

Information Fishing: 即応的な情報検索と持続的な情報提示の統合

Information Fishing: Combining Immediate Information Retrieval and Continuous Information Recommendation

小林 正朋 五十嵐 健夫*

Summary. We introduce a *fishing* metaphor to provide a seamless transition between an instant process of information retrieval (IR) and an automatic process of collecting relevant information. Such a dynamic transition is required for facilitating exploratory IR activities for ambiguous goals, because an IR process that has an imprecise goal often involves two types of interaction: those for obtaining simple, definitive information and those for gathering various, suggestive information. Traditional query-result interfaces are efficient only in the former process, while existing information collecting systems intend only the latter process. Our system provides both definitive responses and suggestive recommendations through a virtual space within which each information unit moves autonomously as a fish and to which the user submits queries as food. Considering the result of our first user study, we improved our prototype system for reducing cognitive load in exploratory activities and discussed sharing the virtual space with other users.

1 はじめに

Rose ら [3] によれば, Web 検索の目的には (a) 特定の情報を求めるのほか, (b) 広い知見を集める, (c) 行動のための示唆を得る, といったものがある. しかし通常の検索エンジンは (a) に主眼を置いて設計されているため, (b)(c) の目的のために用いるとクエリの入力や情報の整理に関してユーザに高い負荷を強いてしまう. 一方, Goromi[9] や Memorium[11] といった「流し読み」を目的としたシステムは利用時の作業負荷が少ないものの, システムによって示唆を得たユーザが特定の情報の必要性に気付いたとしてもその検索要求に対応することができない. したがって状況に応じて通常の検索エンジンを併用しなければならないが, 頻繁な手段の切り換えはユーザの滑らかな思考活動を妨げてしまう恐れがある. (b)(c) に分類される検索目的の多くでは従来型の機械的な検索作業と総合的で曖昧な検討作業の両方が必要になると予想されるため, 我々はそれらを複合的に提供するシステムの実現を試みる.

前段に述べた目標を達成するため, 本稿では仮想的な空間と仮想的な生態系を介して情報検索を行なう機構を考える. 具体的には, 被検索文書が「魚」となって空間内を泳ぎ回り, ユーザの入力したクエリは「餌」として魚を誘う. 個々の魚は周囲の魚に作用を及ぼしながら自律的に動作するほか, ユーザからのフィードバックを受け付ける. 結果として,

魚は仲間同士で群れを形成するなどの複雑な挙動を示す. なお, 魚群の分布は仮想的な海の風景として視覚化される (図 1).

空間性の導入により, 情報を一元的な基準でソートことなく複数の情報の間に存在する複雑な関連性を表現することができるほか, 並行する検索作業をクエリの配置によって整理することができるようになる. このシステムはユーザの思考の流れに応じて複数の検索作業が重層的に発生する状況での利用を想定しているため, このような特徴は有用である.

また, 生態系の導入により「食餌と獲物」の形でクエリ 結果型の単純な検索活動を, 「移動と繁殖」の形で自動的な情報提示機構を表現することができる. このシステムは餌が投入されると主に検索エンジンとして, 時間が経過すると主に情報を流動的に提案するシステムとして働く. 両機能間の遷移は連続的である. さらに, 魚釣りメタファは検索に娯楽性を与え, 単調作業の繰り返しに陥りがちな (b)(c) の場面においてユーザを心理的に支援する.

我々のシステムは, 特に種々の創造的なプロセスにおける初期段階で有効だと考えられる. 必要な情報の固まっていない時点においては, 精密な情報よりもむしろ広範に渡る雑多な情報および新鮮な発想が求められるからである. また, 仕事の合間の休憩時間においても有効であろう. 現在でもしばしば息抜きとして Web 検索を楽しむことがあるが, クエリの入力に注力するあまり本来の仕事が疎かにしてしまうという事故が後を絶たない. しかし, 我々のシステムでは明示的な入力がなくとも結果が変化し続けるため, 仕事中はバックグラウンドでシステム

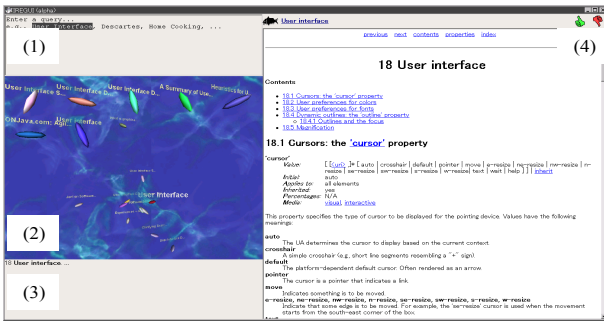


図 1. 作成した Web 検索システムのインターフェイスは (1) クエリ入力領域, (2) 魚群可視化領域, (3) プレビュー領域, および (4) フィードバックボタンから成る.

を動作させておき, 決まった休憩時間のみ結果を確認するという使い方が可能となる.

2 ユーザインターフェイス

図 1 は我々の Web 検索システムのスクリーンショットである. 画面の左側の領域は検索インターフェイスの本体であり, 右側の領域は検索された文書を閲覧するウィンドウである. 検索インターフェイスには (1) クエリ入力領域, (2) 魚群シミュレーション可視化領域, および (3) 捕獲された魚のプレビュー領域が含まれる. また, 画面右側の上部には捕獲された魚に対してフィードバックを返すボタン (4) がある. (2) は海面のイメージを表現しており, 海中には餌および魚の群れの姿が観察される. この画面では, 餌は文字列, 魚はキャプションとして文書タイトルの付随した楕円体として可視化されている. 魚群分布の変化はアニメーションによって表現される. ユーザは適当なタイミングで海面をクリックまたはドラッグすることで餌や魚を投入, 捕獲することができる. 初期入力としてはキーワード (餌) を投入できるほか, 外部から文書をドラッグ・アンド・ドロップすることにより新たな文書 (魚) を放流することも可能である. 一定の範囲内に複数の餌を投入すれば複合検索となり, 集まった魚は餌や別の魚との関連性に応じて空間内に分布する. より離れた位置に複数の餌を投入した場合はそれぞれの検索が独立並行して進行する.

図 2 はあるユーザがクエリを入力してから結果を得るまでの一連の流れである. ここでは, ジェスチャ志向の単純な操作系が採用されている. なぜなら, 通常の GUI システムが機能の高度化とともに複雑で厳密なものになりやすいのに対して, 「魚釣り」のメタファは単純性および曖昧性を要件としているからである. 大まかには, 釣り上げるジェスチャ (クリック) や網を投げるジェスチャ (ドラッグ) を介して餌や魚を扱う. また, マウスホイールの回転によ

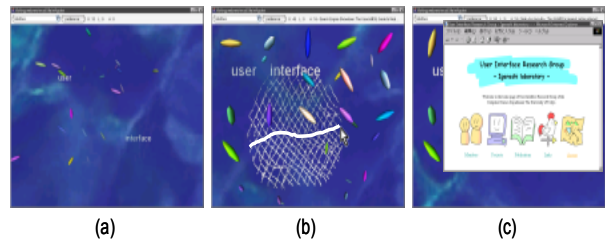


図 2. (a) クエリを入力, (b) 投網して魚を捕獲 (マウスドラッグ), (c) 魚をポイントして閲覧.

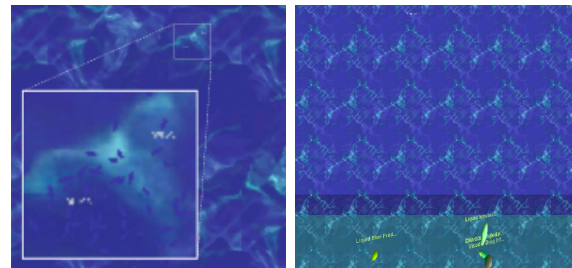


図 3. 鳥瞰表示では総合的な魚群の分布が視認され (左), 生け簀 (下部の領域) には捕獲した魚を保管できる (右).

て視点の高さが変わる (図 3・左). これにより, 生態系全体を鳥瞰したり, 特定の群れに注目したりすることができるようになる. また, 捕獲した魚を生け簀の中に保存しておく機能があり, これをブックマークとして利用することができる (図 3・右).

3 実装

作成した Web 検索システムはユーザインターフェイスのほか, 魚群シミュレーションおよび外部検索インターフェイスより構成されている (図 4). ここで, 魚群シミュレーションは次の原則にしたがう:

1. クエリは餌であり, 被検索文書は魚である.
2. 魚は餌と仲間を好み, それに近づく.
3. ユーザは文書の評価を介して対応する魚の生命力を上下させることができる.

生命力の高い魚は繁殖する. このとき, 子は親の関連文書である (詳細は後述). ユーザはクエリとして餌を投入するほか, 捕獲した魚にフィードバックを与えて再び放流することで生態系の繁殖過程を緩やかに制御することができる.

個々の魚は仮想的な三次元空間中に分布, 周囲の魚や餌の影響を受けながら自律的に動作する. 上記の原則により, 魚は仲間同士あるいは餌の付近に群れを成すこととなる. また, 子を産んだ魚の生命力は減衰係数 α で低下する. 子の生命力は親の生命力に減衰係数 β を乗じた値となる. なお, 魚と餌を共

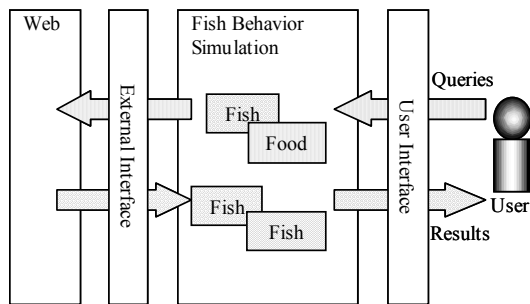


図 4. Web 検索システムは魚群シミュレーション, ユーザインタフェース, 外部検索インタフェースから成る.

通の原則のもとで動作させるため, 餌についても便宜的に同様の繁殖性や生命力を定義する. 餌の産む子はそのクエリに応じた関連文書であり, 通常の検索クエリに対する検索結果に対応する. 生命力は魚または餌の大きさとして可視化される. また, 特に関連の強い魚同士は似た色合いで表現される.

魚 F は文書本体とメタデータで構成される. このメタデータには F の仮想位置 x , 仮想速度 v , 生命力 V , 選り好みリスト S が含まれている. S の要素は他の魚または餌 F' へのポイントと選り好み係数 s とのペアである. s は実数値を持ち, 正の値は F が F' を好んでいること, 負の値は嫌っていること, 0 はいずれでもないことを表現している. 例えば図 5 に示される状況において, 魚 F_1 は F_2 を好んでいる一方, F_2 は F_1 を若干嫌っている. また, F_1 および F_2 と F_3 との間にはいかなる関連もない.

外部検索インタフェースは, Web から広く文書を収集, 魚群シミュレーションのために必要となる前述のメタデータを付加する役割を担う. 具体的には, 定期的に (あるいはユーザの動作に応じて) 既存の Web 検索エンジン呼び出し, その時点での魚群の中で特に生命力の高い 1 匹以上の魚の子または餌に近寄る魚を生成する. また, 生命力の弱い魚を消去する. 外部検索エンジンに与えるクエリ文字列は, 当該の餌の文字列あるいは魚の URL やタイトル (もしくは後述するタグ) の中から一定の割合で選択される. また, 外部検索エンジンが関連文書検索の機能を提供している場合には, 魚の繁殖に際して一定の割合でその機能も活用する. 前述の「親は子を産むために生命力を消費し, また子の生命力は親よりも低い」という原則により, 外部検索が同一のクエリを繰り返し使用することはなく, 子孫の生成が際限なく発散してしまうこともない. 現在の実装では, 外部検索エンジンとして Google Web API[14] または Yahoo! Web Service[15] を使用可能である.

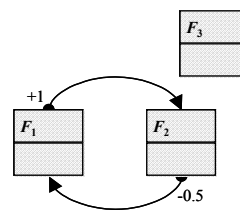


図 5. 魚 F_1 が F_2 を強く好む一方 ($s = +1$), F_2 は F_1 を若干嫌っている ($s = -0.5$). F_3 は無関係な第三者である.

4 魚群の振舞いの定義

4.1 選り好み係数

前節で導入した選り好み係数を決定する方法として, 当初の実装では以下の 3 つの基準を用いた: 魚は親を好む: 定義により個々の魚は各々の親 (関連文書に対応する魚または検索クエリに対応する餌) に強く関連しているため, 選り好み係数 $+1$ を持つ. 魚は兄弟を好む: 同一のクエリにより外部検索エンジンを介して生成された魚同士は兄弟として選り好み係数 $+0.1$ を持つ. 魚はハイパーリンクされた魚を好む: ハイパーリンクを正の関連付けであると解釈する. 文書がハイパーリンクを介して別の文書に言及しているならば, 前者に対応する魚は後者に対応する魚に対して選り好み係数 $+0.2$ を持つ.

これらの基準は文書の内容とは無関係に定義されているため, HTML 文書に限らずあらゆるタイプのリソースに対応可能である (ただし, HTML 文書以外はハイパーリンクの起点になれない). 具体的に実現され得る関連付けの種類は, 使用する外部検索エンジンで利用可能な検索クエリの種類に依存する.

4.2 魚群の挙動

我々のシステムは, 生物群のシミュレーション手法である Boids[2] を情報可視化に適用した Proctorら [4] の方法をさらに拡張して, 仮想三次元空間内における魚群の動作を制御する. すなわち, 以下の 3 つの原則に基づく (図 6): 衝突回避: 個々の魚は, 衝突を防ぐべく直近の魚から離れる. 速度調和: 個々の魚は, 近傍の魚と等速 (向き, 速度) で動作すべく速度を調整する. 追跡: 個々の魚は, 好みの魚や餌に近寄る. この動作は近傍の魚に対する選り好み係数 s の値によって決定される.

実際には, 魚 F に式 (1) の仮想加速度 A を与えることで上記の動作を実現する.

$$A = a \sum_i (x - x_i) - b(x - C) - c \sum_i s_i (x - x_i) \quad (1)$$

ただし, x は F の仮想位置, x_i および s_i はそれぞれ

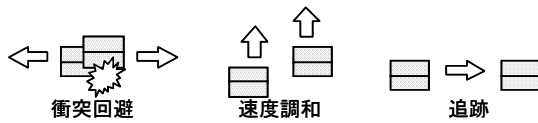


図 6. 魚は互いに衝突を避け、速度を合わせ、餌や仲間を追う。

れ F の選り好みリストの第 i 要素の指す魚または餌の仮想位置および選り好み係数, C は近傍の魚および餌の重心, a, b, c は定数である. また, 魚が極度に密集してしまう状態を避けるため, s_i を時刻 t の関数として変動させる (現在の実装では単純な正弦振動). この変動は出力に適度なゆらぎを与える効果もある.

上記の加速度を与えられた結果, 魚 F は近傍の魚と群れを成しながら餌や仲間を追跡するようになる. ここで, 選り好み係数 s_i が非対称であるということが重要である. F_i から F_j への係数と F_j から F_i への係数は異なるため, 対称かつ静的な係数行列によって「群れ」を表現する Proctor らのモデルと比較してより複雑な振舞いを示すことが保証される. これにより, 魚の「群れ」の中に新たな繋がりを見出す機会をユーザに与え, 同時に検索作業の魅力を高めるものと期待される.

5 実験

本手法 (以下, Fishing) と Google (以下, Google) を検索結果の多様性と娯楽性の観点から比較するとともに, 後の改良へ向けた知見を得るため実験を行った. 上記の観点に基づいたのは, 創造的な思考活動では情報の多様性が有用だとされており [1], また前述のように単調な繰り返し行動を心理的に支えるためには作業の面白さが必要だと考えられるからである. なお, Fishing の外部検索エンジンとしては Google を使用した.

5.1 被験者およびタスク

計算機科学を専攻する 22 歳から 25 歳までの日本人学生 8 人 (すべて Google の常用者) を被験者とした. 被験者に与えられたタスクは以下の通りである:

1. 与えられた 4 テーマ (「チューリング機械」「隠れマルコフモデル」「ギザのピラミッド」「ナスカの地上絵」) について関連文書をそれぞれ検索して知見を深める.
2. そのうち 2 テーマについては Fishing, 残りは Google を用いて作業する.
3. 制限時間は 1 テーマあたり 10 分とする.
4. 関連キーワードを紙の上にメモしながら作業を行なう.

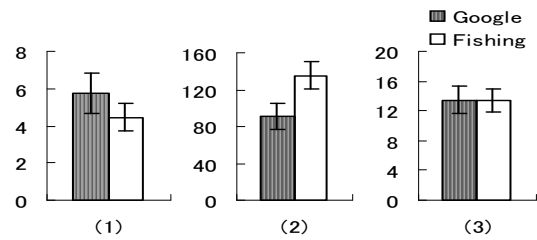


図 7. 1 テーマあたりの (1) 入力クエリ数, (2) 被検索文書数, (3) 被閲覧文書数の平均値および標準誤差.

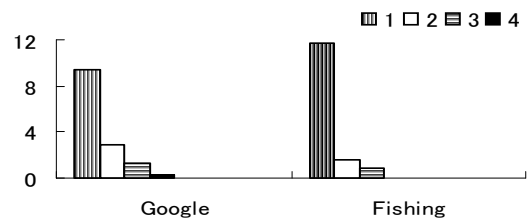


図 8. 閲覧文書の重複度分布 (1 テーマあたり).

5. 作業後に, 上記のメモを見ながらそれぞれのテーマについて得た知見をまとめる.

また, 実験後に両手法に対する印象についてアンケート形式で調査を行なった. 作業の順序は被験者間で平衡するように調整した.

5.2 結果

5.2.1 閲覧文書の多様性

図 7 は 1 テーマあたりの (1) 入力クエリ数, (2) 検索結果として画面に出力された文書の総数, (3) 実際に被験者が閲覧した文書数を示している. また, 図 8 は閲覧した文書の重複度を表している. 重複度とはある文書を閲覧した被験者の数であり, 文書 A を 8 人すべての被験者が閲覧していたとき文書 A の重複度は 8, 一人の被験者のみが閲覧したとき 1 となる.

Fishing では入力クエリ数が約 23% 減少したものの, その差は有意ではなかった ($p > .05$). 一方, 検索文書数は約 48% 増加した ($p < .001$). また, 閲覧文書数には大きな変化がなかった ($p > .05$). 閲覧文書の重複度の平均値は Fishing の 1.26 に対して Google は 1.52 であり, 約 20% 減少した ($p < .05$). 閲覧可能な文書数の増加と重複度の減少から, 少なくとも検索結果の多様性は増していると言える.

5.2.2 主観評価

図 9 は正確さ, 素早さ, 楽しさ, 難しさの各点について, 被験者が主観的に 5 段階評価を行なった結果である.

Fishing では正確さが約 1.3 ポイント, 素早さが

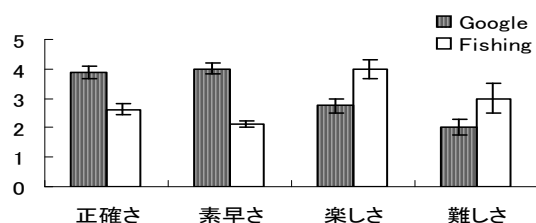


図 9. アンケート調査による主観評価の平均値および標準誤差。

約 1.9 ポイント減少する一方 (ともに $p < .001$), 楽しさは約 1.3 ポイント増加した ($p < .01$). 難しさは約 1.0 ポイント増加したが, 差は有意ではなかった ($p > .05$). また, 各被験者が実験中に作成したメモおよび実験後にまとめた知見に関して, 使用手法による明確な特徴の相違はなかった。

6 実験結果に基づく改良

6.1 可視化レベルの切り換えと即応性の実現

実験では娯楽性について想定した通りの向上が見られたものの, 正確さおよび素早さについては悪化している。

実験後に募った自由回答のコメントを総合すると, 正確さが低く評価された理由は検索自体の精度よりもむしろ文書の抜粋を一覧できない (Google ではできる) ことによる手掛かり不足にあったようだ。実際, 本手法に対して「クエリと無関係な検索結果が増える」という意見を持つ被験者はなかったが, 2人の被験者が「魚の内容が捕獲してみるまで不明」を短所として挙げた (ただし「思いもよらない結果が楽しい」を長所とした被験者も 1 人いた)。

手掛かり不足は, 魚群の可視化に際して個々の文書の内容よりも文書間の集約的な関連性の表現に注力したことによる。しかし, 実際の検索作業においては今回の実験と同様に「手掛かり不足」がユーザの思考活動を阻害してしまう可能性を否定できないため, 文書の内容を詳細に表示するモードと文書集合を抽象的に俯瞰するモードの両方を備えるべきである (具体的な表現にはユーザの認知負荷を高めるという短所もあるため [7], 持続的な思考補助を意図した我々のシステムには依然として抽象的なモードが必要である)。モードはマウスホイールによるズーム操作に連動して遷移する (図 10)。

なお, 素早さが低く評価された原因は外部検索のタイミングにあることが明らかである。実験では外部検索を固定周期で行なう単純な実装を用いたため, クエリを投入してから最初の結果が現れるまでに無視できない待機時間を要した。この欠点は, 外部検索タイミングを (a) 餌の投入時, (b) 餌または魚の肯定評価時, (c) 固定周期の複合とすることで単純

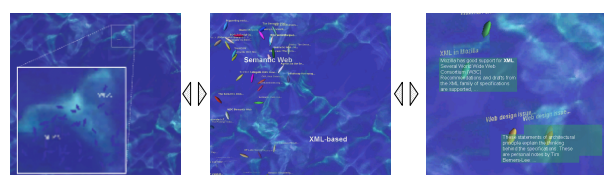


図 10. 魚影のみ視認可能な鳥瞰 (左), タイトルが表示される通常 (中), 内容の抜粋も表示される詳細 (右)。

に解決される。

以上の改良により, 我々のシステムは単純な検索作業において一般的な検索エンジンとほぼ同等の応答性と手掛かりをユーザに提供できるようになった。一方, 持続的な情報収集作業においては状況に応じて可視化の抽象度を高めることで特定の内容に捕われず情報の全体を俯瞰することができる。この性質が有効に機能する場面として, 特にサブディスプレイや focus-plus-context screen [8] のような機構によりバックグラウンド動作時にも随時ユーザの周辺視野に映るような利用形態を想定している。

6.2 タグ付けによる新たなインタラクション

捕獲した魚とのインタラクションとして, 実験で使用したシステムでは肯定ないし否定の評価を与えることのみ可能だったが, より高度な操作としてタグ付け機構を追加した。付加されたタグは (a) 外部検索のクエリとして, (b) 詳細モードで表示される手掛かり情報として, より適切な指標を提供する。また, タグを用いて 4.1 節に述べた選り好み係数の定義に第 4 の基準「タグの重複する魚同士はその重複率を選り好み係数として持つ」を追加した。なお, タグ付けはプレビュー画面で任意の語を選択あるいは入力することによって行なう。この拡張により, (a) 特定の餌を任意の魚を関連付ける, (b) ブックマークとしての生け簀を整理するなどの操作が可能となった。

6.3 マルチユーザ化への展望

さらなる発展として, 仮想生態系を複数のユーザで共有する利用形態を考える。餌や魚に加えて, ユーザも「釣り人」として空間中に分布するのである。個々のユーザは各々独自にクエリの投入や魚の捕獲, タグ付けや放流を行なうことができるものとする。

シングルユーザ環境の場合, 我々のシステムは釣りとしての情報検索プロセスと生け簀としてのブックマーク機構を実現するが, マルチユーザ環境ではそれらに加えて (a) ユーザからユーザへの情報提供や (b) ブックマークの相互参照といった仕組みが成立するはずである。(a) は Google Answers [13] やはてな [16] などで実現されているいわゆる人力検索であり, (b) は del.icio.us [12] やはてなブックマーク [17] などの提供するソーシャルブックマーク

サービス (SBS) と似た目的を持つが、我々のシステムではユーザ自身が仮想空間内に分散している。ユーザ同士の知識領域や興味の不一致から性能が低下する恐れのある人力検索や SBS において、ユーザを自然にクラスタリングすることができる空間メタファは有効に働く可能性が高い。また、検索履歴やブックマーキングは極めて私的な情報であるため実際の運用ではプライバシー保護の問題を避けられないが、それについても空間の概念を活用した簡潔なプライバシーコントロール UI の提供を計画している。

7 関連研究

空間性を備えた情報提示手法としては AIDE[6] や InfoLead[5] が挙げられる。前者は複数のキーワードを平面上に配置することで用語間の関連性を視覚化する。後者は複数の Web ページを空間中に配置することでユーザの Web 体験を演出する。これらは情報検索を意図したものではなく、また出力は流動的でないが、空間的な位置関係によって情報間の関連性を視覚化しようとしている点で我々のシステムと共通の目的を持つ。

受動性および流動性に主眼を置いた情報提示手法としては前述の Memorium や Goromi がある。また、WebFarm[10] は生物メタファを利用した先進的なブックマーク管理手段を提供する。これらはいずれも外部検索を介して情報集合を随時更新する機構を備えているが、能動的な検索作業を想定したシステムではないため、そのインタラクションは主として受動的なものに限られる。我々のシステムは、受動的な情報収集から能動的な情報検索へ、またその逆の連続的な遷移をサポートしようとするものである。

8 まとめ

本稿は、仮想空間中の仮想生態系を介して単純なクエリ 結果型の検索作業と自動的な情報提示プロセスとを混合する手法について述べた。我々のシステムは Web 検索システムの出力に空間性、流動性および確率性を付加し、相互に関連する複数の検索作業の並行や多様性のある示唆的な情報の獲得を支援する。また、このシステムは釣りや生け簀のメタファによって情報検索とブックマーキングの機能を提供する。今後、仮想空間の共有などの発展を図るとともに、情報検索と情報提示との連携に関して詳細な評価実験を行なう予定である。

謝辞

実験に協力していただいた被験者に謝意を述べる。なお、本研究の一部は IPA 平成 15 年度未踏ソフトウェア創造事業 (未踏ユース) の支援を受けた。

参考文献

- [1] A.F. Osborn, *Applied Imagination: Principles and Procedures of Creative Problem-Solving*, Scribner's, New York, 1957.
- [2] C.W. Reynolds, Flocks, Herds and Schools: A Distributed Behavioural Model, *Computer Graphics*, Volume 21, Number 4, pp.25-34, 1987.
- [3] D.E. Rose and D. Levinson, Understanding User Goals in Web Search, In *Proceedings of the 13th conference on World Wide Web*, pp.13-19, 2004.
- [4] G. Proctor and C. Winter, Information Flocking: Data Visualisation in Virtual Worlds Using Emergent Behaviours, *Lecture Notes in Computer Science*, Volume 1434, pp.168-176, 1998.
- [5] K. Arai, T. Mutou, and A. Kanai, InfoLead - A New Concept for Cruising Navigation Technology, In *Proceedings of the 2002 Symposium on Applications and the Internet*, pp.162-167, 2002.
- [6] K. Mase, Y. Sumi, and K. Nishimoto, Informal conversation environment for collaborative concept formation, *Community Computing: Collaboration over Global Information Networks*, John Wiley & Sons, 1998.
- [7] L. Dabbish and R.E. Kraut, Controlling Interruptions: Awareness Displays and Social Motivation for Coordination, In *Proceedings of the ACM Conference on Computer Supported Cooperative Work*, pp.182-191, 2004.
- [8] P. Baudisch, N. Good, and P. Stewart, Focus Plus Context Screens: Combining Display Technology with Visualization Techniques, In *Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology*, pp.31-40, 2001.
- [9] 大坪五郎, Goromi-Web 上の情報を「流し見」する方法, 第 12 回インタラクティブシステムとソフトウェアに関するワークショップ論文集, pp.107-110, 2004.
- [10] 日野洋一郎, 中島伸介, 小山 聡, 田中克己, Web-Farm: 動物メタファを用いた Web ブックマーク再利用機構, 第 14 回データ工学ワークショップ論文集, 2003.
- [11] 渡邊 恵太, 安村 通晃, Memorium: 眺めるインタフェースの提案とその試作, 第 10 回インタラクティブシステムとソフトウェアに関するワークショップ論文集, pp.99-104, 2002.
- [12] del.icio.us, <http://del.icio.us/>
- [13] Google Answers, <http://answers.google.com/>
- [14] Google Web API, <http://www.google.com/apis/>
- [15] Yahoo! Web Service, <http://developer.yahoo.net/>
- [16] はてな, <http://www.hatena.ne.jp/>
- [17] はてなブックマーク, <http://b.hatena.ne.jp/>