

よいウェブサイトを構築するためには、一貫性を維持することが重要であると指摘されている。リンク構造の一貫性に着目し、一貫性を乱す不規則構造を発見する方法を研究したのが本研究である。これまでのリンク構造の研究が、特定の構造とのずれを発見したり、全体構造に関する指標を発見するものであるのに対して、特定の構造を定めずに、個々の不規則構造を発見する方法を提案していることが、本研究の特徴である。

本研究は、6章からなる。第1章では、ウェブサイトにおいて、一貫性が重要であることを文献により示した上で、リンク構造の一貫性に着目し、その一貫性を乱す不規則構造を検出することを本研究の目的とすることを述べている。第2章では、ウェブのリンク構造を有向グラフで表現したウェブグラフ、集合の所属関係をグラフで表現した所属関係グラフを説明した後で、ウェブグラフを集合の所属関係グラフとみなすことで、集合の外延性公理をウェブグラフの頂点の同値性の判定に適用することを提案している。これは大胆な飛躍である。このように適用することの正しさは、証明できる性質のものではないが、このように適用することに意味のある例が示され、また、これに基づいた分析方法により第5章で有効な結果を得ていることから、一定の妥当性があると認められる。第3章は数学的な説明に充てられている。ウェブグラフには循環構造があるため、ウェブグラフを集合の所属関係グラフと見なすことを一般のウェブグラフに適用可能とするためには、循環構造を許す集合論、つまり集合 A が A の要素であるようなことが可能な集合論が必要となる。そのような集合論は非有基的集合論と呼ばれるが、代表的なものとして、4つの定式化が知られている。そのなかから、ウェブグラフをそれぞれの所属関係グラフとみなした時の妥当性の検討により、Aczel の非有基的集合論を採用したことが説明され、Aczel の非有基的集合論を本研究の目的のために十分な範囲で簡略化したものが定義されている。次に、集合のまわりの構造を近似的に表す高階ランクと呼ばれる指標が提案される。高階ランクは近似の極限が Aczel の非有基的集合論と一致する指標であり、その提案は評価できる。第4章では、ウェブ構造を分析する具体的な方法として、簡約度分析と高階ランク分析の2つの方法を提案している。簡約度分析は、弧に関する不規則構造を発見する方法であり、高階ランク分析は頂点に関する不規則構造を発見する方法である。これらの分析方法のなかで、第1章で直感的に説明された不規則構造に、厳密な定義を与えている。第5章では、第4章で導入した2つの方法を、HWB、Google、Yahoo、外務省、首相官邸のウェブサイトに適用し、不規則構造の発見に成功している。ここで発見される不規則構造は第5章の意味のものであるが、第1章の意味の不規則構造との対応については、ウェブサイトの更新履歴の分析や、人工的なデータでの検討などを行っており、一定の説得力が認められる。第6章では結論として、構造の特徴や共通性の抽出への応用可能性に言及している。

本研究は、ウェブの構造の分析に集合論を適用するという独創的なアイデアにより、あらかじめ目標とする構造を設定することなく、不規則構造を発見できることを示した。その実現においては、非有基的集合論の適切な選択、高階ランクの数学的定式化と性質の分析、ウェブの性質を考慮に入れた分析方法の開発、ウェブサイトのデータ収集と分析方法の適用、分析結果の評価など、緻密な検証作業が行われており、信頼できる結果が得られている。本研究は、非有基的集合論の、構造の分析への適用の可能性を拓いたものとも云え、その寄与は大きいものと判断される。よって、本審査委員会は本研究が博士（学術）の学位を授与するにふさわしいものと認定する。