Toward A Reinforcement-Learning-Based System for Adjusting Medication to Minimize Speech Disfluency

Pavlos Constas, Vikram Rawal, Matthew Honorio Oliveira, Andreas Constas, Aditya Khan, Kaison Cheung, Najma Sultani, Carrie Chen, Micol Altomare, Michael Akzam, Jiacheng Chen, Vhea He, Lauren Altomare, Heraa Muqri, Asad Khan, Nimit Amikumar Bhanshali, Youssef Rachad, Michael Guerzhoy

University of Toronto

Abstract and Introduction

We propose a hypothetical Reinforcement Learning-based system for helping physicians adjust medications for people with speech disfluency. We focus on a proof-of-concept of the system. We show via simulation that, for plausible parameters of medication/patient interactions and for our disfluency detection system, that is possible.

Regulatory and clinical issues would need to be addressed before actually implementing such a system. However, we show that such a system can in principle be implemented, and automatic disfluency detection can work well enough to support such a system.

On a high level, the system works as follows: the person with a speech disfluency interacts with a system that measures their speech fluency; the system tries to assign different medications in different doses to the person to minimize their speech disfluency. Using Reinforcement Learning (RL), the system trades off exploring to find the best medication combination for the person and exploiting effective medication combinations that have been found.

Data Collection

We compile a dataset of of psychiatric medications and the range of their effecs. We sourced audio samples from publicly available YouTube vlogs, focusing on individuals discussing their experiences with mental health issues. Utilizing a combination of keywords related to depression and anxiety, we curated a collection of 195 audio clips from 19 channels. Data was rated in two stages. Data from the first stage was not used in our experiments. The round was used for acquainting raters with the variation of disfluency observed in the dataset. At the end of this phase, each rater was privately given summary statistics regarding their ratings in the round, including the mean and standard deviation of their ratings across audio samples, as well as a spreadsheet containing a measure of their bias for each audio sample (where bias is the distance of their rating from the mean rating across all raters for that audio sample). This process was aimed at allowing raters to recognize possible inconsistencies and biases in their "internal model" of disfluency. The ratings from the second stage were the finalized ratings that would be used for fine-tuning our disfluency-detection system. The ratings were standardized, as described below.

Rater Performance Analysis

We analyzed rater consistency in assessing speech disfluency using the regression model:

$$r_{ij} = \alpha_i + \beta_j + \varepsilon_{ij}$$

where r_{ij} is the rating for clip j by rater i, α_i represents rater bias, β_j is the true disfluency level of the clip, and ε_{ij} is the error term. Our validation using a split dataset approach showed an RMSE of $0.9/6.0$, indicating moderate rater consistency.

Disfluency Pipeline

We designed a pipeline to assess speech disfluency by transcribing audio with Whisper [3], tagging disfluency with an Auto-Correlational Neural Network (DT-ACNN) [1], and predicting disfluency scores using fine-tuned GPT-2.

Results of the Disfluency Pipeline

Our system achieved promising accuracy in disfluency prediction, with validation showing a standard error of approximately $0.15/6$. These results demonstrate our approach's effectiveness in disfluency measurement.

Figure 1. Learning curves: disfluency prediction regression task.

Reinforcement Learning

We apply RL to adjust medications for patients with speech disfluency, aiming to minimize these symptoms effectively by exploring and exploiting different medication regimens based on simulated patient responses.

In our study, we employed the LinUCB algorithm [2] to optimize medication regimens for patients with speech disfluency. This approach allows for personalized medication management by continuously learning and adapting to each patient's unique response patterns.

Patient Simulation

We model a medication administration environment, aiming to determine the most effective medication regimen for people experiencing depression, anxiety, insomnia, and resulting speech fluency issues. The person's health state evolves based on Hidden Markov Models (HMMs). Each health issue (depression, anxiety, insomnia) has its unique HMM, governing how the patient's state progresses. The patient's observed speech fluency is also influenced by these health states.

RL Results

In our simulations, we run the RL algorithm and keep track of disfluency over time. We inject noise into the algorithm's simulated measurements of disfluency to simulate the fact that our disfluency detection system does not measure disfluency perfectly.

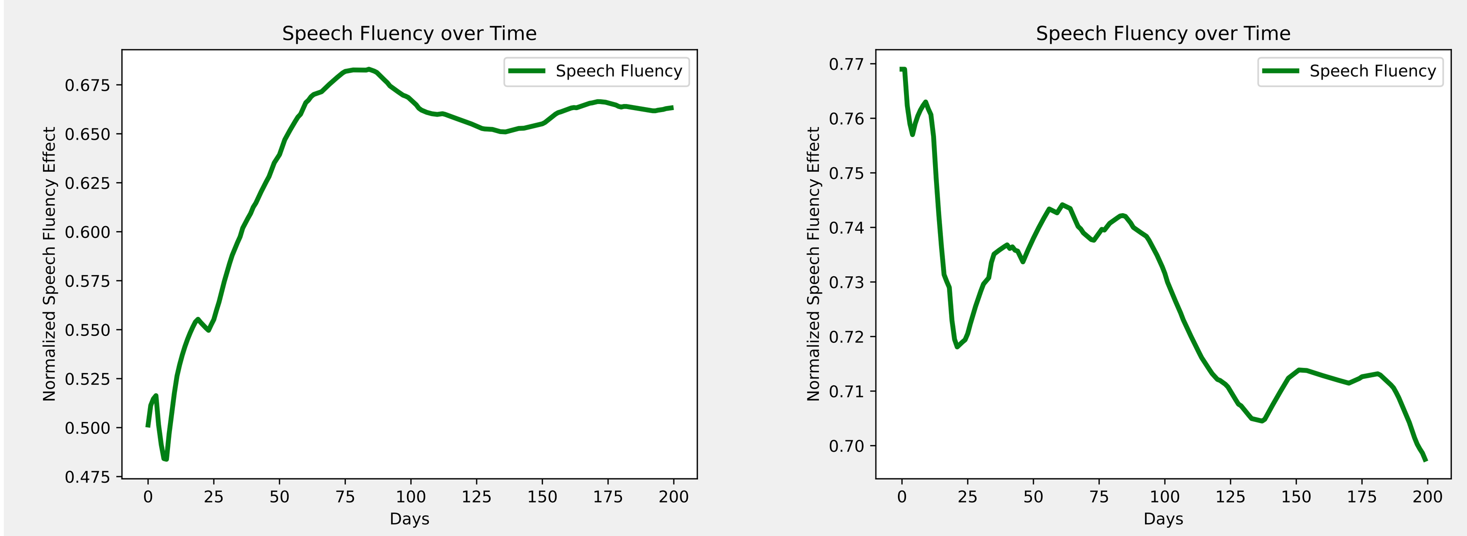

Figure 2. Left: Convergence to higher speech fluency. Right: Lack of convergence to higher speech fluency.

Our preliminary results indicate that if speech disfluency can be measured to within 10% of the true score and the medications have plausible properties, then reinforcement learning is a possible method to adjust medication dosage to minimize speech disfluency.

Ethical Considerations

The ethical implications of using an algorithm automatically adjusting medications to improve speech fluency must be taken into consideration. Such automation exists, in cases like insulin pumps, and promises improved treatment. However, it requires rigorous clinical trials and ongoing informed patient consent due to significant privacy and ethical concerns. Patients must fully understand the system's workings, risks, and benefits. Moreover, the reliance on sensitive data and the necessity for continuous oversight highlight the need for strict safeguards.

References

[1] Paria Jamshid Lou, Peter Anderson, and Mark Johnson. Disfluency detection using auto-correlational neural networks. In Ellen Riloff, David Chiang, Julia Hockenmaier, and Jun'ichi Tsujii, editors, *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 4610–4619, Brussels, Belgium, October-November 2018. Association for Computational Linguistics.

[2] Elliot Nelson, Debarun Bhattacharjya, Tian Gao, Miao Liu, Djallel Bouneffouf, and Pascal Poupart. Linearizing contextual bandits with latent state dynamics. In *Uncertainty in Artificial Intelligence*, pages 1477–1487. PMLR, 2022.

[3] Alec Radford, Jong Wook Kim, Tao Xu, Greg Brockman, Christine McLeavey, and Ilya Sutskever. Robust speech recognition via large-scale weak supervision. In *International Conference on Machine Learning*, pages 28492–28518. PMLR, 2023.

 guerzhoy@mie.utoronto.ca