

# Chappie - A Semi-automatic Intelligent Chatbot

**Bibek Behera**

Magictiger

Indiranagr Double Road

Bangalore 560038, India

bibek@magictiger.com

## Abstract

The personal assistant, a human is inefficient and takes a long time to process a single request such as booking tickets, ordering food, getting beauty-services etc. This reduces the throughput and business performance drastically. Obviously there is an increase in the demand of chat automation because a) it removes the human factor and b) it can give a 24 hour service which will have a multiplicative effect on the revenue generation. The chatbot, Chappie was born as a business requirement to automate the personal assistant or concierge. Presently Chappie is being used as a routing agent wherein it can classify the requirement of user into one of the services provided by business based on the first few chats and then transfer it to an agent expert in that service. It uses natural language processing (nlp) to analyse chats and extracts intent of the user with a score similar to the likes of WIT<sup>1</sup>. Then it uses this information and AIML(Artificial Intelligence Mark-up Language) to make a conversation with the user. This is the marked difference between Chappie and existing chatbots like ALICE(Shawar and Atwell, 2003), which work solely on AIML. Hence we describe it as intelligent compared to other bots. Once it understands the intent and gets other information like name and email id, it transfers the chat to

a human agent. Thus it is a semi-automatic chatbot that switches to a manual mode once it is able to understand the intent of the user and the kind of service required by her. Presently we let Chappie go live once an hour everyday and we record the responses as well as test cases for classification. So far the results have been encouraging since people have not been able to distinguish Chappie from a human being. Also it has been able to route correctly on several occasions. Once we obtain a decent amount of accuracy in routing and chat coherence, it will be deployed on a 24 hour basis.

## 1 Introduction

We are a service provider company and we claim to fulfill all the requirements of customer be it ordering food, booking tickets or paying electricity bills. The chat agents are expert in their fields. But they are inefficient and they lack the continuity required in increasing throughput in enhancing business. They need breaks and relaxation time which can be harmful to business specially a start-up. So it is natural that business wants to change it orientation from humans to automation. Chappie was born as a requirement from business side and desire for efficiency and automation.

As far as automation of user experience is concerned, they are three ways. First one is the flipkart/amazon way wherein the customer has to search

<sup>1</sup><https://wit.ai>

for his/her product and choose it herself before initiating payment. In this case, user will get a mechanical experience and far from the user experience that a company provides. Many people in Asia and in other countries/continents are not very computer friendly and will prefer chat to get orders fulfilled. Those are our potential targets. The second way is to automate chats in a way totally oblivious to the customer. In this way, she gets the user experience while nullifying the defects of a human agent atleast partly. If something goes wrong, the bots route it to a human agent. But we assume that most of the services we are dealing with like booking flights or paying electricity bills can be automated. So the second is a hybrid approach which has both automation and human element. The third approach is to totally automate the chat. This is the biggest challenge and has not been achieved satisfactorily by any known system whether it be Siri or Iris. So there are 3 ways to automation :-

- mechanical(flipkart)
- semi-automatic(Chappie)
- completely automatic(Futuristic/fictional chatbot - like Jarvis shown in the scifi movie Ironman)

In this paper, we discuss a semi-automatic intelligent chatbot called Chappie. The entire paper concentrates on the bot only and not the human element of the overall system. The long term idea is to slowly get rid of humans by improving on algorithms and design of system. This way we can move to the futuristic chatbot which will be completely automated and simultaneously gives a seamless user experience.

In the following sections, we try to understand existing bots and their limitations in section 2. Then we try to define an intelligent chatbot in section 3 and set criteria. Afterwards, we demonstrate the working of Chappie in section 4 and how it fulfils the criteria one by one. Next we describe the algorithm briefly in section 5 followed by an example in section 6. Finally we have a result section 7 followed by conclusion and future works in section 8.

## 2 Background of chatbots

The advent of chatbots has created a new dimension to AI research. Chatbots are intelligent interfaces that can make conversation in a coherent manner. They were made to replace human beings as chat agents. The conventional chatbots use AIML, a pattern matching xml parser, as a response system. AI community have tried to construct chatbots like ALICE and ELIZA. These are generic chatbots. There are others like chatbots for cultural heritage(Pilato et al., 2005), security training(Kowalski et al., 2013), blind high school students(Bigham et al., 2008). There have been discussion on the utility of chatbots(Shawar and Atwell, 2007).

All these chatbots function similarly. If we consider Alice, it basically generates an AIML file from a dialogue corpus with patterns and templates(Shawar and Atwell, 2003). The limitation begins from the presence of a corpus which assumes all knowledge comes from previous dialogue done by human agents. Secondly generating AIML from a corpus cannot guarantee a coherent chat because there is a fear of getting repetitive statements, which will worsen the user chat experience. There are two defects in existing chatbots - lack of intelligence and similar responses for repetitive statements due to the pattern matching nature of AIML.

There are works which try to understand the semantics(Augello et al., 2009). But the responses are generated using statistics. In this case the responses are memorized. Also people have tried to capture semantics through different ways of knowledge representation(Pilato et al., 2012). This seems an overfitting approach because humans have a generalist approach towards understanding concepts. Different knowledge representation have an inherent undecidability problem of application in which context thus making the system unnecessarily complicated.

## 3 What is an intelligent chatbot?

A chatbot like ALICE lacks AI in the true sense. Firstly, if a bot wants to mimic a human, it should not memorise every conversation in database. An intelligent chatbot must be powered by AI/NLP to reply coherent messages atleast from the business point of view. ALICE has no way to understand the intent of the user. That is why these existing

192	chatbots are not good enough for replacing human beings. They are over reliant on AIML. There is a demand for an interface between AIML and the chatbot such that it understands chats, filters entities and then generates text in AIML readable format in a smart way. Thus, it can convert its knowledge to a response custom made for AIML. So AIML will become a layer that is at the lowest level of the chatbot brain and is used just an interface to reply in a natural language. Thats how we have designed Chappie which smartly extracts all sorts of information such as name, intent, mail, city, etc and generates a coherent response to user.		
193			
194			
195			
196			
197			
198			
199			
200			
201			
202			
203			
204			
205	Secondly, an intelligent bot will never respond in the same way if user is sending same text. A random response as customary in AIML is not the right way to deal with repetitions. The bot should have a counting mechanism to know that there is repetition in incoming messages and if this behaviour continues further chats should be routed to a human agent in a polite manner or the chat terminated in case of completely automated bots. The ability to count can be achieved only if the bot has a system to understand the intent of the client and if it maintains a state diagram or a milestone status with a counter for each state. In this manner, if it encounters the same state on two or more consecutive occasions, it can safely route the user to a human agent or terminate.		
206			
207			
208			
209			
210			
211			
212			
213			
214			
215			
216			
217			
218			
219			
220	Thirdly, response cannot be generated based on probability. That will cause havoc because sentence generation should be accurate and AIML, that is self curated and not generated by pattern extraction algorithms, is the proper way to respond. System cannot entirely depend on AIML, but it cannot get rid of AIML. AIML is a very powerful device when it comes to generating responses in natural language.		
221			
222			
223			
224			
225			
226			
227			
228	The three important criterion of an intelligent chatbot are:-		
229			
230			
231			
232			
233			
234			
235			
236			
237			
238			
239			
240			
241			
242			
243			
244			
245			
246			
247			
248			
249			
250			
251			
252			
253			
254			
255			
256			
257			
258			
259			
260			
261			
262			
263			
264			
265			
266			
267			
268			
269			
270			
271			
272			
273			
274			
275			
276			
277			
278			
279			
280			
281			
282			
283			
284			
285			
286			
287			
288			
289			
290			
291			
292			
293			
294			
295			
296			
297			
298			
299			
300			
301			
302			
303			
304			
305			
306			
307			
308			
309			
310			
311			
312			
313			
314			
315			
316			
317			
318			
319			
320			
321			
322			
323			
324			
325			
326			
327			
328			
329			
330			
331			
332			
333			
334			
335			
336			
337			
338			
339			
340			
341			
342			
343			
344			
345			
346			
347			
348			
349			
350			
351			
352			
353			
354			
355			
356			
357			
358			
359			
360			
361			
362			
363			
364			
365			
366			
367			
368			
369			
370			
371			
372			
373			
374			
375			
376			
377			
378			
379			
380			
381			
382			
383			
384			
385			
386			
387			
388			
389			
390			
391			
392			
393			
394			
395			
396			
397			
398			
399			
400			

n is the number of milestones. For each state, we generate an AIML pattern and a response. Now we are ready to initiate the chat. Every chat may have some information and hence a response will be generated based on the state and counter associated with the state. The response is asked in a manner so as to elicit information from the user and reach the next milestone. Once we have completed the conversation, routing is initiated and chat is transferred to a human agent. If there are repetitive messages, routing is done immediately without waiting for completion of conversation. This way we prevent impostors from jamming our system.

- Input:- A set of messages
- Output:- A json containing information such as name, email, intent, etc.

## 6 A sample chat

Lets say for the routing bot or greetings bot, we can have five milestones like welcome, name, email, intent, routing. Using these milestones, we can have a tree with  $5! = 120$  leaves. Each leaf will denote a state. We have demonstrated a test case in Figure 1. If a user says “hi”, we dont have any information, so we welcome her and ask her requirement. In the next iteration if she says “I want a chicken biryani” we get the category as food with a score and sub-category is “chicken biryani” and intent is “want”. Now we ask her name and if she says “xyz”, we have all the information and hence we route the chat and some notes to the human agent. He can then rate the conversation which will be used as feedback to our classifier. In the future we can deal with wrongly classified chats to improve accuracy of the system.

## 7 Results

We used a manually labelled training corpus of 5000 sentences. We obtained decent accuracy of 65% with a bag-of-words based multiclass Naive Bayes Classifier and even SVM followed closeby with 67%. Then we filtered unnecessary sentences like hi, hello, welcome, etc and our corpus reduced to 1500 sentences. This time our accuracy increased to 92%.

We have not evaluated Chappie because it has gone live recently. In future, we will be able to give

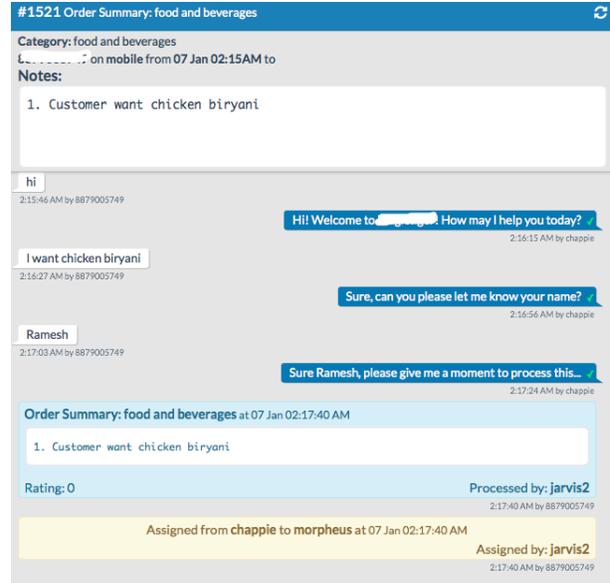


Figure 1: Chappie is the chatbot. Jarvis is the supervising bot. Once Chappie finishes the chat it informs Jarvis which then routes the chat to human agent Morpheus. Based on the chat experience human agent can give a rating of 0-5. There is also a note at the top which summarises the overall chat that “Customer wants chicken biryani”. It also assigns the category as food-and-beverages.

a numerical score to Chappie. At the surface level, it’s working smoothly. People have not been able to identify it as a bot.

## 8 Conclusion

Through Chappie, we are trying to redefine chat experience in an automated manner. The novelty lies in the way we define our system as not merely a response generator but an intelligent interface to a response generator. Then we try to bring counting as a way to avoid repetitions. Overall Chappie is performing decently, but it needs more sophisticated algorithms to extract intent and classify chats more accurately. In the future, we will retain the milestone format but we need to scale Chappie to handle the entire conversation. So there can be milestones within milestones. Overall Chappie will be a cooperation of bots with a supervisor bot called Jarvis sitting on top of domain expert bots very much like the hierarchy of chat agents in our company.

384	<b>References</b>	432
385		433
386	Agnese Augello, Giovanni Pilato, Giorgio Vassallo, and	434
387	Salvatore Gaglio. 2009. A semantic layer on semi-	435
388	structured data sources for intuitive chatbots. In <i>Com-</i>	436
389	<i>plex, Intelligent and Software Intensive Systems, 2009.</i>	437
390	<i>CISIS'09. International Conference on</i> , pages 760–	438
391	765. IEEE.	439
392	Jeffrey P Bigham, Maxwell B Aller, Jeremy T Brudvik,	440
393	Jessica O Leung, Lindsay A Yazzolino, and Richard E	441
394	Ladner. 2008. Inspiring blind high school students to	442
395	pursue computer science with instant messaging chat-	443
396	bots. In <i>ACM SIGCSE Bulletin</i> , volume 40, pages	444
397	449–453. ACM.	445
398	Stewart Kowalski, Katarina Pavlovska, and Mikael Gold-	446
399	stein. 2013. Two case studies in using chatbots for	447
400	security training. In <i>Information Assurance and Secu-</i>	448
401	<i>rity Education and Training</i> , pages 265–272. Springer.	449
402	Giovanni Pilato, Giorgio Vassallo, Agnese Augello,	450
403	Maria Vasile, and Salvatore Gaglio. 2005. Expert	451
404	chat-bots for cultural heritage. <i>Intelligenza Artificiale</i> ,	452
405	2(2):25–31.	453
406	Giovanni Pilato, Agnese Augello, and Salvatore Gaglio.	454
407	2012. A modular system oriented to the design of ver-	455
408	satile knowledge bases for chatbots. <i>ISRN Artificial</i>	456
409	<i>Intelligence</i> , 2012.	457
410	Bayan Abu Shawar and Eric Atwell. 2003. Using dia-	458
411	logue corpora to train a chatbot. In <i>Proceedings of the</i>	459
412	<i>Corpus Linguistics 2003 conference</i> , pages 681–690.	460
413	Bayan Abu Shawar and Eric Atwell. 2007. Chatbots: are	461
414	they really useful? In <i>LDV Forum</i> , volume 22, pages	462
415	29–49.	463
416		464
417		465
418		466
419		467
420		468
421		469
422		470
423		471
424		472
425		473
426		474
427		475
428		476
429		477
430		478
431		479