

会話プログラムの方針転換

なんとなくいきづまりを感じていた会話プログラムでしたが、友人に「うずら」を教えてくださいました。

google で「うずら」と「irc」で検索するとひっかかります。

irc 上で会話をするロボットということですが、どうやら、ちゃんと学習するらしい。7年ほど前にはじめられたようで、データもたくさん保持しているようです。羨ましい...さて、会話ログの一部がのっていて、それを見ると、確かに絶妙な返事をしながら会話が成立しているように見えます。

あまりに気になったので、その irc のチャンネルを覗いてみることに。

ずいぶん、よく出来てるとは思いますが、... .. やっぱ難しいですね。

当然のように、形態要素解析だけでは限界がありますので、やはり構文解析をきちんと行なって意味ネットワークを構築しないとイケないのかも

でも、どこまで持てばいいのか、とか

実際に文を生成する際にどの情報(ネットワーク)を利用するのか、とか

に対して妙案は浮んでないのですが。

それよりも、他の人の言われるように、何を重視するかというのが一番大事なのですが。

ところで、irc っていうのはいいアイデアだな~と思います。

やっぱり一対一で話しをしていて相手に何かを覚えてもらうことには無理があると思います。

量の問題も当然ありますが、なによりも、

複数での会話を傍受していれば、なんとなく旨い文章を捻りだすように学習することが可能になる気がします。

そういうわけでちょっと方針転換して、irc のロボットベースの会話プログラムを作っていくことにしようかな、とか

言語も C はやめて Ruby へ移行。

しかし web をいろいろ見て周ると同じようなことをやっている人がいっぱいいる~。

差別化の鍵は、やっぱり構文解析かな ... うさんくさい実装でも ... できるのかな？

ircd

というわけで、自分のマシンの FreeBSD 上で実験のために ircd をあげることに

irc-2.10.3 をコンパイル/インストールするまではよかったです

ircd.conf が書けない。何回やっても文法チェック (chkconf) でおこられる。

error をみると delimitate の記号が : から % に変更されたいらしい。

IPv6 のためらしいのですが、example の記述も変更しておいて欲しかったなあ

なんか、とって時間を費してしまいました。