CHAPTER 6 : REFERENCES

[1] N. Gupta, G. Tur, D. Hakkani, G. Riccardi and M. Gilbert, *The AT&T Spoken Language Understanding System*. IEEE Trans. Speech Audio Process, Jan. 2006, vol. 14, no. 1, pp. 213-222.

[2] A.M.Pasca, and S.M. Harabagiu, *High Performace Question/ Answering*. New Orleans, Louisiana, USA. SIGIR, 2001.

[3] A.L. Gorin, G. Riccardi, and J.H. Wright, *How may I help you?*, Speech Commun., 1997, vol. 23, pp. 113-127.

[4] E. Bril, S. Dumais, and M. Banko, *An analysis of the AskMSR Question-Answering System*, In MNLP2002.

[5] A. Berger, and V.O. Mittal, *Query - Relevant Summarization using FAQs*, Proceedings of the 38th Annual Meeting on Association for Computational Linguistics - Stroudsburg, PA, USA, 2000.

[6] Apache Software Foundation (May 2012). "*Apache Lucene*" [Online].Available: http://lucene.apache.org/java/docs/index.html.

[7] S. Seneff, *TINA: A natural language system for spoken language applications*, Computat. Linguist., 1992, vol. 18, no.1, pp. 61-86.

[8] W. Ward, and S. Issar, *Recent improvements in the CMU spoken language understanding system*, in Proc.ARPA HLT Workshop, Mar. 1994, pp. 213-216.

[9] J. Dowding, J.M. Gawron, D. Appelt, J. Bear, L. Cherny, R. Moore,and D. Moran, *Gemini: A natural language system for spoken language understanding*, in Proc. ARPA Workshop on Human Language Technology, Princeton, NJ, Mar. 1993.

[10] S.M. Harabagiu, M.A. Pasca, and S.J. Maiorano, *Experiments with Open-Domain Textual Question Answering*, 18th International Conference on Computational Linguistics - Saarbrucken, Germany, COLING 2000.

[11] G. Miller, *WordNet a lexical database*, Communications of the ACM,1995.

[12] W. Lehnert, *The Process of Question Answering*, Lawrence Erlbaum,1978.

[13] M. Klaus et al., *JGAP - Java Genetic Algorithms and Genetic Programming Package*, URL: http://jgap.sf.net.

[14] C.D. Manning, P. Raghavan, and H. Schtze, *Introduction to Information Retrieval*, Cambridge University Press, NewYork, NY, 2008.

[15] D. Jurafsky, and J.H. Martin, *Speech and Language Processing (2nd Edition)*, Prentice-Hall, Inc., Upper Saddle River, NJ, 2009.

[16] L. Hirschman, and R. Gaizauskas, *Natural language question answering: The view from here*. Journal of Natural Language Engineering, Special Issue on Question Answering, Fall2001.

[17] G. Salton, *The SMART Retrieval System: Experiments in Automatic Document Processing*. Prentice Hall. 1971.

[18] D. Molla, and J.L. Vicedo, *Question Answering in Restricted Domains: An Overview*, Special Section on Restricted-Domain Question Answering, Association for Computational Linguistics, 2007.

[19] N.F. Noy, and D.L.McGuinness,*Ontology development 101: A guide to creating yourfirstontology*.Technical Report KSL-01-05,Knowledge Systems Laboratory, 2001.

[20] S. Auer, C. Bizer, G. Kobilarov, J.Lehmann, and Zachary Ives, *DBpedia: A Nucleus for a Web of Open Data*, In 6th Int'l Semantic Web Conference, Busan, Korea, 2007.

[21] (May 2012).*Wikipedia*, [Online].Available:http://www.wikipedia.org/

[22] W Roush, *The story of Siri, from birth at Sri to acquisition by Apple: Virtual personal assistants go mobile*. Xconomy, June 2010.

[23] (May 2012). *TREC*, [Online]. Available:http://trec.nist.gov

[24] (May 2012). *Wolfram Alpha*,[Online].Available:http://www.wolframalpha.com

[25] (May 2012). *IBM Watson*,[Online].
Available:http://www.aaai.org/Magazine/Watson/watson.php

[26] (May 2012). *True Knowledge Evi*,[Online].Available:http://www.evi.com/

[27] (May 2012). *Google*,[Online].Available:http://www.google.com/

[28] B. F.,Green, A. K.Wolf, C. Chomsk y, andK. Laughery, *Baseball: An* automatic *question answerer*. In Proceedings Western Computing Conference, 1961, vol 19,pp. 219–224.

[29] WA Woods, *Lunar rocks in natural English: Explorations in natural language question answering*. Linguistic structures processing, North-Holland, New York. 1977.

[30] W. Robert, N.C. David, M. Luria, J. Martin, J. Mayfield, and D. Wu, *The Berkeley Unix Consultant project*. Computational Linguistics, 1994,14(4),pp.35–84.

[31] (May 2012). *Paraphase Detection*, ACL web,[Online].

Available:http://aclweb.org/aclwiki/index.php?title=Paraphrase_Identification_(State_of_the_art).

[32] S. Fernando, and M. Stevenson, *A semantic similarity approach to paraphrase detection*, Computational Linguistics UK.11th Annual Research Colloquium. 2008.

[33] A. Finch, Y.S. Hwang, and E. Sumita, *Using machine translation evaluation techniques to determine sentence-level semantic equivalence*, "Proceedings of the Third International Workshop on Paraphrasing", Jeju Island, South Korea, 2005,pp. 17-24.

[34] A. Islam, and D. Inkpen, *Semantic similarity of short texts*, Proceedings of the International Conference on Recent Advances in Natural Language Processing, Borovets, Bulgaria, 2007, pp. 291-297.

[35] Z. Kozareva, and A. Montoyo, *Paraphrase identification on the basis of supervised machine learning techniques*, Advances in Natural Language Processing: 5th International Conference on NLP.Turku, Finland, 2006,pp. 524-533.

[36] R. Mihalcea, C. Corley, and C. Strapparava, *Corpus-based and knowledge-based measures of text semantic similarity*, Proceedings of the National Conference on Artificial Intelligence, Boston, Massachusetts, 2006, pp. 775-780.

[37] L. Qiu, M.Y. Kan, and T.S. Chua, *Paraphrase recognition via dissimilarity significance classification*, Proceedings of the Conference on Empirical Methods in Natural Language Processing,2006, pp. 18-26.

[38] V. Rus, P.M. McCarthy, M.C. Lintean, D.S. McNamara, and A.C.Graesser, *Paraphrase identification with lexico-syntactic graph subsumption*, FLAIRS 2008, pp. 201-206.

[39] R. Socher, E.H. Huang, J. Pennington, A.Y. Ng, and C.D. Manning, *Dynamic pooling and unfolding recursive auto encoders for paraphrase detection*, "Advances in Neural Information Processing Systems 24".2011.

[40] S. Wan, M. Dras, R. Dale, and C. Paris, *Using dependency-based features to take the "para-farce" out of paraphrase*, Proceedings of the Australasian Language Technology Workshop (ALTW 2006), pp. 131-138.

[41] D. Demner-Fushman, and J. Lin, *Knowledge extraction for clinical question answering: Preliminary results.* In Workshop on Question Answering in Restricted Domains. 20th National Conference on Artificial Intelligence, Pittsburgh, PA, 2005,pp. 1–9.

[42] E.T.K. Sang, G.Bouma, and M. Rijke, *Developing offline strategies for answering medical questions.* In Workshop on Question Answering in Restricted

Domains.20th National Conference on Artificial Intelligence (AAAI-05), Pittsburgh, PA, 2005, pp. 41–45.

[43] Y. Niu, and G. Hirst, *Analysis of semantic classes in medical text for question answering.* In Workshop on Question Answering in Restricted Domains. 42nd Annual Meeting of the Association for Computational Linguistics (ACL-2004), Barcelona, Spain, 2004, pp. 54–61.

[44] H. Yu, C. Sable, and H.R. Zhu. *Classifying medical questions based on an Evidence Taxonomy*. In Workshop on Question Answering in Restricted Domains. 20th National Conference on Artificial Intelligence(AAAI-05), Pittsburgh, PA, 2005,pp. 27–35.

[45]F. Benamara, *Cooperative question answering in restricted domains: The WEBCOOP Experiments*. In Workshop on Question Answering in Restricted Domains.42nd Annual Meeting of the Association for Computational Linguistics (ACL-2004), Barcelona, Spain, 2004, pp. 31–38.

[46] K. Barker, V.K. Chaudhri, S.Y. Chaw, P. Clark, J. Fan, D. Israel, S. Mishra, B.W. Porter, P. Romero, D. Tecuci, and P.Z. Yeh, *A Question-answering system for AP chemistry: Assessing KR&R technologies. In Principles of Knowledge Representation and Reasoning:* Proceedings of the Ninth International Conference, Whistler, Canada, 2004,pp. 488–497.

[47] D. Lenat, and R.V. Guha, *Building Large Knowledge-Based Systems: Representation and Inference in the Cyc Project*. Addison-Wesley,1990.

[48] B. Porter, J. Lester, K. Murray, K. Pittman, A. Souther, L. Acker, and T. Jones, *AI research in the context of a multifunctional knowledge base: The botany knowledge base project.* Technical Report, 1988, Department of Computer Sciences, University of Texas at Austin.

[49] P. Cohen, R. Schrag, E. Jones, A. Pease, A. Lin, B. Starr, D. Easter, D. Gunning, and M. Burke, *The DARPA high performance knowledge bases project.* AI Magazine, 1988, 19(4), pp.25–49.

[50] R. Schrag, M. Pool, V. Chaudhri, R.C. Kahlert, J. Powers, P. Cohen, J. Fitzgerald, and S. Mishra, *Experimental evaluation of subject matter expert-oriented knowledge base authoring tools.* In Proceedings of the PerMIS Workshop: Measuring the Performance and Intelligence of Systems, Gaithersburg, MD. 2002, pp. 272–279.

[51] R. Zajac, *Towards ontological question answering*. In Proceedings of ACL,Workshop on Open Domain QA, Toulouse,2001.

[52] D. Klein and C.D. Manning, *Conditional structure versus conditional estimation in NLP models*, EMNLP '02 Proceedings of the ACL-02 conference on Empirical

methods in natural language processing. 2002. vol: 10, pp. 9-1.

[53] (May 2012). *Microsoft Research Paraphrase Corpus*, [Online]. Available: http://research.microsoft.com/en-us/downloads/607D14D9-20CD-47E3-85BC-A2F65CD28042/default.aspx.

[54] C. Leacock, and M.S. Chodorow, *Combining local context and WordNet similarity for word sense identification.* In Fellbaum, C. (Ed.), Wordnet: An Electronic Lexical Database, MIT Press,1998,pp. 265–283.

[55] P. Resnik, *Using information content to evaluate semantic similarity in a taxanomy*. In International Joint Conference for Artificial Intelligence (IJCAI-95), 1995,pp. 448–453.

[56] Lin, *An information-theoretic definition of similarity*. San Francisco, In *ICML 1998*, pp. 296–304.

[57] J.J. Jiang, and D.W. Conrath, *Semantic similarity based on corpus statistics and lexical taxonomy*. In ROCLINGX, Taiwan. 1997.

[58] S. Banerjee, and T. Pedersen, *Extended gloss overlaps as a measure of semantic relatedness.* In *IJCAI 2003*, pp. 805–810.

[59] F. Lauer, and G. Bloch, *Incorporating prior knowledge in support vector machines for classification: A review*. Neurocomputing vol. 71, Issues 7–9, March 2008, pp. 1578–1594.

[60] V. Vapnik, *Statistical Learning Theory*. John Wiley and Sons, Inc., New York, 1998.

[61] R.E. Schapire, M. Rochery, M. Rahim, and N. Gupta, *Boosting With Prior Knowledge for Call Classification*, IEEE Transactions on Speech and Audio Processing, vol. 13, No 2, March 2005.

[62] B. Zadrozny, J. Langford, and N. Abe,*Cost-Sensitive Learning by Cost-Proportionate Example Weighting*, ICDM 2003.

[63] (May 2012). *LibSVM*, [Online]. Available:http://www.csie.ntu.edu.tw/~cjlin/libsvm/

[64] C.W Hsu, C.C. Chang, and C.J. Lin, *A Practical Guide to Support Vector Classification.*

[65] (May 2012). *Standford Dependency Parser* Available: http://nlp.stanford.edu/software/lex-parser.shtml

[66] G. Riccardi, and A.L. Gorin, "*Spoken language adaptation over time and state in a natural spoken dialog system*" IEEE Trans. Speech AudioProcess,2000, vol. 8, no. 1, pp. 3–10.

[67] A. Jönsson, and N. Dahlbäack, "*Talking to a computer is not like talking to your best friend*," in Proc. 1st Scandinavian Conf. Artificial Intelligence, Tromsø, France, 1988.

[68] A.R. Diekema, O.Yilmazel, and E.D. Liddy. "*Evaluation of restricted domain question-answering systems*". In Workshop on Question Answering in Restricted Domains.42nd Annual Meeting of the Association for Computational Linguistics (ACL-2004),Barcelona, Spain,2004.pp. 2–7.

[69] G.D. Fabbrizio, D. Dutton, N. Gupta, B. Hollister, M. Rahim, G. Riccardi, R. Schapire, and J. Schroeter, "*AT&T help desk*" in Proc. Int. Conf. Spoken Language Processing, Denver, CO, Sep. 2002.

[70] P. Haffner, G. Tur, and J. Wright, "*Optimizing SVM's for complex call classification"* in Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing, Hong Kong, 2003.

CHAPTER 7 : APPENDIX A - RULE BASED CLASSIFIER

It is important to cluster the knowledge base into service-type based clusters due to the following reasons

- Most common queries are often common to most of the service types but the answers to those queries
- User seems to sufficiently explain the service issue but the service type in the initial input

It will be shown in the following sections that sub-dialog models to collect clarifying information when service type is missing in the user's initial input can be developed strongly based on a service type-based classification rule base.

Let the user's initial input, after input normalization, be defined by the vector

$$\mathbf{x} = [w_1, w_2, w_3, \cdots, w_i, \cdots, w_n]$$

where, $w_i, i = 1, 2, 3, \cdots, n$ are the words in the input.

7.1 Major Services

For convenience, the broadband services and standard telephony services are considered as major services. For reasons given in Section 7.2, in here, the terms salient to the auxiliary services such as VoIP/MoIP, hosting services, VPN, and messaging services (email, SMS, MMS, email to fax, and so forth) are ignored. Major services may further be classified as wireless and wired (copper and fiber) services.

First, the field vectors are defined. They consist of application specific field vocabulary words and phrases. The internet or broadband field vector is given by

$$\boldsymbol{w}_{Internet} = \boldsymbol{w}_{Internet1} \cup \boldsymbol{w}_{Internet2}$$

where,

$w_{Internet1}$

$= [internet, broadband, www, authenticate, auth, bandwidth, bitrate, bit\ rate,$
$datarate, bps, speed,\ \ dataWiFi, WiFi, modem, sync, ppp, splitter, router,$
$relocate, relocation]$

$\boldsymbol{w}_{Internet2}$

$= [data, upload, download, IP, web, website, webpage, URL, brows, browser,$
$domain, FTP, search\ engine, TCPIP, ethernet, telnet, LAN, WAN, PC, computer,$
$laptop, tablet, iPad, notebook, webcam, webcast, codec]$

Internet field vector is broken up into two to make it possible to write better sub-dialog models. For instance, if $w_i \in \boldsymbol{w}_{Internet1}$only, then it is often required to know the type of broadband service to answer the query. Queries with $w_i \in \boldsymbol{w}_{Internet2}$, may be directed to a Q-A common pool "Internet".

Telephony field vector is

$\boldsymbol{w}_{Telephony}$

$= [telephony, telephone, phone, call, voice, ring, dial, tone, handset, number,$
$caller, talk, IDD, CallBack, SurePage, leave\ message, answering\ machine,$
$preselect, pre\ select, preselection, override, long\ distance]$

7.1.1 Major Wireless Services

The two major wireless services are wireless internet (wireless broadband) and wireless telephony. Wireless-service field vector is

$\boldsymbol{w}_{WirelessService}$

$= [wireless, mobile, cell, cellular, handphone, hand\ phone, cellphone, smartphone,$
$smart\ phone, iPhone, PDA, SIM, USIM, GPRS, GPRS2G, GSM, reception, signal,$
$roaming, antenna, PUK, PUC]$

Wireless-broadband field vector is

$\boldsymbol{w}_{WirelessBB}$

$= [HSPA, HSDPA, WCDMA, GPRS3G, GPRS\ 3G, dongle, USB\ modem, USB\ stick, WAP, APN, tethering]$

Classification rules for wireless services are

- Mobile Broadband: $\boldsymbol{w}_{WirelessBB} \vee (\boldsymbol{w}_{WirelessService} \wedge \boldsymbol{w}_{Internet})$
- Mobile Telephony: $(\boldsymbol{w}_{WirelessService} \wedge \boldsymbol{w}_{Telephony}) \wedge \neg \boldsymbol{w}_{Internet} \wedge \neg \boldsymbol{w}_{WirelessBB}$

For convenience, user inputs with $w_i \in \boldsymbol{w}_{WirelessService}$ may also be considered to be belonging to Mobile Telephony category. They may include the queries on the mobile carriers, mobile phones, roaming etc.

7.1.2 Major Wired Services

Major wired services are the two wired broadband services, ADSL and fiber optics, and wired telephony.

Copper-line field vector is

$\boldsymbol{w}_{CopperLine}$

$= [copper, landline, land\ line, wireline, fixed\ line, wire, receiver, PSTN, landphone, land\ phone, home\ phone]$

ADSL field vector is

$\boldsymbol{w}_{ADSL}$

$= [ADSL, ADSL+, ADSL1, ADSL2, ADSL2+, DSL, HSDSL, SHDSL, microfilter, micro\ filter]$

Fiber field vector is

$\boldsymbol{w}_{Fiber}$

$= [fiber, fibre, optic, optical, NBN, NBNCo, Opticomm, ONT, OLT, build\ drop]$

Classification rules for wired services are

- Fiber Broadband: $\boldsymbol{w}_{Fiber}$
- ADSL Broadband: $\boldsymbol{w}_{ADSL} \vee (\boldsymbol{w}_{CopperLine} \wedge \boldsymbol{w}_{Internet})$
- Wired Telephony: $(\boldsymbol{w}_{CopperLine} \wedge \boldsymbol{w}_{Telephony}) \wedge \neg \boldsymbol{w}_{Internet} \wedge \neg \boldsymbol{w}_{WiredBB}$

where, $\boldsymbol{w}_{WiredBB} \in (\boldsymbol{w}_{ADSL} \cup \boldsymbol{w}_{Fiber})$. For convenience, any user inputs with $w_i \in \boldsymbol{w}_{CopperLine}$ may also be considered to be belonging to Wired Telephony category.

7.2 Auxiliary Services

VoIP/MoIP, VPN, hosting, and messaging services are considered as auxiliary services. The messaging services considered here are: SMS, MMS, IMS, email, voicemail, email to SMS, email to fax, and FoIP. The knowledge base of the question answering system will have two main modules, namely, the major-services-module and the auxiliary-services-module. User queries that contain salient terms related to the auxiliary services, irrespective of the rest of the content words and phrases, will be directed to the auxiliary services module and served there. Hence, these questions will not be passed on to the major services module.

Field vector, $\boldsymbol{w}_{Messaging}$ consists of the terms that are common to most of the messaging services considered here.

$$\boldsymbol{w}_{Messaging} = [address\ book, inbox, outbox, message\ box, spam, send\ message, receive\ message, retrieve\ message, delete\ message, listen\ message, store\ message, save\ message, delivery\ report]$$

The SMS, email, voicemail, and facsimile field vectors, respectively, are defined as follows

$$\boldsymbol{w}_{SMS} = [SMS, SMSs, SMSes, short\ message, text\ message, exeSMS, webSMS, MMS, MMSs, MMSes, multimedia\ message, multi\ media\ message, multipart\ message, IMS, IMSs, virtual\ mobile\ number, VMN, VMNs, texting]$$

$$\boldsymbol{w}_{Email} = [email, mail, exemail, webmail, SMTP, IMAP]$$

$\boldsymbol{w}_{Voicemail} = [voicemail, voice\ mail, voice\ message, VMS, VMSs, VMSes]$

$\boldsymbol{w}_{Fax} = [fax, facsimile, FoIP]$

Classification rule for messaging services are

- Messaging Services: $(\boldsymbol{w}_{Messaging} \vee \boldsymbol{w}_{SMS} \vee \boldsymbol{w}_{Email} \vee \boldsymbol{w}_{Voicemail} \vee \boldsymbol{w}_{Fax}) \wedge \neg \boldsymbol{w}_{Hosting}$

where, $\boldsymbol{w}_{Hosting}$ is the hosting services field vector that will be defined shortly. Field vectors for VoIP and MoIP services are given below

$\boldsymbol{w}_{VoIP}$
$= [VoIP, IP\ telephony, voice\ over\ IP, internet\ voice, internet\ telephony,$
$broadband\ telephony, voice\ over\ BB, VoBB, ATA, telephone\ adapter,$
$phone\ adapter, DID, virtual\ telephone\ number, virtual\ phone\ number]$

$\boldsymbol{w}_{MoIP} = [MoIP, mVoIP, mobile\ VoIP, API]$

Classification rule for VoIP service is

- VoIP Service: $\boldsymbol{w}_{VoIP}$
- VoIP Service (possibility): $\boldsymbol{w}_{Internet} \wedge \boldsymbol{w}_{Telephony}$

Possibilities are the scenarios that need to be verified through a brief dialog with the user.

Classification rule for MoIP service is

- MoIP Service: $\boldsymbol{w}_{MoIP}$
- MoIP Service (possibilities): $\boldsymbol{w}_{VoIP} \wedge (\boldsymbol{w}_{WirelessService} \vee \boldsymbol{w}_{SMS} \vee \boldsymbol{w}_{Email})$

Classification rules for hosting and VPN services, respectively, are

- Hosting Service: $\boldsymbol{w}_{Hosting}$
- VPN Service: $\boldsymbol{w}_{VPN}$

where,

$\boldsymbol{w}_{Hosting} = [webhosting, webspace, web\ space, DNS, CMS, domain \wedge host\ ,$ $create \wedge (web \vee website \vee webpage \vee homepage \vee home\ page), server]$ and $\boldsymbol{w}_{VPN} = [VPN, (virtual \vee private) \wedge (network \vee LAN)]$.

7.3 Control Strategy

Functional block diagram of the classifier control strategy is given in Figure 7-1. The main blocks therein are

- A pool of Q-A pairs common to all services
- Auxiliary services module that contains Q-A pairs specific to auxiliary services
- Main services classifier and sub-dialogs
- Main services module that contains Q-A pairs specific to the main services and unrelated to the auxiliary services

Once a user input has been received, first of all, the algorithm checks whether the query is on any of the service issues that are common to all the services. The rule-base to identify the queries that fall into this category is based purely on the terms salient to service issues and independent of the service type as follows.

- Service issues common to all services: $\boldsymbol{w}_{CommonIssues}$

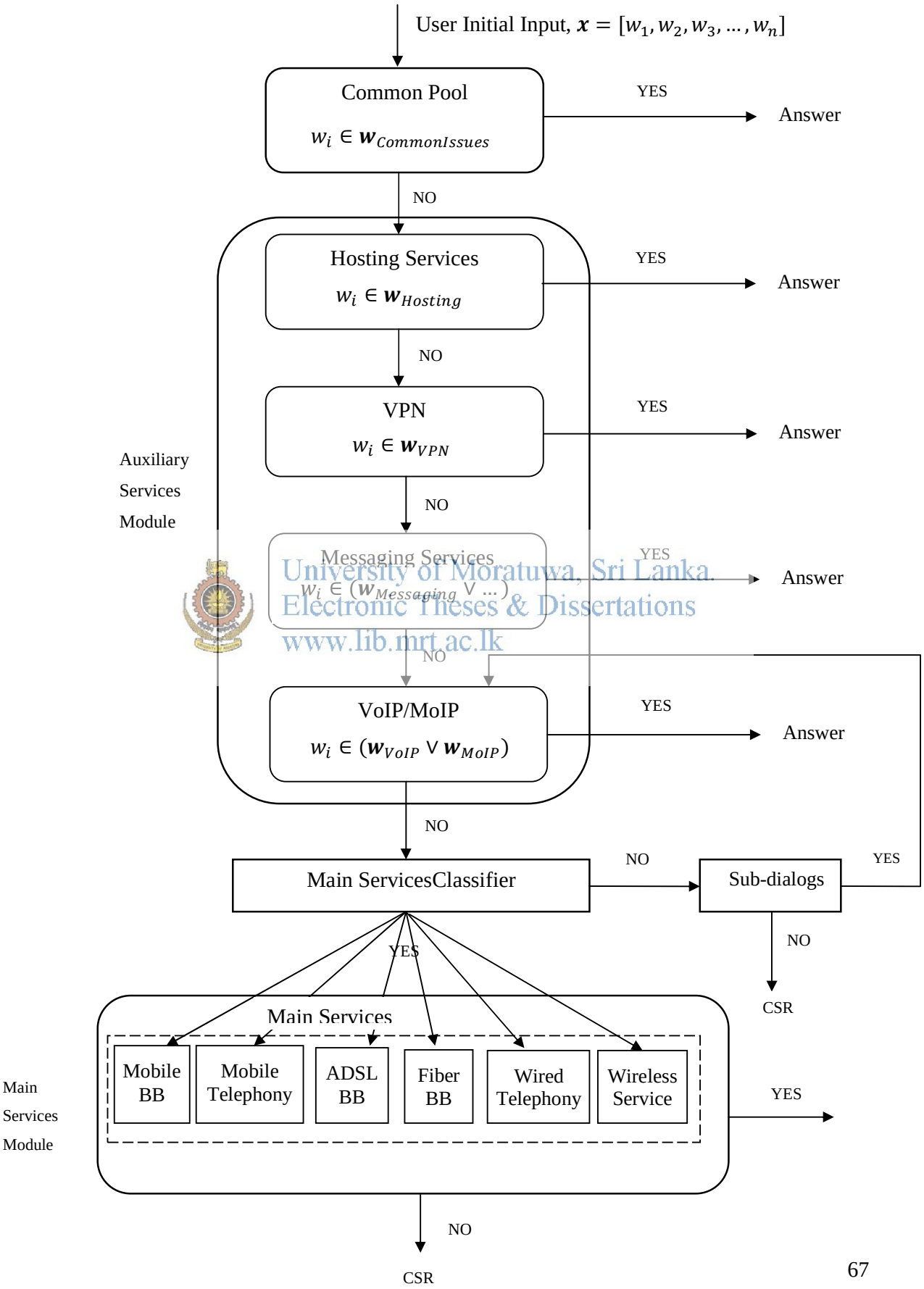
where, algorithm marks questions as common, if one or more rules mentioned below is satisfied.

R1: What ∧ billing ∧ (cycle ∨ period ∨ method)
R2: (how ∨ when) ∧ (bill ∨ charge)
R3: (pro rata ∨ prorata) ∧ (calculate ∨ calculation)
R4: Administration ∧ (fee ∨ charge)
R5: (service ∨ credit card) ∧ surcharge
R6: (want ∨ access ∨ read ∨ understand ∨ receive ∨ retrieve ∨ obtain ∨ previous ∨ old) ∧ (bill ∨ invoice)
R7: (mode ∨ method ∨ option) ∧ payment
R8: AMEX card
R9: (change ∨ edit ∨ update ∨ modify) ∧ (payment ∨ account) ∧ (details ∨ information)

R10: (adjust ∨ change) ∧ billing ∧ date
R11: (delay ∨ late ∨ overdue ∨ outstanding ∨ fail ∨ reject ∨ decline ∨ dishonor ∨ overdrawn ∨ unbilled) ∧ (payment ∨ invoice ∨ bill ∨ charge ∨ fee)
R12: Interim
R13: Excess usage
R14: Insufficient ∧ (fund ∨ money)
R15: Refund ∨ credit account back ∨ credit money back ∨ return money
R16: (merge ∨ combine ∨ connect) ∧ (account ∨ ID ∨ IDs ∨ invoice ∨ bill)
R17: (master ∨ one ∨ single) ∧ (account ∨ invoice ∨ bill)
R18: (early ∨ contract) ∧ (cancel ∨ cancellation ∨ terminate ∨ termination) ∧ (fee ∨ charge)∨ETC
R19: (contact ∨ speak) ∧ (billing ∨ sales ∨ support ∨ service ∨ department ∨ section ∨ division ∨ exetel ∨ details)
R20: (transfer ∨ change) ∧ ownership
R21: (How ∨ want ∨ need) ∧ (cancel ∨ terminate ∨ unsubscribe ∨ quit ∨ …) ∧ (ADSL ∨ fiber ∨ broadband ∨ BB ∨ mobile ∨ exemail ∨ mail ∨ email ∨ hosting ∨ VoIP ∨ SMS ∨ service ∨ connection ∨ internet ∨ telephone ∨ plan)

Quite in contrast to the main services classifier, the classification rules in the auxiliary service module are checked sequentially as shown in Figure 7-1. Handling auxiliary services related queries separately in a separate module simplifies the classification problem. For instance, isolating the messaging services related questions prevents them from being distributed across the five main services clusters. This simplifies the overall classification problem as well as the development process of sub-dialog models.

Figure 7-1: Control Strategy of the Classifier



7.4 Sub-dialog Models

Experience has shown that the user, in the initial input, often describes the service related issue sufficiently but the underlining service type, which is often required to generate the most appropriate answer. To overcome this problem, sub-dialogs are required to collect clarifying information from the user. This section discusses how such sub-dialogs may be generated based on the classification rules and the field vectors introduced in previous sections.

Note that the sub-dialogs are activated if and only if the user initial input does not satisfy any of the classification.

Sub-dialog 1: $w_i \in \boldsymbol{w}_{Internet1}$

Sub-dialog 1-1: $w_i \in [\text{sync}, \text{ppp}, \text{splitter}, \text{router}, \text{relocate}, \text{relocation}]$ (***Note***: Most probably on a subscribed service; it can either be ADSL or Fiber)

Cal: Are you referring to an ADSL (copper line) or fiber internet connection?

User: (ADSL/copper)/Fiber → ADSL/Fiber

Wireless/mobile → Mobile BB

Default → ADSL

Sub-dialog 1-2: $w_i \in [\text{modem}]$ (***Note***: Whether it's on a subscribed service or otherwise is immaterial; Most probably ADSL. It can remotely be Wireless BB or Fiber)

Cal: Are you referring to an ADSL modem (wired copper line) or USB stick used for Wireless Internet?

User: ADSL/wired/copper/land → ADSL

Wireless/dongle/USB/stick → Mobile BB

Fiber/fibre/ONT → Fiber

Default → ADSL

Sub-dialog1-3:

$w_i \in$

[internet, broadband, www, authenticate, auth, bandwidth, bitrate, bit rate, datarate, data rate, bps, speed, WiFi, Wi − Fi] (***Note***: It can be on a subscribed or not both; Most probably ADSL. It can be any BB service)

Cal: Kindly let me know the type of internet service you are referring to: ADSL (wired: Copper line), Fiber (wired: Fiber link), or Wireless Internet

User: fiber → fiber

Wireless/mobile → Mobile BB

ADSL/copper → ADSL

Default → ADSL

Sub-dialog 2: $w_i \in \boldsymbol{w}_{Telephony}$

Cal: Are you referring to standard mobile telephony, landline telephony, or VoIP?

User: Mobile/Landline/VoIP → Mobile Telephony/Wired Telephony/VoIP

Sub-dialog 3: $w_i \notin \boldsymbol{w}_{ServiceType}$ where $\boldsymbol{w}_{ServiceType}$ is the vector of all the service-type-related salient terms.

Cal: Please specify the service type you are referring to: Wireless Internet, Mobile Telephony, Landline Telephony, ADSL Broadband, Fiber Broadband, Internet Telephony (VoIP)

User: Wireless Internet/Mobile Telephony/Landline Telephony/ADSL Broadband/Fiber Broadband/Internet Telephony (VoIP) → Mobile BB/Mobile Telephony/Wired Telephony/ADSL BB/Fiber BB/VoIP

Internet → Sub-dialog 1

Telephony → Sub-dialog 2

Sub-dialog 4: VoIP possibility $w_i \in \boldsymbol{w}_{Internet} \wedge \boldsymbol{w}_{Telephony}$

Cal: Are you referring to Internet Telephony (VoIP) service?

User: Yes/MoIP → VoIP/MoIP

W1 w2 w3 …. → Return to the Main Services Classifier