

F_038

コンテンツエージェントモデルに基づく Web ページの自動編成システムの試作 Implementing an Automatic Web Page Composition System based on a Content Agent Model

浅見昌平† 大園忠親† 伊藤孝行† 新谷虎松†
Shouhei Asami, Tadachika Ozono, Takayuki Ito, Toramatsu Shintani

1 はじめに

閲覧者のニーズに則した Web ページを自動的に編成するためには、Web ページへのアクセスの解析が必要である。アクセスログの収集は、一般的には、困難になる場合が多い。例えば、HTML の動的生成を行う Web ページにおいて、アクセスログの解析は困難である。また、同一 URL で処理が進むシステムの場合、閲覧者の動きがアクセスログに記録されないため、どのページが重要であるか解析できない。この問題を解決するためには、システムに閲覧者のアクセス解析機能を持たせる必要がある。

例えば、Web Watcher[1] は閲覧時の振る舞いを観察することによって、その閲覧者の嗜好を推定する。しかし、このようなシステムでは、ニュース記事や広告が変化して、閲覧者の行動の意味が変化した場合には対応できない。

本研究では、Web ページとサーバ間で動作するマルチエージェントの働きによって、閲覧者のプロフィールを作成する。Web コンテンツ自体に意味を持たせ、その Web コンテンツへの閲覧者の行動を観察する。閲覧者にとって有用でない Web コンテンツは、別の Web コンテンツとリアルタイムに入れ替わり、Web ページは閲覧者のニーズに応えるよう変化する。閲覧者のプロフィールはその編成過程で作られる。Web ページは閲覧者の行動に合わせて変化する結果、個別化していく。このマルチエージェント環境をコンテンツエージェントモデルと定義し、閲覧者のニーズ分析が可能な Web ページ自動編成システムの基盤として提案する。

2 コンテンツエージェントモデル

2.1 モデルの構成

コンテンツエージェントモデル (Content Agent Model) は、Web ブラウザと Web サーバを動作環境とするマルチエージェントシステムである。図 1 にコンテンツエージェントモデルの概要を示す。ここでは、HTML 要素をコンテンツ、コンテンツにエージェントとしての手続きを付加したものをコンテンツエージェント (Content Agent) と定義する。全てのコンテンツエージェント情報を格納する場所をエージェントデータベース (Agent DataBase) と定義し、Web サーバ上に設置する。また、Web サーバにある知識を格納するデータ構造を黑板と定義する。

2.2 コンテンツエージェントの機能

コンテンツエージェントはそれぞれ Web ページに表示するコンテンツを持つ。自分の持つコンテンツが閲覧者にとって必要かどうか判断するために、各コンテンツエージェントは if-then 形式の状態-行動ルールを持つ。if 部にあたる状態は黑板にある知識とマッチングを行う。

知識とは、閲覧者の行動を観測した監視エージェントが行動結果を記録したもの、閲覧者の訪問履歴、等の閲覧者の情報である。

エージェントデータベースには、全てのコンテンツエージェントの情報が XML 形式で格納されている。このデータベースからエージェントのルール集合を引き出し、ルールベースシステムと同様に推論を行う。

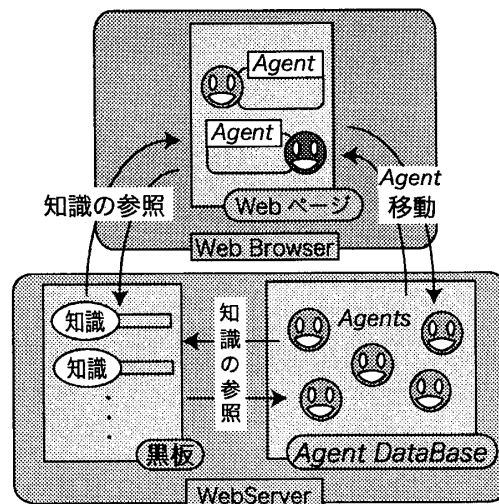


図 1: コンテンツエージェントモデルの概要

表 1: コンテンツエージェントの持つデータと機能

データまたは機能	サーバ (本体)	ページ (ミラー)
名前	○	○
コンテンツ	○	○
ルール	○	×
推論機構	○	×
フィードバック	×	○

コンテンツエージェントはページ上で動作するために、エージェントの名前とコンテンツを継承したエージェントを、ページへ送る。送られるエージェントをミラーと呼ぶ。また、ミラーは Web ページで起こったイベント情報を本体へフィードバックする機能を持つ。サーバとページ上におけるコンテンツエージェントのデータ構造と機能を表 1 にまとめる。

エージェントが推論した結果、発火したルールの行動として次のものが挙げられる。

- 黑板への知識の書き込み
- 閲覧者にとってのコンテンツの価値を変更
- サーバから Web ページへコンテンツエージェントのミラーを送信
- Web ページのミラーを削除

これらの行動は閲覧者が Web ページを閲覧している最中に行われる。

3 Web ページ自動編成機構

本章では、コンテンツエージェントモデルを基盤にした、Web ページ自動編成システムについて述べる。Web ページの自動編成とは、コンテンツを閲覧者のニーズに合わせて、Web ページをリアルタイムに変化させる機構である。ここでは、Web ページは閲覧者にとって最適なコンテンツが表示されるように自動で自身のコンテンツを編成する。

† 名古屋工業大学 知能情報システム学科

3.1 システムの機能

本システムは、閲覧者の行動や情報に合わせて、Web ページの構成を自動的に変更させる機能を持つ。Web ページにアクセスした閲覧者は、コンテンツのクリック、ページのスクロール、コメントの投稿などの行動を取る。この行動結果に基づき、コンテンツエージェントが推論を行って、Web ページを編成する。

本システムでは式 (1) のように、閲覧者の行動に価値を定義することが可能である。

$$V = v(a, c, t) \quad (1)$$

V を閲覧者の行動の価値とすると、関数 v は閲覧者の行動 a 、コンテンツ c 、及び時間 t から行動の価値を返す関数である。この式により、閲覧者の行動結果からコンテンツエージェントは報酬を受け取り、推論に反映させる。

コンテンツエージェントモデルが Web ページの個別化において基盤となる理由を述べる。まず Web ページを構成するコンテンツ単位で、表示、非表示のルールを適用することが可能であることが挙げられる。これは閲覧者の個別ページの管理が、コンテンツの選択情報のみで出来るというメリットがある。ページ全体の情報を保持しなくても、コンテンツの情報とプロフィール情報のみで Web ページを構成できる。

また、コンテンツエージェントのルールで閲覧者のプロフィールが作成できるという点でメリットがある。閲覧者の興味があると判断されたコンテンツのカテゴリを分類すれば、その閲覧者がどの分野に興味をもっているのかが推測できる。興味を持っていると判断する度合いもルールの中で設定することができるため、閲覧者の興味を反映した個別ページの構成は細かく決定できる。

4 ニュースサイトの構築への応用

自動編成されるコンテンツとして、ニュース関連の RSS を集めたニュースサイトを構築した。図 2 は本システムで構築した、自動編成するニュースサイトである。記事は、タイトルと要約、全文表示を要求するボタン、記事への評価を選択するボタンで構成される。

サーバのデータベースには、ニュース記事のコンテンツが格納されている。ニュース記事は予めジャンルごとに分類されていて、XML のタグで識別できる。

閲覧者の行動の価値は式 (2) を用いて求める。これは、式 (1) を閲覧者のニュース記事への行動に対して定義したものである。

$$V = \begin{cases} c & \text{全文表示, 良評価のクリック} \\ -1 & \text{悪評価のクリック, 時間 } t \text{ の経過} \end{cases} \quad (2)$$

価値 V は、閲覧者が行動する、または時間 t が経過する度に、コンテンツエージェントへ報酬として渡される。式 (2) において c は、コンテンツごとに定義されているコンテンツの価値である。コンテンツが全文表示、良評価のクリックを受けたときに、そのコンテンツの価値 c だけそれを持つコンテンツエージェントへ渡される。

本ニュースサイトにおいて、コンテンツの主なルールでは、コンテンツエージェントが持つ報酬を条件とするルールを構築する。例えば、次のようなルールで Web ページの自動編集を行う。

- 一定量の報酬があるならば、そのコンテンツエージェントと同じジャンルを持つコンテンツエージェントをページに送る。
- 報酬が負の値になったならば、そのコンテンツエージェントをページから削除する。

Sports NEWS

MLB = イチロー1安打、井口が左足首ねんざ

米大リーグ (MLB) は3日、各地で14試合を行い、マリナーズのイチロー外野手が本拠地で行われたエンゼルス戦に1番・右翼で先発出場し、4打数1安打を記録した。城島健司捕手は7番・捕手で先発し、同じく4打数1安打。マリナーズは1-7で敗れた。

全文を読む
良評価 悪評価

NHL = ブロンガー、マイティダックスへ移籍

北米アイスホッケーリーグ (NHL) のオイラーズは3日、マイティダックスとの交換トレードで、ディフェンスマンのクリス・ブロンガー (31) を放出した。

全文を読む
良評価 悪評価

図 2: 自動編成するニュースサイト

5 まとめ

本システムの関連研究として Adaptive Web Sites[2] と、Active Web[3] が挙げられる。Adaptive Web Sites は、Web サーバのアクセスログから利用者のアクセスパターンを解析し、インデックスページを再構成する。また、Active Web はページ制作者が閲覧者に対するルールを設定することで、条件ごとに Web ページを個別化することに成功している。どちらの研究も、Web ページの個別化によって閲覧者のニーズに応えるページを生成している。しかし訪問するまでの情報で解析されたニーズは、個別化されたページ内で得られるとは限らない。分析すべきは、個別化されたページ内でどこにニーズがあるか、である。本研究は、閲覧者のニーズに応える個別ページを作るために、Web ページの動的編成という手法を用いる。

本稿では、コンテンツエージェントモデルに基づく Web ページ自動編成システムについて述べた。コンテンツエージェントモデルは、動的に生成される Web ページにおいて、閲覧者のニーズを分析可能な枠組みとなる。Web ページにおいて、コンテンツ単位で閲覧者の行動を調べるため、ニーズの高いジャンルの分類が容易にできる。閲覧者の行動結果によって編成されるページは、その閲覧者のニーズに応えられるコンテンツの構成になる。

今後の課題として、閲覧者の興味分析を適切に行えるように、ルールを構築する手法を実現する必要がある。閲覧者の行動結果の検証や、コンテンツのカテゴリ分類など、ルールの条件を整備する仕組みが必要となる。

参考文献

- [1] T. Joachims, D. Freitag, and T.M. Mitchell, "Web Watcher: A tour guide for the World Wide Web," Proc. 15th International Joint Conference on Artificial Intelligence (IJCAI'97), pp.770-77, 1997.
- [2] Mike Perkowitz, Oren Etzioni: "Towards adaptive Web Sites: Conceptual framework and case study," Artificial Intelligence 118 (2000)245-275
- [3] 清光英成, 竹内淳記, 田中克己: "Active Web - アクティブルールの XML 表現に基づく Web コンテンツの個別化とアクセス管理," 情報処理学会 Vol.2000, No.69, pp.383-390