

Web情報探索前後のコンセプトマップの分析: 探索者によるノード同定結果を用いて

江草由佳^{t1} 齋藤ひとみ^{t2} 中島諒^{t3}
高久雅生^{t4} 神門典子^{t5} 三輪眞木子^{t6}

本研究では、探索前後の知識の変化をとらえる評価手法の開発を目指し、探索実験前後のコンセプトマップの比較分析によって、知識変化がとらえられるかどうかを検討した。事前、事後コンセプトマップの統合ツールの開発したうえで、事前と事後のコンセプトマップ間の共通ノード判定方式による違いを分析した。実験参加者は2つの探索条件(発散・収束)に分けられ、2つのトピック(環境・旅行)の探索を行った。分析の結果、条件やトピックによってマップの特徴が異なることが明らかになった。環境課題よりも旅行課題の方がノード・リンク・リンクラベル数が多かった。これらの結果は、トピックや探索条件が利用者の探索行動や情報の選択に影響を与えた可能性を示唆している。また、実験参加者のノード同定による分析を行い、実験参加者による同一性判定を加味した場合、共通ノード数に平均で2ノード分の差があることが分かった。

Analysis of Concept Maps between before and after Search: Using Identified Nodes with Participants

YUKA EGUSA,^{t1} HITOMI SAITO,^{t2} RYO NAKASHIMA,^{t3}
MASAO TAKAKU,^{t4} NORIKO KANDO^{t5}
and MAKIKO MIWA^{t6}

In this study, we investigate the changes in concept map representations of a topic during their exploratory search on the Web by comparing the concept maps depicted before and after each users searching. We have developed a tool to merge two concept maps for the analysis. This paper discusses how a method to identify a common node in pre- and post-search maps effects the analysis results. Participants were instructed to search for and gather Web pages for the task of writing a magazine article on two given topics, ecology and trip, while working in either of two conditions: divergent and convergent. A comparative analysis of the concept maps between the pre- and post-search maps indicated

that the users in the divergent condition changed their knowledge widely and the users in the convergent condition changed their knowledge deeply. The results also showed that there were more nodes, links, and link labels in the trip task than in the ecology task.

1. はじめに

探索型検索 (Exploratory Search) とは、探索者が学習しながら情報ニーズを明確化しつつ、問題解決を進めるプロセスである¹⁾。現在、探索型検索を支援するシステムの開発が進められている。それにともなっ、探索型検索システムの評価方法の開発も必要になってきた。伝統的な情報検索システムであれば、クランフィールド実験にはじまった情報検索システム評価の仕組みが整いつつあるが、探索型検索システムの評価方法はいまだ模索が続いている。

探索型検索システムは伝統的な情報検索システムとは異なり、検索結果のランキング文書リストを再現率・適合率で測る評価ではなく、探索者がいかに学習し、知識を変化させたかを評価することがキーとなる。しかし、これまで提案されてきた探索型検索システムの評価方法は、直接的に知識の変化を図る方法ではなく、システムログや発話、インタビュー、アンケートに基づき、利用者がシステムを使ってうまく情報を検索できているか、また探索的な検索が由来しているかによって評価する方法であった²⁾。

そこで本研究では、探索前後の知識の変化をとらえる評価手法の開発を目指し、探索前後のコンセプトマップの比較分析することによって、知識変化がとらえられるかどうかを検討した。

-
- ^{t1} 国立教育政策研究所 教育研究情報センター
National Institute for Educational Policy Research
 - ^{t2} 愛知教育大学 教育学部
Aichi University of Education
 - ^{t3} 住友電工情報システム株式会社
Sumitomo Electric Information Systems Co., Ltd.
 - ^{t4} 物質・材料研究機構 科学情報室
National Institute for Materials Science
 - ^{t5} 国立情報学研究所 / 総合研究大学院大学
National Institute of Informatics / Graduate University for Advanced Studies
 - ^{t6} 放送大学 ICT 活用・遠隔教育センター
The Open University of Japan

る知識にどのような違いが出るかについては、明確な結果は示されていない⁸⁾。

著者らは、これまで、探索のトピックや探索条件が明らかに異なる状況として、環境問題と旅行というトピック、収束と発散という探索条件を設定し、探索の前後のコンセプトマップの比較分析を行ってきた。その結果、条件やトピックによってマップの特徴が異なることが明らかになっており⁹⁾、コンセプトマップによって知識の変化を図ることができると可能性を示唆している。本稿ではこのコンセプトマップを用いた実験の分析をさらに進め、とりわけ、探索者自身による探索前後のコンセプトマップ中のノードの同一性判定を用い、より精緻なコンセプトマップ分析を目指すとともに、評価手法としての考察を試みる。

2. 実験

2.1 実験デザイン

実験は、探索トピック（環境問題・日帰り旅行）を被験者内要因とし、探索条件（発散・収束）を被験者間要因とする2×2の混合計画で実施した。探索のトピックは、雑誌の編集者として働いているというシナリオを設定し、「環境問題」は様々な環境問題を紹介する連載記事を、「日帰り旅行」はいろいろな東京からの日帰り旅行を紹介する連載記事を書くための情報を集めるためにWeb検索をするものであった。

また探索範囲は、各回の連載で扱う内容を幅広く調べる「発散」と1回分の連載で扱う内容を詳細に調べる「収束」を設定した。2つの探索トピックについて、それぞれ発散・収束の課題を用意した。

2.2 実験参加者

学部生32名（男性16名、女性16名）が実験に参加した。参加者の専攻は、文学、法学、理学、教育学、工学、政治経済学、商学、理工学など多様であった。

2.3 情報検索課題

課題は、雑誌の編集者という状況を設定し、連載記事のアイデアのための情報を収集するというものであった。発散条件では、各回の連載記事で扱う内容を網羅的に集めるという指示を行った。また収束条件では、連載1回分の記事を書くための詳細な情報を集めるという指示を行った。

各トピックの探索条件ごとの課題文の一部を表1に示す。下線はトピックや探索条件によって異なる箇所を示しており、波線はトピック間、直線は探索条件間での違いを示している。どちらの課題においても、15分間で必要な情報を集めること、収集したページはブックマークに登録することが求められていた

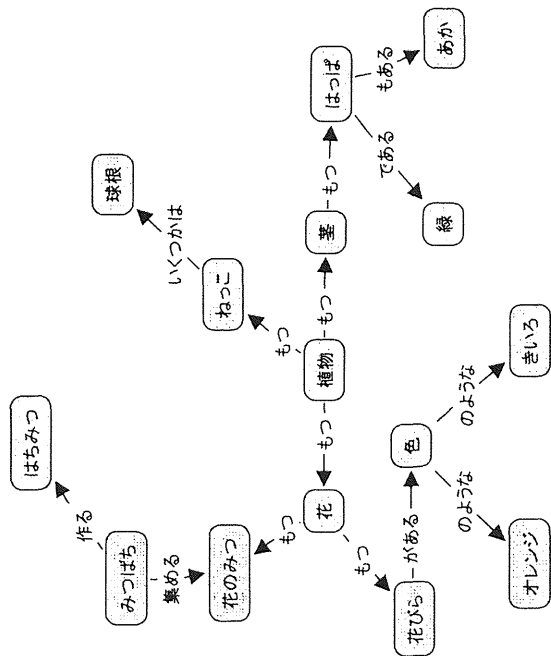


図1 コンセプトマップ

コンセプトマップとは、物や出来事に対する知識やアイデアを単語やフレーズとしてごとばに表し、それらのことばをその概念ごとに整理し、視覚的に表現したものである¹⁰⁾。コンセプトマップは、モノやコトを表すノード、ノード同士をつなぐリンク、ノードとノードの関係を表すリンクラベルで構成される。図1は、「植物」に関するコンセプトマップの例である。植物や花、茎などの丸で囲まれた名詞がノード、矢印がリンク、矢印の上にかかれたもつ、のようなのがあるなどがリンクラベルを表す。

コンセプトマップは、これまで学習者の知識や理解を測定する指標として利用されてきた。授業やコースの前後でコンセプトマップを比較すると、事後のマップの方が形状が複雑になり、ノードやリンク、リンクラベルの数が増加することや、誤概念が減り有用な概念が増加することなどが報告されている^{4),5)}。

情報検索の分野においても、探索の前後でコンセプトマップを比較した研究が行われている。それらの研究では、分野による違いや分野に対する経験がコンセプトマップの形状などに違いを与える可能性、長期的なスパンでの比較、行動とマップとの関係、シナリオの違いが影響を与える可能性が示されている^{6),7)}。しかし、どんな情報を収集するかによって、得

表 1 情報検索課題

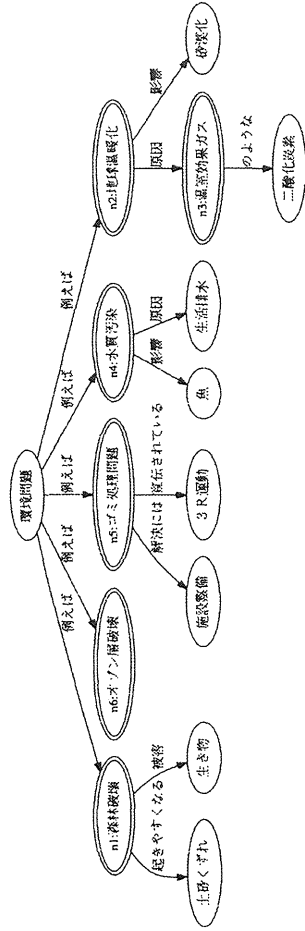
	環境問題	東京からの日帰り旅行
共通	あなたは、全国紙の新聞社でニューズ雑誌の編集者として働いています。あなたは編集長から、次号から始まる様々な環境問題を紹介する連載記事の情報を集めるよう指示を受けました。	あなたは、大手の出版社で旅行雑誌の編集者として働いています。あなたは編集長から、次号から始まる様々な東京からの日帰り旅行を紹介する連載記事の情報を集めるよう指示を受けました。
発散	編集長は、会議で各回の連載記事で採り上げる内容を示しています。そのため連載1回分の記事に関する詳細な情報よりも、各回の記事でどんな問題を抱えているかを議論するための多様な情報を必要としています。	編集長は、会議で各回の連載記事で採り上げる内容を示しています。そのため連載1回分の記事に関する詳細な情報よりも、各回の記事でどんな問題を抱えているかを議論するための多様な情報を必要としています。
収束	編集長は、会議で連載1回分の記事について具体的な記事の内容を示しています。そのため複数の旅行に関する情報よりも、特定の旅行についての記事を書くための詳細な情報を必要としています。	編集長は、会議で連載1回分の記事について具体的な記事の内容を示しています。そのため複数の旅行に関する情報よりも、特定の旅行についての記事を書くための詳細な情報を必要としています。
共通	もちろん、読者の興味を引くものが良いと考えています。サーチャージンを使って情報を探し、編集長が求める情報を収集してください。Web ページはブックマークに登録してください。	もちろん、読者の興味を引くものが良いと考えています。サーチャージンを使って情報を探し、編集長が求める情報を収集してください。Web ページはブックマークに登録してください。

2.4 手続

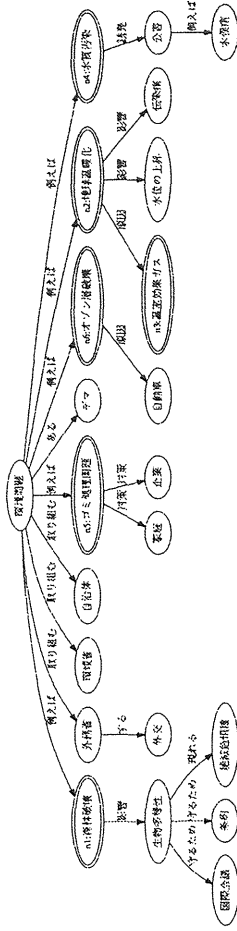
実験参加者は、事前アンケートとして、Web サーチャージンを用いた日常の検索経験やインターネットの利用経験について回答した。次にコンセプトマップの書き方についての指示と練習に取り組んだ。

課題についての指示の後で、トピックについてのコンセプトマップを10分間記入した。コンセプトマップの記入には、トピックを表す語「環境問題」「東京からの日帰り旅行」が中心にノードとして記載された用紙を使用した。次にWeb 検索を15分間行った。検索終了後、Web 検索で得た知識を踏まえて、コンセプトマップを事前と同様に10分間記入した。また課題アンケートとして、課題についての既有知識、興味、課題の難易度、情報収集の難易度、課題の感想を回答した。課題の指示から課題アンケートまでをトピックごとに2回繰り返した。

最後に、2つの課題の比較評価、検索前後での知識の変化や事前と事後のコンセプトマップを比較したときの感想などについて回答した。最後のアンケート回答時に、実験参加者に事前と事後のコンセプトマップのコピーを配布し、事前と事後のコンセプトマップで同じ意味を持つ概念を赤ペンで囲み、対応する概念同士に同じ番号を振る作業を行った。



(a) 事前コンセプトマップ



(b) 事後コンセプトマップ

図 2 ある参加者の環境問題のコンセプトマップ

3. 分析手法

3.1 ノード、リンク、リンクラベル数の分析

コンセプトマップのノード・リンク・リンクラベルについて、事前と事後のマップに含まれるそれぞれの数をカウントした。図 2 は参加者が作成した環境問題についてのコンセプトマップである。図の二重線で囲まれたノードは、参加者が同定判定において同じ意味をもつと判断したノードであり、記号が同じものが同じと判断されたノードを示す。図 2(a) の場合、ノード数は 15、リンク数は 14、リンクラベル数は 14 となり、図 2(b) の場合、ノード数は 23、リンク数は 22、リンクラベル数は 22 となる。

3.2 共通ノード

検索前後のコンセプトマップの比較をするために、事前と事後の共通部分を算出する方法

を検討した。そこで、ノードをキーとして共通部分を算出できないうかが検討をした。

検索前後のコンセンプトマップでの共通ノードの候補には以下の3種が考えられる。

赤丸ノード 実験参加者によって、事前と事後のマップについて、同じ意味をもつ概念を赤

ペンで囲み、同意判定を行ったノード

ルートノード コンセンプトマップにはじめから中心に書いてあるルートノード

同文字列ノード ノードラベルの文字列が完全一致するノード

同文字列ノードはルートノードを含む。また、赤丸ノードの中には、同文字列ノードでもあるノードがある場合がある。赤丸ノードは、探索者自身による判定であるため、記述の揺らぎや同義語などを含めもともと正確な同じ意味を持つ概念を同定できる可能性がある。しかし、見落としなど判定者のミスの可能性もありうる。同文字列ノードは機械処理で同定でき、分析コストを抑えられる点がメリットである。しかし、同じ意味を持つであっても記述の揺らぎや同義語などの場合、同一概念の共有ノードとして扱えない場合など、取りこぼしがありうる点がデメリットとなる。また逆に、同じ文字列であっても違う文脈で扱われ異なる意味を持つ場合であっても、同じ概念ノードとなつてして扱われてしまうという欠点もある。そこで、共通ノードを赤丸ノードと同文字列ノード両方を含むと定義することも可能である。

システム評価の手法の開発という観点では、分析コストの低さや分析の際の揺らぎの低さが重要になってくる。そこで、赤丸ノード、同文字列ノードに大きな違いがあるか、それとどの定義でもさほど差はないのかについて検証した。

3.2.1 共通ノード数の分析

共通ノードの種類を区別し数種類のカウンタ方法によって結果に影響するかどうか分析した。共通ノードの定義として、以下の3つの定義をそれぞれ用いた場合、それぞれについて検討した。

定義1 同文字列ノードのみを共通ノードとする

定義2 赤丸ノードのみを共通ノードとする

定義3 同文字列ノードと赤丸ノードを合わせたものを共通ノードとする

4. 統合コンセンプトマップの可視化ツール

事前コンセンプトマップと事後コンセンプトマップとを統合したマップを可視化するツール

VizCmap^{*1}を開発し、分析にあたっては適宜、このツールを参照しながら事前・事後コンセンプトマップを比較した。図2は、図3に示した事前コンセンプトマップと事後コンセンプトマップを統合したコンセンプトマップである。以下はこの図を例として説明する。

二重枠線で表現したものは全て共通ノードである。ルートノード(dot表現としては、最初に出現するルートノードであり、かつ、idがid0のノード)は青の二重線で囲んだ矩形ノードとして表現している(ラベル「環境問題」のノード)。赤丸ノード(dotの表現としては、ラベルが「n1:」といったコロロン記号の前に書かれたプレフィックス表記で同一性を表現したノード)は赤の二重太線で囲んで表現している(例:「森林破壊」のノード)。ルートノード、赤丸ノード以外の同文字列ノードは黒の二重枠線で表現している。

ルートノード、赤丸ノード、同文字列ノードを手掛かりにして、共通リンク、共通リンクラベル、消失リンク、消失リンクラベル、新規ノード、新規リンク、新規リンクラベルも表現する。両端がルートノード、赤丸ノード、同文字列ノードのいずれかである場合は、共通リンクとし、二重線の矢印で表現している(例:「地球温暖化」から「音質効果ガス」へのリンク)。共通リンクのリンクラベルが事前と事後マップで同文字列となつている共通リンクラベルは下線付で表現している(例:「地球温暖化」から「音質効果ガス」へのリンクのリンクラベル「原因」)。事前コンセンプトマップにのみ出現し、事後コンセンプトマップに現れなかった「消失ノード」は点線の枠線で囲まれ(例:「土砂くずれ」)、「消失リンク」は点線で(例:「森林破壊」から「土砂くずれ」へのリンク)、「消失リンクラベル」は、グレー色で表現している(例:「森林破壊」から「土砂くずれ」へのリンクのリンクラベル「起きやすくなる」)。それ以外の実線で囲まれたノードで黒文字で表現されているものは、事後コンセンプトマップにのみ出現する「新規ノード」(例:「生物多様性」)、「新規リンク」(例:「生物多様性」から「国際会議」へのリンク)である。同様に、実線のリンクは「新規リンクラベル」である(例:「生物多様性」から「国際会議」へのリンクのリンクラベル「守るため」)。

図3の統合コンセンプトマップをみると、この事例の場合、ルートノードから直接リンクがあるノードに赤丸ノードが多いこと、共通リンクや共通リンクラベルも共通リンクノードと同様にあることがわかる。

*1 <http://cres.jpn.org/?VizCmap>

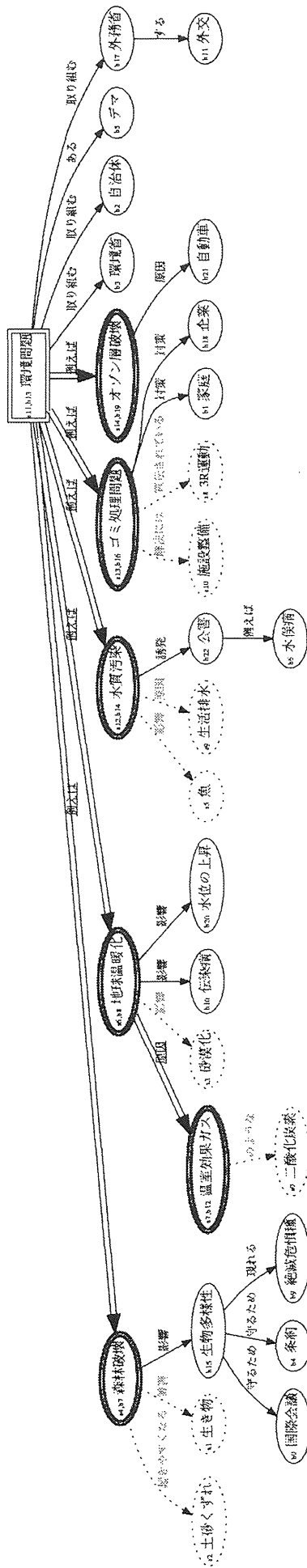
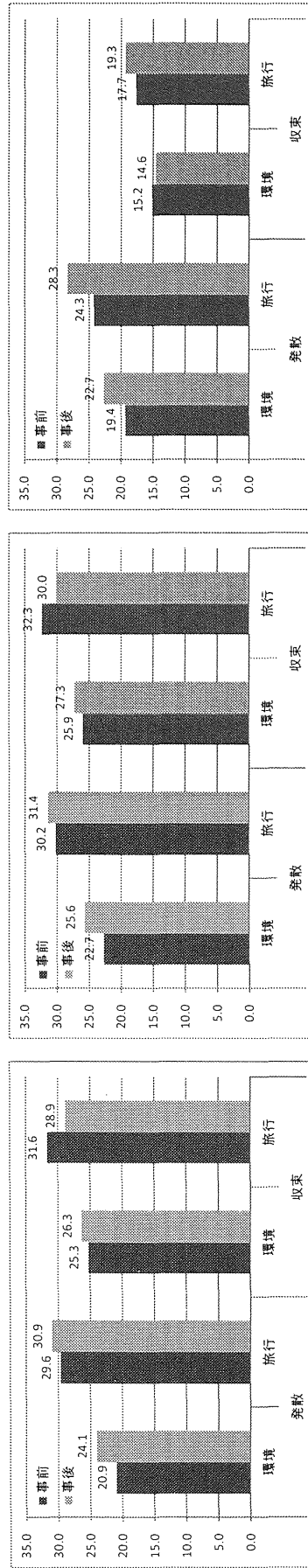


図3 統合コンセプトマップ



(a) ノード

(b) リンク

(c) リンクラベル

5. 結果

5.1 ノード、リンク、リンクラベル数の分析結果

図4は検索トピックおよび探索範囲条件ごとのノード、リンク、リンクラベル数の平均を示している。探索範囲条件、検索トピック、検索前後の3要因混合分散分析の結果、ノードとリンクについては検索トピックと検索前後の交互作用がみられ(ノード: $F(1,30)=4.38, p < .05$;

リンク: $F(1,30)=3.36, p < .10$), 検索前後に関わらず環境問題より日帰り旅行の方がノード・リンク数が多かった(ノード事前: $F(1,30)=30.00, p < .01$; ノード事後: $F(1,30)=14.10, p < .01$; リンク事前: $F(1,30)=21.26, p < .01$; リンク事後: $F(1,30)=10.46, p < .01$). また環境問題は検索前後でノード・リンク数が増加した(ノード: $F(1,30)=7.48, p < .01$; リンク: $F(1,30)=5.03, p < .05$). 探索範囲条件による違いは見られなかった。

リンクラベル数は、ノード・リンク数とは異なり、探索範囲条件、検索トピック、検索前後で

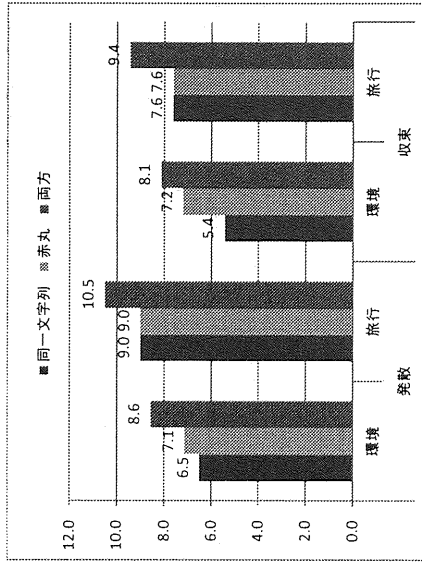


図 6 共通ノード数の平均

表 2 共通ノードの定義による共通ノード数, 新規ノード数, 消失ノード数

定義	共通ノード数	新規ノード数	消失ノード数
定義 1 (同一文字列)	7.1	19.9	19.29
定義 3 (同一文字列+赤丸ノード)	9.1	16.96	16.18

ノード数の定義としては、1) 同一文字列ノードによるもの、2) 赤丸ノードによるもの、3) 2つの定義の和集合をとったものをそれぞれ示す。旅行トピックでの定義1と定義2の平均値に差が無い点を除けば、おおよその傾向として、定義1から定義3に向かうにつれて、共通ノード数が増えていることが分かる。全体を平均したときの共通ノード数は、定義1 (同一文字列条件) と定義2 (赤丸ノード条件) で0.6 (8%) 増加し、定義1 (同一文字列条件) と定義3 (同一文字列条件+赤丸条件) では2.0 (28%) 増加する (表2)。共通ノード数が増える一方で、新規ノード数と消失ノード数はそれぞれ7.9 (15%)、7.1 (16%) 減少する (表2)。

なお、表2に見られるとおり、全体的な傾向として、共通ノードよりも新規ノードと消失ノードのほうが7~8も多いことから、検索前後でマップに書かれた知識は大きく変化している。

6. まとめ

分析の結果、検索の前後でコンセプトマップが大きく変化する、また検索トピックや

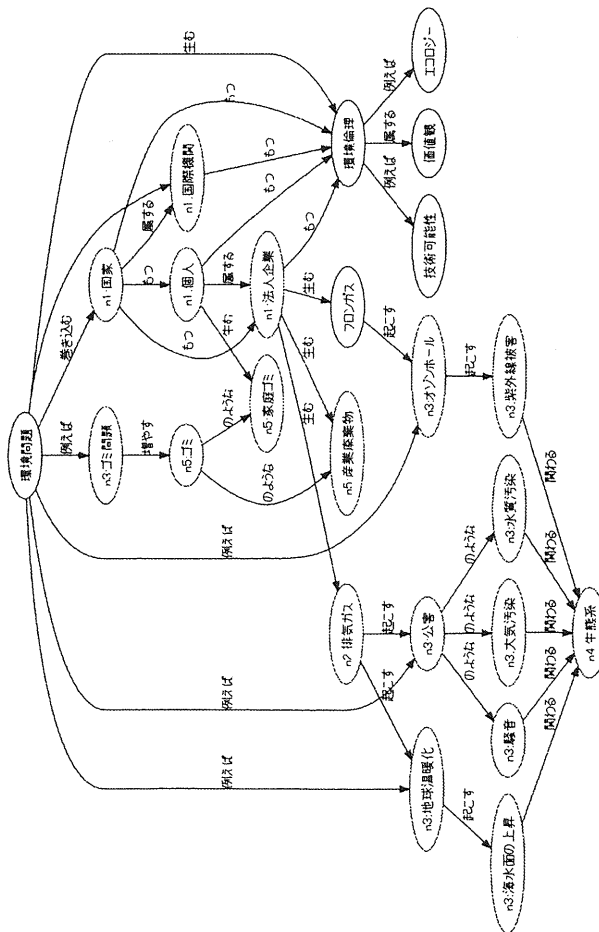


図 5 赤丸を付ける基準が広めなコンセプトマップの例

の主効果が見られた (探索範囲条件: $F(1,30)=4.74, p<.05$; 検索トピック: $F(1,30)=15.10, p<.01$; 検索前後: $F(1,30)=6.34, p<.05$)。探索範囲条件では収束より発散の方が、検索トピックでは環境問題よりも日帰り旅行の方が、検索前後では事前より事後の方がリンクラベル数が多かった。

5.2 探索者自身による共通ノード判定の結果

どのような赤丸ノードがあるか事前に確認したところ、実験参加者により、同一性判定に違いがある場合が見受けられ、同じ番号の赤丸が見られたコンセプトマップの例である。例えば、{ 国家、個人、法人企業、国際機関 } という4つの概念のノードを同一と判定していた。このようにノードの同一性判定には若干の個人差が見られたが、検索トピックや探索範囲条件による違いは見受けられなかった。

5.2.1 共通ノード数の分析結果

図 6 は検索トピックおよび探索範囲条件ごとの共通ノード数の平均を示している。共通

探索範囲条件によって、マップの特徴に違いがあることが明らかになった。まずトピックでは、環境よりも旅行の方がノード・リンク・リンクラベル数が多いことが分かった。また、環境よりも旅行のほうが赤丸ノード数が多いことも分かった。これらの結果から、同じトピックでも探索範囲が異なることによって参加者の収集する情報が異なり、それがマップにも反映された可能性が示唆される。

また、共通ノードの判定に探索者自身の判定による結果を加えることにより、共通ノード数は全体で28%増えることが分かり、コンセプトマップ間の比較の際には表記揺れによる影響等を考慮する必要があることが示唆された。しかし、実験参加者によっては、赤丸ノードを多めに同定する場合もみられ、コンセプトマップを使ったシステム評価手法の開発という点では、教示など実験の実施方式に配慮が必要であることも分かった。

参 考 文 献

- 1) White, R. W. and Roth, R. A.: *Exploratory Search: Beyond the Query-response Paradigm (Synthesis Lectures on Information Concepts, Retrieval & Services)*, Morgan and Claypool Publishers (2009).
- 2) White, R. W., Marchionini, G. and Muresan, G.: Evaluating exploratory search systems: Introduction to special topic issue of information processing and management, *Inf. Process. Manage.*, Vol.44, No.2, pp.433-436 (2008).
- 3) J.D. ノヴァック, D.B. ゴーウイン: 子どもが学ぶ新しい学習法概念地図法によるメタ学習, 東洋館出版社 (1992).
- 4) Rebich, S. and Gautier, C.: Concept Mapping to Reveal Prior Knowledge and Conceptual Change in a Mock Summit Course on Global Climate Change, *Journal of Geoscience Education*, Vol.53, No.4, pp.355-365 (2005).
- 5) Meagher, T.: Looking Inside a Student's Mind: Can An Analysis of Student Concept Maps Measure Changes in Environmental Literacy?, *Electronic Journal of Science Education*, Vol.13, No.1, pp.1-28 (2009).
- 6) Penanen, M. and Vakkari, P.: Students' Conceptual Structure, Search Process, and Outcome while Preparing a Research Proposal: A Longitudinal Case Study, *Journal of the American Society for Information Science and Technology*, Vol.54, No.8, pp.759-770 (2003).
- 7) Cole, C., Lin, Y., Leide, J., Large, A. and Beheshti, J.: A Classification of Mental Models of Undergraduates Seeking Information for a Course Essay in History and Psychology: Preliminary Investigations into Aligning their Mental Models with Online Thesauri, *Journal of the American Society for Information Science and Technology*, Vol.58, No.13, pp.2092-2104 (2007).

8) Egusa, Y., Saito, H., Takaku, M., Terai, H., Miwa, M. and Kando, N.: Using a Concept Map to Evaluate Exploratory Search, *IIX2010: Proceedings of the Third Symposium on Information Interaction in Context*, New York, NY, USA, ACM, pp.175-184 (2010).

9) 齋藤ひとみ, 中島 諒, 江草由佳, 高久雅生, 寺井 仁, 神門典子, 三輪眞木子: 情報検索の前後におけるユーザーの知識構造の変化: コンセプトマップを使った分析, 情報知識学会誌, Vol.21, No.2, pp.137-142 (2011).