

(19)  
(12)

(KR)  
(B1)

(51) 。 Int. Cl. <sup>7</sup>  
G06F 17/27

(45)  
(11)  
(24)

2002 01 10  
10 - 0320348  
2001 12 27

(21) 10 - 1999 - 0044751  
(22) 1999 10 15

(65) 2001 - 0037310  
(43) 2001 05 07

(73)

31

(72)

31

31

31

(74)

:

(54)

5

1a

, 1b 1a

2

3

4

5

, , ,

가

, ,

가

, 가 ,

,

,

가

가

, (b) , (a)  
, (c)  
, (d) 가

; (b) (a)  
, ; (c) 가 ; (e) (b) (d) ;  
(d) 가 가

1a 1b  
1a 'Z' , 'V' , 'z'  
'Zv\*Vz' '르'  
1b  
'ZV\*V르' 가 '르'

2 HI , DI  
가 가  
2

2 , '르' 가 'ZV'  
'ZV\*ZV르'  
'ZV'

3 , 'ZV\*'  
(300 ).  
'ZV\*ZV르'  
'ZV\*ZV르'  
'ZV\*'



(Viterbi search)

가

가  
가  
가

4

(Part - of - Speech tagging)

( , , , ) , ( , , )  
( , )

(a) , ,

; (b)

; (c)

가

(a) (a1)  
; (a2)

; (a3)

(a3)

; (a4) (a1)

(a)

; (b)

; (c)

; (d) 가

가

; (e) (b)

(d)

(d)

( , m = e<sub>0</sub> e<sub>1</sub> e<sub>2</sub> ... e<sub>n</sub>

, e

$$\frac{\Pr(t|m)}{\Pr(t)}$$

, #

$$\frac{\Pr(t|m)}{\Pr(t)} = \frac{\Pr(e_0|e_1|e_2|...|e_n)}{\Pr(e_0|e_1|e_2|...|e_n)} = \frac{\Pr(e_0|e_1|e_2|...|e_n)}{\Pr(e_0|e_1|e_2|...|e_n)}$$

, f<sub>t</sub> 가 t

(57)

1.

(a)

(b)

(c)

(d)

2.

1

3.

1

(b)

(b1)

(b2)

(b3)

(b4) (b1) (b3)

4.

- (a) ;
- (b) , ;
- (c) , 가 ;
- (d) 가 ;
- (e) (b) (d) 가

5.

4 , (d)  $\frac{\Pr(t|m)}{\Pr(t)}$

$$\frac{\Pr(t|m)}{\Pr(t)} \approx \Pr_t(e_0|\#, \#) \cdot \Pr_t(e_1|\#, e_0) \cdot \prod_{i=2}^n \Pr_t(e_i|e_{i-2}, e_{i-1}) \cdot \Pr_t(\#|e_{n-1}, e_n)$$

$$\Pr_t(e_i|e_{i-2}, e_{i-1}) \approx f_t(e_i|e_{i-2}, e_{i-1}) + f_t(e_i|e_{i-1}) + f_t(e_i)$$

, e , # , f\_t 가 t ( , m=e\_0 e\_1 e\_2 ... e\_n ) .

6.

- (a) , ;
- (b) ;
- (c) 가 가

7.

6 ,

;

;

,

가

8.

6 , (a)

(a1) ;

(a2) , , ;

(a3) , ,

;

(a4) (a1) (a3) 가

9.

가

(a) ;

(b) , , ;

(c) , 가 ;

(d) 가 ;

(e) (b) (d) 가

10.

9 , (d)

$$\frac{\Pr(O|m)}{\Pr(O)}$$



$$\frac{\Pr(t|m)}{\Pr(t)} \approx \Pr_f(e_0|\#\#\#) \cdot \Pr_f(e_1|\#,e_0) \cdot \prod_{i=2}^n \Pr_f(e_i|e_{i-2},e_{i-1}) \cdot \Pr_f(\#|e_{n-1},e_n)$$

$$\Pr_f(e_i|e_{i-2},e_{i-1}) \approx f_i(e_i|e_{i-2},e_{i-1}) + f_i(e_i|e_{i-1}) + f_i(e_i)$$

$e_0 e_1 e_2 \dots e_n$ ,  $e$ ,  $\#$ ,  $f_t$  가  $t$  ( ,  $m=$  )

1a

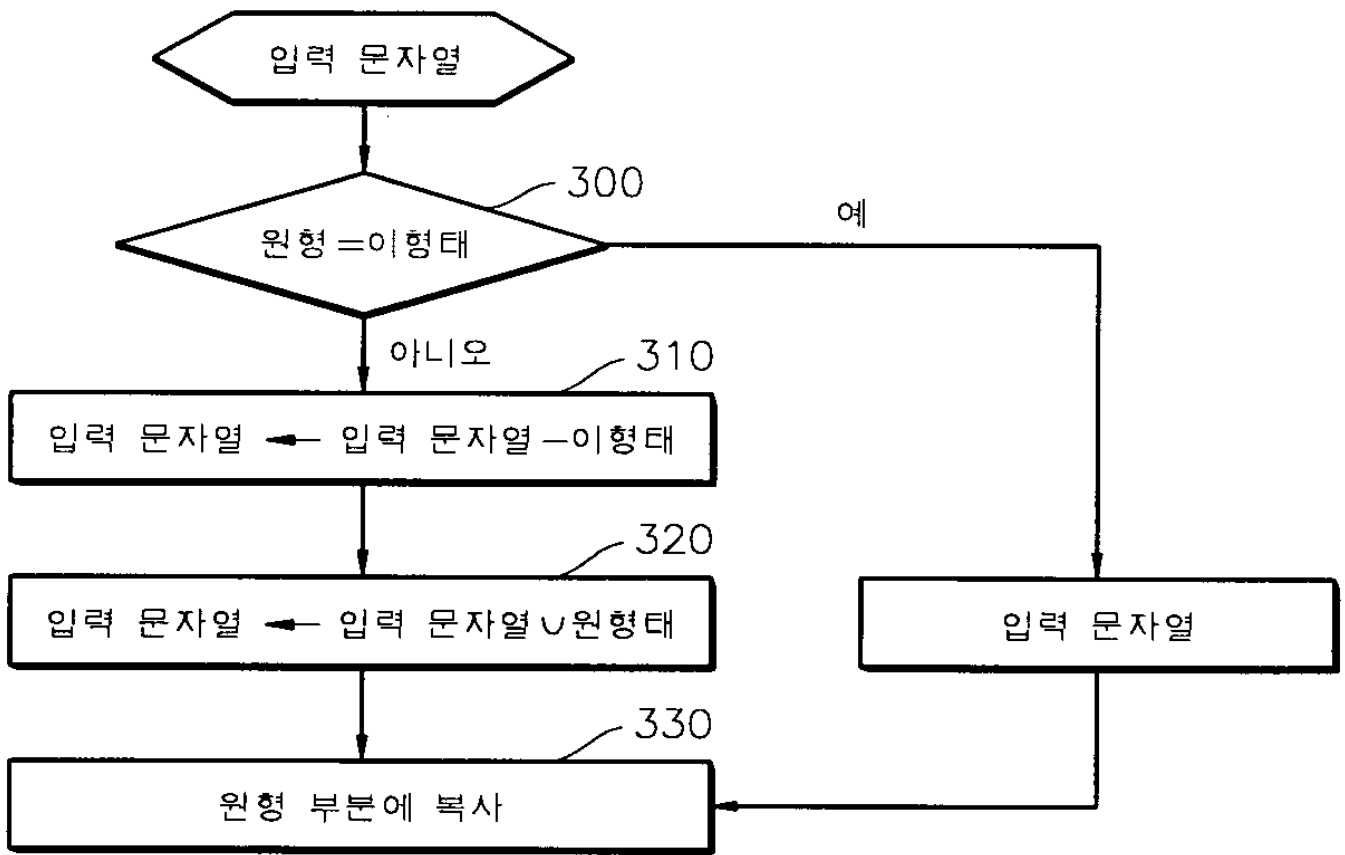
정규화 표현 문자	설명
Z	자음
V	모음
z	'ㄹ'을 제외한 자음 받침
*	임의의 문자

1b

정규화 표현	설명
ZV*갈	'갈'로 끝나는 모든 형태소
ZV*Vz	'ㄹ'을 제외한 자음으로 끝나는 모든 형태소
ZV*Vㄹ	'ㄹ'로 끝나는 형태소

원형	이형태	접속정보
HIㄹ <ZV*갈>	(ZV*갈)	[규>어>]
HIㄹ <ZV*갈>	(ZV*가)	[불>]
HIㅅ <ZV*ZVㅅ >	(ZV*우)	[불>]
HIㅅ <ZV*ZVㅅ >	(ZV*위)	[축약>]
HIㅅ <ZV*ZVㅅ >	(ZV*와)	[축약>]
DIㅅ <ZV*젓>	(ZV*젓)	[규>]
DIㅅ <ZV*젓>	(ZV*저)	[불>어>]
DIㄷ <ZV*들>	(ZV*들)	[규>]
DIㄷ <ZV*들>	(ZV*들)	[불>어>]

3



4

$$m = e_0 e_1 e_2 \dots e_n$$

$$\frac{\Pr(t | m)}{\Pr(t)} \approx \Pr_t(e_0 | \#, \#) \Pr_t(e_1 | \#, e_0)$$

$$\prod_{i=2}^n \Pr_t(e_i | e_{i-2}, e_{i-1}) \Pr_t(\# | e_{n-1}, e_n)$$

$$\Pr_t(e_i | e_{i-2}, e_{i-1}) \approx f_t(e_i | e_{i-2}, e_{i-1}) + f_t(e_i | e_{i-1}) + f_t(e_i)$$

$m$ :형태소

$e$ :음절

$\#$ :형태소경계

$f_t$ :품사가  $t$ 일경우 빈도수

5

