

解説

ネットワーク社会を支援する新しい知能メディア技術

4. ネットワークエージェントによる情報収集と流通

Network Agents for Information Gathering and Circulation by Fumio HATTORI (Computer Science Laboratory, Communication Science Laboratories, Nippon Telegraph and Telephone Corp.).

服部 文夫¹

¹ 日本電信電話(株) コミュニケーション科学研究所 計算機科学研究部

1. はじめに

インターネットおよびWWWの急速な拡大発展によって、誰もが電子化情報に簡単にアクセスできるようになった。しかし、一方で情報の海の中から、必要な情報を探しだしてくることは、きわめて困難なことになっている。ネットサーフィンという格好がよいが、現実にはいくら時間があっても足りない状況になってしまっている。

ネットワークエージェントは、人間の代理としてネットワーク(ここでは主としてインターネットを対象に考える)上の情報の収集や流通を支援するソフトウェアエージェントである。(インターネットエージェントといってもよい。)

エージェントという用語の定義についてはいろいろな視点、意見があるが¹⁾、詳しい議論については省略する。ここでは応用の視点からとらえ、実現方法のいかんにかかわらず、利用者の代理または支援をするソフトウェアという程度に考えていただきたい。また、本稿ではネットワークそのものの管理や通信そのもののサポートを行うエージェント(たとえばOracle Mobile Agentなど)については取りあげない。

ネットワークからの情報収集、流通をモデル化してみると、概ね以下のような形態に分類できると考えられる。ネットワークエージェントの応用はこのそれぞれの形態の中で考えられる(図-1参照)。

(1) 情報の所在の発見

欲しい情報がネットワーク上のどこにあるかを探す。利用者の要求条件がはっきりしているか否かでさらに2つに分けられる。

(1-1) リソース検索

利用者の要求条件がはっきりしている場合である。オンデマンドで巨大なWWW空間の探索を行うのは困難であるので、事前にリソースの所在(URL)と内容とを対応づけたインデックス(ディレクトリ)を作成することになる。このディレクトリを自動的に作成するのがWeb Robotであり、ネットワークエージェントの一種と考えてよいであろう。

(1-2) ブラウジング

一方、利用者の要求条件がはっきりしない場合には、利用者自身がウェブブラウザを使って情報空間を散策しながら、興味のある情報を発見する。ウェブブラウザをインテリジェント化して、ユーザの好みや特性を学習することにより、ユーザのブラウジングを支援する研究が行われている。

(2) 情報のアクセス

情報の所在が明らかになったら、そのサーバにあって情報の中身にアクセスすることになる。このときに、利用者の要求や好みに合致するように、動的に情報を編集して渡すようなエージェントが考えられる。知的ブラウザとは逆に、ウェブサーバをインテリジェント化するアプローチである。

(3) 情報の統合

情報が1カ所だけで得られるとは限らないので、複数のサーバにアクセスして、それらの情報を利用者の要求条件に合致するように統合することが必要になる。これを自動化あるいは支援するエージェントが考えられる。(1)ではキーワード程度の要求条件に基づいて検索を行っていたのに対し、ここでは多様なサーバの情報の中身を理解して加工する処理が要求される。

(4) 情報の流通

上記の(1)~(3)は情報の利用者側からのアクセ

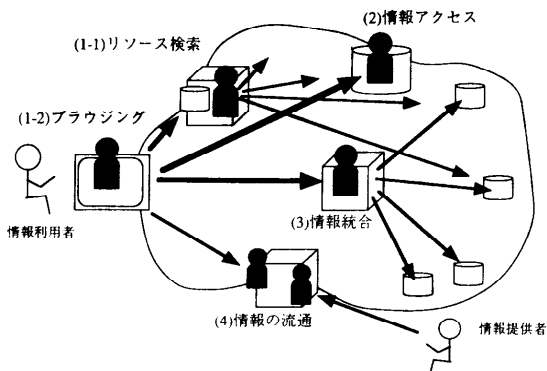


図-1 ネットワークからの情報収集, 流通

スのしやすさを向上させるという視点から研究が進められているが、逆の視点として情報を提供者から利用者へといかに情報をスムーズに流通させるかということが考えられる。実は、情報の提供者側にも情報が有効に活用されうる相手に絞って配布したいという要求が当然ある。これは現在のWWWの仕組みではなかなか難しいことではあるが、このような情報流通を支援する1つの方法として、マッチメイキングがある。これは利用者側と提供者側の出会いを支援するものである。

以下、2章ではそれぞれの形態におけるネットワークエージェントの事例を紹介する。3章では、このようなネットワークエージェントを実現するための技術について考察するとともに、今後の課題について述べる。

2. ネットワークエージェントの事例

2.1 Web Robot

WWWは巨大なハイパメディア空間であり、あらゆる分野にまたがる多種多様な情報が公開されている。一方で、WWWにおける情報発信は一方的でかつ管理者がいっさい存在しないことから、大規模なネットワークの中からユーザが欲する情報を探し出すのは非常に困難である。Web Robotは自律的にリソースを探索し、インデックスを作成することによって、ユーザからの検索要求に迅速に回答することを可能とするものである。Lycos, Alta Vistaをはじめ、現在では非常に数多くのWeb Robotが実現されている²⁾。

Web Robotの一般的な構成を図-2に示す。リソース探索ロボットは図-3のようなアルゴリズムによってWWWのハイパリンクをたどり、カ

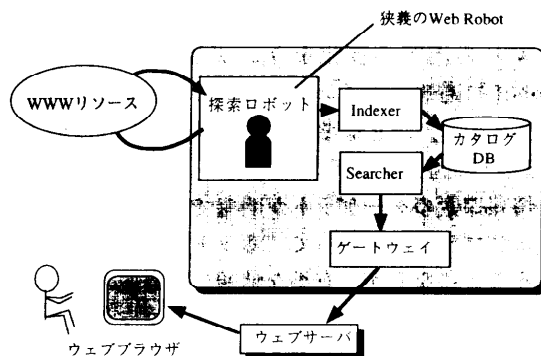


図-2 Web Robotの一般的な構成

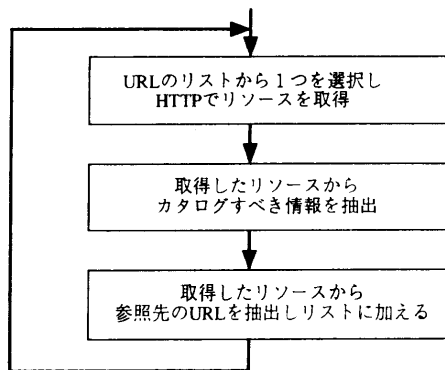


図-3 探索ロボットのアルゴリズム

タログDBに記録する。このとき、どのように探索するか(一般的には幅優先探索)、どのような情報を抽出するか(ヘッダ、タイトル、本文、...)によって、Web Robotの特徴が出てくる。日本においても数多くのWeb Robotが実現されている³⁾。たとえば、TITAN⁴⁾は多言語による検索や、探索過程をサポートするインタラクティブインタフェースを実現している点に特長がある。

最近のようにWWWが巨大化してくると、Web Robotを使ってもなかなか利用者の意図するような結果が得られないことも多い。また、ダイナミックな更新についていけないという問題もある。林はこのような問題に対応するWeb Robotの最近のトピックとして、Web Robotの統合とインデックスメンテナンスをあげている⁵⁾。前者は複数のWeb Robotに対して並行して検索を行うメタロボットに関するもので、MetaCrawler⁶⁾が著名である。WWWの内容の多様性に対して、Web Robotを専門化するというアプローチが出てきていることから(たとえば個人のホームページを探索するAhoj!⁷⁾)、1つの方向を示すものといえよう。後者は、WWWが発

信のみの一方通信であるという欠点に対して、WWW の内容の変化をウェブサーバから Web Robot に通知するインタフェースをもたせようというもので、現在標準化の検討が進められている。

2.2 知的ブラウザ

情報探索のターゲットが明確でない場合、通常ウェブブラウザを使って“ネットサーフィン”をすることになるが、有益な情報に行き着くことはなかなか難しい。そこで、ユーザのブラウジングの様子を学習し、少しでも利用者の好みや特性にあったブラウジングを支援することが考えられる。

Letizia⁸⁾は P. Maes のインタフェースエージェント⁹⁾の流れを汲むもので、ユーザのブラウジングの様子を観察して興味を学習する。学習が進むと、ユーザがアクセスしたページから、次のリンクを先読みし、ユーザの興味にあいそうなリンクを推奨することによってブラウジングを支援する。Webwatcher¹⁰⁾は正確には代理サーバ(proxy server)として動作し、Letizia と同様にユーザがアクセスしたページのリンクの中で、利用者の特性にマッチするリンクを推薦するとともに、さらに類似の情報をもつリンクを追加して表示する。Balabanovic らによる研究¹¹⁾も類似の研究である。文献 10) や 11) はゴールやアクセスしたページについての評価をユーザに入力させる点が Letizia と異なるが、逆に効率的な学習が可能となる。

2.3 知的サーバ

知的サーバの場合は前の知的ブラウザとは逆に、情報を提供するウェブサーバ側にエージェントがいる。最近、個人用にカスタマイズしたページの提供機能をもつサーバが出現してきている(たとえば文献 12))。知的サーバはこれを一歩進めて、ユーザの好みを学習し、適切な情報を提示する。一種の情報フィルタリングを行っていると考えてもよいだろう。

Firefly¹³⁾は MIT の P. Maes らの研究をもとに開発されたもので HOMR と呼ばれていたが、現在では商用サービスとして提供されている。Firefly は音楽や映画の情報提供サービスを行っているが、ユーザの評価結果に基づいてユーザの嗜好を学習し、好みに合いそうな CD/映画を紹介してくれる(図-4)。さらに、Firefly は類似する

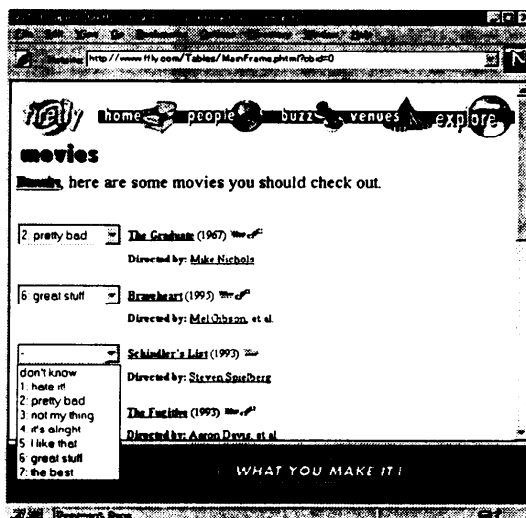


図-4 Firefly¹³⁾のアクセス例

ユーザをグループ化してコミュニティを作成し、似た人の評価結果を活用するとともに、ユーザ間で口コミ情報の共有を可能にしている。

2.4 仲介エージェント

ユーザと情報提供サーバの間になつて、情報を仲介するエージェントが考えられる。このエージェントは情報の探し方や情報の所在に関する知識をもっており、それに基づいてユーザのリクエストする情報をネットワーク中から収集してくる。ユーザはサーバの所在やアクセス方法を知る必要はなく、単に要求条件を提示すると、仲介エージェントがそれを解釈し、適切なサーバを探してそのサーバに合った方法で代行アクセスする。さらに、複数のサーバからの結果を統合して編集し、ユーザの要求にマッチした形で返却することになる。

Internet Softbot¹⁴⁾は Washington 大の Etzioni らのグループによって開発が進められているソフトウェアエージェントである。文献 14) では Internet 上の人と情報のやりとりをする際に、Softbot が ftp, telnet, mail, netfind, archie, gopher などの知識を使ってユーザの仕事を支援する例が書かれている。同じ Etzioni のグループで開発された ShopBot¹⁵⁾は、複数のオンラインショッピングの商品情報を集めて比較リストを作成し、ユーザに提供するものである。ShopBot ではショッピングサイトのアクセス方法の知識を ILA¹⁶⁾という方法で自動的に学習している。ILA

は各サイトに書かれているであろう属性(たとえば商品の名前, 価格など)とアクセス例を与えることで, 各サイトのアクセス方法を学習する。

Takahashi ら¹⁷⁾は, イエローページを基本となる知識ベースと考え, WWW化されたイエローページサーバへのアクセスを起点に, 関連する情報をいろいろなサーバから収集してくるシステムを提案している(図-5)。Knoblock らの研究¹⁸⁾では情報収集のプランを立てるところに重点が置かれている。CIG Searchbots¹⁹⁾はマルチエージェントシステムで協調的に情報収集を行う。

この種のシステムにはどの程度インテリジェントかは別として, 商用システムをめざしてサービスを開始しているものも多い(BargainFinder²⁰⁾など)。

2.5 マッチメーカ

前節までのアプローチは, 情報利用者側が情報の探すというものであったが, ここでのアプローチは視点を変えて, 情報利用者と情報提供者の出会いを支援しようというものである。

Yenta²¹⁾は同じような興味をもつ人々を相互に紹介し, グループの形成を支援する。Yentaのユーザは自分の代理としてのエージェントをもち(エージェントはユーザの個人情報からその好みを知る), エージェントは相手を探して紹介する。Cyber Yenta²²⁾はその商用バージョンである。同様に Kasbah²³⁾は物を売りたいという情報と買いたいという情報をそれぞれエージェントにもたせ, 条件の合う相手を探し合うものである。

3. ネットワークエージェントの実現技術

ネットワークエージェントの実現に必要な技術と今後の課題について述べる。なお, 情報検索や情報フィルタリングなど, 情報そのものを取り扱うための技術に関してはここでは取りあげない。たとえば文献(24), (25)などを参照されたい。Wooldridge らはエージェントの備えるべき特性として, 以下の4つをあげている²⁶⁾。

- 自律性(autonomy)
- 社会性(social ability)
- 適応性(reactivity)
- 自発性(pro-activeness)

これらの項目にあてはめて, ネットワークエージェントの実現技術について考察してみよう。

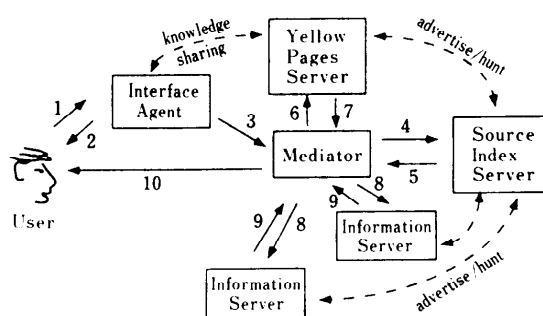


図-5 イエローページを起点とした検索(文献17)より引用)

(1) 自律性

Web Robot や仲介エージェントは, 利用者の直接の指示なしに自律的に動作する必要がある。実行時の環境に応じて処理を柔軟に変更しやすいよう, インタプリティブに実行が可能なスクリプト言語(perl, Tcl, Lisp など)で実装されることが多い。また, リモートサイトでの実行のための言語として Java, Telescript²⁷⁾などが出現している。将来的には, エージェント自身の意図などを含んだ, より高度の自律性を実現するために, Agent Oriented Programming²⁸⁾のような内部モデルをもつ言語が必要になってくると予想される。

(2) 社会性

仲介エージェントは情報の非均質性に関する問題を解決する必要がある。すなわち, エージェントはさまざまなサーバのコンテンツのアクセス方法やそのセマンティクスを知る必要がある。現状ではこれを知識として与える知識指向の手法が主流であるが, 将来的には Agent Communication Language²⁹⁾などの最低限のレベルの標準化と, セマンティクスの理解のベースとなるオントロジ(共通概念知識体系)の構築や流通が必要となろう。また, ShopBot のエージェント自身にこれらの知識を学習させようというアプローチ¹⁶⁾も興味深い。このような社会性を実現する技術は, マッチメーカなどのほかのエージェントでも, サービス範囲が広がってくると必須になってくると思われる。

(3) 適応性

知的ブラウザや知的サーバでは, 利用者の好みや特性に適応することが本質的な課題である。学習技術としては, 少ない事例から効率的に学習する必要があるので難しい面も多いが, 一方で情報

アクセスというドメインに限定されているので、うまく知識を与えてやれば有効な学習が可能となる。また、情報やユーザの特性を表現するモデル化の方法がかなり重要となってくる。Web Robot の専門化にみられるように、コンテンツに応じてさらにドメインを細分化するのも 1 つの方向であろう。

(4) 自発性

Web Robot が自発的にリソース情報を収集しているのは当然として、マッチメーカーなどもエージェントが相手を自発的に探してくることが必要になる。しかし、効率の面や、一時 Web Robot で問題になったようなトラフィックの集中によるネットワーク全体に与える影響の問題など、検討すべき課題は多い。エージェントシステムの社会システムとしての解析など、新しいアプローチが必要になると考えられる。

4. おわりに

インターネットにおける情報の収集や流通を、情報の発見、情報のアクセス、情報の統合、情報の流通の 4 つの局面にモデル化し、それぞれにおけるエージェントの応用を、事例を中心にみてきた。また、これらのエージェントの構築にあたっての技術的な課題を考察した。

ネットワークへのエージェントの応用はまさにホットな分野であり、まだまだ多くの応用が考えられる^{*}。本稿では情報収集、流通以外の応用、たとえばメールの配送、電子商取引などへの応用についてはふれなかったが、これらを含め、今後のサイバースサイエティの実現に向けて、エージェント技術の果たすべき役割は大きいものと考えられる。

エージェント技術はまだ未成熟であり、ここで紹介してきた事例についても、名前だけが先走っている面もある。しかし、これらの応用を通して技術的な要求条件が明確化され、真に役に立つ技術が開発されていくことが、エージェント技術の健全な発達につながるのではないかと考える。

^{*} 本稿での内容を含め、エージェントの応用に関するリソースの探索の出発点としては、文献 30)～32)などを推奨する。

参 考 文 献

- 1) 石田：エージェントを考える，人工知能学会誌，10(5)，pp.663-667 (Sep.1995).
- 2) The Web Robots Database, <http://info.webcrawler.com/mak/projects/robots/active.html>
- 3) Japanese Search Engines, <http://www.atrium.com/search.html>
- 4) 林，菊井，鷺崎，砂場：WWW 情報空間における Resource Discovery と Navigation 支援，信学技報，AI95-31 (Sep.1995).
- 5) 林：探索ロボットに基づく WWW サーチエンジン，1996 年電子情報通信学会情報・システムソサイエティ大会講演論文集，pp.580-581 (1996).
- 6) <http://metacrawler.cs.washington.edu:8080/>
- 7) Shakes, J., Langheinrich, M. and Etzioni, O. : Ahoy!: The Homepage Finder, <http://ahoy.cs.washington.edu:6060/>
- 8) Lieberman, H. : An Agent that Assists Web Browsing, Proc. of IJCAI-95, pp.924-929 (1995), <http://lieber.www.media.mit.edu/people/lieber/Lieberary/Letizia/Letizia-Intro.html>
- 9) Maes, P. : Agents that Reduce Work and Information Overload, CACM, 37(7), pp.30-40 (July 1994).
- 10) Armstrong, R., Freitag, D., Joachims, T. and Mitchell, T. : Webwatcher: A Learning Apprentice for the World Wide Web, Proc. of the Symposium on Information Gathering from Heterogeneous, Distributed Environments, AAAI Press (1995), <http://webwatcher.learning.cs.cmu.edu:8080/project-home.html>
- 11) Balabanovic, M. and Shoham, Y.: Learning Information Retrieval Agents: Experiments with Automated Web Browsing, Proc. of the Symposium on Information Gathering from Heterogeneous, Distributed Environments, AAAI Press (1995).
- 12) <http://www.netscape.com/custom/index.html>
- 13) <http://www.ffly.com/>
- 14) Etzioni, O. and Weld, D. : A Softbot-Based Interface to the Internet, CACM, 37(7), pp.72-76 (July 1994).
- 15) Doorenbos, B., Etzioni, O. and Weld, D.: ShopBot, <http://www.cs.washington.edu/research/shopbot/>
- 16) Perkowitz, M. and Etzioni, O.: Category Translation: Learning to Understand Information on the Internet, Proc. IJCAI-95, pp.930-936 (1995).
- 17) Takahashi, K., Nishibe, Y., Morihara, I. and Hattori, F. : Collecting Shop and Service Information with Software Agents, Proc. of PAAM96, pp.587-596 (1996).
- 18) Knoblock, C.A. and Arens, Y. : An Architecture for Information Retrieval Agents, AAAI Spring Symposium on Software Agents (1994).
- 19) Decker, K., Lesser, V., NagendraPrasad, M.V.,

- and Wagner, T.: An Architecture for Multi-agent Cooperative Information Gathering, Proc. CIKM Workshop on Intelligent Information Agents (1995),
<http://dis.cs.umass.edu/research/searchbots.html>
- 20) BargainFinder, <http://bf.cstar.ac.com/bf/>
- 21) Foner, L. and Maes, P. : Yenta, <http://foner.www.media.mit.edu/people/foner/Yenta>
- 22) Cyber Yenta : <http://boston.com/thefix/mind/cyber/fmccyb.html>
- 23) Chavez, A. and Maes, P.: Kasbah: An Agent Marketplace for Buying and Selling Goods, Proc. of PAAM96, pp.75-90 (1996).
- 24) 住田, 三池: 知的情報検索の動向, 人工知能学会誌, 11(1), pp.10-16 (Jan.1996).
- 25) 森田, 速水: 情報フィルタリングシステム, 情報処理, 37(8), pp.751-757 (Aug.1996).
- 26) Wooldridge, M. and Jennings, N.R.: Agent Theories, Architectures, and Languages: A Survey, in Wooldridge and Jennings (Eds.) Intelligent Agents, Lecture Notes in Artificial Intelligence 890, Springer-Verlag (1995).
- 27) White, J.E.: Telescript Technology: The Foundation for the Electronic Marketplace, White paper, General Magic Inc. (1994), <http://www.genmagic.com/Telescript/>
- 28) Shoham, Y.: Agent-Oriented Programming, Artificial Intelligence, 60(1), pp.51-92 (1993).
- 29) Finin, T., McKay, D., Fritzson, R. and McEntire, R.: KQML as an Agent Communication Language, Proc. CIKM '94, ACM Press (1994).
- 30) <http://www.cs.umbc.edu/agents/>
- 31) <http://www.cl.cam.ac.uk/users/rwab1/agents.html>
- 32) <http://ai.iit.nrc.ca/subjects/Agents.html>
 (平成8年11月5日受付)



服部 文夫 (正会員)

1950年生。1973年早稲田大学工学部電子通信学科卒業。1975年同大学院工学研究科電気工学専攻修士課程修了。同年日本電信電話(株)入社。現在、NTTコミュニケーション科学研究所計算機科学研究部長。主に、データベース管理システム、エキスパートシステム、エージェントシステムなどの研究開発に従事。電子情報通信学会、人工知能学会、AAAI各会員。