Browsing Large HTML Tables on Small Screens

Keishi Tajima[†] **Kaori Ohnishi**[‡]

† Kyoto University

‡ Japan Advanced Institute of Science and Technology

(Supported by IPA Exploratory Software Project)

ACM UIST'08 **at Monterey, CA USA**
22 Oct. 2008

School of Informatics
Kyoto University

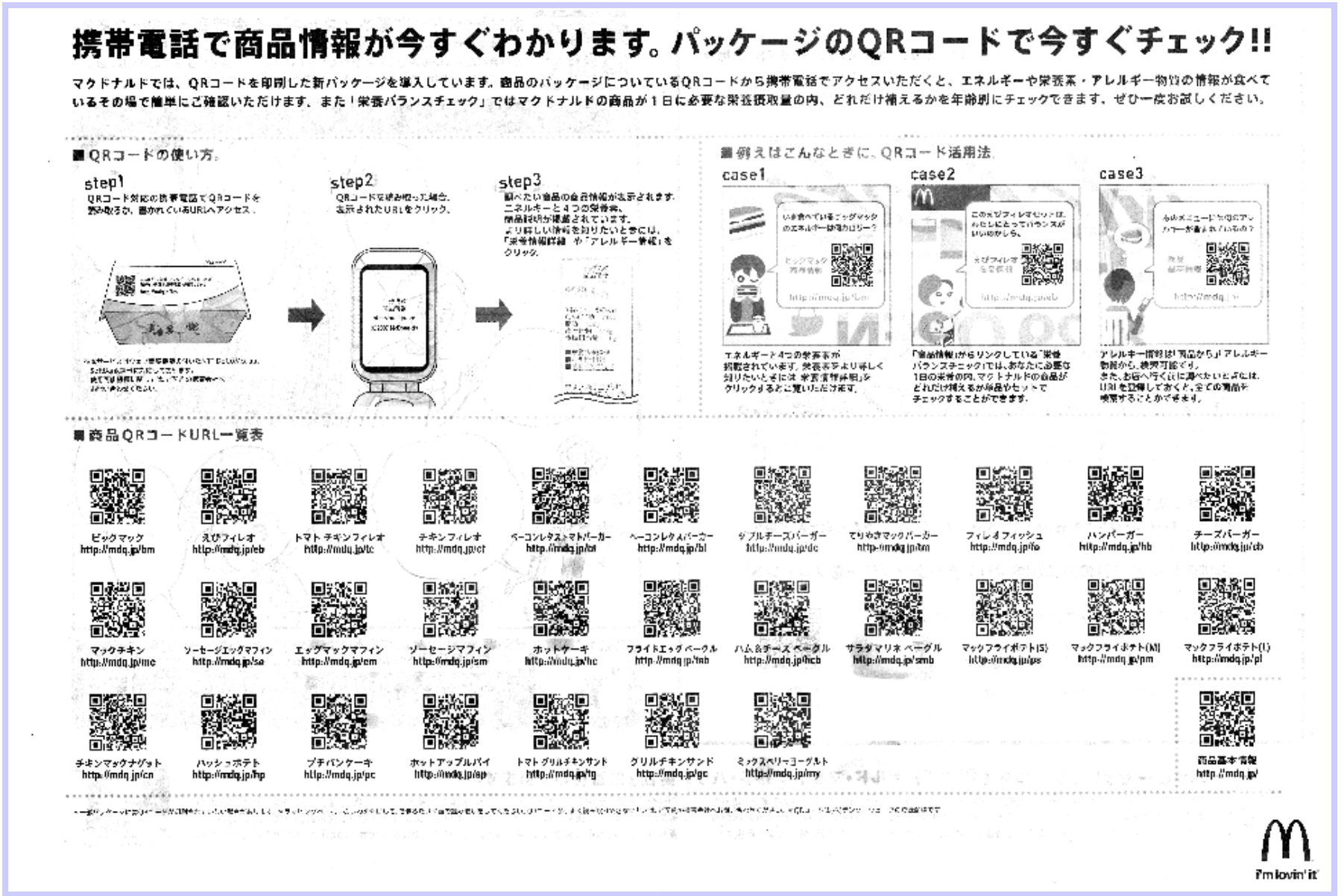
What is the Problem?

Small screen size is a problem in mobile Web access...

When is screen size really a problem?

- **text data** (e.g., latest news)
  - **sequential data** → you can read by sequential scrolling
  - **2-dimensional page layout** is the problem
- **maps, images**
  - inherently **2-dimensional** data
  - one of the most difficult data types on small screens
- **tables**
  - also **2-dimensional**
  - also one of the most difficult types on small screens

School of Informatics
Kyoto University



携帯電話で商品情報が今すぐわかります。パッケージのQRコードで今すぐチェック!!

マクドナルドでは、QRコードを印刷した新パッケージを導入しています。商品のパッケージについているQRコードから携帯電話でアクセスいただくと、エネルギーや栄養素・アレルギー物質の情報が食べているその場で簡単にご確認いただけます。また「栄養バランスチェック」ではマクドナルドの商品が1日に必要な栄養摂取量の内、どれだけ補えるかを年齢別にチェックできます。ぜひ一度お試しください。

■QRコードの使い方

step1
QRコード対応の携帯電話でQRコードを読み取るか、直接入力されているURLへアクセス。

step2
QRコードを読み取った場合、表示されたURLをクリック。

step3
調べたい商品の栄養情報が表示されます。
エネルギーと4つの栄養素、
商品詳細が掲載されています。
より詳しい情報を知りたいときには、
「栄養情報詳細」や「アレルギー情報」を
クリック。

■例えばこんなときに、QRコード活用法

case1
エネルギー4つの栄養素が掲載されています。栄養素をより詳しく知りたいときには 栄養情報詳細 をクリックするとご覧いただけます。

case2
「商品情報」からリンクしている「栄養バランスチェック」では、あなたに必要な1日の栄養の内、マクドナルドの商品がどれだけ補えるか商品をセットでチェックすることができます。

case3
アレルギー情報は「商品から」「アレルギー物質から」検索可能です。
また、お気に入り登録に進みたいときには、1回を登録しておくと、全ての商品を検索することができます。

■商品QRコードURL一覧表

ビックマック http://mdq.jp/bm	えびフィレオ http://mdq.jp/eb	トマト チキンフィレオ http://mdq.jp/tc	チキンフィレオ http://mdq.jp/cf	ベーコンレタス・トマトバーガー http://mdq.jp/bt	ベーコンレタスバーガー http://mdq.jp/bl	ダブルチーズバーガー http://mdq.jp/dc	てりやきマックバーガー http://mdq.jp/tm	フィレオフィッシュ http://mdq.jp/fo	ハンバーガー http://mdq.jp/hb	チーズバーガー http://mdq.jp/cb
マックチキン http://mdq.jp/mc	ソーセージエッグマフィン http://mdq.jp/se	エッグマックマフィン http://mdq.jp/em	ソーセージマフィン http://mdq.jp/sm	ホットケーキ http://mdq.jp/hc	フライドエッグ ベーグル http://mdq.jp/feb	ハム&チーズ ベーグル http://mdq.jp/hcb	サラダマリネ ベーグル http://mdq.jp/smb	マックフライポテト(S) http://mdq.jp/ps	マックフライポテト(M) http://mdq.jp/pm	マックフライポテト(L) http://mdq.jp/pl
チキンマックナゲット http://mdq.jp/cn	ハッシュポテト http://mdq.jp/hp	プチパンケーキ http://mdq.jp/pc	ホットアップルパイ http://mdq.jp/ap	トマト グリルチキンサンド http://mdq.jp/tg	グリルチキンサンド http://mdq.jp/gc	ミックスベリーヨーグルト http://mdq.jp/my				商品基本情報 http://mdq.jp/

I'm lovin' it

(a tray sheet, McDonald's Company (Japan), Ltd.)

School of Informatics
Kyoto University

What is the Problem?

Do we access tables from mobile phones so often?

⇓

Yes

- **train timetables**
- **movie schedules**
- **stock markets**

  ⋮

- **Many Web sites shows search results in tables.**

Tables are important in mobile Web access.

School of Informatics
Kyoto University

Goal of this Research

Tables are:

- **frequently accessed in mobile Web access**

  while

- **especially difficult to browse on small screens**

⇓

We develop interaction techniques for browsing HTML tables on small screens.

School of Informatics
Kyoto University

Basic Structure of Tables

Item	Cal	Sugar	Fat	···	Milk	Wheat	Egg
Hamburger	100	5.5	9	···	-	√	√
Cheeseburger	200	6.5	12	···	√	√	√
Potato(S)	230	0.0	11	···	-	√	-
Potato(M)	380	0.0	19	···			
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Gigaburger	540	8.8	29	···	-	√	√

- concisely represent two contexts of each cell by its horizontal and vertical positions.
- "cal of Hamburger is 100, sugar of Hamburger is 5.5, . . . , cal of Cheeseburger is 200, . . . "

School of Informatics
Kyoto University

Basic Structure of Tables

Item	Cal	Sugar	Fat	···	Milk	Wheat	Egg
Hamburger	250	5.5	9	···	-	√	√
Cheeseburger	300	6.5	12	···	√	√	√
Potato(S)	230	0.0	11	···	-	√	-
Potato(M)	380	0.0	19	···			
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Gigaburger	540	8.8	29	···	-	√	√

attribute names: describe types of values in each column

School of Informatics
Kyoto University

Basic Structure of Tables

Item	Cal	Sugar	Fat	···	Milk	Wheat	Egg
Hamburger	250	5.5	9	···	-	√	√
Cheeseburger	300	6.5	12	···	√	√	√
Potato(S)	230	0.0	11	···	-	√	-
Potato(M)	380	0.0	19	···			
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Gigaburger	540	8.8	29	···	-	√	√

attribute names: describe types of values in each column
key values: have unique values that briefly explain the meaning of each row

We call them headers.

School of Informatics
Kyoto University

Browsing Modes

We provide three browsing modes.

1. **Normal Mode (default):**
   - **ordinary representation + various useful functions**
2. **Record Mode:**
   - **Show each row/column in a record-like format**
   - **When we randomly move around, we may get lost.**
3. **Cell Mode:**
   - **Show each cell together with its row/column headers, so we never get lost.**

School of Informatics
Kyoto University

DEMO

Typical Tasks and Corresponding Operations

Read a cell at the intersection of a given row & column
⇒ record mode or cell mode
Read cells in a given row (or column)
⇒ record mode or cell mode
Read tables row by row (or column by column)
⇒ record mode or cell mode
Randomly move and read cells at various positions
⇒ cell mode
Compare two (or more) given rows (or columns)
⇒ hiding unnecessary rows/columns in normal mode
Find where in the table a given value appears
⇒ intra-page search and click the cell to see its headers in normal mode

School of Informatics
Kyoto University

Technical Issues

We need to detect logical structure in tables.

1. **Which cells are row/column headers?**
   (attribute names or key values)
2. **Semantics of composite cells?**
   - **In normal mode:**
     We need to know hierarchical attributes.
   - **In record/cell mode:**
     We decompose composite cells in accordance with their semantics.

School of Informatics
Kyoto University

Semantics of Composite Cells

Neighbors Sharing Same Values

Item	Cal	Sugar	Fat	···	Allergen		
					Milk	Wheat	···
Hamburger	150	2.5	5.8	···	-	-	···
Cheeseburger	160	2.5	6.0	···	-	-	···
Potato(S)	160	2.5	6.0	···	-	✓	···
Potato(M)	250	3.8	9.1	···			
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Gigaburger	280	4.3	13.0	···	-	-	···

School of Informatics
Kyoto University

Semantics of Composite Cells

Neighbors Sharing Same Values

Item	Cal	Sugar	Fat	···	Allergen		
					Milk	Wheat	···
Hamburger	150	2.5	5.8	···	-	-	···
Cheeseburger	160	2.5	6.0	···	-	-	···
Potato(S)	160	2.5	6.0	···	-	√	···
Potato(M)	250	3.8	9.1	···	-	√	···
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Gigaburger	280	4.3	13.0	···	-	-	···

Decompose and Copy the Contents

Semantics of Composite Cells

Cells Representing Aggregation of Neighboring Cells

Number of Professors

year	assist. prof.	assoc. prof.	prof. (non-tenure)	prof. (tenure)
2002	31	21	23	
2003	28	21	14	10
2004	30	19	13	12

- very rare
- indistinguishable without understanding the contents

School of Informatics
Kyoto University

Semantics of Composite Cells

Hierarchical Attributes

Item	Cal	Sugar	Fat	···	Allergen		
					Milk	Wheat	···
Hamburger	150	2.5	5.8	···	-	-	···
Cheeseburger	160	2.5	6.0	···	-	-	···
Potato(S)	160	2.5	6.0	···	-	√	···
Potato(M)	160	2.5	6.0	···	-	√	···
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Gigaburger	280	4.3	13.0	···	-	-	···

School of Informatics
Kyoto University

Semantics of Composite Cells

Hierarchical Attributes

Item	Cal	Sugar	Fat	···	Allergen	Allergen	···
Item	Cal	Sugar	Fat	···	Milk	Wheat	···
Hamburger	150	2.5	5.8	···	-	···	···
Cheeseburger	160	2.5	6.0	···	-	···	···
Potato(S)	160	2.5	6.0	···	-	√	···
Potato(M)	160	2.5	6.0	···	-	√	···
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Gigaburger	280	4.3	13.0	···	-	···	···

inappropriate decomposition...

Semantics of Composite Cells

Hierarchical Attributes

Item	Cal	Sugar	Fat	···	Allergen Milk	Allergen Wheat	··· ···
Hamburger	150	2.5	5.8	···	-	-	···
Cheeseburger	160	2.5	6.0	···	-	-	···
Potato(S)	160	2.5	6.0	···	-	√	···
Potato(M)	250	3.8	9.1	···	-	√	···
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
Gigaburger	280	4.3	13.0	···	-	-	···

Concatenate Multi-Level Attributes

School of Informatics
Kyoto University

Semantics of Composite Cells

Headings, Delimiters, Footers, etc.

Phone Directory

Name	ext.	Email
Accounting		
A. Smith	2850	smith
J. Doe	2851	jdoe
Human Resource		
T. Johnson	3010	tj
D. Adams	3011	adams

Do Not Decompose

School of Informatics
Kyoto University

Semantics of Composite Cells

Nested Tables

Aquarium Admission Fee

Category	Fee		
Adult			$13
Child			$4
Group	> 20	Adult	$11
		Child	$3
	> 100	Adult	$10
		Child	$2

Replace with a Link to a Separate Table

Automatic Detection of Logical Structure

Overview

We use:

- structure of composite cells
  - to detect attribute names
  - to detect semantics of composite cells
- uniqueness of values
  - to detect attribute names
  - to detect keys

Automatic Detection of Logical Structure

Step1: Attribute Detection Based on Composite Cells

- We defined valid table structure by a set of rules.
- We search the attribute/body border satisfying them.

<table>
<tr><th colspan="3">category</th><th>weight</th><th>max</th><th>min</th></tr>
<tr><td rowspan="6">mail</td><td rowspan="2">type 2</td><td>card</td><td>2g to 6g</td><td colspan="2" rowspan="2"></td></tr>
<tr><td>double postal</td><td>4g to 12g</td></tr>
<tr><td rowspan="2">type 1</td><td>regular</td><td>up to 50g</td><td></td><td></td></tr>
<tr><td>irregular</td><td>up to 4kg</td><td colspan="2" rowspan="4"></td></tr>
<tr><td colspan="2">type 3</td><td>up to 1kg</td></tr>
<tr><td colspan="2" rowspan="2">type 4</td><td>up to 1kg</td></tr>
<tr><td>up to 3kg</td></tr>
<tr><td rowspan="3">parcel</td><td colspan="2">normal/braille</td><td>up to 30kg</td><td colspan="2" rowspan="3"></td></tr>
<tr><td colspan="2">book parcel</td><td rowspan="2">up to 3kg</td></tr>
<tr><td colspan="2">for handicapped</td></tr>
</table>

School of Informatics
Kyoto University

Automatic Detection of Logical Structure

Step2: Attribute/Key Detection by Value Uniqueness

- We find a set of columns/rows that gives unique values.
- We give priority to those closer to top/left edge.

	year	first name	last name	age
⋆	freshman	David	Adams	18
⋆	freshman	David	Jones	19
⋆	sophomore	Mary	Adams	19
⋮	⋮	⋮	⋮	⋮

Automatic Detection of Logical Structure

Step3: Semantics of Composite Cells

- cells spanning the entire row/column
  → captions etc.

- composite cells in the attributes
  → structured attributes

- composite cells in the body without corresponding partitions in the header
  → nested tables

- the others
  → merged cells

School of Informatics
Kyoto University

Experiment

- **We collected 80 tables, and our method can detect correct headers for 76 of them (95%).**
- **Our method fails when:**
  - **headers appear only at the bottom or only at the right edge**
  - **tables have some illogical structure**
  - **it has some types of nested tables**

Future Work

1. Evaluation of the proposed browsing modes
   - Can users complete the task in shorter time?
   - Detected attributes/keys really help users?
   - Do we need record mode?
2. detection of logical tables described by `<div>` tags (e.g., Google)
3. Improvement of automatic detection methods

and more...