

Émergence de phénomènes linguistiques par apprentissage par renforcement à plusieurs agents

Proposition de thèse CIFRE avec
Facebook AI Research (FAIR) &
Institut Jean-Nicod, École Normale Supérieure

Théo Matussi re

Novembre 2018

Mots-cl s :  volution du langage , apprentissage par renforcement, syntaxe, phonologie, protocole de communication, compositionnalit 

Depuis quelques ann es l’av nement de m thodes permettant d’associer des objets math matiques bien connus (les vecteurs de \mathbb{R}^d)   des mots, des phrases ou m me des images permet des avanc es scientifiques significatives dans de nombreux domaines. Ces nouvelles m thodes, connues du public sous le nom de *Deep Learning*, ou apprentissage profond, ont permis des am liorations majeures dans la plupart des t ches classiques du Traitement Automatique des Langues (TAL, ou *Natural Language Processing*, NLP en anglais) comme la traduction, la r solution de r f rences, la reconnaissance d’entit s nomm es ou la mod lisation du dialogue.

Ces m thodes sont motiv es par une approche statistique du langage, en partie en raison de l’hypoth se distributionnelle qui inf re le sens d’un mot en fonction des mots avec lesquels il apparait r guli rement. L’estimation de ces cooccurrences de mots entre eux est le ressort le plus courant pour d velopper ces m thodes, elle est permise par l’analyse de grands corpus de textes. Bien qu’elles obtiennent des succ s ind niables, il existe un regain d’int r t pour des mod lisations du langage qui prennent en compte sa nature fondamentalement interactive. Cette initiative a obtenu des r sultats int ressants par le pass , notamment au regard des questions sur l’apparition et les  volutions du langage et de ses ph nom nes associ s. Parmi ces r sultats, on trouve notamment Kirby et al. (2008), o  les auteurs montrent la viabilit  de leur m thode it rative qui aboutit   des ph nom nes linguistiques  mergents. D’autres travaux comme Steels (1997)  taient la pertinence de cette approche. Plus r cemment les avanc es de l’apprentissage par renforcement (*Reinforcement Learning*, RL), qui ont permis la victoire de l’algorithme AlphaGo au jeu de Go en 2016 (Silver et al., 2017), ont  t  mobilis es pour simuler l’ mergence de protocoles de communication entre agents au sein de jeux collaboratifs. Les r sultats de ces exp riences ne sont pour l’instant pas convaincants puisque les

systèmes de communication qu'ils obtiennent ne sont pas proprement complexes, et assez éloignés de nos langages humains.

Néanmoins, cette direction est prometteuse puisque la puissance de modélisation que permettent les méthodes de RL autorise la complexification des environnements et des agents. Enrichir les environnements, complexifier les tâches et bénéficier d'agents plus « intelligents » pourrait permettre l'observation de langages plus complexes. Notre démarche s'inscrit dans la continuité de ces travaux récents que nous détaillons dans la suite. Nous exposons notamment l'originalité de notre approche et les expériences et réflexions que nous comptons mener, puis l'organisation logistique du travail à venir.

1 Objectifs & intentions

La question de l'origine de nos langages humains est une question centrale de la philosophie et des sciences en général. L'hypothèse d'une évolution graduelle de systèmes de communication semble aujourd'hui la plus probable, en opposition aux théories Chomskiennes d'une apparition soudaine du langage dû à une mutation génétique et la naissance soudaine de l'« organe du langage » \diamond^1 . L'investigation sur la nature de cette évolution, sur les trajectoires qu'ont pu emprunter les langues humaines pour arriver au niveau de complexité que nous leur connaissons est l'objet de différentes disciplines, parmi lesquelles les sciences cognitives, la psychologie et la philosophie. La richesse du travail effectué par les chercheurs de ces disciplines est mis en avant par une initiative portée par [Roberts \(2018\)](#) qui compile les avancées faites autour de cette question sous la forme d'un graphe de causalité entre cause supposées et phénomènes linguistiques résultants.

Nous proposons une lecture originale de cette question. La recherche d'invariants universels aux langages humains s'est souvent révélée infructueuse ([Evans and Levinson, 2009](#)), ainsi nous prendrons comme dénominateur commun minimal aux langues humaines l'existence d'une syntaxe, autrement dit d'une forme de compositionnalité, i.e. de l'existence de mots *fonctions*. Notre postulat est que cette compositionnalité est une réponse optimale au problème de la communication *sous contraintes*. Si [Roberts \(2018\)](#) fournit de nombreuses pistes quant à l'établissement desdites contraintes, nous avons isolés cinq contraintes fondamentales qui feront l'objet de nos expériences :

Contrainte	Description
Expressivité	Capacité à décrire son environnement dans toutes ses potentialités.
Concision	Communiquer coûte de l'énergie.
Mémoire	Le vocabulaire doit être fini.
Simplicité	Le langage doit avoir une complexité de Kolmogorov minimale.
Robustesse au bruit	Le canal de communication n'est pas parfaite.

L'idée originale étant qu'expressivité et concision sont des contraintes poussant dans des sens opposés et nécessitant la naissance d'un mécanisme comme celui de la com-

1. need back up

positionnalité. Les discussions successives avec les encadrants proposés pour cette thèse ont ensuite fait naître la nécessité d’une perspective moins simpliste sur le nombre et la complexité des contraintes qui s’appliquent de fait à nos environnements.

Nos premières expériences construirons de façon incrémentale sur l’état de l’art afin d’y ajouter ces contraintes. La première contrainte que nous testerons est celle de l’ajout de bruit dans les canaux de communication, car elle est la plus facile d’accès. Elle est de plus motivée par des recherches antérieures sur les matrices de confusion entre consonnes pour des locuteurs d’une langue donnée [Miller and Nicely \(1955\)](#). L’idée étant d’imposer des perturbations similaires à celles que l’on peut trouver dans nos langues humaines (matrices quasi-symétriques et quasi-diagonales) lors de l’envoi des symboles disponibles aux agents. C’est à dire qu’une matrice aléatoire quasi-symétrique et quasi-diagonale B décidera $P[s_j \text{ reçu} \mid s_i \text{ envoyé}] = B_{i,j}$, influant par là l’utilisation ou non de certains symboles et favorisant la redondance selon l’importance du bruit introduit. L’hypothèse sous-jacente étant que la redondance est une propriété commune à nos langues humaines, et qu’elle participe de la complexité, et pourquoi pas ultimement de la compositionnalité. L’ajout de bruit dans le canal de communication n’est pas une idée originale, et nous la trouvons en fait dans un des premiers articles mêlant RL et émergence de communication organisée avec [Jorge et al. \(2016\)](#), bien qu’ils introduisent le bruit comme une astuce de régularisation pour la convergence de leur algorithme (*Curriculum learning*), là où elle nous servira à proprement parler à influencer le langage.

La possibilité de donner aux agents une marge de manœuvre sur B est aussi envisageable, quoique les contours de cette expérience additionnelle devront être ajustés à la lumière des résultats de la première. Le protocole exact de cette première expérience nécessite une présentation succincte des méthodes actuellement utilisées dans l’état de l’art que nous fournissons dans la section suivante.

Les expériences pour tester l’apport des contraintes d’expressivité (richesse de l’environnement, contrôlable) et de concision (fonction de coût dépendant de la longueur des productions) sont faciles à intégrer à n’importe quel protocole d’expérience. La mémoire est également un facteur sur lequel nous pouvons facilement jouer puisqu’il s’agit de modifier nos modèles pour en intégrer plus ou moins. Ainsi ces paramètres seront testés sur chacune des tâches que nous développerons. Il sera en particulier intéressant d’étudier les interactions entre ces contraintes, selon le degré d’intensité que nous leur imposerons.

Seule la contrainte de simplicité reste non tractable...En particulier la contrainte de simplicité que nous définissons avec la complexité de Kolmogorov, définie comme la longueur minimale du programme permettant de produire la langue — dans le cas d’une langue naturelle cela équivaudrait à la taille nécessaire à la description de sa grammaire — pose de sérieux problèmes. Il est en général très dur de calculer la complexité de Kolmogorov, et l’on peut tout au plus seulement en donner des bornes inférieures dans la plupart des cas.

Cette contrainte est liée à la nécessité de transmettre le protocole de communication à une nouvelle génération d’agents, mais elle peut aussi être vue comme un invariant

cognitif, comme suggéré par [Chater and Vitányi \(2003\)](#).

La question de la *nécessité* de la compositionnalité dans le contexte de ces contraintes est sans doute la question centrale au travail de cette thèse. Elle nécessite l'établissement d'une définition rigoureuse et d'une compréhension profonde de ses expressions. Ses manifestations peuvent être trompeuses ; prenons l'exemple de la séquentialité comme utilisation implicite de prédicats logiques. Par là nous entendons, dans le jeu de construction de Wittgenstein, la compréhension de l'énonciation consécutive de différents objets comme la nécessité de leur conjonction. C'est-à-dire que la phrase « marteau, planche, scie » soit entendu comme le message signifiant la volonté de tous ces objets « marteau ET planche ET scie », et non de l'un d'entre eux comme ce serait le cas pour l'utilisation de la séquentialité pour le choix « marteau OU planche OU scie ».

Détecter si une phrase fait partie d'une grammaire est un problème complexe, qui nécessite des algorithmes sophistiqués et une certaine quantité de données pour fonctionner correctement. Détecter la structure même d'une grammaire est donc un défi de taille. Plus simplement on peut citer [Kirby et al. \(2008\)](#), qui ont introduit une mesure naïve de complexité pour tenter de deviner la naissance ou non de phénomènes de nature compositionnels en détectant les sauts de continuité pour la distance Levenstein et dans l'espace sémantique. Cette mesure est trop simpliste puisque l'existence des périphrases la rend obsolète (forte distance de Levenstein, faible distance sémantique), mais elle figure sans doute un point de départ intéressant.

Cette introspection sur la nature de la compositionnalité est au programme du travail de thèse. La perspective de réussir à faire émerger ce phénomène n'a rien d'évident, et il existe quelques résultats négatifs quand aux conditions de réalisations de ce phénomène, notamment [Lake and Baroni \(2017\)](#) dans le cas d'une modélisation utilisant des réseaux de neurones classiques mais habituellement parmi les plus performants pour la modélisation du langage. Nous espérons pouvoir donner des suites à ces travaux, et dans le cas où nous ne parviendrions pas à donner naissance à de véritables systèmes compositionnels nous pourrions néanmoins simuler l'apparition de phénomènes linguistiques d'intérêt comme la phonologie ou la morphologie, qui sont des phénomènes de vocabulaire. \diamond^2

2 Formalisation du protocole & état de l'art

Nous sommes intéressés par la façon dont des populations d'agents utilisent un ou plusieurs canaux de communication afin de résoudre une tâche partagée (collaboration) ou individualisée (compétition). Ce canal de communication peut être discrétisé, c'est le cas de la communication par symboles, ou bien continu dans le cas de la communication par vecteurs de \mathbb{R}^d . Si la communication par vecteur peut fonctionner, nous nous intéresserons surtout au cas de la communication par symbole, puisque nos langages fonctionnent en discrétisant des éléments sonores et en les réassemblant pour former les mots et phrases du langage. Si la question de savoir comment nous parvenons justement

2. not too clear, but do you get what I'm aiming at ?

à discrétiser les sons est une question active en sciences de la cognition, nous partirons simplement de l’observation de la nature foncièrement discrète du langage.

Le comportement des agents peut être modélisé par des algorithmes classiques, comme cela a pu être fait dans le cas de Kirby et al. (2008); Steels (1997); Beuls and Steels (2013) ou bien par des méthodes d’apprentissage. Nous nous intéresserons à ce dernier cas, en particulier aux agents modélisés par l’apprentissage par renforcement (RL).

Apprentissage par renforcement

Fondamentalement le RL se base sur la théorie des jeux de Markov : (S, A, P, R) avec : S un ensemble d’états fini, A un ensemble d’actions fini, $P[s' | s, a]$ la probabilité d’être en s' en ayant actionné a depuis l’état s , et $R(s', s)$ la fonction de récompense. L’idée étant de trouver une politique $\pi : S \rightarrow A$ optimale qui permette de maximiser, avec $0 < \gamma < 1$ une préférence pour le futur, la fonction suivante, appelée fonction de valeur :

$$V_{\pi}(s) = \mathbb{E} \left[\sum_{t \geq 0} \gamma^t R(s_{t+1}, s_t) \mid s_0 = s \right]$$

Il existe plusieurs approches pour déterminer π^* la politique optimale, comme l’itération sur la politique ou sur la fonction de valeur, ou encore les estimations Monte Carlo. Cependant ces techniques ne sont envisageables que dans le cas d’espaces d’états-actions réduits. Les méthodes d’apprentissage par descente de gradient permettent de considérer des ensembles plus vastes, faisant progresser la discipline en incorporant des jeux toujours plus complexes. C’est dans cette catégorie d’algorithmes que se trouvent d’une part le célèbre AlphaGo (Silver et al., 2017) et d’autre part les algorithmes qui nous intéressent dans le cas des jeux référentiels, que nous présentons maintenant.

État de l’art

Le premier type de tâche qui est investigué relève des jeux de signaux ou de référence (*signaling games*, Lewis (1969)), à la façon du jeu de langage des maçons de Wittgenstein (1953). Dans ce contexte on cherche à plonger des agents dans un environnement simple comportant un nombre fini d’objet, ces derniers devant réussir *in fine* à se comprendre lorsqu’ils désignent l’un ou l’autre des objets de ce monde. Si ces jeux de référence ont une longue histoire et de nombreuses modélisations, un des premiers articles à employer les méthodes de RL est Lazaridou et al. (2016). Les résultats reportés par les auteurs sont encourageants, avec notamment la confirmation que des agents modélisés par des réseaux de neurones rudimentaires sont capables de se coordonner.

Les méthodes de RL sont difficiles à faire converger, notamment dû à une forte variance des estimateurs, et nécessitent d’impressionnantes quantités de données. Afin de pallier à ces déficiences de récents articles Mordatch and Abbeel (2017); Havrylov and Titov (2017) ont utilisés une nouvelle méthode, l’astuce du Gumbel-Softmax, afin d’échapper à la forte variance des algorithmes dérivés de REINFORCE (Williams, 1992).

Si certains articles ([Mordatch and Abbeel, 2017](#); [Havrylov and Titov, 2017](#)) donnent des éléments encourageants quant à la capacité de faire converger les agents en leur faisant utiliser plusieurs symboles à la fois, les conclusions exactes sur la nature de ces habilités restent floues.

Protocoles originaux

Une des pistes importantes de la thèse sera la conception de tâches complexes nécessitant l'utilisation de langages complexes pour pouvoir être surmontées. Nous sommes naturellement portés vers des jeux qui favoriseront l'émergence de mots fonctions. Le cadre exact de conception de ces jeux se situe plutôt dans le cadre de la théorie des jeux, ce qui nécessitera un travail bibliographique.

Une opportunité intéressante serait de faire jouer de façon conjointe des agents à un jeu portés sur la conjonction, et à un autre jeu porté sur le choix. On espérerait ainsi forcer l'apparition de quantificateurs logiques ET, OU pour pouvoir désambiguer l'utilisation de la séquentialité. Le cadre exact de ce jeu reste à définir.

Un jeu intéressant est celui du labyrinthe, avec un observateur omniscient devant communiquer à un agent récepteur le chemin donnant vers la sortie. Le mode de communication optimal est évidemment séquentiel ; « gauche-gauche-droite-gauche-... », mais on peut le compresser en introduisant les nombres ; « 2-gauche-droite-2-gauche ». Peut-être qu'en contraignant fortement le coût énergétique de communication nous pourrions forcer ce phénomène.

3 Institutions & considérations pratiques

Le laboratoire d'intelligence artificielle de Facebook (le *Facebook AI Research*, FAIR) a pour but de développer des systèmes à des niveaux d'intelligence humaine en faisant avancer l'état de l'art autour de l'Intelligence Artificielle. La recherche qui y est produite couvre tous le spectre des activités liées à l'IA : théorie, optimisation, algorithmes, applications et infrastructures. L'intégralité du travail du laboratoire est publié dans les revues académiques de premier plan, le code développé rendu publique et les résultats reproductibles. Le FAIR s'est en particulier démarqué récemment en étant à l'origine des articles majeurs traitant de l'émergence de protocoles de communication dans des sociétés d'agents intelligents, avec notamment ([Lewis et al., 2017](#); [Lazaridou et al., 2016](#); [Sukhbaatar et al., 2016](#); [Evtimova et al., 2017](#); [Lee et al., 2017](#)).

L'Institut Jean Nicod (IJN) est sous la tutelle universitaire de l'École normale supérieure (ENS) et de l'École des hautes études en sciences sociales (EHESS). Si la discipline de base est la philosophie analytique, l'unité héberge également des recherches de pointe en linguistique et en sciences sociales, le thème unificateur étant l'esprit humain et la nature des représentations (linguistiques, mentales, et sociales). À l'ENS, l'IJN est rattaché à la fois au département de philosophie et au département d'études cognitives (DEC). L'institut héberge notamment des chercheurs dont le travail est centré autour de questions fondamentales sur la nature et l'utilisation du langage [Egre and Ó Madagáin \(2018\)](#) et

des traits qu’il peut partager avec d’autres espèces [Schlenker et al. \(2016\)](#). Ces sujets et cette prise de distance par rapport à une conception autocentrée du langage est une richesse qui saura nourrir les réflexions et les hypothèses que nous souhaitons tester sur cette thèse.

Ces deux institutions, pour tout ce qui les séparent, sont à même d’apporter tous les éléments techniques et théoriques nécessaires au succès de cette entreprise qui se trouve à l’intersection d’un certain nombre de disciplines. Le FAIR offre l’opportunité de côtoyer les chercheurs et ingénieurs du meilleur niveau mondial en RL et en NLP, un atout majeur dans la production de code efficace et du meilleur niveau. Toujours parce que notre travail aura cette importante composante computationnelle, l’apport du FAIR sera essentiel grâce aux capacités de calculs sans rivales qu’il offre à ses chercheurs. L’IJN de son côté pourra offrir le cadre théorique qui permettra l’inspection des nouvelles formes de communication que nous espérons faire émerger. L’absence de cadre formel dans lequel inspecter ou détecter des formes de compositionnalité nécessite la conception d’une théorie adéquate. L’expérience des chercheurs de l’Institut permettra en outre de mettre nos travaux en relation avec ce qui est fait en linguistique évolutionnelle et en sciences cognitives.

Nous proposons le plan suivant, tableau [1](#), pour l’évolution de la thèse. Produire deux articles sur une thèse est sans doute conservateur, mais articles ambitieux. Le reste correspond à ce qui a été décrit en section 1.

Temps	Travaux théoriques	Travail expérimental
1^{ère} year	bibliographie classification des formes de compositionnalité	expériences de canaux bruités
2nd year	classification des formes de compositionnalité inspection & classification des langages publication article sur canaux bruités	expérimentations jeux complexes
3rd year	écriture thèse publication article jeux complexes	communication multi-canaux

TABLE 1 – Roadmap

Références

- Katrien Beuls and Luc Steels. 2013. [Agent-Based Models of Strategies for the Emergence and Evolution of Grammatical Agreement](#). *PLoS ONE*, 8(3) :e58960.
- Nick Chater and Paul Vitányi. 2003. [Simplicity: a unifying principle in cognitive science?](#) *Trends in Cognitive Sciences*, 7(1) :19–22.
- Paul Egre and Cathal Ó Madagáin. 2018. [Concept Utility](#).
- Nicholas Evans and Stephen C. Levinson. 2009. [The myth of language universals: Language diversity and its importance for cognitive science](#). *Behavioral and Brain Sciences*, 32(05) :429.
- Katrina Evtimova, Andrew Drozdov, Douwe Kiela, and Kyunghyun Cho. 2017. [Emergent Communication in a Multi-Modal, Multi-Step Referential Game](#). *arXiv :1705.10369 [cs, math]*. ArXiv : 1705.10369.
- Serhii Havrylov and Ivan Titov. 2017. Emergence of Language with Multi-agent Games : Learning to Communicate with Sequences of Symbols. *arXiv preprint arXiv :1705.11192*.
- Emilio Jorge, Mikael Kågebäck, Fredrik D. Johansson, and Emil Gustavsson. 2016. [Learning to Play Guess Who? and Inventing a Grounded Language as a Consequence](#). *arXiv :1611.03218 [cs]*. ArXiv : 1611.03218.
- S. Kirby, H. Cornish, and K. Smith. 2008. [Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language](#). *Proceedings of the National Academy of Sciences*, 105(31) :10681–10686.
- Brenden M. Lake and Marco Baroni. 2017. [Generalization without systematicity: On the compositional skills of sequence-to-sequence recurrent networks](#). *arXiv :1711.00350 [cs]*. ArXiv : 1711.00350.
- Angeliki Lazaridou, Alexander Peysakhovich, and Marco Baroni. 2016. Multi-agent co-operation and the emergence of (natural) language. *arXiv preprint arXiv :1612.07182*.
- Jason Lee, Kyunghyun Cho, Jason Weston, and Douwe Kiela. 2017. [Emergent Translation in Multi-Agent Communication](#). *arXiv :1710.06922 [cs]*. ArXiv : 1710.06922.
- David Lewis. 1969. *Convention : A philosophical study*. John Wiley & Sons.
- Mike Lewis, Denis Yarats, Yann Dauphin, Devi Parikh, and Dhruv Batra. 2017. [Deal or No Deal? End-to-End Learning of Negotiation Dialogues](#). In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 2443–2453, Copenhagen, Denmark. Association for Computational Linguistics.

- George A Miller and Patricia E Nicely. 1955. An analysis of perceptual confusions among some english consonants. *The Journal of the Acoustical Society of America*, 27(2) :338–352.
- Igor Mordatch and Pieter Abbeel. 2017. Emergence of Grounded Compositional Language in Multi-Agent Populations. *arXiv preprint arXiv :1703.04908*.
- Sean Roberts. 2018. [CHIELD: Causal hypotheses in evolutionary linguistics database](#). In *Proceedings of the 12th International Conference on the Evolution of Language (Evolang12)*. Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika.
- Philippe Schlenker, Emmanuel Chemla, Anne M. Schel, James Fuller, Jean-Pierre Gauthier, Jeremy Kuhn, Dunja Veselinović, Kate Arnold, Cristiane Căsar, Sumir Keenan, Alban Lemasson, Karim Ouattara, Robin Ryder, and Klaus Zuberbühler. 2016. [Formal monkey linguistics](#). *Theoretical Linguistics*, 42(1-2) :1–90.
- David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, et al. 2017. Mastering the game of go without human knowledge. *Nature*, 550(7676) :354.
- Luc Steels. 1997. [The Synthetic Modeling of Language Origins](#). *Evolution of Communication*, 1(1) :1–34.
- Sainbayar Sukhbaatar, Arthur Szlam, and Rob Fergus. 2016. [Learning Multiagent Communication with Backpropagation](#). *arXiv :1605.07736 [cs]*. ArXiv : 1605.07736.
- Ronald J Williams. 1992. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine learning*, 8(3-4) :229–256.
- Ludwig Wittgenstein. 1953. [Philosophical Investigations](#).