

アルゴリズムの平均感度解析

吉田 悠一

国立情報学研究所 情報学プリンシプル研究系 教授

どんな研究？

- アルゴリズムの安定性を保証したい
- 安定性を数学的に定式化した**平均感度**の提案
- アルゴリズム解析を平均感度の観点からやり直す！

何がわかる？

- アルゴリズムの出力精度と平均感度
の間のトレードオフ
- 平均感度を低く抑えることで、社会
において安心して使えるアルゴリズム
を提供

背景・目的

アルゴリズムを安心して使うには…

アルゴリズムを用いた意思決定や知識発見は既に社会の一部（例：人員割当・商品の推薦・実験結果の解析）

もしアルゴリズムの出力が不安定だと…

- 人員の割当を頻繁に変更するのは**非効率**
- 一貫性がない推薦は**信用を失う**
- 科学的な主張を行うためには**再現性が必要**

安定性の定式化

これまで提案されてきた安定性の指標には不満が多い

- 綺麗に定式化されておらず比較・改善が難しい
- 精度だけで出力については安定性を担保しない
- 要求が強すぎて使いにくい

理論的にも応用的にも使いやすい指標はないか？

研究内容（方法・結果・結論）

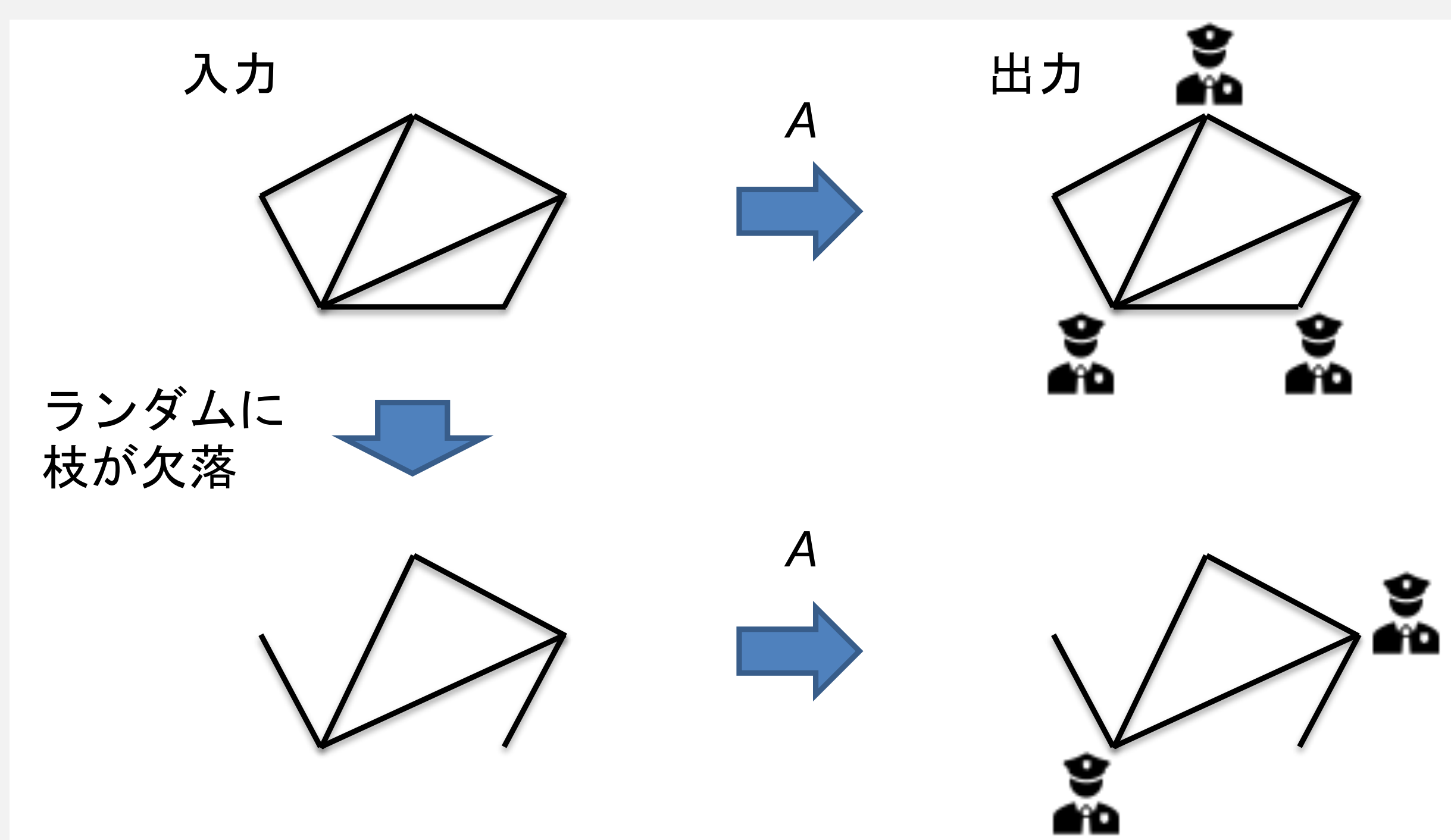
平均感度

入力が少し変化したとき出力解がどれだけ変化するかを**平均感度**として定式化（下図）

アルゴリズムの解析を平均感度の観点からやり直す！

- これまでのアルゴリズムの平均感度は高い？低い？
- 今の速度・精度を保ちつつ平均感度を低くできる？

頂点被覆問題におけるアルゴリズムの平均感度
(頂点被覆：グラフの全ての枝を監視員で監視する)



元の入力と欠落後の入力とで配置が三人変更
(二人削減 + 一人追加)

平均感度 := 変更数の枝の欠落に関する平均

平均感度に理論保証のあるアルゴリズム

多くの問題で近似性能をほとんど犠牲にせずに平均感度を小さくできる！

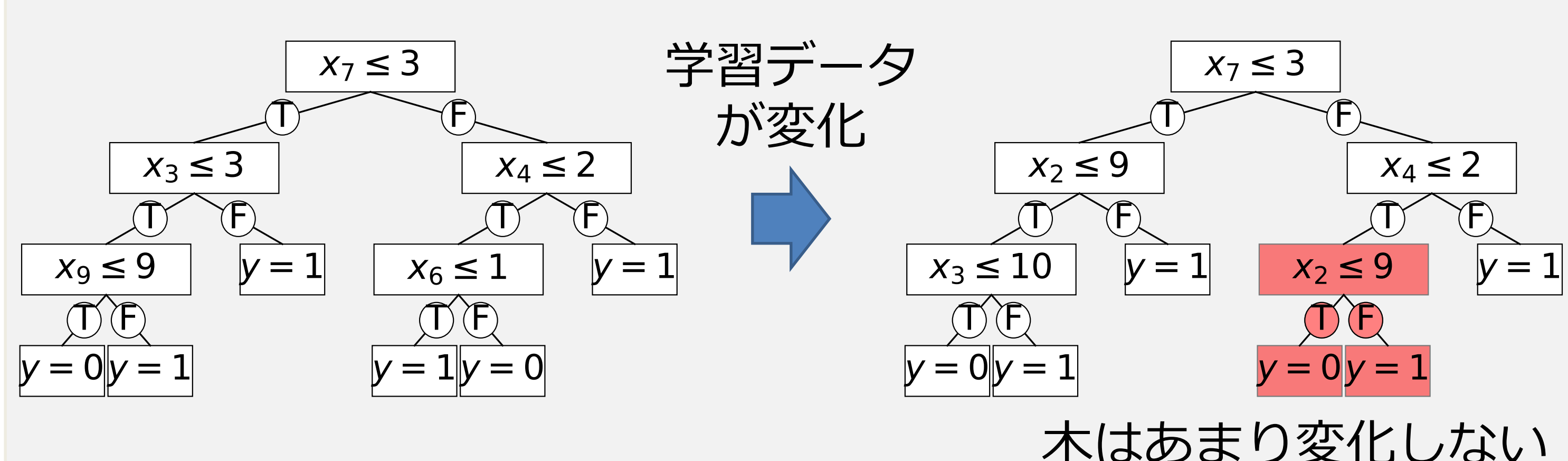
問題	近似度	平均感度
最小全域木	1	$O(1)$
最大マッチング	$1/2$	$O(1)$
	$1 - \epsilon$	$\exp \exp \exp O(\epsilon^{-1})$
ナップサック問題	$1 + \epsilon$	$O(\log^3 n / \epsilon)$

決定木学習

学習データから決定木（下図）を構築
決定木は解釈性の高さが利点 ⇒ 安定に構築したい

高精度と低平均感度を両立するアルゴリズムを提案

決定木：点 x からラベル y を予測



木はあまり変化しない！