インプットメソッド

かな漢字変換の
話題です

インプットメソッドとは？

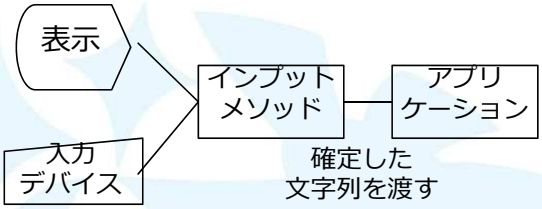
- キーボードから直接入力できる文字を
  使う人たち　vs
  直接入力できない文字を使う人たち
  - 直接入力できない ⇒ 何か方法が欲しい
- で、考えること
  - 変換をどこで行うか、どう組み込むか
  - どう取り替え可能にするか
  - どうやって変換するか ← ここでは扱わない

(脱線) インプットメソッド

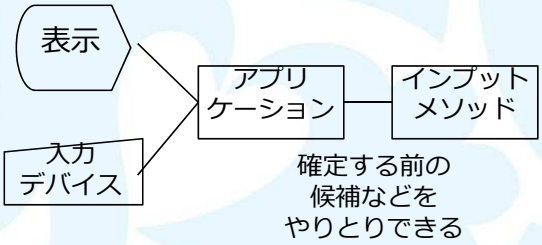
- (実は) キーボードに限らない
  - 音声
  - 手書き文字認識
  - スマートフォン等でのさまざまな(漢字)入力

フロントエンド・バックエンド

- フロントエンド方式
  - 変換してからアプリへ渡す

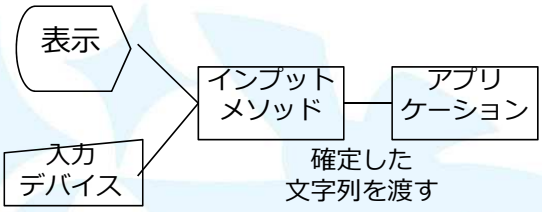


- バックエンド方式
  - 変換前にアプリが関与する
    変換方法・候補選択や
    表示の方法・位置などを
    アプリが制御できる

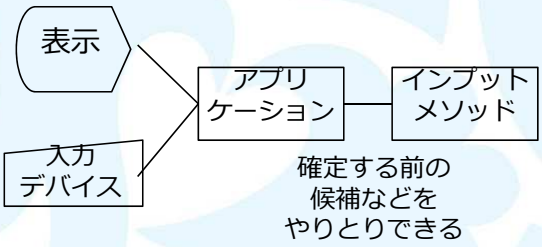


フロントエンド・バックエンド

- フロントエンド方式
  - 変換してからアプリへ渡す
  - 単純だが制限が多い



- バックエンド方式
  - 変換前にアプリが関与する
    変換方法・候補選択や
    表示の方法・位置などを
    アプリが制御できる
  - アプリの負担が多くなる



変換エンジン

- 例：(ローマ字)かな漢字変換エンジン
  - キーボードからローマ字 ⇒ かな ⇒ 漢字
  - 同一システム内でも複数エンジン選択可能
    Windowsでも、Microsoft IME (これもOS付属版とOffice付属版で細かい点が異なる）の他、他社有料のATOKやGoogle日本語入力などが使える
  - LAN上でエンジンを共有
    LAN内でPCを移っても同じエンジン＋辞書が使える
  - インターネット上でエンジンを共有
    どこで端末を移っても同じエンジン＋辞書が使える

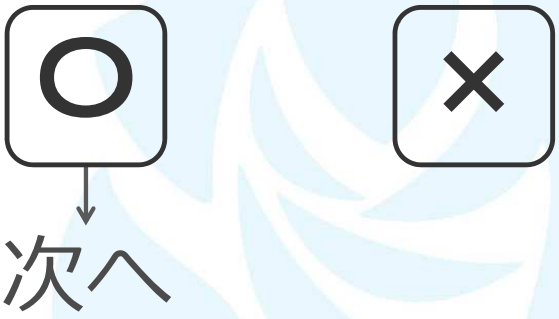
(脱線) かな漢字変換の方法？

- 要求：
  - 高速応答性 ～ あまり複雑な処理はできない
  - 小さい ～ プログラムやデータが妥当な大きさ
- 辞書かルールか？
  - なるべく辞書に頼る ⇒ 辞書が膨大になる
  - なるべくルールに頼る ⇒ 例外処理が多くなる
- 自然言語処理(構文解析・意味処理)か？
  - 本格的な自然言語処理は複雑すぎる
  - 統計的処理、特に確率遷移モデルがよく使われている
  - 誤変換ゼロは非常に難しい

ここまでのまとめ

1. インプットメソッドの位置
   フロントエンド vs バックエンド
   (大きなシステムでは)バックエンドが普及
2. 変換エンジン
   ネットワーク化・ネットワーク透過
   ← 辞書や設定を共通に使えるように

インプットメソッドが何か理解できましたか？



次へ