

[別紙2]

審査の結果の要旨

論文提出者氏名 尾暮拓也

本論文は専門的な分野における電子化文書を専門的な知識を利用して検索する手法についてオントロジー工学を利用し、また検索アルゴリズムとしてベクトルモデルを概念レベルに拡張したモデルを用いる手法の性能と特性を検証しているものである。

第1章は序論で、キーワード検索システムを利用して文書検索を行う場合と専門家による文書の推薦を受ける場合を想定して比較し、専門家の持つ専門知識が果たす役割について考察を行い、専門家が問題領域を同定する過程を意味ネットワークモデルを用いて定性的にシミュレートし、有効な文書推薦を行うためには検索課題を持つ人間と専門家それぞれの持つ意味ネットワークが相似でなければならないとし、これから、専門分野の検索システムには専門知識を内蔵することが有効であるとし、専門知識を扱う枠組みとしてオントロジー工学を応用することを提案している。

第2章はオントロジーの概要について述べている。オントロジー研究の歴史的な経緯を紹介して、自動推論に用いられる記号に明確な意味定義を与えることが必要と考えられるようになった経緯や、記号に明確な意味定義を与えることによって人工知能システムや人間など知能機械一般のコミュニケーションが円滑になるという利点をまとめている。またさまざまな研究者によるオントロジーの定義を紹介し、本論文で提案される手法でオントロジーとして利用できる知識の必要十分条件を与えている。

第3章は情報検索の概要について、代表的な検索モデルの紹介のほか、適合性という観念を科学的に扱う問題の本質的な難しさ、実験方法の有効性と限界を中心に説明している。

第4章は関連研究をまとめており、文書検索を単語レベルから概念レベルに移行させることを目的としたアプローチを紹介している。また東京大学図書館で契約する論文検索サービスの検索結果を根拠にして、現段階では明示的にオントロジーを利用した情報検索モデルは報告されていないとしている。

第5章は提案する文書検索手法を解説しており、ベクトルモデルにオントロジーの概念項目を応用した概念レベルのベクトルモデルを提案している。また活性伝播モデルを用いてこのベクトル要素を決定する方法を具体化している。このモデルでは活性伝播率と呼ぶチューニングパラメータを導入している。

第6章は人工知能分野の文書を用いたクランフィールド型実験の結果を報告しており、再現率、精度と11点平均精度のそれぞれの評価基準でキーワード検索よりも明らかな性能改善が認められたとしている。具体的にはキーワード検索では再現率が最大0.4であったのに対し、人工知能オントロジーを利用した今回の手法で0.9付近まで改善すること、精度においても全般的に提案する手法が上回ること、11点平均精度では1.3倍以上の性能改善が観察されたとしている。

第7章は同様の条件のクランフィールド型実験によって活性伝播率を変化させた場合の性能の変化を報告しており、活性伝播率の最適値が存在したことからこの手法を利用するためにはチューニングが必要であろうこと、オントロジーの任意性などから最適値はこの手法が利用される条件に依存するであろうことを考察している。さらに実験では活性伝播率をオントロジーの構造に応じて2種類に区別しており、これらを別々に変化させた結果から、実験に用いたオントロジーの各抽象度の概念が検索性能に寄与した割合を概算で求めている。この考察として低い抽象度の付近の概念をより充実させることによって今回の実験よりも検索性能が向上する可能性を示唆している。

第8章は研究の展望を述べており、チューニングパラメータを機械学習させることが論理的に可能であることを説明し、また検索に用いているオントロジーの代わりに機械学習した概念体系を用いる方法を提案している。

第9章は結論である。

以上を要すれば、本論文はオントロジーを利用した新しい文書検索手法を提案してその有効性を明らかにするなど、これからの発展が見込まれる知的な文書検索の分野の基礎技術に関する知見を得ており、システム量子工学の発展に寄与するところが少なくない。よって本論文は博士(工学)の学位請求論文として合格と認められる。