

Evaluating Use Cases Suitability for Conversational User Interfaces

Pedro Ferreira and André Vasconcelos

INESC-ID, Instituto Superior Técnico, Avenida Rovisco Pais 1, Lisbon, Portugal

Keywords: Use Cases, Chatbot, Conversational User Interface, Natural Language Understanding and Dialog Systems.

Abstract: The developments in Natural Language Understanding (NLU) are enabling tasks that were typically performed interacting with humans to be now performed interacting with dialog systems, using the same natural language. Dialog systems can also be used in alternative to more traditional graphic user interface (GUI) applications. A review of the intrinsic differences and benefits of humans interacting with dialog systems in alternative to other humans or GUI applications is performed. It is also reviewed the types of use cases that are now being performed by chatbots. This paper aims to identify the factors that influence the selection of use cases suitable for conversational user interfaces, enabling organizations to make more informed decisions regarding chatbots implementations. The factors identified are grouped in three categories: (i) general factors, (ii) factors to be considered to implement a chatbot over a human operator; and (iii) factors that should be considered when implementing a chatbot over a traditional GUI application. Finally, an assessment to the scheduling a medical appointment use case is performed, using the defined factors. This use case is considered suitable to a conversational user interface according to the factors.

1 INTRODUCTION

The developments in the artificial intelligence field have been responsible to the rise of new, more intelligent systems. Specifically, the developments in the Natural Language Processing, drive the development of chatbots. Chatbots are systems that interact with the user using natural language, as if the user were talking with another human. Today people are using chat platforms as one of the main channels of communication, using applications such as Facebook Messenger or WhatsApp. The heavy usage of chat platforms allied with the developments in NLP create a favourable scenario to organizations to offer their services using conversational user interfaces. Services can be accessed directly from the chat platforms the users already use, in a more natural way, instead of requiring the users to install a specific app or access the organization website.

The goal of this paper is to create an evaluation tool that can be used by any business to evaluate the suitability of a use case to be implemented in a chatbot.

2 BACKGROUND AND RELATED WORK

2.1 Dialog Systems

Conversational agents or *dialog systems* are programs that communicate with users in natural language. This kind of systems can be classified in two categories (Jurafsky and Martin, 2008):

- **Task-oriented Dialog Agents:** are designed for a particular task and set up to have short conversations to get information from the user to help complete the task. These include the digital assistants that can give travel directions, control home appliances, find restaurants, or help make phone calls or send texts.
- **Chatbots:** Chatbots are systems designed for extended conversations, set up to mimic the unstructured conversational characteristic of human-human interaction, rather than focused on a particular task. These systems often have an entertainment value. Chatbots are also often attempts to pass the Turing test. Chatbots can also have some practical uses such as testing theories of psychological counselling.

The word “chatbot” is often used in the media and in industry as a synonym for conversational agent

(Jurafsky and Martin, 2008). In this paper the term “chatbot” is used in that same more general sense. In reality, the kind of systems explored in this paper are typically task-oriented dialog agents, even though we may refer them using the word “chatbot” instead of “task-oriented dialog agent”.

It is also important to notice that even though dialog systems communicate with users in natural language, other form of GUI elements are often used such as predefined quick replies that the user can click in order to make the interaction faster and easier.

2.2 Natural Language Processing

Natural language processing (NLP) is a subfield of computer science concerned with using computational techniques to learn, understand, and produce human language content (Hirschberg and Manning, 2015). Some applications of NLP include: *information extraction*, transforming unstructured data found in texts into structured data (Jurafsky and Martin, 2008); *conversational agents*, that aid human-machine communication (Hirschberg and Manning, 2015); or *machine translation*, the use of computers to automate the process of translating from one language to another, aiding human-human communication (Hirschberg and Manning, 2015) (Jurafsky and Martin, 2008).

The factors that have allowed the development of NLP in the last years twenty years, according to (Hirschberg and Manning, 2015), are: (i) increase in computing power, (ii) the availability of large amounts of linguistic data, (iii) the development of successful machine learning methods, and (iv) a richer understanding of the structure of human language and its deployment in social context.

2.2.1 Natural Language Understanding in Dialog Systems

There are various possible structures to represent the meaning of linguistic expressions. Modern task-based dialog systems are based on a domain ontology, a knowledge structure representing the kinds of intentions the system can extract from user sentences (Jurafsky and Martin, 2018). The ontology defines a frame-based representation, with one or more frames, each a collection of slots, and defines the values that each slot can take.

Dialog agents typically have a natural language understanding module. NLU is responsible for the semantic parsing of user utterance, i.e., it gives semantic meaning to user utterances. This module is responsible for selecting the appropriate frames and

filling the slots of the beforementioned domain ontology structure. This module objective is therefore to extract three things from the user’s utterance (Jurafsky and Martin, 2018):

- **Domain Classification:** if the systems is not single-domain, there is the need to determine what domain is the user referring to.
- **Intent Determination:** what general task or goal is the user trying to accomplish. For example, the task could be to Find a Movie, or Show a Flight, or Remove a Calendar Appointment.
- **Slot Filling:** extract the particular slots and fillers that the user intends the system to understand from their utterance with respect to their intent.

Consider the sentence “Book me a table for two for Friday night at Sushi Place”. The NLU module would recognize the domain as “restaurant”; the intent as “book table” and would fill the time slots with “night” and “Friday”; the restaurant name slot as “Sushi Place”; and finally, the slot for the number of seats as “two”.

The domain and intent determination are usually treated as a semantic utterance classification (SUC) problem and the slot filling as a sequence labelling problem (Zhang and Wang, 2016).

Possible methods used by for domain/intent recognition and slot filling include: (i) hand-written rules; (ii) semantic grammars, that are context-free semantic grammar in which the left-hand side of each rule corresponds to the slot names; and (iii) supervised machine learning, using a training set that associates each sentence with the semantics, we can train a classifier to map from sentences to intents and domains, and for slot filling a sequence model can be used (Jurafsky and Martin, 2018).

Training machine learning models requires having access to rare expertise, large datasets, and complex tools, which presents a barrier to smaller companies (Raman and Tok, 2018). The availability of NLU services in the cloud has powered the widespread use of chatbots.

2.3 Uses of Chatbots

There are several tasks that can be performed by chatbots. This set of tasks make possible a panoply of use cases that can be supported by bot interaction. The main tasks performed by a chatbot are: send alerts; take action; retrieve information and answer questions.

It is possible to identify some categories of uses cases that are already being implemented by some companies taking advantage of the previously

mentioned set of tasks. (Shevat, 2017) identifies the following major use cases.

2.3.1 Conversational Commerce

Conversational bots offer the ability to order and explore services or goods directly by a conversational interface. This has the advantage of removing the need to install the specific app of the business, calling services directly from the preferred conversational channel.

2.3.2 Bots for Business

Bots are not only being used for external customers but also internally by businesses, in order to support their internal business processes. This can improve productivity by facilitating short, contextual and actionable tasks.

2.3.3 Notification Bots

Notification through chat platform are being used as an alternative to more traditional notifications such as email. One of the advantages of this type of use is that traditional notifications usually redirect the user to another platform, while chatbots, taking advantage of its “take action” capability, can perform some related action directly from the conversational UI. For instance, a client can receive some appointment reminder, and can confirm/cancel it directly from the chat platform.

2.3.4 Bots as Routers between Humans

Bots can also help connecting humans to humans. By using chatbots we can by identifying the intent of the user and connect him or her to the most suitable human. This can act as a replacement of the traditional IVR systems, providing a more natural experience until a human takes the place.

2.3.5 Customer Service and FAQ Bots

Bots helping answering questions to clients are a very common use case. The fact that most questions are usually asked several times, and the answer is standard, make bots a great way to replace humans in this repetitive task, reducing costs and usually being faster.

2.3.6 Productivity and Coaching

There are also various examples of bots focused on reminders, to-do lists, personal and team task management completion, that help people be more

productive. Bots can also be used for personal coaching to assist in areas such as weight loss and finance. By providing a more personal experience, bot users are more willing to provide information to the bot than to fill forms in an app.

2.3.7 Third-Party Integration Bots

Third party integrations make possible to bring external apps used in someone workflow to some chat app like slack. This avoids that users have to context-switch between apps to gather the information needed to their workflow.

2.3.8 Games and Entertainment Bots

Bots are also being used in the entertainment area by using a conversation as a fun activity. One of the advantages of using bots in this area, is that bots can reengage with the users and encourage them back to the service in a less intrusive and more customizable and friendly way than app notifications, for example.

2.3.9 Brand Bots

Business are using the chat channels to create brand awareness and engagement. One of the major incentives to use bots in this area is app fatigue, creating this way a new way to engage with users.

2.4 Chatbots, Humans and GUI Applications

This section explores what are the chatbots benefits, when replacing human interaction and traditional graphical user interface applications (GUI). It is also explored the intrinsic differences of the interactions.

2.4.1 Human-human vs Human-Chatbot Interaction

There are some differences in the way that people interact with a bot compared to a human. A study (Hill et al., 2015) concluded that people communicate with the chatbot for longer durations, using shorter messages, than they did with another human. Additionally, human-chatbot communication used simpler vocabulary than what is found in conversations among people and exhibited greater profanity. Factors such as number of words per conversation, shorthand terms, and emoticons were found to have no statistically significant differences.

The usage of chatbots in some scenarios bring advantages over humans, namely (Janarthanam, 2017):

- *Consistency*: chatbots can be consistent in services, which is important in certain sectors and may be hard to achieve with human operators.
- *Scalability*: chatbots can easily scale up to handle periods of unregular increased traffic, which is much harder with human operators.

With good design and implementation, Accenture (Accenture, 2016) reports more than 80% of chat sessions resolved by a chatbot, that would otherwise been a human in a chat session or call.

2.4.2 Human-Traditional GUI vs Human-Chatbot Interaction

A report (Ask et al., 2016) by Forrest identifies the following factors that foster chatbots adoptability over traditional applications:

- *Chatbots Promise a More Convenient and Natural User Interface*: Typically, users must go to the process of discover, download, and install apps. Then, apps provide touch graphical interfaces to help consumers perform tasks. The experience isn't natural, but it is effective. Conversations offer a more natural experience.
- *Mobile Moment Ownership is Plateauing for Enterprises*: Mobile is the first screen for consumers; however, consumers use only 25 to 30 apps on average each month and spend 88% of their time in just five downloaded apps.
- *Heavy use of Instant Messaging Platforms*: Consumers spend 78% of their time on smartphones within apps. The median usage of instant messaging apps is 21.47 minutes per day among users of those apps and the pace of adoption is accelerating.

The fact that we are living an app fatigue moment, allied with the heavy usage of messaging apps, present an opportunity to replace traditional applications with chatbots available on the messaging applications that users are already using.

3 CHATBOT USE CASE EVALUATOR

In order to enable the evaluation of use cases, several factors are identified, reflecting the characteristics a use case should have in order to be appropriate to be implemented in a chatbot.

Chatbots lie between human operators and traditional graphical user interfaces (GUI) applications. In one hand, they can be used in the place of a human, offering a similar way of

interaction, by using natural language. On the other hand, they can also be used instead of a traditional application, replacing a traditional graphical user interface with natural language.

The factors are divided in three major groups:

1. **General Factors**: general factors that are essential to be considered to assure the suitability of the use case to be implemented in a conversational UI.
2. **Factors over GUI Application**: this group of factors reflect characteristics of a use case that can indicate that a chatbot is more adequate to expose it, instead of a traditional GUI application.
3. **Factors over Human**: factors that reflect characteristics of a use case that can indicate that the use case would benefit from being implemented by a chatbot instead of a human operator.

The analysis of such factors for each category yielded the following result.

3.1 General Factors

- **Business Rules Well Defined**: Chatbots perform better solving specific requests were the process to solve it is standard (Sengupta and Lakshman, 2017). This facilitates the creation of the flow of the conversation based on that business rules.
- **Integration with Existing Systems**: concerns if it's possible to integrate the bot with the organization systems, via existing APIs. This factor guarantees that the chatbot can access the business logic and data required to the use case in question.

3.2 Factors over GUI Applications

- **Multiple Steps or Multiple Input Parameters** (Accenture, 2016): A simple traditional UI might be more practical to use cases that are simple and require only one step, but for tasks that require several user data, using NLU we can sometimes get all the information that the user would input in a form, for instance, in a simple sentence. Consider the sentence "Can you rebook my flight to Madrid to the following Monday after 3pm and get me a window seat". A traditional GUI would require the user to insert the different pieces of information in different steps of the process, while a chatbot would recognize all the information parameters directly from the natural language sentence. This presents one of the main advantages of chatbots using NLU.

- **Notifications Required:** Messaging applications already include an efficient and functional push notification system, which is available by default without any additional implementation effort (Klopfenstein et al., 2017).
- **Authentication Required** (Klopfenstein et al., 2017): Usually, for each new application, users must create a new account to be uniquely identified. With bots, user authentication is not necessarily needed. The messaging platform used already provides a reliable identification of the user. Users are uniquely identified by default. This reduces the effort asked to the user to start using the service, not requiring to create an additional account.

3.3 Factors over Humans

- **High Volume, Simple Tasks, Performed by Humans:** For simple, well defined, repetitive tasks, a chatbot can be more suitable than a human, in the sense that is more economical and frees the HR for another tasks (Accenture, 2016).
- **Consistency Required:** For use cases that is important consistency in the performance, i.e., the use case must be performed the same way in every occurrence, chatbots can be more suitable than a human operator (Janarthanam, 2017). Bots are intrinsically more consistent than human operators.
- **Scalability Required** (Janarthanam, 2017): some use cases have unstable loads of requests from users. Bots can scale-up to fulfil the requests. Using human operators, is hard to handle sudden increases of requests.

4 APPLICATION

In this section, the previously defined factors will be applied to a concrete use case, in order to demonstrate how these factors can be used in practice.

The use cases is in the context of MedClick, a company in the healthcare field, that will provide a one-stop platform to book a medical appointment in a fast and user-friendly way, across multiple medical service providers

The use case of Scheduling an Appointment is assessed.

4.1 Scheduling Appointment Use Case

4.1.1 Use Case Definition

In order to schedule an appointment, the user must select the speciality needed; choose one of the

available doctors; and finally choose one of the available time slots. This use case is typically performed with clerks, either by phone or in person, or in the case of some clinics, an application is available to the user schedule the appointment independently. Users are notified of the appointment close to the scheduled date, in order to confirm the presence or optionally reschedule.

4.1.2 Use Case Evaluation

In this section an analysis to the suitability of the scheduling an appointment use case is performed, using the factors defined in section 3.

General Factors: Scheduling an appointment is a common use case, that has business rules well defined. In the case of MedClick, an API is available in order to request all the information needed to perform this use cases, including list of doctors, available time slots, and scheduling the appointment.

Table 1: General factors assessment for scheduling appointment UC.

Factor	Assessment
Business Rules Well Defined	Yes
Integration with existing Systems	Yes

Factors Over GUI Application: Scheduling an appointment has multi input parameters, namely the desired medical specialty, the name of the doctor and the time slot. The identification of the user must also be known. The user must be notified close to the date of the appointment, in order to confirm its presence or, optionally, reschedule the appointment.

Table 2: Factors over GUI application assessment for scheduling appointment UC.

Factor	Assessment
Multiple Steps/ Multiple Input parameters	Yes
Notifications Required	Yes
Integration with existing Systems	Yes

Factors Over Humans: Hospitals and clinics usually address high volume requests for appointment scheduling. The volume of requests may vary in an unpredictable way, requiring scalability. This use case is still commonly performed by human operators.

Table 3: Factors over Humans assessment for scheduling appointment UC.

Factor	Assessment
High Volume, Simple Tasks, performed by humans	Yes
Consistency required	No
Scalability required	Yes

4.1.3 Evaluation Conclusions

The fact that this use case meets the two general factors indicates the viability to implement it in a chatbot. Furthermore, it is possible to conclude that the use case is might benefit from the implementation in a chatbot over a traditional GUI application, meeting the three factors. It might also be adequate to implement it in a chatbot over a human.

5 CHALLENGES IN DIALOG SYSTEMS AND NATURAL LANGUAGE PROCESSING

It is important to consider that the Natural Language Processing (NLP) field is in constant development but faces some challenges that are still open to date. The quality of a dialog system is linked to the quality of its natural language understanding module. It is therefore important to acknowledge current issues in NLP.

NLP must deal with the ambiguity of natural languages, i.e., the multiple meaning that the same sentence or word can have and with linguistic variability, i.e., the fact that the same idea can be expressed in multiple forms.

Co-reference is another challenge in NLP. It is a core task in NLP far for being solved despite the significant progress observed on learning-based coreference research (Ng, 2017).

When considering dialog systems that require speech recognition, other challenges arise such as speaker variability, channel variability and environment variability (Petkar, 2016).

This section references only some of the problems in NLP. Even though improvements in these problems are active research topics, they must be considered in the sense that they are not fully addressed and may compromise the quality of the dialog agent.

6 CONCLUSIONS

Conversational user interfaces lie between the traditional user interfaces interactions and human interaction using natural language. Chatbots present an opportunity to automate use cases performed by human operator, offering the same natural way of communication, and can also be used in place of traditional applications offering a more natural interaction. It is important to evaluate if a use case is suitable to the characteristics of conversational UI before deciding to implement in a chatbot over a GUI application or human. In this paper were identified factors that can be used to facilitate this evaluation, aiming to contribute to more informed decisions and more successful chatbot implementations. The application of this factors to the particular use case of scheduling a medical appointment indicates this use case as a good candidate to implement in a chatbot.

7 FUTURE WORK

In order to evaluate the use case evaluator, a chatbot for the use case of scheduling a medical appointment will be developed. Users will interact with the bot and also with a traditional GUI application for the same end of scheduling an appointment. Metrics of both interactions will be compared, such as efficiency (measured in time), and task completion success. The goal is to determine if the use case selected by the evaluator is indeed appropriate to a chatbot, when compared with a traditional GUI application.

ACKNOWLEDGEMENTS

This work was supported by national funds through Fundação para a Ciência e a Tecnologia (FCT) with reference UID/CEC/50021/2019 and by the European Commission program H2020 under the grant agreement 822404 (project QualiChain)

REFERENCES

- Accenture, 2016. *Chatbots in Costumer Service*, s.l.: s.n.
- Ask, J., Facemire, M. and Hogan, A., 2016. *The State Of Chatbots*, s.l.: s.n.
- Hill, J., Ford, R. and Farreras, I., 2015. Real conversations with artificial intelligence: A comparison between human-human online conversations and human-

- chatbot conversations. *Computers in Human Behavior*, Volume 49, pp. 245-250.
- Hirschberg, J. and Manning, C., 2015. Advances in natural language processing. *Science*, 349(6245), pp. 261-266.
- Janarthanam, S., 2017. *Hands-On Chatbots and Conversational UI Development*. s.l.:Packt Publishing.
- Jurafsky, D. and Martin, J., 2008. *Speech and Language Processing*. 2nd ed. New Jersey: Prentice Hall.
- Jurafsky, D. and Martin, J. H., 2018. *Speech and Language Processing (Third Edition draft)*. s.l.:s.n.
- Klopfenstein, L., Delpriori, S., Malatini, S. and Bogliolo, A., 2017. *The Rise of Bots: A Survey of Conversational Interfaces, Patterns, and Paradigms*. s.l., s.n., pp. 555-565.
- Ng, V., 2017. Machine Learning for Entity Coreference Resolution: A Retrospective Look at Two Decades of Research. *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, pp. 4877-4884.
- Petkar, H., 2016. A Review of Challenges in Automatic Speech. *International Journal of Computer Applications*, 151(3), pp. 23-26.
- Raman, A. and Tok, W. H., 2018. *A Developer's Guide to Building AI Applications*. s.l.:O'Reilly Media.
- Sengupta, R. and Lakshman, S., 2017. *Conversational Chatbots – Let's chat*. [Online] Available at: <https://www2.deloitte.com/content/dam/Deloitte/in/Documents/strategy/in-strategy-innovation-conversational-chatbots-lets-chat-final-report-noexp.pdf> [Accessed 25 September 2018].
- Shevat, A., 2017. *Designing Bots*. Birmingham: O'Reilly Media.
- Zhang, X. and Wang, H., 2016. A Joint Model of Intent Determination and Slot Filling for Spoken Language Understanding. *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, pp. 2993-2999 .