

個人的主観に基づくハイパーリンク構造 日記データからの連想構造の分析

Associative Structures Analysis of the Hyperlink Type Diary Data based on the Individual Subjectivity

奥 健太 河村 竜幸 河野 恭之 木戸出 正継
Kenta OKU Tatsuyuki KAWAMURA Yasuyuki KONO Masatsugu KIDODE

奈良先端科学技術大学院大学 情報科学研究科
Graduate School of Information Science, Nara Institute of Science and Technology

We propose a system that supports to arrange experiences of a human's lifetime, which are recorded by a wearable computer. When we record our lifetime experiences, they should be arranged them to retrieve and browse them. We consider that these experiences are arranged by associative structures. In this paper, to find the best way to arrange them, we analyze his/her diary data, which has hyperlink structures.

1. はじめに

近年、ユビキタス/ウェアラブルコンピューティングの時代を迎え、記憶媒体の小型化、大容量化、記録装置の充実により、体験記録の実現が可能となってきた。体験記録とは、人が生涯にわたり「みたもの」、「きいたもの」、「かいたもの」などの体験をすべて記録することである。我々はこのような体験記録として収集されたデータ（以降、体験記録データ）は、後に閲覧されたり、検索されたりすることに意味があると考え、自己の過去の体験を回想したり、他者と体験を共有することによりコミュニケーションを活性化させたり、さらには、未来の行動決定の際に過去の体験を再利用したりすることを可能とさせるためである。そこで、体験記録データを記録、収集するだけでなく、そのデータを整理することに着目する。我々は体験記録データを「生涯体験アルバム」として整理することを考え、そのアルバムの作成を支援するシステムの構築を目的とする。

生涯体験アルバムの持つべき構造として連想構造を考える。一般に人間は連想的に記憶を想起することが多く、その記憶想起過程を模倣することにより、直感的なデータの検索が可能となることが考えられる。しかし、現段階においては、連想構造とは具体的にどのような構造であるのかは明確ではない。そこで、本稿ではまずその連想構造について分析する。分析対象データとして、ハイパーリンク構造の日記データを用いる。これは日記を含め、写真や体験記などの種々の体験記録データがハイパーリンクにより関連付けられているデータであり、そのハイパーリンクの関係をみることで連想構造について分析する。

2. 生涯体験アルバム：コンセプト

先行研究により、体験記録データを収集するための装置や方式が充実しつつある。上岡ら [Ueoka 03] は、ウェアラブルコンピュータを装着することで、臨場感のある個人の主観的体験を客観的に記録する方式の研究を行っている。美崎ら [Misaki 04] は、個人の体験（見たもの、書いたもの）を、カメラやスキャナなどによりすべてデジタル化している。それらを住宅に設置された多数のディスプレイによるスライドショーにより閲覧

している。自己の体験を常時受動的に閲覧することにより、記憶想起活動が活性化することを示した。

これら先行研究が個人の体験の記録および収集を目的としているのに対し、我々はこれらの閲覧を容易にするために、記録、収集された体験記録データを整理することを目的としている。その体験記録データから生涯体験アルバムの作成を支援するシステムを検討する。生涯体験アルバムの持つべき構造として連想構造を考える。図1は、生涯体験アルバムによる体験記録データの連想的な閲覧のイメージである。図1の例では、テニスラケットに関するデータ（例えば写真など）をきっかけとし、それから連想的にテニスに関するデータが提示され、共にテニスを楽しんだ仲間であったり、試合やその大学などの思い出が提示される。さらに連想が進むと、仲間と共に遊んだ思い出や、大学での授業などが提示されることになる。このような連想的な閲覧により、テニスラケットなどのようなふとしたきっかけから、意外なデータが発見されることが期待される。

生涯体験アルバムを閲覧することにより、過去の思い出を連想的に回想することによる楽しみを得たり、このアルバムを他者と共有することにより、コミュニケーションの活性化につながることを期待される。また、楽しみの提供だけでなく、記憶想起支援にもつながる。さらには、過去の体験をふり返ることにより、未来の行動の予測や決定などにつながることを期待される。

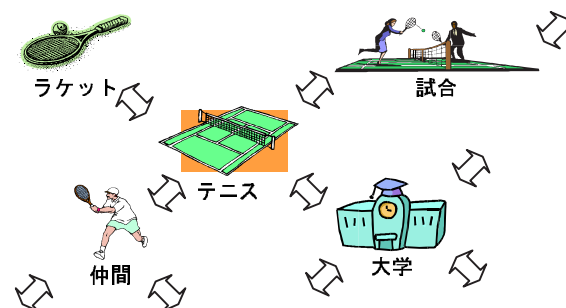


図1: 生涯体験アルバムによる連想閲覧イメージ

連絡先: 奥 健太, 奈良先端科学技術大学院大学 情報科学研究科, kenta-o@is.naist.jp

3. 分析概要

3.1 分析対象データ

分析対象データとして、ハイパーリンク構造の日記データを用いる。このデータには日記記述者の一生分の個人体験、つまり、生涯を通して「みたもの」や「かいたもの」がデジタルデータとして存在する。ここで、「みたもの」とは、写真などの静止画および書籍、雑誌などの紙媒体上の情報であり、「かいたもの」とは、日記をはじめ手紙や手帳、原稿などである。対象データは、表 1 に示すように、31,940 のテキストデータおよび 602,938 の静止画データから構成され、それぞれが相互にハイパーリンクにより関連付けられている。なお、プライバシー情報となるデータは削除されており、削除データ数は、テキストデータが 1,518、静止画データが約 4,240 である。それぞれ、データ全体の 4.5%、0.7% にあたる。図 2 は、ハイパーリンク構造の日記データのイメージ図である。ある日記ファイル中に、その日の体験に関連する他の日記データや、テキストデータ、静止画データへのハイパーリンクが貼付されている。2 章で挙げたイメージを例とすると、日記中にラケットという語句が出現し、それに関連し、テニスをやっていたときの写真や試合の内容に関する記述、またテニス仲間の名前やプロフィールなどが記述されているというものである。本稿では、このデータ間のハイパーリンクの関係をみることによって、人間の思考過程に基づく連想構造について分析する。

表 1: 分析対象データ (2004/12/14 現在)

データ形式	内容例	データ数	削除データ数
テキスト	日記, 手紙, 手帳, 原稿など	31,940	1,518(4.5%)
静止画	写真, 書籍, 雑誌など	602,938	4,240(0.7%)

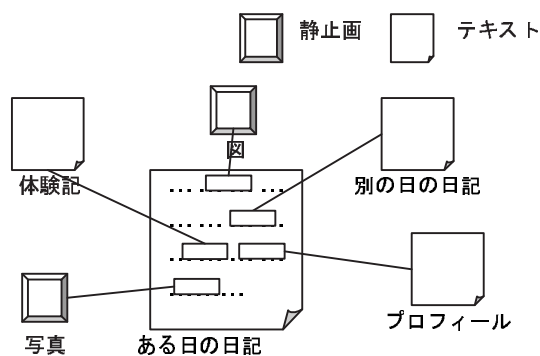


図 2: ハイパーリンク構造日記データイメージ

3.2 分析の流れ

ハイパーリンク構造の日記データを対象に、そのハイパーリンクの関係を以下に示す 2 方向からみることにより、連想構造について分析する。

1. 順方向リンク分析
2. 逆方向リンク分析

まず被験者の思考過程にしたがって、ハイパーリンクが貼付される対象となるデータや、そのデータを貼付する意図を見

ることにより、連想構造の概観を把握する。そのために順方向リンク分析を行う。順方向リンク分析では、日記データに注目して、その日記データと、そこからリンクされている対象データとの関係を分析する。つづいて、連想構造の概観からさらに深く分析するために逆方向リンク分析を行う。逆方向リンク分析では、ハイパーリンクの対象データに注目して、そのデータのリンク元を参照することにより、リンクが貼付されたきっかけ、さらにそのきっかけと対象データとの関係を分析する。なお、順方向リンク分析の対象データ数は 53、リンク数が 914 であり、逆方向リンク分析の対象データ数は 30、リンク数が 592 である。

4. 順方向リンク分析

4.1 分析内容

順方向リンク分析は、日記データ中に出現するハイパーリンクに注目し、

- リンク契機語句：ハイパーリンクを貼付するきっかけとなった日記中の語句
- リンク先データ種別：ハイパーリンクの対象となったデータの種別 (表 2 参照)
- リンク貼付意図：ハイパーリンクを貼付した意図 (表 3 参照)

をそれぞれ抽出し、それらを表 2,3 にしたがって分類する。順方向リンク分析のためのデータ分類の流れを以下に示す。

1. 日記データ中のハイパーリンクの抽出
2. ハイパーリンクの契機となる語句の抽出
3. リンク先データの種別分類
4. リンク意図抽出、分類

まず、日記データ中に貼付されているハイパーリンクを抽出する。つづいて、そのハイパーリンクの契機となる語句を抽出する。この抽出はリンク契機語句として断定できるものもあれば、推定によるものもある。例えば、日記中のある語句に隣接してハイパーリンクが貼付されていれば、それはそのハイパーリンクの契機語句として断定できるが、文につづいてハイパーリンクが貼付されている場合、その文中に出現するいずれかの語句が契機語句であるということ推定しなければならない。次に、ハイパーリンク先のデータの内容に基づいて、そのデータの種別を表 2 のように分類する。さらに、ハイパーリンクを貼付した意図を推測し、表 3 のように分類する。この分類を各ハイパーリンク、各日記データに対して行い、その統計を取ることで分析を行う。

表 2: リンク先データの種別分類項目

上位分類	下位分類
日記	-
体験記	個人体験, 人物・組織とのやりとり
メモ	人物, 組織, 場所, 物品, 参照, 予定
資料	イベント, 記事, 物品
作品	書籍, 漫画・劇, 音楽, 物品
システム	ファイル, メール, プログラム, サイト

表 3: リンクの貼付意図の分類項目

リンク先データ	分類項目
日記	体験日, 追記日, 対象日, 関連日
その他	体験, 関連, 利用, 参照, 対象, 等価

4.2 分析結果

分析結果の一例を表4に示す。表4は、分析対象データにおける各リンク先データ種別の度数と各リンク貼付意図の度数とのクロス統計を取ったものである。表4は、各リンク先データ種別に対するリンク貼付意図の割合を表している。例えば、日記データにおいては、日記中に記述されたある体験をした日として貼付されたリンクが日記データ全体の.570あり、日記中に出現した語句に関連する日として貼付されたリンクが日記データ全体の.204あることを示している。ここでは、リンク先データ種別として表2、リンク貼付意図として表3のように、それぞれ分類している。分析結果のうち「日記」における「体験日」の割合が.570、「体験記」における「体験」の割合が.781、などのように、リンク先データ種別ごとにリンク貼付意図の割合に偏りが生じていることが見られる。

4.3 考察

分析結果より、リンク先データ種別により、おおよそのリンク貼付意図が決まってくる、ということがいえる。この順方向リンク分析より、連想構造には何らかの連想契機となるものがあり、それからある対象事物が連想されるというものであり、さらにそれらの間には何らかの連想関係が存在する、という概観を得た。この概観から連想構造は以下の3要素から構成されるということがいえる。

- 連想契機：連想のきっかけとなる事物
- 連想対象：連想契機により、連想される対象事物
- 連想関係：連想契機と連想対象との間の関係

このような連想構造の概観を基に、それをさらに深く分析するため、次節に示す逆方向リンク分析を行う。

表4: 順方向リンク分析結果(一例)

リンク先データ種別	リンク貼付意図					
	体験日	関連日	体験	利用	参照	等価
日記	.570	.204	.005	.000	.000	.012
体験記	.000	.000	.781	.000	.094	.047
メモ	.007	.007	.007	.000	.523	.141
作品	.000	.000	.000	.000	.069	.701
計	.264	.095	.062	.026	.132	.133

5. 逆方向リンク分析

5.1 分析内容

逆方向リンク分析は、リンク先データに注目し、先に示した連想構造の3要素：連想契機、連想対象、連想関係をそれぞれ抽出し、それらを表5,6にしたがって分類する。それらの統計を取り、その結果から連想構造について分析、考察する。逆方向リンク分析のためのデータ分類の流れを以下に示す。

1. 対象データの種別分類
2. 参照元データ中のリンク契機語句の抽出、分類
3. 対象データとリンク契機語句との関係の抽出、分類

まず、連想対象となるデータに対して表5にしたがってデータ種別の分類を行う。つづいて、その参照元データから連想契機となるキーワードを抽出し同様に種別の分類を行う。さらに、それら連想契機と連想対象の間に見られる連想関係を抽出し、表6にしたがって分類する。

ここでは、連想契機となりうるものとして、日記中のキーワードとし、4W1D:When, Where, Who, What, Do (日付, 場所, 人, 物, 体験など)に基づいて分類されるものとする。また、連想対象となりうるものとして、テキストデータおよび静止画データとする。テキストデータには、ある日の日記データやある事物に対するメモや記録、資料、記述者の見解などが記述されている。静止画データには、記述者が撮影した写真、紙媒体のスキャンデータなどがある。それらは、同様に4W1Dに基づいて分類されるものとする。連想契機および連想対象の抽出を行いながら、逐一それぞれの種別として適合する分類項目を作成した。表5は、その4W1Dに基づく分類項目例である。同様に、連想契機と連想対象との間の関係を推測し、それを両者間の連想関係として逐一適合する分類項目を作成した。表6は、連想対象がWho:人物の場合の連想関係の分類項目例である。表5,6にしたがって分類を行い、その統計を取ることにより分析を行う。

表5: 4W1Dに基づく種別分類項目例

4W1D	上位分類	下位分類
When	絶対	日付, 一般, 特定日
Where	地域	国, 観光地, 地名
	施設	空港, オフィス, 大学
Who	人物	一般, 著者, 歌手
	属性	個人属性, 所属
	企業	一般, 出版社, 提供者
What	資料	原稿, レポート, 書類
	作品	小説, 漫画, 音楽
	書籍	雑誌, 参考書, 写真集
	物品	電子機器, 小物
	イベント	学会, 発表会, 講演会
	情報	サイト, 生活
Do	体験	日常, やりとり, 作業

表6: 連想関係の分類項目例(連想対象:Who)

連想対象	第1分類	第2分類	第3分類
Who	関係	等価関係者	等価人物, 人物関係者
		共有者	遭遇者, 同伴者
	体験	主体者	利用者, 体験者
		構成者	出演者, 参加者
	属性	所有者	所属者, 物品所有者
		著作者	執筆者, 開発者
対象	対象者	対象者	

5.2 分析結果

分析結果を表7に示す。表7は、4W1Dの各分類項目が連想契機および連想対象として用いられた割合である。割合にばらつきがあるものの、ほぼ全体に及んで割合が見られる。また、各種別における割合に偏りが見られる。例えば、「人物」に対する連想契機割合および連想対象割合がそれぞれ.168,.384であるのに対し、「書籍」に対してはそれぞれ.017,.016である。

また、連想契機と連想対象との関係を見るために、各連想契機種別の度数と各連想対象種別の度数とのクロス統計を取った。その統計結果の一例を図3,4に示す。図3は、連想契機を中心とし、それから連想される対象の割合を示したものである。ここでは、人物とイベントを連想契機とした例を示している。一方、図4は、連想対象を中心とし、それを連想する契機の割合を示したものである。ここでは、人物と体験を連想対象とした例を示している。

さらに、連想契機と連想対象との間には何らかの連想関係が存在する。その連想関係の割合の一例を表8に示す。

表 7: 連想契機キーワード種別割合 (一例)

4W1D	上位分類	連想契機割合	連想対象割合
When	絶対	.011	.043
Where	地域	.038	.001
	施設	.099	.000
Who	人物	.168	.384
	属性	.072	.120
	企業	.108	.015
What	資料	.019	.000
	作品	.070	.004
	書籍	.017	.016
	物品	.048	.001
	イベント	.105	.034
Do	体験	.172	.226

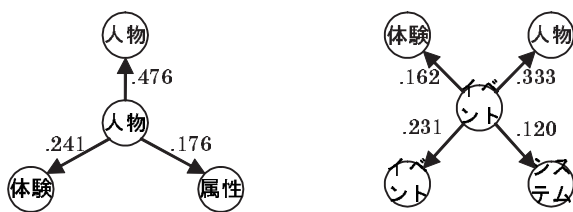


図 3: 連想契機を中心とした統計結果 (一例)

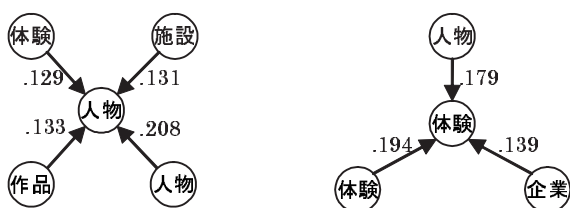


図 4: 連想対象を中心とした統計結果 (一例)

5.3 考察

分析結果より、以下のことがいえる。

- 4W1D に基づくあらゆる事物をきっかけとして連想がなされる
- 4W1D に基づくあらゆる事物が対象として連想される

また、各種別において割合の偏りが生じていることから、

- 連想契機となりやすい種別、そうでない種別とが存在する
- 連想対象となりやすい種別、そうでない種別とが存在する

ということがいえる。このことから、種別ごとに連想契機率および連想対象率が存在すると考えられ、

- 連想契機率が高いものほど、そのキーワードをきっかけとした連想想起の活性度が高い
- 連想対象率が高いものほど、そのデータが何らかのきっかけにより連想される機会が多い

ということがいえる。これらは、人により異なり、個人的主観に基づくものであるといえる。今回の分析対象データの例では、人物や体験をきっかけとし連想想起がされやすく、また人物や体験が連想されやすいと考えられる(表7参照)。

各連想契機種別の度数と各連想対象種別の度数とのクロス統計結果を示している図3,4より、この割合が連想契機と連想対象との連想強度であるといえ、この強度が強ければ、その

連想契機からその連想対象が連想されることが多いといえる。この割合もまた個人的主観に基づくものであるといえる。今回の分析対象データの例では、人物からまた人物が連想されやすく、イベントから人物やまたイベントが連想されやすい、ということがいえる。また、体験や人物からまた人物が連想されやすく、体験が連想されやすい、ということがいえる。

さらに、連想関係の割合の一例を示している表8から、連想契機と連想対象により連想関係がほぼ一意に決まるということがいえる。今回の被験者の例では、企業を契機として人物が連想された場合には、その人物がその企業に所属しているという関係がほとんどであるということがいえる。

表 8: 連想関係割合 (一例)

連想契機		連想対象		連想関係	
上位分類	下位分類	上位分類	下位分類	分類	割合
企業	一般	人物	一般	所属者	.947
		体験	やりとり	所属者	.923
作品	小説	著者	人物	著者	1.000
イベント	学会	参加者	人物	一般	.929

6. おわりに

生涯体験アルバム作成支援システムのプロトタイプ設計を検討する。提案システムでは、体験記録データ整理過程として日記記述過程を考える。日記記述を支援しながら、体験記録データを連想構造にしたがって整理する。ここで、体験記録データとしては一般的に、テキストデータや静止画データ、動画データ、音声データなど様々な種類のデータが存在すると考えられる。これらの体験記録データが連想対象となりうるものとして考えられる。また、体験記録データに加え、日記中に出現するキーワードは連想契機となりうるものとして考えられる。さらに、連想強度は連想構造を構築するために重要な要素であり、この設定のばらつきが個人的主観を表すものと考えられる。支援システムに必要な機能として、大きく以下の2点が挙げられる。

- 連想関係の自動構築および連想強度の設定
- 連想契機から連想される対象データの自動提示

システムの利用イメージを以下に示す。ユーザが一日の体験を日記に記述する際、その日や過去の体験記録データをハイパーリンクにより貼付する。システム側は、ハイパーリンクの対象データやその契機となった日記中の語句から、連想契機と連想対象との間に連想関係を自動的に構築する。さらに、ハイパーリンクが貼付される頻度などから連想強度を設定する。また、記述中の日記にある語句が出現した際に、その語句を契機として連想される体験記録データを自動的に推測し、そのデータを提示する。

今後は、ここで示した機能を具体化し、体験記録データの整理過程である日記記述を支援することにより、個人的主観に基づく体験記録データ間の連想関係の構築方法および連想強度の設定方法などを検討していく必要がある。

参考文献

[Ueoka 03] 上岡 玲子, 広田 光一, 廣瀬 通孝, ウェアラブルコンピュータによる主観的体験とその展開, ヒューマンインタフェース学会研究報告集, Vol.5, No.4, pp.65-68, 2003.
 [Misaki 04] 美崎 薫, 河野 恭之, 「記憶する住宅」~55万枚のデジタルスキャン画像の常時スライドショー・ブラウジングによる過去の蘇りの実際, 情報処理学会(編), インタラクシオン 2004, pp.129-136, 2004.