

観光情報に関する概念形成セッションの提案と考察

北海道情報大学	*齋藤 一	SAITO Hajime
北海道大学大学院	金城 伊智子	KINJYO Ichiko
北海道大学大学院	長尾 光悦	NAGAO Mitsuyoshi
1004631 北海道大学大学院	大内 東	OHUCHI Azuma

1. はじめに

従来、観光に関する情報は、ホテルの経営者や、旅行代理店といったサービスの提供者側から発信されていた。このような情報は、経営戦略や情報提供媒体との利害関係の影響を受けやすいと考えられる。しかし、今日のインターネットの発達により、旅行者の実際の体験に基づく情報を発信、収集することが容易にできるようになってきている。

これまで本研究では、概念地図と構造モデリングを用いた概念形成支援法を提案し、その協調学習への応用を検討してきた[1]。もし、旅行者の持つ概念構造と、旅行サービスの提供者の概念構造が概念地図として比較することができたならば、旅行者の効果的なプランニングに利用できるだけでなく、サービス提供者の新たな戦略に利用できると考えられる。また、ホテル経営者が、利用者が発信する季節による観光情報の概念地図の違いから、利用者がどの時期に何を求めているかを的確に知ることができると期待している。

本論文では、上記を実現するための観光情報に関する概念形成セッションを提案し、その第一ステップとして大量の Web 情報からキーワード抽出し、効果的に可視化する方法を示す。

2. 概念形成セッション

ある概念を理解しようとするユーザのブラウジングは、以下のようにまとめることができる[2]。

- (1) サーチエンジンを使って、目標概念に関連のある Web ページを検索する。
- (2) 検索された Web ページのうち、役に立ちそうなページを見て理解する。
- (3) その Web ページにおいて、未知の概念を目

標概念として(1)にいく。

しかし、役に立ちそうなページの判断や、次の目標概念の決定は、ユーザの情報活用能力に依存する。

本研究では、サーチエンジンによって収集された大量の観光に関する Web 情報から、概念形成に役立つと考えられる概念(キーワード)を自動抽出し、それらの間の関連度を二次元空間(キーマップ)に表示する方法を提案する。本研究では、この「キーマップ」と「概念地図」を利用した概念形成支援プロセスの実行を「概念形成セッション」と呼ぶ。以下に概念形成セッションの概要を示す。

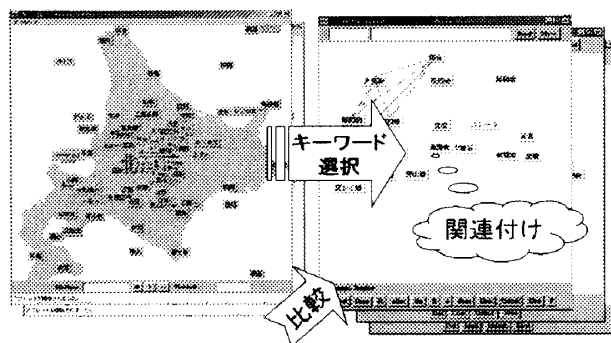


図1 概念形成セッションの概要

【概念形成セッション】

1. 問題領域の設定
2. 具象化：
 - 2.1 キーマップの作成
 - 2.2 キーワードの選択
 - 2.3 キーワード間の擬順序関係^{注1}の決定
3. 構造化：
 - 3.1 レベル分割、階層構造の抽出
4. 概念地図の描画
5. 概念地図の比較と戦略の導出

注1) 擬順序関係は、旅行プランの順序関係や、Web ページのリンクの親子関係を表すものとする。

3. キーマップによるキーワードの可視化

本論文では、概念形成セッションの問題領域を「北海道と観光」と設定し、「2.1 キーマップの作成」と「2.2 キーワードの選択」のプロセスについて詳しく述べる。

3.1 キーマップの作成手順

キーマップの作成手順を以下に示す。

1. Web 情報の収集
2. キーワードベクトルの作成
3. 不要語リストの作成
4. 関連度の計算
5. キーマップの描画

Web 情報の収集には、一般の検索サイトを利用する。北海道と観光についての概念形成を行うためには、収集された Web 情報中に必ず北海道の観光情報が記述されていなければならない。そこで、実験では、検索エンジンが収集した Web 情報の中から、目視によって、HTML 文書のソースファイルを確認し、北海道と観光に関係のない記述や、冗長する文書を削除した 100 件のホームページを利用する。

次に、集められた情報から、形態素解析ソフト[3]によって、一般名詞および固有名詞のみを抽出し、その頻度をカウントすることで、キーワードベクトルを作成する。ただし、一般名詞も含むことから概念形成には不相当であると考えられる名詞も数多く出現することから、不要語リストを作成し、これらを削除する。

キーワード間の関連度は、キーワード間の共起関係を表すものとして、キーワードベクトル間の内積を利用する。この内積の値を関連度としてキーマップの描画に利用する。

キーマップの描画には、[4]の方法を主に利用する。これにより、ランダムに初期配置されたキーワードは、関連度の高いもの同士が、近い位置へ、低いキーワード同士が、遠い位置へ移動する。

利用者は、このキーワードの移動の軌跡によって、キーワードの関連度（共起関係）を視覚的に知ることができる。

図 2 において、最高頻度のキーワードである「北海道」を中心に、共起するキーワードの幾つかのク

ラスタ（「スキー」「ゴルフ」「カヌー」などのクラスターや、「小樽」「函館」などの札幌周辺都市のクラスター等）ができています。

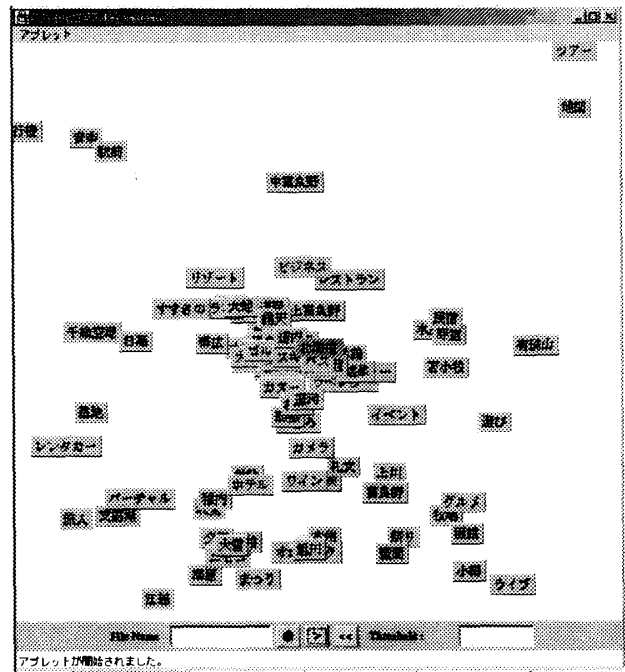


図 2 「北海道」と「観光」のキーマップ

4. まとめと今後の課題

本論文では、観光情報に関する概念形成セッションを提案し、キーマップの具体例を示した。今後は、キーワードまたはクラスター間の（推移的）関係を抽出する方法や、構造モデル（概念地図）の比較手法の検討を行い、概念形成セッションの完成を目指す。

参考文献

- [1] Hajime Saito, Azuma Ohuchi, and Takashi Maeda, "On the Use of a Structural Modeling for Collaborative Learning using the Concept Map", Proceedings of the 2nd Joint International Workshop, ORSJ Hokkaido Chapter and ASOR Queensland Branch, pp.277-284, 2000.
- [2] 山田 誠二, 村田 剛志, 北村 泰彦, "知的 Web 情報システム", 人工知能学会誌, Vol.16, No.5, pp495-502, 2001.7.
- [3] <http://chasen.aist-nara.ac.jp/>
- [4] K. Hori, "A system for aiding creative concept formation", IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, vol.24, no.6, pp.882-894, 1994.