

(19)
(12)

(KR)
(B1)

(51) 。 Int. Cl. ⁶
G06F 17/30

(45)
(11)
(24)

2002 11 01
10 - 0359233
2002 10 18

(21) 10 - 1999 - 0028638
(22) 1999 07 15

(65) 2001 - 0009983
(43) 2001 02 05

(73) 1가 21

(72) 108 - 1404

가 236 - 1 ICU 229

(74)
:

(54)

가

URL

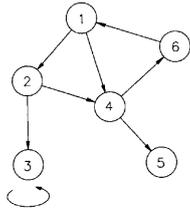
URL

URL

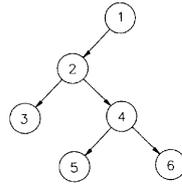
URL

URL

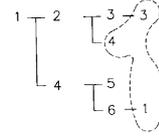
[도 3a]



[도 3b]



[도 3c]



1 가 ,
 2 ,
 3a URL URL
 ,
 3b A - ,
 3c 3b ,
 4a 4d ,
 5 ,
 6 .
 <
 100 : 120 :
 160 : 220 :
 240 : 242 : URL
 246 : 260 :
 300 : 310, 320 : URL

: (a) ; (b)
 ;
 (c) URL 가 URL,
 ; (d) URL URL UR
 ; (e) ; (e) URL URL
 ; (f) ; (g)
 ; (h) ; (i)

1 (10) 가
 가 URL
 URL

URL((10)) URL 가 1 URL
 (20)

, 2 , 1 , 2
 (100) 2 (200)

(200) (100) (120) (140) ,
 L (160) (140) HTM
 (200) (200) (120) (200)가
 0) , HTML (download) (12
 (200) 4a 가

(160) (140)
 20) (160) , , (200) , (1
 (300) (160) (thread)
 , 4b (160)
 (162, 164, 166) (168) 가 (162) URL
 (164) 가 (166)

(tree)

가 3 () (

(162) URL (120) URL URL
(164) URL 가 URL
(166) 3 URL
가 (200) 4c (162, 164, 166) (120) URL (168)
, "JVM" (200)

(200) (fetch module : FM)(220), (search module : SM)(240),
(collection module : CM)(260)

(220) URL 가 (310 320) H
TML URL URL
URL (220) 3

3a (220) (10) URL
(1) (2) (4)가 (2) (3) (4)가
(3) 가 (4) (5) (6)가
(6) (1)가 가)
가 (N N - 1 가 가) (An
(Anti-hierarchical Edge : A -)
A -
가 가 URL A -

(Depth First Search : DFS)

HTML 가
BFS 가 (Se
mantic Loss) 가
가 가

3b (220) 3a (10) A - (Anti
 - hierarchical Edge)
 , A - 가 (222)
 , 가 , A - 가 , A - 가
 . A - 가 A - 가
 3c (1) (2) (4)
 가 (222) (1, 2, 4)가 (2) (3)
 4)가 (4) (222)
 (3) 가 (3)
 (2) (222)
 (5)가 (242) (5)가 (4), (5)
 (4) (4) (6) (6) (242)
 (6) (1) (222) 가 3b
 가 URL 가 URL 가

3b
 A -
 (220) (160)
 (240) (220) (300) 가 URL
 (content - based search) 가 (260)
 (240) URL (242), (syntax analyzer) (246)

URL (242) HTTP URL
 , HTML 가 , HTTP
 mail, ftp, gopher , HTTP 가 , URL
 가 , HTTP
 가 , ps, doc, ppt MINE
 HTML URL (242) HTTP
 URL
 (246)
 (220) 가 URL
 HTML

(260) URL , (240) URL (100) URL (100) URL (120) URL (200) URL (100) URL (260) HTML .

5

(100) (120)가 URL (100) (120) (140) (510) (520) (140) (120) (160)(4b) .

(530) , (160) (162, 164, 166) URL, (200) URL, (168) URL가 "javasoft.com/products/jdk. . ." ; 가 "JVM" ; (125) URL가

가 4c "3" ,

(540) , (220), (240) (260)

(220) URL 가 URL (3 300 400) HTML (550), (240) (242) (220) URL HTTP (246) URL 가

560). (240) URL (242) (220) URL HTTP (246) URL 가

(570) , (240) (260) (260) URL (260) HTML (240) 가 가 "3" (260) URL (590).

L URL URL HTML URL UR
URL 4d URL
가

6 , 가
(response time) 가

가 가
URL

URL
(Concurrency) URL

(57)

1.

ation) URL 가 URL(Uniform Resource Loc

(a) URL URL ;

(b) URL URL ;

(c) ;

(d)

2.

1 URL ,

URL

URL

;

URL

URL

;

URL

URL

,

URL

;

URL

URL

,

URL

3.

(a)

;

(b)

;

(c)

가

URL

;

(d)

URL,

;

(e)

URL

URL

;

(e)

URL

URL

;

(f)

;

(g)

(a)

(f)

;

(h)

;

(i)

4.

3

,

,

URL

,

URL

URL

URL

5.

3

,

URL :

URL URL ;

URL URL ;

URL URL , URL ;

URL URL , URL .

6.

5 ,

URL ,

URL HTTP URL ;

URL 가 가 .

7.

1 ,

URL HTML .

8.

1 ,

9.

가 가

:

URL ;

가 URL, URL URL , URL URL

;

:

URL

URL

;

;

10.

9 , , URL URL , URL ,
URL

11.

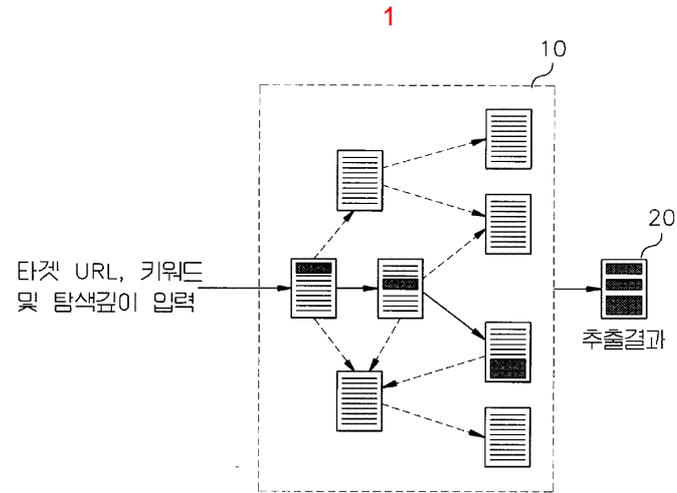
10 , URL URL , .

12.

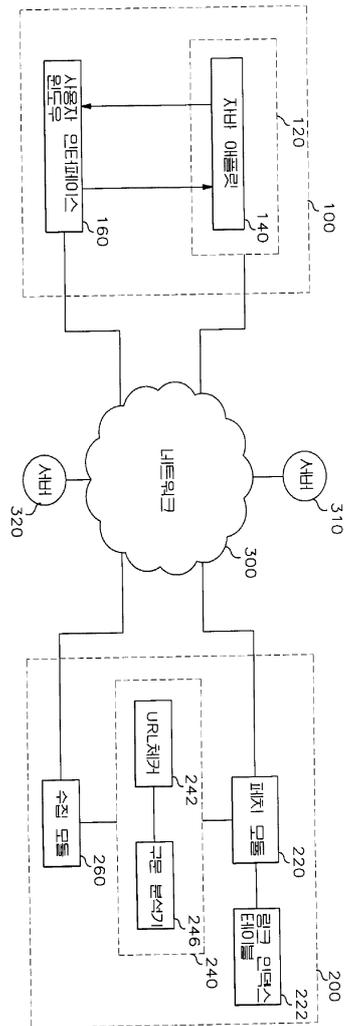
11 , , HTTP URL ;
URL HTML

13.

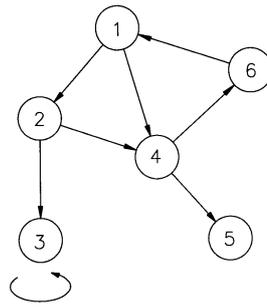
9 URL URL , .



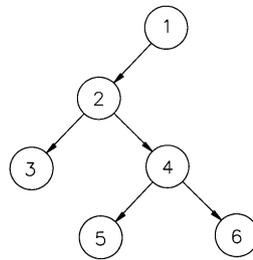
2



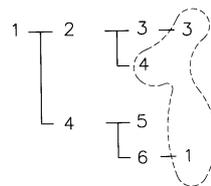
3a



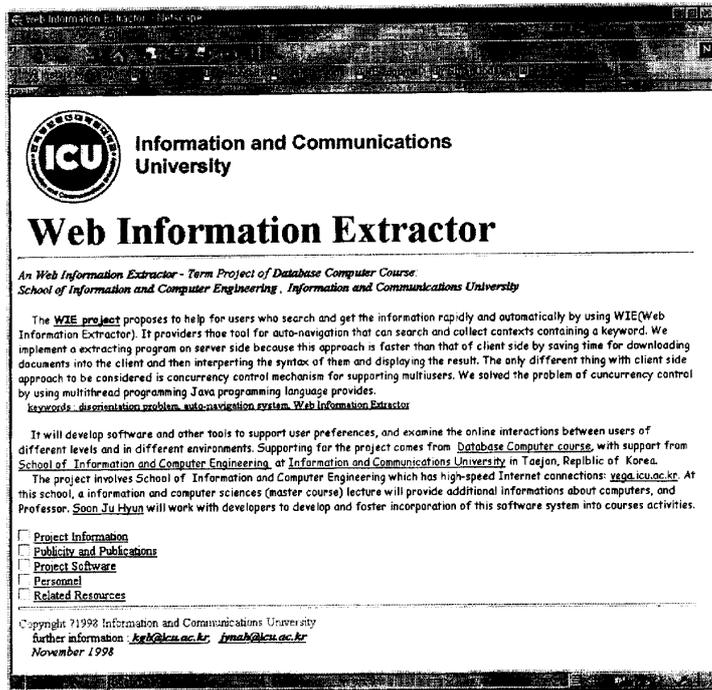
3b



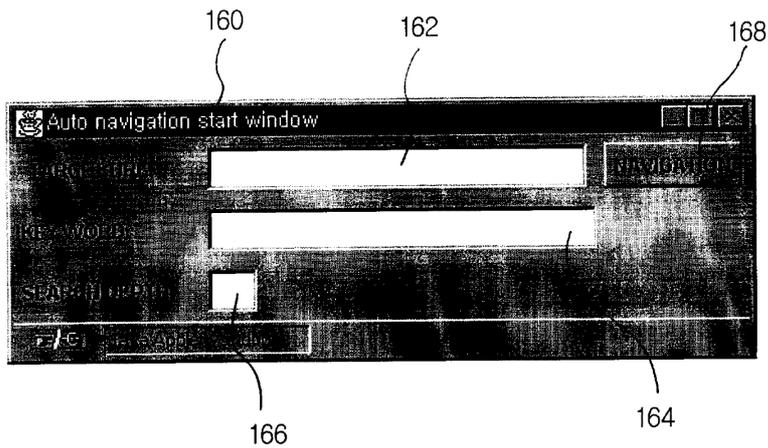
3c



4a

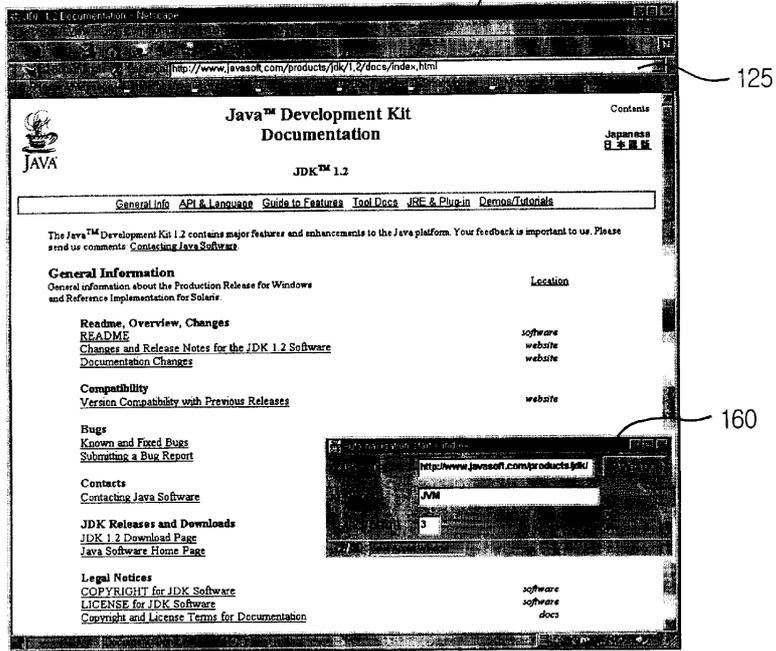


4b

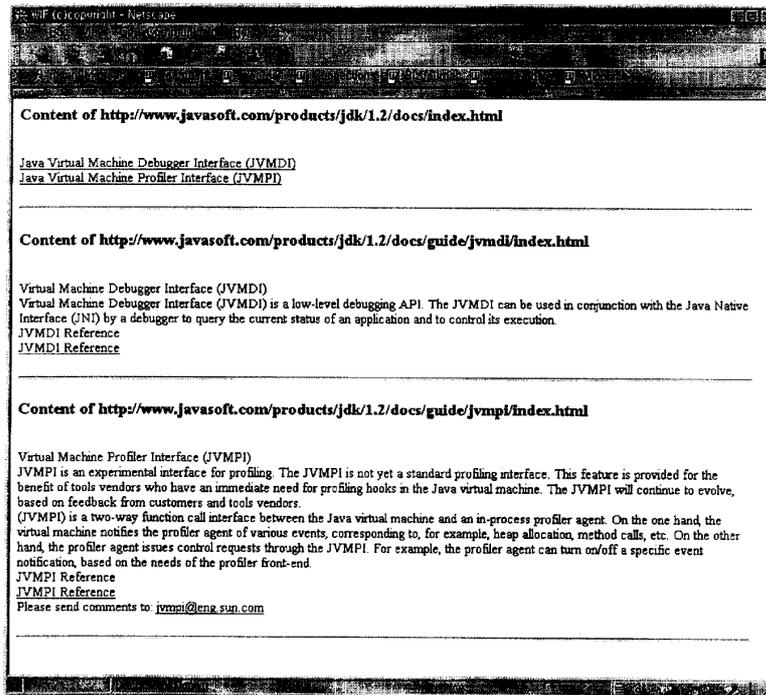


4c

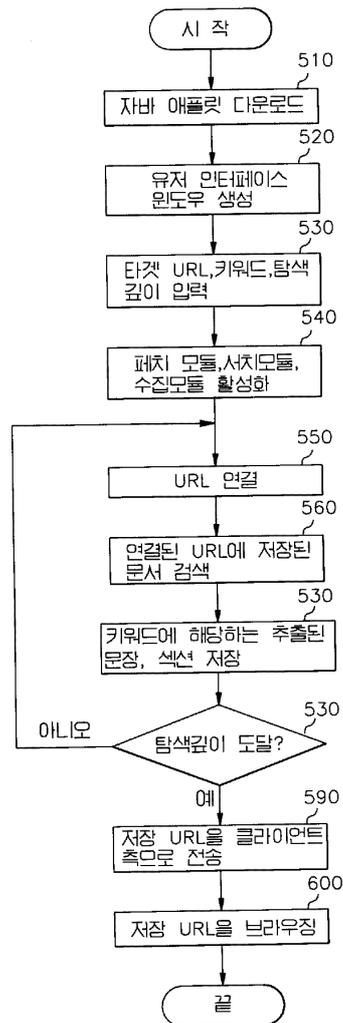
120



4d



5



6

