

審査の結果の要旨

氏 名 ノバック ジョセフ ロバート

本論文は「Advancing Finite State-based Spoken Language Processing Systems: Theoretical Principles and Practical Techniques (和題：有限状態制御の音声言語処理系の先鋭化：理論と実装)」と題し、全7章から成る。有限状態制御に基づく音声言語処理系は、個々のモジュールを統一的な視点から実装できるため、高度な最適化・統合化が可能である。本論文は、重み付き有限状態トランスデューサ (Weighted Finite State Transducer, WFST) による音声言語処理系を効率的に開発できるツールキットを構築すると共に、WFST を使った音声認識系及び書記素音素変換系を実装し、更にはその精度を向上させる手法を提案し、その有効性を実験的に検証した論文である。

第一章は「Introduction (序章)」と題し、近年その構成がますます複雑になる音声言語処理系、自然言語処理系に対して、全てのモジュールを統一的な視点で実装でき、より効率的な開発・最適化が可能である重み付き有限状態トランスデューサを紹介し、本論文が取り扱う問題についてその概要を述べると共に、論文構成を示している。

第二章は「Weighted Finite-State Transducers (重み付き有限状態トランスデューサ)」と題し、その理論的枠組と基礎的項目を説明している。

第三章は「Overview of Automatic Speech Recognition (自動音声認識の概要)」と題し、音声を認識する (音声を単語列に変換する) 技術が必要とするモジュール・機能について解説している。特徴量抽出、音響モデリング、言語モデリング、そしてデコーディング、各モジュールの概要を述べている。

第二章、第三章を受け、第四章は「WFST-based Automatic Speech Recognition(WFST に基づく自動音声認識)」と題し、第三章で説明した各モジュール・機能が、第二章で説明した WFST によって、どのように実装され、最適化され、系として統合されるのかについて論じている。WFST に基づく音声認識系は近年多くの研究者が参入している分野であるが、ここでは、どの実装系でも共通に見られる技術について論じている。具体的には、Grammar acceptor, Lexicon transducer, Context-dependency transducer, Silence class

transducer について紹介し、その統合手法も紹介している。この章では更に複数モジュールの統合とその最適化について、筆者が構築したツールキットを使って実験的に検討し、本ツールキットの動作確認を行うと共に、従来のデコーダとの比較を通して、本ツールキットの有効性を実験的に示している。

第五章は「Dynamic Grammars and Lookahead Composition (動的文法と先読み合成)」と題し、WFST に基づく音声認識系における、デコーディングの処理効率を向上させる方法について理論的に論じ、実験的にその有効性を示している。WFST はデコーディング処理が行なわれる前に、全ての仮説をメモリー上に静的に展開する手法として導入されたが、タスク規模が拡大するにつれ、動的な仮説展開手法 (on-the-fly composition) の必要性が叫ばれている。ここでは、先読みを動的に行う手法を提案し、その有効性を実験的に検証した。

第六章は「Grapheme-to-Phoneme Conversion in the WFST Framework (WFST フレームワークにおける書記素音素変換)」と題し、WFST を使った書記素音素変換の精度向上について検討している。書記素音素変換は、スペルのどの部分が各音素に対応するのかを推定する技術であるが、発音辞書 (単語スペルと音素列との対応表) を用いて実装される。従来構築された変換系を WFST に基づいて実装するとともに、再帰型ニューラルネットワーク型の言語モデルの利用、ベイズリスク最小化基準に基づいた変換系の実装などを通して従来システムよりも高い精度を示すことに成功している。

第七章では本論文を総括し、今後の課題について述べている。

以上要するに本論文は、Weighted Finite-State Transducer (WFST) に基づく音声言語処理系構築に対して、音声言語処理を専門としない研究者でも使えるツールキットを構築し、それを使って実装した音声認識系においてより簡素かつ効果的な動的仮説展開法を提案した。更に、WFST に基づく書記素音素変換系を実装すると共に精度向上を実現しており、情報理工学に貢献するところが少なくない。

よって本論文は博士 (情報理工学) の学位請求論文として合格と認められる。