

SuperSQL クエリ作成支援ツール

近藤 寛一[†] 遠山 元道^{††}

^{††} 慶應義塾大学理工学部情報工学科 〒 223-8522 神奈川県横浜市港北区日吉 3-14-1

E-mail: [†]hiro@db.ics.keio.ac.jp, ^{††}toyama@ics.keio.ac.jp

あらまし 近年の高度情報化社会の拡大に伴い、様々な情報が関係データベースに格納されており、その情報を有益に利用したいというニーズも高まっている。慶應義塾大学遠山研究室で開発、研究されている SuperSQL は関係データベースの出力結果を構造化し、多様なレイアウト表現を指定することを可能とした SQL の拡張言語である。これによりユーザは SuperSQL クエリを指定するだけで、欲しい情報を多様なレイアウトを考慮した HTML などのメディアデータとして直接得ることが出来るようになった。しかしながら、多様なレイアウトを実現するにはクエリ自体も複雑なものとならざるを得ず、SuperSQL クエリを書くという作業自体が複雑化してしまうことも少なくなかった。そこで本研究ではこの SuperSQL クエリ生成を支援する GUI ツールを提案し、そのプロトタイプの実装と評価を行う。

キーワード SuperSQL, マルチメディア処理, 問い合わせ処理

Graphical querying interface for SuperSQL

Hirokazu KONDO[†] and Motomichi TOYAMA^{††}

^{††} Department of Information and Computer Science, Faculty of Science and Technology,
Keio University

Hiyoshi3-14-1, Kouhoku-ku, Yokohama-shi, Kanagawa, 223-8522 Japan

E-mail: [†]hiro@db.ics.keio.ac.jp, ^{††}toyama@ics.keio.ac.jp

Abstract Recently, with expansion of a highly information society, various information is stored in the relational database and the needs of making good use of the information are also increasing. SuperSQL is the SQL extensible language, which can specify output structure and various layout expression. So the user only specify a SuperSQL query, he can acquire media data which allow various layout directly. But it needs complex query to specify various layout expression, writing SuperSQL query become difficult. In this paper, we develop a tool which support write a SuperSQL query, and evaluate it.

Key words SuperSQL, Multimedia, Query Processing

1. ま え が き

あらゆる情報がデジタルデータで飛び交う今日、関係データベースは様々な情報システムの中核となり多くの重要な情報を蓄積している。それゆえに関係データベースに蓄積された情報をできるだけ簡単に、印刷したり、HTML ファイルとして WWW のブラウザから利用したり等と、有益に活用したいというニーズも拡大している。

SuperSQL は関係データベースの出力結果を構造化し多様なレイアウト表現を可能とする SQL の拡張言語である。これによりユーザは SuperSQL クエリを指定するだけで関係データベースの情報を多様なメディアデータとして直接得ることが可能となった [3]。さらに近年の SuperSQL の研究によって、出力可能なメディアの種類が増加するとともに、得られるメディア

データの品質も大きく向上した。またより多様で実用的なレイアウトの指定 [5] なども進められており、得られるメディアデータの実用性は大きくなっている。

しかしながら、多様なレイアウトを実現するには SuperSQL クエリ自体も複雑化せざるを得ず、SuperSQL クエリを書くという作業自体が一般ユーザにとって難易度の高いものとなってしまった。そこで本研究では煩雑になりユーザの負担となってしまった SuperSQL クエリの作成を支援するツールを提案し、そのプロトタイプの実装と評価を行った。評価の結果ユーザの負担となっていた SuperSQL のクエリ作成時の問題点の多くを、本ツールの使用によって軽減できることが示された。

以下、2 章において SuperSQL について説明し、3 章において従来の技法の問題点を述べ、本ツールの必要性を示す。4 章において本ツールの概要を、5 章でその中心となるレイアウト

指定について説明する。6章において本研究の評価・検討を行い、最後に結論を述べる。

2. SuperSQL

本章では本ツールの母体となる SuperSQL について簡単に述べる。前述の通り、SuperSQL は SQL を拡張したメディアデータへの直接出力を実現する言語である [1] [2]。その質問文は SQL の SELECT 句を GENERATE< media >< TFE > の構文を持つ GENERATE 句で置き換えたものである。ここで < media > は出力媒体を示し、HTML、XML、Excel、 \LaTeX 、PDF などの指定ができる。また < TFE > はターゲットリストの拡張である Target Form Expression を表し、結合子、反復子などのレイアウト指定演算子を持つ一種の式である。

2.1 結合子

結合子はデータベースから得られたデータをどの方向(次元)に結合するかを指定する演算子であり、以下の3種類がある。括弧内はクエリ中の演算子を示している。

- 水平結合子 (,)

データを横に結合して出力。

例: Name, Tel

name	tel
------	-----

- 垂直結合子 (!)

データを縦に結合して出力。

例: Name! Tel

name
tel

- 深度結合子 (%)

データを3次元方法へ結合。出力がHTMLならばリンクとなる。

例: Name % Tel

name

 →

tel

2.2 反復子

反復子は指定する方向に、データベースの値があるだけ繰り返して表示する。また反復子はただ構造を指定するだけでなく、そのネストの関係によって属性間の関連を指定できる。例えば

[科目名]! , [学籍番号]! , [評点]!

とした場合には各属性間に関連はなく、単に各々の一覧が表示されるだけである。一方、ネストを利用して

[科目名 ! [学籍番号 , 評点]]!

とした場合には、その科目毎に学籍番号と評点の一覧が表示されるといったように、属性間の関連が指定される。以下、その種類について述べる。

- 水平反復子 ([,])

データインスタンスがある限り、その属性のデータを横に繰り返して表示する。

例: [Name],

name1	name2	...	name10
-------	-------	-----	--------

- 垂直反復子 ([!])

データインスタンスがある限り、その属性のデータを縦に繰り返して表示する。

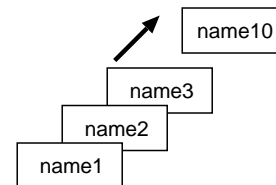
例: [Name]!

name1
name2
...
name10

- 深度反復子 ([] %)

データインスタンスがある限り、その属性のデータを奥行き方向(例:HTMLではリンク、PDFではページ変え)に繰り返して表示。

例: [Name]%



2.3 装飾子

SuperSQLでは関係データベースより抽出された情報に、文字サイズ、文字スタイル、横幅などの情報を付加できる。これらは装飾演算子(@)によって指定することが可能である。

<属性名>@[<装飾指定>]

装飾指定は”装飾子の名称 = その内容”として指定する。複数指定するときは各々を”,”で区切る。

2.4 関数

SuperSQLでは様々な関数が定義されていて、以下のようなフォーマットで一般の属性と同じようにTFE部で指定できる。

<関数名>(引数)

データベースに格納されたデータではない一般的な文字列を表示する verb 関数や、データベースに格納されたデータが画像へのパスであることを示し該当する画像を表示させる imagefile 関数などの他、出力先メディアに特有のものも含め多数が定義されている。

3. SuperSQL クエリ作成時の問題点

SuperSQLではクエリを書くだけで関係データベースの出力を複雑なメディアデータとすることを可能としたが、複雑な出力を求める際にはクエリ自体も長く複雑なものになってしまう。本章ではSuperSQLクエリを作成する際に、特にユーザーにとって負担となっていた点について検証する。

3.1 クエリ入力時の問題点

SuperSQLクエリを手作業により作成する際に、ユーザーの負担として考えられるのは以下のような点である。

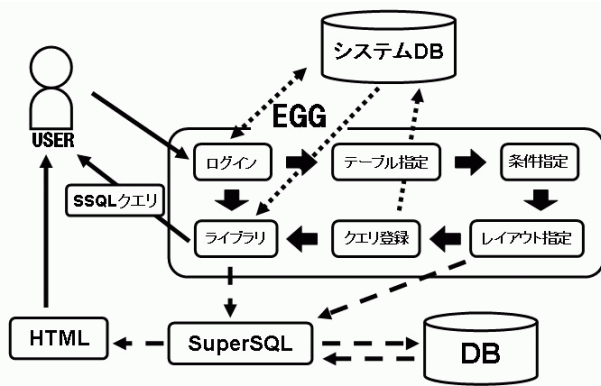


図 1 システムフロー

- スペルミス
- 括弧の対応間違い (括弧の閉じ忘れなど)
- 括弧の書き間違い (大括弧と中括弧)
- 生成されるメディアデータをイメージしにくい
- 構文を良く理解する必要がある
- エラー発生時に該当箇所を特定するのが困難

これらの問題はクエリが長くて複雑になればなるほど、より大きな問題としてユーザの負担となる。手作業の場合、クエリが長い程タイプミスや括弧の種類や対応付けを間違える可能性が増大するためである。さらに、これらにより実際にエラーが発生した場合にもエラー発生箇所を特定することは非常に困難で、ユーザにとっては大きな負担となっている。

3.2 その他の問題点

ユーザが SuperSQL クエリから作成されるメディアデータをイメージすることは困難であり、苦勞してクエリを作成しても出力されるメディアデータのレイアウトが意図したものと異なることも少なくない。また、特に装飾子の指定時などにはわずかな調整を繰り返し行うことが多くある。このような場合には、クエリの修正、SuperSQL の実行、作成されたファイルの確認、といった作業を意図するデータが得られるまで繰り返し行う必要があった。

また、多種のクエリを作成した場合には自分でうまく管理しないと、どのクエリがどのようなメディアデータを出力するものなのかわからなくなってしまうことも多い。特に、間が大きく空いた場合やデータを複数人で管理する場合などではクエリを効率良く再利用できれば作業効率も上がると考えられる。

4. システム概要

本研究では前章で提示した問題を解決とユーザ負担の軽減、初心者であっても SuperSQL を利用できるようにすることを目的として、GUI による SuperSQL クエリ作成支援ツールの開発を行った。より多くのユーザが手軽に利用できるシステムとすることを目指し、Web アプリケーションとして PHP により実装した。本章ではシステムの基本的な流れと、特徴的なステップの概要について述べる。

4.1 システムフロー

システムの基本的な流れは図 1 のようになる。なお、本研

究では本システムを EGG(Easy SuperSQL query Generator with GUI) と名付けた。

• ログイン

あらかじめ利用するユーザやデータベースをシステムデータベースに登録しておき、それらの情報を用いてシステムにログインする。

• テーブル指定

ログインしたデータベースに存在するテーブルの一覧が表示される。ユーザはそのデータ内容を確認しながら、実現したいメディアデータの作成に必要なテーブルを選択する。

• 条件指定

指定したテーブルに対して、SQL の WHERE 句以下にあたるクエリの条件部分の指定を行う。

• レイアウト指定 (5. 章)

レイアウトイメージ (5.1) を見ながら、実際にメディアデータのレイアウトを作成していく。同時にユーザの指定したレイアウトイメージを実現する SuperSQL クエリが EGG により自動的に生成される。

• クエリ登録・ライブラリ

完成したクエリはコメントと共にライブラリに登録し管理される。ユーザはライブラリから必要となるクエリを選択し適宜ダウンロードして利用する。

5. レイアウト指定

レイアウト指定部の画面イメージは図 2 のようになる。画面は出力可能属性表示部、レイアウトイメージ表示部、SuperSQL クエリ表示部の三つの部分から構成される。レイアウト指定の基本的な流れは以下ようになる。

- (1) 出力可能属性表示部からユーザの意図するレイアウトに必要な属性を選択する
- (2) 選択した属性を現在のレイアウトイメージに対してどのように反映させるか (データの反復方向、既出の属性との相対関係、位置関係) を指定する

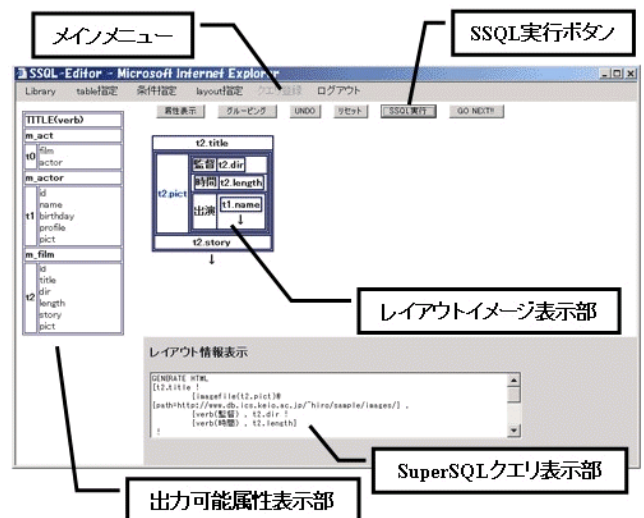


図 2 レイアウト指定画面

(3) 指定内容がレイアウトイメージに反映され、同時に SuperSQL クエリ表示部にクエリが生成、表示される

(4) 上記 (1)~(3) の繰り返し、グルーピング動作、属性名クリックからの装飾子指定等によって、意図するレイアウトイメージを作成する

レイアウトイメージに対する全ての操作は、システムによって動的にクエリに変換され SuperSQL クエリ表示部に反映される。また、随時 SuperSQL 実行ボタンをクリックすることで、その時点でのクエリで SuperSQL を実行しそれにより生成されるメディアデータの確認を行うことができる。

5.1 レイアウトイメージ

レイアウトイメージは本研究で導入した SuperSQL クエリのレイアウト情報を視覚的に表現したものである。このレイアウトイメージの導入によって、ユーザは生成されるメディアデータのレイアウトを視覚的にイメージしながら作業を進めることが可能となった。

レイアウトイメージ例を図 3 に示す。図 a) がレイアウトイメージを示したものであり、属性を示す四角とその反復方向を示す矢印から構成される。この例は「属性 name を縦方向に反復して表示」という意味を示し、生成されるメディアデータは同図 b) のようになる。両図を比較するとレイアウトイメージが出力結果を直感的に表現できていることがわかる。

新たな属性をレイアウトイメージに反映させていく際に指定する項目とそれに対して指定可能な値を表 1 に示す。最初の属性に対しては反復方向のみを、以降の属性に対しては表中の項目全てを指定して、レイアウトイメージを作成する。これら

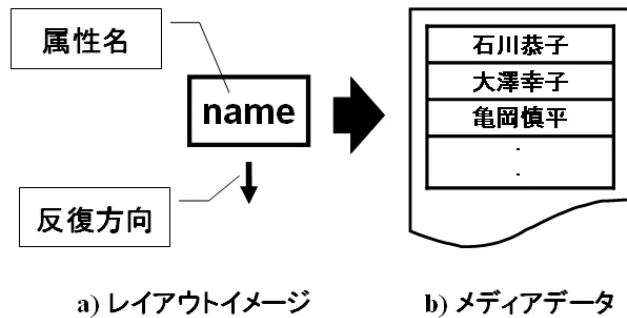


図 3 レイアウトイメージ例

表 1 属性に指定可能な値

項目	指定可能な値
相対関係	一対一 直前に指定した属性に対して一対一の関係がある場合
	一対多 直前に指定した属性に対して一対多の関係がある場合
	リンク関係 直前に指定した値をクリックしたリンク先に属性を配置する場合
連結方向	縦方向 指定済みの属性に対して縦に連結する
	横方向 指定済みの属性に対して横方向に連結する
反復方向	相対関係で一対多を指定した場合のみ指定する
	縦方向 多値属性値を縦方向に反復して表示 横方向 多値属性値を横方向に反復して表示

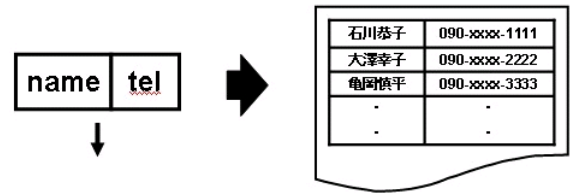


図 4 一対一：横方向

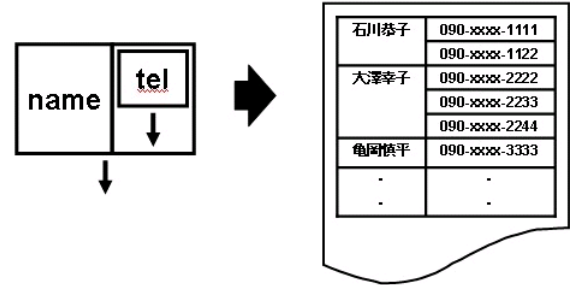


図 5 一対多：横方向

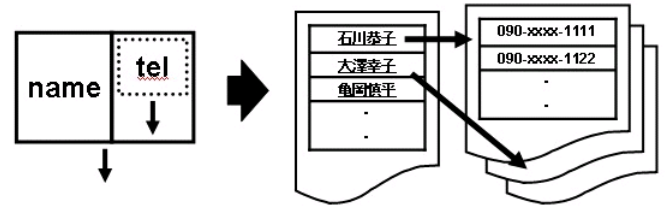


図 6 リンク：縦方向

の指定によってレイアウトイメージを生成した例と、それより生成されるメディアデータを図 4, 5, 6 に示す。それぞれ図 3 に属性を追加したものである。一対多を指定した場合には、その前に指定した属性によってグルーピングして出力される。リンクの場合にはリンク先に表示される属性が破線の四角によって表現される。このようにレイアウトイメージを導入することで、実際に生成されるメディアデータを直感的にイメージすることが可能となった。

また、レイアウトイメージには属性だけでなく任意の文字列を使用することもできる。これは SuperSQL の Verb 関数の機能によるものであるが、これを利用するには出力可能属性表示部最上部から属性ではなく、Verb を選択する。任意の文字列について指定する項目も属性の時と同じものである。

各属性に対して修飾子を指定する場合には、レイアウトイメージ上の属性名をクリックする。これにより装飾子指定ウィンドウが開き、ここで装飾子を指定することが出来る。装飾子を指定されている属性は、レイアウトイメージ上で異なる色で表示され視覚的に確認できる。

5.2 グルーピング

SuperSQL では結合子の種類が変わった際、表示済みの属性性を一つの塊として結合される。例えば図 7 左の図に対し、縦方向の結合を行った場合には同中央図のようになってしまう。ここで同右図が実際にユーザの得たいレイアウトである場合、SuperSQL では t0.b から t0.c ままで中括弧でくくる必要

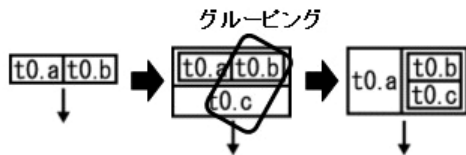


図7 グルーピング概念

がある。

これを本システムで実現するのがグルーピングである。この場合には t0.b から t0.c までをグルーピングすると図7右図のようなレイアウトイメージを得られることになる。このグルーピングを用いることで柔軟にレイアウトを表現することが可能となる。

5.3 SuperSQL 実行

実行ボタンをクリックするとその段階でのクエリで SuperSQL を実行し、作成される HTML ファイルを確認することができる。この HTML とレイアウトイメージを随時確認しながら、ユーザは好みのレイアウトを作成していく。なお、表示するデータ数が増大した場合にはこの実行にかかる時間も無視することができないため、SuperSQL 実行の際のデータ数を任意の数に制限してメディアデータを作成することが可能である。

6. 評価・検討

今回実装したツールの有用性を検証するために評価実験を行った。本章ではその方法を示し、結果について考察、検討していく。

6.1 実験

実験用データベースと徐々に難易度の上がると考えられるクエリ四題を用意する。この四題の概要は以下のようなものである。

- Q1. 単テーブルから列を表示するレイアウト
- Q2. 二つのテーブルを結合して、一対多属性含んだレイアウト
- Q3. 多対多の結合から、グルーピングを含んだレイアウト
- Q4. 多対多の結合、グルーピング、リンクを含んだレイアウト

被験者にはシステムの使用方法、データベースの内容を説明した後で、出力してもらいたいメディアデータのイメージを渡し、それを実現するクエリをシステムを利用して作成してもらった。被験者は SuperSQL を知っている五名、SuperSQL は知らないが SQL は知っている五名とした。また、SuperSQL を知っている被験者には手作業によるクエリ作成も行ってもらいシステムを利用した場合との比較も行った。

6.2 評価

6.2.1 SuperSQL クエリ構文知識レベルへの依存性

図8は SuperSQL を知っているユーザと知らないユーザのクエリ作成にかかった時間の平均値を比較したものである。図9、10は平均と各被験者の分布を示したものである。この比較によって本システムの SuperSQL のクエリ構文の知識レベルに対する依存性を測定した。

SuperSQL を知っている者、知らない者の場合で大きな違い

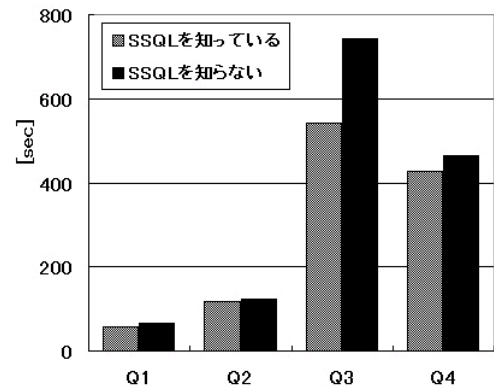


図8 SuperSQL 知識レベル別測定時間

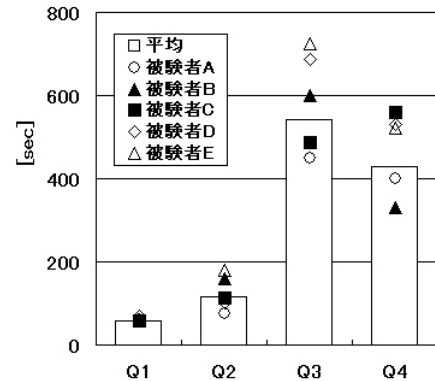


図9 SuperSQL を知っている人

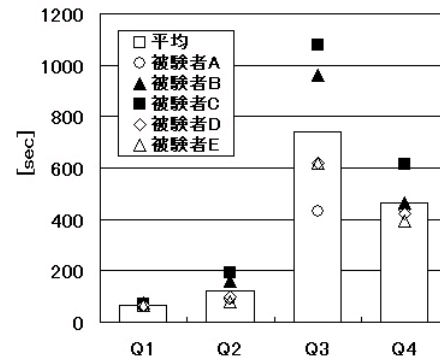


図10 SuperSQL を知らない人

は見られなかった。これにより本システムがユーザの SuperSQL のクエリ構文の知識レベルに依存することなくクエリ作成を支援出来たことがわかる。

6.2.2 SuperSQL クエリ生成に要する時間

図11は手作業とシステム利用時のクエリ作成にかかった時間を比較したものである。手作業時に比べて作業時間をかなり短縮できていることがわかる。今回の実験の被験者は皆 SuperSQL について非常に高い知識を持った者であったが、それでもこれだけの時間を短縮できた。対象を SuperSQL の知識があまりないものとすればより顕著な結果を得たことが推測できる。また手作業の場合には、クエリ自体にエラーが発生する可能性は否定することが出来ないために、そのデバッグに要する時間も推測することができない。ここに未知の時間を取ら

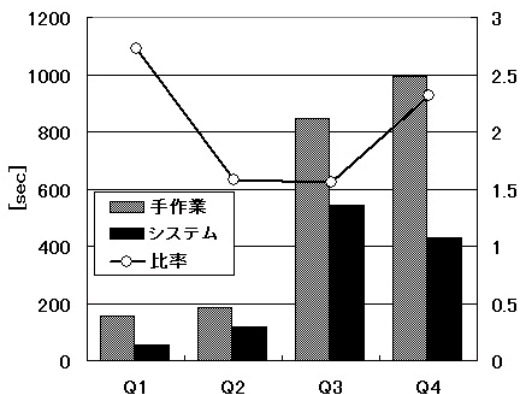


図 11 クエリ作成にかかる時間の比較

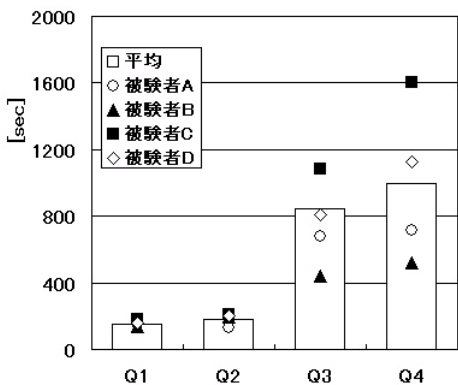


図 12 手作業によるクエリ作成時間

れる危険性のない点も本システムの利点である。

さらに、Q 3 と Q 4 は難易度的には Q 4 の方が上ではあり、手作業時には実行にかかる時間も増加しているが、システムを利用した場合の実行時間は短縮している。これはユーザがシステムに慣れたことによるものであろう。つまりユーザの熟練度に依存する部分がシステムでは大きいことがわかる。これによってユーザの熟練度に応じてシステム利用時の作業時間はより短縮していくことが期待される。

被験者の意見としては以下のような点が挙げられた。

- タイプミスなどによるクエリ自体のエラーは起きないので無駄な時間をデバッグに割かれることがない
- 完成データをイメージしながらできる
- 結果をリアルタイムで確認できるのがわかりやすい
- 括弧の対応関係などクエリを意識することなく作成できる
- 慣れるまでわかりにくい部分もある

6.2.3 一般ユーザによる評価

一般的なユーザの見地からの本システムを考察するために、学生 80 名に SuperSQL と本システムを利用してもらいアンケートを実施した。まず「SuperSQL を利用して不便を感じる点はどんなところか?」というアンケートを行った。集計結果は図 13 の様になった。結果を見てわかるように、被験者のあげる問題点のほとんどが SuperSQL のクエリ作成時に関するものであり、特に構文に関するものやエラーへの対所の困難さに

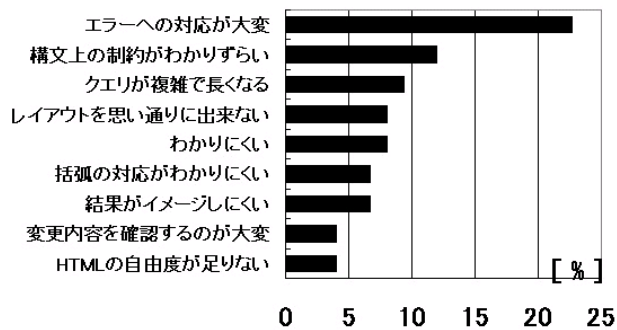


図 13 SuperSQL の問題点

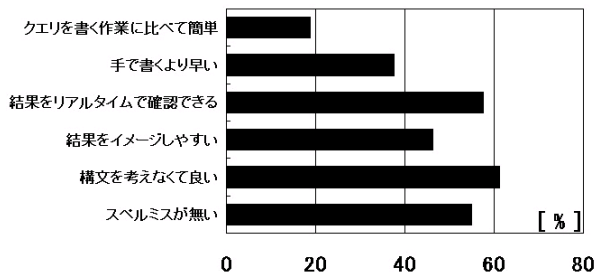


図 14 本システムの利点

についての意見が多く見られた。これは本システムの必要性を明らかにしたものであるとすることができる。

次に本システムを利用しての利点を挙げてもらったところ、結果は図 14 のようになった。SuperSQL を利用して不便と感じた点として挙げられた点のほとんどが本システムの利点として挙げられていることがわかる。特に、構文を意識する必要のない点や、結果をリアルタイムで参照することが出来る点については 50%以上の被験者が利点として挙げている。本システムがユーザの負担となっていた SuperSQL クエリ作成時の問題点の多くを解決することが出来たということが出来る。

6.3 検 討

実験の結果を見てわかる通り、SuperSQL を知っているユーザ知らないユーザを問わず入力支援を行うことができ、その際のクエリ作成による時間も短縮することが出来た。これにより既に SuperSQL を利用しているユーザにとっても、これから利用しようとするユーザにとっても本システムが有益であることを示した。

さらに従来の手作業による問題点のほとんどの部分も本システムを利用することで解決することが出来た。表 2 に本システムの利用により既存の問題点をどのように解決できたのかをまとめる。

しかしながら、慣れるまで操作に戸惑うユーザもみうけられた。操作の複雑さに対する批判的な意見も多くあった。よりわかりやすいインタフェースの研究が必要である。また、今回のシステムでは出力イメージを確認できるメディアを HTML に限定したが、その他のメディアファイルを作成しリアルタイムに内容を確認できるようにすればより有益なシステムとすることができる。

表 2 本システムによる問題点の解決

問題点	解決	方法
スベルミス	◎	ユーザが手作業によって SuperSQL クエリを入力することが無いのでスベルミスは発生しない
括弧の対応間違い	◎	ユーザが手作業で SuperSQL クエリを入力することが無いので対応は考えなくて良い
括弧の書き間違い	◎	同上
結果をイメージしにくい	○	レイアウトイメージを導入することで結果をイメージしながらの構築が可能となったが、リンクがある場合などイメージがわかりにくくなることもある
構文を良く理解する必要がある	◎	システムが自動で正しい SuperSQL クエリを生成するので構文を考える必要が無い
エラー箇所の特定が困難	◎	システムが自動で正しい SuperSQL クエリを生成するのでエラーは発生しない
変更内容を確認するのが困難	◎	常時 SuperSQL コマンドを実行して生成内容を確認することができる
クエリ管理が煩雑	◎	システムに作成した SuperSQL クエリを登録しておくことで、内容の確認、ファイル管理が容易になる

- [3] T. Seto, T. Nagafuji and M. Toyama, "Generating HTML Sources with TFE Enhanced SQL," *Proc. ACM Symposium on Applied Computing(SAC '97)*, pp. 96-105, 1997
- [4] 永藤, 瀬戸, 遠山, "TFE による HTML ソースの動的生成," *DEWS*, pp. 37-42, 1996
- [5] 岡部康矢, 遠山元道, "冗長なネストの除去による SuperSQL レイアウトの高品質化," *DBWS*, pp. 419-426, 2002

7. おわりに

本研究では SuperSQL のクエリ作成にかかるユーザの負担を軽減することを目的として、視覚的にレイアウトイメージを指定していくことでそのレイアウトを実現する SuperSQL クエリを自動的に生成するツールを実装した。これにより一般的なユーザであっても、直感的な操作のみで関係データベースの情報をメディアデータに変換するクエリを作成することが可能となった。

しかし、条件の指定や、非常に複雑なレイアウトの指定を行う場合などは、ユーザが戸惑う場面や、レイアウトイメージと実際に生成されるメディアデータの差異が大きくなってしまう場面も確認された。そこで、より直感的でわかりやすいユーザインタフェースの提供や、出力メディア特有の装飾子への対応などについて検討していく必要がある。

今後の情報化社会のさらなる発展により、データベースの必要性は強まり、蓄積される情報の量、重要性もより拡大していくと思われる。そのような中、有益な情報を簡単な操作でドキュメントとして得たり、HTML として提供したりしたいというニーズも当然拡大するだろう。それを実現する SuperSQL であるが、本ツールがその幅広い普及の為の窓口となることを期待したい。

文 献

- [1] SuperSQL: <http://ssql.db.ics.keio.ac.jp/>
- [2] Motomichi Toyama, "SuperSQL: An Extended SQL for Database Publishing and Presentation," *Proceedings of ACM SIGMOD '98 International Conference on Management of Data*, pp. 584-586, 1998