

## 文書検索システムの動的抄録提示インタフェースの評価

酒井哲也      三池誠司      住田一男

(株) 東芝 研究開発センター

膨大な文書検索結果の全文を読み所望の文書を選出したり情報を得たりする労力の軽減のために、我々は、検索した文書の抄録を自動生成して提示するインタフェースを開発した。これは、ユーザーによる「詳しく」「簡単に」の指示に応じて、抄録の任意の部分の長さを変更することを可能とする。このインタフェースの評価実験を行った結果、文書の要不要の判定時間は、原文のみを提示した場合に比べ80%程度に短縮されることがわかった。この代償として、平均では判定の質が原文のみを提示した場合の90%程度に低下してしまうが、内容の深い理解を必要としない判定問題においては判定の質を保持することができた。

## Evaluation of the Abstract Presentation Interface for a Document Retrieval System

Tetsuya SAKAI      Seiji MIIKE      Kazuo SUMITA

Research and Development Center, Toshiba Corporation  
1, Komukai Toshiba-cho, Saiwai-ku, Kawasaki 210, Japan

In order to lighten the burden of browsing through numerous retrieved documents to select relevant documents or to obtain useful information, we have developed a user interface for generating and presenting abstracts of the retrieved documents. The interface enables the user to alter the length of any part of the abstract by entering DETAIL or BRIEF commands. Experiments show that, using this interface, the time for judging the relevance of each retrieved document can be reduced to 80% in comparison to using an interface that only presents the full-text to the user. Although the quality of relevance judgment was on average lowered to 90% of that achieved in the case of full-text presentation, it was not affected when deep understanding of the contents was not required.

## 1 まえがき

近年、計算機パワーの飛躍的向上を背景に、文書の全文を検索対象とした全文検索システムの実用化が進みつつある。しかし、漏れもなく誤りもない所望の文書の検索を行うことは非常に困難であり、ユーザーに始めの検索結果から必要あるいは不必要な文書を選ばせて検索要求を修正し、再検索を行うという作業により、徐々に所望の文書が得られていくというのが実状である。電子化文書の普及などに伴い、検索対象となる文書の量が膨大になればなるほど、検索結果の全文を読んで所望の文書を選出したり情報を得たりする労力は大きくなると考えられる [三輪 92]。

我々はこのようなユーザーへの負荷を軽減するために、検索した文書の抄録を自動生成して提示し、かつその詳細度をユーザーとの対話により調節するインタフェースを開発した [住田 94, Miike94]。本報告では、このインタフェースの有効性を検証するために行った実験について述べる。

## 2 動的抄録提示インタフェース

文書検索システム BREVIDOC [住田 94] に搭載されている動的抄録提示インタフェースは、文書構造解析 [三池 93, 住田 93, Ono94] により事前に抽出された文書の章だてや文間の修辞関係を参照しながら、リアルタイムに抄録を生成する。より具体的には、文書構造解析により各文に付与されたペナルティ値を参照し、閾値よりもペナルティ値が小さい文のみを抜粋表示する。ユーザーにより抄録をより詳しく、あるいは簡単にしよう指示が出されるたびに、この閾値を調節する。また、抄録と並べて原文も提示し (図 1)、原文画面中では、抄録のために抜粋されている文 (以後、抄録文と呼ぶ) を他の文とは別の色で表示している。

ブラウジングの負荷軽減のために、予め定めた項目について原文から情報を抽出し簡単な文を生成する研究もあるが [安藤 94]、対象分野が限定されている。このような情報抽出によるアプローチは、作りつけの抄録を提示した場合と同様、文書の主題に関する要不要 (relevance) の判定には有効であると考えられるが、個々のユーザーの用途に合わせた要不要 (usefulness)

[Ingwersen92] の判定には不向きであると考えられる。これに対して、動的抄録提示インタフェースは、ユーザーとの対話により必要最少限の文を提示しながら、主題に限らない任意の情報を迅速に提供することができると考えられる。

## 3 実験

### 3.1 目的

今回の実験は、動的抄録提示インタフェースが文書の主題に関する要不要 (relevance) の判定に与える影響を調べることを目的とする。抄録の提示により要不要の判定時間が短縮される一方で、判定を行った時点で抄録が重要な文を含まなかったり、重要でない文を含んでいたりするために誤判定が起こる可能性があるため、判定時間と共に判定の正確さをも測定することにより本インタフェースの有効性を検証する。

### 3.2 方法

動的抄録と原文とを並べて提示する本インタフェースの効果を、原文のみを提示するインタフェースの場合と比較する。このために、本インタフェース (図 1) に加えて、抄録画面が表示されない以外は全く同等の機能をもつ実験用インタフェースを用意した (図 2)。以後、前者を (a)、後者を (b) と呼ぶ。いずれの場合も、被験者は以下の手順で実験を行った。

- 表 1 のいずれかの与えられた問題を読む。
- 表 1 の検索語を検索語入力画面に入力し [住田 94]、全文検索の結果を表示させる。
- 提示された文書群中の各文書が問題に該当するかどうかを、(a) 抄録と原文、あるいは (b) 原文のみを見ることにより判定する。
- 該当するかどうかの判定に要した時間 (文書群中の各文書の判定に要した時間の総和) と、判定再現率 = (被験者が該当するとした正解文書数) / (正解文書数)

<p>ゆらく日本的雇用と暮らし(930601)</p> <p>◆</p> <p>その中で、望ましい流れを選び育ててゆく作業が、暮らしやすい社会をつくることに結びつく。不況でゆらいでいる日本の雇用についても、そうした視点からの冷静な洗い直しが必要だろう。</p> <p>現場と事務職との差をつけずに処遇することで、現場のやる気を引き出す人事管理も、特徴の一つといえよう。その背景には、企業を大家族時代の家と同じにみる「企業一家主義」がある。理想的な親子や兄弟を頭において、運命共同体的に「ウチの会社」のために働くことをよしとするのである。労使双方の意識が終身雇用を前提としなくなりつつあるという点では、そうかも知れない。しかし、代償として、根深かった組織への帰属意識もまた薄れてくることを忘れてはならない。こうした転換期は、個人と組織とのあり方を見直す好機なのだ。だが、その企業一家意識も今後は弱まってゆくに違いない。働く人の側に沿った多様な適尺杖を用意できる組織が望ましい。この方向で新たな理念づくりを努める企業に未来は開けてくる。</p>	<p>ゆらく日本的雇用と暮らし(930601)</p> <p>◆</p> <p>いつの世にも、さまざまな流れが渦巻いている。その中で、望ましい流れを選び育ててゆく作業が、暮らしやすい社会をつくることに結びつく。不況でゆらいでいる日本の雇用についても、そうした視点からの冷静な洗い直しが必要だろう。</p> <p>日本の雇用システムについては、「三種の神器」がよく引き合いに出される。終身雇用と年功序列、企業別組合だ。現場と事務職との差をつけずに処遇することで、現場のやる気を引き出す人事管理も、特徴の一つといえよう。</p> <p>その背景には、企業を大家族時代の家と同じにみる「企業一家主義」がある。理想的な親子や兄弟を頭において、運命共同体的に「ウチの会社」のために働くことをよしとするのである。</p> <p>ところが、不況でホワイトカラーが余ってしまい、管理職への露骨な肩たたきさえ目につく。そのため、日本の雇用は崩壊との声が高まっている。労使双方の意識が終身雇用を前提としなくなりつつあるという点では、そうかも知れない。</p> <p>しかし、代償として、根深かった組織への帰属意識もまた薄れてくることを忘れてはならない。そうならば、仕事一筋の会社人間的な生き方は減ってくる。</p> <p>こうした転換期は、個人と組織とのあり方を見直す好機なのだ。これまで個人より組織の方を大切にしすぎなかったか。敗戦後の廃墟から復興に力を注いだ時代はそれでもよかった。けれども、いまではひずみが明らかだ。「三過き家族」の実態さえなかなか改めにくい。勉強し過ぎの子ども、働き過ぎの親、暇があり過ぎるお年寄り、にみられる不均衡だ。</p> <p>いまや企業には体力が付き、国の経済力は大きくなった。個人の暮らしをもっと充実させるのは当然のことだ。</p> <p>それに、「三種の神器」といっても、通用されるのは大企業と官庁の男性労働者だけである。過半数を占める中小企業の働き手はもちろん、急増している女性労働者にも縁が薄い。終身雇用が崩れるとしても、働く人全体からすれば、一部のことにはすぎない。しかも、人生八十年の長寿社会を迎えて、やっと定着しかけた六十歳「定年」なども終身雇用などではありえ</p>
--	--

図 1: (a) 抄録と原文を提示するインタフェース

<p>ゆらく日本的雇用と暮らし(930601)</p> <p>◆</p> <p>その中で、望ましい流れを選び育ててゆく作業が、暮らしやすい社会をつくることに結びつく。不況でゆらいでいる日本の雇用についても、そうした視点からの冷静な洗い直しが必要だろう。</p> <p>現場と事務職との差をつけずに処遇することで、現場のやる気を引き出す人事管理も、特徴の一つといえよう。</p> <p>その背景には、企業を大家族時代の家と同じにみる「企業一家主義」がある。理想的な親子や兄弟を頭において、運命共同体的に「ウチの会社」のために働くことをよしとするのである。</p> <p>ところが、不況でホワイトカラーが余ってしまい、管理職への露骨な肩たたきさえ目につく。そのため、日本の雇用は崩壊との声が高まっている。労使双方の意識が終身雇用を前提としなくなりつつあるという点では、そうかも知れない。</p> <p>しかし、代償として、根深かった組織への帰属意識もまた薄れてくることを忘れてはならない。そうならば、仕事一筋の会社人間的な生き方は減ってくる。</p> <p>こうした転換期は、個人と組織とのあり方を見直す好機なのだ。これまで個人より組織の方を大切にしすぎなかったか。敗戦後の廃墟から復興に力を注いだ時代はそれでもよかった。けれども、いまではひずみが明らかだ。「三過き家族」の実態さえなかなか改めにくい。勉強し過ぎの子ども、働き過ぎの親、暇があり過ぎるお年寄り、にみられる不均衡だ。</p> <p>いまや企業には体力が付き、国の経済力は大きくなった。個人の暮らしをもっと充実させるのは当然のことだ。</p> <p>それに、「三種の神器」といっても、通用されるのは大企業と官庁の男性労働者だけである。過半数を占める中小企業の働き手はもちろん、急増している女性労働者にも縁が薄い。終身雇用が崩れるとしても、働く人全体からすれば、一部のことにはすぎない。しかも、人生八十年の長寿社会を迎えて、やっと定着しかけた六十歳「定年」なども終身雇用などではありえ</p>	<p>ゆらく日本的雇用と暮らし(930601)</p> <p>◆</p> <p>いつの世にも、さまざまな流れが渦巻いている。その中で、望ましい流れを選び育ててゆく作業が、暮らしやすい社会をつくることに結びつく。不況でゆらいでいる日本の雇用についても、そうした視点からの冷静な洗い直しが必要だろう。</p> <p>日本の雇用システムについては、「三種の神器」がよく引き合いに出される。終身雇用と年功序列、企業別組合だ。現場と事務職との差をつけずに処遇することで、現場のやる気を引き出す人事管理も、特徴の一つといえよう。</p> <p>その背景には、企業を大家族時代の家と同じにみる「企業一家主義」がある。理想的な親子や兄弟を頭において、運命共同体的に「ウチの会社」のために働くことをよしとするのである。</p> <p>ところが、不況でホワイトカラーが余ってしまい、管理職への露骨な肩たたきさえ目につく。そのため、日本の雇用は崩壊との声が高まっている。労使双方の意識が終身雇用を前提としなくなりつつあるという点では、そうかも知れない。</p> <p>しかし、代償として、根深かった組織への帰属意識もまた薄れてくることを忘れてはならない。そうならば、仕事一筋の会社人間的な生き方は減ってくる。</p> <p>こうした転換期は、個人と組織とのあり方を見直す好機なのだ。これまで個人より組織の方を大切にしすぎなかったか。敗戦後の廃墟から復興に力を注いだ時代はそれでもよかった。けれども、いまではひずみが明らかだ。「三過き家族」の実態さえなかなか改めにくい。勉強し過ぎの子ども、働き過ぎの親、暇があり過ぎるお年寄り、にみられる不均衡だ。</p> <p>いまや企業には体力が付き、国の経済力は大きくなった。個人の暮らしをもっと充実させるのは当然のことだ。</p> <p>それに、「三種の神器」といっても、通用されるのは大企業と官庁の男性労働者だけである。過半数を占める中小企業の働き手はもちろん、急増している女性労働者にも縁が薄い。終身雇用が崩れるとしても、働く人全体からすれば、一部のことにはすぎない。しかも、人生八十年の長寿社会を迎えて、やっと定着しかけた六十歳「定年」なども終身雇用などではありえ</p>
--	--

図 2: (b) 原文のみ提示するインタフェース

表 1: 実験に用いた問題

検索対象：朝日新聞社説			
文書群	正解文書数 / 文書数	検索語	問題「X について主に述べたもの」
A	5/10	雇用	X= 雇用のあるべき姿
B	4/16	円高	X= 円高対策
C	7/11	核	X= 核保有問題
D	7/17	選挙区制	X= 選挙区制改革
E	5/21	憲法	X= 憲法擁護を唱える政党や政治家
F	2/22	水	X= 水質汚濁
G	5/17	戦争	X= 戦争責任
H	2/16	欧州	X= 欧州統合問題

表 2: 実験の条件

	(a) 抄録+原文	(b) 原文のみ
文書群 A,B,C,D	被験者 (1)(2)(3)(4)	被験者 (5)(6)(7)(8)
文書群 E,F,G,H	被験者 (5)(6)(7)(8)	被験者 (1)(2)(3)(4)

表 3: 実験結果

4 回の平均→ 文書群↓	判定時間 (秒)			判定再現率 (%)			判定適合率 (%)		
	(a)	(b)	(a)/(b)	(a)	(b)	(a)/(b)	(a)	(b)	(a)/(b)
A	331	525	63.0	80	80	100	100	96	104
B	506	663	76.3	56	88	64	79	95	83
C	275	358	76.8	71	71	100	100	100	100
D	441	610	72.3	64	71	90	92	94	98
E	636	656	97.0	25	50	50	45	62	73
F	290	351	82.6	100	100	100	75	71	106
G	432	435	99.3	80	75	107	89	100	89
H	383	407	94.1	100	100	100	92	92	100
	平均 82.7			平均 89			平均 94		

判定適合率=(被験者が該当するとした正解文書数)/(被験者が該当するとした文書数)  
を記録する。ここで、正解文書とは問題を用意した筆者自身が問題に該当すると判定したものである。

例えば、表1の文書群Aは、「雇用」という検索語を含む文書10件からなる。検索システムにこの検索語を入力するとこれらの文書が検索されるので、この中から「雇用のあるべき姿について主に述べたもの」を被験者に選ばせる。

被験者は8人とし、表2のように、各被験者に(a)と(b)で同じ問題を与えないように配慮した。この表から、例えば被験者(1)は文書群A,B,C,Dに対して抄録付きの実験を行い、文書群E,F,G,Hに対して抄録なしの実験を行ったことがわかる。また、(a)の場合、動的抄録の圧縮率の初期値は30%とし、被験者に「詳しく」「簡単に」のボタンを押して抄録の長さを調節することを許した。

### 3.3 結果

表3に実験結果を示す。文書群A,C,F,Hについては、動的抄録提示インタフェースにより、判定再現率及び適合率を損なうことなしに判定時間が短縮されるという結果が得られている。一方、文書群B,D,E,Gについては、判定時間は短縮されているものの判定再現率や適合率が若干犠牲にされている。

## 4 考察

文書群によって実験結果に差が出た原因としては、抄録の質の差、問題の難易度の差、個人差などが考えられる。今回は、個人差はないものとみなし、抄録の質の差及び問題の難易度の差の2つについて考察する。

### 4.1 抄録の質の差

文書群A,C,F,Hと文書群B,D,E,Gの実験結果に差をもたらした要因として第一に考えられるのは、抄録の質のばらつきである。すなわち、文書群A,C,F,Hの抄録の質に比べて文書群B,D,E,Gの抄録の質が劣っていたために、判定再現率や適合率に悪影響が出たとい

う仮説を立てることができる。そこで、抄録の質と今回の結果の相関の有無を調べるために、今回実験に用いた文書の抄録の質を文書集合毎に評価した。

以下に、抄録の質の評価の手順を示す。

- 画面にちょうど収まるくらいに抄録を「詳しく」し、この時の文数Nを数える。
- 正解作成者に原文を与え、重要文をN文選ばせる。
- 抄録文N文と重要文N文を比較し、(抄録文でありかつ重要文である文の数)/Nを算出する。

このようにして評価された抄録の質の、文書群毎の平均値を表4に示す。

表4: 抄録の質

文書群	検索語	抄録の質平均 (%)
A	雇用	62.3
B	円高	50.1
C	核	53.9
D	選挙区制	52.7
E	憲法	57.1
F	水	52.5
G	戦争	49.9
H	欧州	49.4
		平均 53.5

なお、ここでの抄録の質は高いとは言い難いが、人手で作成した抄録の質を同じ方法により評価しても平均6割程度であることが報告されているので[Ono94]、ある程度実用にたえるものであると考えられる。

文書群A,C,F,Hの抄録の質は平均54.5%、文書群B,D,E,Gは52.5%で、有意差は見られない。これより、抄録の質のばらつきが判定再現率及び適合率に直接影響を与えているわけではないことがわかる。

### 4.2 問題の難易度の差

文書群A,C,F,Hと文書群B,D,E,Gの判定再現率及び適合率に差が出たのは、問題の難易度のばらつきのせいである可能性がある。

判定再現率及び適合率を損なわずに判定時間を短縮することができた文書群のうち、Cの適合率及びF,Hの再現率は抄録なしの場合で既に100%になっているので、これらは選定の比較的容易な問題であると考えられる。例えば、Fの文書群には「水をさす」という表現を含む文書が含まれているが、このような文書を「水質汚濁について主に述べたもの」ではないとして排除することはユーザーにとって容易である。これに対し、例えば再現率の低下が著しかったEを見てみると、問題文が「憲法擁護を唱える政党や政治家について主に述べたもの」と他よりも複雑であり、選定の難易度が高い。このため、抄録なしの場合でも再現率は50%という低い値をとっている。Fのような問題とは違い、字面を眺めるだけでなく内容を把握することが要求される問題であると言える。このように考えると、字面を眺めるような浅いレベルで要不要の判定ができる平易な問題では判定再現率及び適合率に影響が出ないが、判定にある程度深い内容理解を必要とするような問題では抄録の質の影響が判定再現率及または適合率の低下という形で現れる傾向があるように思われる。ただし、今回の実験では抄録なしの場合のGの適合率が100%であったにもかかわらず判定の質の低下が見られており、また難易度に客観性・一貫性をもたせることは困難であるため、この傾向が一般的なものであるかどうかはわからない。

## 5 まとめ

検索結果の文書を効率的にブラウズするための動的抄録提示インタフェースの評価実験を行った。文書の要不要の判定時間は、原文のみを提示した場合に比べ80%程度に短縮されることがわかった。この代償として、平均では判定再現率及び適合率が原文のみを提示した場合の90%程度に低下してしまう。しかし、内容の深い理解を必要としない比較的容易な文書の選定問題においては判定の質を保持することができた。今後も、高い再現率及び適合率を実現する検索手法と共に、効果的な検索結果の提示の仕方を探求していく予定である。

謝辞 本研究を進める上で有益な議論をして頂いた(株)東芝 研究開発センターの竹林洋一、平川秀樹、小野顕司、本谷秀堅、伊藤清司、山中太市郎、梶浦正浩、酒匂孝之諸氏、(株)東芝 東京システムセンターの武田公人、伊藤悦雄、野村浩一諸氏に感謝致します。

## 参考文献

- [安藤 94] 安藤他：“新聞記事からの情報抽出システム — 指定情報の抽出と多言語文章による提示 一,” 第8回人工知能学会全国大会 24-3, pp671-674, 1994.
- [三池 93] 三池他：“文書の構造解析に基づく文書情報検索,” 情処学会情報学基礎研究会 FI-31-6, pp39-46, 1993.
- [三輪 92] 三輪真木子：“データベースサーチャーの視点,” 情報処理 Vol.33 No.10, pp1162-1170, 1992.
- [住田 93] 住田他：“対話的文書検索のための文書構造解析,” 情処学会自然言語処理研究会 NL-97-11, pp71-78, 1993.
- [住田 94] 住田他：“対話的抄録生成機能を持つ文書検索システム,” 情処学会ヒューマンインタフェース研究会 HI-52-3, pp17-24, 1994.
- [Ingwersen92] Ingwersen, P.: “Information Retrieval Interaction,” Taylor Graham, London, 1992.
- [Ono94] Ono, K., Sumita, K. and Miike, S.: “Abstract Generation Based on Rhetorical Structure Extraction,” *Proc. COLING '94*, pp343-348, 1994.
- [Miike94] Miike, S., Itoh, E., Ono, K. and Sumita, K.: “A Full-Text Retrieval System with a Dynamic Abstract Generation Function,” *Proc. SIGIR '94*, pp.152-161, 1994.