Learning informative rna sequence embeddings for reinforcement learning-based multiple sequence alignment

Connor Clark-Baba, Basil Jancso-Szabo & Leon A. Li *
Department of Computer Science
McGill University
Montreal, Quebec, Canada
{connor.clark-baba,leon.a.li, Basil.jancso-szabo}@mail.mcgill.ca

Parham Ghasemloo
Department of Computer Science
McGill University
Montreal, Quebec, Canada
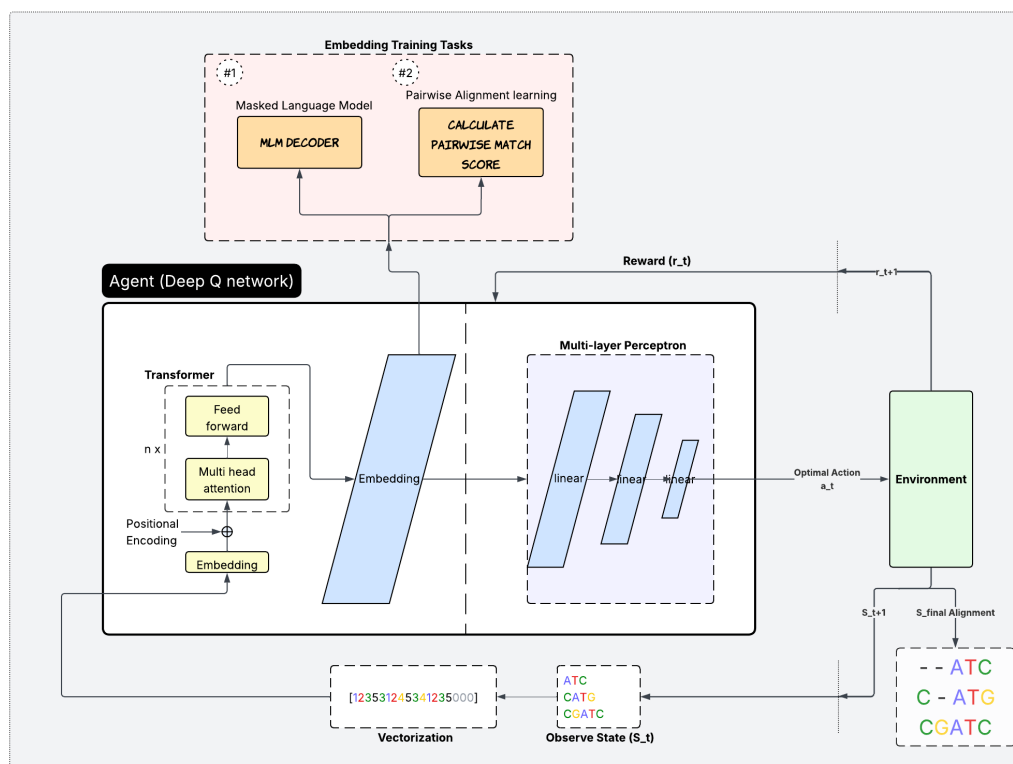{parham.ghasemloogheidari}@mail.mcgill.ca

Figure 1: Overview of the proposed Transformer-based Deep Q-learning framework for multiple sequence alignment and embedding learning. The agent integrates a multi-head self-attention mechanism within a Transformer encoder and a multi-layer perceptron for alignment quality estimation. Two auxiliary embedding training tasks, masked language modeling (MLM) and pairwise alignment learning (AL), are employed to contextualize sequence representations and capture underlying structural relationships among sequences.

---

*Connor Clark-Baba, Basil Jancso-Szabo, Leon A. Li are student contributors to this work. Parham Ghasemloo serves as the mentor with supervising capacity only.

1 BACKGROUND AND MOTIVATION

Multiple sequence alignment (MSA) of RNA sequences is a fundamental step in phylogenetic inference and RNA structure prediction. It identifies homologous positions across sequences by introducing gaps that maximize correspondence according to a predefined scoring function. Standard MSA tools such as ClustalW, MAFFT, and T-COFFEE rely on heuristic or progressive algorithms to approximate optimal alignments for what is known to be an NP-hard problem (Gaad et al., 2023). With recent advances in artificial intelligence, the community has begun exploring reinforcement learning (RL), particularly deep reinforcement learning (DRL), and natural language processing (NLP) approaches that model RNA sequences as language-like entities and the alignment process as a sequence of strategic, stepwise decisions analogous to a sequential decision problem (Liu et al., 2023; Dotan et al., 2025). Recent work has shown that combining self-attention and positional encodings with DRL can improve alignment performance, highlighting the benefit of richer sequence representations within the learning framework (Liu et al., 2023). However, the embedding components in such models are typically integrated directly into the RL pipeline rather than pre-trained through language-modeling objectives that strengthen positional and contextual understanding (Akiyama & Sakakibara, 2022). By capturing sequence relationships, this work aims to enhance the RL agent's ability to generate accurate and biologically meaningful RNA alignments.

2 RESEARCH QUESTION AND METHODOLOGY

To limit the scope of this project, we will adopt the DRL implementation of Liu et al. (2023) to investigate the research question: Can RNA sequence embeddings pretrained with language-model objectives enhance the performance of RL-based MSA? We will evaluate three model variants. **(1) Baseline Model:** the original *Liu et al.* deep Q-network with jointly trained embeddings, reproduced using the authors' training setup and data. **(2) Frozen-Pretrained Model:** the baseline Q-network paired with a pretrained RNA encoder trained on two complementary objectives: a *masked language modeling (MLM)* task to capture contextual dependencies among bases, and a *modified alignment-learning (AL)* task to encourage alignment-consistent representations without explicit structural information (Akiyama & Sakakibara, 2022). The encoder will remain frozen during RL training to isolate the effect of representation quality. **(3) Fine-Tuned Model:** initialized with the same pretrained encoder but jointly fine-tuned with the Q-network to evaluate how pretraining interacts with RL optimization. Comparing the performance of these three configurations will reveal whether language-model-based pretraining improves alignment accuracy and learning efficiency. The novelty of this work lies in integrating language-model pretraining with DRL for RNA MSA.

3 HYPOTHESIS AND EXPECTED RESULTS

We hypothesize that pretraining RNA sequence encoders with language-model objectives will improve the performance of RL-based MSA algorithms. Specifically, embeddings trained through MLM and modified AL tasks are expected to provide richer sequence representations, enabling the RL agent to make more accurate alignment decisions. We anticipate that **model (2), (3)** will outperform the baseline model in alignment accuracy and convergence speed, with **model (3)** achieving the best overall balance of accuracy and training efficiency.

4 EXPERIMENTAL DESIGN

Experiments will use the DRL implementation of Liu et al. (2023) as a base, evaluated across the three model variants mentioned above. Training and inference will be performed on GPU servers available through the Waldispühl lab. To ensure computational feasibility, the training dataset from Liu et al. (2023) will be filtered to include only sequence groups with a total input length of up to 150 bases. This configuration requires approximately 2–3 GB of VRAM during training and under 1 GB for inference. Model performance will be evaluated using sum-of-pairs (SP) and column score (CS), along with training and prediction times to assess efficiency. Convergence will also be monitored to evaluate stability across models (Gaad et al., 2023).

5 PROJECT TIMELINE AND ROLES

All students will contribute equally to all phases of the project, while the mentor provides high-level guidance and feedback. The project will be conducted in three phases: by October 31, completion of training for model (1); by November 15, completion of training for models (2) and (3); and by November 30, evaluation, visualization, and final analysis. The writing component will be developed continuously throughout the project.

REFERENCES

Manato Akiyama and Yasubumi Sakakibara. Informative rna-base embedding for rna structural alignment and clustering by deep representation learning. *NAR Genomics and Bioinformatics*, 4 (1):lqac012, 2022. doi: 10.1093/nargab/lqac012.

Edo Dotan, Elya Wygoda, Noa Ecker, Michael Alburquerque, Oren Avram, Yonatan Belinkov, and Tal Pupko. Betaalign: a deep learning approach for multiple sequence alignment. *Bioinformatics*, 41(1):btaf009, 2025. doi: 10.1093/bioinformatics/btaf009.

Chaimaa Gaad, Chadi Mohamed-Amine, Mohamed Sraitih, and Ahmed Aamouche. Exploring reinforcement learning methods for multiple sequence alignment: A brief review. *BIO Web of Conferences*, 75:01004, 2023.

Yuhang Liu, Hao Yuan, Qiang Zhang, Zixuan Wang, Shuwen Xiong, Naifeng Wen, and Yongqing Zhang. Multiple sequence alignment based on deep reinforcement learning with self-attention and positional encoding. *Bioinformatics*, 39(11):btad636, 2023. doi: 10.1093/bioinformatics/btad636.