

生命科学知識を得るために入力される自然言語クエリの構文解析

原 忠義¹

建石 由佳²

Jin-Dong Kim³

宮尾 祐介¹

¹ 国立情報学研究所

² 工学院大学

³ ライフサイエンス統合データベースセンター

自然言語クエリ文の構文解析

- 多くの生命医学の知識・データベースに「容易」かつ「柔軟」に問い合わせたい
 - 複雑なクエリ言語を用いてではなく、自然言語で
 - “*Find articles about function of FancD2*”

自然言語
クエリ文

構文解析

解析木

変換

クエリ言語



自然言語クエリ文の構文解析

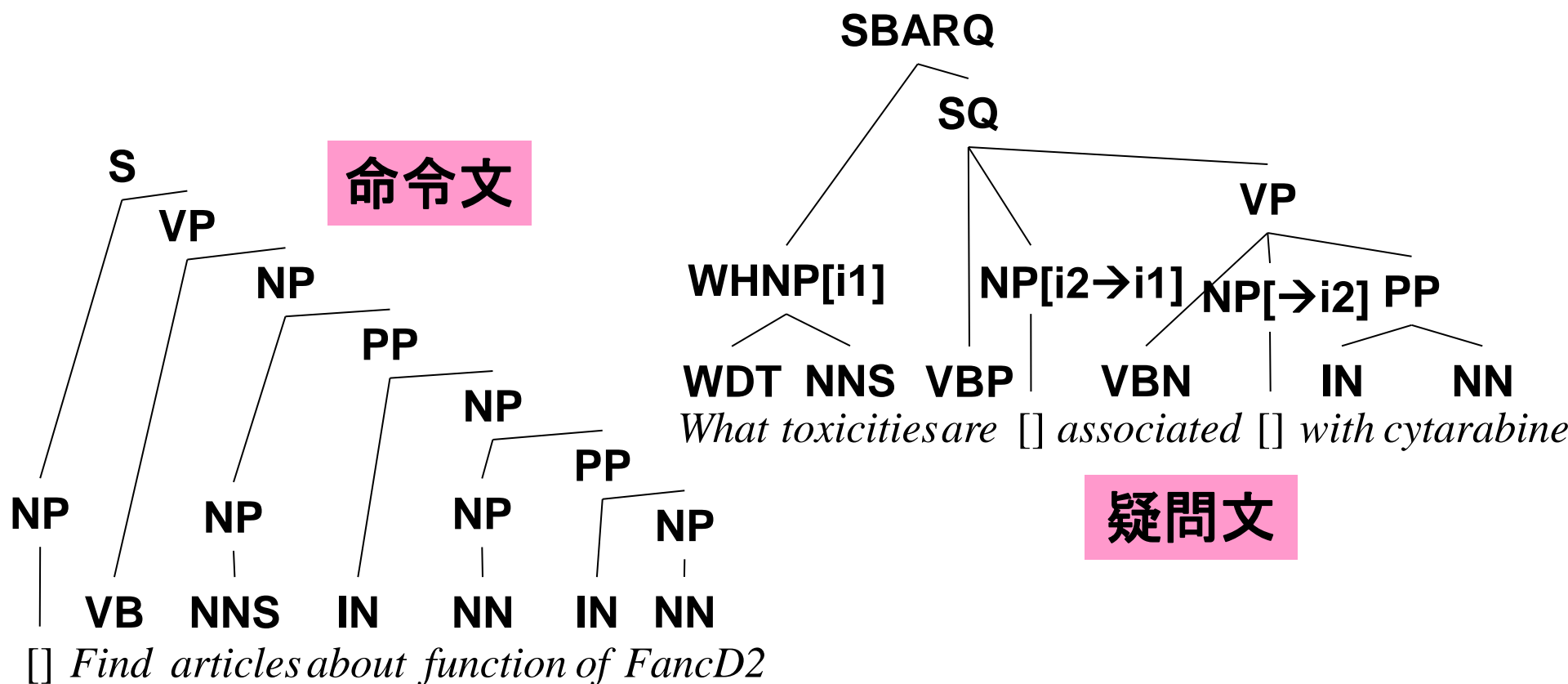
- 多くの生命医学の知識・データベースに「容易」かつ「柔軟」に問い合わせたい
 - 複雑なクエリ言語を用いてではなく、自然言語で
 - “*Find articles about function of FancD2*”



最先端の構文解析器が正しくクエリ文を解析できるのかどうか調査する

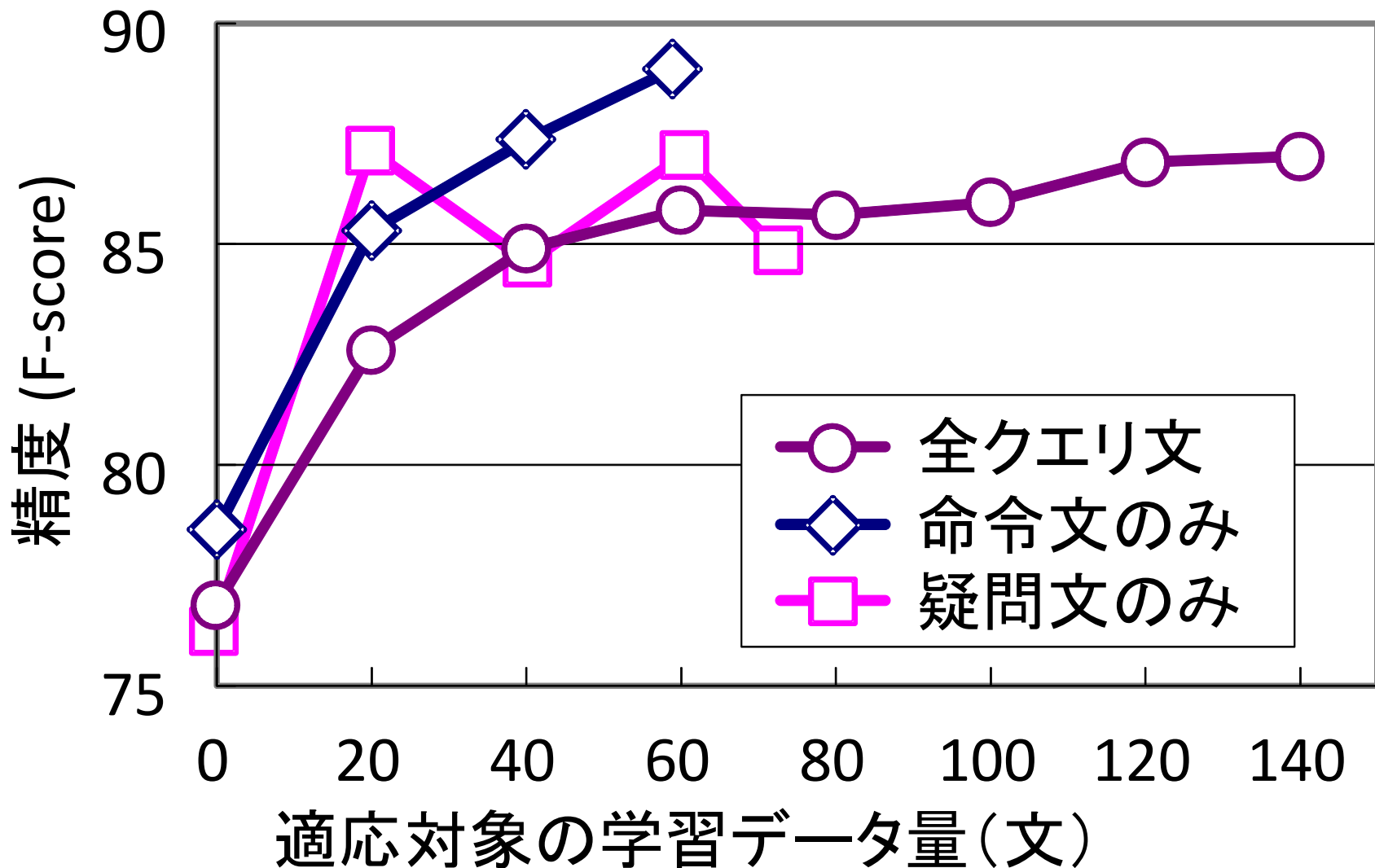
Step1: クエリ文の木構造アノテーション

- Genomics Track 2004-07 から 196 のクエリ文
– 命令文 82 文, 疑問文 98 文, その他 16 文



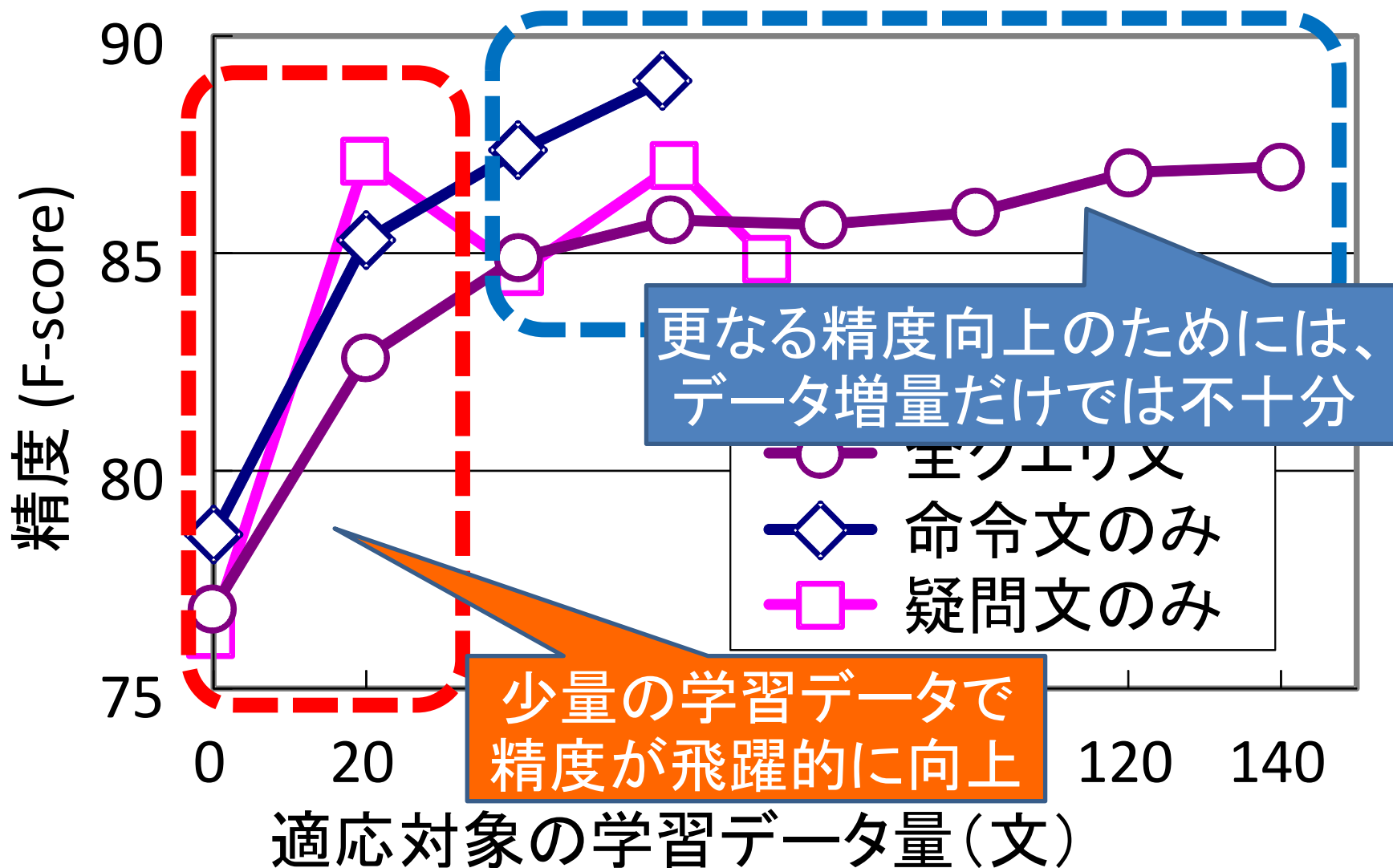
Step2: 構文解析精度の調査

構文解析精度



Step2: 構文解析精度の調査

構文解析精度



ポスターでは分析の詳細をお見せいたします