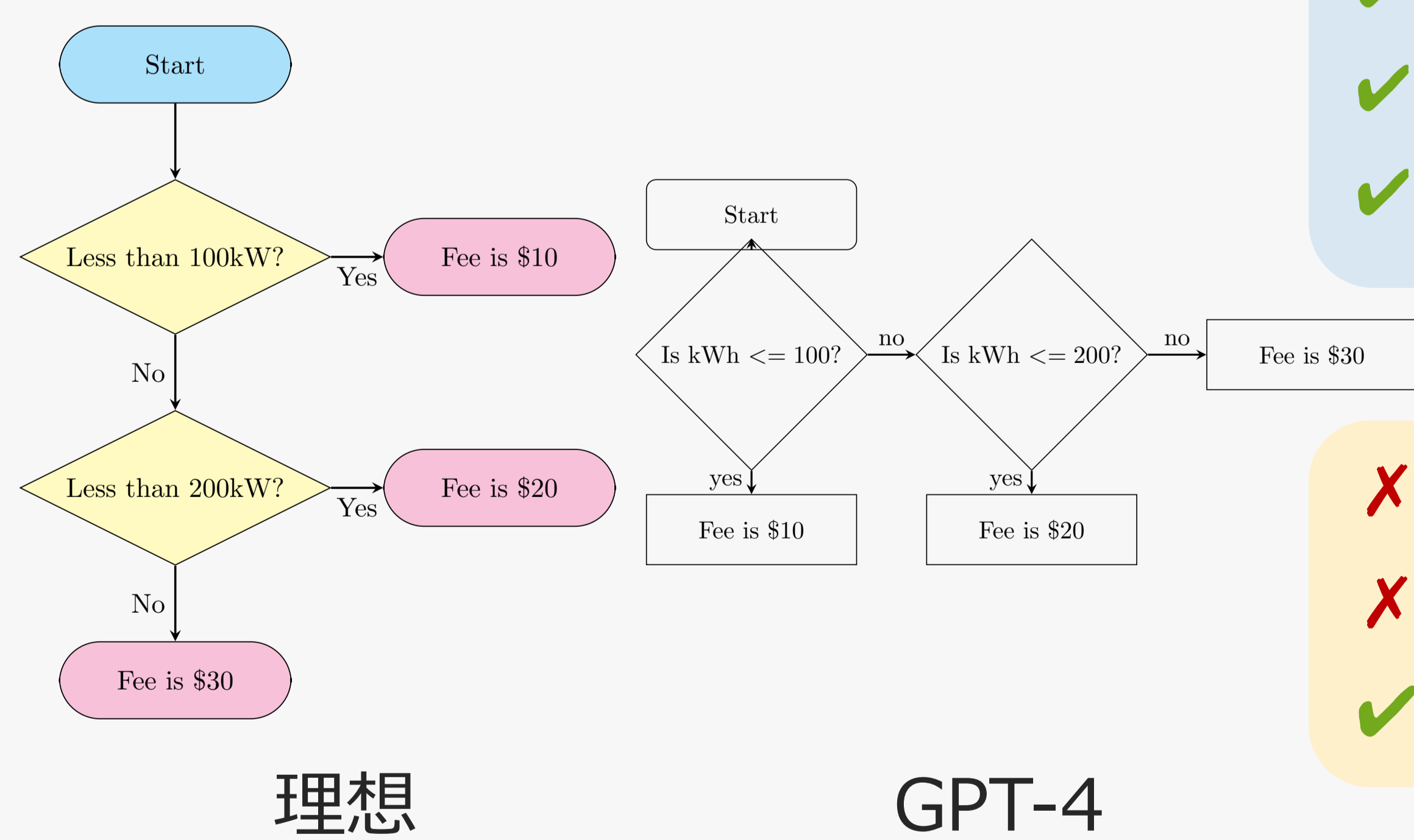


- ✓ ノードの数
- ✗ パスの接続*
- ✓ パスのラベル

* コードでは正しいが画像では間違っている

- ✗ スペース
- ✓ 色使い
- ✓ 一貫性

A flowchart for checking whether the fee will be \$10, \$20, or \$30 with the border as 100 kW or 200 kW



- ✓ ノードの数
- ✓ パスの接続
- ✓ パスのラベル

- ✗ スペース
- ✗ 色使い
- ✓ 一貫性

今後の展望

- 具体的な評価指標の作成
→ **手法によらない**指標を目指す
- 既存のモデルの限界を調査
→ できる / できない を細かく