

Java を対象としたソフトウェア部品検索システム SPARS-J の実験的評価

梅森 文彰[†] 西 秀雄[†] 横森 励士[†] 山本 哲男^{††} 松下 誠[†]
楠本 真二[†] 井上 克郎[†]

[†] 大阪大学大学院情報科学研究科 〒560-8531 大阪府豊中市待兼山町1-3

^{††} 独立行政法人科学技術振興機構 〒332-0012 埼玉県川口市本町4-1-8

E-mail: †{umemori,h-nisi,yokomori,t-yamamt,matusita,kusumoto,inoue}@ist.osaka-u.ac.jp

あらまし ソフトウェア部品検索システムの実現により、ソフトウェアの再利用や、対象ソフトウェアの理解を容易に行うことができると考えられる。本論文では、我々の研究グループが構築した Java ソフトウェア部品検索システム SPARS-J について実験的評価を行う。実験においては、部品の取得を想定した検索について、一般的な検索システムとの適合率に関する比較を行うとともに、SPARS-J の順位付け性能に関する評価を行う。さらに、企業のソフトウェア開発現場に適用することで本システムの有効性を検証する。

キーワード ソフトウェア部品、ソフトウェア検索、ソフトウェア再利用、Java、実験的評価

An Experimental evaluation of Java Software Component Retrieval System SPARS-J

Fumiaki UMEMORI[†], Hideo NISHI[†], Reishi YOKOMORI[†], Tetsuo YAMAMOTO^{††}, Makoto MATSUSHITA[†], Shinji KUSUMOTO[†], and Katsuro INOUE[†]

[†] Graduate School of Information Science and Technology, Osaka University
1-3, Machikaneyama-cho, Toyonaka-shi, Osaka 560-8531, Japan

^{††} Japan Science and Technology Agency

4-1-8, Honmachi, Kawaguchi-shi, Saitama 332-8531, Japan

E-mail: †{umemori,h-nisi,yokomori,t-yamamt,matusita,kusumoto,inoue}@ist.osaka-u.ac.jp

Abstract Software component retrieval system facilitates software reuse and software understanding. In this paper, we evaluate the effectiveness of SPARS-J (Software Product Archive, analysis and Retrieval System for Java). On the experiments, we compare SPARS-J with other retrieval systems, and we evaluate about each ranking method realized in SPARS-J. In addition, we apply it to the actual software development environment.

Key words Software component, Software retrieve, Software reuse, Java, Experimental evaluation

1. はじめに

近年のソフトウェアの大規模化と複雑化に伴い、開発されたソフトウェアの有効な二次利用が極めて重要になりつつある。ここで言うソフトウェアの二次利用とは、開発されたソフトウェアに関する直接的な活動や、開発されたソフトウェアの稼働と対比させたもので、ソフトウェアの再利用や生成されたソフトウェアの管理など、開発作業の効率化を目指す試みを指す。

既存のソフトウェアの有効的な活用方法として、ソフトウェアの部品化があげられる。ソフトウェアの部品化を行い、個々の部品の関連性を把握することで、適切な部品を組み合わせ

新たなソフトウェアを開発(ソフトウェア部品の再利用)したり、保守作業においてソフトウェアを把握(ソフトウェアの理解)したりすることが容易となる。実際にオブジェクト指向設計開発はソフトウェアを部品化する手法として注目を集めており、再利用に関する研究も盛んである[2],[5]。

我々の研究グループでは、Java ソースコードの集合を対象としたソフトウェア部品検索システム **SPARS-J** (*Software Product Archive, analysis and Retrieval System for Java*) に関する研究を行っている[10],[13]。SPARS-J では、ソースコード内のクラスをソフトウェア部品とみなして、利用関係や類似といったソースコード特有の特性を考慮しながら大量のソフト

ウェア部品に対する検索機能を提供する。さらに、検索結果表示の際にソフトウェア部品に関する詳細な情報を併せて提供することで、検索キーに関連したソースコードの効率的な検索機能を実現している。このシステムを用いることで、ソフトウェア部品の知識が無い開発者も有用なソフトウェア部品やそれに付随する有益な情報を平易に入手できる。しかし、SPARS-Jに関する評価は行われておらず、システムの有効性の評価が必要である。

本研究では、SPARS-Jについて、ソフトウェア部品検索システムとしての実験的評価を行う。実験ではまず、多くの検索目的で用いられること多い Web ページ検索エンジン Google [3] や、信頼性の高い全文検索システム Namazu [9] について SPARS-J との比較実験を行う。比較においては、ある目的に応じた部品の取得を想定した検索を行い、上位で取得された文書の適合率をもとに評価を行う。

次に、SPARS-Jの順位付け性能に関する評価を行う。SPARS-Jでは、順位付けを行うための手法として、利用関係に基づいた評価、各文書における検索キーの出現頻度に基づく評価を用いている。各順位付け手法を評価するために、ユーザが最適だと思う順位と実際に得られた順位との差を求めることで、各評価付け手法の特徴や、最適な評価手法を調査する。

最後に、SPARS-Jを企業における実際のソフトウェア開発現場に適用し、被験者に対して実施したアンケート結果に関する考察を行う。SPARS-Jが検索システムとしてだけでなく、部品の管理や理解を支援するシステムとしても有効であることを確認する。

以降、2.節でSPARS-Jについて簡単に説明する。3.節では、各評価実験の概要および結果を紹介し、4.節で実験的評価を通しての考察を行う。最後に、5.節でまとめと今後の課題について述べる。

2. SPARS-J

2.1 概要

SPARS-J (*Software Product Archive, analysis and Retrieval System for Java*) は、Java 言語 [6] で記述されたソースコードを対象としたソフトウェア部品検索システムである。一般にソフトウェア部品 (Software Component) は再利用できるように設計された部品とされる [5] が、SPARS-J ではクラスを部品の単位とみて、各クラスのソースコードに対する検索機能を実現している。実際の利用では、SPARS-J は、Web ブラウザを介し入力された検索キーに対して、対応する索引キーを持つクラスを抽出し、その一覧を出力する。検索結果は検索キー内の単語の部品における出現頻度の計算や、利用関係を評価する手法 [4] により順位付けされる。

また、検索結果においては、あらかじめ抽出した部品間の利用関係や類似部品に関する情報に基づいて、部品単体の情報だけでなく、その部品に関連する部品の情報も提供する。これにより、各部品の利用例の取得や、ソフトウェアの構造の理解などを支援することができ、効率的な再利用や保守を行なうことが可能となる。

2.2 順位付け手法

一般に検索を行う場合、検索キーと合致する適合文書は複数存在することが考えられる。そのため、検索結果を提示する際には適合文書に対して順位付けを行い、適切な順位で表示することが必要となる。検索システムにおいては、検索キーと各文書の適合度を計算し、その適合度に基づいて順位付けを行う手法が一般的である。

しかし、ソフトウェア部品においては、クラスの継承やメソッド呼び出しなどのソフトウェア部品間の関係が存在し、それらは抽出可能である。そのソフトウェア部品の特性を生かすために、我々は、検索キーと適合度が高い部品を上位に位置づけることが重要であると同様に、よく利用されている部品を上位に提示することも重要であると考え。なぜなら、そのような部品は利用例も多く、部品の再利用や理解をスムーズに行い易いと考えられるからである。

そのため、我々は、SPARS-Jの実現に際し、検索キーと部品の適合度を表す指標と共に、よく利用されている部品かを定量的に評価するための指標を実現した。以下、それぞれについて説明を行う。

2.2.1 Keyword Rank 法

自然言語文書検索システムで一般的に用いられている手法では、検索対象となる文書集合から各文書の特徴を表すキーワードである索引語を抽出し、索引語の集合によってその文書の内容を近似する。登録するそれぞれの文書の特徴を的確に表すように付与された、索引語の集合などからなる情報のことを索引キーと呼ぶ。検索キーは検索者の要求の内容を近似しているため、検索キーと各文書の索引キーとの適合度を測り、その適合度に基づいて各文書を順位付けする手法がよく用いられる。

SPARS-Jでは、検索キーと部品の適合度を表す指標の算出手法として、情報検索の分野で一般的に用いられる **TF-IDF** 法 [7] に基づいた手法を採用した。TF-IDF 法は任意の部品中における特定の索引語の出現頻度 TF (*Term Frequency*)、および特定の索引語を含む部品数の逆数 IDF (*Inverse Document Frequency*) の値をもとに索引語の重みの正規化を行う。その後、検索キー内の単語毎に各部品における索引語の重みを計算し、それに基づいて検索キーと各部品の適合度を算出する。

SPARS-Jでは、索引語の重みを算出する際に、出現した単語の部品内での役割に応じて、異なる重みを与えている。経験的にクラス定義名やメソッド定義名など、その部品の概念を表す索引語に対して大きな重みを設定することで、与えられた検索キーに適合する結果を得やすいことが分かっている。以下では、後述する CR 法に対応させて、本手法を **KR 法** (*Keyword Rank 法*) と呼び、測定した適合度の値を KR 値と呼ぶ。

2.2.2 Component Rank 法

我々はこれまでに、ソフトウェア部品間の利用関係に基づいて順位付けする手法 (*Component Rank 法*, **CR 法**) を提案している [4], [15]。CR 法では、十分な時間が経過し利用関係が収束した部品の集合に対して、各部品間に存在する利用関係に基づいてグラフおよび行列を構築し、構築された行列に対して繰り返し計算を行う事で各部品を評価する。求められる値は、開

発者が利用関係に沿って参照を行うと仮定した場合の各部品の参照されやすさを表しており、よく利用される部品や、重要な部品から利用される部品の順位は高くなる。CR 法によって測定した適合度の値を CR 値と呼ぶ。

実際の部品の集合には多数のコピーや類似した部品が存在しており、異なるシステムをまたいで全く同じ、もしくはほとんど似た部品が現れる場合、それらの部品が再利用されたものと推測できる。そのため、CR 法では類似した部品を部品群としてまとめ、解析の単位としている。こうすることで、それぞれの類似した部品への利用関係が一つの部品群への利用関係とみなされるため、コピーされた部品への評価を高くすることが可能となる。類似部品の判定方法は [8] で示している。

2.2.3 順位統合

SPARS-J の CR 法、KR 法はそれぞれ別の観点から、部品の順位付けを行っている。この順位付けを統合することで、それぞれが持つ検索アルゴリズムの特徴を考慮した検索が可能であると考えられる。このような順位統合は、複数の検索システムで同一の検索対象集合の検索を行ない、その結果を統合することで検索結果を返す、メタ検索システムで利用される手法で、いくつかの順位統合の手法が提案されている。

SPARS-J では CR 法、KR 法の順位統合手法として、Borda の手法 [1] を採用している。Borda の手法では、各順位に対して評価点を割り当て、その合計値をもとに最終的に統合された順位を求める。SPARS-J の CR 法、KR 法それぞれで得られた順位を評価点とし、2 つの評価点を足した合計点を昇順に並べ直した順位を統合後の検索結果とする。

例えば、表 1 は CR 法、KR 法、および Borda の手法で統合された (CR+KR) 結果を、表 2 はそれに対応した各部品の評価点を表しており、表中の A, ..., J はそれぞれ部品である。部品 A は CR で 1 位、KR で 4 位であり、評価点は 5 点となる。よって部品 A は順位を統合した結果 2 位となる。

以下では、この CR 法、KR 法の出す順位統合によって順位を求める手法を CR+KR 法と呼び、得られた順位を CR+KR による順位と呼ぶ。

表 1 Borda の手法

	CR	KR	CR+KR
1 位	A	C	C
2 位	E	F	A
3 位	C	G	E
4 位	I	A	F (3 位)
5 位	J	B	B
6 位	B	D	I (5 位)
7 位	H	I	G
8 位	F	E	J
9 位	G	H	D
10 位	D	J	H (9 位)

表 2 評価点

部品	CR	KR	CR+KR
A	1	4	5
B	6	5	11
C	3	1	4
D	10	6	16
E	2	8	10
F	8	2	10
G	9	3	12
H	7	9	16
I	4	7	11
J	5	10	15

3. 評価実験

本節では SPARS-J について、ソフトウェア部品検索システムとしての実験的評価を行い、システムの有効性を検証する。

3.1 実験目的

(1) 一般的な検索システムとの順位付け性能の比較
この評価実験では、一般的な検索システムと比較して、SPARS-J が適合する部品をどれだけ十分に開発者に提供できるかを評価する。比較においては、ある目的に応じた部品の取得を想定した検索を行い、上位で取得された部品の適合率をもとに評価を行う。SPARS-J を用いることで迅速に目当てのソフトウェア部品を検索でき、ソフトウェアの再利用に SPARS-J が有効であることを示す。

(2) SPARS-J 内で利用される各順位付け手法の比較
2.2 節で示したとおり、SPARS-J は CR 法と KR 法、更にそれらを統合した CR+KR 法という 3 種類の順位付け手法を実現している。この実験では、適合率やユーザが最適だと思う順位と実際に得られた順位との差を求めることで、各評価付け手法の特徴や、最適な評価手法を調査する。

(3) 企業内のソフトウェアに対する SPARS-J の適用
SPARS-J はただ単に検索機能を提供するだけでなく、データベース内における利用関係や類似部品等、部品に関する様々な情報を提供することができ、部品の管理や理解を支援するシステムとしても有効であると考えられる。この実験では、SPARS-J を企業におけるソフトウェア開発現場に適用し、実際の開発者・管理者に対してアンケートを実施することで、それらの情報が実際の開発現場ではどのように役立てられるのかを確認する。

以下では評価における指標について説明をした後に、各実験内容についてその詳細と実験結果を示す。

3.2 実験における評価尺度

実験では、以下の評価尺度を用いる。

- 適合率

検索された文書の中で適合している文書数の割合を指す指標で、適合率が高いほど検索結果に不要な文書を含んでおらず、よい検索結果を示しているといえる。

- ndpm (Normalized Distance-based Performance Measure) 法 [14] による評価

順位付け手法の比較に用いられる手法で、文書間の順位付けの差を表現する手法である。詳細は後述する。

3.2.1 ndpm 法 [14]

ndpm 法とはユーザ・プリファレンスの概念を用いて、文書間の順位付けの差を表現する手法である。ユーザ・プリファレンスとは、そのユーザの好みを明示的に表現したものである。ここでは、文書におけるユーザプリファレンスを、文書を二つ一組にして比較した場合の、「どちらの文書がより有用である（あるいは目的に適合する）か」を記述したものとす。

この時、 D を文書の有限集合とすると、ユーザプリファレンスは、 $d, d' \in D$ について、 D 上の二項関係 \succ によって定義され、文書集合に対するユーザの各判断は、文書間の相対的な順序付けとみなすことができる。

[14] では、この文書間の相対的な順序に着目し、ユーザの順位付けとシステムの順位付けの距離に基づいてシステムの性能を計測する手法として ndpm 法が提案されている。ndpm は以下のように計算することができ、0~1 の値に正規化される。

ndpm の計算方法

$d, d' \in D$ の全ての組み合わせに対して、 $d \succ_1 d'$ かつ $d' \succ_2 d$ となる d, d' の組数を m とし、 n を文書集合 D の要素数とすると

$$ndpm(\succ_1, \succ_2) = \frac{m}{nC_2}$$

値が小さいほど、ユーザの順位付けとシステムの順位付けの差が少なく、システムの順位付けが妥当で、ユーザの理想的な順位付けに沿ったものであることを示している。[14] では、この手法が順序的尺度の測定手法として妥当であることを確認している。

3.3 評価実験 1: 一般的な検索システムとの順位付け性能の比較

実験では、以下の2つの検索エンジンを用いて SPARS-J による検索結果の比較を行った。

- Google [3]
- Namazu [9]

Google は多くの検索目的に用いられる有用な Web ページ検索システムである。実際には Web ページ上でソースコードの解説を行っているものも多く、ソフトウェアを検索するときには Google を用いることが多い。また、Namazu はフリーウェアでありながら幅広い導入実績を持つ、信頼性の高い全文検索システムである。

一般に、Web ブラウザを介して提供される検索システムにおいては、検索結果として一番最初に表示されるページに該当部品があれば、利用者はそれを利用するが、なければ利用者は下位を見に行くよりも、クエリを変更して再検索することが多い。

そこで比較実験では、ある目的に応じた部品の取得を想定した検索を行い、上位で取得された部品の適合率をもとに評価を行う。具体的には、システムが順位付けしたときの検索結果上位 10 件の部品に対する適合率を計算する。SPARS-J を用いることで迅速に目当てのソフトウェア部品を検索でき、ソフトウェアの再利用に SPARS-J が有効であることを確認する。

検索対象データベース

実験においては、JDK1.4 や、SourceForge.net [12] などのオープンソース開発コミュニティなどで公開されているソースコード、約 14 万ファイル (約 15 万クラス) から SPARS-J のデータベースを構築した。

また、Namazu についても、データベースを自分で構築することができるため、SPARS-J のデータベースに登録するソフトウェア部品群と全く同じものを用いた。

それに対して、Google に対する検索では、一般に公開されている検索エンジンに対してクエリを入力し、得られた結果を Google の結果とみなした。

なお、SPARS-J における検索結果は類似部品を一つのグループにまとめているため、それらの部品群は一件の検索結果として扱われている。そのためある一つの部品が検索結果の上位に順位付けされている場合、その類似部品はいくつあったとしても順位付けに影響しない。そこでそれに対応した処置として、

Google ではインデント表示されている類似ページを一件とみなし、Namazu の検索結果は手作業で類似部品をまとめ、一件として扱った。

3.3.1 実験結果

上記の検索対象データベースに対して、表 3 に示した検索キーワードで検索した結果を評価した。各キーワードの検索結果に対する各検索システムの適合率の計算結果を示す。表中の K1, ..., K10 は表 3 の検索キーワードの番号に対応している。

3.3.2 分析・評価

適合率について SPARS-J と Google, Namazu とで、対応のある平均値の差の検定 [11] を有意水準 5% で行ったところ、有意な差が存在した。このことから SPARS-J は、ソフトウェア検索において、総合検索エンジンや文書検索システムと比較して明らかに優れており、部品検索システムとして機能を十分に果たすと期待できる。

表 3 検索キーワード

	検索キーワード
K1	quicksort
K2	binarysearch
K3	clock applet
K4	applet textarea
K5	randum number generate
K6	stack push pop
K7	chat server client
K8	classfile dump
K9	zip deflate
K10	write read inputstream outputstream

3.4 評価実験 2: SPARS-J 内で利用される各順位付け手法の比較

SPARS-J におけるそれぞれの順位付け手法である、CR 法、KR 法、CR+KR 法の比較を行う。SPARS-J の場合、順位付け手法を変えても、部品の検出手法は変わらない。そのため得られる検索結果は、検索キーと合致する索引キーを持つ部品の集合において、その順位付けのみが変わったものである。

そこで実験では、実験 1 と同じ検索を行い、各順位付け手法における検索結果上位の適合率を求めると共に、検索結果からユーザにとっての最適な順位を作成し、その最適な順位との比較を行う。具体的には、システムが順位付けしたときの検索結果上位 10 件の部品に対する適合率を計算するとともに、ndpm 法を用いて、実際に得られた順位と最適な順位との差を計算することで、各評価付け手法の特徴や、最適な評価手法を調査する。

検索対象となるデータベースおよび検索キーワードは、評価実験 1 と同じものを利用した。

3.4.1 実験結果

SPARS システムの各順位付け手法に対して、表 3 に示された各キーワードの検索結果に対する適合率の計算結果を表 4 に、ndpm 値の計算結果を表 5 に示す。

3.4.2 分析・評価

適合率について対応のある平均値の差の検定を行ったところ、KR法とCR+KR法はCR法と比較して有意水準5%で有意な差が存在した。また、ndpm値の比較についても同様に対応のある平均値の差の検定を行ったところ、CR法とCR+KR法は、KR法と比較して有意水準5%で有意な差が存在した。これらのことから、CR法はKR法と比較して、全体的にユーザの想定する順位付けがされていると考えられ、KR法は上位に適合部品が多く存在しやすいと考えられる。また、CR+KR法はCR法とKR法の両方の性質を受け継ぎ、それぞれCR法、KR法より順位付けの性能が改善していると考えられる。

表4 適合率の評価結果

	CR	KR	CR+KR	Google	Namazu
K1	1	1	1	0.7	0.9
K2	1	1	1	0.4	0.6
K3	0.5	0.5	0.5	0.3	0.4
K4	0.4	0.9	0.8	0.3	0.6
K5	0.4	0.4	0.4	0.1	0.3
K6	0.2	0.2	0.2	0	0.1
K7	0.9	1	1	0.3	0.4
K8	1	0.8	1	0.1	0.2
K9	0.6	0.7	0.7	0.4	0.4
K10	0.5	0.7	0.7	0.4	0.7
Ave.	0.65	0.72	0.73	0.3	0.46

表5 ndpm法による評価結果

	CR	KR	CR+KR
K1	0.036078	0.048104	0.037004
K2	0.193676	0.260840	0.221344
K3	0.133333	0.116667	0.091667
K4	0.122908	0.199672	0.189121
K5	0.207827	0.191633	0.194332
K6	0.183755	0.184040	0.159717
K7	0.080519	0.103247	0.079870
K8	0.047059	0.109244	0.052101
K9	0.190476	0.304762	0.257143
K10	0.209524	0.323810	0.266667
Ave.	0.143373	0.178014	0.140611

3.5 評価実験3:企業内のソフトウェアに対するSPARS-Jの適用

この実験では、企業の協力を得てSPARS-Jを実際の開発現場において適用した結果を紹介する。協力した企業は食品製造業で、商品の在庫や出荷管理、社内の事務手続きに関するソフトウェアシステムを構築している。

我々は、実際の業務における検索システムを用いたソースコード管理支援を目的として、蓄積された企業内のソフトウェアの一部(約2500クラス)を用いてSPARS-Jのデータベースを構築し、適用した。被験者は企業内のソフトウェア開発者7名で、その中にはソフトウェアの部品化や標準化の担当者も含む。

2週間の運用の後、その7名に対して、本システムの使いやすさに関してアンケートを行った。本システムの使いやすさに関して行ったアンケートの項目とその回答結果を下の表6に示す。回答は1から5までの5段階評価で、5が一番良い評価であり、逆に1が一番悪い評価となっている。また表中の空白はその項目に対して無回答であったことを示す。また、表の一番右の列はその項目において最も回答が多かった評価値(最頻値)を表す。

表6 アンケート内容とその回答

回答者	回答(良54321悪)							最頻値
	A	B	C	D	E	F	G	
使い勝手について								
パッケージブラウザの利用	5	4	4	5	5	3	3	5
類似部品のグループ化	5	2	4	5		4	3	5,4
被利用クラスの参照	5	5	5	5		5	5	5
利用クラスの参照	5	5	5	1		5	5	5
メトリクス値	4	2		1	1	4	5	4,1
ソースコードのダウンロード	3	5		1	5	2	5	5
時間的コストの削減	5	3		3	5	4	1	5,3
ソフトウェア品質の向上	3	3		5	3	4	1	3
企業内のソフトウェア把握	5	5	3	1	3	2	1	5,3,1
表示について								
検索結果一覧	5	3	4	4	5	3	5	5
ハイライト表示	5	5	3	5	5	5	5	5

3.5.1 分析・評価

表6のアンケート結果から、SPARS-Jにおける「パッケージブラウザの利用」、「類似部品のグループ化」、「利用クラスの参照」、「被利用クラスの参照」「検索結果一覧」、「ハイライト表示」は、実際の開発現場の開発者や管理者から高い評価を得ていることがわかる。

中でもパッケージブラウザと利用クラスの参照、被利用クラスの参照という機能は、ソフトウェアシステムの全体像の把握という目的に非常に合致しており、非常に便利であったという意見が多かった。具体的な感想では、SPARS-Jを用いることで、「部品を利用しているアプリの把握」ができ、「部品改訂時の影響範囲の調査」がしやすいと評価された。また、検索時間が短くレスポンスがよいため非常に利用しやすいという意見もあった。

4. 実験の考察

4.1 一般的な検索システムとの順位付け性能の比較

実験結果として得られた上位の適合率を比較すると、SPARS-JはJavaのソフトウェア部品検索を想定した場合、GoogleやNamazuより優れていると言える。

GoogleはWebページ検索システムであり、ソフトウェア部品検索に特化していない。そのため、検索対象範囲が広すぎて検索結果が絞りきれず、ソフトウェア部品検索として必要のないものが多数上位に順位付けされ、適合率が低くなったと考えられる。NamazuはSPARS-JのKR法と同様に評価値にTF-IDF法を用いている。しかしながら、上位の適合率におい

て有意差が見られたのは、SPARS-J では構文解析を行い索引語の種類によって重み付けを行っているためであると考えられる。索引語の重み付けを行うことで、重要な位置に出現している場合には高い評価を与えることができる。実験の結果から、ソースコード中で単に出現回数が多いだけで上位に順位付けするのは、誤りであるということがわかる。

4.2 SPARS-J 内で利用される各順位付け手法の比較

対応のある平均値の差の検定により、各手法によって得られた適合率、ndpm 値において手法間で有意差が存在した。実験より、CR 法は全体的にユーザの想定する順位に近いという点で優れており、一方、KR 法は検索結果上位における適合率の点で優れているという性質を持つと考えられる。

また、CR+KR 法によって順位付けされた結果は、CR 法の持つ「ユーザの想定する順位で部品が並びやすい」という性質と、KR 法の持つ「検索結果上位に適合部品が存在しやすい」という性質を持つものであった。CR+KR 法は CR 法と KR 法の両方の性質を受け継ぎ、CR+KR 法により順位付けの性能が改善できると考えられる。これは、CR 法か KR 法の片方の順位付けでのみ高順位に順位付けされても CR+KR 法による順位付けでは上位には順位付けされず、CR 法と KR 法の両方で高順位に順位付けされた部品のみが上位に順位付けされるからであると考えられる。

4.3 企業内のソフトウェアに対する SPARS-J の適用

SPARS-J の実際の利用後のアンケートでは、検索結果だけでなく、それに付随した要素である、「パッケージブラウザの利用」、「ハイライト表示」「利用クラスの参照」、「被利用クラスの参照」が役に立ったという回答が多かった。中でもパッケージブラウザと利用クラスの参照、被利用クラスの参照という機能は、ソフトウェアシステムの全体像の把握という目的に非常に合致しており、非常に便利であったという意見が多かった。SPARS-J は、ただ単に再利用時に部品を取得したいという目的だけでなく、既存の部品の保守や管理という目的でも非常に有用であると考えられる。

アンケート時に得られた要望としては、日本語の対応など検索面での性能向上と共に、利用者に応じた検索範囲の制限機能の実現などが挙げられた。これは、実際の開発現場ではコーディングを社外に発注することが多く、様々な立場の開発者がシステムの作成に関わっていることを示している。組織における実利用を考えた場合、単に検索機能や使い勝手だけではなく、セキュリティなどの面を考える必要があることがわかったのは非常に有意義であったと思う。現在は、日本語対応など、要望のあった機能の一部を実現している。

5. まとめと今後の課題

本研究では、ソフトウェア部品検索システム SPARS-J の実験の評価を行った。実験では、SPARS-J と他の検索システム Google, Namazu との比較や、SPARS-J 内の各順位付け手法の比較を行った。さらに、実際の企業の開発現場に適用し、アンケートによる定性的な評価を行った。

評価の結果、SPARS-J は比較した他の検索システムと比べ

て検索結果上位における適合率で有意な差を見ることができた。また、SPARS-J 内における各順位付け手法の違いにおいては、検索結果上位の適合率や、ユーザの理想順位との差の比較において 2 つの順位付け手法にそれぞれ特徴が見られ、さらにそれらを組み合わせた順位付け手法においては単独で順位付けを行うより順位付けの性能の向上が見られた。さらに、アンケートの結果から、SPARS-J はソフトウェアの再利用を目的とした検索だけでなく、ソフトウェアシステムの保守などを目的とした、部品の管理や把握に非常に有効であるという結果が得られた。

これらを総合すると、SPARS-J はソフトウェア部品検索システムとして高度な順位付けの性能を持ち、ソフトウェア開発時の部品再利用の促進・支援や保守時における部品の把握を支援するのに有効で、その結果、ソフトウェア品質の向上支援に利用でき、過去の資産の共有に役立つと考えられる。

今後の課題として以下が挙げられる。

- 他のソフトウェア部品検索システムとの比較
- 順位付け性能以外の定量的な評価

謝辞 本研究は、独立行政法人科学技術振興機構計算科学技術活用型特定研究開発推進事業 (ACT-JST) の支援を受けている。

文 献

- [1] J. C. Borda: "M'emoire sur les 'elections au scrutin", *Histoire de l'Acad'emie Royale des Sciences*, (1781).
- [2] C. Braun, in John J. Marciniak, editor", *Encyclopedia of Software Engineering*, Vol. 2, John Wiley & Sons, pp. 1055-1069, (1994).
- [3] "Google", <http://www.google.com/>.
- [4] K. Inoue, R. Yokomori, H. Fujiwara, T. Yamamoto, M. Matsushita, S. Kusumoto: "Component Rank: Relative Significance Rank for Software Component Search", *Proceedings of the 25th International Conference on Software Engineering (ICSE2003)*, pp. 14-24, Portland, Oregon, U.S.A., May 6-8, (2003).
- [5] I. Jacobson, M. Griss and P. Jonsson: "Software Reuse", *Addison Wesley*, (1997).
- [6] J. Gosling, B. Joy, G. Steele, and G. Bracha: "The Java Language Specification", Addison-Wesley, (2000).
- [7] 北, 津田, 獅々堀: "情報検索アルゴリズム", 共立出版, (2002).
- [8] 小堀, 山本, 松下, 井上: "類似度メトリクスを用いた Java ソースコード間類似度測定ツールの試作", 電子情報通信学会技術研究報告, SS2003-2, Vol. 103, No. 102, pp. 7-12, (2003).
- [9] "Namazu Project", <http://www.namazu.org>.
- [10] 西, 梅森, 山本, 横森, 松下, 楠本, 井上: "Java ソフトウェア部品解析・検索システム SPARS-J の構築", 電子情報通信学会技術研究報告, SS2003-23, Vol. 103, No. 481, pp.43-48, (2003).
- [11] 芝, 渡部, 石塚 編: "統計用語辞典", 新曜社, (1984).
- [12] "SOURCEFORGE.net", <http://sourceforge.net/>.
- [13] 山本, 横森, 松下, 楠本, 井上: "利用頻度に基づくソフトウェア部品の解析・検索システムの提案", 電子情報通信学会技術研究報告, SS2002-17, Vol. 102, No. 329, pp. 13-18, (2002).
- [14] Y. Y. Yao: "Measuring Retrieval Effectiveness Based on User Preference of DocumentS", *Journal of the American Society for Information Science*, Vol. 46, No. 2, pp. 133-145, (1995).
- [15] 横森, 藤原, 山本, 松下, 楠本, 井上: "利用実績に基づくソフトウェア部品重要度評価システム", 電子情報通信学会論文誌 D-I, Vol. J86-D-I, No.9, pp.671-681, (2003), and Technical Report of SE Lab, Dept. of Computer Science, Osaka University, SEL-Nov-21-2002, Nov (2002).