We created the RSM model, which uses the words in the human utterance as inputs in a bag of words (TF*IDF and a mix of uni-grams and bi-grams). With the bag of words, a Support Vector Machine (SVM) with a linear kernel predicts the next word in the machine's response. It then adds the predicted word into the bag of words in order to predict the next word. The process is repeated until the <EOS> tag is predicted, indicating the end of the machine's response.

A purely statistical and recurrent model for building conversations

Data:
1. Data set 1: a small portion of the Friends script (data set 1)
2. Data set 2: a script of both Season 1 and 2 together (data set 2)
3. Data set 3: a script with Seasons 1, 2 and 3 (data set 3).

Logic Model
This model translates human utterances into first-order logic formulas and enforces logic rules in order to draw inferences and responses from the facts provided by the human.

N-gram Search
Given a starting n-gram, this model finds the next n-gram that is most likely to follow the previous n-gram in the sequence until the response reaches a certain predetermined length.

Comparison RSM with the top chatbots available (data set 3)
RSM
CleverBot
ALICE
Siri

How RSM improves as the data set increases

GCT Scores broken down into conversation topics (using data set 2)

General Conversation
Philosophical Questions
Feelings
Closing

Applying RSM with different models (using data set 2)

Multinomial Naive Bayes
Logistic Regression
Linear SVM

Granular Conversation Test (0-4):
We developed our own test that breaks down the conversation into 4 components:
1. Semantics
2. Structure
3. Context
4. Feeling