

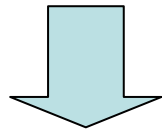
魂の通う人工無能(?)

分散支援型人工知能(仮)

栗原一貴

分散支援型人工知能(仮)とは

- 機械だけでは解決できない問題を, 細かく分割して多数の**人間の知識**により解決するシステム. cf. Folding@home 多数の計算機力で解決
- 参加する一人一人には, 全体で何を行っているのかわからないほど問題が分割されている.
- 機械には難しいが, 一般人なら誰でも持っている能力を少しずつ借りる.
cf. 遠隔医療システム ←少数のプロの能力



二つのアプローチ

- 自然言語処理. テキスト解析を簡単な国語の問題に.
- 多数の人間の「総意」をもとに機械に「人格」を与える.

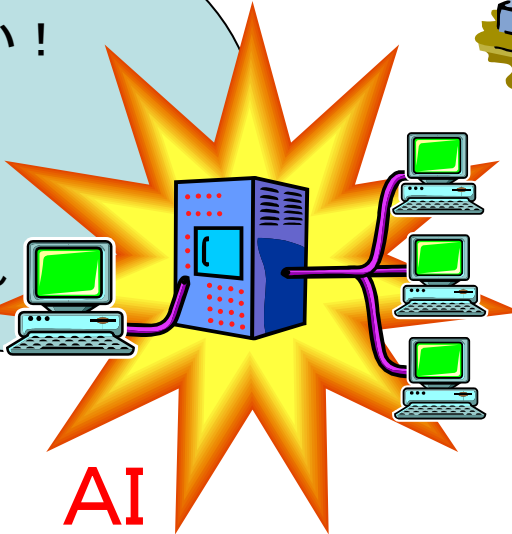
システムの概要図

(1) 機械が自分の力でできるところまでやる

(2) もし自分だけでは解決不能な場合は、
何人かのサポーターを選び、問題形式で問いかけ

「Time flies like an arrow.」
の品詞付けができない！

- ①時は矢の様に飛ぶ
- ②時蠅は矢が好き
- ③矢のように蠅を測れ



①じゃないかな

①じゃないかな

今忙しい.

(3) 結果を反映, 学習する

サポーター(人間)

チャットソフトのような常駐クライアント
が走っている。
もしくは携帯電話.

誰が買いたい(利用したい)か

- サーバ側(人工知能)
 - 安い人件費で, 大量の自然言語テキストから偏りなく情報収集し整理したい人.
 - ロボットに知能を与えたい人.
 - 国のプロジェクト, 企業, 研究機関
- クライアント側(サポーター)
 - オフィスワークに従事しており, 片手間でちょっとした小遣い稼ぎをしたい, 中卒程度の国語力を持っている一般人.
 - 携帯電話を持つ学生(◎高校生, ○大学生, △中学生)

1. 自然言語処理への応用

自然言語処理の流れと 各段階における妥当な問題形式

アナログ資料

画像取り込み

OCR

デジタル資料

形態素解析

構文解析

意味解析

文脈解析

・この文字列には何と書かれていますか？

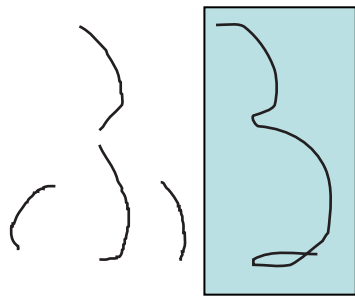
・この文(句, 節)を品詞分解しなさい。

・この文(句, 節)の構造として正しいものを選びなさい。

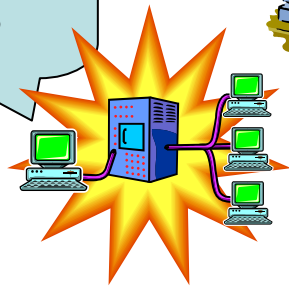
・この文(句, 節)におけるこの単語の意味を選びなさい。

・この指示語の内容を答えなさい。

・この文章Aは次の文章Bと関係がありますか？



ふるー？
ふろー？
ふ3ー？

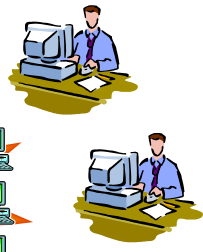
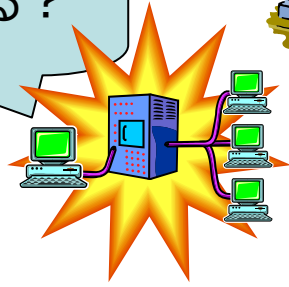


以下の文字を読み
なさい。
ふるー

- ① ふるー
- ② ふろー
- ③ ふ3ー
- ④ もっと前後の文
字が見たい

およいでから
はしる

泳いでから走る？
泳いで蚊等は知る？



単語の区切り方と
して適切なものを選
びなさい。
「およいでから
はしる」

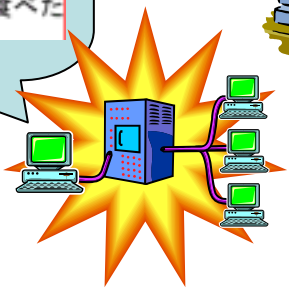
- ① 泳いでから走る
- ② 泳いで蚊等は知
る
- ③ もっと前後の文
が見たい

ただの漢字変換の問題か。

太郎は私と
カレーを食べた。

太郎
私 は カレー を 食べた

太郎は 私
カレー を 食べた



文の構造として適切なものを選びなさい。
「太郎は私とカレーを食べた。」

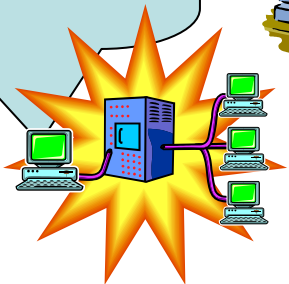
① 太郎
私 は カレー を 食べた

② 太郎は 私
カレー を 食べた

③ もっと前後の文が見たい

それはどう考えてもおかしい。

滑稽だ？
奇妙だ？
不合理だ？



次の「おかしい」の意味で適切なものを選びなさい。
「それはどう考えてもおかしい。」

① 滑稽だ

② 奇妙だ

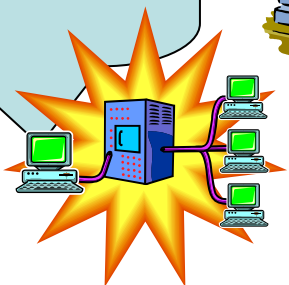
③ 不合理だ

④ もっと前後の文が見たい

あまりに短すぎて文脈がわからない場合もある。

作文

この作文（もしくはその一部）は、賛成意見？ 反対意見？

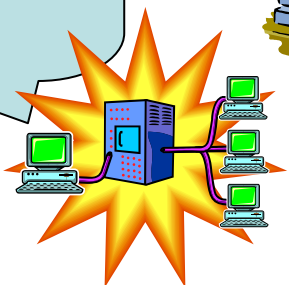


以下の文章内容は、「～」という意見に賛成ですか。反対ですか。それとも曖昧ですか。

- ① 賛成
- ② 反対
- ③ 曖昧
- ④ もっと前後の文が見たい

作文

「効率がよいから全員強制の方がよい」という論旨かな？



以下の文章内容は、「効率の面で全員強制の方がよい」という論旨ですか。

- ① Yes
- ② No
- ③ 曖昧
- ④ もっと前後の文が見たい

どのくらいの長さの文章を一人に読ませるのか。

メリット(?)

- 「現代語」のコーパスなら、実際に大勢の現代人に解析してもらった方がいい。偏見が無く公平で、生きた情報だから。

どんな問題なら適用できるのだろう。

- ある程度機械で処理できる。
- どこが機械に難しいか機械自身知っている。
- 人間ならその問題を容易に解決できる。

+

- 問題を細かく分割することができる。

文章の解析だけではなく生成は？

2. ロボットの知能への応用

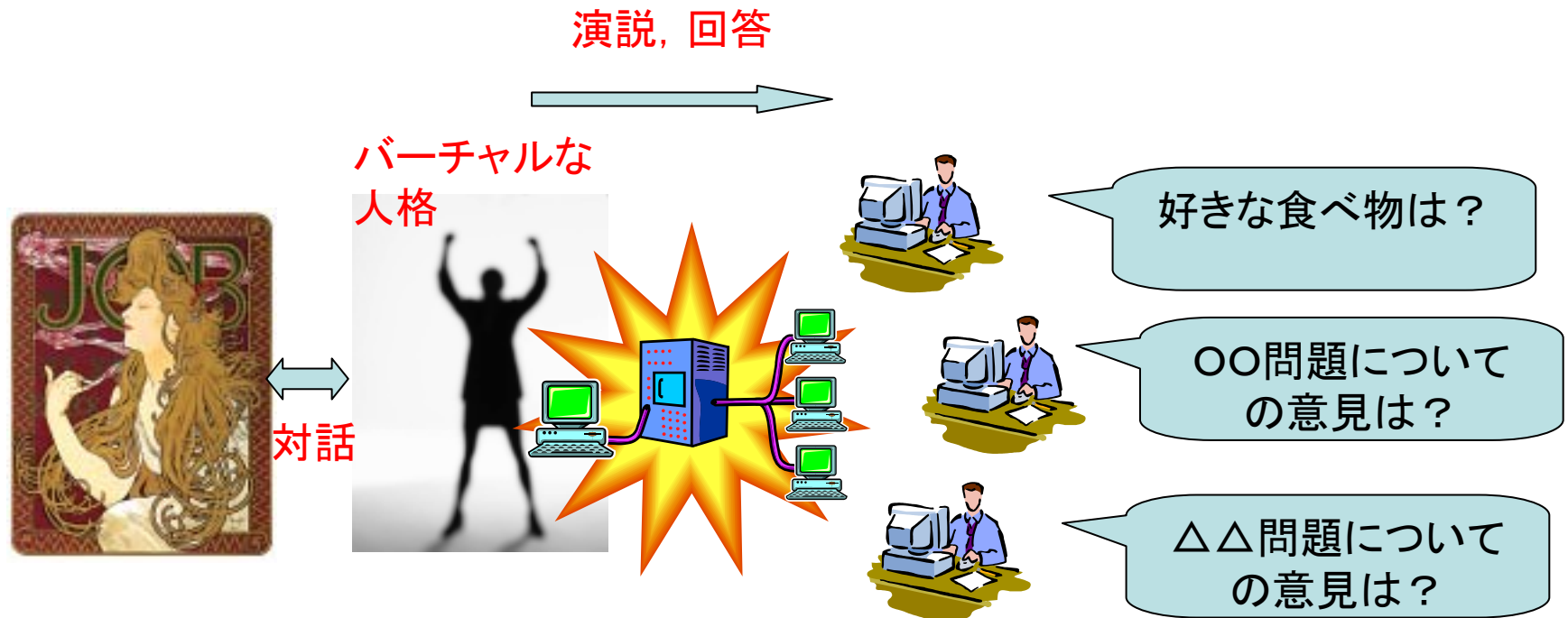
人工知能に対する私見

- ロボットに人格を認めるのは学者ではなく、社会全体.
- 仕組みはどうあれ、たくさん人間がだまされてしまえばそこに人格はある.
- 質問に厳密に答えられなくても、個性の範疇になる.
- 無根拠にいい加減なことを答えるロボットは信用されないが、本システムでは、発言の根拠を背後の人間がもっている.

ロボットに人格・個性を与える



要人と一般市民の対話番組



- 大統領
- 社長
- アイドル

- 市民
- 社員
- 熱烈ファン

このシステムは・・・

- 皆の「総意」を性格にもつ知能ができる。
- 一人一人のサポーターの影響は小さいので、あたかも新しい知能が誕生したように見える

終

システムの人件費

- 【事例】企業がアンケート処理係として一人の人間を1ヶ月雇う
 - 20万円(大卒初任給程度), 160時間(8h/day, 5day/week, 4week)
 - 一人の意志には偏りがある
 - 分散すれば偏りなく並列処理で作業できる
 - 個人情報情報の漏洩は問題. うまく隠すか当たり障りないものを.

単純に報酬を分配したときの労働時間

	月給	労働時間
1人	20万円	160時間
100人	2000円	96分(3.2分/day)
1000人	200円	9.6分

時給
1250円

この辺が片手間の小遣い稼ぎとして妥当か

もはやボランティア

本システムが完成した場合の 仮想運用デモ

出来たらよいと思っていることリスト

- 現代語解析済みコーパス作り
- 文脈マッチング型検索
- 対話システム
- (翻訳システム)

現代語解析済みコーパス作り

- 現在，自然言語処理のための解析済みコーパスは人手により作られている。（憶測）
- 「現代語」のコーパスなら，実際に大勢の現代人に解析してもらった方がいい．公平で，生きた情報だから．
- 自然言語研究のためのコーパスを自然言語処理＋人手で作る（！？）
- 将来的には，人手に頼る部分を減らして行きたいので，システム自身の教育のためにも．

文脈マッチング型検索

- ある文章Aが，ある文章Bに関係しているかどうか調べる.
- 国語の作文問題の添削，苦情メール・アンケート処理，長い文章が「要は賛成なのか，反対なのか」を判定する 等.
- 文脈解析の問題

例：国語の作文試験の添削支援

問. 「中学生のボランティア活動への参加の強制について①賛成②反対のどちらかの立場をとって論じなさい。」

(実際に栃木県の中学生向け模擬試験に出題された問題より)

- 文章内容があらかじめ指定された賛成, 反対と一致しているか.
- 「効率」という観点で論じていると減点.
- 誤字, 脱字
- 起承転結の段落構成か.

(採点基準の指針の例)

対話システム

- リアルタイム性が要求される
- 厳密に答えられなくても、個性の範疇になる
- 無根拠にいい加減なことを答えるロボットは信用されないが、本システムでは、発言の根拠を背後の人間がもっている。

翻訳システム（可能性の示唆のみ）

- 翻訳ページにある「テキスト翻訳」は逐語訳的で性能が悪い
- しかし訳し先の言語を母国語とする人には、大体何が言いたいかが理解できる場合がある
- →ちゃんとした文章に整形してもらおう
- 結局2ヶ国語話せる人が必要になるような気がする。

やはり問題となるプライバシー

- 個人情報をお細切れにして散布するやうなもの.
- 人名を偽名に変換する処理は可能か？

難問：

松井選手，メツじゃマツイ。（AERA地下鉄中刷りより）

【松井でないと成立しないし，そもそも形態素解析が難しい。】

問題点

- サポーターがまじめに働いていることをどうやって判断するか
 - 時々正解のある普通の問題を出し、成績をつけて判断
- 性能の向上の為に
 - 一つの質問を複数の人間に投げかけ、返答の傾向を手がかりにする
 - 一つの知能を皆でサポートしているという意識を盛り上げ、コミュニティを作る.
 - 複数の知能を並行動作・競争させ、(どう判断するかはともかく)成績がよかった場合は配当を出す

サッカーのサポーターと株式市場をお手本に