Universidad del País Vasco

Euskal Herriko Unibertsitatea

Conversational Agents

ATAI

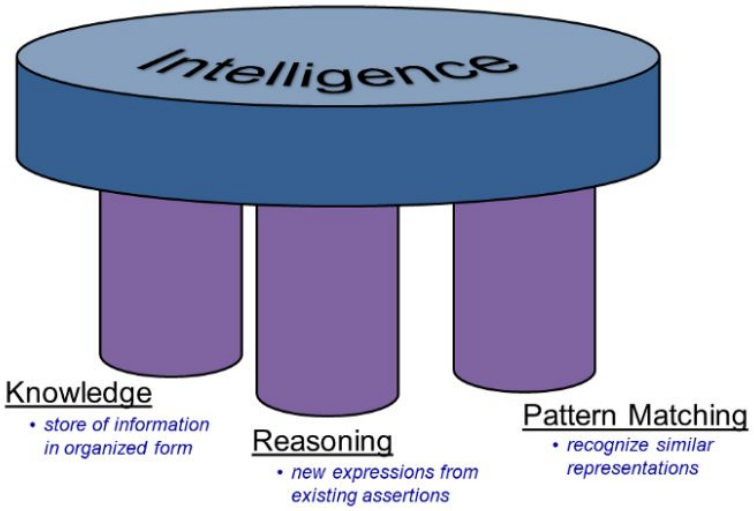
Xabier de Domingo Yesa
Mikel González Saiz
Jon Pérez Etxebarria

Pillars of AI

As to establish a baseline concerning what artificial intelligence is, we should explain what intelligence on its own. Take for one Knowledge, someone could know as many facts as you would like but if they are unable to connect any of them together we wouldn't consider that being to be intelligent based on that knowledge alone.

On the other hand anyone can tell an intelligent being needs to be able to reason, but of what use is reason if you have no facts to reason with. At this point it should be apparent that intelligence comes from a combination of qualities.

In the field of cognitive science intelligence is decomposed in a particular way, we can call this the pillar of intelligence.



Two of this we have already mentioned broadly but I will attempt to give a small definition for each of this.

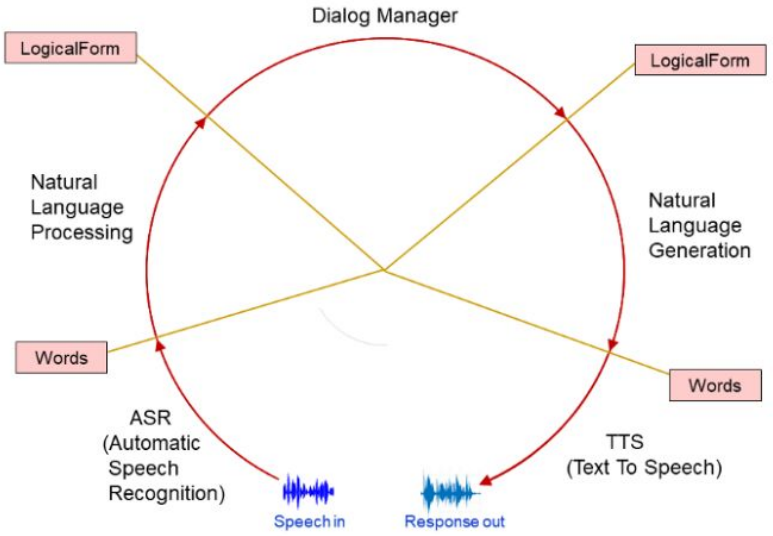
Knowledge: facts, data, skills, procedures and beliefs organized in such a fashion they can be accessed when needed.

Pattern Matching: the ability to generalize across specific cues and instances of data.

Reasoning: given some assertions and knowledge to derive new assertions.

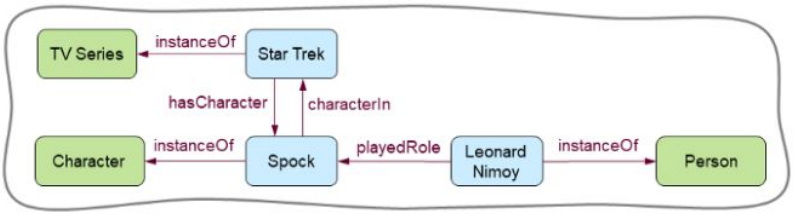
Architecture of a Conversational Agent

Once we have a working definition of intelligence we can have a look at how a conversational agent works.



We can observe there are many steps from the intake of information from the human until he derives an appropriate answer. The first step is speech recognition, a common way to do this is the use of recurrent neural networks, this will convert the waves of sound into words. Then a natural language processing module will need to interpret your words and convert them into a logical form.The dialog manager will take that logical form and decide what the answer should be, for example searching for your question and getting the answer in logical form. The second to last step is for a natural language generation unit to convert that logical form into language we are more used to. And the last will get those words and use TTS(Text to speech) output sounding words.

One of the cutting edge technologies in intelligent conversational agents lies in Natural language processing and the dialog manager modules. The collections of facts are represented in knowledge graphs. Below we can see an example.



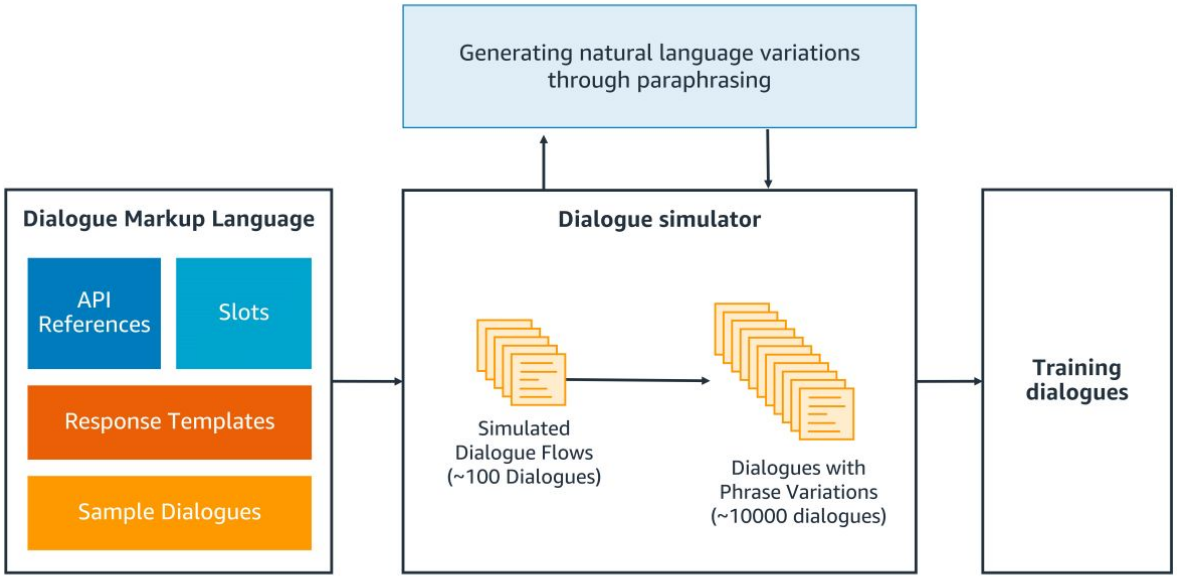
A graph consists of nodes. In a knowledge graph the nodes represent entities and the links represent relations between these entities. These can be constructed based on different rules. In these cases we have colored in green entities or classes and tokens in blue. Dozens of knowledge graphs have been constructed over the years according to an array of more or less tightly constrained knowledge ontologies

Alexa Conversations

Enables customers to interact with Alexa in a natural and conversational manner. It's powered by two major structures: a Dialogue simulator for data augmentation and a Conversation modeling architecture to train deep-learning based models to improve the conversations beyond the sample dialogues.

Dialogue Simulator

Building high performing deep-learning models requires a lot of data, which are usually hard to find, and that’s the purpose of the Dialogue simulator. It automatically generates sample dialogues and even dialogues that are hard or not so common to occur.



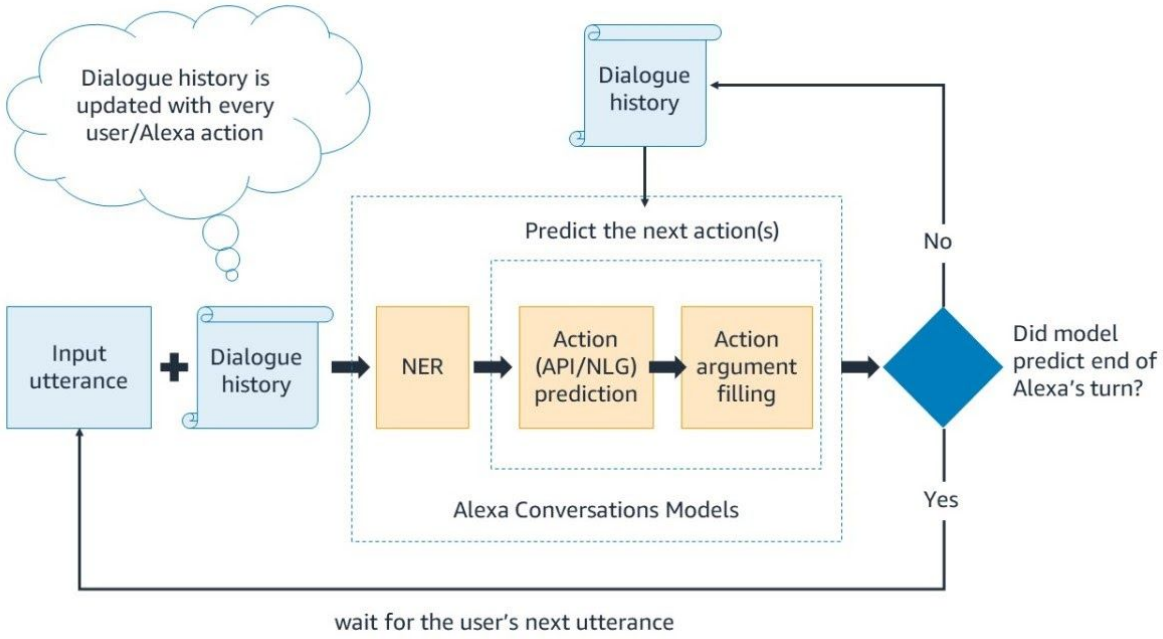
The inputs for this model are APIs, slots and associated catalogues for slot values and response templates, and with that it generates dialogues in two steps.

In the first one, the simulator generates dialogue variations that represent different paths a conversation can take using two agent-simulators one representing the customer and the other for Alexa.

The second step is injection of language variations into the dialogue path, like alternate expressions for the same things. Some of these are generated by the response templates while others just paraphrasing.

Modeling Architecture

A natural conversational experience could follow any one of a wide range of dialogue patterns, so does the Conversation modeling architecture improving dialogue simulators and components offering proactively related functionalities as well as a robust support if a conversation changes its topic midway.



The modeling architecture is separated into three different models: Named Entity Recognition model (NER), Action Prediction model (AP), and the Argument Filling model (AF).

The models are built by combining supervised training techniques on the training data provided by the Dialogue simulator.

The NER model identifies slots in each utterance, it's sequence-tagging model.

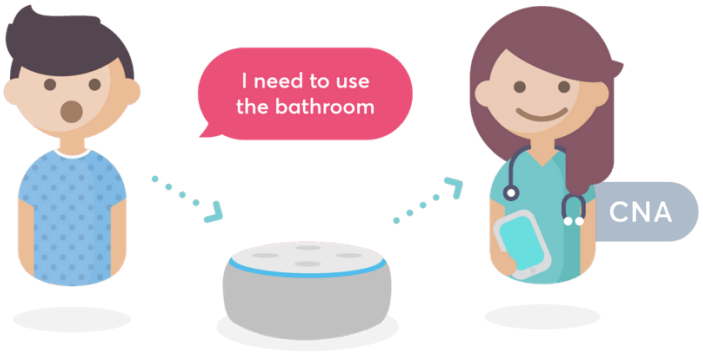
The AP model predicts the optimal next action for Alexa to take, such as calling an API or responding explicit information.

The AF model fills in the argument values for the API and response templates by looking at the whole conversation looking for context.

Usage of Conversational Agents

Talking and controlling smart devices aren't the only usage for conversational agents. Over the last years, the conversational agents have been used in many other areas, such as health care and learning.

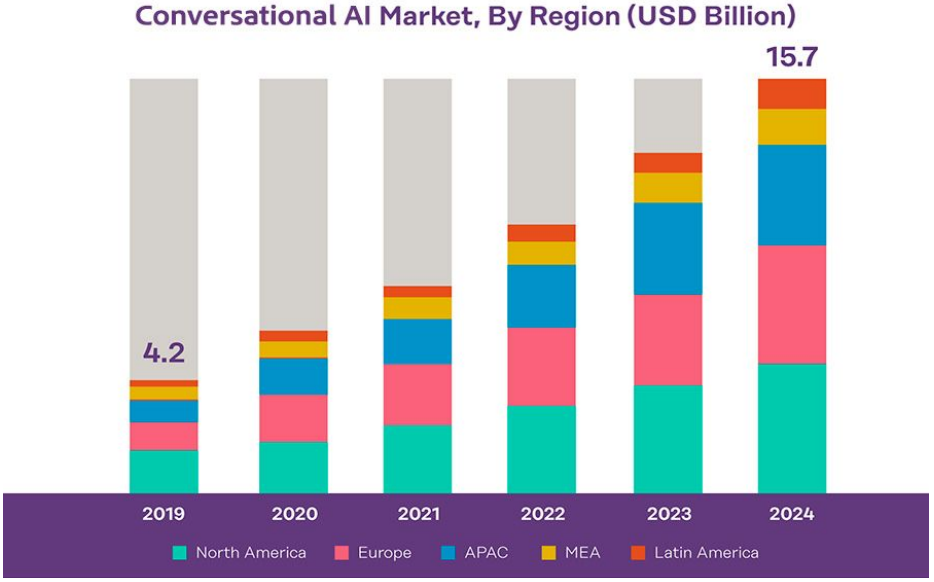
The conversational agents can be used in health care for attending to the patients, asking them for their requests, and then, notifying the nurses. This can help the employees to prioritise a request over another, so the urgent or the important ones can be attended first.



Moreover, mental health services could also be an area in which conversational agents could be useful. A digital therapy could provide an interactivity that emulates therapeutic conversation, and also choice and control over session content and intensity. However, existing digital therapies have suffered from some problems, like a low adherence and some concerns about the efficiency without human support.

Learning could also be improved with the usage of conversational agents. These tools could not just teach the student, but also learn from him, motivating the student so he can see how the conversational agent got smarter because of his teaching.

Nevertheless, these usage are not common these days. As this market is predicted to grow from $4.2 billion in 2019 to $15.7 billion by 2024, the conversational agents are expected to improve in the next few years, so they can fulfil these tasks much better than today.



Bibliography

Pillars of AI:
How Do Conversational Agents Answer Questions? | by Eric Saund

Alexa Conversations:
Science innovations power Alexa Conversations dialogue management

Usage of Conversational Agents:
Intelligent Conversational Agents in Healthcare: Hype or Hope?
Towards an Artificially Empathic Conversational Agent for Mental Health Applications: System Design and User Perceptions
Conversational Agents in the Treatment of Mental Health Problems: Mixed-Method Systematic Review
Aiva Health: Virtual Health Assistant
Using Conversational Agents To Support Learning By Teaching
Conversational AI Statistics: State of Chatbots in 2020