Overview of the Rational Speech Act Model

Victorien Marette

1 Introduction

The Rational Speech Act Model is a linguistics model introduced by Goodman & Frank in 2012 [1]. This model describes a situation where a speaker knows the state of the word and tries to convey it to a listener using utterances (messages he can send, following a language rules). His goal is to maximize the chance of the listener guessing the right state. In this model, the speaker and the listener recursively reason about each other's mental states to infer the meaning of utterances and generate utterances in response [2].

To define a RSA model, you need to define:
- A set of all states of the word possible named C. In this model, we consider only a finite number of states. We can then describe C as $[1, n]$.
- A set of all the utterances named U. In this model, we consider only a finite number of utterances. We can then describe U as $[1, n']$.
- A probability distribution P on C. It represents the prior beliefs of the speaker on the probability of each state.
- A meaning function $M : U \times C \to \{0, 1\}$. This gives 1 if the utterance can describe the state in the language.
- A cost function $k : U \to \mathbb{R}$. This gives the cost of every utterance for the speaker.
- $\alpha \in [0, +\infty[$, this will define the speed at which the Speaker adapts.
The notations used here are from the article from Yuan, Arianna, Will Monroe, Yunru Bai and Nate Kushman [2] as it is a more recent and comprehensive take on the model.

From those parameters, we can build two sequences of sets of probability distribution, L and S. L_n is a sequences of sets of probability distribution on C indexed by U. S_n is a sequences of sets of probability distribution on U indexed by C.
To simplify notations, for $u \in U$ and $t \in C$, we will write $(L_n)_u(\{t\})$ as $L_n(t|u)$ and $(S_n)_t(\{u\})$ as $S_n(u|t)$.
L_n represents the mental state of the listener after having inferred n time on the mental state of the speaker. S_n represents the mental state of the speaker after having inferred n+1 time on the mental state of the listener.
$L_n(t|u)$ is the probability that given utterance u, the Listener will guess the state of the world t. $S_n(u|t)$ is the probability that given the state of the world t, the Speaker will speak utterance u.

Formally we define them as:

$$\forall u \in U, \forall t \in C, \begin{cases} L_0(t|u) = \frac{M(u,t)P(t)}{\sum_{t' \in C} M(u,t')P(t')} & \\ L_n(t|u) = \frac{S_{n-1}(u|t)P(t)}{\sum_{t' \in C} S_{n-1}(u|t')P(t')} & , \forall n \in \mathbb{N}^* \\ S_n(u|t) = \frac{\exp(\alpha(\ln(L_n(t|u)) - k(u)))}{\sum_{u' \in U} \exp(\alpha(\ln(L_n(t|u')) - k(u')))} & , \forall n \in \mathbb{N} \end{cases}$$

Note that is possible to do this model by initializing S_0 first.

The Listener is Bayesian. By definition, the formula for his beliefs is the fallowing:

$L_n(t|u) = \frac{P(u|t)P(t)}{P(u)}$.

For listener n, $P(u|t) = \begin{cases} S_{n-1}(u|t) & \text{if } n \neq 0 \\ M(u,t) & \text{if } n = 0 \end{cases}$

With the law of total probability we know that $P(u) = \sum_{t' \in C} P(u|t')P(t')$.
Combining those two fact we get the formulas for the listener.
The formula for the Speaker is a softmax function on the utility $U(u,t) = \ln(L_n(t|u)) - k(u)$. This makes it more likely to convey the utterance that maximize U. The bigger the α the more likely he is to take this utterance.

2 RSA simulation

In this part, we will try recreating results from "Understanding the Rational Speech Act model."[2] article.

As U and C are finite we can represent S_n and L_n as matrixes where $(MSpeaker_n)_{(i,j)} = S_n(i|j)$ and $(MListener_n)_{(i,j)} = L_n(j|i)$

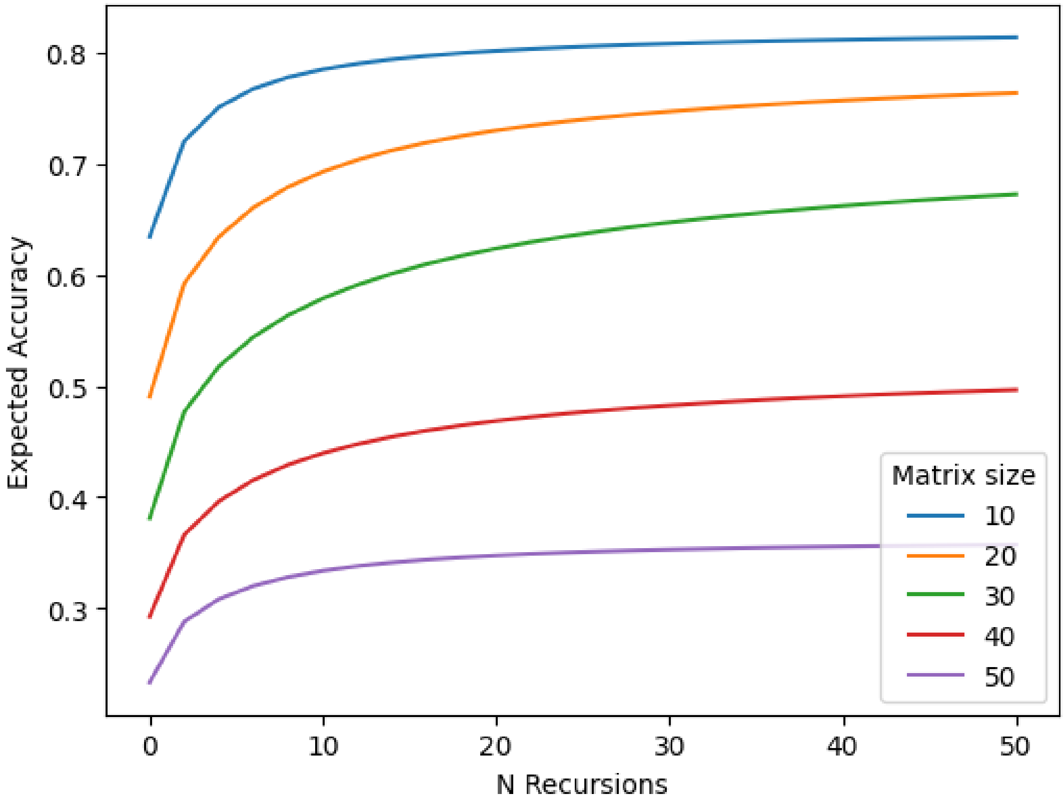
The simulation will be about expected accuracy. It represents the likelihood of the listener to guess right the state of the world after n recursions. In the case where the prior belief probability function is uniform and there are q utterances and world states, expected accuracy is given by this formula: $\frac{1}{q}tr(MListener_{n-1}^T \times MSpeaker_n)$. Proof of this formula is in "Understanding the Rational Speech Act model."[2].

We want to see how the number of recursions and the size of the matrix, affect the expected accuracy. To do this, we generate 200 meaning matrixes with 0 and 1, for each matrix size we choose, here {10, 20, 30, 40, 50}. Then for each meaning matrix we simulate recursion to see how the expected accuracy evolves.

The results (Figure 1) indicate that recursions improves the expected accuracy and does it in very few recursion. In short, recursions can quickly make it clearer which state it is. We can also see that the more utterances and states the less the expected accuracy. This result supports the results of "Understanding the Rational Speech Act model."[2].

The code of the simulation is here: CODE

Figure 1: Results of the simulation



References

[1] Michael C. Frank, Noah D. Goodman. 2012. "Predicting Pragmatic Reasoning in Language Games" *Science* 336

[2] Yuan, Arianna, Will Monroe, Yunru Bai and Nate Kushman. 2018. "Understanding the Rational Speech Act model." Cognitive Science