

# Table des matières

|          |  |          |
|----------|--|----------|
| <b>1</b> | <b>Conception du système</b>   | <b>2</b> |
| 1.1      | Introduction . . . . .   | 2        |
| 1.2      | Architecture du système . . . . .                                    | 2        |
| 1.2.1    | Couche utilisateur . . . . .   | 2        |
| 1.2.2    | Couche système . . . . .   | 2        |
| 1.3      | Module de reconnaissance automatique de la parole . . . . .          | 2        |
| 1.3.1    | Architecture du module . . . . .                                     | 2        |
| 1.3.2    | Modèle acoustique . . . . .  | 2        |
| 1.3.3    | Modèle de la langue . . . . .  | 3        |
| 1.4      | Module de compréhension automatique du langage naturel . . . . .     | 3        |
| 1.4.1    | Architecture du module . . . . .                                     | 3        |
| 1.4.2    | L'analyse sémantique . . . . .                                       | 3        |
| 1.4.3    | Analyse sémantique basée sur les grammaires de dépendances . . . . . | 3        |
| 1.4.4    | Analyse sémantique avec apprentissage automatique . . . . .          | 3        |
| 1.5      | Module de gestion du dialogue . . . . .                              | 3        |
| 1.5.1    | Architecture du module . . . . .                                     | 3        |
| 1.5.2    | Les ontologies du système . . . . .                                  | 3        |
| 1.5.3    | Les simulateurs d'utilisateurs . . . . .                             | 4        |
| 1.5.4    | Modèles d'apprentissage . . . . .                                    | 4        |
| 1.6      | Module de génération du langage naturel . . . . .                    | 4        |
| 1.7      | Conclusion . . . . .   | 4        |

To correct

To explain more deeply

# Chapitre 1

## Conception du système

### 1.1 Introduction

Parler de la structure générale du chapitre

### 1.2 Architecture du système

Essayer d'être aussi précis que possible tout en laissant assez de détails à être expliqués dans les sections suivantes

#### 1.2.1 Couche utilisateur

Ce que l'utilisateur peut voir comme entrée/sortie et les interfaces qui lui sont accessibles

#### 1.2.2 Couche système

Ce que l'utilisateur ne voit pas et fait donc partie du système interne

### 1.3 Module de reconnaissance automatique de la parole

#### 1.3.1 Architecture du module

#### 1.3.2 Modèle acoustique

Données d'apprentissage

Fichiers audios avec leurs transcriptions en texte

## **Type du modèle**

Réseau de neurones profond avec apprentissage de bout en bout

### **1.3.3 Modèle de la langue**

#### **Données d'apprentissage**

Parler de la procédure de récolte des données (provenances, qualité, quantité ...)

## **Type du modèle**

N-grams

## **1.4 Module de compréhension automatique du langage naturel**

### **1.4.1 Architecture du module**

### **1.4.2 L'analyse sémantique**

### **1.4.3 Analyse sémantique basée sur les grammaires de dépendances**

### **1.4.4 Analyse sémantique avec apprentissage automatique**

#### **1.4.4.1 Les données d'apprentissage**

Construction et enrichissement du corpus

#### **1.4.4.2 Modèle utilisé**

## **1.5 Module de gestion du dialogue**

### **1.5.1 Architecture du module**

### **1.5.2 Les ontologies du système**

Here talk about the graph encoder

Explorateur de fichiers

### **1.5.3 Les simulateurs d'utilisateurs**

Explorateur de fichiers

### **1.5.4 Modèles d'apprentissage**

L'agent coordinateur

Les sous agents

## **1.6 Module de génération du langage naturel**

Reste à décider

## **1.7 Conclusion**

## Table des figures

# Bibliographie

- [1] D. A. Norman, *The design of everyday things*. New York : Basic Books, 2002.
- [2] R. Knotte, A. Janson, L. Eigenbrod, and M. Söllner, “The what and how of smart personal assistants : Principles and application domains for is research,” in *Multikonferenz Wirtschaftsinformatik (MKWI)*, 2018.
- [3] H. Gellersen, *Handheld and Ubiquitous Computing : First International Symposium, HUC'99, Karlsruhe, Germany, September 27-29, 1999, Proceedings (Lecture Notes in Computer Science)*. Springer, 1999.
- [4] S. J. Russell and P. Norvig, *Artificial Intelligence : A Modern Approach*. Pearson Education, 2 ed., 2003.
- [5] T. Dingler, “Cognition-aware systems as mobile personal assistants,” in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing Adjunct - UbiComp '16*, ACM Press, 2016.
- [6] E. Luger and A. Sellen, “”like having a really bad PA”,” in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems - CHI '16*, ACM Press, 2016.
- [7] R. Trappl, ed., *Your Virtual Butler*. Springer Berlin Heidelberg, 2013.
- [8] A. Purington, J. G. Taft, S. Sannon, N. N. Bazarova, and S. H. Taylor, “”alexa is my new BFF”,” in *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems - CHI EA '17*, ACM Press, 2017.
- [9] B. R. Cowan, N. Pantidi, D. Coyle, K. Morrissey, P. Clarke, S. Al-Shehri, D. Earley, and N. Bandeira, “”what can i help you with?”,” in *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services - MobileHCI '17*, ACM Press, 2017.
- [10] J. Imtiaz, N. Koch, H. Flatt, J. Jasperneite, M. Voit, and F. van de Camp, “A flexible context-aware assistance system for industrial applications using camera based localization,” in *Proceedings of the 2014 IEEE Emerging Technology and Factory Automation (ETFA)*, IEEE, sep 2014.
- [11] A. Janson and M. T. de Gafenco, “Engaging the appropriation of technology-mediated learning services - a theory-driven design approach,” in *ECIS*, 2015.
- [12] “Apple shares examples of siri’s third-party app integration on ios 10.” <https://www.idownloadblog.com/2016/09/01/apple-siri-ios-10-app-integration/>. (Accessed on 10/29/2018).
- [13] “macos sierra review : Hey siri, where did my files go? - six colors.” <https://sixcolors.com/post/2016/09/sierra-review/>, 2016. (Accessed on 10/29/2018).
- [14] T. M Mitchell, “The discipline of machine learning,” 01 2006.

- [15] S. Kotsiantis, I. Zaharakis, and P. Pintelas, “Machine learning : A review of classification and combining techniques,” *Artificial Intelligence Review*, vol. 26, pp. 159–190, 11 2006.
- [16] H. Barlow, “Unsupervised learning,” *Neural Computation*, vol. 1, no. 3, pp. 295–311, 1989.
- [17] B. Sonali, “Research paper on basic of artificial neural network,” p. 1, 2014.
- [18] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the inception architecture for computer vision,” *CoRR*, vol. abs/1512.00567, 2015.
- [19] A. Graves, A. Mohamed, and G. Hinton, “Speech recognition with deep recurrent neural networks,” in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 6645–6649, May 2013.
- [20] D. Yu and L. Deng, *Automatic Speech Recognition*. Springer London, 2015.
- [21] M. Velay and F. Daniel, “Seq2seq and multi-task learning for joint intent and content extraction for domain specific interpreters,” *CoRR*, vol. abs/1808.00423, pp. 3–4, 2018.
- [22] P. Liu, X. Qiu, and X. Huang, “Recurrent neural network for text classification with multi-task learning,” *CoRR*, vol. abs/1605.05101, 2016.
- [23] F. Rosenblatt, *Perceptron simulation experiments (Project Para)*. 1959.
- [24] F. Murtagh, “Multilayer perceptrons for classification and regression,” *Neurocomputing*, vol. 2, pp. 183–197, jul 1991.
- [25] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, pp. 436–44, 05 2015.
- [26] Z. C. Lipton, “A critical review of recurrent neural networks for sequence learning,” *CoRR*, vol. abs/1506.00019, 2015.
- [27] C. Olah, “Understanding lstm networks – colah’s blog.” <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>, Août 2015. (Accessed on 02/19/2019).
- [28] S. Hochreiter, “The vanishing gradient problem during learning recurrent neural nets and problem solutions,” *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.*, vol. 6, pp. 107–116, Apr. 1998.
- [29] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural Comput.*, vol. 9, pp. 1735–1780, Nov. 1997.
- [30] D. Jurafsky and J. H. Martin, *Speech and Language Processing, 2nd Edition*. Prentice Hall, 2008.
- [31] Z. Ghahramani, “Hidden markov models,” ch. An Introduction to Hidden Markov Models and Bayesian Networks, pp. 9–42, River Edge, NJ, USA : World Scientific Publishing Co., Inc., 2002.
- [32] J. G. Kemeny and J. Laurie Snell, “Markov processes in learning theory,” *Psychometrika*, vol. 22, pp. 221–230, Sep 1957.
- [33] R. Rabiner and B. H. Juang, “An introduction to hidden markov models,” *IEEE ASSP Magazine*, vol. 3, pp. 4–16, 1986.
- [34] G. D. Forney, “The viterbi algorithm,” *Proceedings of the IEEE*, vol. 61, pp. 268–278, March 1973.
- [35] J. Bloit and X. Rodet, “Short-time viterbi for online hmm decoding : Evaluation on a real-time phone recognition task,” pp. 2121 – 2124, 05 2008.
- [36] P. Milhorat, S. Schlögl, G. Chollet, B. , A. Esposito, and G. Pelosi, “Building the next generation of personal digital assistants,” 03 2014.

- [37] F. Al-Anzi and D. AbuZeina, “Literature survey of arabic speech recognition,” pp. 1–6, 03 2018.
- [38] S. Narang and M. D. Gupta, “Speech feature extraction techniques : A review,” 2015.
- [39] P. M. Chauhan and N. P. Desai, “Mel frequency cepstral coefficients (mfcc) based speaker identification in noisy environment using wiener filter,” in *2014 International Conference on Green Computing Communication and Electrical Engineering (ICGCCCE)*, pp. 1–5, March 2014.
- [40] D. D. O’Shaughnessy, “Linear predictive coding,” *IEEE Potentials*, vol. 7, pp. 29–32, 1988.
- [41] H. Rahali, Z. Hajaiej, and N. Ellouze, “Robust features for speech recognition using temporal filtering technique in the presence of impulsive noise,” *International Journal of Image, Graphics and Signal Processing*, vol. 6, pp. 17–24, 10 2014.
- [42] W. Ghai and N. Singh, “Literature review on automatic speech recognition,” *International Journal of Computer Applications*, vol. 41, pp. 42–50, March 2012. Full text available.
- [43] I. Deng, G. Hinton, and B. Kingsbury, “New types of deep neural network learning for speech recognition and related applications : An overview,” pp. 8599–8603, 10 2013.
- [44] R. Glass and M. Kyle McCandless, “Automatic acquisition of language models for speech recognition,” 10 1994.
- [45] B. Roark, M. Saraclar, and M. Collins, “Discriminative n-gram language modeling,” *Comput. Speech Lang.*, vol. 21, pp. 373–392, Apr. 2007.
- [46] B. Liu and I. Lane, “Attention-based recurrent neural network models for joint intent detection and slot filling,” *CoRR*, vol. abs/1609.01454, 2016.
- [47] C.-W. Goo, G. Gao, Y.-K. Hsu, C.-L. Huo, T.-C. Chen, K.-W. Hsu, and Y.-N. Chen, “Slot-gated modeling for joint slot filling and intent prediction,” in *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics : Human Language Technologies, Volume 2 (Short Papers)*, (New Orleans, Louisiana), pp. 753–757, Association for Computational Linguistics, June 2018.
- [48] M. Schuster and K. Paliwal, “Bidirectional recurrent neural networks,” *Trans. Sig. Proc.*, vol. 45, pp. 2673–2681, Nov. 1997.
- [49] J. Chorowski, D. Bahdanau, D. Serdyuk, K. Cho, and Y. Bengio, “Attention-based models for speech recognition,” *CoRR*, vol. abs/1506.07503, 2015.
- [50] Y. Wang, Y. Shen, and H. Jin, “A bi-model based rnn semantic frame parsing model for intent detection and slot filling,” in *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics : Human Language Technologies, Volume 2 (Short Papers)*, (New Orleans, Louisiana), pp. 309–314, Association for Computational Linguistics, June 2018.
- [51] t. . A. M. D. P. Richard Bellman” *Indiana Univ. Math. J.*, *fjournal* = “*Indiana University Mathematics Journal*”, vol. 6, pp. 679–684, 1957.
- [52] H. Chen, X. Liu, D. Yin, and J. Tang, “A survey on dialogue systems : Recent advances and new frontiers,” *SIGKDD Explor. Newsl.*, vol. 19, pp. 25–35, Nov. 2017.
- [53] D. Goddeau, H. Meng, J. Polifroni, S. Seneff, and S. Busayapongchai, “A form-based dialogue manager for spoken language applications,” vol. 2, pp. 701 – 704, 11 1996.
- [54] S. Young, M. Gašić, S. Keizer, F. Mairesse, J. Schatzmann, B. Thomson, and K. Yu, “The hidden information state model : A practical framework for pomdp-based spoken dialogue management,” *Comput. Speech Lang.*, vol. 24, pp. 150–174, Apr. 2010.



- [55] K. J. Åström, “Optimal control of Markov Processes with incomplete state information,” *Journal of Mathematical Analysis and Applications*, vol. 10, pp. 174–205, January 1965.
- [56] M. Henderson, B. Thomson, and S. J. Young, “Deep neural network approach for the dialog state tracking challenge,” in *SIGDIAL Conference*, pp. 467–471, 2013.
- [57] C. Lee, S. Jung, K. Kim, D. Lee, and G. G. Lee, “Recent approaches to dialog management for spoken dialog systems,” *JCSE*, vol. 4, pp. 1–22, 2010.
- [58] J. Henderson, O. Lemon, and K. Georgila, “Hybrid reinforcement/supervised learning of dialogue policies from fixed data sets,” *Comput. Linguist.*, vol. 34, pp. 487–511, Dec. 2008.
- [59] T.-H. Wen, M. Gasic, N. Mrksic, L. M. Rojas-Barahona, P. hao Su, S. Ultes, D. Vandyke, and S. J. Young, “A network-based end-to-end trainable task-oriented dialogue system,” in *EACL*, pp. 438–449, 2017.
- [60] I. V. Serban, A. Sordoni, Y. Bengio, A. Courville, and J. Pineau, “Building end-to-end dialogue systems using generative hierarchical neural network models,” in *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, AAAI, pp. 3776–3783, AAAI Press, 2016.
- [61] G. Weisz, P. Budzianowski, P. hao Su, and M. Gax0161ix0107, “Sample efficient deep reinforcement learning for dialogue systems with large action spaces,” *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 26, pp. 2083–2097, 2018.
- [62] D. Bertsekas, *Dynamic Programming & Optimal Control, Vol II : Approximate Dynamic Programming*. Athena Scientific, 01 2012.
- [63] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. A. Riedmiller, A. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, “Human-level control through deep reinforcement learning,” *Nature*, vol. 518, pp. 529–533, 2015.
- [64] J. Schatzmann, B. Thomson, K. Weilhammer, H. Ye, and S. J. Young, “Agenda-based user simulation for bootstrapping a pomdp dialogue system,” in *HLT-NAACL*, pp. 149–152, 2007.
- [65] K. Georgila, J. Henderson, and O. Lemon, “Learning user simulations for information state update dialogue systems,” in *INTERSPEECH*, pp. 893–896, 2005.
- [66] H. Cuayáhuatl, S. Renals, O. Lemon, and H. Shimodaira, “Human-computer dialogue simulation using hidden markov models,” *IEEE Workshop on Automatic Speech Recognition and Understanding, 2005.*, pp. 290–295, 2005.
- [67] S. Chandramohan, M. Geist, F. Lefèvre, and O. Pietquin, “User simulation in dialogue systems using inverse reinforcement learning,” in *INTERSPEECH*, p. 10251028, 2011.
- [68] R. Evans, P. Piwek, and L. Cahill, “What is nlg?,” in *Proceedings of the Second International Conference on Natural Language Generation*, pp. 144–151, 2002.
- [69] E. Reiter and R. Dale, “Building applied natural language generation systems,” *Natural Language Engineering*, vol. 3, pp. 57–87, Mar. 1997.
- [70] C. Labbé and F. Portet, “Towards an abstractive opinion summarisation of multiple reviews in the tourism domain,” *CEUR Workshop Proceedings*, vol. 917, pp. 87–94, 01 2012.
- [71] N. Dethlefs, “Context-sensitive natural language generation : From knowledge-driven to data-driven techniques,” *Language and Linguistics Compass*, vol. 8, pp. 99–115, 03 2014.
- [72] J. Yu, E. Reiter, J. Hunter, and C. Mellish, “Choosing the content of textual summaries of large time-series data sets,” *Natural Language Engineering*, vol. 13, pp. 25–49, Mar. 2007.

- [73] A. Gatt and E. Krahmer, “Survey of the state of the art in natural language generation : Core tasks, applications and evaluation,” *J. Artif. Int. Res.*, vol. 61, pp. 65–170, Jan. 2018.
- [74] M. Theune, E. Klabbers, J. R. De Pijper, E. Krahmer, and J. Odijk, “From data to speech : A general approach,” *Natural Language Engineering*, vol. 7, pp. 47–86, Mar. 2001.
- [75] G. Angeli, C. D. Manning, and D. Jurafsky, “Parsing time : Learning to interpret time expressions,” in *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics : Human Language Technologies*, NAACL HLT ’12, (Stroudsburg, PA, USA), pp. 446–455, Association for Computational Linguistics, 2012.
- [76] M. A. K. Halliday and C. M. I. M. Matthiessen, *Introduction to Functional Grammar*. London : Hodder Arnold, 3 ed., 2004.
- [77] W. C. Mann and C. M. I. M. Matthiessen, “Nigel : A systemic grammar for text generation.,” 1983.
- [78] J. A. Bateman, “Enabling technology for multilingual natural language generation : The kpml development environment,” *Nat. Lang. Eng.*, vol. 3, pp. 15–55, Mar. 1997.
- [79] M. Elhadad and J. Robin, “An overview of surge : a reusable comprehensive syntactic realization component,” in *International Natural Language Generation Workshop*, pp. 1–4, 1996.
- [80] I. Langkilde-Geary, “Forest-based statistical sentence generation,” in *ANLP*, pp. 170–177, 2000.
- [81] A. Belz, “Automatic generation of weather forecast texts using comprehensive probabilistic generation-space models,” *Nat. Lang. Eng.*, vol. 14, pp. 431–455, Oct. 2008.
- [82] D. Espinosa, M. White, and D. Mehay, “Hypertagging : Supertagging for surface realization with ccg,” in *ACL*, pp. 183–191, 2008.
- [83] T. C. Ferreira, I. Calixto, S. Wubben, and E. Krahmer, “Linguistic realisation as machine translation : Comparing different mt models for amr-to-text generation,” in *INLG*, pp. 1–10, 2017.
- [84] T.-H. Wen, M. Gasic, N. Mrksic, P. hao Su, D. Vandyke, and S. J. Young, “Semantically conditioned lstm-based natural language generation for spoken dialogue systems,” in *EMNLP*, pp. 1711–1721, 2015.
- [85] A. Sordoni, M. Galley, M. Auli, C. Brockett, Y. Ji, M. Mitchell, J.-Y. Nie, J. Gao, and W. B. Dolan, “A neural network approach to context-sensitive generation of conversational responses,” in *HLT-NAACL*, pp. 196–205, 2015.
- [86] R. Goyal, M. Dymetman, and E. Gaussier, “Natural language generation through character-based rnns with finite-state prior knowledge,” in *COLING*, pp. 1083–1092, 2016.