

Table des matières

Table des matières	1
1 Conception du système	5
1.1 Introduction	5
1.2 Architecture du système	5
1.3 Module de reconnaissance automatique de la parole	7
1.4 Module de compréhension automatique du langage naturel	11
1.5 Module de gestion du dialogue	15
1.6 Module de génération du langage naturel	30
1.7 Conclusion	31
Table des figures	32
Bibliographie	33
To correct	
To explain more deeply	

Introduction générale

"Nous entrons dans un nouveau monde. Les technologies d'apprentissage automatique, de reconnaissance de la parole et de compréhension du langage naturel atteignent un niveau de capacités. Le résultat final est que nous aurons bientôt des assistants intelligents pour nous aider dans tous les aspects de nos vies."

— Amy Stapleton,, *Opus Research*

Plus de quarante ans se sont écoulés depuis la présentation du premier assistant virtuel commandé vocalement par la compagnie IBM. Déjà à cette époque là, ce fut présenté comme une révolution technologique. Un programme qui pouvait reconnaître 16 mots et les chiffres de 0 à 9. Quelques générations plus tard, nous nous retrouvons avec des assistants capables de reconnaître, comprendre et parler plusieurs langues. Ces assistants intelligents sont la nouvelle génération d'intelligences artificielles capables de s'intégrer dans nos vies personnelles et professionnelles.

Un domaine qui à particulièrement émergé est celui de l'aide à la manipulation d'un ordinateur personnel. Selon certains experts, l'époque où nous utilisons encore le clavier, la souris et l'écran est une étape de transition. Le futur se trouve dans l'exploitation de la parole comme moyen de communication principal avec les machines. Et ce futur est proche, la course à la réalisation d'assistants qui excellents dans plusieurs domaines a commencé il y a quelques années avec l'entrée de grandes compagnies comme Google et Apple dans le secteur. Par la suite, de très grands efforts ont été fournis dans le but d'améliorer l'expérience d'utilisation de ces assistants. Les chercheurs ainsi que les industriels se sont orientés vers cette solution très rapidement.

En tant que novice dans ce secteur qu'est la personnalisation des services électroniques, l'Algérie devra rapidement se positionner pour s'incorporer dans l'évolution de ces technologies. Nous avons donc été motivés par l'envie de nous initier à ce domaine, ainsi que de motiver d'autres compatriotes scientifiques à continuer sur cette voie. Nous pensons que les plus ambitieux des projets commencent avec des contributions à petites échelles. Notre assistant aura pour but d'améliorer l'expérience d'utilisation d'un ordinateur. Ceci en effectuant des tâches rudimentaires, efficacement et sans réel effort hormis l'énonciation de la requête. En utilisant des techniques d'intelligence artificielle d'actualité comme la reconnaissance automatique de la parole, la compréhension du langage naturel et l'apprentissage par renforcement, ce projet se veut assez ambitieux et vise à faire gagner du temps à tout utilisateur d'un ordinateur de bureau ou portable.

Dans cette optique, nous présenterons notre étude de la littérature concernant les assistants personnels intelligents. Nous passerons en revue les aspects théoriques qui sont utilisés dans des solutions considérées comme état de l'art du domaine. Nous nous baserons sur ces techniques pour la conception des modules de notre système tout en les adaptant à nos besoins.

Ce mémoire se constitue de quatre chapitres. Le premier chapitre sera consacré à la présentation des assistants virtuels intelligents. Le deuxième chapitre se focalisera sur l'étude des travaux existants traitant de notre sujet. Le troisième chapitre entamera l'étude conceptuelle. Le quatrième et dernier chapitre présentera les évaluations du système réalisées, suivies d'une discussion des résultats obtenus. Enfin nous présenterons notre interface pour l'application et nous conclurons ce travail avec une conclusion générale.

Table des figures

1.1	Architecture générale du système Bethano	6
1.2	Architecture du module de reconnaissance de la parole (ASR)	8
1.3	Architecture du modèle DeepSpeech [5]	9
1.4	Processus de génération du corpus pour le modèle de langue	10
1.5	Architecture du module de compréhension automatique du langage naturel (NLU)	12
1.6	Schéma de transformation de trame sémantique en graphe	16
1.7	Schéma de l'architecture multi-agents pour la gestion du dialogue	17
1.8	Schéma représentant l'apprentissage des agents parents avec les simulateurs des agents feuilles	17
1.9	Schéma global du gestionnaire de dialogue	18
1.10	Graphe de l'ontologie de dialogue	19
1.11	Schéma de transformation d'une action en graphe	20
1.12	Graphe de l'ontologie de l'exploration de fichiers	21
1.13	Schéma de transformation d'une action de demande de création de fichier en graphe	22
1.14	Diagramme de décision de l'action à prendre	24
1.15	Schéma représentant un encodeur de graphe	27
1.16	Schéma représentant un encodeur séquentiel de graphe	27
1.17	Schéma de l'apprentissage d'un encodeur séquentiel de graphe	28
1.18	Schéma du réseau DQN	29
1.19	Schéma du réseau DQN relié avec l'encodeur directement	30
1.20	Schéma de fonctionnement du générateur de texte	31

Bibliographie

- [1] wfuzz. <https://github.com/xmendez/wfuzz/blob/master/wordlist/general/common.txt>. (Consulté le 14/06/2019).
- [2] G A. Rummary and Mahesan Niranjan. On-line q-learning using connectionist systems. *Technical Report CUED/F-INFENG/TR 166*, 11 1994.
- [3] Tom Bocklisch, Joey Faulkner, Nick Pawlowski, and Alan Nichol. Rasa : Open source language understanding and dialogue management. *CoRR*, abs/1712.05181, 2017.
- [4] Chih-Wen Goo, Guang Gao, Yun-Kai Hsu, Chih-Li Huo, Tsung-Chieh Chen, Keng-Wei Hsu, and Yun-Nung 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, June 2018. Association for Computational Linguistics.
- [5] Awni Y. Hannun, Carl Case, Jared Casper, Bryan Catanzaro, Greg Diamos, Erich Elsen, Ryan Prenger, Sanjeev Satheesh, Shubho Sengupta, Adam Coates, and Andrew Y. Ng. Deep speech : Scaling up end-to-end speech recognition. *CoRR*, abs/1412.5567, 2014.
- [6] Thomas N. Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. In *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*, 2017.
- [7] Yujia Li, Daniel Tarlow, Marc Brockschmidt, and Richard S. Zemel. Gated graph sequence neural networks. *CoRR*, abs/1511.05493, 2016.
- [8] Bing Liu and Ian Lane. Attention-based recurrent neural network models for joint intent detection and slot filling. *CoRR*, abs/1609.01454, 2016.
- [9] Mitchell Marcus, Grace Kim, Mary Ann Marcinkiewicz, Robert MacIntyre, Ann Bies, Mark Ferguson, Karen Katz, and Britta Schasberger. The penn treebank : Annotating predicate argument structure. In *Proceedings of the Workshop on Human Language Technology*, pages 114–119, Stroudsburg, PA, USA, 1994. Association for Computational Linguistics.
- [10] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin A. Riedmiller, Andreas Fidjeland, Georg Ostrovski, Stig Petersen, Charles Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dharmashan Kumaran, Daan Wierstra, Shane Legg, and Demis Hassabis. Human-level control through deep reinforcement learning. *Nature*, 518, 2015.

- [11] Jost Schatzmann, Blaise Thomson, Karl Weilhammer, Hui Ye, and Steve J. Young. Agenda-based user simulation for bootstrapping a pomdp dialogue system. In *HLT-NAACL*, 2007.
- [12] Svetlana Stoyanchev and Michael Johnston. Knowledge-graph driven information state approach to dialog. In *AAAI Workshops*, 2018.
- [13] Michael Wessel, Girish Acharya, James Carpenter, and Min Yin. *OntoVPAAAn Ontology-Based Dialogue Management System for Virtual Personal Assistants : 8th International Workshop on Spoken Dialog Systems*. 01 2019.