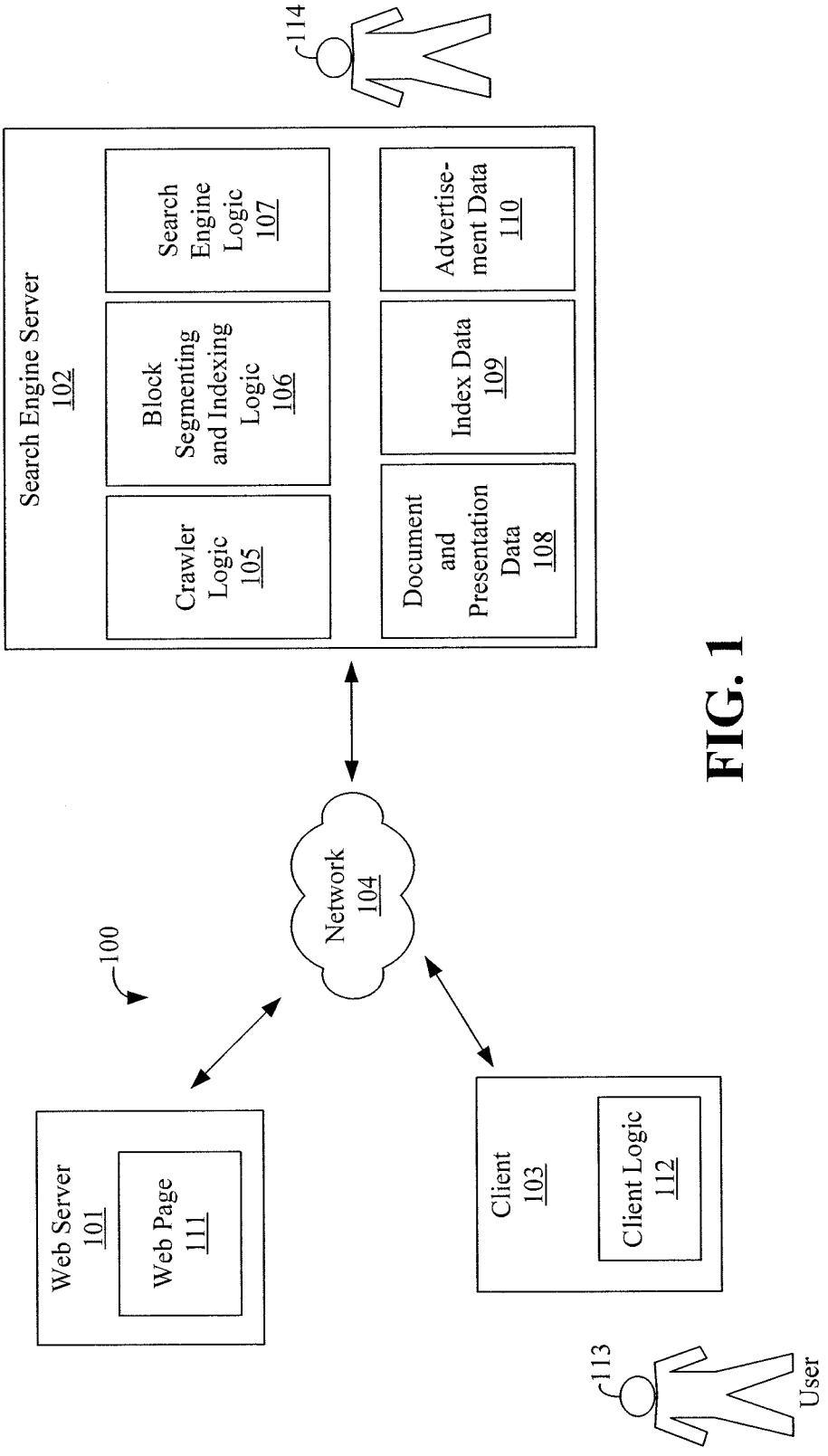

114
Search Engine Server
102
Search Engine Logic
107
Block Segmenting and Indexing Logic
106
Crawler Logic
105
Advertise-ment Data
110
Index Data
109
Document and Presentation Data
108
Network
104
100
Web Server
101
Web Page
111
Client
103
Client Logic
112
113
User

FIG. 1

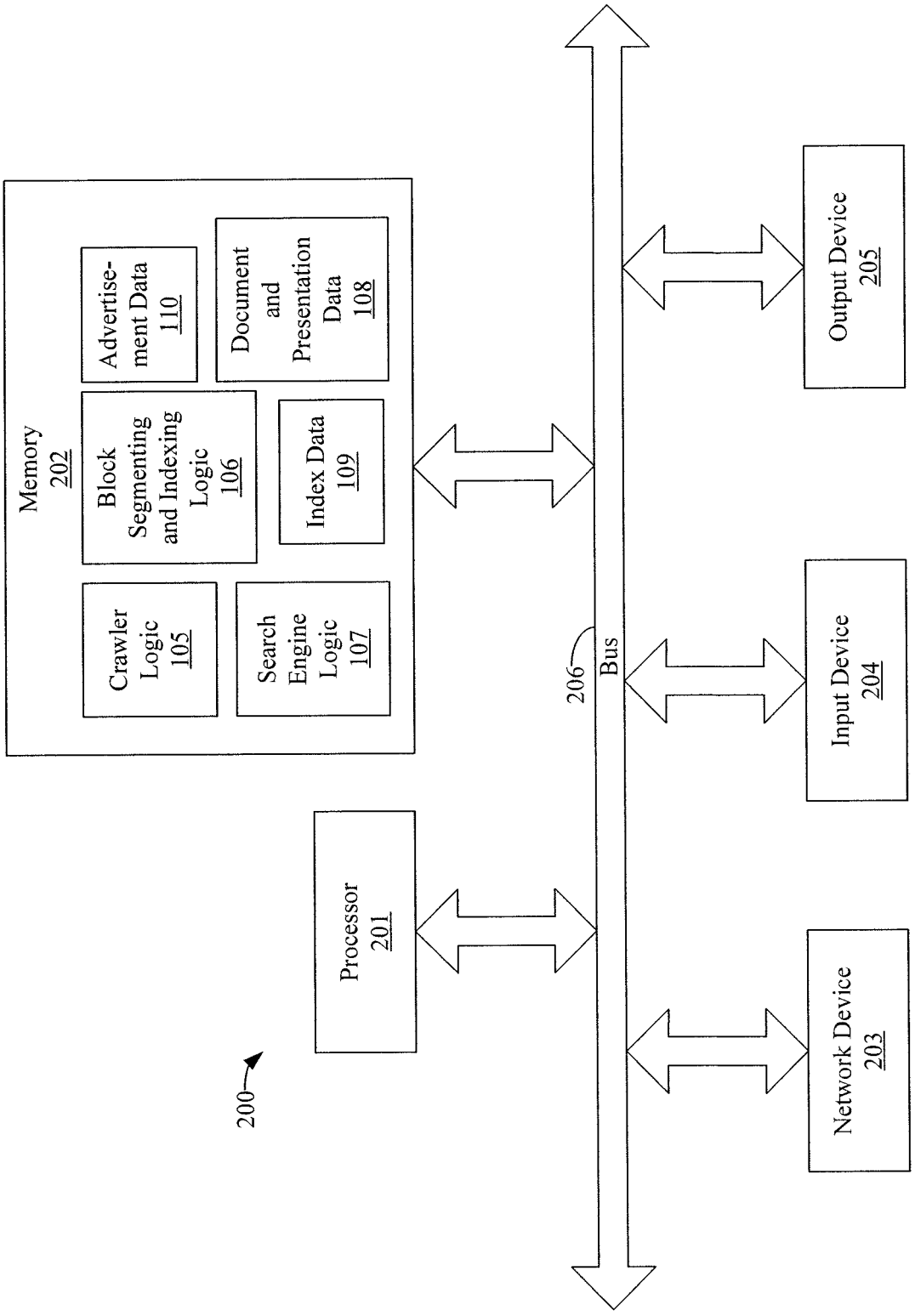
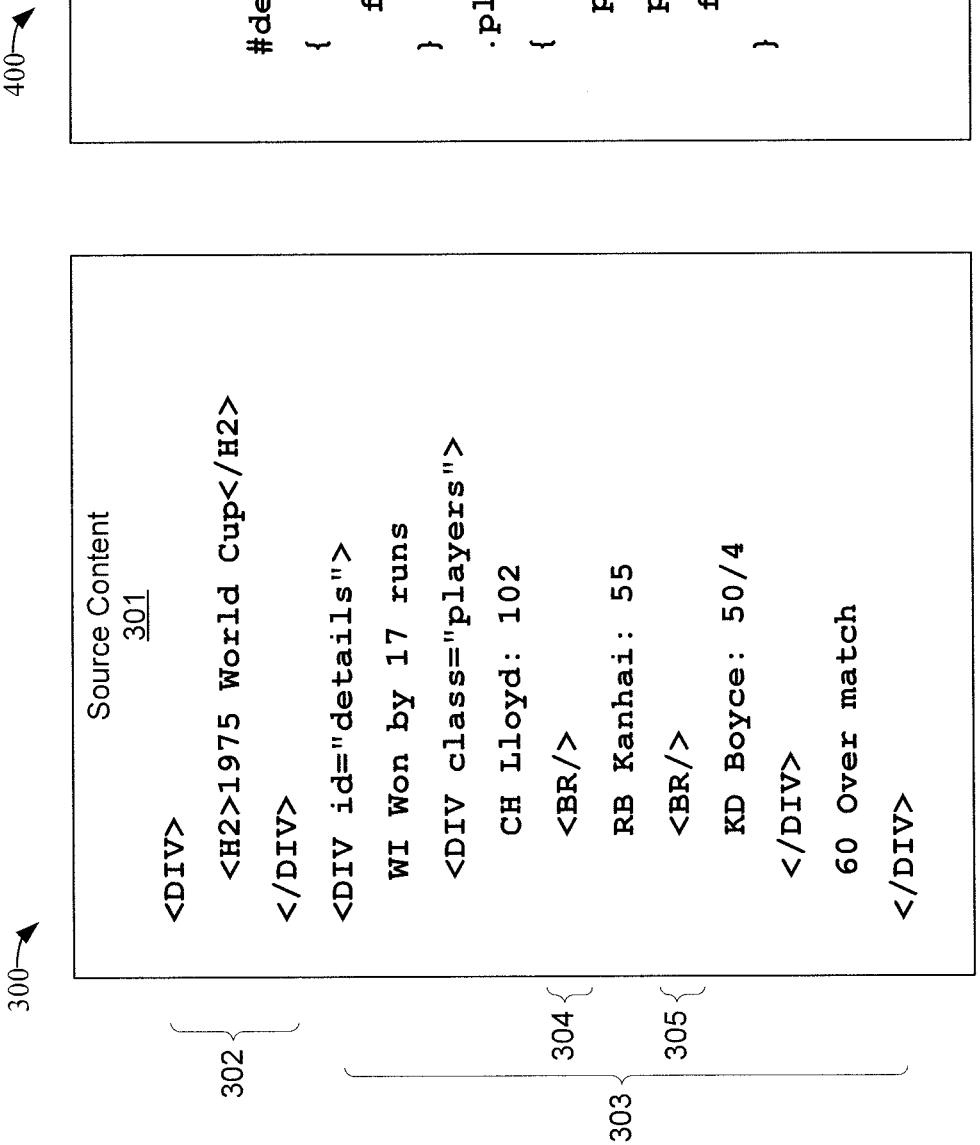

200
Processor 201
Memory 202
Crawler Logic 105
Search Engine Logic 107
Block Segmenting and Indexing Logic 106
Index Data 109
Advertise-ment Data 110
Document and Presentation Data 108
206
Bus
Network Device 203
Input Device 204
Output Device 205

FIG. 2


300
Source Content
301
<DIV>
<H2>1975 World Cup</H2>
</DIV>
<DIV id="details">
WI Won by 17 runs
<DIV class="players">
CH Lloyd: 102

RB Kanhai: 55

KD Boyce: 50/4
</DIV>
60 Over match
</DIV>
302
303
304
305

FIG. 3

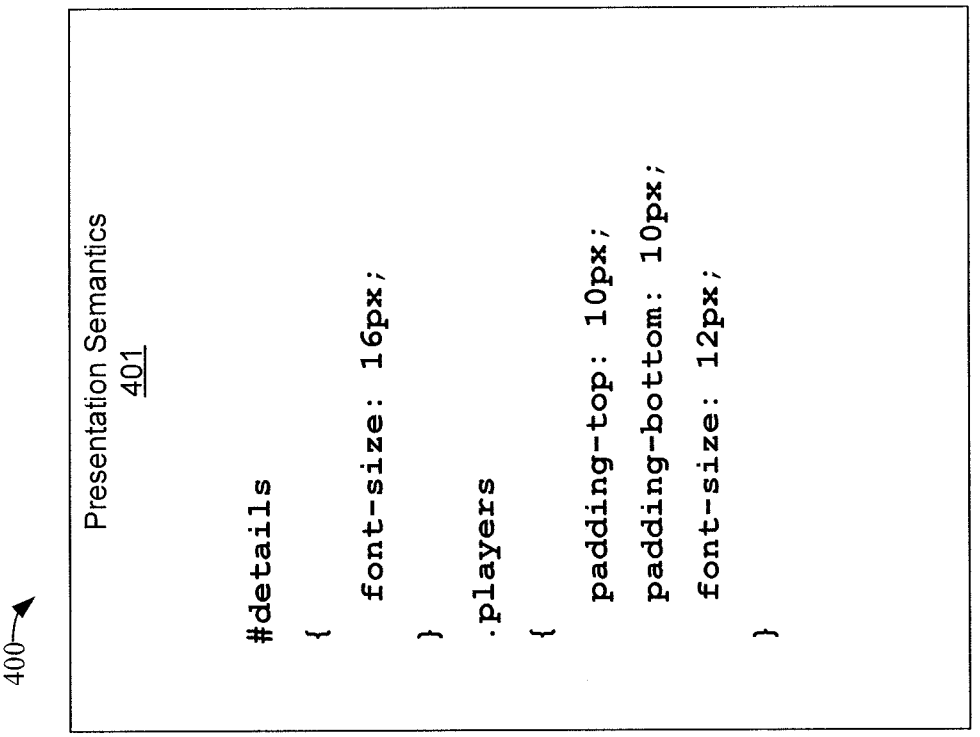

400
Presentation Semantics
401
#details
{
font-size: 16px;
}
.players
{
padding-top: 10px;
padding-bottom: 10px;
font-size: 12px;
}

FIG. 4

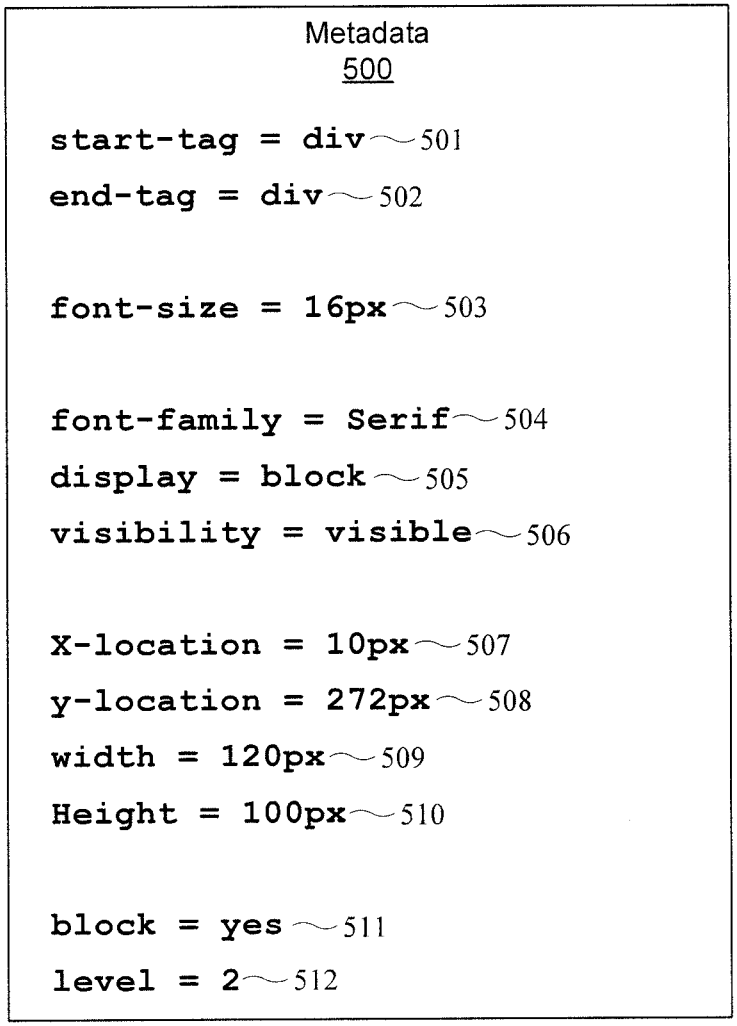


FIG. 5

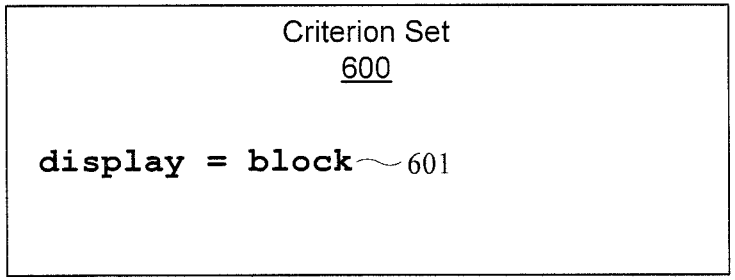


FIG. 6

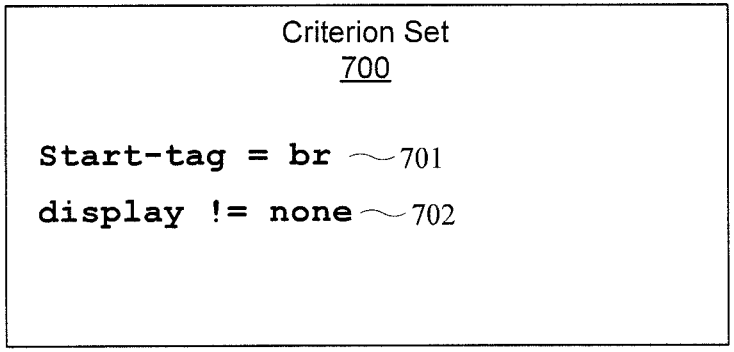


FIG. 7

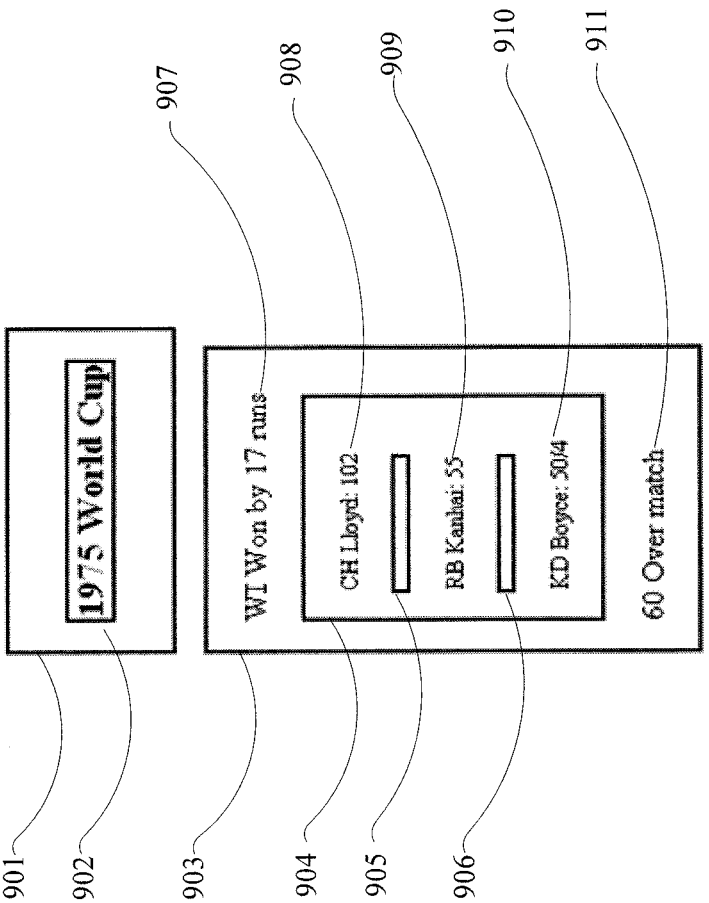

1975 World Cup
WI Won by 17 runs
CH Lloyd: 102
RB Kanhai: 55
KD Boyce: 504
60 Over match
901
902
903
904
905
906
907
908
909
910
911

FIG. 9

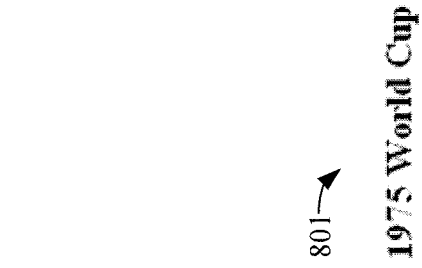

801
1975 World Cup
WI Won by 17 runs
CH Lloyd: 102
RB Kanhai: 55
KD Boyce: 504
60 Over match

FIG. 8

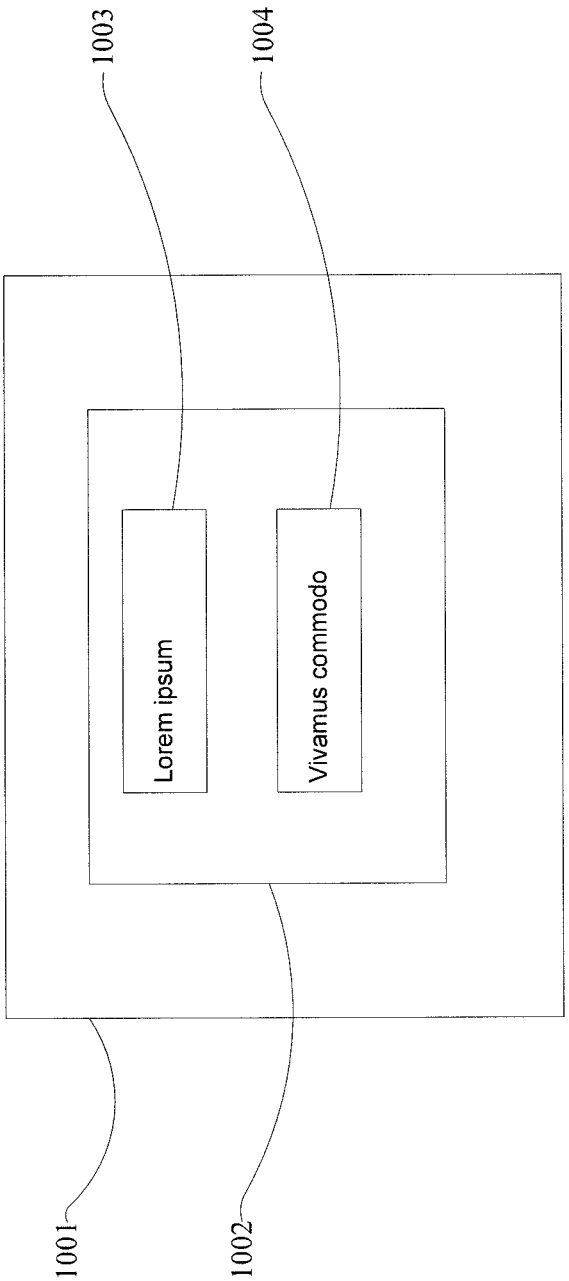


FIG. 10

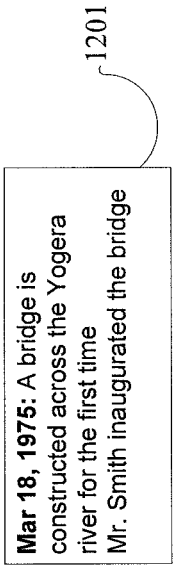

1201
Mar 18, 1975: A bridge is constructed across the Yogera river for the first time
Mr. Smith inaugurated the bridge

FIG. 12

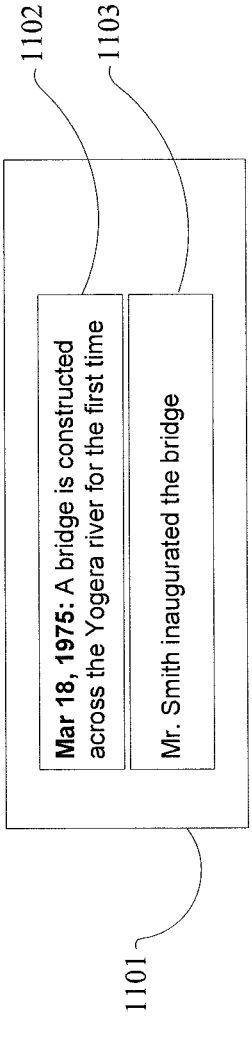

1102
1103
Mar 18, 1975: A bridge is constructed across the Yogera river for the first time
Mr. Smith inaugurated the bridge
1101

FIG. 11

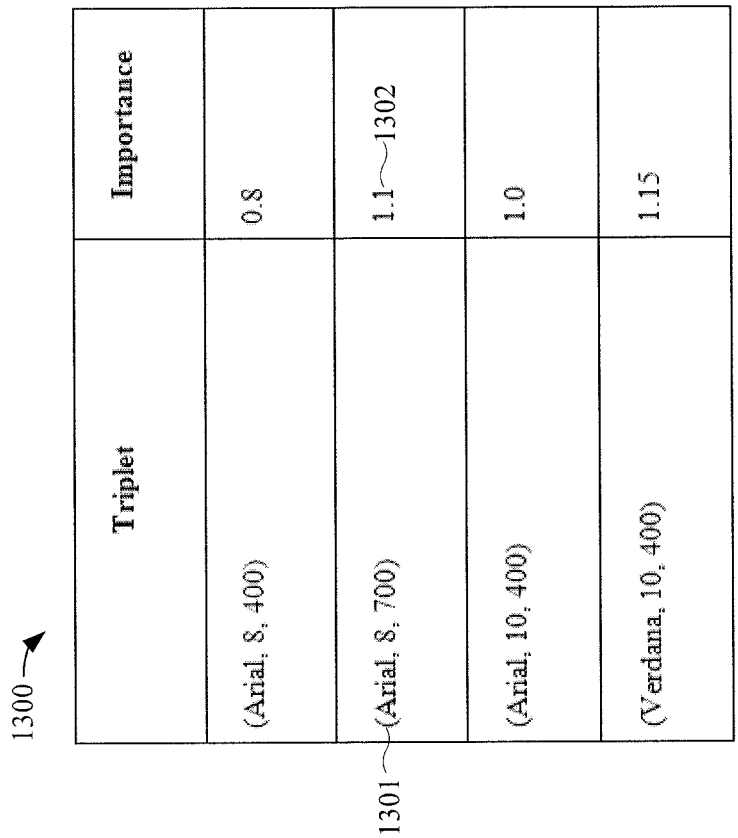

1300
Triplet
Importance
(Arial, 8, 400)
0.8
(Arial, 8, 700)
1.1
1302
1301
(Arial, 10, 400)
1.0
(Verdana, 10, 400)
1.15

FIG. 13

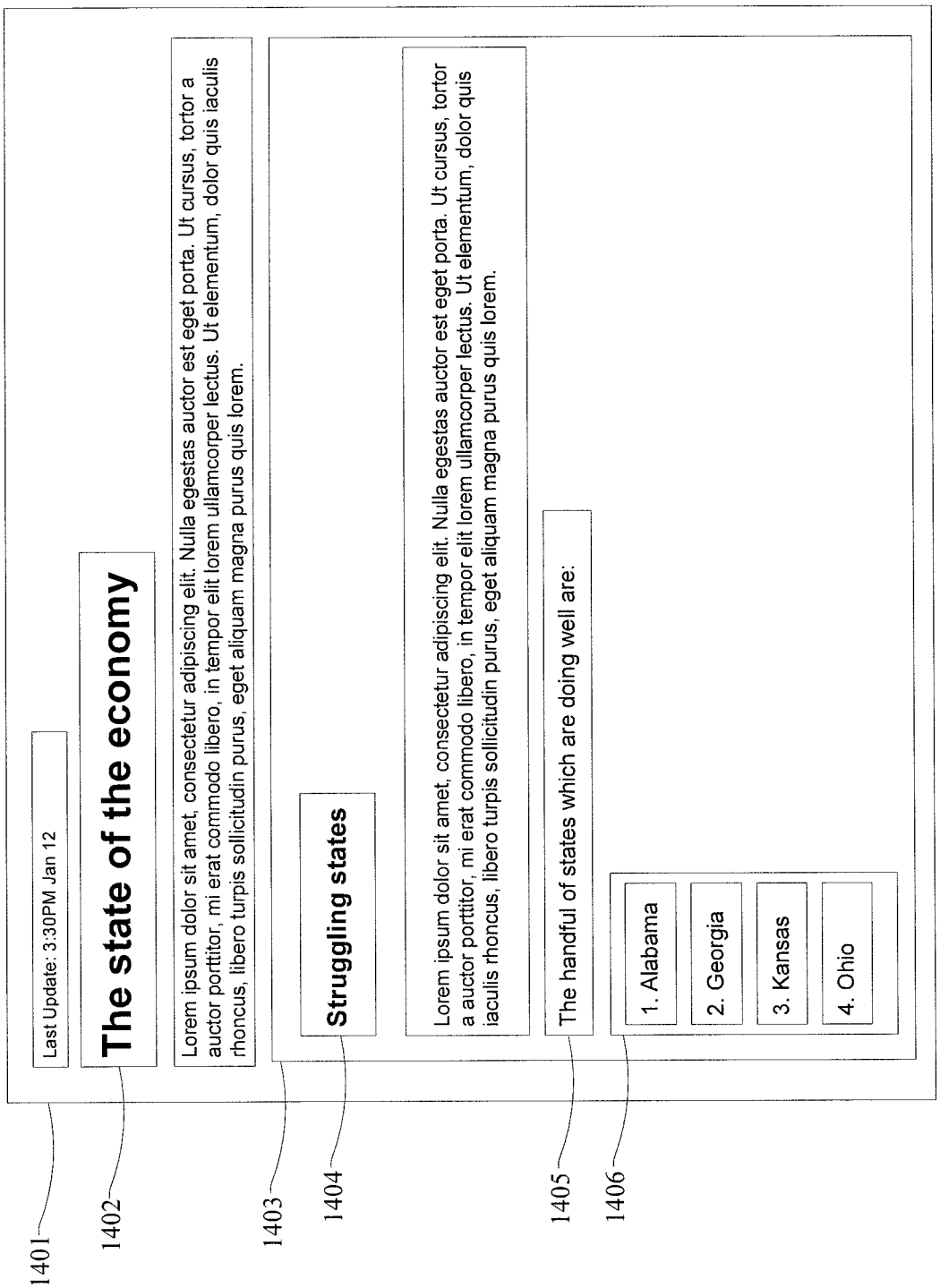

Last Update: 3:30PM Jan 12
The state of the economy
Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nulla egestas auctor est eget porta. Ut cursus, tortor a auctor porttitor, mi erat commodo libero, in tempor elit lorem ullamcorper lectus. Ut elementum, dolor quis iaculis rhoncus, libero turpis sollicitudin purus, eget aliquam magna purus quis lorem.
Struggling states
Lorem ipsum dolor sit amet, consectetur adipiscing elit. Nulla egestas auctor est eget porta. Ut cursus, tortor a auctor porttitor, mi erat commodo libero, in tempor elit lorem ullamcorper lectus. Ut elementum, dolor quis iaculis rhoncus, libero turpis sollicitudin purus, eget aliquam magna purus quis lorem.
The handful of states which are doing well are:
1. Alabama
2. Georgia
3. Kansas
4. Ohio
1401
1402
1403
1404
1405
1406

FIG. 14

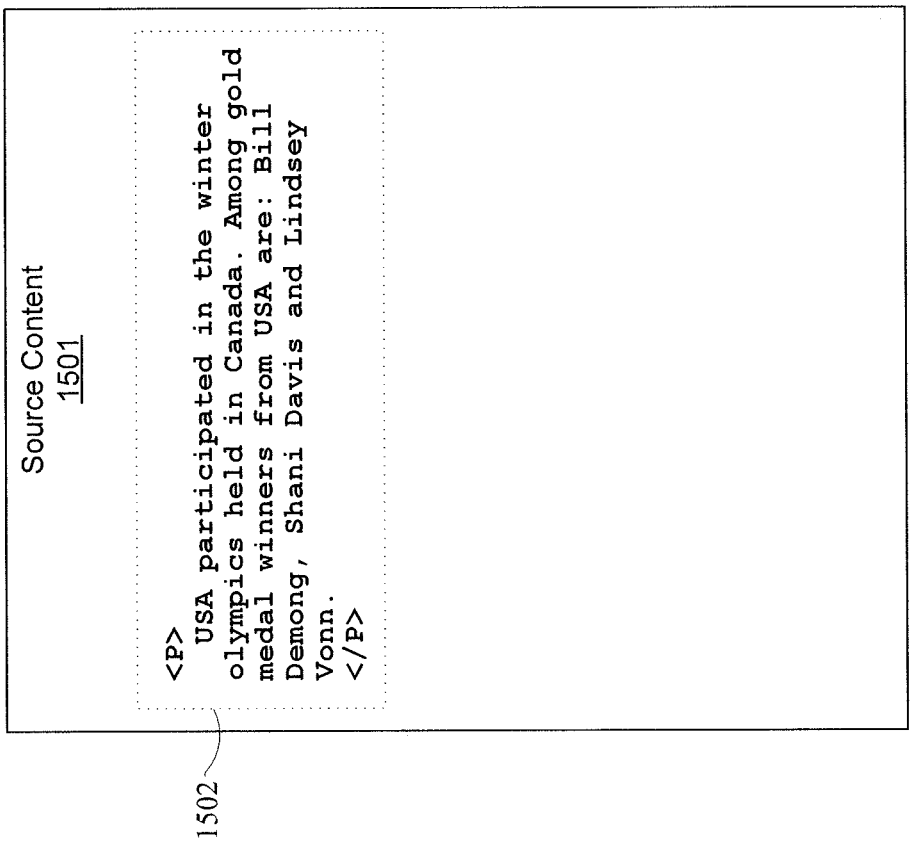

Source Content
1501
<P>
USA participated in the winter olympics held in Canada. Among gold medal winners from USA are: Bill Demong, Shani Davis and Lindsey Vonn.
</P>
1502

FIG. 15

Presentation Semantics
1701

```
.codeWords
{
  font-family: courier;
}
```

FIG. 17

Source Content
1601

```
<P>
  The <span class="codeWords">
text-decoration</span> property
allows certain text effects such as
underlining.
</P>
```

FIG. 16

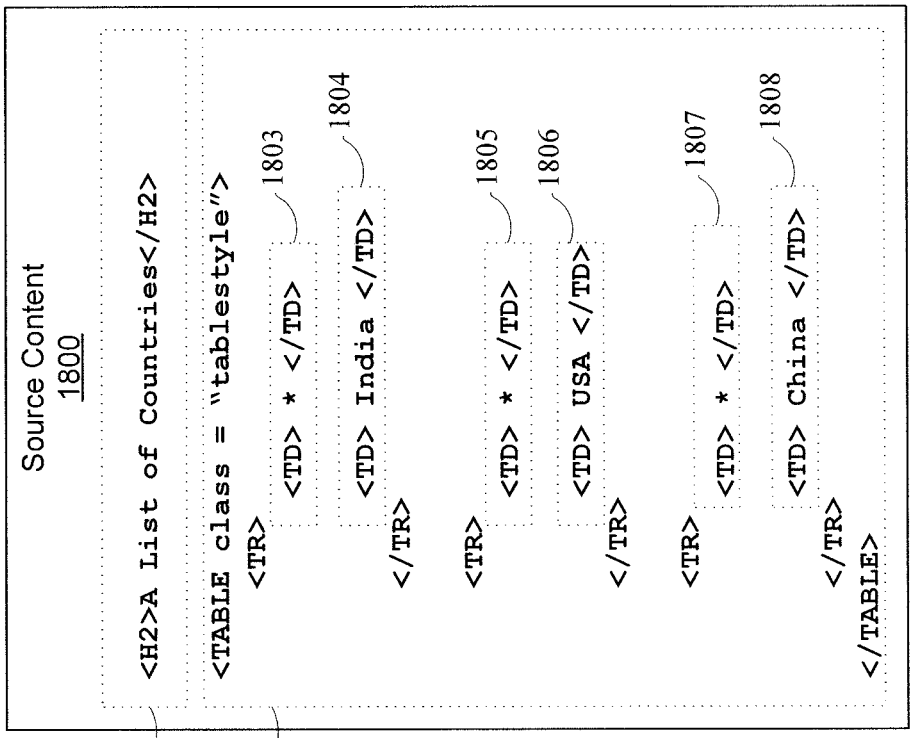

Source Content
1800
<H2>A List of Countries</H2>
<TABLE class = "tablestyle">
<TR>
<TD> * </TD>
<TD> India </TD>
</TR>
<TR>
<TD> * </TD>
<TD> USA </TD>
</TR>
<TR>
<TD> * </TD>
<TD> China </TD>
</TR>
</TABLE>
1801
1802
1803
1804
1805
1806
1807
1808

FIG. 18

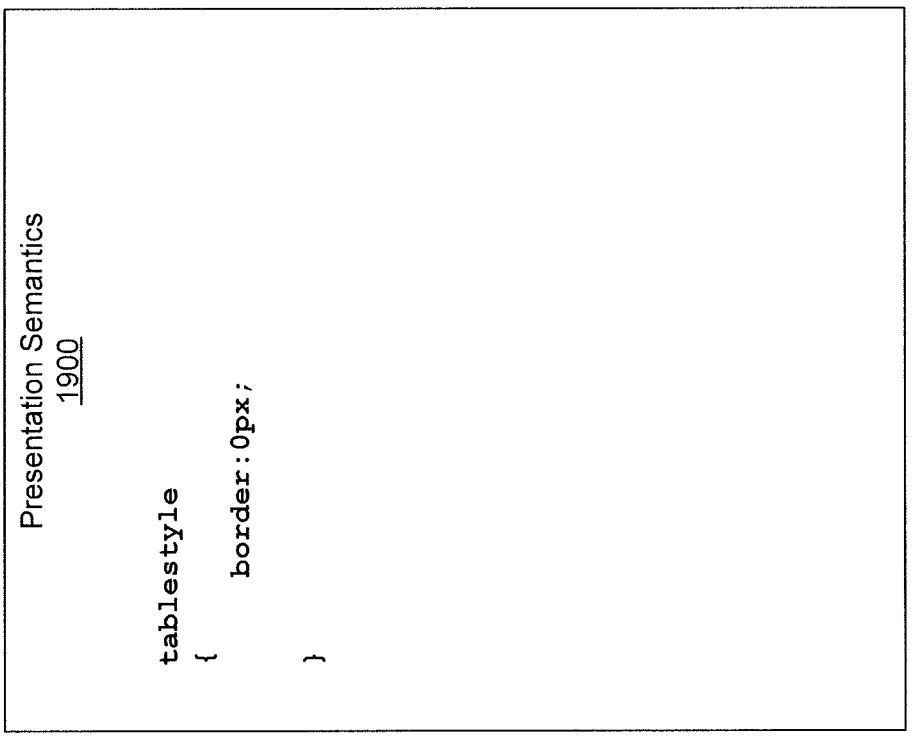

Presentation Semantics
1900
tablestyle
{
border:0px;
}

FIG. 19

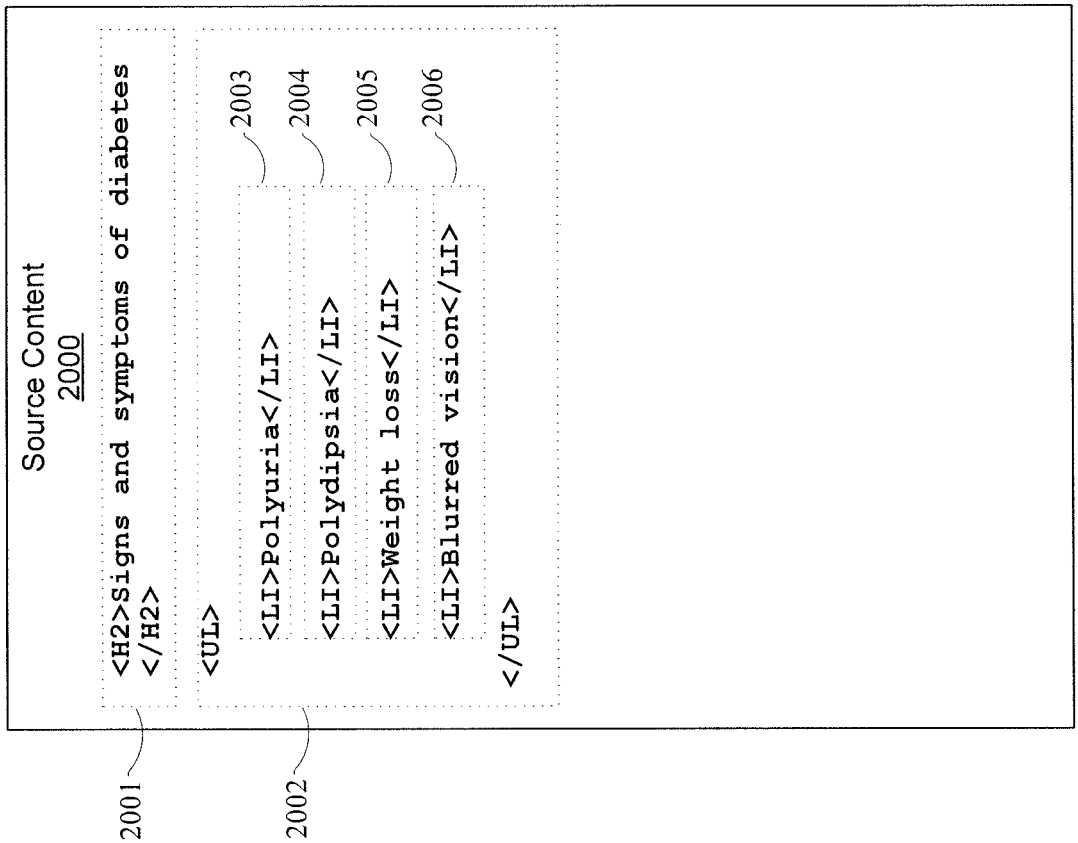

Source Content
2000
<H2>Signs and symptoms of diabetes
</H2>
<UL>
<LI>Polyuria</LI>
<LI>Polydipsia</LI>
<LI>Weight loss</LI>
<LI>Blurred vision</LI>
</UL>
2001
2002
2003
2004
2005
2006

FIG. 20

Source Content
2100

2101

```
<P>

<span class="bold">Diabetes</span>
is a condition in which the body
does not produce enough, or
properly respond to, insulin.
Insulin enables cells to absorb
glucose in order to turn it into
energy. The American Diabetes
Association reported in 2009  that
there are 12 million children and
adults in the United States-12.3%
of the population, who have
diabetes. In addition, the American
Diabetes Association estimated that
there are 6 (nearly one in two)
million diabetics unaware that they
have the disease.

</P>
```

FIG. 21

Presentation Semantics
2200

```
.bold
{
  font-weight: bold;
}
```

FIG. 22

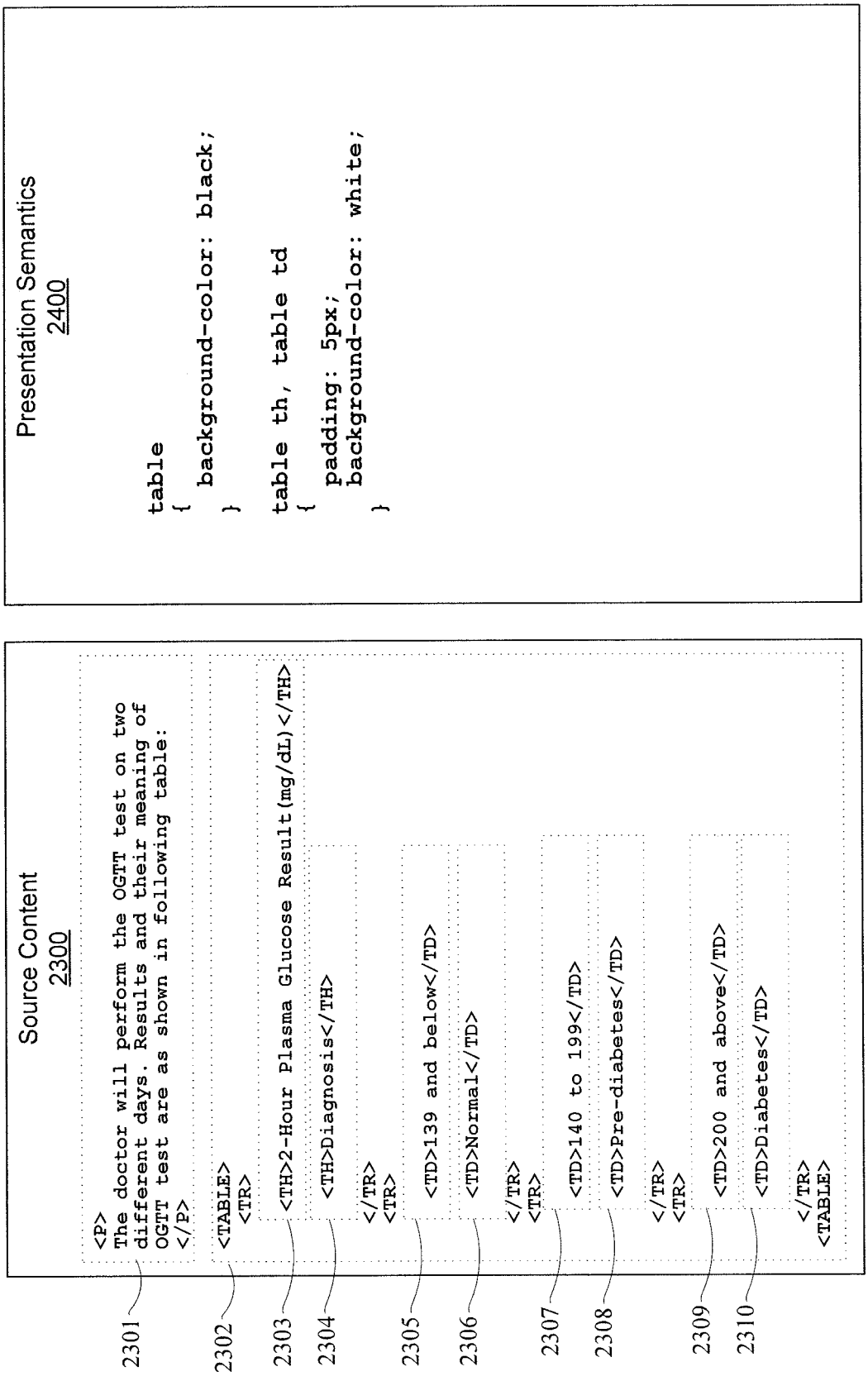

Source Content
2300
<P>
The doctor will perform the OGTT test on two different days. Results and their meaning of OGTT test are as shown in following table:
</P>
<TABLE>
<TR>
<TH>2-Hour Plasma Glucose Result(mg/dL)</TH>
<TH>Diagnosis</TH>
</TR>
<TR>
<TD>139 and below</TD>
<TD>Normal</TD>
</TR>
<TR>
<TD>140 to 199</TD>
<TD>Pre-diabetes</TD>
</TR>
<TR>
<TD>200 and above</TD>
<TD>Diabetes</TD>
</TR>
</TABLE>
2301
2302
2303
2304
2305
2306
2307
2308
2309
2310
Presentation Semantics
2400
table
{
background-color: black;
}
table th, table td
{
padding: 5px;
background-color: white;
}

FIG. 23

FIG. 24

Source Content
2500

```
<P>
The doctor will perform the OGTT test on two different days.
Results and their meaning of OGTT test are as shown in following
table:
</P>
<div class="t">
  <div class="tr">
    <div class="thc">2-Hour Plasma Glucose Result (mg/dL)
    </div>
    <div class="thc">Diagnosis</div>
  </div>
  <div class="tr">
    <div class="tc">139 and below</div>
    <div class="tc">Normal</div>
  </div>
  <div class="tr">
    <div class="tc">140 to 199</div>
    <div class="tc">Pre-diabetes</div>
  </div>
  <div class="tr">
    <div class="tc">200 and above</div>
    <div class="tc">Diabetes</div>
  </div>
</div>
```

2501
2502
2503
2504
2505
2506
2507
2508
2509
2510

FIG. 25

Presentation Semantics
2600

```
.t
{
  display: table;
  border: 1px solid black;
}

.tr
{
  display: table-row;
  border: 1px solid black;
}

.thc
{
  display: table-cell;
  padding: 5px;
  border: 1px solid black;

  font-weight: bold;
}

.tc
{
  display: table-cell;
  padding: 5px;
  border: 1px solid black;
}
```

FIG. 26

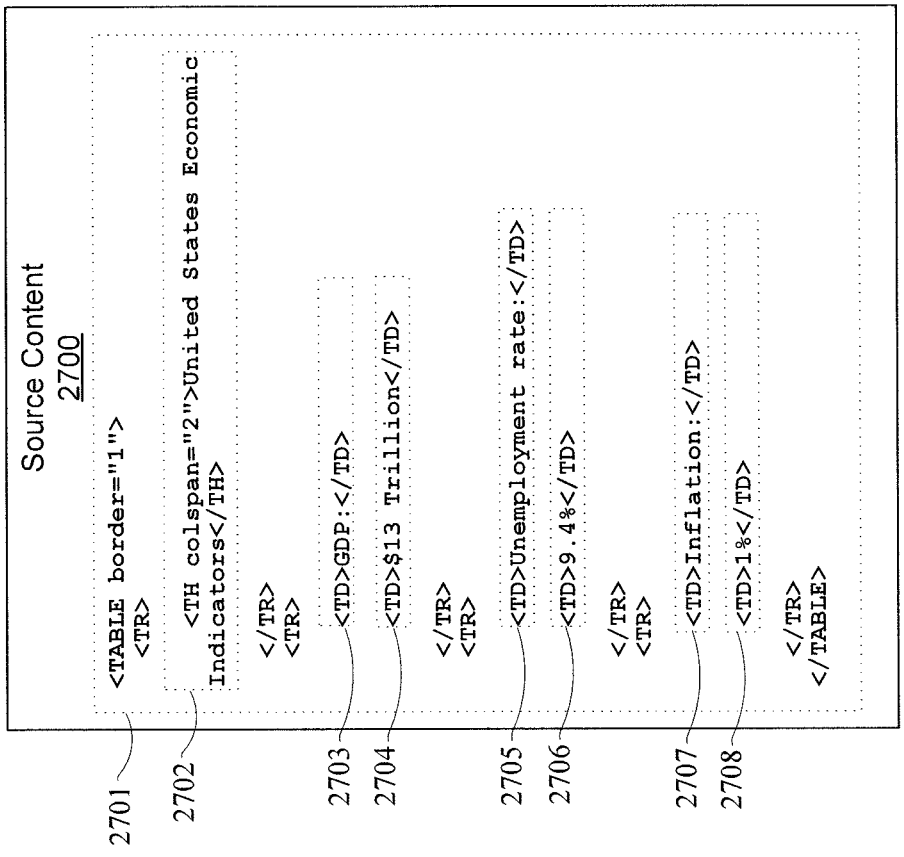

Source Content
2700
<TABLE border="1">
<TR>
<TH colspan="2">United States Economic Indicators</TH>
</TR>
<TR>
<TD>GDP:</TD>
<TD>$13 Trillion</TD>
</TR>
<TR>
<TD>Unemployment rate:</TD>
<TD>9.4%</TD>
</TR>
<TR>
<TD>Inflation:</TD>
<TD>1%</TD>
</TR>
</TABLE>
2701
2702
2703
2704
2705
2706
2707
2708

FIG. 27

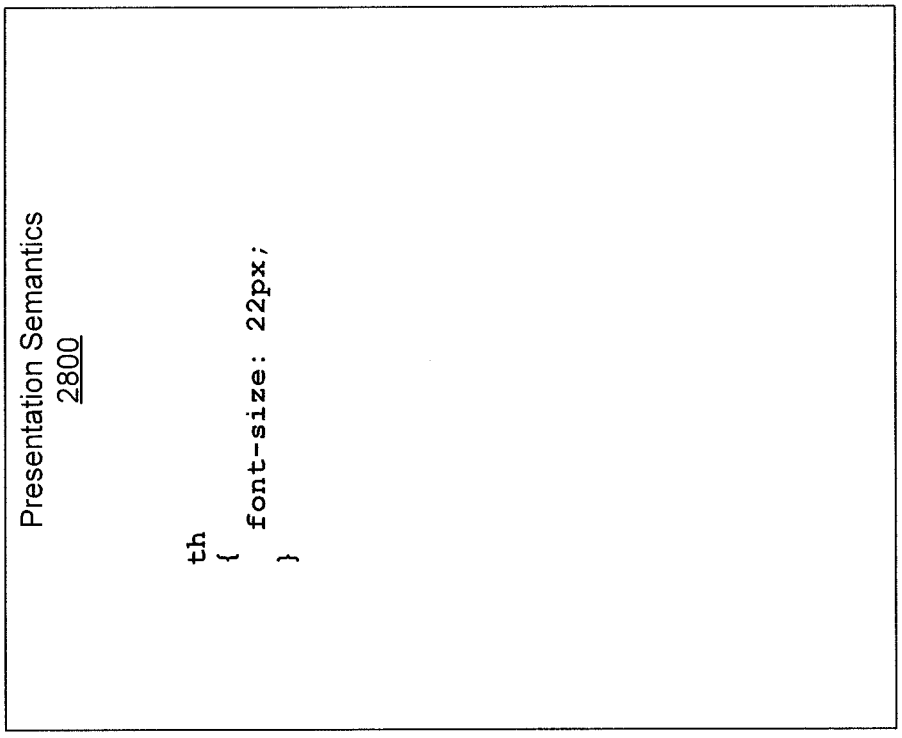

Presentation Semantics
2800
th
{
font-size: 22px;
}

FIG. 28

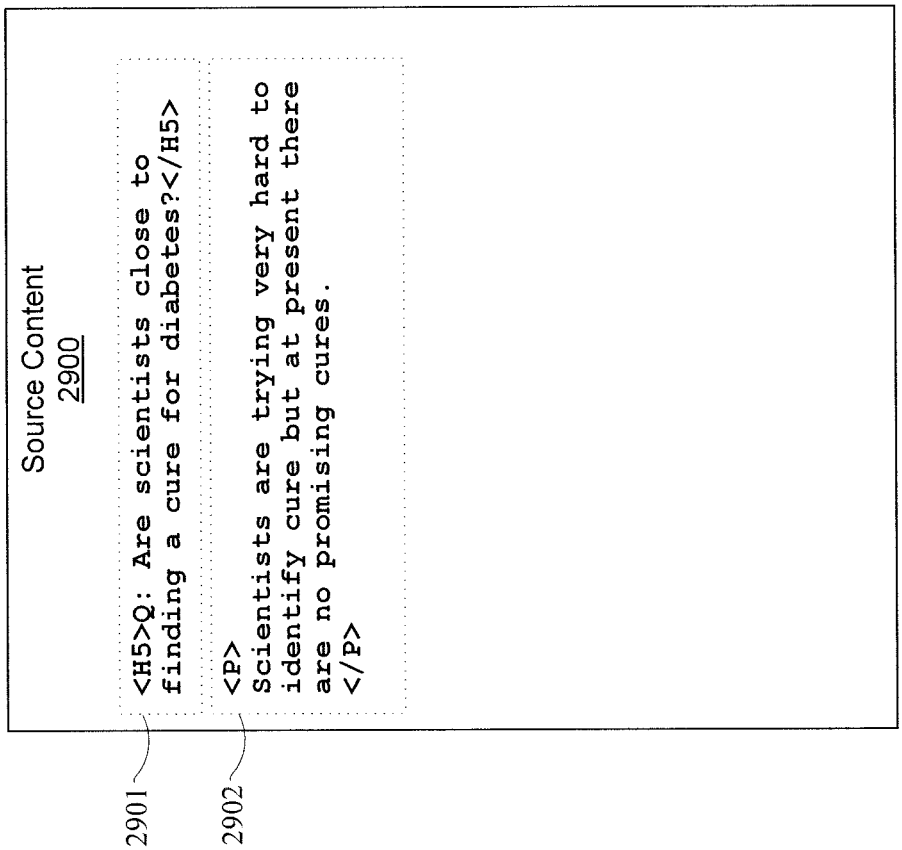

Source Content
2900
<H5>Q: Are scientists close to finding a cure for diabetes?</H5>
<P>
Scientists are trying very hard to identify cure but at present there are no promising cures.
</P>
2901
2902

FIG. 29

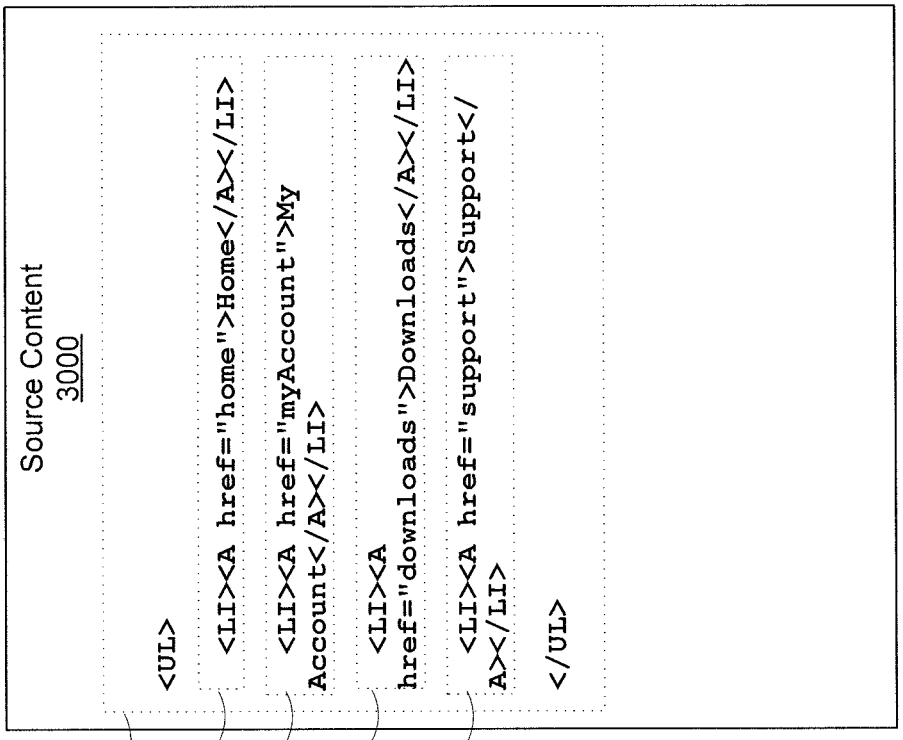

Source Content
3000
<UL>
<LI><A href="home">Home</A></LI>
<LI><A href="myAccount">My Account</A></LI>
<LI><A href="downloads">Downloads</A></LI>
<LI><A href="support">Support</A></LI>
</UL>
3001
3002
3003
3004
3005

FIG. 30

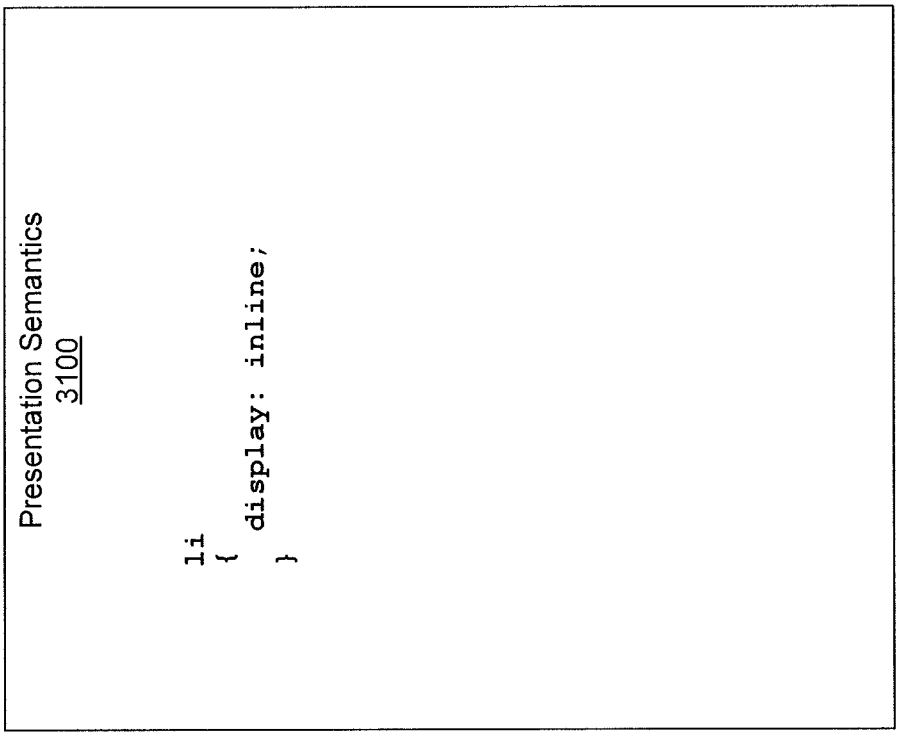

Presentation Semantics
3100
li
{
display: inline;
}

FIG. 31

Source Content
3200

```
<TABLE border="1">
  <TR>
    <TH colspan="2">Method Summary
    </TH>
  </TR>
  <TR class="fixedWidthText">
    <TD>int</TD>
    <TD>getMinutes()</TD>
  </TR>
  <TR class="fixedWidthText">
    <TD>int</TD>
    <TD>getSeconds()</TD>
  </TR>
  <TR class="fixedWidthText">
    <TD>long</TD>
    <TD>getTime()</TD>
  </TR>
</TABLE>
```

3201

3202

FIG. 32

Presentation Semantics
3300

```
.fixedWidthText
{
  font-family: monospace;
}

th
{
  font-size: 22px;
}
```

FIG. 33

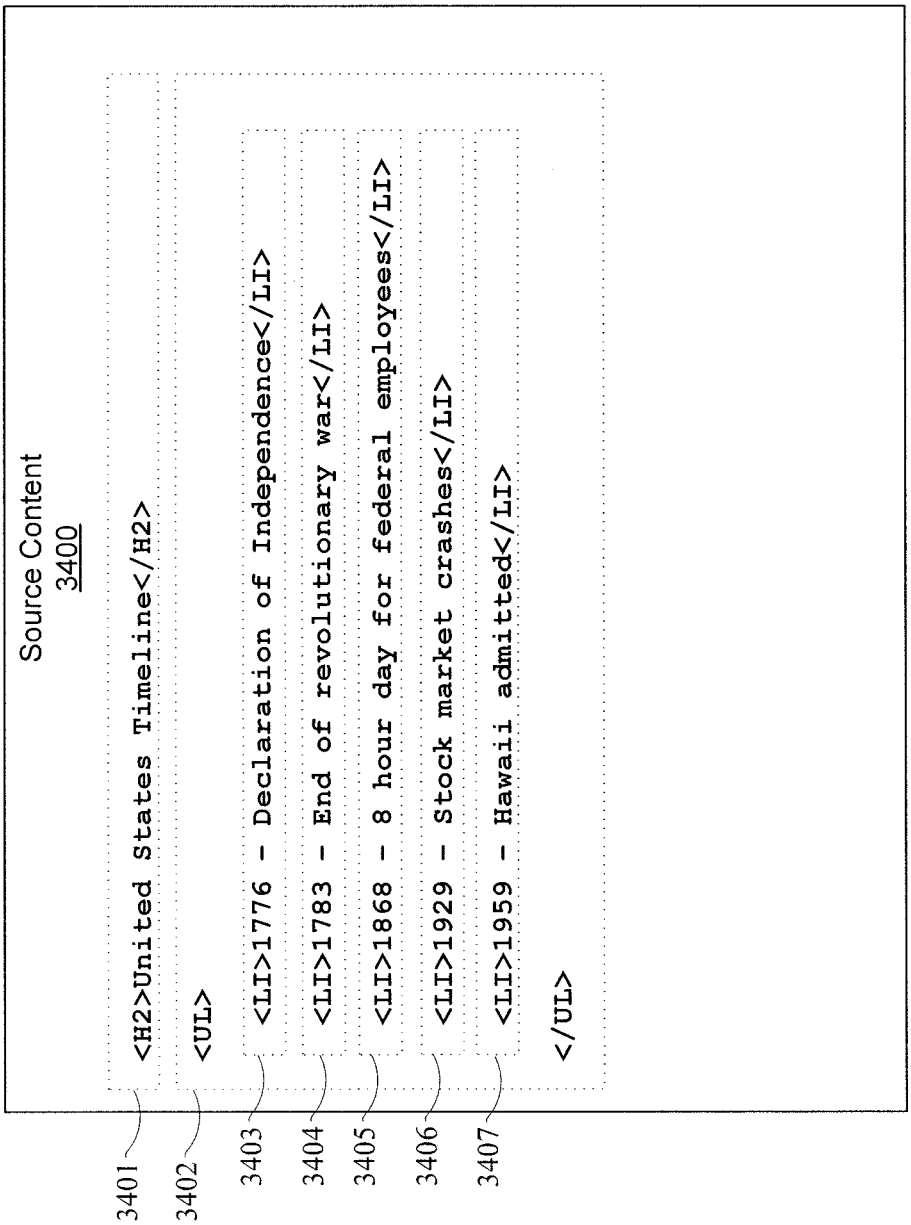

Source Content
3400
<H2>United States Timeline</H2>
<UL>
<LI>1776 - Declaration of Independence</LI>
<LI>1783 - End of revolutionary war</LI>
<LI>1868 - 8 hour day for federal employees</LI>
<LI>1929 - Stock market crashes</LI>
<LI>1959 - Hawaii admitted</LI>
</UL>
3401
3402
3403
3404
3405
3406
3407

FIG. 34

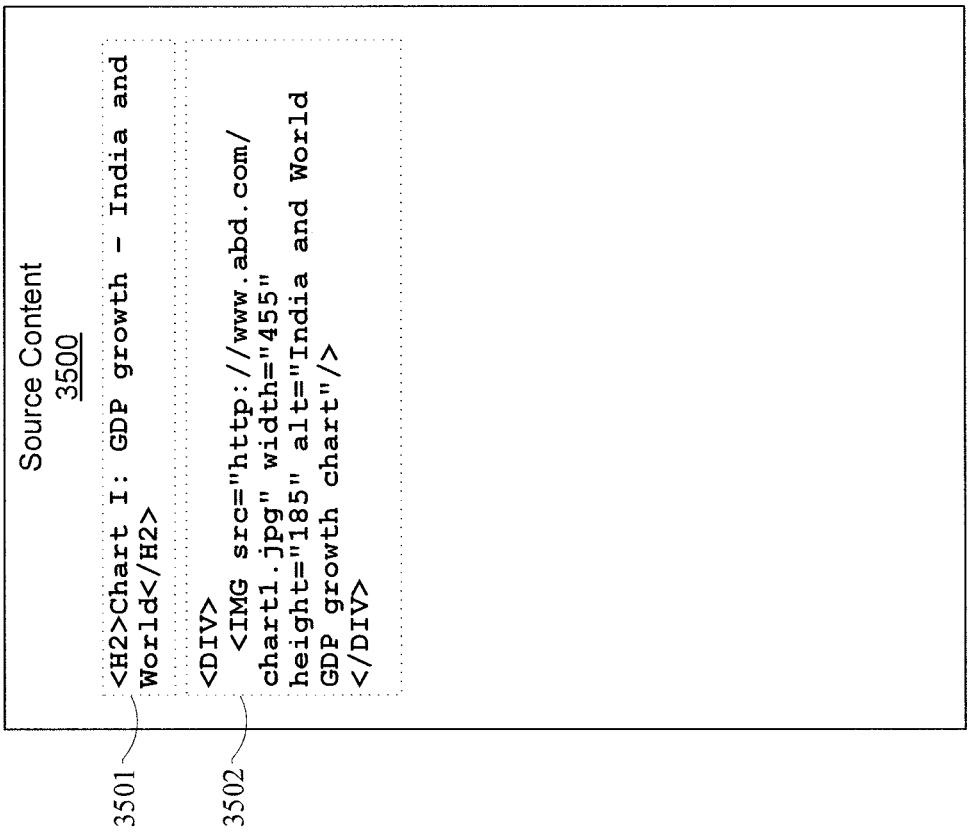

Source Content
3500
<H2>Chart I: GDP growth - India and World</H2>
<DIV>
<IMG src="http://www.abd.com/chart1.jpg" width="455" height="185" alt="India and World GDP growth chart"/>
</DIV>
3501
3502

FIG. 35

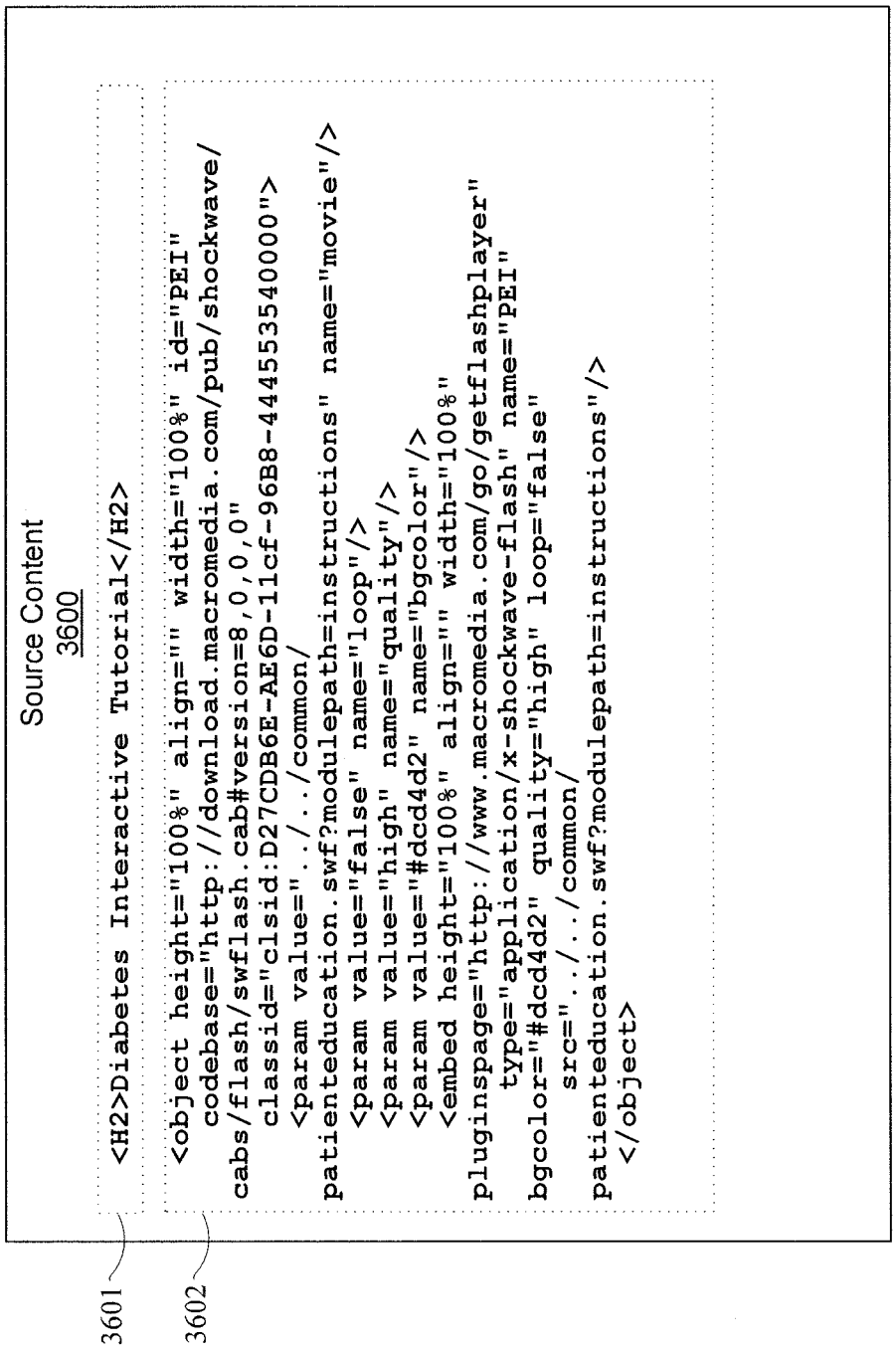

Source Content
3600
3601
3602
<H2>Diabetes Interactive Tutorial</H2>
<object height="100%" align="" width="100%" id="PEI"
codebase="http://download.macromedia.com/pub/shockwave/
cabs/flash/swflash.cab#version=8,0,0,0"
classid="clsid:D27CDB6E-AE6D-11cf-96B8-444553540000">
<param value="../../common/
patienteducation.swf?modulepath=instructions" name="movie"/>
<param value="false" name="loop"/>
<param value="high" name="quality"/>
<param value="#dcd4d2" name="bgcolor"/>
<embed height="100%" align="" width="100%"
pluginspage="http://www.macromedia.com/go/getflashplayer"
type="application/x-shockwave-flash" name="PEI"
bgcolor="#dcd4d2" quality="high" loop="false"
src="../../common/
patienteducation.swf?modulepath=instructions"/>
</object>

FIG. 36

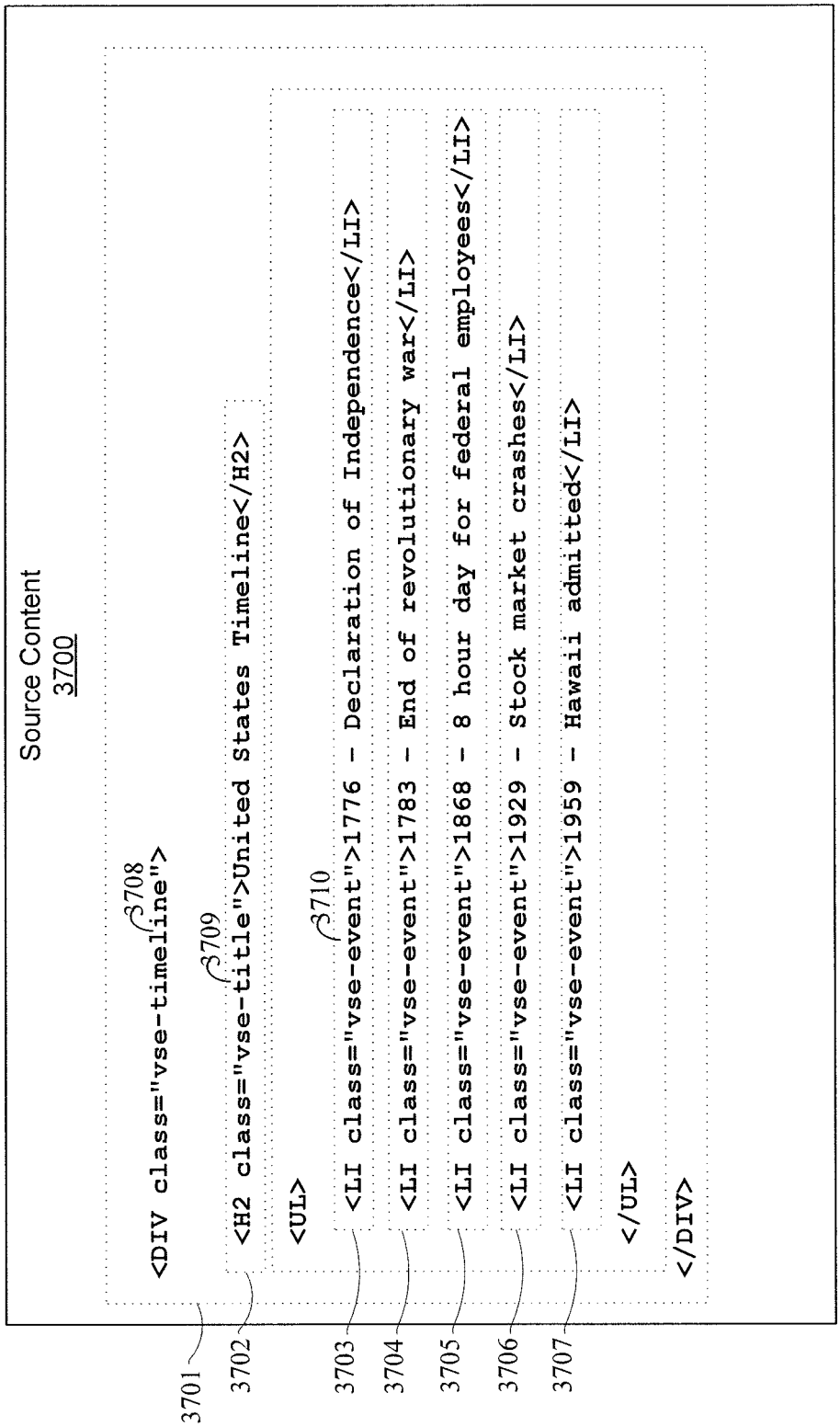

Source Content
3700
<DIV class="vse-timeline">
<H2 class="vse-title">United States Timeline</H2>
<UL>
<LI class="vse-event">1776 - Declaration of Independence</LI>
<LI class="vse-event">1783 - End of revolutionary war</LI>
<LI class="vse-event">1868 - 8 hour day for federal employees</LI>
<LI class="vse-event">1929 - Stock market crashes</LI>
<LI class="vse-event">1959 - Hawaii admitted</LI>
</UL>
</DIV>
3701
3702
3703
3704
3705
3706
3707
3708
3709
3710

FIG. 37

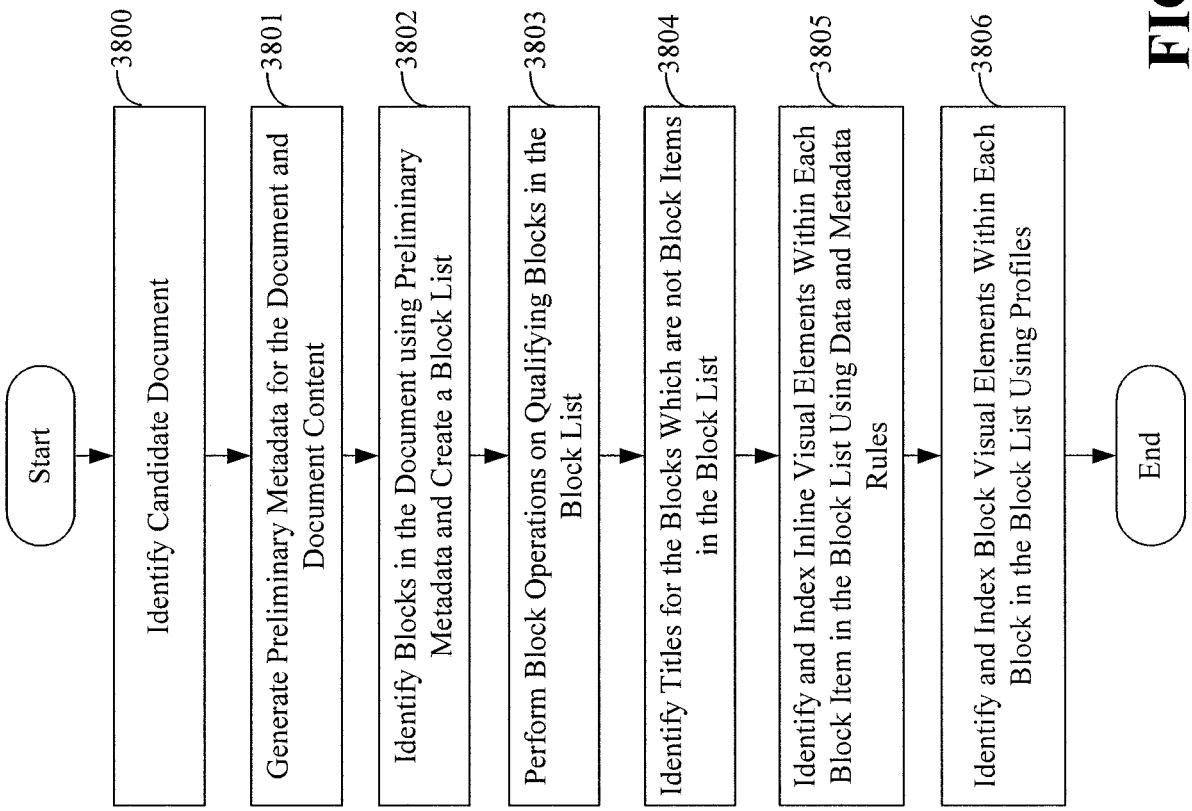

Start
3800
Identify Candidate Document
3801
Generate Preliminary Metadata for the Document and Document Content
3802
Identify Blocks in the Document using Preliminary Metadata and Create a Block List
3803
Perform Block Operations on Qualifying Blocks in the Block List
3804
Identify Titles for the Blocks Which are not Block Items in the Block List
3805
Identify and Index Inline Visual Elements Within Each Block Item in the Block List Using Data and Metadata Rules
3806
Identify and Index Block Visual Elements Within Each Block in the Block List Using Profiles
End

FIG. 38

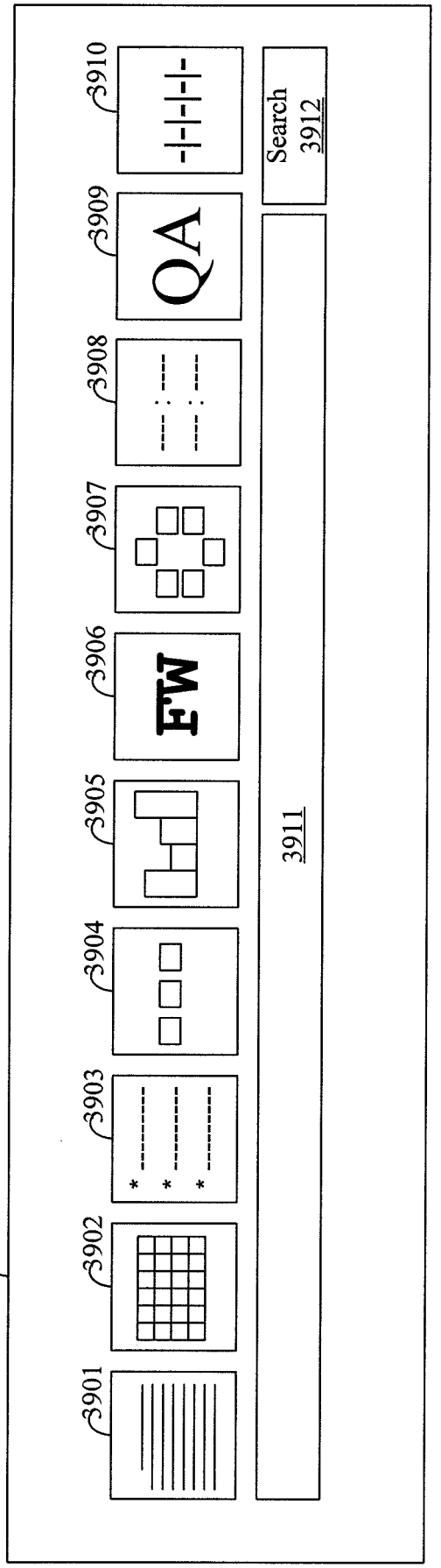

3900
3901
3902
3903
3904
3905
3906
FW
3907
3908
3909
QA
3910
3911
Search
3912

FIG. 39

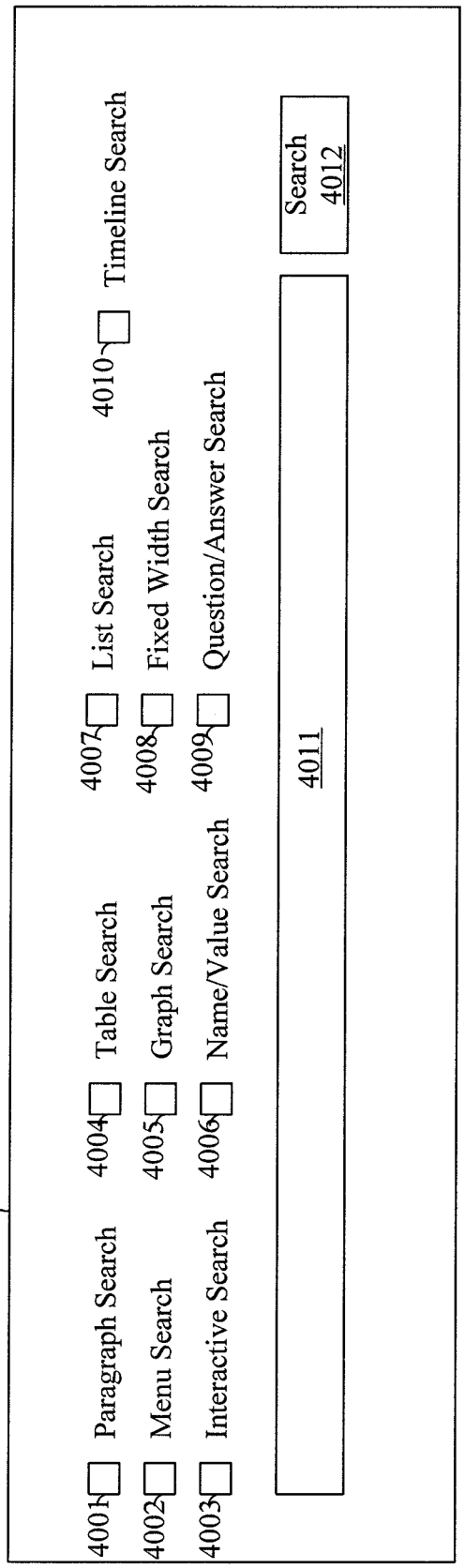

4000
4001 Paragraph Search
4002 Menu Search
4003 Interactive Search
4004 Table Search
4005 Graph Search
4006 Name/Value Search
4007 List Search
4008 Fixed Width Search
4009 Question/Answer Search
4010 Timeline Search
4011
Search
4012

FIG. 40

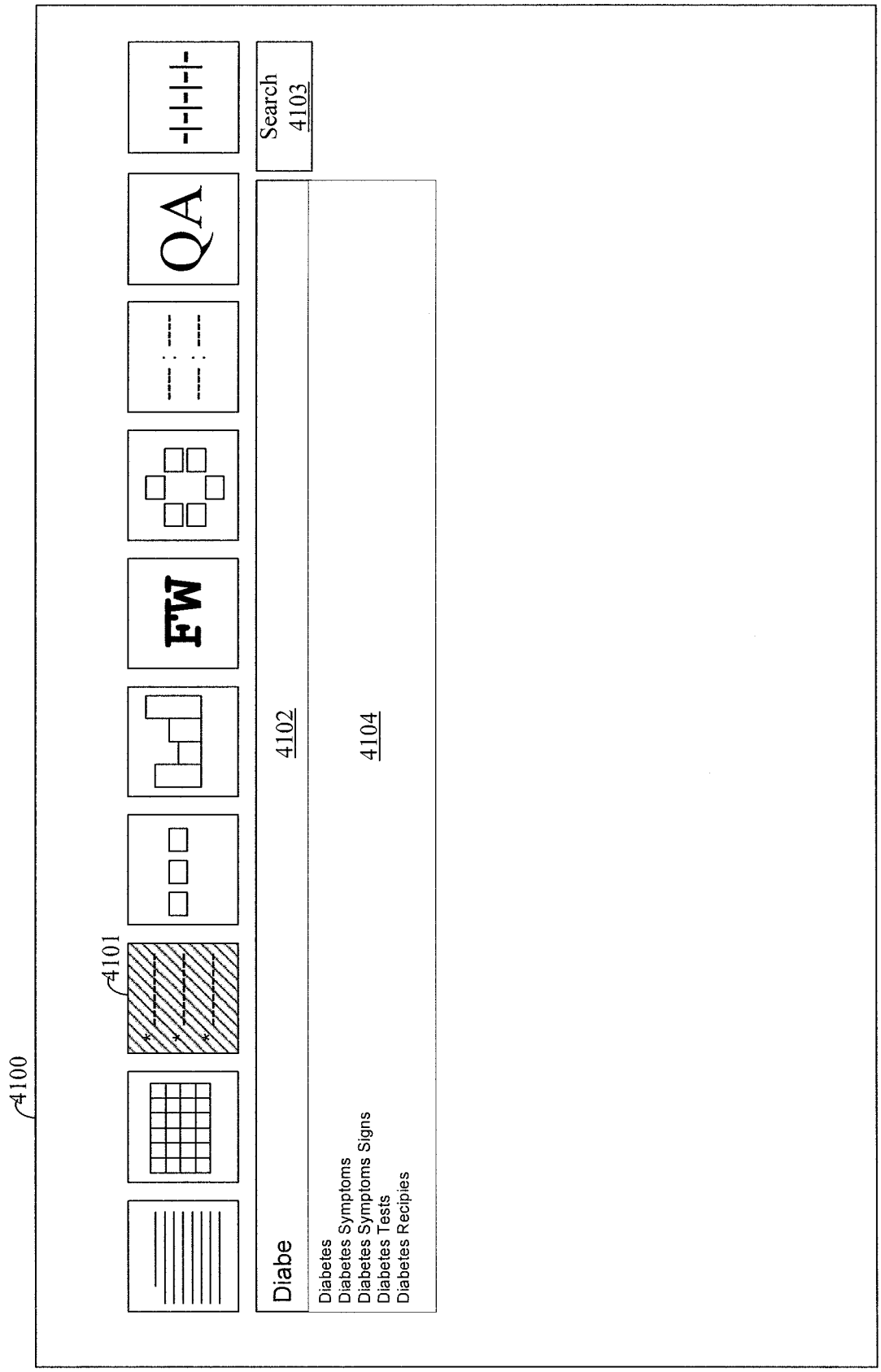

4100
4101
FW
QA
Search
4103
Diabe
4102
Diabetes
Diabetes Symptoms
Diabetes Symptoms Signs
Diabetes Tests
Diabetes Recipies
4104

FIG. 41

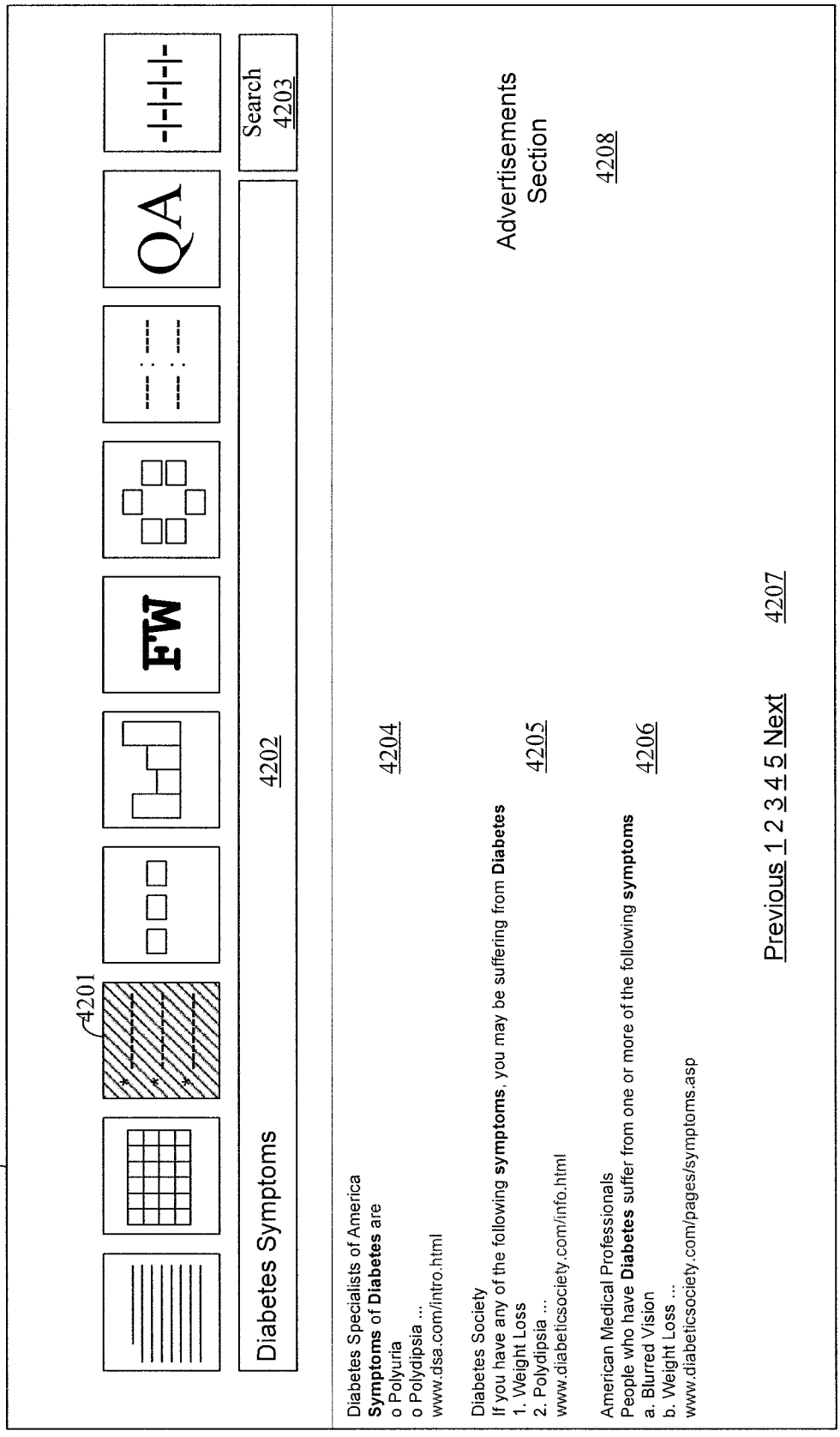

4200
4201
QA
Diabetes Symptoms
4202
Search
4203
Diabetes Specialists of America
Symptoms of **Diabetes** are
o Polyuria
o Polydipsia ...
www.dsa.com/intro.html
4204
Diabetes Society
If you have any of the following **symptoms**, you may be suffering from **Diabetes**
1. Weight Loss
2. Polydipsia ...
www.diabeticsociety.com/info.html
4205
Advertisements Section
4208
American Medical Professionals
People who have **Diabetes** suffer from one or more of the following **symptoms**
a. Blurred Vision
b. Weight Loss ...
www.diabeticsociety.com/pages/symptoms.asp
4206
Previous 1 2 3 4 5 Next
4207

FIG. 42

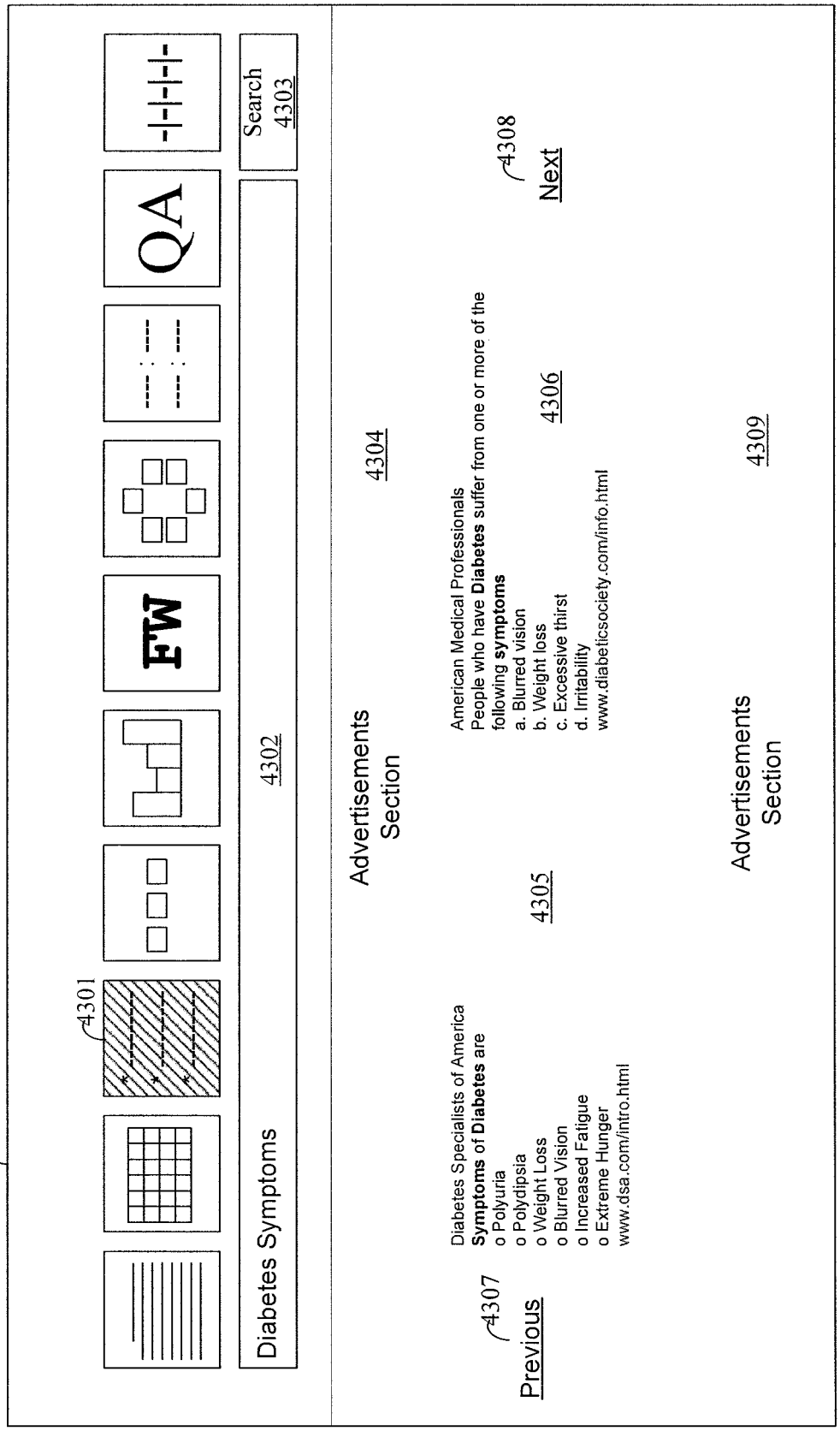

4300
4301
QA
Diabetes Symptoms
4302
Search
4303
Advertisements
Section
4304
Diabetes Specialists of America
Symptoms of **Diabetes** are
o Polyuria
o Polydipsia
o Weight Loss
o Blurred Vision
o Increased Fatigue
o Extreme Hunger
www.dsa.com/intro.html
4305
American Medical Professionals
People who have **Diabetes** suffer from one or more of the following **symptoms**
a. Blurred vision
b. Weight loss
c. Excessive thirst
d. Irritability
www.diabeticsociety.com/info.html
4306
4307
Previous
4308
Next
Advertisements
Section
4309

FIG. 43

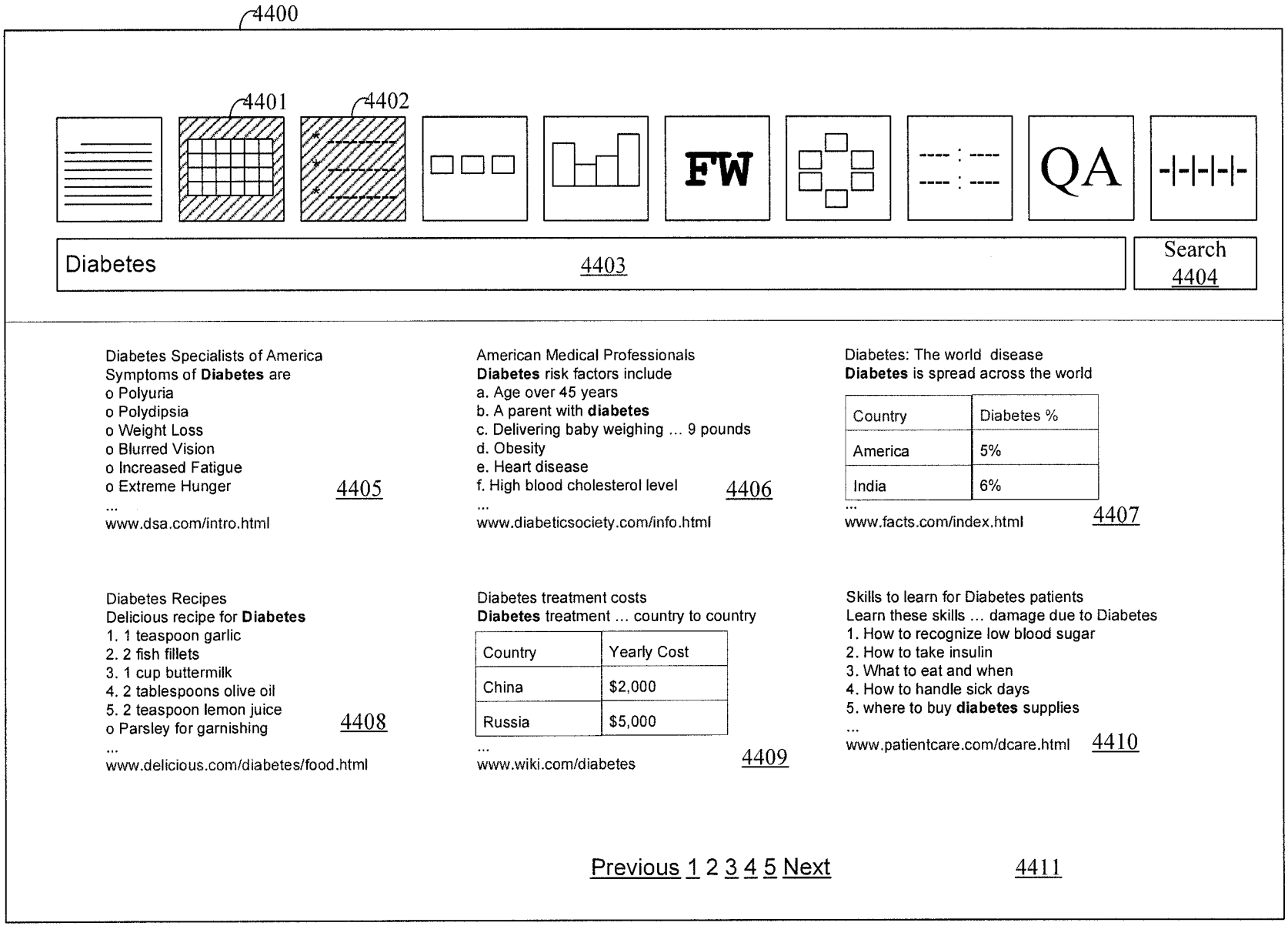

4400
4401
4402
FW
QA
Diabetes
4403
Search
4404
Diabetes Specialists of America
Symptoms of **Diabetes** are
o Polyuria
o Polydipsia
o Weight Loss
o Blurred Vision
o Increased Fatigue
o Extreme Hunger
...
www.dsa.com/intro.html
4405
American Medical Professionals
Diabetes risk factors include
a. Age over 45 years
b. A parent with **diabetes**
c. Delivering baby weighing ... 9 pounds
d. Obesity
e. Heart disease
f. High blood cholesterol level
...
www.diabeticsociety.com/info.html
4406
Diabetes: The world disease
Diabetes is spread across the world
Country | Diabetes %
America | 5%
India | 6%
...
www.facts.com/index.html
4407
Diabetes Recipes
Delicious recipe for **Diabetes**
1. 1 teaspoon garlic
2. 2 fish fillets
3. 1 cup buttermilk
4. 2 tablespoons olive oil
5. 2 teaspoon lemon juice
o Parsley for garnishing
...
www.delicious.com/diabetes/food.html
4408
Diabetes treatment costs
Diabetes treatment ... country to country
Country | Yearly Cost
China | $2,000
Russia | $5,000
...
www.wiki.com/diabetes
4409
Skills to learn for Diabetes patients
Learn these skills ... damage due to Diabetes
1. How to recognize low blood sugar
2. How to take insulin
3. What to eat and when
4. How to handle sick days
5. where to buy **diabetes** supplies
...
www.patientcare.com/dcare.html
4410
Previous 1 2 3 4 5 Next
4411

FIG. 44