

1 Introduction

본 리포트는 논문 Bertsimas, Dimitris, and Alex Paskov. "An exact and interpretable solution to wordle." (2022)의 개괄적 요약입니다. 프로젝트의 사용할 아이디어를 얻기 위한 요약이므로 모든 내용이 포함되어 있지 않습니다.

2 Approach

기본적으로 강화학습을 사용합니다.

- Action은 적을 수 있는 단어 하나로 대응합니다.
- State는 정답 후보의 집합으로 대응합니다.

전이 확률은 다음과 같습니다:

$$p(s_{t+1} \mid s_t, a) = \frac{|\{w \in s_t \mid \text{color}(a, w) \text{ 결과 적용 시 } s_{t+1} \text{ 가 가능한 단어 집합이 됨}\}|}{|s_t|}$$

여기서, $\text{color}(a, w)$ 는 답이 w 일 때, 단어 a 를 추측했을 때 반환되는 색상 정보입니다.

이것을 이용하면 다음의 벨만 방정식을 세울 수 있습니다:

$$V_t^*(s_t) = \min_{a \in \mathcal{A}} \left[1 + \sum_{s_{t+1}} p(s_{t+1} \mid s_t, a) V_{t+1}^*(s_{t+1}) \right]$$

이것을 동적계획법 top-down 방식으로 계산하면 됩니다.

3 Optimize

논문에서는 최적화를 위해 다음 커팅을 추가로 적용하였습니다:

- 5번째 질의 이후 집합의 크기가 1보다 크면 $V_t(s_t)$ 를 ∞ 로 처리, 1이면 1로 처리
- $|s_t| = 1$ 이면 $V_t(s_t)$ 를 1로 처리, $|s_t| = 0$ 이면 불가능 처리
- 가능한 모든 상태에서 $V_{t+1}^*(s_{t+1} \mid a_1) \leq V_{t+1}^*(s_{t+1} \mid a_2)$ 이면 a_2 를 $t+1$ 번째 액션 공간에서 제외
- 확률이 높은 s_{t+1} 부터 계산