



情報・システム研究機構 若手研究者集会

最先端構文解析とその周辺

統計数理研究所

2012/12/19 (水)

主催：持橋大地 (統計数理研究所)
宮尾祐介 (国立情報学研究所)



Twitter / tarowatanabe: @neubig 構文解析は、自然言語処理の王道です。 ...

https://twitter.com/tarowatanabe Twitter, Inc. Google

Google ISM Realtime Tools

検索

 **Taro Watanabe**
@tarowatanabe

フォローする

@neubig 構文解析は、自然言語処理の王道です。構文解析を理解でけへんかったら、研究が進みません。

返信 リツイート お気に入りに登録

1 件のリツイート 3 FAVORITES

2012年9月13日 - 4:45 · このツイートをサイトに埋め込む




Taro Watanabe (tarowatanabe)さんはTwitterを使っています

https://twitter.com/tarowatanabe Twitter, Inc. Google

Google ISM Realtime Tools

検索

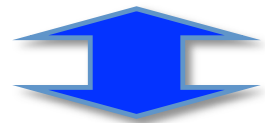
 **Graham Neubig** @neubig 9月13日
よく考えると、自然言語処理で一番有名な先生は結構な割合で構文解析の研究から始まったなあ。Manning, Eisner, Klein, Resnikなど。日本ならMatsumoto, Kurohashiもそう (Tsujiは確か最初はMT?)。
開く

 **Taro Watanabe** @tarowatanabe 9月13日
[@neubig](#) Knightは構文解析の中のunification。
開く

 **Graham Neubig** @neubig 9月13日
[@tarowatanabe](#) 調べたらPereiraとJurafskyもそうみたいですな。
開く

3つの構造

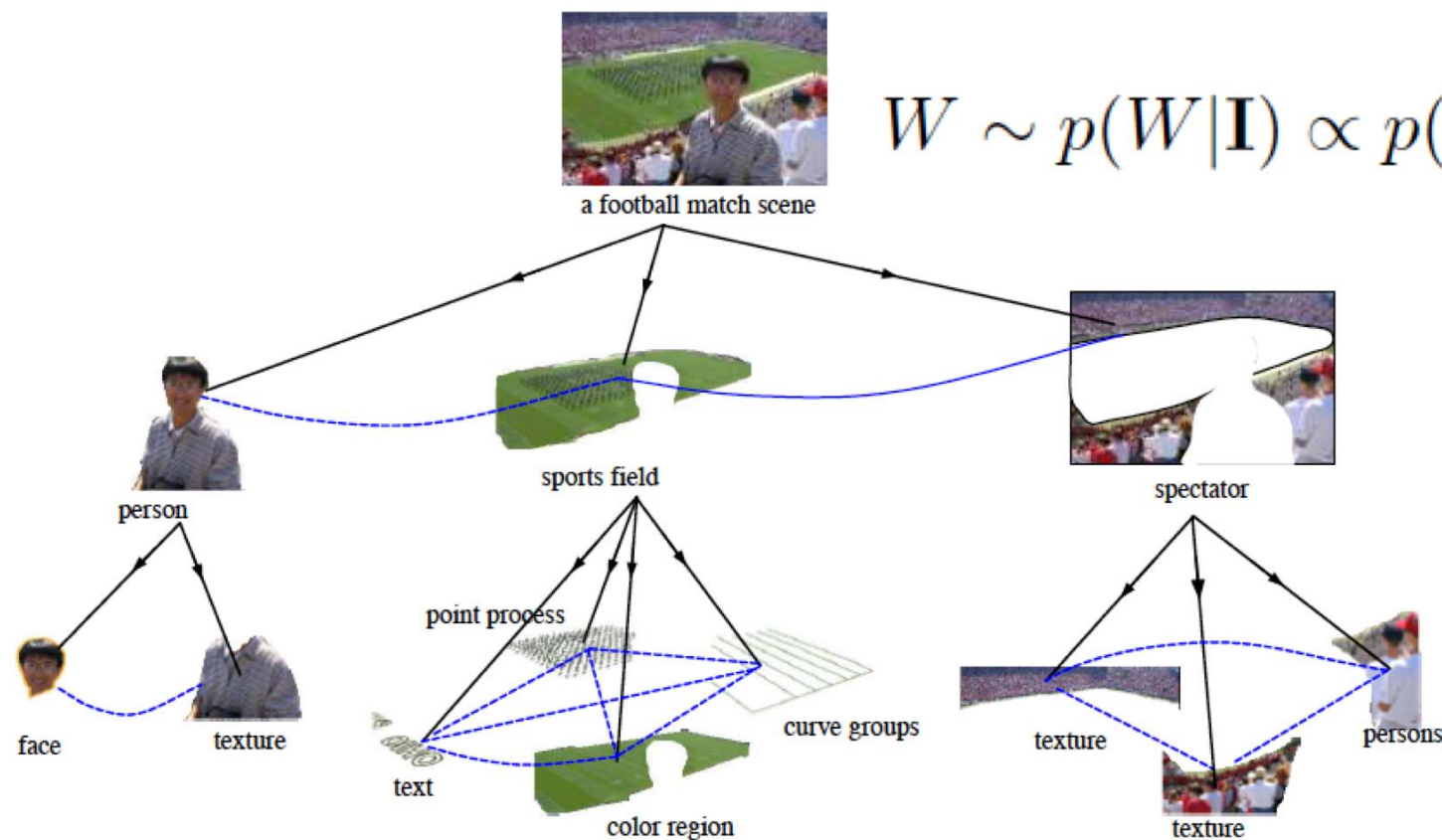
- 木構造 (Trees) ← 人工知能
 - PCFG
- 系列モデル (Sequences) ← 情報理論
 - n-gram
- 分布モデル (Distributional) ← 統計学
 - PLSA, LDA



- Branching process
- Markov process
- Poisson process

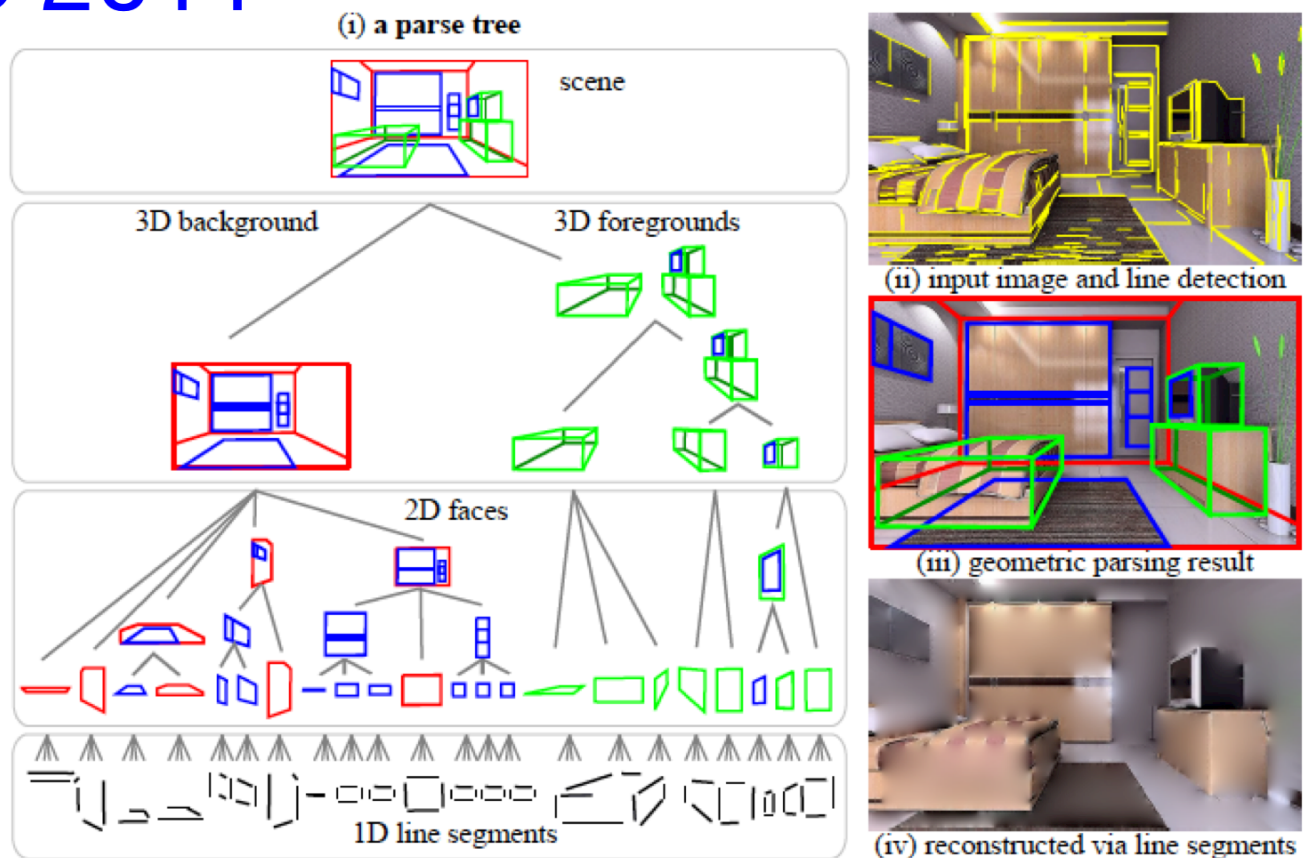
他の可能性も？

Image Parsing (Zhu+, 2005)



- “Image Parsing: Unifying Segmentation, Detection, and Recognition”, ICJV, 2005

NIPS 2011

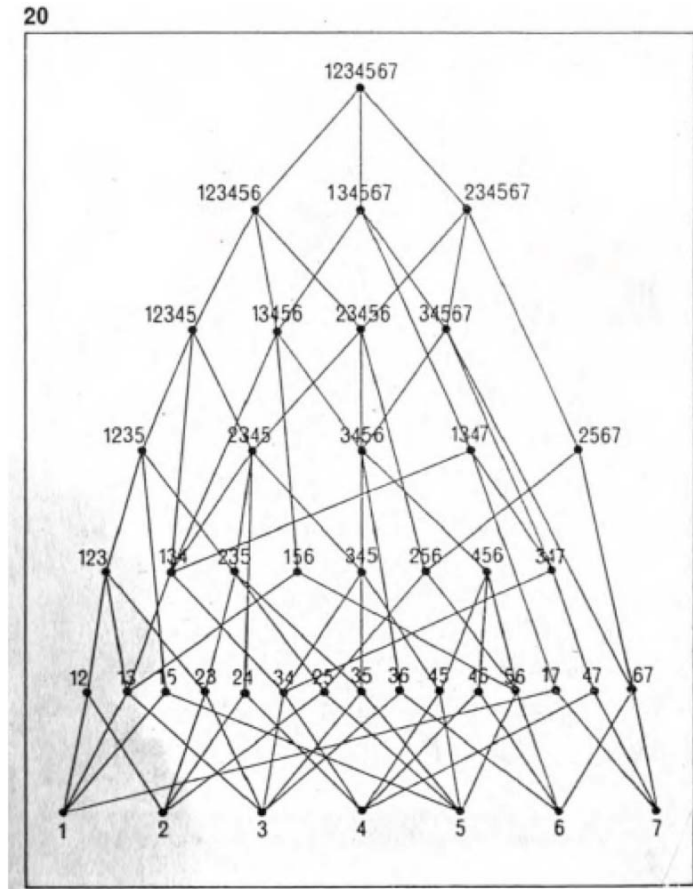


$$E(pt|I) = \sum_{v \in V^{OR}} E^{OR}(A_T(Ch_v)) + \sum_{v \in V^{AND}} E^{AND}(A_G(Ch_v)) + \sum_{\Lambda_v \in \Lambda_I, v \in V^T} E^T(I(\Lambda_v))$$

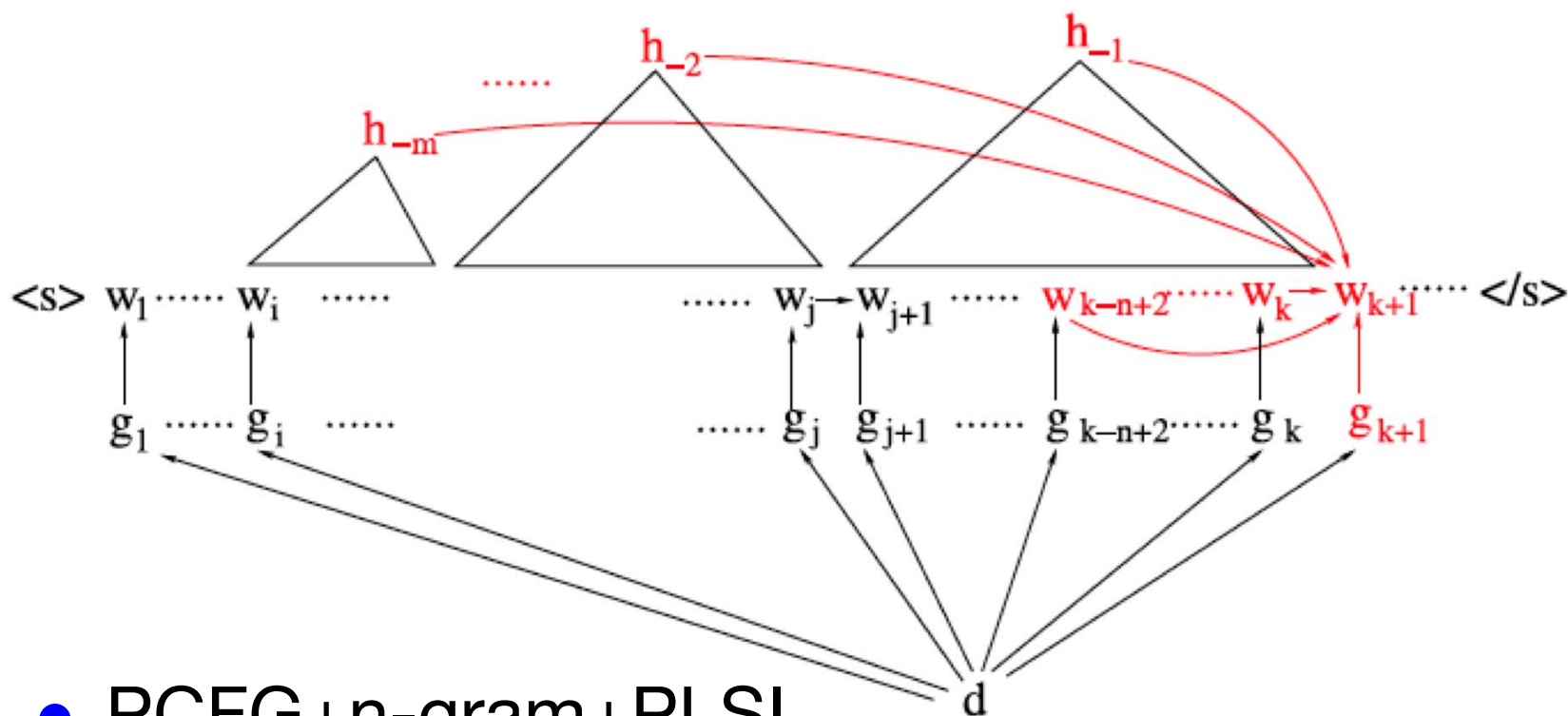
- “Image Parsing via stochastic scene grammar”, Zhao&Zhu

「都市はツリーではない」

- Christopher Alexander の有名な言葉 (1965)
- “言語はツリーではない”？
– もちろんツリーだが..

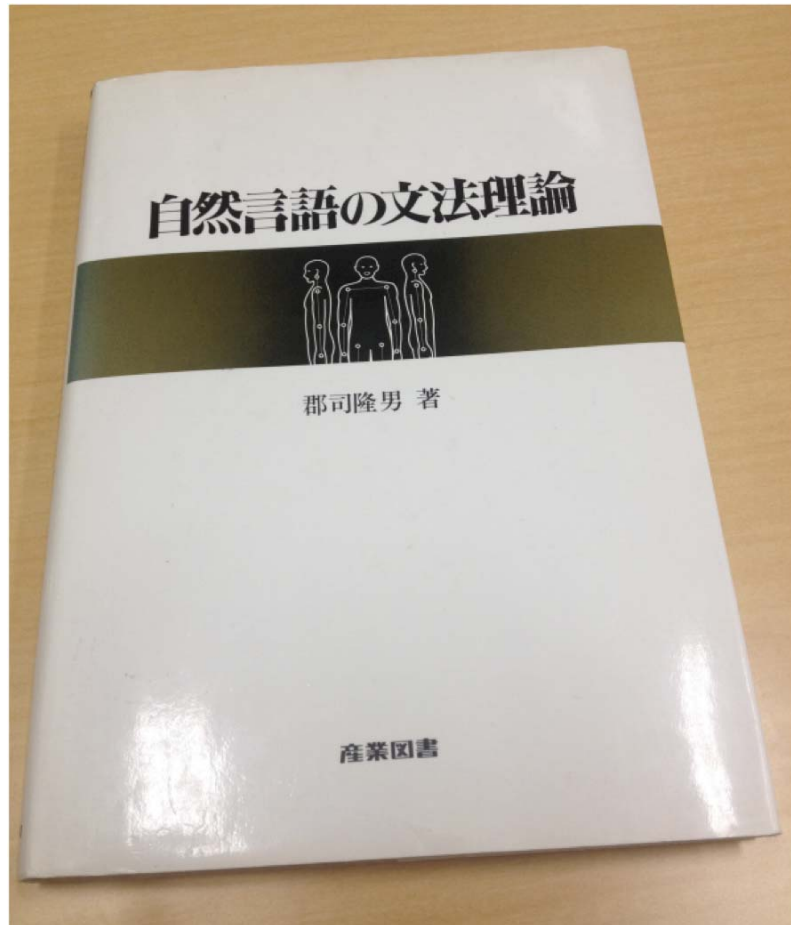


Latest Composite Language model



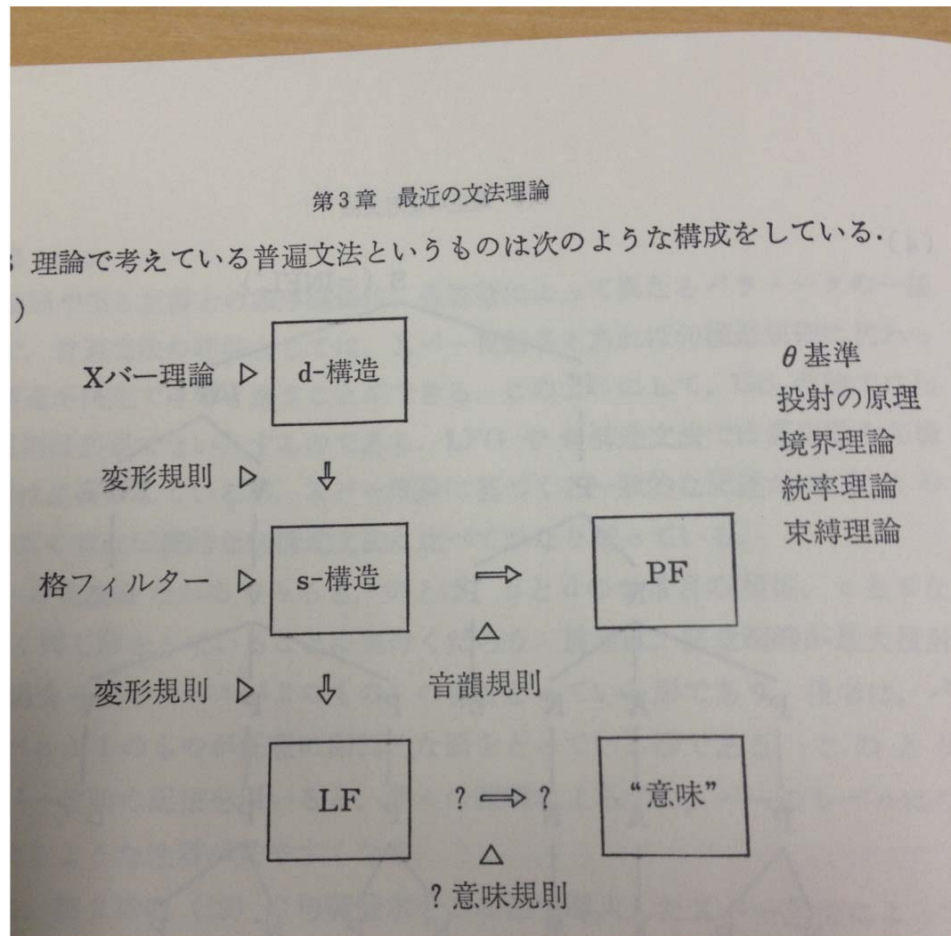
- PCFG+n-gram+PLSI
 - Shaojun Wang+, “A scalable distributed syntactic, semantic, and lexical language model”, *Computational Linguistics*, 2012

言語の科学



- 学部生の頃、最初に読んだ本
- 言語が科学になる！
(実際にずっと読んでみると..)

文法理論の確率モデル化



- 「深層構造」からの「生成」
- 「変形操作」
- 「音韻規則」

単なる理論から、
具体的な確率モデル
へ



今日の講演

- 宮尾祐介氏 (国立情報学研究所)
「自然言語処理における構文解析と言語理論の関係」
- 進藤裕之氏 (NTTコミュニケーション科学基礎研究所)
「統計的手法による文法モデリングと言語解析」
(休憩)
- 吉仲亮氏 (京都大学 情報学研究科)
「文脈自由木文法の分布学習」
- 能地宏氏 (東京大学 情報理工学研究所)
「教師なし構文解析の進展」



諸注意

- 休憩時には、外にコーヒーとお菓子が準備してありますので、ご自由にご利用下さい
- 大会議室の中(のみ)は、飲食禁止となっております
 - － 部屋の外や1Fロビーの椅子をご利用下さい