The BIMbot: A Cognitive Assistant in the BIM Room

19

Ivan Mutis, Adithya Ramachandran, and Marc Gil Martinez

Abstract

Today, collaborative environment method is of with widespread use among project stakeholders. They benefit project planning in a variety of ways, including by enabling team members to build stronger relationships, enhance communication, and perform efficient planning, to name several. The collaboration occurs in sessions that immerse stakeholders in an environment commonly referred to as a BIM room—a shared space that enables project stakeholders, such as architects, general contractors, structural and MEP trades, and other specialized knowledge actors, to physically or virtually meet and to establish constant presence. The BIM room is a medium for stakeholders (BIM-room participants) to more accurately and efficiently make informed decisions on end to end construction problems. This project is aimed at investigating the use of information technology as a mediating mechanism to facilitate sharing meanings of expressions and to assist stakeholders in effectively finding relevant information that connects to their intent in the BIM-room. This research proposes the creation and implementation of a cognitive assistant to project stakeholders: BIMbot. The BIMbot is an agent that will have the ability to simulate a conversation or a messaging exchange with a present actor. From the actor-BIM-bot exchange and having an order, command, or request, BIMbot will carry common functions for the actors within the BIM room like retrieve the current version of family-objects of the BIM; load, filter, and view section(s) of interest; automate object placement; etc. BIMbot is designed to produce significantly more efficient interaction of collaborative meetings in the BIM room.

Keywords

Cognitive intelligent agent • BIM room • Pre-design phase • Generative models

19.1 Introduction

When Building Information Modelling (BIM) was first introduced, few could have imagined the large impact that this term would reach in the years that followed. BIM has gone from being a barely known academic development to capturing the attention of the whole architecture, engineering and construction (AEC) industries [1] (1). But beyond technology, BIM unfolds a whole new set of design and construction capabilities with regard to traditional practices (2). These workflows encourage all project stakeholders to collaborate from initial design stages, creating changes in roles and relationships amongst them. Therefore, BIM can be better understood as a process towards innovative practices rather than a mere set of tools and technologies. Today, designers and construction stakeholders from several disciplines are required to collaborate together in temporary teams [2]. The communication and collaboration is key from the early stage normally called pre-design phase, where the design team can set new relationships faster. This exchange is described in Fig. 19.1. Then, BIM aids an integrated design and construction procedure that needs to be supported by Integrated Project Delivery (IPD). Its methods consist of carrying out early collaboration amongst the stakeholders who are involved in the project. The logical

I. Mutis (✉) · A. Ramachandran · M. G. Martinez
Illinois Institute of Technology, Chicago, IL 60616, USA
e-mail: imutissi@iit.edu

I. Mutis and T. Hartmann (eds.), *Advances in Informatics and Computing in Civil and Construction Engineering*,
https://doi.org/10.1007/978-3-030-00220-6_19

way of thinking would lead us to claim that the teams are more confident when computer tools embrace coordination and collaboration [2]. Nevertheless, the principal proposition in the academic and professional literature on BIM are seriously concerned on issues involving the creation, exchange and management of data; shared geometries and software interoperability, which are valuable for getting teams to work in shared computing environments. In this paper, our attention is focused on the dialogue and conversations that constitute collaboration, especially those in technologically supported environments. Dossick and Neff [3] carried out a study to conclude that clean technology, although highly contributes to a better project documentation, planning, scheduling and also providing the project with data-rich, object based intelligent and parametric digital representations of the facility, still brings a lot of inefficiencies with the dynamic and those "messy" activities needed to support problem-solving [2].

This is where our approach takes part in the BIM process. An artificial cognitive assistant called BIMBot, a cognitive assistant that approaches to the requirements of messy talk since it has been created as an active and interacting tool. The idea is to incorporate this artificial assistant to the team members of the collaboration team. We hypothesize that this tool will fulfill our goal to reduce the inefficiencies in the pre-design process, since the stakeholders will be able to shorten the actions of the conversations by interacting with the cognitive agent which will rapidly respond to queries and actions, by retrieving those tasks in order to survey require tasks as a result of that messy talk.

To implement a cognitive assistant, a qualitative research method will be carried out to obtain the correct data. This is performed since we want to understand meanings of different stages, tools, and opertaions of the BIM process. Also, we wish to collect information through describing and understanding experience, ideas and beliefs of BIM experts that can add value to our collected data. To develop inferences and collect data, we review a myriad of research papers, BIM magazines, BIM standards and other experiences that we can find in the web. Reviewing papers will certainly point us in the right direction to arrive on the right inferences but it will not help us in gathering the data we need for this research. This is because no real conversations (data) in huge volumes are recorded in research papers. As mentioned, pre-design phase is studied and focused on the clashes/interferences phase. In Table 19.1 is described the most frequent ones in BIM room. So once the BIM software detects those interferences, the BIM participants involved in the BIM meetings, usually resort to solving interferences verbally. It is the verbal conversations during the pred-design phase that act as data for BIMBot. Therefore, by providing this data to the BIMbot, the final solution will enable to support another similar discussion in the near future, either in the same team or another project. As a validation, we will propose to set up an innovative experience in

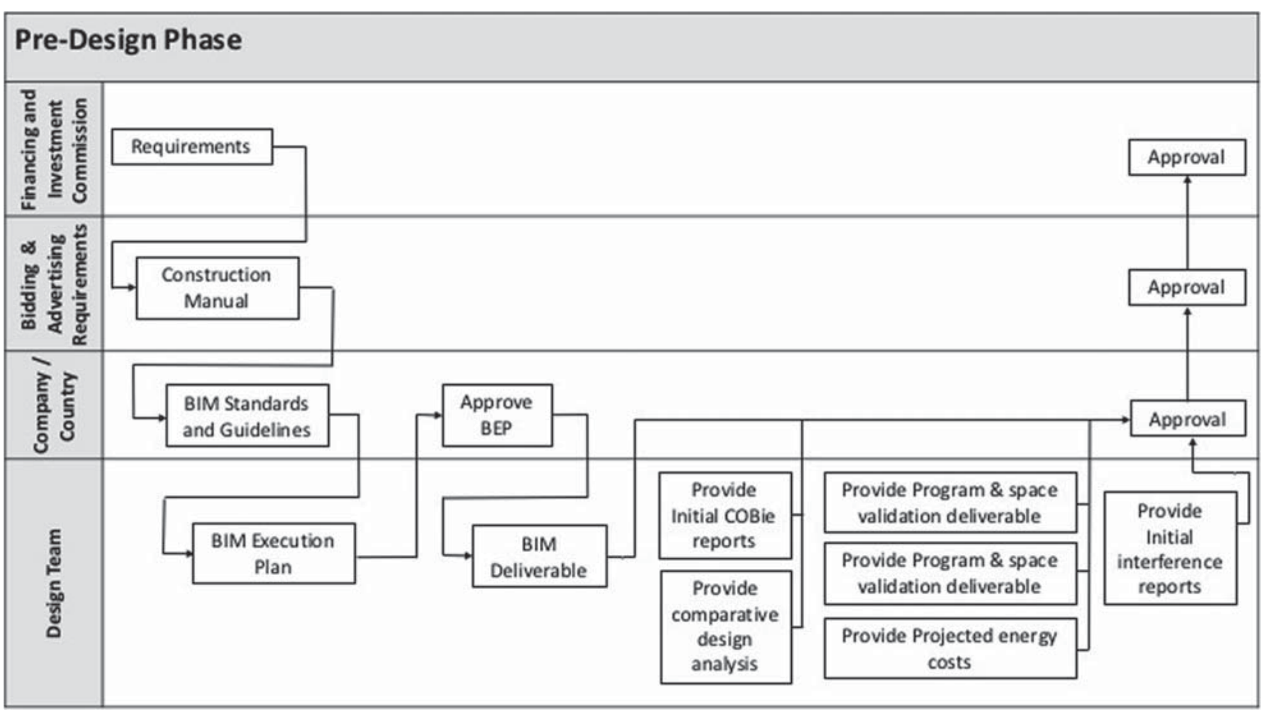

Fig. 19.1 Based on Georgia Tech BIM requirements for architects, engineers and contractors

Table 19.1 Level two interferences. Georgia Tech

Casework	Electrical fixtures, devices, conduit, raceways, homeruns, cable trays, piping systems and accessories
Furnishings	Electrical fixtures, devices
Structure (columns, beams, Framing, etc.)	Specialty equipment, electrical equipment, fixtures, devices, conduit, raceways, homeruns, cable trays, piping systems and accessories
Ductwork and piping	Floors, specialty equipment, electrical equipment, fixtures, devices

the BIM room, so that we will carry out a selection of a design team to work closely with our BIMbot. Thus, they will be able to check the ability of this cognitive assistant and see whether or not it is a useful tool to incorporate in a regular BIM room. So that, it will let the assistant find the most relevant information from any other conversation. The cognitive assistant that we have developed is currently capable of retrieving information on desired models. This prototype is the first version of many to follow in the aim of achieving our final goal which is elaborated in Sect. 19.5.

19.2 Related Work

The construction of a project is a complex process that involves activities of various nature carefully organized in time and space. If building information modelling wanted to represent the building process in all its depth, it would have to integrate the work of engineers, architects, contractors and all disciplines involved. This was the main reason why BIM software titles started to develop own versions focused on a certain area of the AEC industry. Several research papers have been focused on the collaboration and coordination among the actors involved in a construction project. This collaboration is carried out in a natural environment called BIM room [4], which is a unique virtual place, accessible and operable from the internet. The main purpose is to manage the BIM model, the information that the BIM generates (graphic and non-graphic), the evolutionary processes of the project, the roles of the agents involved and their communications in a structured and directional way towards the achievement of the project's objectives [5].

The CDE (Common Data Environment) [2] provides a collaborative environment where teams can share their work and files for the Information Management process. The CDE could consist of more than one system configured to meet the requirements defined in BS 1192: 2007 (United Kingdom). This document presents the management process within the CDE necessary to facilitate the delivery of the necessary project information. As it is a unique environment where different disciplines, interests and collaborative actions converge, this system should facilitate interoperability between the different software platforms and be based on standard formats of free access.

Researchers quickly realized about the lack of efficiency on the BIM room and studied the integration of technology into the collaboration by using ethnographic observation and one-on-one interviews with project participant [3]. They also introduce the concept of Messy talk, which is defined as the interstitial dialogue between and after formally organized agenda items. Finally, it is understood that although technology is a useful tool that contributes to a better project documentation, planning, scheduling and also providing the project with data-rich, object based intelligent and parametric digital representations of the facility, still brings a lot of inefficiencies with the dynamic and those "messy" activities needed to support problem-solving [2]. Thus, an artificial cognitive agent to develop and enhance this task is needed if it is created taking into account the previous work conclusions.

19.3 Methodology

This section highlights the reason for choosing to implement our cognitive assistant in Pre-design phase (Sect. 19.3.1) and the architecture of our cognitive assistant (Sect. 19.3.2).

19.3.1 Pre-design Phase

The collaboration begins in the pre-design phase, where stakeholders are involved in an environment commonly named BIM-room. Project stakeholders are architects, general contractors, structural and MEP trades, among other specialized knowledge actors, who physically or virtually share the same space to establish constant presence. In this phase, it is discussed several issues and clashes that may occur among all the stakeholders, who communicate during the sessions through a conversationally verbal exchange. Nevertheless, this communication is not yet well supported by BIM technology, as indicated by Dossick and Neff in their paper [3], that still needs to be adapted to the concept of "messy talk" carried out in the pre-design phase. Thus, based on the lack of efficiency in the discussion phase, where drawings or notes are still used instead of technology, BIMbot will indeed serve as a tool to put technology and collaboration together. Therefore, our first approach is carried out in the pre-design phase, firstly starting on the clashes and interference discussions, as it is explained and illustrated below.

19.3.1.1 Conflicts/Interferences at the Pre-design Phase

The first step of our approach is to get familiarized with the current practices of IPD and the clashes or interferences that normally arise among the design team in the pre-design phase. IPD, is a project delivery method distinguished by early collaboration between cross-functional teams through all phases of design, fabrication, and construction. By entering this type of contractual agreement, teams are able to collaboratively harnesses the talents and insights of all participants to optimize project results, reduce waste, and maximize efficiency through all phases of design, fabrication, and construction. The entire process, from concept to construction, is defined by early substantive engagement by all key stakeholders.

In this process three levels of interferences were identified [6]. In the first level are the ones considered critical to the design process. These interferences have been assigned the highest priority and should be rectified within the model as soon as possible. In the second level are those which are considered important to the design and construction process [6]. These interferences have been assigned the greater priority and should be rectified during project meetings during design. Lastly, the interferences that belong to the third level are reported interferences that, while are considered important to the correctness of the model, will generally be changing on a regular basis throughout the design and construction process. These interferences have been assigned a lower level of priority but should be rectified before the phase submission of the models [6]. Thus, we conducted an analysis of the second level of interferences since they are marked as the greatest priority. The most frequent clashes cases are found in Table 19.1, where each row works independently and means that normally the elements in the first column get in conflict with the elements of the second. With this analysis, the aim of providing data in a Q&A format to the BIMbot will be achieved.

19.3.2 Cognitive Agent—BIMBot's Architecture

Rule based models are easy to create but they do not work well as they function purely on pattern matching. If there are no patterns found to a given query, then Rule based bots do not generate a response. In addition, as pattern matching functions on rules (like first order logic rules), it becomes a tedious task to write rules to cover all scenarios [7]. As this was seen difficult, introduction of Machine Learning paved way to create intelligent bots. The bots are called intelligent as they have the capability to learn from data or previous conversation. There are two type of models that make use of Machine Learning and exploit its advantages. (1) Retrieval-based models, (2) Generative models. Retrieval-based models do not create any new responses. These models generally select a response from a pool of responses based on the question it received. On the other hand, Generative models are more intelligent. They can create their own responses based on the question they receive. Generative models learn the structure of sentences through training. Generative models require huge amount of data for training purposes. If these models are trained, they tend to outperform Retrieval based models when it comes to seeing queries that have not been seen before [7]. Currently, majority of the research is on Generative models using Deep learning techniques. This is because Deep learning architectures like Sequence to Sequence are apt for text generation. Inspired from the success of generative models, we have created our cognitive agent using a generative model along with a rule-based model which runs on pattern matching [8, 9]. The upcoming sections will elaborate on four major components of BIMBot's architecture: (1) Corpus (2) Neural Machine Translation(NMT) (3) Rule-based Model (4) NLP Engine (Fig. 19.2).

19.3.2.1 Corpus

One of BIMBot's important component is Corpus, which is a collection of texts/data in an unstructured format. It is pivotal for the architecture as it acts as a source of training data for BIMBot. There are two steps involved in building the right corpus for our cognitive agent. The first one was gathering data for the Corpus and the last one being, cleaning the gathered data.

The first step is to source data for training purposes and they are as follows: (1) Cornell Movie Corpus, (2) Reddit and (3) BIM room conversations (from different Civil companies). The reason to aggregate data from three different sources is to have both open domain knowledge (general knowledge) and to have closed domain knowledge (BIM related knowledge).

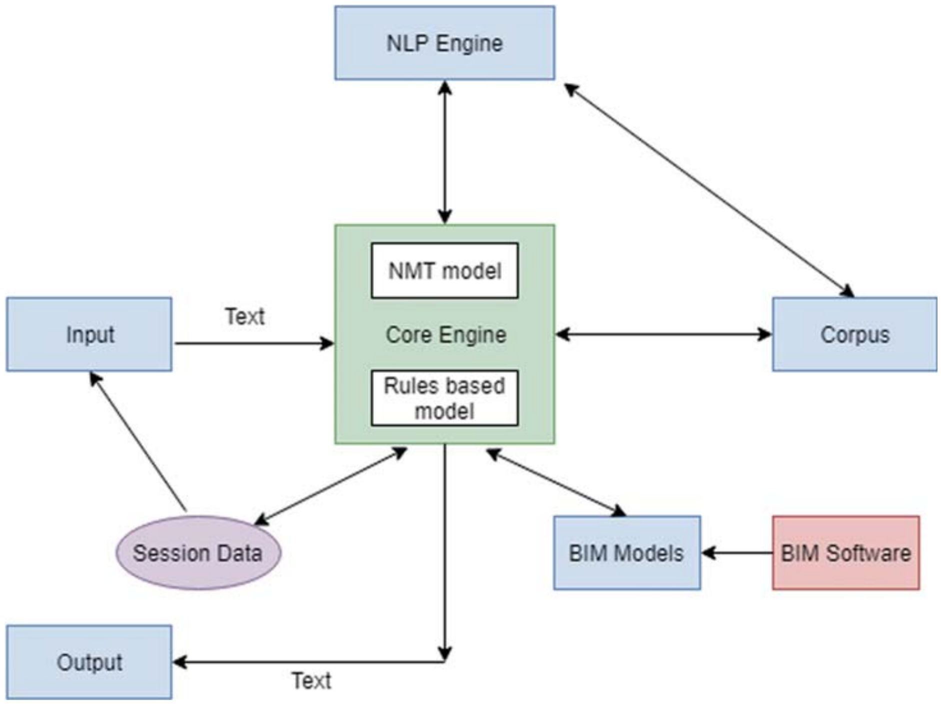

Fig. 19.2 BimBot's architecture

Table 19.2 Cleaned data—Question/Answer format

Question	Answer
I cannot ignore the verdict of my council	Surely you can do anything you want
Can I go out to play?	After you are done with your work

Moving to the last step of our two-step process, the data when first gathered is usually unstructured. A thorough cleaning of data is done before the data is used for training the cognitive agent. The cleaning process happened in two steps. The first step was to structure the sentences in Question/Answer format as described in Table 19.2. Separate python scripts were written for reddit dataset and Movie Cornell Dataset to structure the data. Following this step, we identified profane words in the structured data. Sentences with such words (profane words) led us to truncate the Question/Answer pair. This was done to ensure that the cognitive agent responded in a polite, patient manner and not otherwise.

The BIM room conversations were acquired in two formats: (1) Text format and (2) Audio Format. We converted the audio format to text by using IBM Watson's speech to text web application. After converting the audio clips to text, we aggregated the data from both the formats to manually structure them in Question/Answer format. This was done manually because the quantity of data was low and we wanted to carefully handle sensitive information, should there be any.

19.3.2.2 Neural Machine Translation - Generative Model

For this research purpose, we have implemented BIMBot, our cognitive agent, through Neural Machine Translation (NMT) which falls under the umbrella of sequence to sequence models [10, 11]. Neural Machine Translation is also known as encoder-decoder architecture. Though Neural Machine Translation model differ in terms of architecture, we have gone with the natural choice of implementing BIMBot's encoder and decoder using Recurrent Neural Network (RNN) [12]. Usually, encoders and decoders are implemented using RNNs. Humans have persistence when it comes to thoughts and thinking. When humans make decisions, it is usually based on their previous thoughts. Conventional Neural Networks don't function like the humans do and this is a major drawback. With the introduction of Recurrent Neural Networks(RNN), this issue is addressed. RNNs have loops enabling them to hold information.

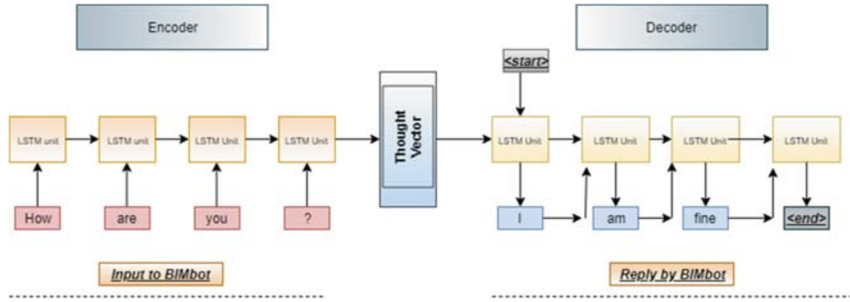

Fig. 19.3 Sequence to sequence model

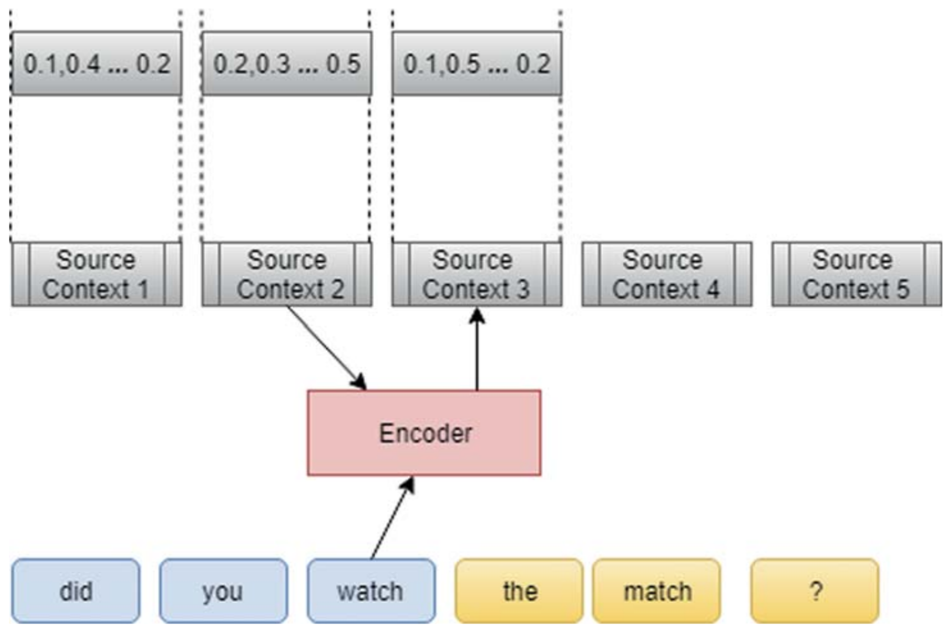

Fig. 19.4 Encoder phase

RNN models vary in terms of (1) "directionality—unidirectional or bidirectional", (2) "depth—single or multi-layer" and (3) "type—a Long Short-term Memory (LSTM), or a gated recurrent unit (GRU)" [13]. For our research, keeping long term dependencies in thoughts, we have opted for an RNN model which is bi-directional, multi layered and has LSTM units. Long term dependencies are important because, as the context/conversation between actors and BIMBot increases, BIMBot should understand the context to generate appropriate responses [14–16] (Fig. 19.3).

The NMT system of BIMBot reads the source sentence with the help of an encoder to build what is called a thought vector as explained in [13]. The thought vector is a "sequence of numbers that represents the sentence meaning". This thought vector is then processed by a decoder to produce the correct response.

Decomposing this learning process, there are two phases involved: (1) Encoder and (2) Decoder. In the Encoder phase (Fig. 19.4), the source sentence from the Corpus is sent to the Encoder to generate "Source Context using the present word and the previous Source Context". The Source Context represents the words in numbers. Each word has a Source Context which is associated with a series of numbers as illustrated in Fig. 19.4. The representation of numbers to a word is done with the help of an embedding layer. The embedding layer is built on words from the Vocabulary file across higher dimensions (Sect. 19.3.3.4). Words that have similar/common property are found to be on the same dimension of space. For example, parts of speech of words can be found on the same dimension while gender can be found on another dimension.

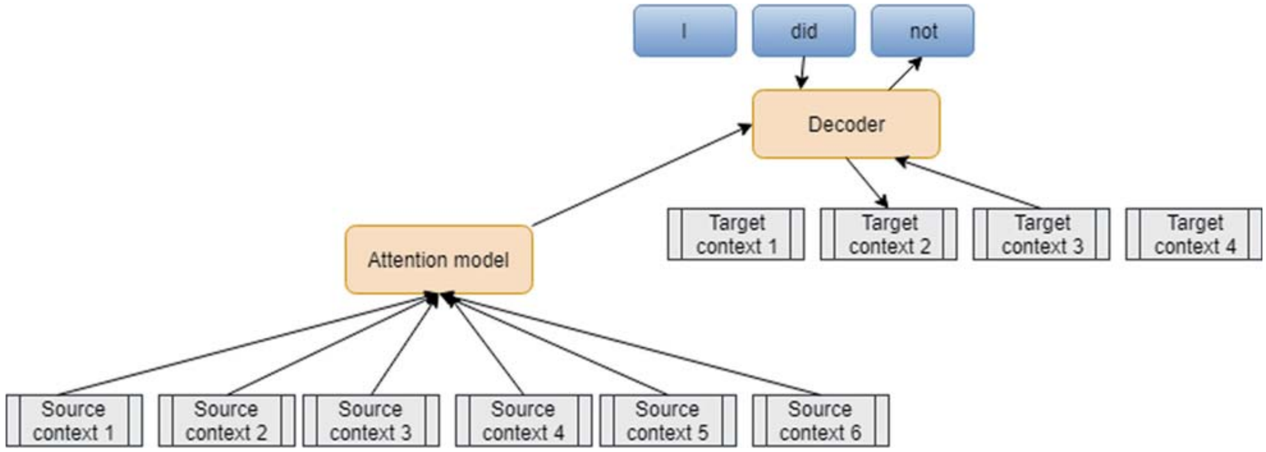

Fig. 19.5 Decoder phase

The target words/responses are generated in the second phase of this process. This is achieved with the help of Target Context, Source Context and the previously generated word (Fig. 19.5). The Target Context contains the status of text generation while the Source Context along with the Attention Model contains the representation of source sentence. The Attention Model calculates attention weights as described in [17] for the Source Context. This representation is called as memory.

At each time step, the Target Context is used to select which part of the memory must be read. This then enables the NMT model to concentrate on the required memory to create appropriate response. The text generation stops when the decoder uses an End-Of-Statement tag [18].

19.3.2.3 Rules-Based Model

The rules-based model is an elaborate text-generation engine which looks for patterns in the input text from the actors "which can serve as hooks to be manipulated and recombined into its responses" [19].

BIMBot offers features like searching IFC models and retrieving IFC mode properties. To have such features, it is necessary to include a rule-based model. Using pattern matching, appropriate rule scripts (majorly based on first order logic) are triggered/called to retrieve properties from IFCxml files. We have created our own version of the IFCxml which replicates the standard IFCxml files. This is done because, IFCxml files cannot be accessed outside BIM software environment. In addition, this provides the capability of addition/deletion of properties at will. Our version of the IFCxml files sit in BIM models which will be accessed by BIMBot on demand to retrieve a model's property/properties.

19.3.2.4 Natural Language Processing(NLP) Engine

The NLP engine consists of tokenizer which splits the texts/data from Corpus into words [20]. The tokenizer helps to create the vocabulary file which is used by BIMBot's core engine during its learning/inference phase.

19.4 Implementation and Experimentation

In our experimental setup, there are two major phases. The first phase being, training, is when our core engine gets trained using the training data from Corpus. The final phase, Inference or testing, is when we test how BIMBot fares against unseen input.

19.4.1 Training

In Sect. 19.3.2.2, we have elaborated on the learning methodology practiced by our cognitive agent to learn one example. Training is the process of repeating the aforementioned procedure for all the conversations in the Corpus.

To train the cognitive agent, we concentrated on two major aspects: (1) Training Data (2) Epoch. The duration/learning will depend on both the aspects. For the cognitive agent to learn diverse topics it is necessary to provide huge volumes of

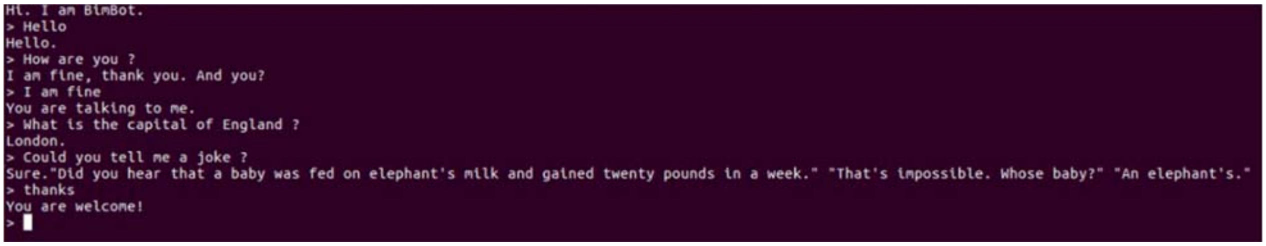

Fig. 19.6 General converstions with BIMBot

data. For the purpose of this experiment, we have gathered around 300,000 pairs of conversations. For the cognitive agent to have a good learning rate, it is necessary to provide the correct number of epochs. An epoch is a forward pass and backward pass of all training data. The cognitive agent could be trained more by increasing the number of epochs, but it is also necessary that we do not overfit/underfit the data by having high/low epochs. We trained our cognitive agent for 36 h on a Nvidia P5000 GPU to create the base model.

19.4.2 Inferences (Results)

After training BIMBot, we were able to get responses for unseen source sentences. This process is called inference. Bimbot is required to understand the semantics of the sentence to gauge if the sentences are general conversations or pertaining to BIM models. BIMBot will understand the semantics and act accordingly as seen in Figs. 19.6 and 19.7.

Currently, there are no standard metrics like precision and recall quantifying the performance of text generation tasks like cognitive agents. This is because, there are far too many ways of responding to a particular question. For such tasks, evaluation is usually done manually. But for our research purpose, we are considering three important metrics: (1) User Engagement, (2) Speed and (3) Functionality.

(1) User Engagement: Cognitive Assistants should be able to initiate conversations and respond to users' requests by keeping the conversations coherent. BIMBot has the capability to keep the conversation coherent up to four questions/responses. When the 5th question/response is provided, our cognitive agent finds it difficult to cduomprehend the question. Therefore, it renders an answer that does not relate to the question. A possible counter to this issue could be changing the NMT's model parameters as discussed in [21].

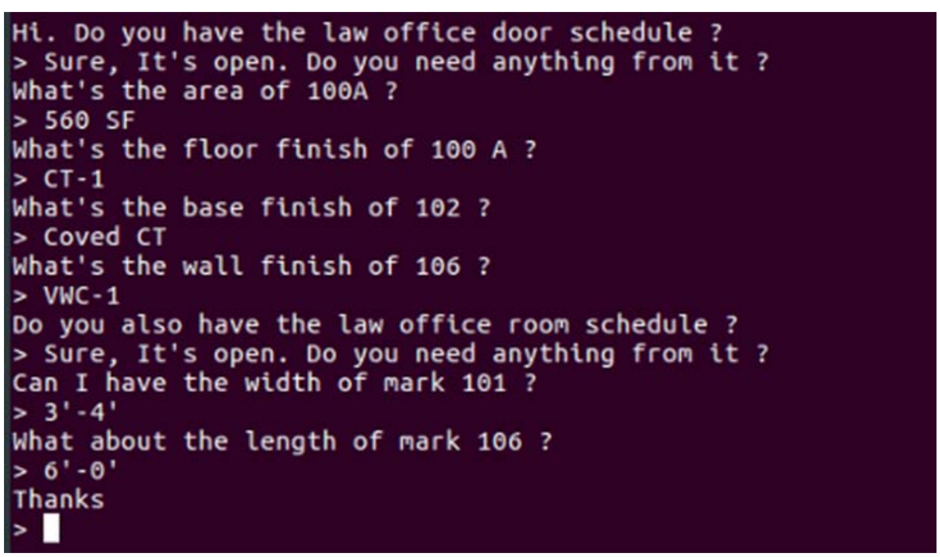

Fig. 19.7 BIM related conversations with BIMBot

(2) Speed: One of the prime reasons to use a cognitive agent is to interact with users for getting desired information instantly. Our Cognitive Assistant is fast and can respond to a user's request in milliseconds.
(3) Functionality: Cognitive Agents should be designed to offer good functionalities to users. BIMBot is designed to retrieve properties from IFCxml models and search if IFC models are found or not. These functionalities help users to easily find the information they are looking for. An example of this is illustrated in Fig. 19.6.

19.5 Conclusion and Future Work

In this paper, we have presented how a cognitive agent could be implemented in the Pre-design phases to increase the efficiency of BIM tasks. Using our agent, architects/engineers work with BIM models in easier fashion by retrieving properties that they desire with ease.

In the future, we would like to venture two interesting paths. The first one of which being, deploying our cognitive agent during a BIM meeting to check the productivity and efficiency it brings about during the Pre-design phase. The second path is to rollout the next version of our cognitive agent which will be an upgrade from the current version. The next version will carry better features like detecting clashes and a better NLP engine with syntactic/semantic analyzer.

References

1. Anon, (n.d.). What is BIM?/NBS. [online] Available at: https://www.thenbs.com/knowledge/what-is-building-information-modelling-bim. Accessed 2018
2. Dossick, C.S., Anderson, A., Azari, R., Iorio, J., Neff, G., Taylor, J.E.: Messy talk in virtual teams: achieving knowledge synthesis through shared visualizations. J. Manag. Eng. **31**(1), A4014003 (2015)
3. Dossick, C., Neff, G.: Messy talk and clean technology: communication, problem solving and collaboration using building information modelling. Eng. Proj. Organ. J. **1**(2), 83–93 (2011)
4. Merschbrock, C., ErikMunkvold, B.: Effective digital collaboration in the construction industry—A case study of BIM deployment in a hospital construction project (2018)
5. Addor, M., Santos, E.: BIM design coordination room infrastructure: assessment of communication activities. In: Computing in Civil and Building Engineering (2014), (2018)
6. Georgia Tech BIM Requirements and Guidelines for Architects, Engineers and Contractors., 1st ed., pp. 1–39. Atlanta, Georgia Institute of Technology (2011)
7. Jack, C.: CHATBOT: Architecture, Design, and Development (2017)
8. Vinyals, O., Quoc, V. Le.: A Neural Conversational Model (2015)
9. Alom, M.Z., Taha, T.M., Yakopcic, C., Westberg, S., Hasan, M., Van Esesn, B.C., Awwal, A.A.S., Asari, V.K. :The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches (2018)
10. Wu, Y., Schuster, M., Chen, Z., Le, Q.V., Norouzi, M. : Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation (2016)
11. Sutskever, I., Vinyals, O., Le, Q.V.: Sequence to Sequence Learning with Neural Networks (2014)
12. Cho, K., Merrienboer, B.V., Bahdanau, D., Bengio, Y. : On the properties of neural machine translation: encoder-decoder approaches. Comp. Sci. (2014a)
13. Google's NMT Documentation—https://github.com/tensorflow/nmt
14. Lipton, Z.C., Berkowitz, J., Elkan, C.: A Critical Review of Recurrent Neural Networks for Sequence Learning (2015)
15. Karpathy, A. : The Unreasonable Effectiveness of Recurrent Neural Networks. http://karpathy.github.io/2015/05/21/rnn-effectiveness/
16. Ramamoorthy, S.: Chatbots with Seq2Seq. Retrieved from http://suriyadeepan.github.io/2016-06-28-easy-seq2seq/
17. Bahdanau, D., Cho, K., Bengio, Y.: Neural Machine Translation by Jointly Learning to Align and Translate (2014)
18. Sennrich, R., Haddow, B., Birch, A.: Neural Machine Translation of Rare Words with Subword Units (2016)
19. Veizenbaum, J.: ELIZA—A Computer Program For the Study of Natural Language Communication Between Man And Machine (1966)
20. Bird, S., Klein, E., Loper, E. : Analyzing text with Natural Language Toolkit
21. Liu, F., Lu, H., Neubig, G. : Handling Homographs in Neural Machine Translation (2018)
22. Yoon, S., Kim, S., Choi, J., Keum, D., Jo, C.: A proposal for using BIM model created in design to construction phase—Case Study on preconstruction adopting BIM. J. KIBIM **5**(4), 1–10 (2015)
23. Shafiq, M., Matthews, J., Lockley, S. : A study of BIM collaboration requirements and available features in existing model collaboration systems. In: International Council for Research and Innovation in Building and Construction (2013)