

集合知を利用した知識評価システムの提案

Proposal of knowledge evaluation system using collective intelligence

相模英太郎[†] 吉田富美男[‡]
Eitaro Sagami[†] Fumio Yoshida[‡]

†長岡技術科学大学 経営情報システム専攻
‡長岡技術科学大学 経営情報系

† Department of Management and Information System Engineering, Nagaoka University of Technology.

‡ Department of Management and Information Systems Science, Nagaoka University of Technology.

要旨

知識の効率的な活用には知識の評価が必要不可欠である。しかし、組織や人の持つ知識の評価は極めて高度な作業であり、基本的には人間が行わざるを得ないため時間や費用がかかるとともに客観性などにも課題が残る。そこで本研究では、客観的な知識評価を自動的に行うシステムを提案する。本システムでは、評価対象が生成した電子化情報を知識と見なしてこれをグラフ構造で表現する。同様に Web 上の電子化された情報を知識群とみなしてグラフ構造で表現し、これらの類似度合を、集合知を利用して算出することによって知識の評価を行う。

1. 研究背景

知識の効率的な活用は企業のみならず、様々な場面において重要である。知識を活用するためには、まずは知識を評価する必要がある。人や組織の持つ知識を評価することによって、知識の比較や、不足している知識の補充などが可能となり、適材適所での活用を検討することができる。

しかし、組織や人の持つ知識を評価するには極めて高度な知識と経験が必要であり、人間が行わざるを得ないが、これには多大な時間と費用が必要であるとともに、評価者によって結果が異なることも日常茶飯事であり、客観性にも課題が残る。

2. 目的

知識に関しては様々な研究が行われているが、少なくとも現時点では、人の持つ知識を正確に表現することさえ困難であり、ましてや人や組織の持つ知識を正確かつ客観的かつ自動的に評価する手法を直ちに確立することは難しいと言わざるを得ない。そこで本研究では、その第一歩として、正確な評価については多少目をつぶり、客観的な知識評価を自動的に行うシステムを提案しその効果を検証することを目的としている。本論文では、本システムの基本アルゴリズムを述べる。

3. 原理

知識の評価は、知識の比較に基づいて行われると考えることができる。そこで本研究では、以下の3つのSTEPで知識の自動的な評価を行うこととした。

- (1) 人や組織の持つ知識をコンピュータが処理できる形式で抽出する。
- (2) 抽出した知識を比較可能な形式で表現する。
- (3) 知識の比較に基づいて評価を行う。

それぞれのSTEPについて検討する。

(1) 知識の抽出

知識の抽出は非常に困難な問題であり、本研究ではこの問題には深入りせず既にデジタル化された情報のみを対象とする。これによって評価の正確性はかなり低減するがやむをえない。

(2) 知識の表現

本研究では、知識を評価することが目的であり、正確な表現ももちろん重要ではあるが、より比較しやすい客観的な表現が望ましいためグラフ構造で知識を表現することとした。グラフ構造は、もっとも基本的なデータ構造の1つであるとともに柔軟性と拡張性を持っており、ノードとフローの種類を増加することによって多様な知識構造を表現できると考えられる。図1に知識の表現例を示す。

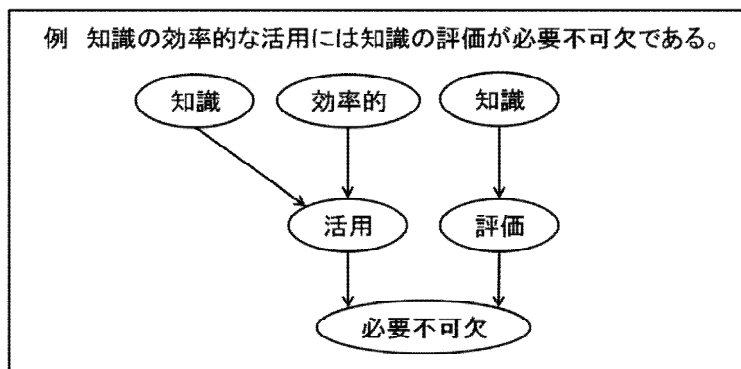


図1 グラフ構造による知識の表現例

(3) 知識の比較と評価

知識をグラフ構造で表現できれば、それぞれのグラフ構造を比較することによって知識を比較することができる。ただ、ある1人の対象者に対して基準となる人を1人用意し、上述のように、一部しか抽出されていない知識を、必ずしも正確とは言えない方法で表現し、比較しても妥当な評価が行われるとは考えにくい。しかし、基準となる人を100万人用意し、その比較結果を統合して評価を行ったとしたら、集合知の仕組み[1]を利用して、より客観性のある評価が行える可能性がある。

4. 基本アルゴリズム

上記の原理に基づいて検討したアルゴリズムの概要を説明する。

4.1. 知識の抽出

評価対象の知識としては、評価対象が生成した電子情報(主にテキスト情報)を利用することとする。評価対象が作成した業務書類やEmail、社内ブログなどを利用できる。

評価基準となる知識としては、集合知を活用するためにも非常に多くの知識が重要であるので、本手法ではWeb ページ上のテキスト情報を活用することとする。Web ページの収集は容易に自動化できるが、最初の段階では、評価の質を向上させるために、評価基準とする Web ページの選定は人間が行うこととする。

4.2. 知識の表現

評価対象および評価基準から抽出したテキストデータについて構文解析、形態素解析を行い、その結果をグラフ構造で表現する。解析結果をどの程度反映させるかによって必要なグラフ構造も変わるが、最初の段階では、単語や文章の係り受けのみを表現することとし、文章中から単語を抽出してノードとし、単語間の係り受けの関係をフローで表現することとする。

評価対象が作成した文章と、評価基準とする全ての Web ページから抽出したテキストデータを全てグラフ構造で表現する。これにより1つの知識は、1つまたは複数のグラフ構造で表現できる。

4.3. 知識の比較と評価

4.3.1 知識の比較

はじめに、以下の手順で評価対象のグラフと評価基準のグラフの類似度を算出する。そして全ての評価基準のグラフについて同様の処理を行う。

類似した単語の抽出

2つの知識表現に同じ単語が含まれているほど、それらの知識は類似していると言える。そこで、2つのグラフのノード同士を比較し、両方のグラフに存在するノードを抽出する。

グラフ構造の比較

単語だけが似ていても、2つの知識表現が類似しているとはいえない。グラフ構造では知識はノードとノードの結びつきによって表現されると考えられる。そこでそれぞれのグラフにおいてノード間の距離を1として抽出したノード間の距離を算出し、その自乗和を2つの知識の類似度とする。

4.3.2 知識の評価

評価対象のグラフと全ての評価基準のグラフの類似度をもとに知識の評価値を算出する。様々な算出方法が考えられるが、各評価基準のグラフとの類似度の平均値を評価対象の評価値とする。

5. まとめ

知識を効率的に活用するために、客観的かつ自動的な知識評価システムの基本アルゴリズムを提案した。知識をグラフ構造で表現して集合知を活用することにより、ある程度の信頼性を持った評価が行える可能性がある。今後は、本システムを実際に構築し、本手法の有効性を検証する予定である。

参考文献

- [1] ジェームズ・スロウィツキー, 「みんなの意見」は案外正しい, 角川書店, 2006.