Zeitschrift: Études de Lettres : revue de la Faculté des lettres de l'Université de Lausanne

Herausgeber: Université de Lausanne, Faculté des lettres

Band: - (1997)

Heft: 3

Artikel: Les défis actuels en synthèse de la parole

Autor: Keller, Eric / Zellner, Brigitte

DOI: https://doi.org/10.5169/seals-870412

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. Find out more

Download PDF: 14.07.2025

ETH-Bibliothek Zürich, E-Periodica, https://www.e-periodica.ch

LES DÉFIS ACTUELS EN SYNTHÈSE DE LA PAROLE

LES DÉFIS ACTUELS EN SYNTHÈSE DE LA PAROLE

1. Introduction

L'arrivée de l'informatique correspond, pour un grand nombre de pratiquants des sciences humaines, au début d'une ère nouvelle. Pour la première fois dans l'expérience humaine, nous avons accès à une machine qui est suffisamment précise, qui peut retenir suffisamment d'informations et qui est suffisamment rapide pour vérifier toute la suite logique et une bonne partie de la portée, d'une hypothèse scientifique en sciences humaines. Finalement, en cette fin de XXe siècle, certaines de nos théories — qui paraissent souvent trop larges, trop englobantes et trop « floues » aux observateurs extérieurs — peuvent être soumises à une évaluation toute aussi rigoureuse que celle traditionnellement appliquée par nos collègues œuvrant dans les sciences dites « plus exactes ». Si un problème peut être formalisé, il peut être vérifié au moyen de l'informatique. Et au contraire de la vérification humaine, la vérification se fera sans erreur — si la procédure de vérification est programmée correctement.

L'utilisation de l'informatique à fins de vérification scientifique s'est avérée particulièrement fructueuse dans le domaine de la linguistique et de la phonétique. Depuis presqu'un demi-siècle déjà, les hypothèses à propos des processus et des structures linguistiques ont atteint des niveaux de complexité tels que seuls les pratiquants les plus déterminés réussissent à suivre leurs multiples formulations et reformulations, tout en tenant compte de l'ensemble de leurs conceptions de départ ainsi que de leurs divers points d'appui. Ainsi, dans le domaine des sciences humaines, la formalisation des règles linguistiques pour une langue — ou pour l'ensemble des langues humaines — représente un domaine propice à la vérification informatique d'hypothèses scienti-

fiques. Grâce à l'ordinateur, il devient possible de regagner un contrôle aussi global que détaillé, sur les informations à rendre disponibles à l'entrée du système de traitement, les processus de transformation à utiliser, ainsi que les contraintes à respecter durant le parcours du traitement informatique. Le processus de formalisation d'un problème devient par voie de conséquence beaucoup plus rigoureux.

Remarquons que les traitements linguistiques se prêtent relativement bien à une vérification informatique. Ceux-ci peuvent être représentés sous la forme d'un système de règles dont la structure a été assujettie à un raffinement continuel pendant presqu'un siècle. De plus, ces règles s'appliquent à un matériel linguistique qui est facilement représenté sur ordinateur, soit sous forme de mots écrits, soit en tant que faisceaux de traits sémantiques ou phonétiques, soit sous forme de mots transcrits phonétiquement, soit encore en tant que représentation matricielle de paramètres acoustiques.

Dans le même ordre d'idées, la parole peut être simulée au moyen d'une synthèse de la parole. Cet exercice est fort intéressant, malgré le fait que le système exécutant informatique (c'est-à-dire, les processeurs, circuits électriques, haut-parleurs) n'est évidemment pas à l'image du système neuro-physiologique humain. Par exemple, au moyen d'une synthèse, nous pouvons vérifier rapidement et aisément si nous avons correctement caractérisé le problème de la synthèse dans sa dimension logique et formelle. En effet, il est possible d'argumenter que l'examen par la synthèse est la vérification la plus sévère possible, car l'oreille humaine bien entraînée à l'écoute de la parole est un juge impitoyable : elle identifie avec certitude un dialecte vaudois ou méridional, elle repère les moindres traces d'un accent non-francophone. De la même façon, elle se rebute de manière presque violente contre les défauts de la synthèse computationnelle de la parole. Cette aptitude nous permet donc de vérifier assez aisément les capacités d'un logiciel à émuler les performances d'un bon lecteur à haute voix, natif de la langue française. On est ainsi en mesure de vérifier rapidement la pertinence et l'adéquation de chacune des règles linguistiques et phonétiques sous-jacentes à un tel système. Il suffit d'ajouter la règle en question au système et de la vérifier sur un ensemble d'énoncés pertinents. On peut le cas échéant, la modifier ou l'enlever. Par cette procédure expérimentale, on peut établir la pertinence et la valeur de chaque règle dans le cadre complet du système.

Mieux encore, on peut soumettre plusieurs conceptions concurrentielles de la génération linguistique au test expérimental. Si une conception réussit à mieux émuler la parole humaine qu'une autre — ou de manière nettement plus efficace — nous sommes amenés à considérer que l'approche en question est mieux étayée. Et c'est en ce sens que plusieurs auteurs de ce volume utilisent la synthèse de la parole. En soumettant plusieurs conceptions concurrentielles au même crible expérimental, ces auteurs ont été amenés à de nouvelles conceptions de certains sous-ensembles de règles. De cette manière, la synthèse devient un outil majeur de notre métier de linguiste et de phonéticien.

2. *Le volume actuel*

Nous présentons dans ce volume cinq auteurs qui partagent cet espace expérimental de l'informatique contemporaine. Dans l'article initial, E. Keller résume l'état actuel de la synthèse de la parole en dessinant les grandes lignes du parcours ces cinquante dernières années dans ce domaine. Il évoque en particulier les travaux accomplis depuis une vingtaine d'années grâce à un processus d'inversion des résultats obtenus à partir d'analyses linguistiques, psycholinguistiques et phonétiques. Il suggère que les améliorations futures de la synthèse se situent au delà des domaines traditionnellement recherchés en sciences de la parole, dans le domaine des marques phonétiques non phonémiques et individuelles, qui enrichissent la parole synthétisée de manière non négligeable.

Dans le domaine de la synthèse, J. Local est un nom bien connu, car il s'agit d'un chercheur qui poursuit de manière systématique et approfondie les fondements de la prochaine génération de synthétiseurs. Dans l'article de ce volume, il met le doigt sur un grand nombre d'exemples de variations phonétiques systématiques, généralement négligées par les analyses phonologiques et phonétiques traditionnelles. Ces variations sont pourtant d'une importance évidente dans le processus de l'amélioration des systèmes de synthèse contemporains.

B. Zellner aborde un domaine de grande importance pour la bonne perception et compréhension des messages synthétiques qu'est celui de la structure temporelle de la parole. En partant d'un examen des fondements physiologiques, psycholinguistiques et sociolinguistiques, elle définit les contraintes temporelles que

la parole doit respecter. Elle motive et circonscrit la logique psycholinguistique sous-jacente à l'algorithme qui en a été dérivé — orientation inhabituelle pour un système de synthèse — et qui a été incorporé dans le système Lausannois de synthèse de la parole.

Dans l'article suivant, S. Werner effectue un survol des systèmes utilisés en synthèse de la parole pour la génération de la fréquence fondamentale, laquelle correspond largement aux variations de la hauteur de la voix. Il s'attarde plus particulièrement sur une approche qui est physiologiquement et expérimentalement motivée, la méthode de Fujisaki. Cette méthode a fait ses preuves par rapport à plusieurs langues et systèmes de synthèse de la parole, y compris celui développé à Lausanne pour le français. Pourtant, la méthode n'est pas encore bien connue en francophonie et mérite bien cette revue détaillée de la question.

Finalement, le volume se clôt avec une contribution de G. Caelen-Haumont et E. Keller portant sur un domaine qui se situe aux frontières de la synthèse actuelle, celui de l'identification du focus sémantique. Ce domaine est clairement appelé à apporter une contribution majeure à l'évolution des systèmes de synthèse futurs, car tout message oral est porteur de sens. Si les systèmes actuels réussissent à transmettre des messages oraux à peu près correctement, ils ne sont généralement pas en mesure d'y relever les éléments sémantiquement saillants. La caractérisation, spécification et automatisation du focus sémantique représentent par conséquent, une des ultimes frontières de notre science.

3. Le contexte

Il nous fait plaisir de présenter ce volume, car il représente la fruit de plusieurs années de réflexion, de discussion et de collaboration internationale. En effet, ces cinq auteurs, provenant de quatre pays (la France, l'Angleterre, la Finlande et la Suisse), partagent une formation fortement orientée vers les sciences phonétiques et la psycholinguistique, en sus de leurs préoccupations de longue durée avec cette partie de l'intelligence artificielle qu'est la synthèse de la parole. Il est donc naturel que nous partagions largement les mêmes suppositions de base en ce qui concerne les défis actuels posés à la synthèse de la parole. Ces partages se sont concrétisés une première fois quand nous nous sommes rencontrés en février 1996 pour présenter ces idées dans

le cadre d'un colloque spécial à l'Université Paris VII (organisateur : G. Boulakia). Suite à un travail intensif d'approfondissement, de discussion et d'édition de nos idées respectives, nous avons le plaisir de nous retrouver ensemble dans ce volume. En lisant ces pages, on peut constater que notre idée centrale concernant les défis actuels posés à la synthèse n'a fait que se concrétiser davantage : loin d'être un problème résolu, la synthèse de la parole n'a traversé que les étapes initiales des défis ultérieurs à relever. Nous sommes encore loin d'une synthèse véritablement naturelle, proche de la production orale humaine. Et en tentant de créer une synthèse véritablement naturelle, nous serons obligés — presqu'inévitablement — de mener de front un processus de découverte particulièrement intéressant.

Novembre 1997

Eric KELLER
Brigitte ZELLNER
Université de Lausanne