

MCC
9651M21

, Lee, Won-Dae, Approximate
Matching Algorithms for ODP Trader, ODP
,
, 1998, 58P, Advisor. Won-Ki Hong,
Text in Korean.

Abstract

An open distributed environment can be perceived as a service market where services are freely offered and requested. Any infrastructure which seeks to provide appropriate mechanisms for such an environment has to include mediator functionality that matches service requests and service offers. One of the functions that will be standardized as part of the ODP activities is the trading function. To search the requested service offers, minimum useful check by the trader is interface signature compatibility. However, signatures are not sufficient to choose a service requested. Properties provide a flexible means of extending service categorization.

In order to search requested service offers, matching service properties is generally based upon exact matching algorithms. This thesis argues that the exact matching algorithms stand in the way of getting a good recall rate of the service offer requested. This thesis presents flexible means of extending service categorization by applying approximate matching algorithms for property matching. This thesis proposes Metaphone algorithm and thesaurus-based search algorithm for the approximate matching algorithm. The matching algorithms have been implemented to a prototype trader and analyzed.

The trader has been experimentally compared with AI-Trader.

ABSTRACT.....	I
.....	II
.....	V
.....	VII
1.	1
1.1	1
1.2	2
2.	4
2.1	4
2.1.1 AI-Trader.....	4
2.1.2 ANSAware	7
2.1.3 DCE	8
2.1.4 OMA/CORBA	9
2.2 X.500	9
2.2.1 Soundex.....	9
2.2.2 Metaphone.....	10
2.3 WWW	11
2.3.1	11
2.4	12
3.	13
3.1	13

3.2	16
3.3	19
4.	19
4.1	19
4.1.1	Soundex.....	20
4.1.2	Metaphone.....	22
4.1.3	Minimum Edit Distance.....	23
4.1.4	Bickel	24
4.1.5	26
4.2	28
4.3	29
4.3.1	30
4.3.2	32
4.3.2.1	Metaphone	32
4.3.2.2	34
4.3.2.3	35
5.	38
5.1	38
5.1.1	38
5.1.2	Soundex Metaphone	40
5.1.3	42
5.1.4	44
5.2	46
5.2.1	Soundex Metaphone	46
5.2.2	48
5.2.3	48
5.2.4	AI-Trader	49
6.	54

.....	55
.....	58
A. Metaphone	58
B.	60

1. IDL	6
2.	6
3.	14
4.	15
5. 가	17
6. Bickel	25
7.	30
8. Metaphone 가 Match	33
9. Match	34
10. Match	36
11. Metaphone	42
12.	43
13.	45
14. AI-Trader	50
15. AI-Trader	50
16. Metaphone	52

17.	53
18.	53

1. Soundex	21
2. Soundex key	21
3. Metaphone	가	23
4. Bickel	25
5.	37

1.

1.1

(Open Distributed Processing)
(architecture),
(requester) (provider),
가
(at compile time) (at run time) (binding)
가
(run time) (dynamic selection)
Open Distributed Processing (ODP) [IS95a] (trading)
가 . ODP ISO 가 [IS95a,
IS95b, IS95c].
ODP
[IS094, Mir91].

3
(exporter) (service offer)
가, 가 가

(properties)

(exact)

[Way95].

가

(recall) [Inr92]

ODP

[IS094]

Metaphone

(thesaurus)

1.2

ODP

Metaphone

ODP

가

가

X.500

[X500a] WWW [TBe95]

ODP

RM-ODP [IS93r]

6 . 1

, 2

X.500

AI-Trader [Alt95a]

WWW

. 3

ODP

, 4

. 5

AI-Trader

. 6

.

2.

AI-Trader

ANSA [ANS93a]

DCE [DCE01]

CORBA [OMG95]

, X.500

WWW

2.1

2.1.1 AI-Trader

AI-Trader [Alt95a]

(type management repository)

(conceptual graph) [Con84]

Interface

Definition Language (IDL)

(flexibility)

. AI-Trader

(a priori)

,
“
”
가

(service offer)

(service request)

DCE [War92]

, ANSA, CORBA [OMG95],
IDL
가

, AI-Trader

, WWW

(cognitive domain)

AI-Trader

AI-Trader

[Con84]

[Alt95b].

“ ” “ ”

, PRINTER, COMPUTER, PAPER INFORMATION

가 .

가 . IDL

(knowledge-based)

PRINTER CORBA IDL

1 .

```

properties PrinterQoS {
    static string Name;
    static int DPI;
    dynamic int QueueLength;
    queryinterface QueryPrinterOoS;
}

```

1. IDL

2 AI-Trader

(notation) B .

[PROPERTY-TYPE:{"PrinterQoS"}] -

-> (STATIC_ATTRIBUTE) -

-> [STRING:{"Name"}],

-> [INTEGER:{"DPI"}],

-> (DYNAMIC-ATTRIBUTE)

-> [INTEGER: {"QueueLength"}]

-> (EVAL-CALLBACK)

-> [INTERFACE-TYPE:{"QueryPrinterQoS"}].

2.

1 2 Name, DPI, QueueLength

가 .

[Alt95a].

AI-Trader , IDL

가

, IDL

, 가

가

2.1.1.2 ANSAware

Advanced Network System Architecture (ANSA) [ANS93a]

ODP ODP

ANSAware [Rov93].

ANSAware(3.0)

[Jea93].

RM-ODP [IS93r]

(naming and relationships)

, ANSAware

[Wa95b].

(structural compatibility)

가

ANSAware [Rov93]

가 ,
,
가
가 (constraints) . ANSAware
가 가 .
ANSAware
ODP

2.1.1.3 DCE

Distributed Computing Environment (DCE) [War92] Open System
Foundation (OSF) , ,

DCE 가 . IDL
. 가 ,
“ (compatible)”
(exact)
(matching)
가 ,

(query)

DCE

가

. ODP

, DCE

ODP

2.1.4 OMA/CORBA

Object Management Group (OMG) Object Management Architecture (OMA) [Rob96]

[OMG92].

Common Object Request Broker

Architecture(CORBA)

OMA RM-ODP

(predicate)

. CORBA

(Interface Repository)

[IS93r]

(Object Relationship)

(behavior

description)

2.2 X.500

2.2.1 Soundex

Soundex

[Knu73]

가

. Soundex

[Knu73]

1918

(OdeII)

(Russell) [Or18]

Soundex 가 ,
Soundex (phonetic)

[Mav95].

‘Hawaii’가 ‘Hawai’

Soundex X.500 가
EAN [NBGS92], DEC Unisys
4 1

2.2.2 Metaphone

(Lawrence Philips) Metaphone
[Phi90], Soundex
(encoding) 가

Soundex

Directory System Agent

(DSA) [X500a]

(100,000) Soundex 가
Soundex
Metaphone

Metaphone 4 1

2.3 WWW

2.3.1

(thesaurus)

()

(Yellow

page)

가

1957

(Doking)

Brownson

[IR90a].

[HBr57].

가

WWW

ODP

2.4

ANSAware, DCE [War92], CORBA [OMG95]

(exact)

(approximate)

2 1

AI-Trader

. AI-Trader IDL

IDL

. AI-Trader

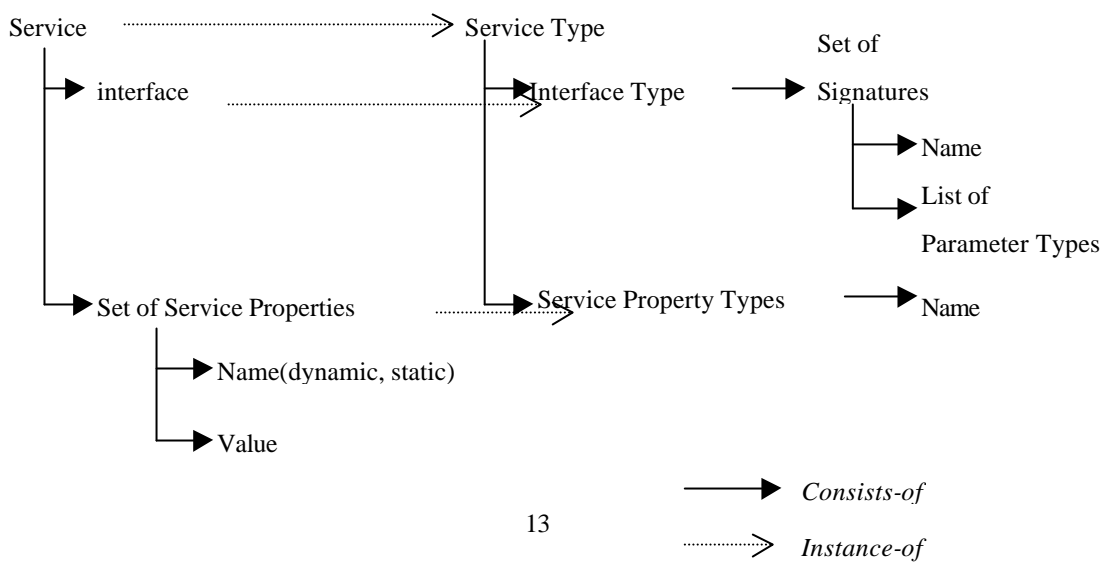
Metaphone

3.

3.1

(instance)

가



3.

3

(QoS)

가 . 가 ,

가 (constraints)

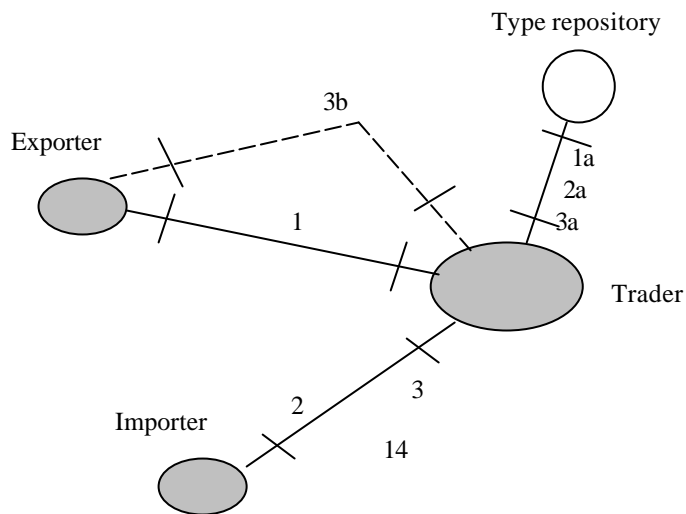
(behavior)

가

(signature)

가

[Mir93].



4.

1. 가

(1a) (validation),

(identifier)

(identifier)

2. 가

(request)

(2a) (constraints) (

) (valid)

3. (requirement

specification)

[IS93r]

[IS93r]

(3a).

(identical and/or sub-type)

,

가

가
, 가 (constraints)
가
.
.
.
(Quality of Service)
, 가
, 가 가
(specified administration)
가
(advertising)
가
(query) , 가 가
(constraints)
(refine) [ANS93b].

3.2

(service categorization)
(flexible means)
[ANS94a].

(value)
[ANS93b].

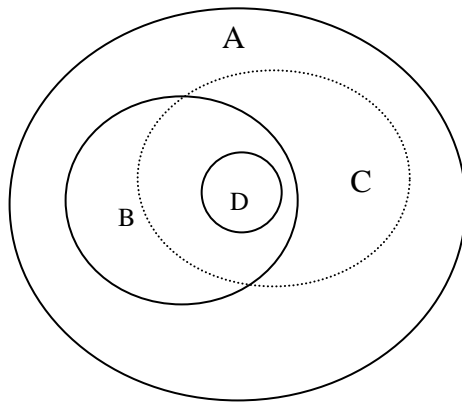
(validity)
가

(correctness)
가

가
가
(correct)

[ANS93b].

5 A
D
(recall) 가 . 가
가
5 B



5. 가


```

(property)          (validity)          (correctness)
[ANS93b]           ,          (exact)
.

      ANSAware          가

,

! traderRef$Export (IFTypeName, Context, PropertyList, ir)

DECLARE {ir} : BankTeller SERVER
Ansa_InterfaceRef ir;

{ir} :: BankTeller$Create(1)
traderRef$Export ("BankTeller", "/banks", "Bank Citibank", ir)

      가          ANSAware

.

! {ir} <- traderRef$Import(IFTypeName, Context, Constraints)

DECARE {ir} : BankTeller CLIENT
Ansa_InterfaceRef ir;

{ir} <- traderRef$Import ("BankTeller", "/banks",
                          "Bank == 'Citybank'")

(PROPERTYLIST)
Bank Citibank .          Bank == Citybank
(constraints)          .

,      5      C      D

```

가 . (C-B)
가 ,
[ANS93b]

3.3

3 2
(service categorization)

(Numeric) 가 .
(string constants) . ,
()
가 ,
가 가 .

4 .

ODP
, 가

4.1

(recall)

(precision) [Inr92]

가

(typographical)

(cognitive)

(phonetic)

4.1.1.1 Soundex

Soundex [Knu73]

1 가

(name)

“Soundex key”

Soundex key 가

Soundex code	Letters
1	B, F, P, V
2	C, G, J, K, Q, S, X, Z
3	D, T
4	L
5	M, N
6	R
0	A, E, H, I, O, U, W, Y

1. Soundex

(soft consonants) 'H', 'W', 'Y' 0

0

2 가

Name	Soundex key
Barker	B626
Knight	K523
Attlee	A34
Livingstone	L152
May	M

2. Soundex key

Soundex 가 (5.2.1).
Soundex key

가 Soundex key
soundex key

Soundex

Soundex 가 , 가

가 . 2.2.1 X.500

[Inr92]

, 가 Soundex
Metaphone
(Michigan)

4.1.2 Metaphone

가 Soundex ,
Soundex . Soundex ,

(grouping) 가
() . Soundex 가 , Metaphone

Metaphone

‘GN’, ‘KN’ ‘PN’
‘N’ . Metaphone 가
가 ,

3

Letter	Code	Conditions
A	A N/A	If first letter Otherwise
B	N/A B	No code if at end of word as in LAMB Otherwise
C	N/A X S K	No code if -SCE-, -SCI-, or -SCY- If -CIA-, or -CH- If -CI-, -CE-, -CY- Otherwise

3. Metaphone 가

Metaphone Soundex 17
(5.2.1).

ODP Soundex
Metaphone .
6734 Soundex key
가 (/6734)

4.1.3 Minimum Edit Distance

Minimum Edit Distance
[Ver88] . (edit
distance) , ,
,
Damerau 가 80 가 1
,
Damerau [Ver88].
, 'clarj'

'clare' 'clark'가
 'clark' 'clare'
 'j'
 'e' 'j' 'k' 가 가
 Soundex Metaphone
 가 가 가
 가
 Mininum Edit Distance

4.1.4 Bickel

Bickel (unigram probabilities)
 [Bic87] Bickel
 (weight)
 가 'Q' 'Z'
 'A', 'S'
 가 (weight) 4

Weight	Letters
3	A, E, I, N, O, S, T
4	D, H, L, R, U
5	C, F, G, M, P, W
6	B, V
7	K, Q
8	J, X, Y
9	Z

19 . , Baxter 16 , Bakefield
Barker 가 가 .

가 , (Likeness value)
Soundex
Metaphone 가 / (TRUE/FALSE)
Bickel .

Bickel Soundex
Metaphone .
Soundex Metaphone 가
가 . Soundex Metaphone
Bickel

Soundex
Metaphone , Soundex
Metaphone ,
(response time)
가 Bickel
가 .

4.1.5

WWW

ODP

Browson

가

[IR90a].

가

가

가

가

가

가

(, ,)

USE UF(used for), NT(narrow term)
BT(broader term), RT(related term)

가

USE UF

. USE , (: Continuous learning USE
lifelong learning), (: Intemperance USE Temperance),
(: Colour USE Color), (:
Linguistics descriptive USE Descriptive linguistics),

(: International Standard Organization USE ISO),
(: USE) .

가 . Soundex Metaphone Bickel Damerau
(logic) 가 ,
, 가 .
(update)가 가 ,
가 .

4.2

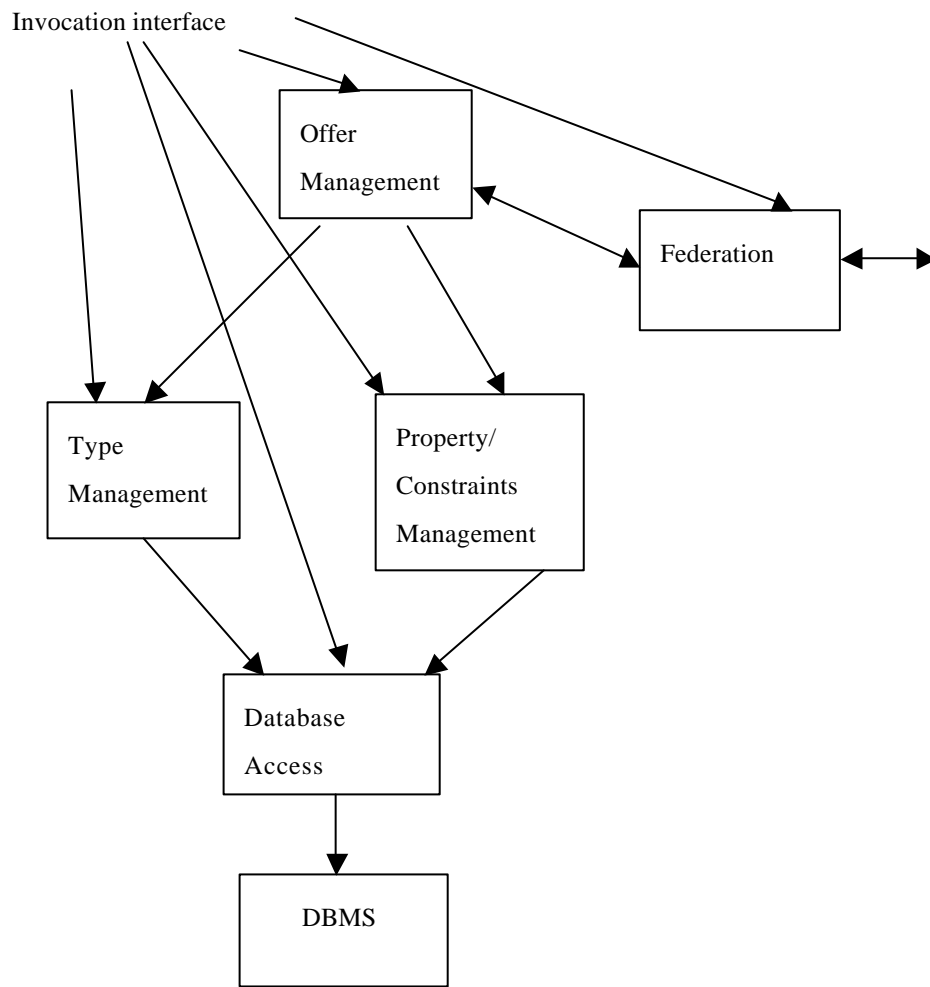
ODP [ISO94] ODP
X.500 WWW
X.500
가 . ODP
(White page)
X.500
WWW

4.3.1

가 , 가

7

[Gom94].



7.

(invocation interface)

가

?

?

?

(interworking)

Offer Management

(operation)

(criteria)

Type Management

가

(engineering viewpoint)

(type definition)

가

(match)

Property/Constraint Management

(matching)

Federation

DataBase Access
(design viewpoint)

DBMS

4.3.2

ODP

가 ,

가

가

8 - 10

7

Property/Constraint Management

가

Match()

. Property/Constraint Management

Parser ,

가

Coordinator 가

(8 - 10)

Coordinator

Match

4.3.2.1 Metaphone

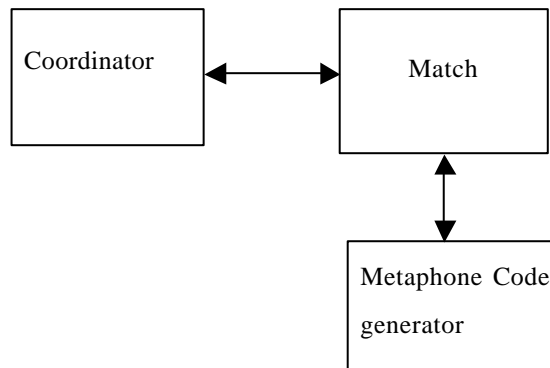
()

Metaphone

(4.1.2)

Metaphone

가 . 가
 Match Metaphone
 가 . Metaphone (Code
 Generator) 가



8. Metaphone 가 Match

Match (Coordinator)
 (Criteria) 가 (PropertyList)
 . (Criteria)
 Metaphone
 Metaphone 가
 (PropertyList) Metaphone
 Metaphone . Metaphone
 Metaphone (match)
 (Boolean) . Match .

```

CORBA::Boolean Match(const ObjTrader::PropertyList& properties,
                      const ObjTrader::PropertyList& criteria);
  
```



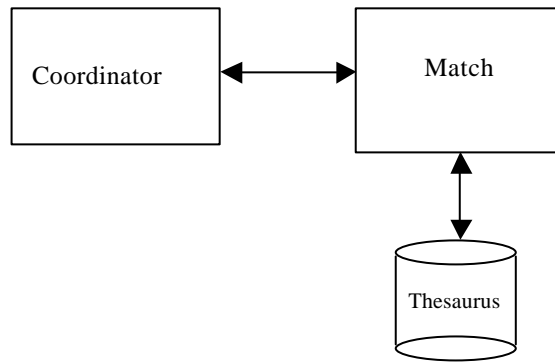
```

Metaphone Code Generator      Match
      Metaphone              Metaphone
      . Metaphone            (Metaphone Code Generator)
      .
ObjTrader::MetaphoneCode      Metaphone_Code_Generator(
                              (const char *) name);
                              ,
                              Metaphone

```

4.3.2.2

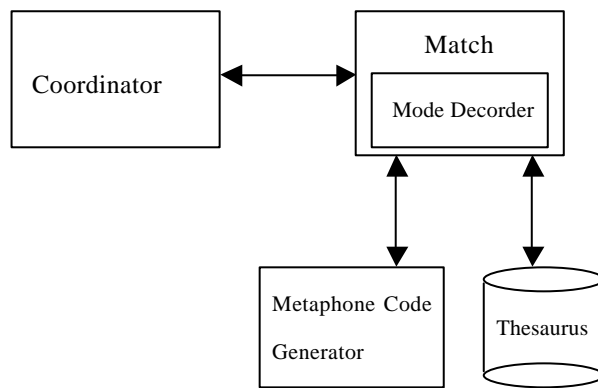
9



9. Match

Match (Criteria) (Coordinator) 가 (PropertyList)

Match (Coordinator)
 (Criteria) 가 (PropertyList)
 (parsing)
 가
 (PropertyList)
 (Boolean)



10. Match

(Mode Decoder)

5 가
 가 Metaphone Metaphone
 가 Thesaurus Thesaurus
 가 Thesaurus

Metaphone (Thesaurus) Metaphone 4.3.2.1
 Metaphone 4.3.2.2

Property Mode	Meaning
Normal	Property can be updated or deleted.
Read_only	Property cannot be updated, but can be deleted.
Fixed_normal	Property can be updated, but not deleted.
Fixed_readonly	Property cannot be updated, but can not be deleted.
Metaphone	Property can be matched by Metaphone.
Thesaurus	Property can be matched by Thesaurus.
Undefined	Not relevant; only used in exceptions.

5.

5 .

5 . 1

1997 CORBA [Ljs97] 'WWW

5 . 1 . 1

가 ,

가 .

가

가 , 가
 . 가 가
 가 가
 가 ,
 가 ()
 , ABC bus Co. 가
 , , 10000
 , ABC bus Co.

```

transport=bus
company=ABC bus Co.
dest=Taegu
fare=10000
speed-class=express

```

가 , 가
 . , 가
 dest=Taegu fare>8000 fare<11000
 (Criteria)
 가
 , ABC bus Co. 가

ABC bus Co.

(bind)

가 .

가

Solaris 2.5 [Sun97] SunSparc
[Sun97] , 가
Sun Sparc 20 [Sun97] .

Ultra sparc 1
Solaris 2.4

OrbixWeb 2.0 [Orweb]

, 가 WWW
Java [Jav97]가 .
C++ .

5.1.2 Soundex Metaphone

(Name) dest 가 .
Soundex [Knu73] Metaphone

가 dest=Taegu dest=Daegu

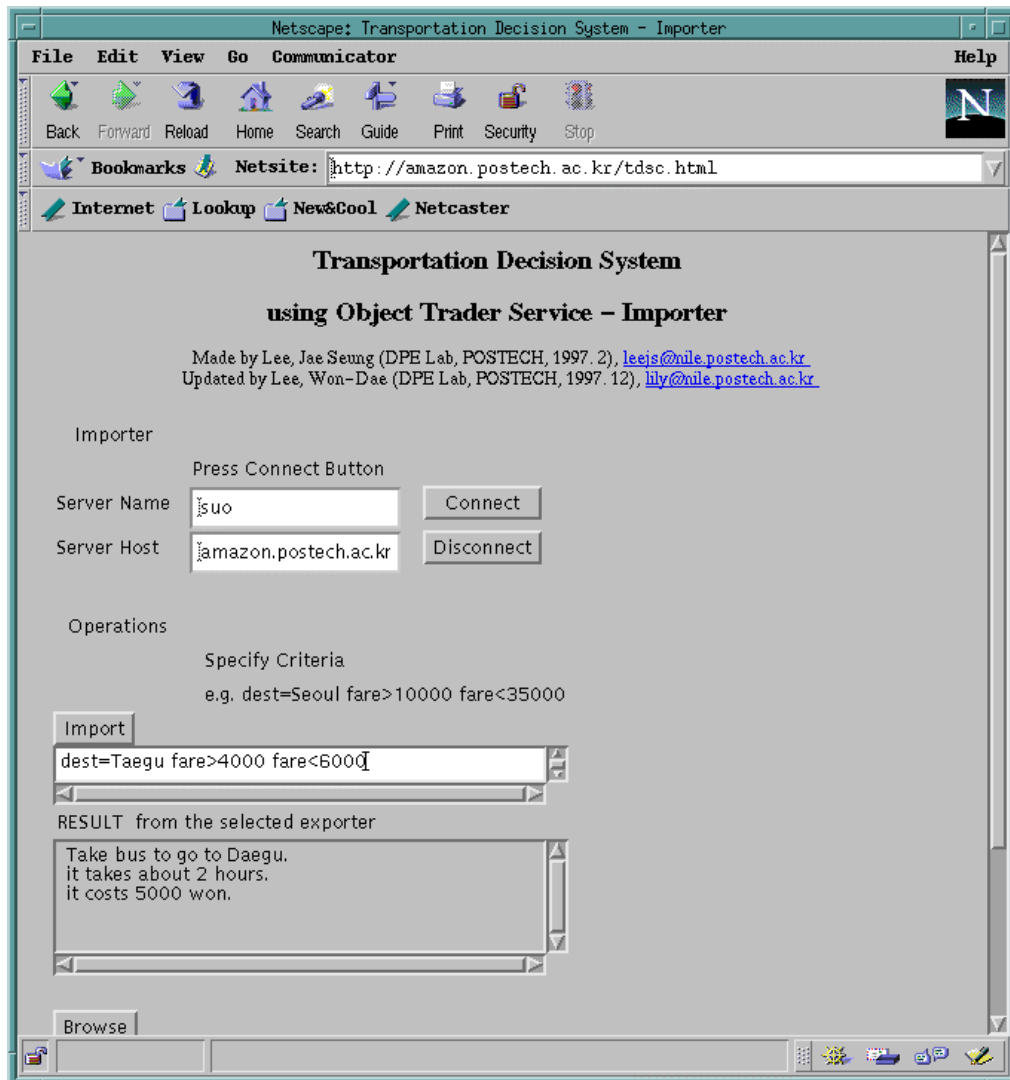
dest=Daegu fare>4000 fare<6000

Metaphone

11

Daegu

Taegu

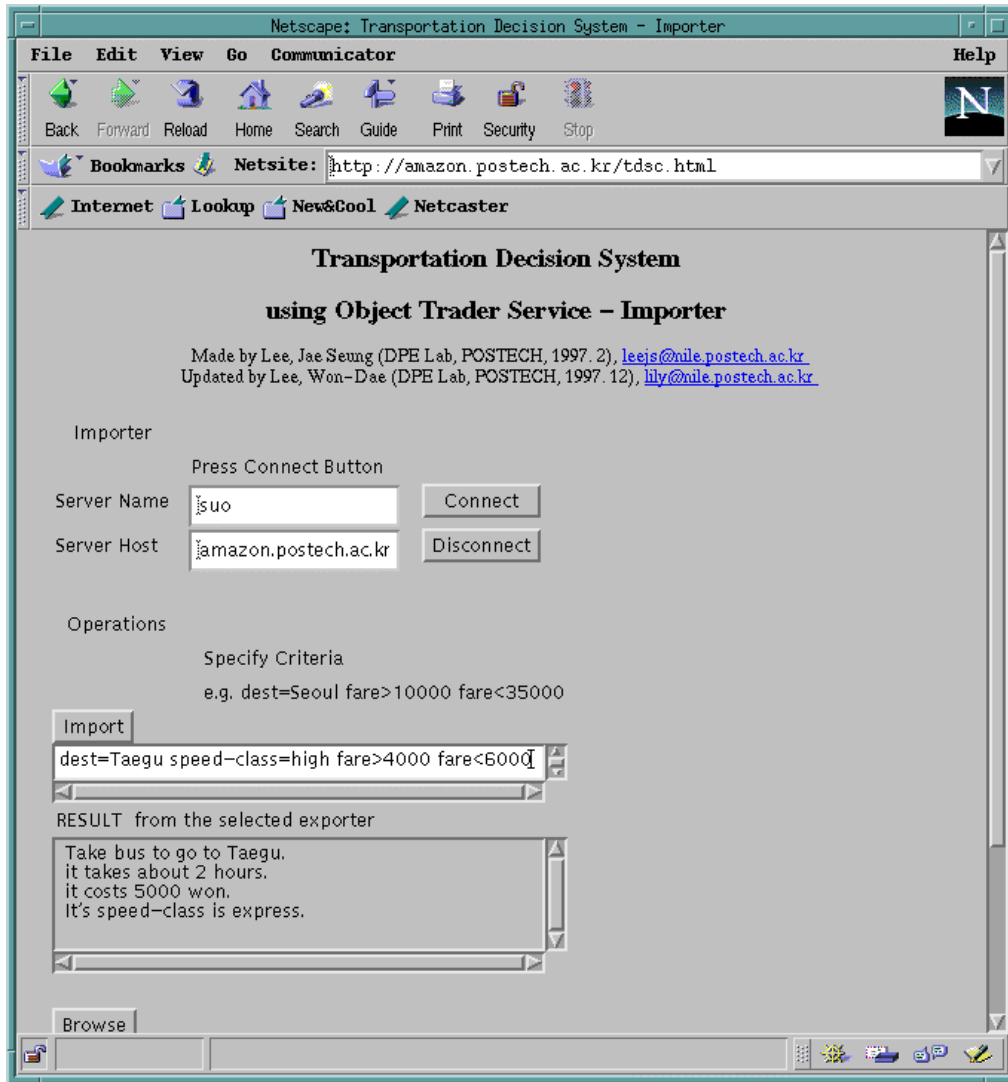


11. Metaphone

5.1.3

(Name) 가 'speed-class'

가 .



12.

(4.1.5)

가 가 , .
 Wordnet [Wor93] .
 가

dest=Taegu speed-class=high fare>4000 fare<6000

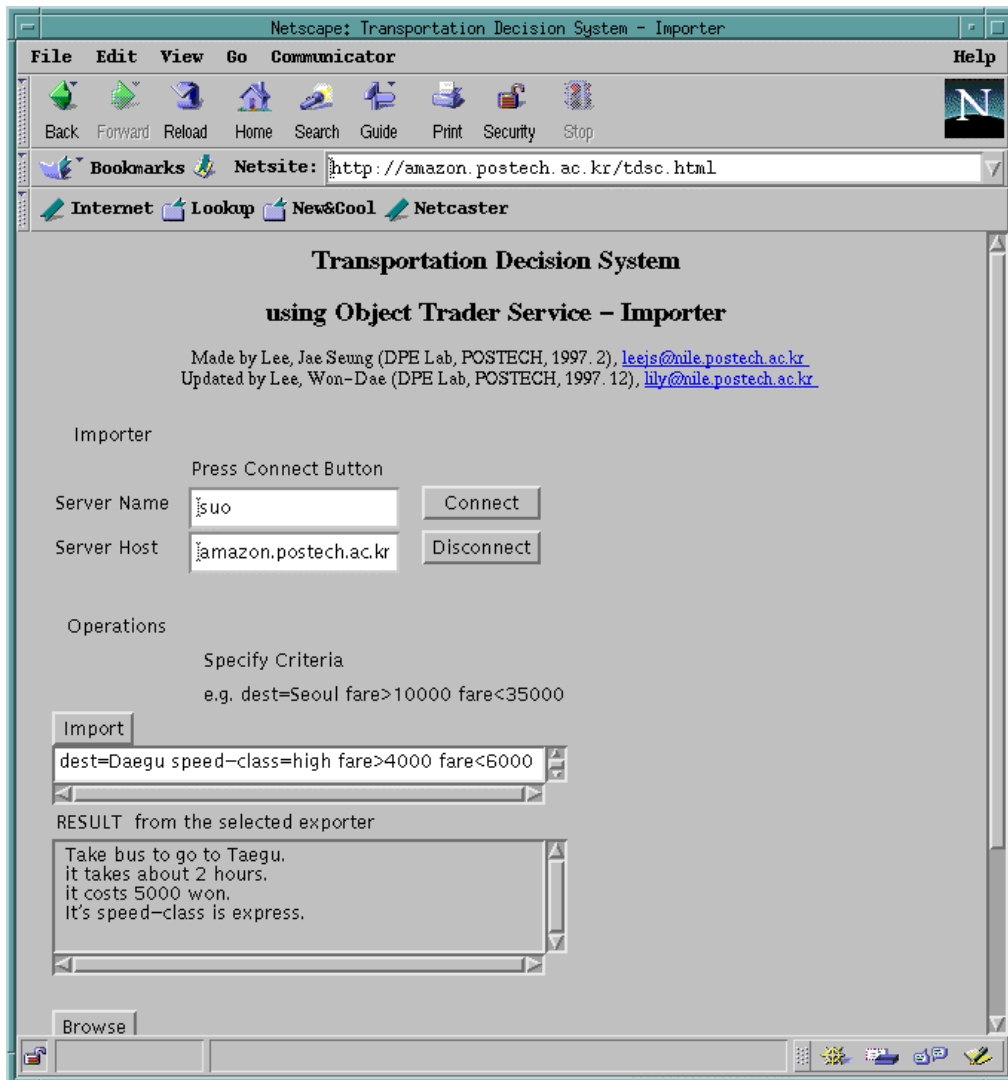
speed-class=high
 12 .
 가 speed-class=express 가 speed-clas=high
 , (RESULT) speed-class=express
 가 .

5.1.4

가

dest=Daegu speed-class=high fare>4000 fare<6000

13 .
 가 dest=Daegu ,
 speed-class speed-class=high .
 Metaphone Daegu Taegu ,
 high express .



13.

5.2

5.2.1 Soundex Metaphone

Soundex Metaphone
 iteration(: for) Big-Oh
 [Alg92] $O(N)$ N

Soundex (recall) [Inr92]
 Soundex
 Soundex 가 ,
 Soundex 가 ,
 Kukich 가 , Soundex (coarse-grained)
 [Kuk92]. Soundex 가 ,
 가 ,
 가
 key 가 . key
 가 key $26*6*6*6$
 key $26*6*6$, key
 $26*6$, key
 26 . key
 (key space) 6734 .
 가 .

$$(26*6*6*6) + (26*6*6) + (26*6) + 26 = 6734 \text{ key spaces}$$

Soundex

'S' , 'Seow' Soundex
 'Suuya', 'Shaha' 가
 'Fan' 'Fawehinmi'
 Soundex 'F5'
 Metaphone . Metaphone
 key (key space) 114,840 keys = 22*17*17*17
 가 22 17 가
 Big-0h
 for 가 0(N) N
 Metaphone , Soundex
 key if (check) Soundex
 , key Soundex
 Metaphone
 가 ,
 가 가
 Soundex Metaphone
 가
 Metaphone
 가

Metaphone

Metaphone

5.2.2

[Alg92].

$O(1)$

가

(update)

가

가

5.2.3

(worst case)

(best case)

가

$O(N) + O(1) = O(N)$

$O(1)$

가

가

ODP

가 가

5.2.4 AI-Trader

AI-Trader

Metaphone

Thesaurus

가

가

: dest=Taegu, time=2, fare=5000

: dest=Daegu, time=2, fare=5000

: dest=Taeku, time=1, fare=5000

: dest=Busan, time=1, fare=3000

(criteria):

dest=Taegu, time=2, fare=5000

14 AI-Trader

15

15

가

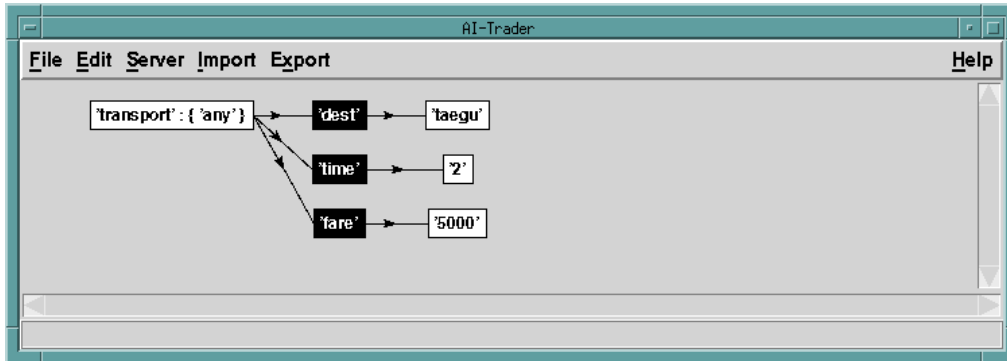
가

100

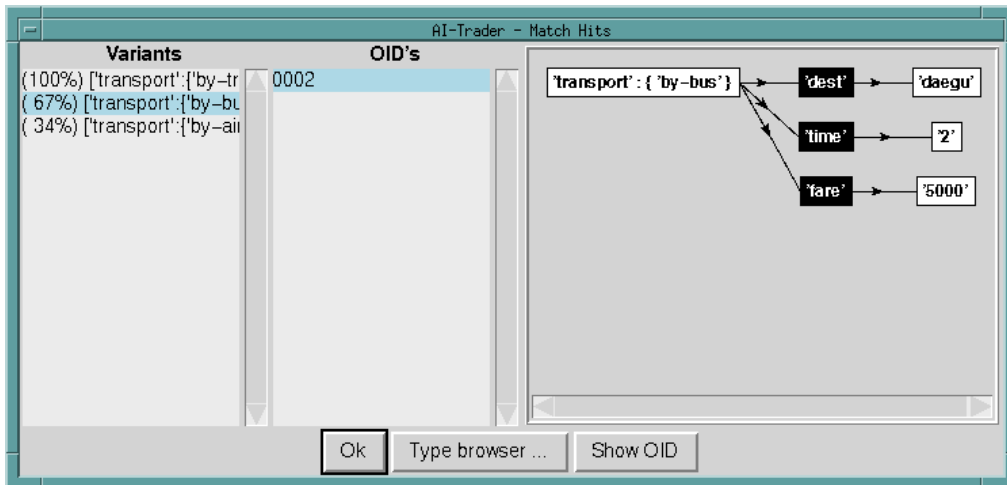
100

(%) = / *

100



14. AI-Trader



15. AI-Trader

16. Metaphone

Metaphone . AI-Trader

, Metaphone dest=taegu

가 .

Metaphone

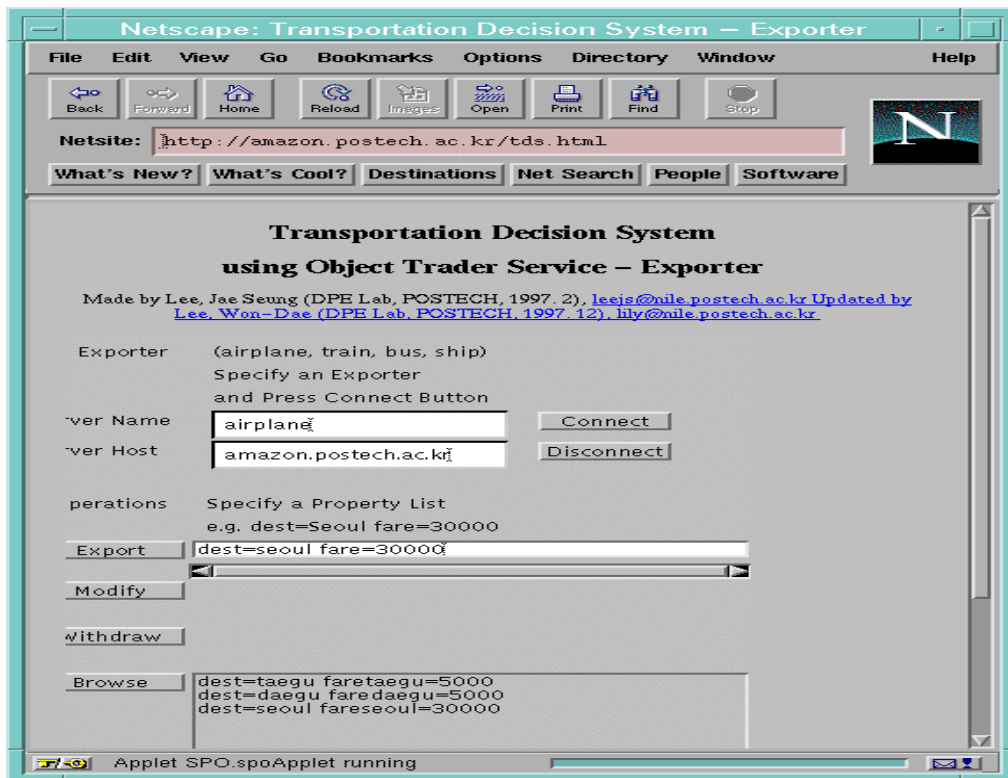
. Metaphone

dest=Seoul 가 dest=capital

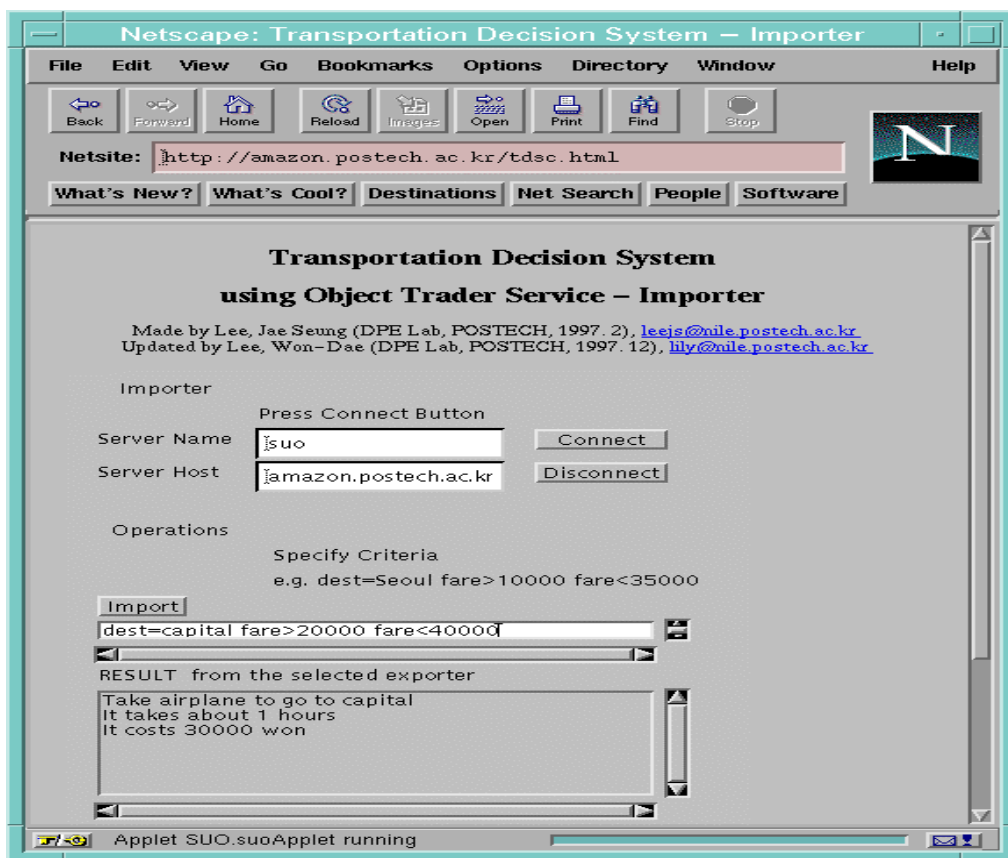
가 .

17

18



17.



18.

6 .

ODP [IS094]
(service categorization)

Metaphone

AI-Trader

Metaphone

(interworking)

가

- Data structures & algorithms, Prentice Hall 1992.
- [Jyh97] " ODP ", , 1997.
- [Jav97] The Source for Java,
Available from URL=<http://java.sun.com/>
- [Knu73] D.E. Knuth. The Art of Computer Programming, volume 3. Addison Wesley, 1973.
- [Kuk92] K. Kukich. "Automatically correcting words in text." ACM Computing Surveys, 24(4):377-439, December 1992.
- [Ljs96] , "CORBA ODP ",
96 7, 23 2 , pp.1073-1084, 1996.
- [Ljs97] "WWW CORBA ",
, 1997.
- [Mav95] B. Marvrogeorge. Soundex history and methods. Genealogy SF, 1995. Available from URL=<http://roxy.sfo.com/~genealogysf/glenda.html>.
- [Mir91] Mirion Bearman and Kerry Raymond, "Federating Traders: An ODP Adventure", Proceedings of the IFIP TC6/WG6.4 International Workshop on Open Distributed Processing, Berlin, Germany, 8-11 October, 1991, publ. North-Holland, pp.125-141.
- [Mir93] Mirion Y. Bearman, CRC for Distributed Systems Technology, Faculty of Information Sciences and Engineering, University of Canberra, "ODP-trader", 1993.
- [NBGS92] G. Neufeld, B. Brachman, M. Goldberg, and D. Stickings. "The EAN X.500 directory service." Internetworking: Research and Experience, 3:55-81, 1992.
- [OMG92] Object Management Group. "Object Management Architecture Guide, second edition", Sept. 1992. Revision 2.0, OMG TC Document 92.11.1.
- [OMG95] OMG, "The Common Object Request Broker: Architecture and Specification Revision 2.0", OMG TC Document, July 1995.
- [OR18] M.K. Odell and R.C. Russell. US patent nos. 1,261,167(1918) and 1,435,663(1922), 1918. US Patent Office, Washington DC.
- [Orweb] OrbixWeb Home Page,
Available from URL=<http://www.iona.com/Orbix/OrbixWeb/index.html>
- [Phi90] L. Philips. "Hanging on the Metaphone." Computer Language, December 1990.
- [Rov96] Robert Orfali, Dan Harkey, Jeri Edwards, The Essential Distributed Objects Survival Guide, John Wiley & Sons, Inc., 1996.
- [Sun97] Sun Microsystems,
Available from URL=<http://www.sun.com/products-n-solutions>
- [TBe95] T. Bernets-Lee, R. Cailliau, A. Luotonen, H. F. Nielsen and A. Secret, "The World-Wide Web", Communications of the ACM, Vol. 37, No. 8, August 1994, pp. 76-82.

- [Typ95] Framingham Coporate Center, "OMG Extended IDL types RFP: Annex B", OMG TC Document 94-11-6. 1994.
- [Ver88] J. Veronis. "Computerized correction of phonographic errors." Computers and the Humanities, 22:43-56, 1988.
- [Way95] Wayne Brookes and Jadwiga Indulska, "A Type Management System for Open Distributed Processing" , Technical Report 285, Department of Computer Science, The University of Queensland, February 1994.
- [War92] Ward Rosenberry, David Kenney, and Gerry Fisher, Understanding DCE, O'Relly & Associates, Inc., September 1992.
- [Wor93] G.A. Miller et al. "Five papers on WordNet". Technical Report CSL Report 43, Cognitive Science Laboratory, Princeton University, March 1993.
- [X500a] International Telecommunication Union, Telecommunication (ITU-T) X.500 series of Recommendation, "Information Technology - Open System Interconnection - The Directory", 1994.

A. Metaphone

Letter	Code	Conditions
A	A	If first letter
	N/A	Otherwise
B	B	
C	N/A	No code if -SCE-, -SCI- or -SCY-
	X	If -CIA- or -CH-
	S	If -CI-, -CE-, -CY-
	K	Otherwise
D	J	If in -DGE-, -DGI-, -DGY-
	T	Otherwise
E	E	If first letter
	N/A	Otherwise
F	F	
G	N/A	No code if in -GH- and not at end or before a vowel
	N/A	If in -GN or -GNED, or if in -DGE-, -DGI-, -DGY-
	J	If before I, or F, or Y and not GG
	K	Otherwise
H	N/A	Silent if after vowel and no vowel follows
	N/A	Silent if after C, G, P, S, T
	H	Otherwise
I	I	If first letter
	N/A	Otherwise
J	J	

*

Letter	Code	Conditions
K	N/A	No code if after C
	K	Otherwise
L	L	
M	M	
N	N	
O	O	If first letter
	N/A	Otherwise
P	F	If before H
	P	Otherwise
Q	K	
R	R	
S	X	If before H or in -SIO- or -SIA-
	S	Otherwise
T	X	If -TIO- or -TIA-
	Zero	If before H
	N/A	Silent if in -TCH-
	T	Otherwise
U	U	If first letter
	N/A	Otherwise
V	F	
W	N/A	Silent if not follow by a vowel
	W	Otherwise
X	KS	
Y	N/A	Silent if not followed by a vowel
	Y	Otherwise
Z	S	

B.

<i>Node type</i>	<i>Name</i>	<i>Description</i>
Concepts	SERVICE-TYPE	가
	INTERFACE-TYPE	
	PROPERTY-TYPE	
	SIGNATURE	
Relations	CONSISTS-OF	/
	IN-PARAMETER	
	OUT-PARAMETER	
	INOUT-PARAMETER	/
	STATIC-ATTRIBUTE	
	DYNAMIC-ATTRIBUTE	
	EVAL-CALLBACK	evaluation callback
	ANNOTATED-BY	