**Intelligent virtual assistant challenge**

**What's in it for me?**

Solve a real world problem, develop your skills, have fun, enhance your resume, meet students from across the Atlantic and earn prizes worth $4,000!

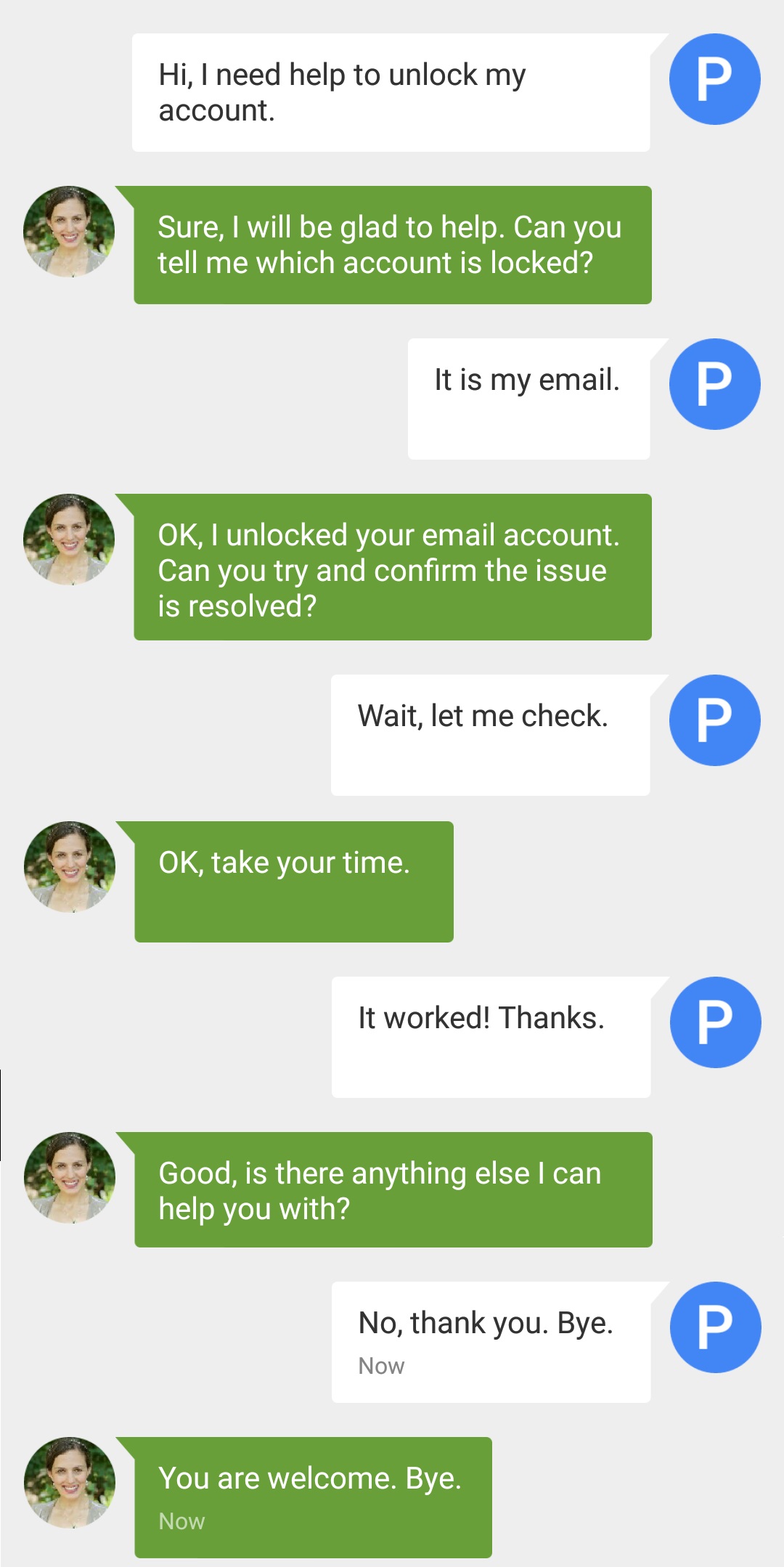
****

**Background and Overview**

Cognitive computing is an exciting technology space. One of the key aspects of cognitive computing is the ability to diagnose what the question or problem is through an interactive conversation and provide guidance on the solution. In the real business world, call centers receive a large volume of calls due to users abandoning existing self-service options. Even when users are presented step-by-step instructions for resolving their issues, they still may have questions on a particular step to help guide them through the solution. This guided assistance within self-service is the goal and it will drive greater success of the system, reducing call volume in our call centers.

**Examples of interactive dialogue**

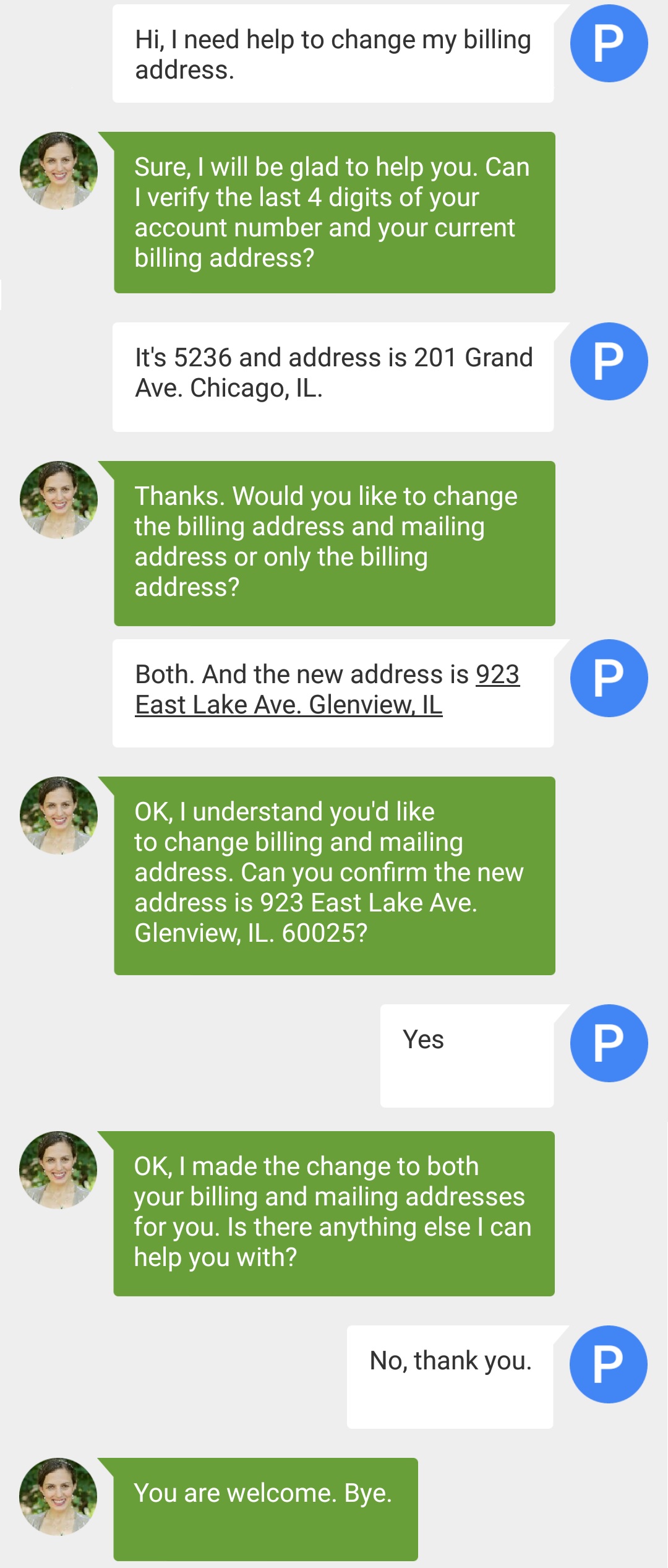
The purpose of following interactive conversation example is to show how this solution could be used in the real world. It is not intent to be the solution specification or exact dialogues.

Example 1. Unlock the user’s email account: (Assume the user has been already authenticated)

Provide guidance to resolve the user’s problem

Asking clarification question and to derive a clear understanding of the user intent

Example 2. Change the user’s address:



Understand the word “both” and associate to context of “billing and mailing addresses”

Gain new domain knowledge; follow a process to verify client account and billing address

**About this challenge**:

The challenge is to leverage natural language processing, ontology-based semantic processing and machine-learning techniques to develop an intelligent virtual assistant that can perform following functions in a basic text conversation. The ideal solution should demonstrate the following capabilities:

1. Understand what type of help the user needs by asking specific clarification questions, evaluate the user’s response and derive a clear understanding of the user’s intent
2. Offer step-by-step solution and guidance to resolve the user’s problem
3. Accept new domain knowledge as input, update the ontology through machine-learning techniques and respond to a new request from the user on that topic

**This challenge is not about**:

* This is not about using the best statistical model or algorithm that is normally measured by the accuracy from the training data set to the test data set. Machine-learning techniques and statistical models can improve the solution in many ways, but it is not required.
* This is not about using ontology-based semantic modeling standards like RDF, WOL, or SPARQL. But it is highly recommended to consider these standards and use them if possible
* This is not about advanced voice-to-text, a human-like avatar, or web application development. A basic command line interface that can simulate a text conversation is good enough.

**Rules**

* Participants may work individually, but we highly recommend working as a team. Team members should not be more than 4 people
* The solution must use natural language processing capability
* The solution must use ontology modeling capability
* The solution must not use any paid license tool, library or data set
* Participants must describe the best use of natural language processing and their ontology model for each specific domain topic
* Participants should use machine-learning techniques to improve the quality of the solution
* Each participating team or individual may submit only one entry.
* Entries must represent original work that has not been previously published or submitted to other challenges.
* Each participating team must send at least one member to present their solution on demo day
* Participants are encouraged to use publicly available data sets, public pre-trained model, SDK, or ontology definitions. Participants are not required to make these data public; however, they are encouraged to do so if possible.
* A team of Allstate representatives will evaluate all entries based on the judging criteria.

**Judging criteria**

The solution must include all four deliverables:

1. Overview and explanation of the solution approach in presentation format; up to 6 slides. Keep in mind “a picture is worth a thousand words”
2. Working computer program in Python (or Java) able to simulate the user and intelligent virtual assistant in a text conversation. Command line interface example will be given in both Python and Java, with “your code goes here” placeholder.
3. Working computer program in Python (or Java), ontology model and machine training data to educate the intelligent virtual assistant specific domain knowledge. (to-do: Two knowledge article will be provided before the kick-off. Participants can re-engineering it to fit their solution)
4. A demonstration of your solution that shows the effortless way to train the intelligent virtual assistant and gain new domain knowledge from your own simulation data sets; then handle different type of assistant tasks.
   1. In your demo, you may, for example, show the initial version of the intelligent virtual assistant can only handle unlocking a user’s email account; but, with additional training, now it change the user address information as well.

**Scoring**:

Max score 100 points

* 20 points - Completeness of all four deliverables. Are they completed? Do they work?
* 10 points – Quality of solution presentation and demo (10 min overview, 10 min demo, 10 min Q&A)
* 15 points – How well does your application understand what type of problem the user has or needs help with? We are looking for succinct clarifying questions gathering the data needed to determine the problem.
* 15 points – How well does your application step the user through the solution to resolve their problem? We are looking for succinct steps with confirmation to ensure the user is completing the tasks.
* 40 points – How easy is it to train your intelligent virtual assistant to gain new domain knowledge, such as follow a process and interactive dialogue?
  + How well does it learn a new conversation flow from your own simulation data sets (15 points)?
  + How well does it learn step-by-step tasks to guide the user through dialogue from your own simulation data sets (25 points)?

**Specific virtual assistant domain knowledge area use cases suggestion:**

* Unlock the user’s email account
* Change the user’s address information

(Above domain knowledge human procedure docs will be provided before the kick-off, it can be used for semantic reasoning and ontology modeling or creating training data set)

**Suggest prerequisites:**

1. Basic Python or Java programming skills
2. Basic understanding of natural language processing. Recommended source:
   * <http://www.nltk.org/>
   * <http://nlp.stanford.edu/software/>
   * <https://opennlp.apache.org/>
3. Basic understanding of semantic-based ontology. Recommended source:
   * <https://www.w3.org/2001/sw/wiki/SPARQL>
   * <https://jena.apache.org/tutorials/sparql.html>
4. Basic understanding of machine-learning techniques:

* <https://www.udacity.com/course/intro-to-machine-learning--ud120>

1. Know how to use <https://github.com/>
2. Open Source license we recommend you use:

* Apache License 2.0
* BSD 3-Clause "New" or "Revised" license
* BSD 2-Clause "Simplified" or "FreeBSD" license
* GNU General Public License (GPL)
* GNU Library or "Lesser" General Public License (LGPL)
* MIT license
* Mozilla Public License 2.0
* Common Development and Distribution License
* Eclipse Public License