

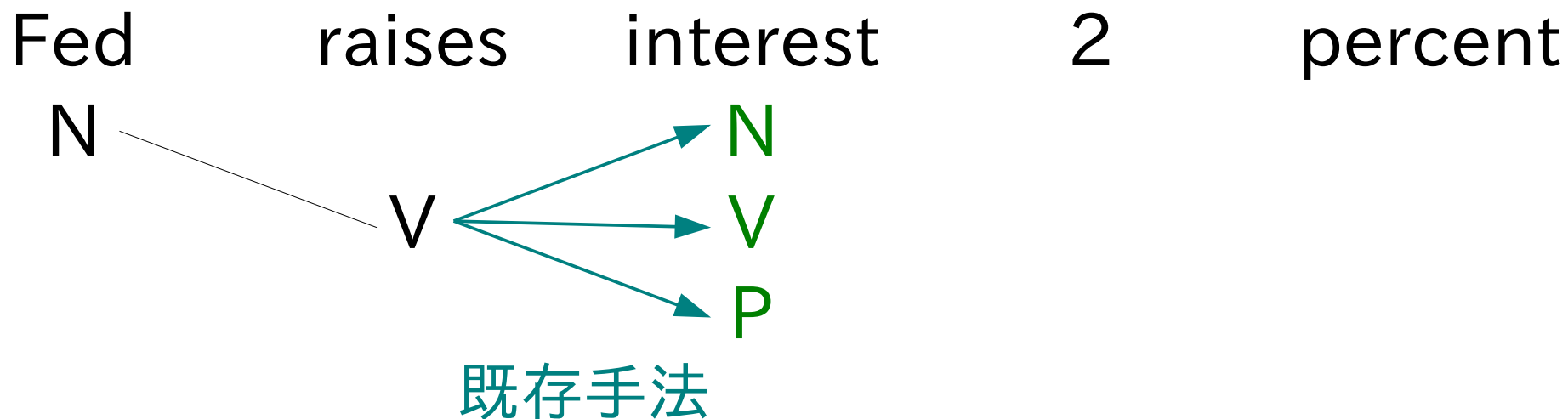
先読みを用いた系列ラベリングにおける 最易優先探索の適用

佐野峻平[†] 三輪誠[‡] 鶴岡慶雅[†] 近山隆[†]

[†]東京大学 [‡]マンチェスター大学

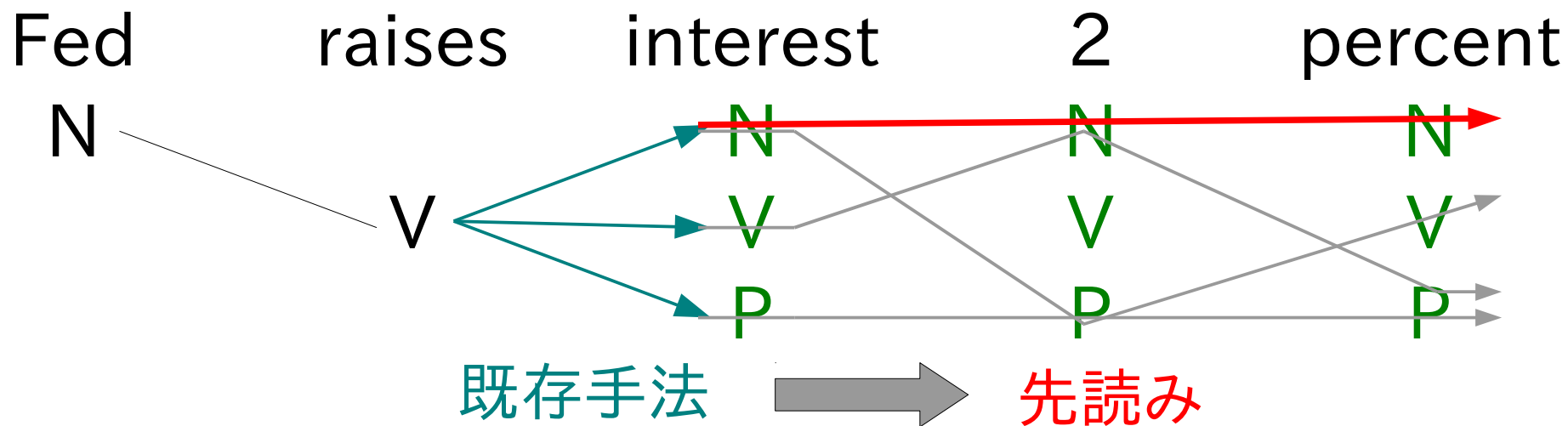
背景

- ▶ 系列ラベリング
POS-tagging, チャンキング, NER
- ▶ 先読み [Tsuruoka et al. 2011]
ラベル付けされていない語のラベル情報も利用



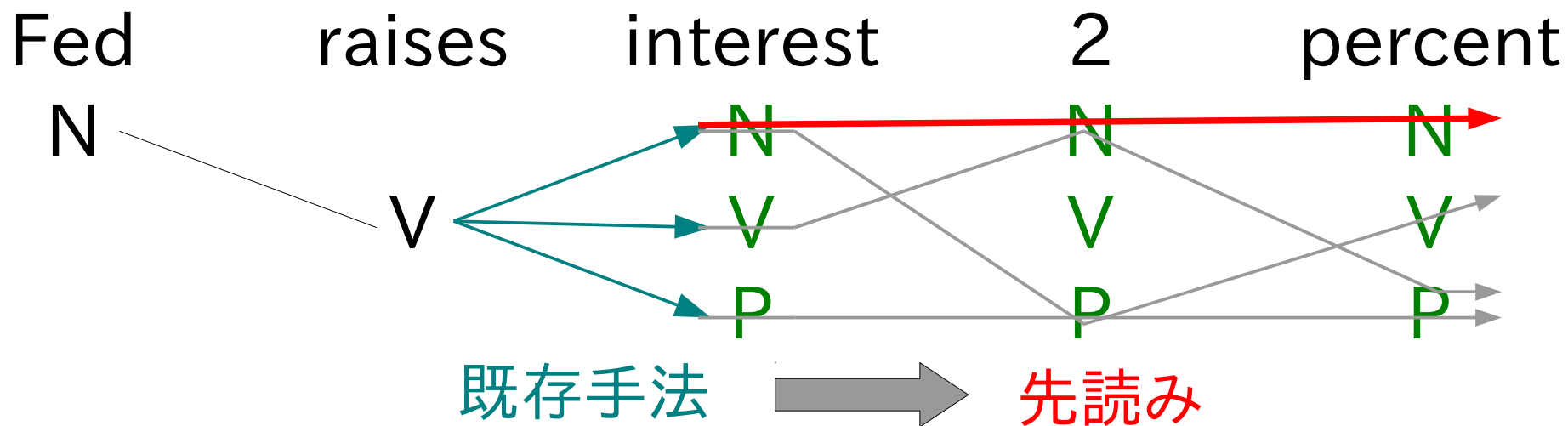
背景

- ▶ 系列ラベリング
POS-tagging, チャンキング, NER
- ▶ 先読み [Tsuruoka et al. 2011]
ラベル付けされていない語のラベル情報も利用



背景

- ▶ 系列ラベリング
POS-tagging, チャンキング, NER
- ▶ 先読み [Tsuruoka et al. 2011]
ラベル付けされていない語のラベル情報も利用

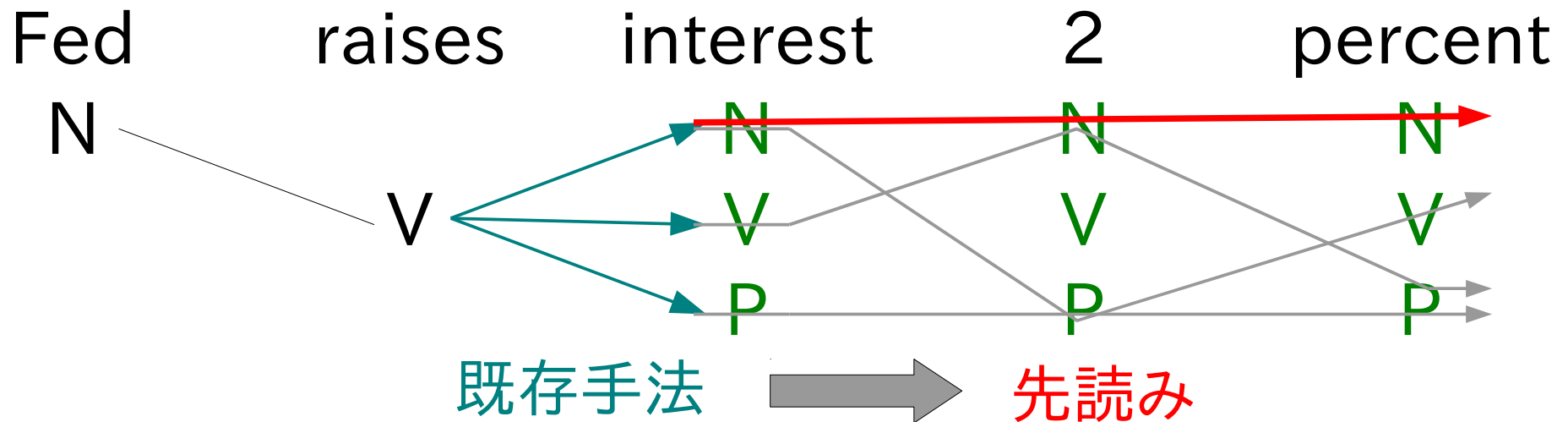


ラベル付けは「前から後ろ」か「後ろから前」

ラベル付け
失敗

背景

- ▶ 系列ラベリング
POS-tagging, チャンキング, NER
- ▶ 先読み [Tsuruoka et al. 2011]
ラベル付けされていない語のラベル情報も利用



ラベル付けは「前から後ろ」か「後ろから前」

ラベル付け
失敗



ラベルの誤りが連鎖的に
後のラベル付けに影響

提案手法

▶ 先読み + 最易優先

ラベル付け確率の最も高い
(語, ラベル) の組

▶ ラベル付け確率の計算に、先読みにより
少し変更を加えたものを利用

▶ **最易な所から順に**ラベル付け

確実な所から
ラベル付け

ラベル付け確率の
計算をより正確に

提案手法

▶ 先読み + 最易優先

ラベル付け確率の最も高い
(語, ラベル) の組

▶ ラベル付け確率の計算に、先読み
に少し変更を加えたものを利用

▶ **最易な所から順に**ラベル付け

確実な所から
ラベル付け

ラベル付け確率の
計算をより正確に

**失敗しにくく
連鎖的な影響も起きにくい
ラベル付け**