



自然言語処理2014

No.12

東京工科大学
コンピュータサイエンス学部
亀田弘之

今日は文法獲得の話しをします。

(確認)構文解析プログラムの作成手順

1. 開発環境の整備

- ① 各種ソフトウェアのインストール
- ② 文字コードの決定 など

2. プログラム作成手順

- ① 言語データの収集
- ② 言語分析 <= 対象言語の知識が必要
- ③ 形式文法の設定 <= 形式言語の知識が必要
- ④ Prolog形式への書き換え <= Prologの知識が必要
- ⑤ NLPプログラムの実行
- ⑥ 構文木を出力するプログラム (draw_term/1)
- ⑦ 動作することの確認
- ⑧ 妥当性の検証 など
- ⑨ 公開へ向けてのドキュメント等の整備

質問

- (1) 言語データをどうやって集める？
- (2) 言語データをどうやって分析する？
- (3) 形式文法への書き換え, 良い手はないの？
- (4) Prolog形式への書き換え自動化できないの？

自分なりの答えを考えてください.

疑問：大量のテキストデータから 言語獲得できないのだろうか？

- できない？
 - なぜできない？
- できる？
 - なぜ(why)?
 - どうやったらできる(how)?

(呟き)人間の子どもように学習(言語学習)できないのかなあ？

問題の分析・整理

- 入力: テキストデータ
- 出力: 言語(体系)に関する知識

幾つかの事例から, 本質的な知識(法則・規則)を発見する!

問題の分析・整理

- 入力: テキストデータ ← 大量のデータが必要
- 出力: 言語(体系)に関する知識
 - ← 記述言語は？
 - ← 抽出・発見方法は？

幾つかの事例から、本質的な知識(法則・規則)を発見する！

問題の分析・整理

- 入力: テキストデータ ← 大量のデータを使おう!
- 出力: 言語(体系)に関する知識
 - ← 記述言語
(Prolog言語 = 1階述語論理)
 - ← 抽出・発見方法(帰納的推論)

幾つかの事例から, 本質的な知識(法則・規則)を発見する!

問題の分析・整理

- 入力: テキストデータ ← 大量のデータを使おう!
- 出力: 言語(体系)に関する知識
 - ← 記述言語
(Prolog言語 = 1階述語論理)
 - ← 抽出・発見方法(帰納的推論)

幾つかの事例から、本質的な知識(法則・規則)を発見する!

問題の分析・整理

- 入力: テキストデータ ← **大量のデータ**を使おう!
- 出力: 言語(体系)に関する知識
 - ← 記述言語
(Prolog言語 = **1階述語論理**)
 - ← 抽出・発見方法 (**帰納的推論**)

幾つかの事例から、本質的な知識(法則・規則)を発見する!

帰納論理プログラミング

(Inductive Logic Programming; **ILP**)

帰納論理プログラミング

(その前に)

いろいろな論理学

- 古典的論理学
 - 命題論理学 (Propositional Logic)
 - 述語論理学 (Predicate Logic)
- 現代論理学
 - 様相論理学 (Modal Logic)
 - 時相論理学 (Temporal Logic)
 - 線形論理学 (Linear Logic)
 - ファジー論理学 (Fuzzy Logic)
 - Paraconsistent Logic
 - その他

論理回路設計, 人工知能, 機械学習,
ソフトウェア検証などで活躍している.

いろいろな推論

- 演繹推論
- 帰納推論
- 発想推論

前提1: 規則「A ならば B」が知られている.

前提2: いま「A」である.

結論: いま「B」である.

- A: すべての人間は死ぬ. 大前提
- B: ソクラテスは人間である. 小前提
- C: ソクラテスは死ぬ. 帰結

- AとBからCを導く → 演繹推論
- AとCからBを導く → 発想推論
- 幾つかのBとCの対からAを導く → 帰納推論

帰納論理プログラミング (ILP)

- 条件設定
 - 背景知識
 - 事例
- 処理
 - 事例を適切に説明することのできる知識を
発見的に獲得する.

このようなことをやってのける方法の1つがILP

ILPシステムの例

- Golem
- Progol
- Alephなど

(注)ソースコードは公開されている.

文法獲得の事例紹介

- 統語規則獲得
 - Prolog, Apleph
- 未知語獲得
 - Prolog (Version 5以降, 発想推論)

(注)配布資料も参照のこと

以上で、今年はおしまいです。

- 次回は1月5日(月)です。
- 良いお年をお迎えください。
A Happy New Year to you!

