

様式 10

論文審査の結果の要旨

報告番号	甲 先 第 297 号	氏 名	神田 峻介
審査委員	主査 北 研二 副査 獅々堀 正幹 副査 泓田 正雄		
学位論文題目			
Space- and Time-Efficient String Dictionaries 空間効率と時間効率の良い文字列辞書			
審査結果の要旨			
<p>大規模なデータを如何に効率的に管理できるかというのは基本的な課題の一つである。また、文書や Web ページ、ゲノムデータなど、あらゆるデータが文字列として表現されているため、効率的な文字列処理のためのデータ構造とアルゴリズムは多くの研究者や技術者によって提案されている。文字列処理のためのデータ構造はいくつか存在するが、文字列の集合を保管するためのインメモリデータ構造である文字列辞書の効率化に取り組んでいる。文字列辞書とは、伝統的に自然言語処理や情報検索において文書中に含まれる語彙を管理するために用いられてきたが、近年では Web 検索エンジンや RDF ストア、NoSQL、地理情報システムなどの大規模なデータセットを扱う用途においても用いられており、その空間効率が重要視されている。実際、静的な用途に限定すれば、既存のデータ構造で十分な空間効率は得られており、元のデータのおよそ 5% のサイズで辞書が実現できるほどに進展している。一方で、時間効率や構築コストとのトレードオフは存在している。データの更新を許す動的な用途に関しては、静的な辞書ほど進展しておらず高い空間効率は得られていない。本論文では、静的な用途と動的用途、それぞれに対して既存の問題を取り除くために、文字列辞書のための新しいデータ構造について研究成果をまとめたものである。</p>			
<p>以上、本研究は文字列辞書の静的な用途と動的用途に対しての既存の問題を取り除くための新しいデータ構造についての研究成果についてまとめられており、提案手法や実験によりその有効性を実証されたものであり、本論文は博士(工学)の学位授与に値するものと判定する。</p>			
<p>なお、本論文の審査には、森田和宏准教授の協力を得た。</p>			