ICLR
International Conference On Learning Representations

Language Technologies Institute

Strategic Planning and Rationalizing on Trees Make LLMs Better Debaters

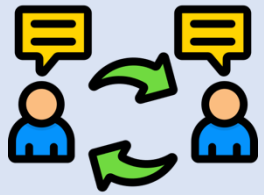
Danqing Wang, Zhuorui Ye*, Xinran Zhao*, Fei Fang, Lei Li*

Carnegie Mellon University

Unique Challenges in Competitive Debate

❖ Interactive Nature

- ➢ Unlike reading a prepared speech, debate requires real-time response to challenges



❖Time Constraint

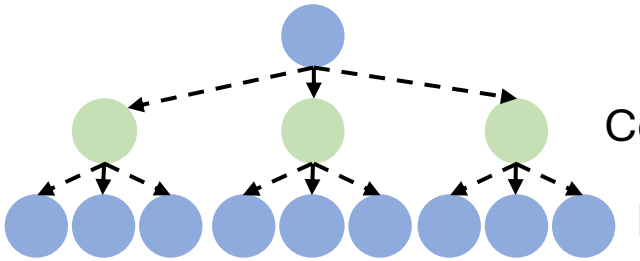
- ➢ Debaters must strategically choose their points and allocate time

❖Participant Engagement

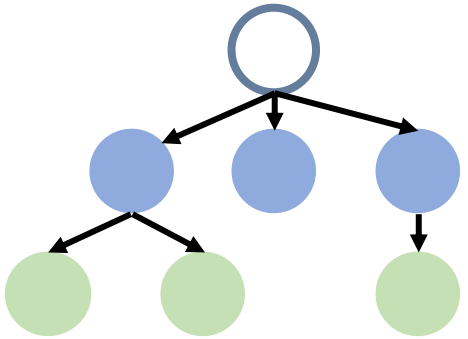
- ➢ The winning rely on both opponent and audience
  - ✓ How the debater defend themselves and attack the opponent
  - ✓ How the debater persuade the audience

How humans solve these challenges?

- Humans rehearse *before the debate*
  - anticipate the counter-claims
  - prepare the potential defense
  - formulate them in a tree-form reasoning



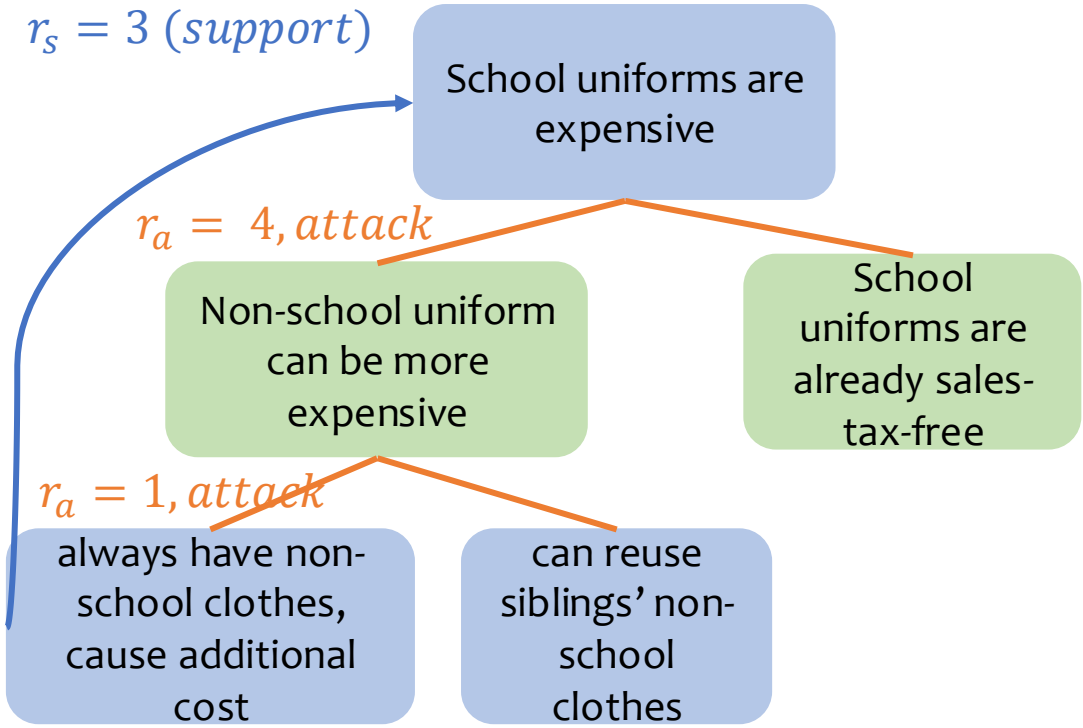
- Humans keep notes *during the debate*
  - keep a structured mental map of which points have been addressed or remain standing
  - adjust their strategies based on the status



TreeDebater: Build the Rehearsal Tree before debate

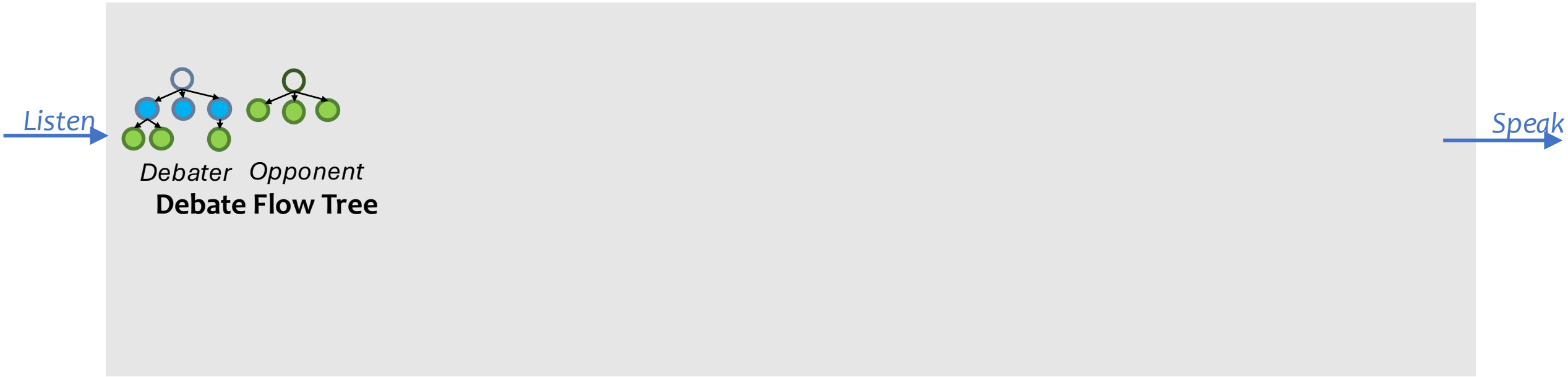
- Calculate the *k-step strength score* in a minimax style
  - the attack and support of its descendants within depth k (*k-subtree*)
- What's the payoff in the worst case
  - opponent use the strongest counter-claim
  - the debater use the best defense

$$f_k(x^l) = f_0(x^l) - \gamma \cdot \max_{x^{l+1} \in \text{Child}(x^l)} f_{k-1}(x^{l+1}),$$



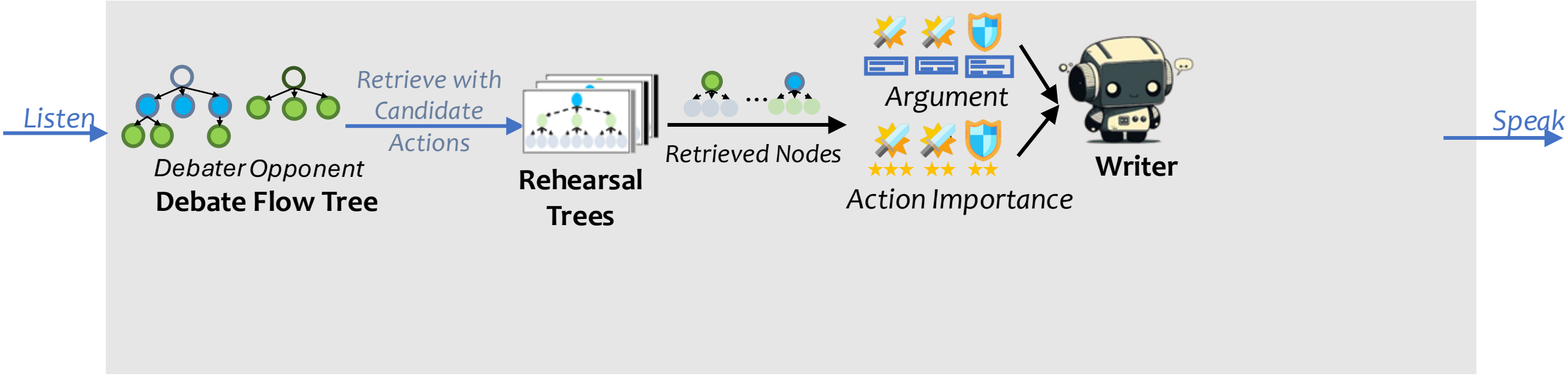
TreeDebater: Update Debate Flow Tree during debate

- After listening, extract the claims from the new statement
- Update the node status in two debate flow trees (debater & opponent)



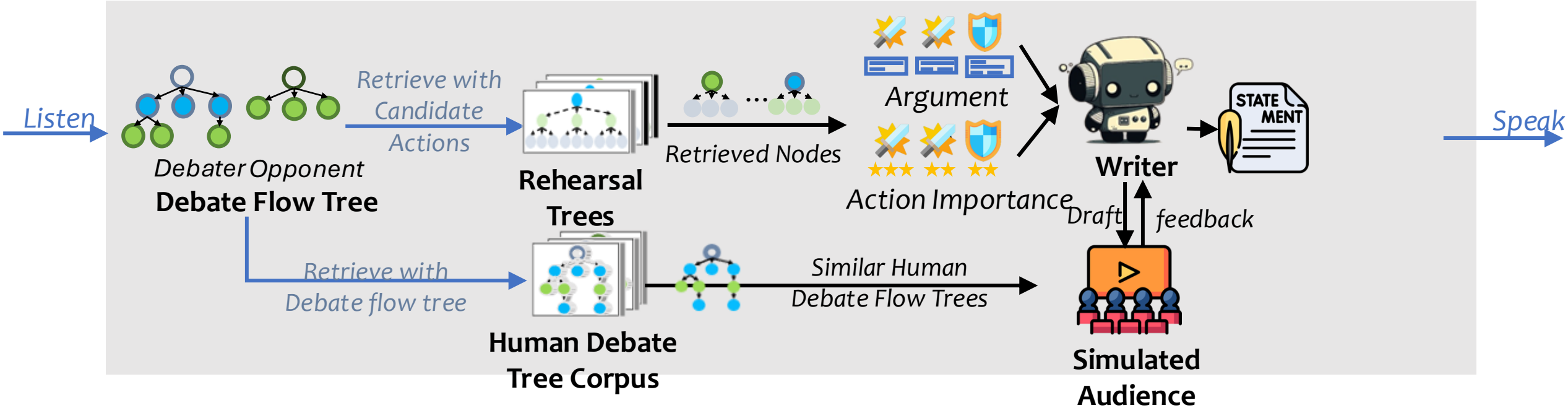
TreeDebater: Retrieve from Rehearsal Tree during debate

- ❖ Identify candidate actions based on node status in Debate Flow Tree
- ❖ Retrieve prepared arguments and strength scores from Rehearsal Tree



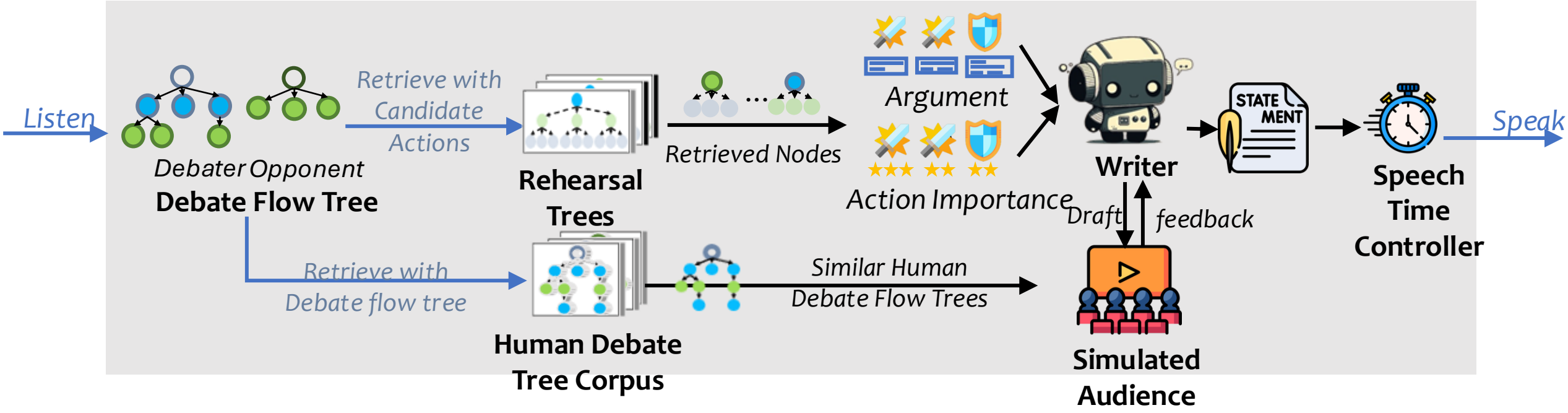
TreeDebater: Refine the statement during debate

- ❖ Draft the statement
- ❖ Retrieve similar human debate flow trees as the reference
- ❖ Simulate audience feedback and refine based on the feedback



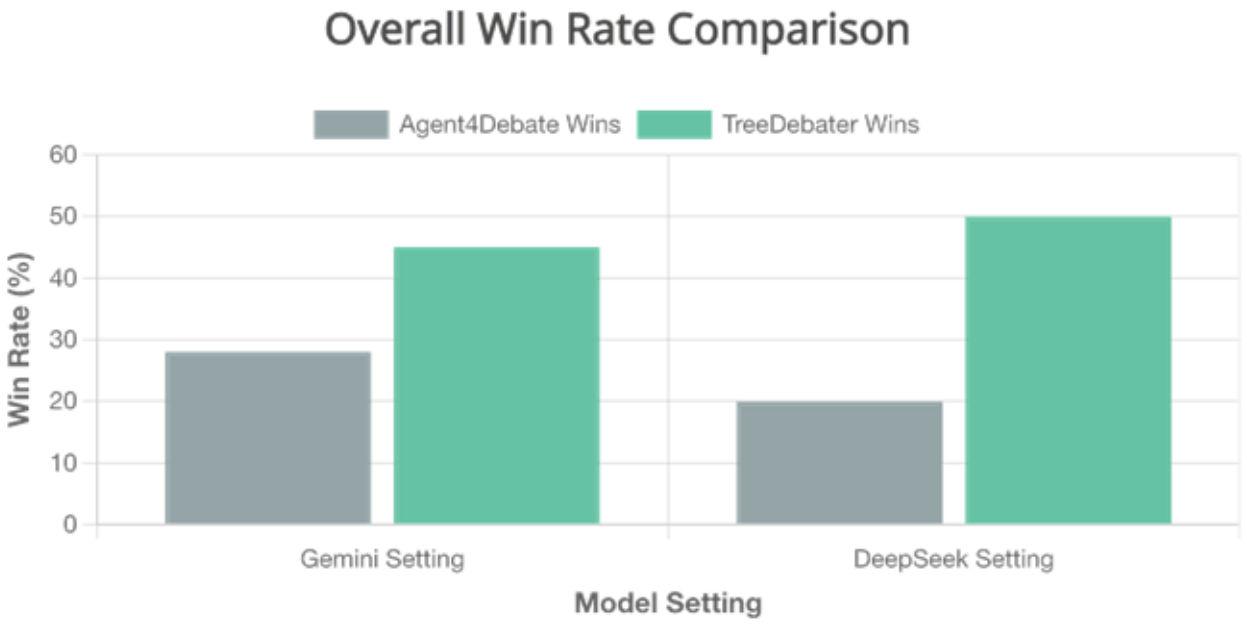
TreeDebater: Control Speech Time during debate

- Use lightweight text-to-speech model to estimate the speech time
- employ the search approach to efficiently find an appropriate word count target for the statement



Our approach shows significant improvements in Opinion Shift Win rates

- Baseline
  - Agent4Debate (multi-agent debate system)
- End-to-End Debate Level Evaluation
  - conduct two debates on each motion with the original and swapped sides.
  - 3 participants are asked to score each stage and vote before and after the debate



Outperform in overall persuasiveness

- Head-to-Head Stage Level Evaluation
  - randomly sample 10 (motion, stance assignment) settings for each stage
  - Given two version for the same side and stage.
  - Ask for the persuasiveness score and preferred version

Gemini Backbone

Overall Persuasiveness Across All Stages

Agent4Debate (Baseline)	→	TreeDebater (Ours)
3.54		3.69

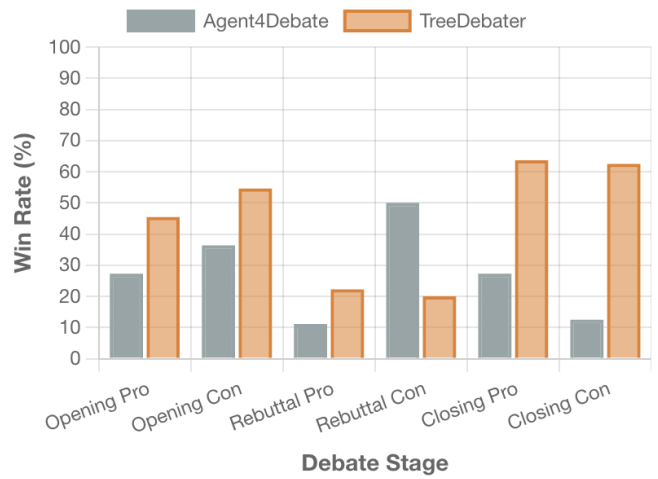
+4.2% improvement in persuasiveness

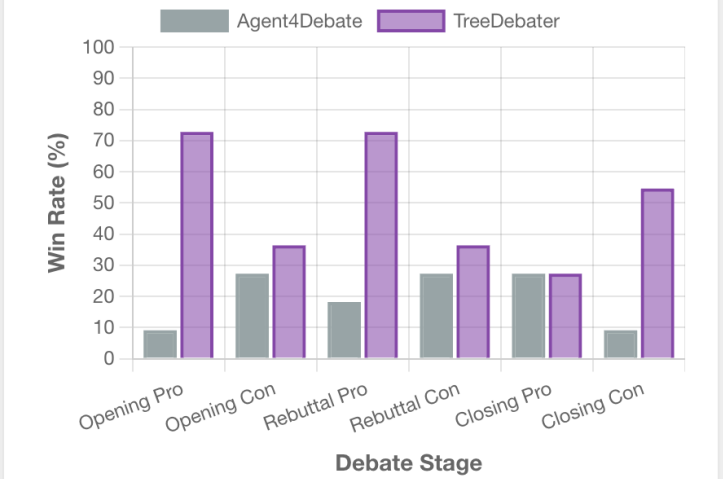
DeepSeek Backbone

Overall Persuasiveness Across All Stages

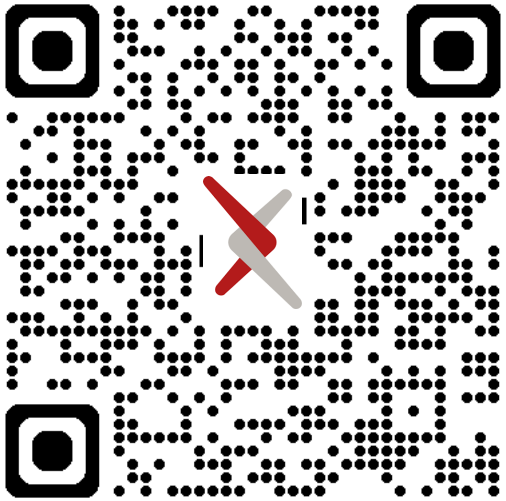
Agent4Debate (Baseline)	→	TreeDebater (Ours)
3.47		4.01

+15.6% improvement in persuasiveness

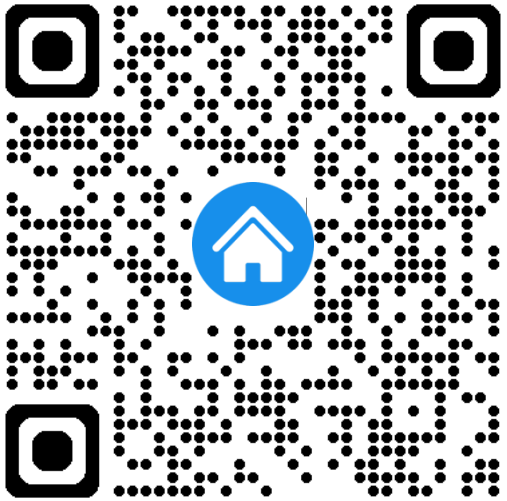




Interested to Learn More?



Paper



Website