

Dictée vocale d'expressions mathématiques pour l'accessibilité de l'écriture des sciences

Participants et affiliations

Lorenzo Brucato¹, Lucas Ondel Yang¹, H  l  ne Bonneau¹, Thomas Mordant², Mathieu Mourichoux³, Jean-Beno  t Bost³, Nicolas Denier¹

1 : LISN, Universit   Paris-Saclay, CNRS

2 : LMO, Universit   Paris-Saclay, CNRS

3 : UMPA,   cole Normale Sup  rieure de Lyon, CNRS



R  sum   du projet

L'objectif principal de ce projet est de produire un syst  me de transcription de la parole en expressions math  matiques. En effet, l'  criture de documents scientifiques est une t  che fastidieuse pour les personnes atteintes de certains handicaps moteurs ou troubles dys. Cela concerne de nombreux   tudiants ou professionnels, plus particuli  rement des chercheurs en math  matiques avec qui ce projet est construit. Cette collaboration int  gre de potentiels utilisateurs dans le processus de conception, ce qui est essentiel dans cette d  marche d'am  lioration de l'accessibilit  . Certains sont aid  s d'assistants pour r  diger en leur nom, mais peu sont form  s aux math  matiques, ou au langage LaTeX permettant de les r  diger. Ce langage est largement utilis   pour la r  daction de contenu scientifique, il a une syntaxe complexe mais n  cessaire car il offre un contr  le pr  cis    ses utilisateurs. Ainsi, l'approche propos  e vise    produire du code LaTeX    partir d'expressions dict  es (Speech2Tex).

Cette preuve de concept est s  par  e en deux   tapes successives :

- D'abord, un syst  me de reconnaissance de la parole (automatic speech recognition, ou ASR) est utilis   (Whisper medium), permettant la transcription de la parole en texte.
- Puis, un syst  me de transducteurs (weighted finite state automata, ou WFST) permet de convertir le texte en symboles du langage LaTeX, et s'assurer de la coh  rence de sa structure. Un autre avantage est que l'on peut en composer plusieurs, ici, un d  finissant le lexique, et un autre une grammaire plus g  n  rale.

L'  tude r  alis  e lors d'un stage de master a pu mettre en   vidence les d  fauts des syst  mes classiques actuels tels que les LLM pour ce type de t  che sp  cifique. Par exemple, ces mod  les ne sont pas entra  n  s pour la reconnaissance de math  matiques, qui a des caract  ristiques particuli  res (prononciation par caract  re plut  t que mot, parenth  ses implicites, etc.). Aussi, de nombreuses ambigu  t  s peuvent survenir, et les transducteurs offrent la possibilit   de mieux les g  rer. Un autre moyen identifi   pour la d  sambigu  sation est de d  tailler la dict  e, en pr  cisant par exemple chaque parenth  se, mais cela ajoute un peu de complexit   pour l'utilisateur.

Un premier jeu de donn  es de 245 locutions a   t   collect   puisqu'aucun n'avait   t   produit dans ce contexte. Des mesures de taux d'erreur pour diff  rentes compositions de syst  mes ont   t   effectu  es. Les meilleures performances sont obtenues avec un syst  me d'ASR id  al (uniquement la partie transducteur, auquel on a fourni les transcriptions exactes), et pour une dict  e d  taill  e. On obtient alors un taux d'erreur au mot de 24%, ce n'est pas encore utilisable mais prometteur. L'int  r  t de l'ajout du syst  me de grammaire a   t   mis en lumi  re, car il am  liore les performances pour une dict  e libre, donc moins contraignante pour l'utilisateur.

Enfin, de nombreuses am  liorations possibles ont   t   relev  es pour produire un syst  me robuste et utilisable. Par exemple, d  finir une syntaxe orale pour la prononciation de math  matiques, ainsi qu'une m  trique d'  valuation plus adapt  e. Il est donc envisag   que ce projet soit poursuivi car l'application de transducteurs s'est r  v  l  e adapt  e    cette t  che.

Une d  monstration est envisag  e, ne n  cessitant pas de mat  riel suppl  mentaire    ce qui sera apport   par les pr  sentateurs du projet.