

G: Interactive String Guessing

原案 : climpet

問題文 : smiken

データセット : climpet

解答 : climpet, hos, smiken, tatyam

解説 : climpet

問題概要

- ‘a’ と ‘b’ のみからなる秘密の文字列 S がある。
- 「ある文字列が S の部分文字列かどうか」を質問できる。 S を特定せよ。

制約

- S の長さは 1000 以下
- 質問回数は 1100 回まで
- ジャッジは適応的ではない。つまり、質問に応じて S が変化することはない。
 - この問題自体は適応的だったとしても解けますが、ジャッジ側の実装が無理でした。

愚直解法

- 質問回数の上限を無視すれば、以下の方法で S を特定できる。
 - T を、はじめ空文字列とする。
 - 次のようにして、 T を左方向にできる限り伸ばしていく。
 - ' a ' + T を質問して Yes なら、 T の先頭に ' a ' を追加。
 - そうでなく、かつ ' b ' + T を質問して Yes なら、 T の先頭に ' b ' を追加。
 - いずれも No なら終了。
 - 同様の方法で T を右方向に伸ばしていく。
 - 最終的に $T = S$ となる。
- S の長さの上限を L ($= 1000$) とおくと、この方法では最大 $2L$ 回の質問が必要となり、誤答となる。

想定解法 (1/2)

- 愚直解法で「 $'b'+T$ を質問」する部分は、言い換えると、 S の左端に到達したかどうかを確認していることに相当する。この質問で No となるのは1回だけであり、毎回質問するのは無駄が多い。
- そこで、 $'a'+T$ を質問して No だった場合、 $'b'+T$ を質問せずに、投機的に $'b'$ を T の先頭に追加することにする。
- これだけだと伸ばしすぎることがある。そこで、ある整数 m に対し、 m 回連続で No だった場合、現在の T を質問する。
 - Yes であればそのまま継続する。
 - No であれば伸ばしすぎていることになるので、どこで止めるべきだったかを二分探索する。
- 右方向についても同様。
 - あるいは、一気に $L (= 1000)$ 文字まで伸ばしたあと、二分探索してもよい。

想定解法 (2/2)

- 以上の解法の質問回数は $L + m + L/m + O(\log L)$ 回となることが示せる。
- $m = \sqrt{L}$ とすれば、 $L + 2\sqrt{L} + O(\log L)$ 回の質問で答えが求められる。
 - 細かな実装にもよるが、1080 回程度あれば十分。

乱択による改善

- ジャッジが適応的でないことを利用して、さらに質問回数を減らすこともできる。
- 左方向に伸ばすとき、先ほどの方法では常に 'a'+T を質問していたが、代わりに 'b'+T を質問しても良い。これをランダムに選ぶことにすると、「端に到達した」以外の理由で m 回連続で No となる回数の期待値は、 $L/2^m$ 回程度に抑えられる。
- $m = \log_2 L$ とすれば、質問回数の期待値は $L + O(\log L)$ 回になる。
 - 実装にもよるが 1020 回ちょっとで可能。

注意事項

- 空文字列や 1001 文字以上の文字列を質問すると即誤答となります。
- 二分探索ではなく「1文字ずつ削っていく」とすると、質問回数の上限を超える場合があります。

統計情報

- AC teams / Trying teams
 - 12 + 4 / 21
- First Acceptance
 - Screenwalkers (64 minutes)