

3K-2

ユーザ主導協働開発環境における Web アプリスケルトン自動生成手法の提案

本田 創大† 米田 多江† 佐々木 淳† 船生 豊†
 †岩手県立大学ソフトウェア情報学部

1 はじめに

昨今、情報化社会の拡大に伴い、エンドユーザ自身による情報システム構築が注目されている [1]。システム利用者であるエンドユーザ自身が情報システムを構築できれば、要求を満たした情報システムを迅速に手に入れられる可能性がある。しかし現状では、エンドユーザは情報システムの開発に関して十分な知識を持っていない、エンドユーザ自身で情報システムを開発できる環境が十分に整っていない等の問題から、エンドユーザのみでの情報システム開発は技術的に困難である。

我々は、近年進展が著しい Web アプリケーションを対象としたエンドユーザとシステム開発者の実用的な協働開発環境の構築を目指している。本研究では、ユーザの基本スキルを活用した、新しいユーザ主導協働開発環境を提案し、その核となる Web アプリケーションスケルトン生成手法を提案する。また、その手法に基づいてプロトタイプシステムを構築し、動作確認を行った。

2 既存研究

三井らによって構築された SE とエンドユーザの協働開発環境「姫神」[2]では、エンドユーザが自然言語で定義した画面と画面遷移、処理手続き等を基に SE がプログラムを記述していくという開発スタイルを採用している。これによりエンドユーザの開発への積極的関与が可能となっている。

しかしながら、SE の負荷が依然として大きい、SE とエンドユーザ間のコミュニケーションのサポートがなされていない、システムの全体像が把握しにくい等の問題が存在している。

3 協働開発環境

我々が提案する Web アプリケーション開発におけるエンドユーザとシステム開発者の協働開発イメージを図 1 に示す。ここでは、(1) エンドユーザが HTML でシステムの画面内容や画面遷移の一部を作成し、(2) その入力からシステムが Web アプリスケルトン (Web アプリケーションにおける画面と画面遷移) を自動生成、(3) スケルトンにシステム開発者がロジックを追加する、(4) エンドユーザは必要に応じて画面内容の編集、類似

ページの追加などを行う、(5) エンドユーザ及びシステム開発者は適宜要求追加、確認等を行う。これにより、システム開発者の負荷軽減、要求伝達の効率化等の効果が期待される。本モデルにおける開発環境には以下の機能が必要となる。

- Web アプリスケルトン生成機能: エンドユーザが作成した HTML ファイル群から、Web アプリスケルトンを自動生成する
- Web アプリスケルトン編集機能: 生成された Web アプリスケルトンをエンドユーザ、システム開発者双方が編集する
- 類似ページ追加機能: 生成されたテンプレートを用い、類似するページを新規に追加する
- コミュニケーション機能: エンドユーザ、システム開発者間での要求伝達等のコミュニケーションをサポートする

本稿では Web アプリスケルトン生成機能についてのみ述べる。

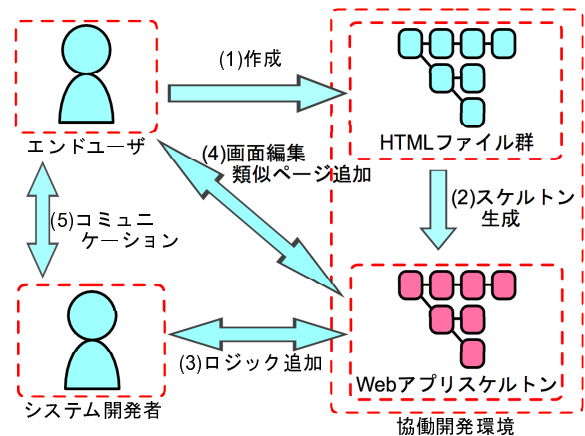


図 1: 協働開発イメージ

4 Web アプリスケルトン生成手法

Web アプリスケルトンは以下の処理によって生成される。

4.1 HTML ファイル構造解析

入力された HTML ファイル群の親子構造を解析する。ページ内のハイパーリンクを正規表現を用いて抽出し、深さ優先探索でページ間の関係を取得する。

4.2 ページ類似度判定

各ページ間の類似度を判定する。類似度の導出には tf-idf 法とベクトル空間モデルを用いた [3]。以下に処理の概要を示す。

A Proposal of Web-appli-skeleton Generating Method in User-initiative Collaborative Development Environment
 Souta HONDA†, Tae YONEDA†, Jun SASAKI†, Yutaka FUNYU†
 † Faculty of Software and Information Science, Iwate Prefectural University,

あるタグ t_i のページ D_j における重要度は、出現頻度 $freq(i, j)$ を用いて式 (1) で表される。また全文書 N におけるタグ t_i の重要度は、タグ t_i が出現する文書数 $Dfreq(i)$ を用いて式 (2) で表される。また式 (3) が最終的なタグの重要度になる。

$$tf_{ij} = \frac{\log(freq(i, j) + 1)}{\log(Terms\ in\ the\ D_j)} \quad (1)$$

$$idf_i = \log \frac{N}{Dfreq(i)} + 1 \quad (2)$$

$$w_j^i = tf_{ij} \times idf_i \quad (3)$$

ページ D_x は各タグを次元に割り当てた多次元空間におけるベクトルとして表現し、式 (4) のように定義すると、ページ D_x とページ D_y 間の類似度は式 (5) で表される。

$$D_x = (w_x^1, w_x^2, \dots, w_x^m) \quad (4)$$

$$sim(D_x, D_y) = \frac{\sum_{k=1}^N w_x^k w_y^k}{\sqrt{\sum_{k=1}^N (w_x^k)^2} \times \sqrt{\sum_{k=1}^N (w_y^k)^2}} \quad (5)$$

4.3 ページのグループ化

類似度の高いページについて、次のアルゴリズムによってグループ化する。

1. 全ページとの類似度の平均が最大のページを親ページとする
2. 親ページとの類似度が最大のページをグループ化する
3. そのグループと類似度が高いページがあればグループを拡張する。無ければそのグループを除外して1に戻る。全ページがグループ化されていれば終了

4.4 テンプレート生成

ページグループ毎にページテンプレートを生成する。グループ化されたページの中の1つをテンプレートとして採用する。

5 評価実験

5.1 実験方法

提案手法に基づき Web アプリスケルトン生成システムを構築し、手法の有効性を確認するための評価実験を行った。実験の手順は次の通りである。

1. 全ページとの類似度の平均が最大のページを親ページとする
2. 親ページとの類似度が最大のページをグループ化する

3. そのグループと類似度が高いページがあればグループを拡張する。無ければそのグループを除外して1に戻る。全ページがグループ化されていれば終了。

対象 Web アプリケーションは以下の2つとし、類似度判定に用いる閾値は初期値として0.8にした。

(対象1) 自作の音楽CD管理システム

- ページ数：15, タグ数：213, タグ種類数：19

(対象2) 大手オークションシステム

- ページ数：16, タグ数：710, タグ種類数：152

5.2 実験結果および考察

実験の結果、想定していた類似度に比べ、対象2では適切にグループ化が行われていたが、対象1では一部グループ化が適切に行われなかった箇所があった。これは tf-idf 法では単語の出現頻度を基に類似度を算出しているため、タグの量と種類が大きく影響したと考えられる。従って、読み込む HTML ファイル群の規模によって類似度判定の閾値を変化させれば、より適切なスケルトン生成が可能になると考えられる。そこで、類似度判定の閾値を対象1では0.92として再度実験を行ったところ、類似度判定、グループ分割ともにほぼ想定通りの結果が得られた。

6 まとめ

本研究では、エンドユーザが作成した HTML ファイル群から Web アプリケーションスケルトンを自動生成する手法を提案し、そのプロトタイプシステムを構築した。また、システムの動作確認と実験評価を行い、システムの有効性を確認した。

今後は、システムの規模による閾値の自動算出や類似度判定の精度向上等、実用化に向けての取り組みを進めていく。また、構築した Web アプリスケルトン生成システムを活用し、実用的な協働開発環境の構築を行っていく。

参考文献

- [1] Chusho, et.al, "A Form-based Approach for Web services by Enduser-Initiative Application Development", SAINT2002 Workshop, pp.169-203(2002)
- [2] 三井ら, "エンドユーザの積極的関与を容易にする Web アプリケーション設計モデルの一考察", 情報処理学会第67回全国大会 (2005)
- [3] 徳永, 情報検索と言語処理, 東京大学出版会 (1999)