

## 平成27年度 情報工学コース卒業研究報告要旨

石川 研究室	氏 名	川端 風貴
卒業研究題目	遷移型依存構造解析を用いた日本語文の語順整序	

日本語は同じ意味を持つ文でも、様々な語順で書くことができる言語であり、比較的自由的な語順を持つ。しかし、完全に任意というわけではなく、同じ意味を持つ文であっても、語順によっては読み難くも読み易くもなる。例えば、次の2つの例文について考える。

1. 彼がついに長年人々が取り組んできた問題を解決した。

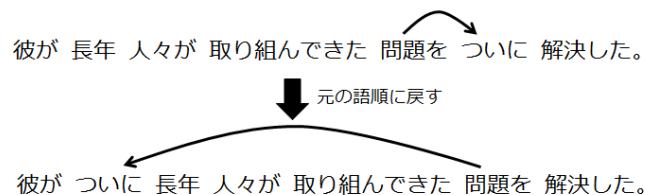
2. 長年人々が取り組んできた問題をついに彼が解決した。

例文1はそのままでは読み難い。しかし、例文2のような語順へと並べ替えることで、読み易い文にすることができる。

読み難い語順を持った日本語文に対して、文節の並べ替えを行うことで読みやすい文へと変換する方法はこれまでにいくつか提案されている。従来手法では、動的計画法であるCYK法をベースとした方法で並べ替えの候補を効率的に導出している。しかし、このアプローチを採用した結果、2文節とそれらに係る文節しか素性として利用することができない。

本論文では遷移型依存構造解析を用いた手法により、より多くの文節の素性に注目して語順整所を行う手法を提案する。遷移型依存構造解析では、最適解は保証されないものの、大域的な素性を容易に扱えるというメリットがある。一方で、探索空間が膨大になるため、これを狭めることが重要になる。提案手法では、入力文が文法上は正しいという仮定のもと、遷移型依存構造解析によって得られる日本語文が、後方修飾性、非交差制約、係り先の唯一性といった日本語の構文的制約を満たすものとなるための制約を設けている。語順整序は、状態に対するShiftとLeftArcと呼ばれる2つの動作の適用により進行するが、日本語の依存関係に関する制約を導入するために、提案手法では、入力文を後方から処理する。これにより、各状態に対して制約に違反する動作を自然な形で記述でき、不要な動作の適用を防ぎ、探索空間を狭めることができる。以下に並べ替え後は日本語の構文的制約を満たすが、並べ替え前に戻した場合に、日本語の構文的制約を満たさなくなるために解候補からは外されてしまう例を示す。

・後方修飾性に違反する候補例



・非交差制約に違反する候補例

