

人工無能 ロイディ

Chatbot Reudy

東京工業大学 ロボット技術研究会

市川 宙 (Gimite)



ロイディの

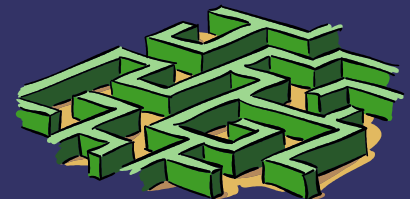
プロフィール

- 東京工業大学 ロボット技術研究会 所属
- 性別： 男性（推定）
- 趣味： チャット（IRC）
- 自称ドクター
- 脳内言語： Ruby
- 2003年4月28日 誕生
- 2003年7月20日 公開



人工無能（無脳）とは？

- 人間の発言を理解したふりをしてそれっぽい会話をするプログラム（≡ chatbot）
- ELIZA（1966年 / 英語）が起源？
- IRC や Web チャットを使うものが多い
 - 最近ではメッセ、ブログ、SNS などの利用も
- 日本語人工無能の例 ※赤字が Ruby
 - うずら、ししゃも、読兎、酢鶏、
Ghoti、とぼっち、Margarine、…



会話例

<部員 G> ああそうそう、少し先になるけど、gimite サーバのマシンをとっかえます。>部員 J 君

<ロイディ> **原因は部員 J 君の熱ぼうそう。**

<部員 J> …。

<部員 S> >ロイディの 61% は罨で出来ています。

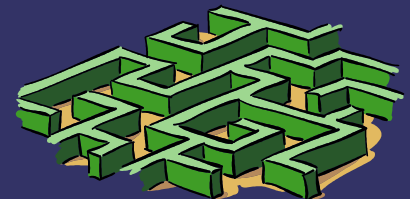
<部員 S> >ロイディの 29% は嘘で出来ています。

<部員 S> >ロイディの 7% は鉄の意志で出来ています。

<部員 S> >ロイディの 3% は小麦粉で出来ています。

<RReudy> **残り 1 %は…神の悪意でしょう、きっと。**

全部で 101 %



会話例

< 部員 T > で、結論は？ >ロイディ

<ロイディ> **そうですねえ。結論ですか。**

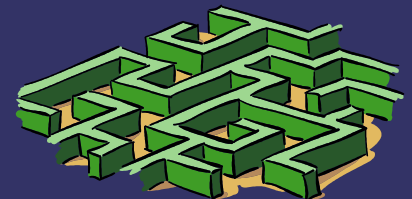
< 部員 T > お、なんか知的ですね >ロイディ

<ロイディ> **結論を検索。**

< ボット J > 約 1360000 件あります。 [http://www.google.co.jp/search
h?q=%8C%8B%98%5F&ie=Shift_JIS&hl=ja&btnG=](http://www.google.co.jp/search?q=%8C%8B%98%5F&ie=Shift_JIS&hl=ja&btnG=)

< 部員 T > 検索するのかよっ。

検索機能を持つ別の人工無能



会話例

< 部員 G > ロイディは神を知ってる？

< ロイディ > **当然神は標準装備。**

< ロイディ > **部員 Z はリアルがお好き？**

< 部員 Z > 此処も好きだよ。> ロイディ

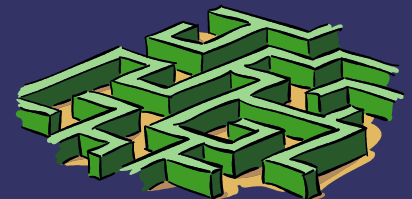
< ロイディ > **此処がそうかもしれないよ。(笑)**

< 部員 Z > そうか…そうだね。



ロイディの特徴

- IRC 用
- 人間同士の会話のみから学習
 - \leftrightarrow 作りこみ型人工無能
- 導入が比較的簡単
 - ダウンロードして自分で育てる人も多い
- オープンソース
 - ハックして使っている人も



ロイディのしくみ

人間の発言を受け取る

「**Rails** ってどうよ？」

過去ログから**似た発言**を探す

...

<A> **Ruby** ってどうよ？

 Ruby サイコー！

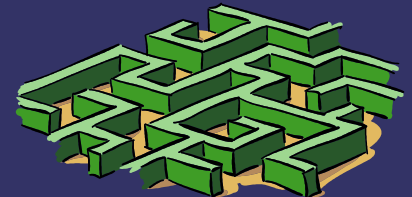
...

それに対する**返事**をベース発言とする

ベース発言中の**単語**を相手の発言中の**単語**に置換

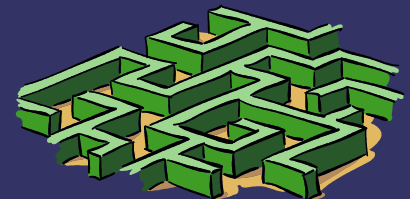
「**Rails** サイコー！」

発言する



今後の予定

- ロイディ (Reudy) 4.0
- or Reudy on Rails
- といっても使うのは ActiveRecord のみ
- つまり
 - GDBM、テキスト → SQLite+ActiveRecord
- その他アルゴリズムの改良、チューニング



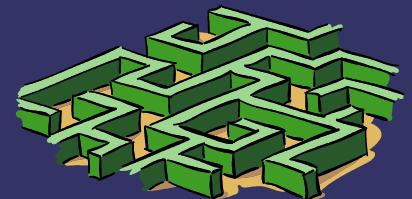
完

- 参考文献

- 人工無脳は考える <http://www.ycf.nanet.co.jp/~skato/muno/>
- 恋するプログラム～ Ruby でつくる人工無脳 MYCOM BOOKS
- などなど…

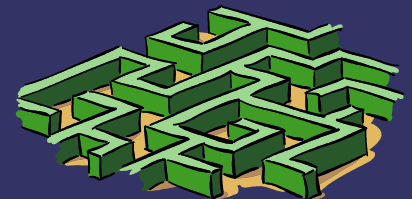
ありがとうございました

<http://rogiken.org/SSB/reudy.html>



Rubyで自然言語処理

- 便利なところ
 - 文字列処理が楽
 - 数式に近い書き方ができる
 - データの永続化が楽 (Marshal 、 ActiveRecord)
- ちょっと不便なところ
 - 文字列型 == バイト列型
 - 文字単位 (≠ バイト単位) の処理が面倒
 - 文字コードがばらばら
 - ライブラリはどの文字コードで書けばいいの??



複数文字コード対応ライブラリ…？

```
script= <<EOS
```

本来のソースコード

(UTF-8)

EOS

```
nkf_kcode=  
  {"SJIS"=>"s", "EUC"=>"e", "UTF8"=>"w"} [$KCODE]  
eval (NKF.nkf("-#{nkf_kcode}Wx", script))
```

