



Are We Conversational Yet?
A Design Study And Empirical Evaluation of
Multi-Turn Dialogues For Virtual Assistants

CS 294S/W Research Project Proposal

Alejandrina G.R., Kat McNeill



What will virtual assistants look like in ~5 years?





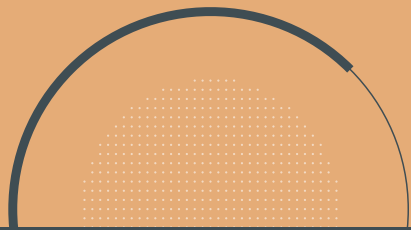
What are current virtual assistants lacking?





Are virtual assistants actually conversational?





Not conversational yet.

And we don't really know if people want dialogues to begin with.



Motivation

User-centered 🧑💻

Infrastructure already built 🙌

Multi-turn dialogues with contextual neural network 👁️👁️

Dialogues = next big thing for virtual assistants? 🤖

Bottleneck for Almond 😞

Definitions

Dialogue: well-formed sequence of turns

State: formal representation of the dialogue, up to a certain point

User state: right after the user speaks

Result state: right after execution Agent state: right after the agent speaks

Abstract State: family of states, as defined by the dialogue state machine

Current state

30 state transactional models.

All transitions are now built with artificial dataset.

Built in a uniform way.

Big claim: 30 transactions cover all the domains.

State-Machine-Based Dialogue Agents with Few-Shot Contextual Semantic Parsers

Giovanni Campagna Sina J. Semnani Ryan Kearns Lucas Jun Koba Sato Monica S. Lam

Computer Science Department
Stanford University
Stanford, CA, USA

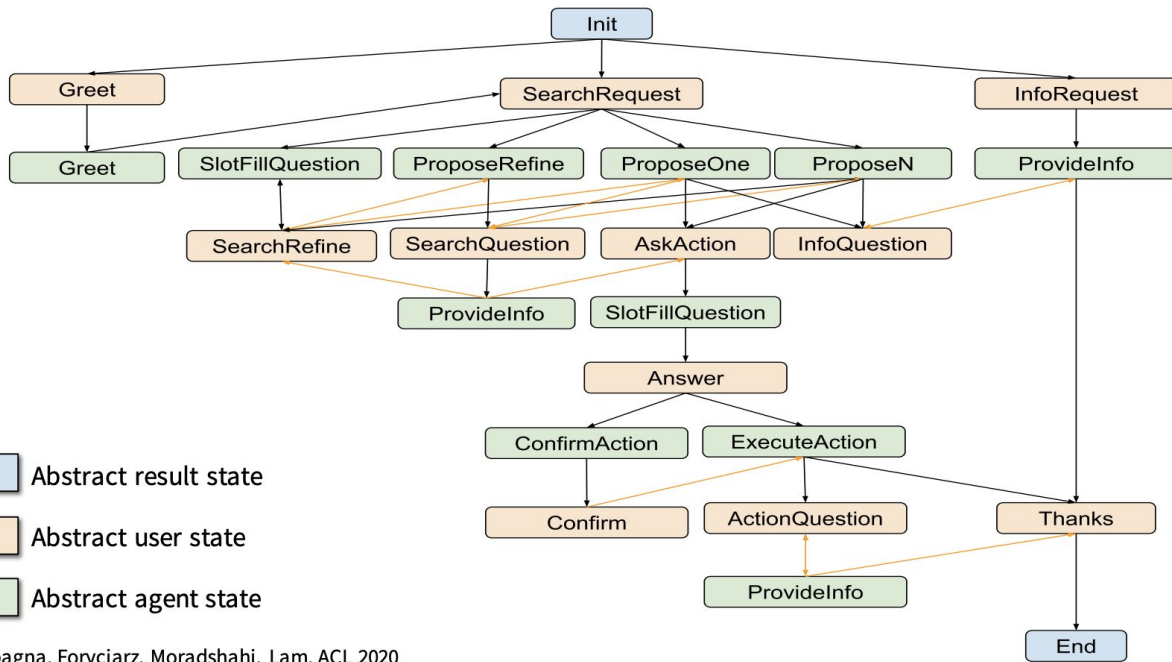
{gcampagn, sinaj, kearns, satojk, lam}@cs.stanford.edu

Abstract

This paper presents a methodology and toolkit for creating a rule-based multi-domain conversational agent for transactions from (1) language annotations of the domains' database schemas and APIs and (2) a couple of hun-

of the tree represents an agent utterance and a small set of *intents* capturing the anticipated user responses (Gao et al., 2018).

Wizard-of-Oz conversations (Kelley, 1984) have traditionally been used to study dialogue state tracking. We realize that they can be put to use to create



- Abstract result state
- Abstract user state
- Abstract agent state

Campagna, Foryciarz, Moradshahi, Lam, ACL 2020

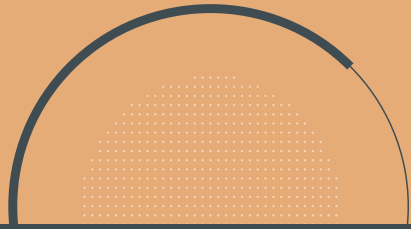
State-Machine-Based Dialogue Agents with Few-Shot Contextual Semantic Parsers

Giovanni Campagna Sina J. Semnani Ryan Kearns Lucas Jun Koba Sato Monica S. Lam
 Computer Science Department
 Stanford University
 Stanford, CA, USA
 {gcampagn, sinaj, kearns, satojk, lam}@cs.stanford.edu

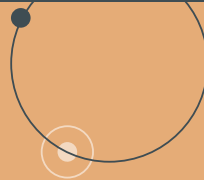
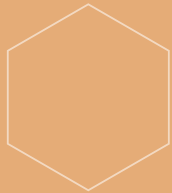
Abstract

This paper presents a methodology and toolkit for creating a rule-based multi-domain conversational agent for transactions from (1) lex

of the tree represents an agent utterance and a small set of *intents* capturing the anticipated user responses (Gao et al., 2018). Wizard-of-Oz conversations (Kelley, 1984) have



Unhappy paths:
How to deal with things we have never seen before?



Unhappy Paths

How will the user respond to unexpected answers?

Can we anticipate how people might do multi turn dialogues?

The first sentence that is not understood the agent is already off track.

Something we have never modeled.

..... **To Learn:**

How do we minimize unexpected answers? Can we crowdsource at scale?

What methodology is best to identify the abstract dialogue acts in unhappy paths?

How do human agent transcripts compare with AI agent transcripts? Can we use a "backoff" scheme?

Can we assume that language variations with the same intent can be handled automatically?

What are the different things that people want to say?



Timeline

01.

Test Almond to get an intuitive feel

02.

Run a small-scale formative study to gauge user responses.
(Use GTP-2 Dialogue?)

03.

Design a crowdsourcing experiment for a small domain



04.

If good results from 3, improve success ratio. Else try another experiment.

05.

Expand dialogue capabilities and iterate.

06.

Final demo/presentation and writeup.



Demo

Results from studies

Expanded dialogue capabilities

Aiming for 50 dialogue transitions (now we have 30)

Sources

- State-Machine-Based Dialogue Agents with Few-Shot Contextual Semantic Parsers.
<https://arxiv.org/pdf/2009.07968.pdf>
- A Formal Language & Data Engineering Approach to Multi-Turn Dialogues
<https://web.stanford.edu/class/cs294s/slides/dialogue-cs294s-2020fall.pdf>
- Genie: A Generator of Natural Language Semantic Parsers for Virtual Assistant Commands
<https://almond-static.stanford.edu/papers/genie-pldi19.pdf>

The top header area features several decorative geometric elements: a large orange circle with a grid of small white dots inside; a white hexagon outline; a grid of small white dots; a white circle with a smaller white circle inside and a small orange dot on its circumference; a white triangle outline with a small orange hexagon inside; and a small orange circle on the left side.

Questions?