Mobile Arabchat: An Arabic Mobile-Based Conversational Agent

Mohammad Hijjawi
Department of Computer Science
Applied Science University
Amman, Jordan

Hazem Qattous
Department of Software Engineering
Applied Science University
Amman, Jordan

Omar Alsheiksalem
Department of Software Engineering
Applied Science University
Amman, Jordan

***Abstract*—The conversation automation/simulation between a user and machine evolved during the last years. A number of research-based systems known as conversational agents has been developed to address this challenge. A conversational Agent is a program that attempts to simulate conversations between the human and machine. Few of these programs targeted the mobile-based users to handle the conversations between them and a mobile device through an embodied spoken character. Wireless communication has been rapidly extended with the expansion of mobile services. Therefore, this paper discusses the proposing and developing a framework of a mobile-based conversational agent called Mobile ArabChat to handle the Arabic conversations between the Arab users and mobile device. To best of our knowledge, there are no such applications that address this challenge for Arab mobile-based users. An Android based application was developed in this paper, and it has been tested and evaluated in a large real environment. Evaluation results show that the Mobile ArabChat works properly, and there is a need for such a system for Arab users.**

Keywords—Conversational Agent; Mobile; ArabChat; Chatterbot and Arabic

I. INTRODUCTION

More than 60 years ago, Alan Turing devised the imitation game (Turing Test) to determine if a computer program could think or imitate the human [1]. Using his game, he tried to prove the machine's capability to act as a human or at least imitating the human through conversations. Turing's game summarised with two separated rooms, in the first room there is a human and in the other room there are a human and a machine. The human in the first room starts text-based natural language conversation with the second party (human or machine) in the second room. The second party (human or machine) will reply to his/her conversations. After certain conversations, the first party (human) should decide who was talking with him either the human or the machine. The game does not check the machine's ability to give the correct answers. Instead, it checks how much the answer close to a human's answer. "Turing's game and people's desire to communicate with computers naturally were the best drivers for the creation of conversational agent" [1].

A Conversational Agent (CA) is an intelligent program that attempts to simulate conversations between the human and machine. CAs can is applicable in many applications such as help desk, information retrieval, education, entertainment, E-commerce and other applications.

From Turing's game time, the researchers compete others to build the most intelligent CA. Their CAs were classified as Embodied CA, Linguistic CA or mixing among them. The embodied CA contains a humanoid character that handle conversations with showing body reactions such as human sounds, facial expressions, and movement of its eyes. Where Linguistic CAs deals with written or/and spoken conversations without to embed the embodied communications.

Different approaches can be used to develop a CA; the most successful common used approach is Pattern Matching (PM). The PM is a technique that use an algorithm to handle user conversations by matching a user's utterance (conversation) against the CA's pre-scripted patterns. These patterns represent the text expression of the expected conversations (sentences). A typical pattern consists of a collection of words, spaces, and wildcards. A wildcard is a symbol used to match a portion of the user's utterance. A number of CAs has been developed using the PM such as in ELIZA [2], InfoChat [3-6], ALICE [7], InCA [8] and ArabChat [9]. To best of our knowledge, there is no Mobile based CA has been developed to handle text-based conversations through mobile devices. Most of the above CAs handle text conversations using a personal computer based application. The following are a brief description for some of those CAs.

ELIZA:

A decade after Turing proposed his test; Joseph Weizenbaum developed a program at MIT to simulate the behaviour of a therapist, called ELIZA [10]. Weizenbaum described it as a program that has the ability to make a natural language conversation with a computer [10]. ELIZA works based on pattern matching, using a few decomposition and re-composition rules. These rules replace some pronouns from the user's utterance with other pronouns and embed them into a response. For instance, it replaces the pronoun "I" from the utterance with pronoun "you" at the suitable position in the response. Consider the following example:

User: My boyfriend made me come here

This user's utterance would match the decomposition rule "my 1 me 2". Where 1 and 2 represent the wildcards that match part of the utterance substrings as follows:

1= boyfriend made

2= come here

Consequently, ELIZA runs the re-composition rule "your 1 you 2". Where 1 and 2 are the matched substrings in the previous decomposition rule. As a result, the ELIZA's response is:

ELIZA**:** your boyfriend made you come here

It is possible to notice from the above example that ELIZA tries to ask questions derived from the user's input to give the impression that it is interested in the user conversation. By using this method, ELIZA tries to keep the conversations going for as long as possible. ELIZA matches the user's input against a series of stored patterns. If ELIZA finds a match then, it replies with part of the response taken from the user's input. Otherwise, ELIZA will reply with some fixed responses such as "Very interesting. Please go on" or "Can you elaborate on that".

ELIZA has many drawbacks. One such drawback, it does not use any logical inference to understand the meaning of the user utterances or even determine the patient's topic. Instead, it uses simple string searches and manipulation and thus ELIZA has no grammatical knowledge just make pronouns swapping. ELIZA replies by rephrasing many of the patient's statements as questions and posing them to the patient in an attempt to encourage them to elaborate. Therefore, ELIZA's responses sometimes make a human feel that he/she is speaking to himself/herself. Also, ELIZA cannot keep the continuity of the conversation, and it is not able to store utterances' information, which might be needed during the same session [11].

ALICE:

Dr. Richard Wallace started developing ALICE in 1995, and continuously improved it over the years [12]. ALICE has won the Loebner Prize three times from the year 2000 to 2004 as the most 'human-like' Conversational Agent. ALICE relies on pattern matching technique to handle a user's conversation, and it uses AIML (Artificial Intelligent Mark-up Language) language to script its knowledge base. The following scripts consider as an example of AIML format [13].

<aiml version="1.0">

<topic name="my topic">

<category>

<pattern> USER INPUT </pattern>

<that>THAT</that>

<template>CHATBOT ANSWER</template>

<category>

<topic>

<aiml>

AIML, which is a specification of XML (Extensible Mark-up Language), organises the scripted knowledge into AIML files. Each AIML file starts with the <aiml> tag to indicate the AIML used version and also includes a series of AIML units called topics and categories. The 'topic' is an optional element (<topic> tag), which has a name, and groups set of categories related to that topic together. Each category (<category> tag) contains a pattern that is matched with the user's utterance and a template that formulates a reply by that category.

ALICE starts normalising the input before matching it. Normalisation attempts to remove punctuation and convert lowercase letters to uppercase. Then, ALICE starts matching the input word by word depending on the depth first search technique to obtain the best pattern matching. The best matching is the longest pattern match in terms of number of keywords [13]. The depth first search finds the first available matched pattern regardless of the possibility of other pattern matches availability, which means it does not guarantee to select the best match [14]. If there is no match, ALICE follows the ELIZA methodology by asking simple questions on common issues to keep the conversation going.

ArabChat:

ArabChat is a Conversational Agent developed in 2012 to handle conversations for the Arabic language [9]. When ArabChat built, it has been taken into consideration two major factors. The first factor, taking the nature of the human conversation that may start with a specific topic and naturally switch between topics. The second factor is considering the Arabic language morphological nature. Also, sometimes speakers might have to remember some spoken topics or the need to retrieve some parts of spoken utterances and to track the sequencing of subtopics.

The ArabChat uses the PM approach for handling the Arabic textual conversations where the ArabChat's framework consists of three main components that are the scripting language, engine and brain. The scripting language used to script the applied domain topics to represent them. ArabChat scripting language is a rule-based language, which depends on a rule structure to handle the expected Arabic conversation. ArabChat scripting language can structure any applied domain into a set of contexts (topics), where each context has many rules. A rule (sub-topic) has many patterns and associated responses. While, the ArabChat's brain is the CA's knowledge base that is used to store the domain's scripts where the engine handles the Arabic user's conversations that target the scripted domain.

The following is an example of a rule in ArabChat:

<AlreadyPaid>

a: 0.4

p:15 have paid already.

p:15 *payment is on @ way.

p:15 *have paid already.

p:15 I have @ paid.

r: Ok. That is fine and thank you

The rule is to confirm that a payment is already paid. Each rule has a unique name ("AlreadyPaid"), and a decimal value (0.4) called "base activation level" that is used to calculate the rule's strength. This strength is used by ArabChat to differentiate between competitor-matched rules to select the best one. The rule that has the highest strength will fire. The rule has many patterns that deal with utterances to confirm that

the payment was paid. A pattern in ArabChat is a collection of characters, spaces and/or wildcards. Each pattern in the mentioned rule has a base strength (p:15), which it used in the pattern strength matching calculation. Then, the calculated matched pattern strength will be inherited to the rule that this pattern belongs to, to compete with other rules.

After firing a rule, the ArabChat enables a scripter to increase the chance for firing other rules for the next expected utterance by promoting them. Such rules might be related to the fired rule, and they are expected to be targeted by the user after the processed utterance. Promoting a rule means increasing the chance of a specific rule to be fired by increasing its activation level. In contrast, ArabChat can degrade the possibility of other rules being fired (after firing a rule) by decreasing their activation level to the minimum (demoting rules). The ArabChat's scripter can kill rules after firing a rule to prevent them from being fired. Also, ArabChat can manage the navigation between contexts through scripted actions. These scripted actions have the ability to forward the processed utterance to other contexts for further processing or move the agent to another context and wait for the next expected utterance.

InCA:

InCA is an assistant conversational character runs on a handled PDA [8]. InCA uses facial animation and speech input/output to handle user's spoken conversations to provide some services such as appointments and weather reports. In this work, the paper discussed the InCA's architecture with focusing on two limitations of a mobile device platform that are limitation computational power and the input module restriction. The InCA contains three main components; the client that runs on the PDA, the server that manages speech recognition and speech syntheses and the third one is the coordinator who is responsible real-time data retrieval. This conversational agent is not text-based CA, and there are no evaluation experiments on it discussed in the mentioned wok [8].

Most of the above CAs used the Pattern Matching approach for handling the conversations. The Pattern Matching approach showed the impression of some intelligence when it used to handle conversations [15]. The PM approach has many features, including that it is easy to understand, and it is a natural language independent. Also, PM based CAs do not require complex pre-processing stage that might require analysing the natural language sentences that require extra time to process. As a result, the PM is not expensive computationally. Therefore, PM based CAs can handle conversations efficiently for large numbers of users in a real-time environment like the Internet [15]. Moreover, the PM resolves a lot of linguistic challenges such as morphological changes (changes occurring in a word when adding affixes to it). These changes can be resolved using the PM through the pattern's wildcards. Also, the pattern's wildcards can resolve grammatical and spelling errors in a user's utterance. Resolving the spelling and grammatical errors from a user's utterance is an important task that help in continuing the conversations between the user and the CA [16]. From the other side, all considerations that the PM faces are the large number of patterns that the scripter should script to cover the CA's domain. However, the PM approach can minimize a large number of needed pattern to the minimum using variety of wildcards. As a result, this paper chooses the Pattern Matching approach to building the Arabic Mobile CA based on the above supporting features. Also, the Mobile device platform still has limitations in the computational power, displaying and wireless network bandwidth that lead to choosing such a light approach.

II. Need for Mobile Arabic CA

In the last decades, the technology has increasingly evolved in a large spread manner. Researchers tries to automate everything through the evolved technology. This automation has different aspects from different fields such as in education, business, and entertainment and even in social societies. For instance, customers for a company prefer to use an online help desk rather than using the traditional help desk through dealing with the company's employees [17]. Using such online help desk has number of features for customers like availability and cost effective. In addition, it is beneficial for an organisation by reducing the operational costs and documenting all customers' requests for easy analysis and data mining.

One of the most technology trend that has been evolved increasingly in the last decade is mobile smartphones development. Mobile technology has the ability to change the way of people communication and impacting positively in their societies and economies [18]. The world has about 7 billion mobile subscriptions which almost equivalent to the number of people in the planet [19]. This caused to change the way that people dealing with business, learning, socializing and with the government transactions [19]. Arabs are not in isolation from this technology evolution. In the Arabic region, the Mobile communication has expanded rapidly and this can be appeared from the number of mobile subscriptions which has been remarkably increased from 126 million to 350 million in the period of 2006 to the 2011 [20]. According to [19], "For more than a decade, an increasing number of Arab countries have realized that the 'knowledge economy' fuelled by wide adoption and availability of information and communication technologies infrastructure will play an essential role in growth and development". The large number of mobile subscriptions and the 'knowledge economy' led the industry to focus more in smart phones manufacturing and thus in its applications development. 42.07% of Arab users use their smartphones to access the internet instead the personal computer or the laptop and 23.71% use tablets to access the internet [19]. Around 52% of Arab users spending from three to seven hours on the Internet daily from different places like home, work, school, mall and in Internet café [19]. Around 47% of Arab users have from 1 to 25 applications in their smartphones [19]. The smartphone penetration in the Middle East region will grow up to 39% by year 2015 and in Jordan by 50% for the same year [21]. According to [19] and [21], around 40% of Arab users use an android based smartphone.

All of these was the driver to develop the Mobile ArabChat for Arab users in general and for Jordanian people specifically in this paper using the Android. The Mobile ArabChat is an intelligent system that can be used in many applications as mentioned above. In this paper, it has been select the Mobile

ArabChat to work as Information Point advisor for a university's students. However, Arab users still have some challenges in using the Internet. Some of these challenges are shown in Figure 1 [19]. A 48.40% of Arab users found that they suffering from some weaknesses of using the internet in terms of its accessibility and connectivity. In addition, as shown in the figure, 44.68% of users were suffered from the cost and 40.75% are suffering from the lack of the Arabic content on the Internet. Moreover, 30.24% of users faced a limited bandwidth when using the Internet in general. For these reasons, it has been selected to propose and develop the Mobile ArabChat in this paper to handle text-based conversations as the text conversations has small size and limited cost when utilising the bandwidth.

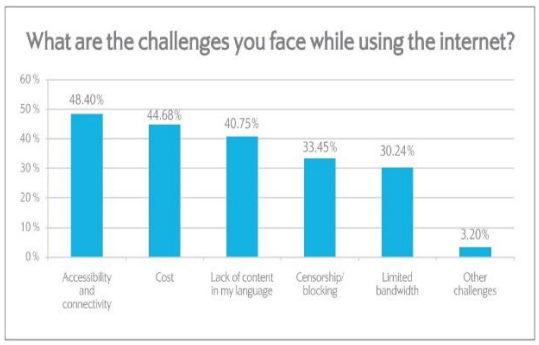

Fig. 1. Arab users challenges when using the Internet[taken from [19]]

III. Mobile ArabChat Framework

The Mobile ArabChat is modelled based upon the ArabChat CA [9]. A new framework has been proposed and developed to issue the Mobile ArabChat to suite the limitations of the mobile device platform. The Mobile ArabChat framework is a rule based CA and contains of number of integrated modules (as depicted in Figure 2) which are scripting engine, scripting language, user interface, temporal memory and knowledge container brain.

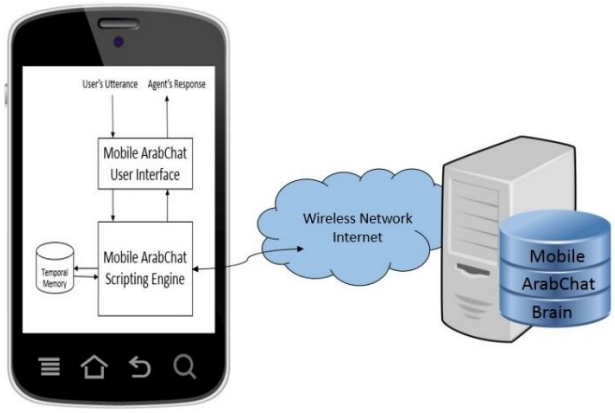

Fig. 2. the Mobile ArabChat's framework

The Mobile ArabChat scripting engine considered as the core module in this framework. Mainly, it is responsible to handle users' conversations by matching them against the scripted patterns. However, it do number of other tasks such as validating the user's utterance to be sure it is in Arabic and valid before proceeding it to the engine to be processed. In addition, the scripting engine can manage and control the conversations especially when it switch among different topics. Moreover, it can encapsulates the Mobile ArabChat's response with some captured information from the user (when needed). The Mobile ArabChat has a temporal memory to store a captured portion of a user's utterance in order to use it later in the Mobile ArabChat's response which give a good impression at a user. In addition, the Mobile ArabChat can save the processing effort and time and clear an ambiguous when the user targeting the same rule indirect by entering for example "I don't understand you" or something similar by re-firing the previous fired rule with generating another response for the same fired rule. The previous fired rule parameters are stored in this temporal memory and the engine checking it before proceeding with the utterance. Before proceeding with the full framework explanation, it is important to explain the hidden part of this framework which is the Mobile ArabChat scripting language. The following subsections explain main modules of the Mobile ArabChat's framework:

A. The Mobile ArabChat scripting language

The scripting language used to script the applied domain topics that covers the Mobile ArabChat service's topic where the engine handles the Arabic user's conversations that target this service's topic. The Mobile ArabChat's brain is the CA's knowledge base that is used to store the domain's scripts.

The Mobile ArabChat uses the PM approach to handle the Arabic textual conversations according to the framework that depicted in figure 2. As mentioned above, the scripting language used to script the applied domain topics in order to represent them. The Mobile ArabChat scripting language is a rule-based language, which depends on a rule structure to handle the expected Arabic conversations. The Mobile ArabChat scripting language structures any applied domain into a set of contexts (topics), where each context has many rules. A rule (sub-topic) has many patterns and associated responses. The Mobile ArabChat scripting language categorises the applied domain's topics through contexts. The domain is the area's topic that the CA will help in such as handling conversations for a company as a help desk employee. Where, a rule is a sub-topic of a context that a user might target in his/her utterance, whereby a pattern is a representation of that utterance which belongs to such a rule. Finally, the responses are the reply to the user's utterance.

In order to handle the conversations, the Mobile ArabChat's scripter should feed/script the applied domain into the Mobile ArabChat's brain categorised as contexts with related rules, patterns and responses. Each context has number of rules to represent the topics inside that utterance. In addition, a context has a "Default rule" to be fired when the utterance targeting the context's topic but no rule in that context has matched that utterance. For instance, if the applied domain was to cover conversations about the tourism places in Jordan, so the contexts might be Archaeological sites, antiquities bazars and therapeutic places such as the Dead Sea. The archaeological sites context might has several rules (sub-topics) such as Petra and Jerash cities. When a user targets tourism in Jordan in general, the engine will reply to the user with a general response about the tourism by firing the rule "default" that belongs to the main context. In contrast, when

the user targets the tourism in Jordan in the Petra city, the engine will reply with a response related to the Petra city and according to his/her utterance by firing the "Petra" rule. The Mobile ArabChat offers number of scripting features that manage the scripting process and control the switching between the scripting topics in order to script a coherent domain. The following considers an example of a rule's structure:

<RegistrationFees>
<رسوم-التسجيل>

a: 0.2
القيمة الإستنادية: 0.2

p: 15 * fees $ registration *
نمط المحادثة: * رسوم $ تسجيل *

p: 15 $ fee * registration #
نمط المحادثة: $ رسوم * تسجيل #

p: 15 $ fee% * registration%
نمط المحادثة: $ %رسوم * تسجيل *

P: 15 * fees $ registration% *
نمط المحادثة: * رسوم $ تسجيل *

الرد: تبلغ رسوم التسجيل 25 دينار كطلب إلتحاق بالجامعه غير مسترده و 300 دينار أردني كرسوم تسجيل فصلية.

r: The registration fee is 25 JD (Jordanian Dinars), paid once and it is non-refundable and 300 JD registration fees for each semester (term).

The above rule (Arabic scripts and translated English scripts) deals with a sub-topic of an applied domain which concerns of the fees of the university registration for a new student. The rule has number of elements which are; rule name ("Registration-Fees"), base activation level (0.2), number of different patterns (started with p:) and a response (started with r:). Each rule has a unique name and a decimal value called "base activation level" which is used to calculate the rule's strength. This strength is used by the Mobile ArabChat to differentiate between competitor-matched rules to select the best matched rule. The rule that has highest strength will fire. A rule has many patterns to deal with upcoming utterance by matching it with the scripted patterns. A pattern in the Mobile ArabChat is a collection of characters, spaces and/or wildcards. Pattern's wildcards that appeared in the above example has different purposes as described below:

Pattern's wildcards types:

1) The wildcard symbol "$" is used to match or represent one word.

2) The wildcard symbol "*" is used to match or represent many words.

3) The wildcard symbol "%" is used to match or represent one character.

4) The wildcard symbol "#" is used to match or represent one digit.

Each pattern in any rule has a base strength (for example p:15), which it used in the pattern strength matching calculation. Calculation the pattern strength depends on number of factors such as the pattern's base strength value, number of matched keywords and the length of the user's utterance. Then, the calculated matched pattern strength will be inherited to the related rule and considered as the new rule's strength and competing with other rules. After firing a rule, the Mobile ArabChat enables a scripter to increase the chance for other rules to be fired for the next expected utterance by promoting them. Such rules might be related to the fired rule and they are expected to be targeted by the user after the processed utterance. Promoting a rule means increasing the chance of a specific rule to fire by increasing its activation level. In contrast, Mobile ArabChat can degrade the possibility of other rules being fired (after firing a rule) by decreasing their activation level to the minimum (demoting rules). The Mobile ArabChat's scripter can kill rules after firing a rule in order to prevent them from being fired. In addition, the Mobile ArabChat can manage the navigation between contexts through scripted actions. These scripted actions have the ability to forward the processed utterance to other contexts for further processing or move the agent to another context and wait for the next expected utterance.

B. The Mobile ArabChat User Interface

The Mobile ArabChat user interface manages displaying the conversations among the two conversations parties as appear in Figure 3. The Mobile ArabChat user interface has been developed using the Android technology which it is now holding more than 40% of Arab mobile smartphones [19, 21]. Designing the user interface was simple and user friendly to accomplish its function with the minimum smartphone resources and with a limited Internet bandwidth usage. Each conversation party has different location side in the interface as most of the mobile based chatting application. In addition, each conversation party utterances has different colours as appeared in Figure 3.

Fig. 3. The Mobile ArabChat user interface

C. The Mobile ArabChat scripting Engine

The Mobile ArabChat scripting engine has number of integrated functions work together in a novel structure. The Mobile ArabChat scripting engine is the core of this framework

and its work according to the PM technique in order to handle users' conversations over the mobile platform.

Once the scripting engine receives an utterance, it starts matching it with patterns of the rules that belong to the current context until all rules of the current context are processed. Initially, the rule that has the highest strength might be fired. However, during the matching process, the engine differentiating between the matched patterns depending on their calculated strengths. The matched pattern that has the highest strength which means having a better match between the pattern and the utterance will inherits its strength to its related rule. Then, the engine will take this calculated pattern strength value and inherited to its rule in order to enable it to compete among other rules in the current context.

In case of no pattern matching the utterance, the scripting engine takes another chance to match the utterance with the patterns of the previously processed context as a precaution step. This precaution step is adopted by the engine to meet the nature of conversations as the consecutive utterance might not have enough explanation in it, by assuming that the previous utterance is already has. If the matching was occurred, the engine will start matching the utterance with the previous context's patterns. Otherwise, a "Default Rule" will fire. A "Default Rule" (DR) is a normal rule but with a higher activation level value (assigned by a scripter) than other rules. A rule that has the highest activation level value means it has a highest strength. A DR's response usually represents a general response for the whole context. Finally, the highest rated rule (regardless if it is a DR or not) will be fired after the conversations' manager is supplied with the needed control information determined by running the fired rule's actions to let the engine moves for which context or to let it remain at the current context and waiting for the next utterance.

D. The Mobile ArabChat Brain:

The Mobile ArabChat brain considered the structured store container that the CA's scripter should feed it with the needed structured scripts to prepare the Mobile ArabChat to handle the Arabic conversations. The Mobile ArabChat brain depends on a DBMS (Data Base Management System), to enable a scripter of doing a familiar scripting, searching, querying, and reporting. The scripter should understand the selected domain before start scripting. Then, the scripter start classifying the applied domain into contexts and associated rules. Finally, he/she should script the needed scripts for each rule in terms of the needed variety of patterns and the suitable responses. The Mobile ArabChat offers number of a friendly interfaces to be used by the scripter to script the selected domain. In addition, it offers other facilities such as logs to store the processed conversations and the non-processed conversations in order to track it and analyse it later.

IV. EVALUATION

The Mobile ArabChat evaluation methodology is comprised of two main approaches: namely, objective approach and subjective approach. The objective approach will be applied using an light automatic evaluation method called RMUT(Ratio of Matched Utterances to the Total) [9] and manually through analysing the Mobile ArabChat logs. Where, the subjective approach will be performed with recourse to human judgment using the user's questionnaire.

The Mobile ArabChat was deployed on the ASU (Applied Science University) local server and 57 students has been asked to deploy the Mobile ArabChat system into their mobiles handsets. Through the ASU internal wireless network, the students can able to access the Mobile ArabChat. The selected scripted domain is to handle Arabic conversations related to two aspects which are the courses fees of ASU and the total credit hours for each course in ASU. The Mobile ArabChat handled 743 utterances from the 57 users with an average of 13 utterances per user. The following subsections discuss the evaluation from the two approaches side:

A. The objective approach evaluation

This evaluation aims to test the Mobile ArabChat performance and the functionality of its main components; mainly the engine. This evaluation will determine whether or not the scripting engine is doing its tasks properly such as recognising patterns' wildcard, matching utterances successfully and the ability to navigate among the scripted contexts. The most challenge in CAs is how to evaluate the CA or to calculate a user satisfaction automatically [14]. However, automating such a task is complex as an utterance has very rich linguistic information especially for a sematic language such as the Arabic.

This evaluation will be done automatically by determining the RMUT (Ratio of Matched Utterances to the Total) of the Mobile ArabChat users. The RMUT is automatically calculated per user by the Mobile ArabChat once a user session is closed and it is automatically calculates the ratio of number of matched utterances to the total utterances for each user (a user's session) based on the following equation [9]:

$$\text{RMUT} = \frac{\text{Number of matched utterances}}{\text{Total number of utterances}} \quad (1)$$

The evaluation results show that the average RMUT for the 57 users of the Mobile ArabChat is 83.2%. This means that around 83% of users' conversations has been matched. Initially, this result might be good but for most accurate result a manual analysing for the Mobile ArabChat logs must be done. The meaning of a matched utterance does not always equal to a successful response to a conversation. When the Mobile ArabChat matching an utterance and fired a related rule, the conversation can be considered as a success conversation. But when it matched an utterance and fired unrelated rule, the conversation consider unsuccessful as the reply to the user was incorrect. This conflict might be occurred when two rules shared the same keywords or shared part of the scripted topic. According to [9], the RMUT is not expected to give a full picture about the user satisfaction, but it used to test whether or not the scripting engine is performing its tasks properly in terms of its ability of matching utterances successfully and gives overview of the quality of scripts. On the other hand, it gives a general overview of scripting engine's performance [9].

After the manual analysing job for the logs has been finished, it has been revealed that the actual percentage of successful conversations is 78.64%. This percentage is more accurate and this is caused due to firing unrelated rules or

speaking with the Mobile ArabChat outside the scripted domain and that's appeared after the manual testing. However, this reducing in the percentage does not mean any fault in the engine's work but it is clear for the need of an extra effort in the scripting process itself and a user's commitment to speak inside the selected topic. Given this, it can be conclude that the Mobile ArabChat achieved a reasonable performance by its ability of handling/automating the Arabic conversations successfully.

B. The subjective approach evaluation

The subjective approach evaluation will be performed with recourse to human judgment using the user's questionnaire. The evaluation will be conducted by asking Mobile ArabChat users to give their opinion about various aspects of using it. In addition, the evaluation aims to ask the users about their satisfaction of using such a mobile based CA instead of using a desktop based CA. The subjective evaluation aims to enable users to evaluate the Mobile ArabChat user interface, usability, naturalness, the applied domain coverage, speed, availability of similar mobile based Arabic agent, and user general satisfaction. The questionnaire has 15 questions designed to meet the above mentioned evaluation aims. For each aim, a number of questions have been assigned to determine the user opinions concerning them. For each question in the questionnaire, a user has 3 options from which to select his/her degree of approval or disapproval for the asking issue. These options are "موافق" ("Agree"), "محايد" ("Neutral"), "غير موافق" ("Disagree"). The following are the questionnaire questions:

- "واجهة النظام كانت مناسبة جدا" "The user interface was suitable".
- "كان النظام قادر على إجابتك على جميع إستفساراتك" "The agent was able to answer all your utterances".
- "أجوبة النظام كانت واضحة ومفهومة." "The agent responses were clear and understandable".
- "لم تواجهك أية مشاكل فنية عند إستخدامك النظام" "You experienced no technical problems whilst using the agent".
- "الوقت المستغرق من النظام للرد على استفساراتك كان مناسبا" "The elapsed time taken by the agent was reasonable".
- " تفاعل النظام معك كان واقعي وحقيقي شبيه بتفاعل الانسان من حيث الأجوبة وردود الفعل" "The interaction with the agent was realistic and believable".
- " صعوبة التخاطب مع الجامعة عبر الهاتف والبريد الالكتروني وصعوبة الوصول لمعلوماتك المطلوبة عبر موقع الجامعة الالكتروني جعلك تلجأ لإستخدام هذا النظام" "The difficulty of contacting the university by phone or email, and accessing your needed information on the university website were the reasons to use the Mobile ArabChat".
- "لقد ساهم النظام في توفير جهدك و وقتك." "The agent saves you time and effort".
- " لايوجد خدمة مثيلة باللغة العربية لأي جامعة ,كلية, أو لشركة و أيضا لايوجد نظام اسئلة وأجوبة باللغة العربية عبر الموبايل" "There is no Arabic university, college or company offering the same services over the Mobile platform, even there is no question answering system in Arabic for Mobile users".
- "كان النظام يشجعك بالاستمرارعلى الحديث" "The agent encourages you to carry on with the conversation".
- "تقييمك الإجمالي للنظام بأنه ممتازا" "Your overall rating for this service is excellent".
- "سوف تنصح أصدقائك باستخدام هذا النظام" "You will recommend your friends to use the Mobile ArabChat system".
- " أنت تفضل استخدام هذا النظام بدلا عن التحدث مع الشخص المسؤول في الجامعة" "You prefer to use Mobile ArabChat rather than speak with a human advisor".
- "سوف تعيد إستخدام النظام في المستقبل" "You will re-use this service in the future".
- " هل تفضل استخدام هذا النظام عبر هاتفك الذكي عن استخدامه مثيل له عبر جهاز الحاسوب؟" "Do you prefer to use the Mobile ArabChat than using such a system on a personal computer"

1) ArabChat PH2 subjective evaluation results

All users have been answered all questions and submitted their questionnaire and the results as shown in Table 1.

TABLE I. THE MOBILE ARABCHAT QUESTIONNAIRE RESULTS

Question number	"Agree" distribution (Percent)	"Neutral" distribution (Percent)	"Disagree" distribution (Percent)
1	53 (92.9%)	2 (3.5%)	2 (3.5%)
2	47 (82.4%)	7 (12.2%)	3 (5.2%)
3	41 (71.9%)	8 (14%)	8 (14%)
4	56 (98.24 %)	1 (1.75%)	0 (0%)
5	54 (94.7%)	2 (3.5%)	1 (1.75%)
6	15 (26.3%)	7 (12.2%)	35 (61.4%)
7	57 (100%)	0 (0%)	0 (0%)
8	51 (89.47%)	5 (8.77%)	1 (1.75%)
9	57 (100%)	0 (0%)	0 (0%)
10	33 (57.9%)	20 (35.1%)	4 (7%)
11	42 (73.68%)	12 (21%)	3 (5.2%)
12	41 (71.9%)	7 (12.2%)	9 (15.8%)
13	52 (91.2%)	5 (8.77%)	0 (0%)
14	51 (89.47%)	4 (7%)	2 (3.5%)
15	55 (96.5%)	2 (3.5%)	0 (0%)

According to the above subjective evaluation results that shown in Table 1, the mobile based user interface was evaluated using the first question and 92.9% of users agreed that the user interface was suitable. Where the Mobile

ArabChat usability was evaluated through three questions in the questionnaire which are 4, 7, and 8. 98.24% of users agreed that they experienced no technical problems while using the Mobile ArabChat. 100% of users agreed that difficulty contacting the university by phone or email, as well as difficulty accessing their needed information on the university website were the reasons that caused them to use the Mobile ArabChat. Finally, 89.47% of users agreed that the agent saved them time and effort.
The naturalness of the Mobile ArabChat has been evaluated through three questions: 3, 6, and 10. 71.9% of users agreed that the ArabChat's responses were clear and understandable. Only 26.3% of users mentioned that ArabChat's interaction was realistic and believable. 57.9% of users disagreed with the notion that ArabChat encouraged them to carry on with their conversation. This inability to encourage further conversations might be due to the response scripting, which fails to encourage users to continue conversations after firing certain rules.

The applied scripted domain for Mobile ArabChat is simple and it is used only to test the CA. However, 82.4% of users (question number 2) agreed that Mobile ArabChat was able to provide all of their requested information regarding the two covered scripted topics as mentioned above. Regarding the Mobile ArabChat interaction speed the interaction speed of ArabChat has been evaluated through item number 5. 94.7% of users agreed that the elapsed time taken by the Mobile ArabChat to handle their utterances was reasonable. Where the availability of similar Mobile CAs was evaluated through item number 9. All users agreed that there is no Arabic university, college or company offering the same service. This high percentage carries a meaning behind it which the Mobile ArabChat might be considered the first Mobile CA responsible for handling user utterances in the Arabic language. The general satisfaction of the Mobile ArabChat users was evaluated through item numbers 11, 12, 13, and 14. 77.2% of users agreed that their overall rating for ArabChat was excellent, while 71.9% agreed to recommend Mobile ArabChat to their friends. 91.2% of users prefer to use ArabChat rather than speak to a human advisor and 89.47% of users confirmed they would use ArabChat for future needs. Finally, Most of users (96.5% of them) prefer to use the Mobile ArabChat instead of using such a system via their personal computers. This means that the good Mobile penetration ratio that affects the Arab countries changed their life of style indeed by depending on their smartphones for all types of communication.

V. CONCLUSION

This paper has discussed the mobile based Arabic Conversational Agent called Mobile ArabChat. The Mobile ArabChat framework comprises mainly of a novel scripting engine and a rule-based scripting language structured in a novel way to handle topics of the conversations. Topics of conversation classified into contexts. Each context contains rules that themselves consisted of patterns and associated textual and action-based responses. The Mobile ArabChat handled user's conversations using the pattern's matching technique, by matching the user's utterance against scripted patterns through navigation the utterance into the novel scripts structure. The matched patterns that belong to different rules compete among each other based on their matching strengths. The pattern that has the highest strength will inherit its strength value to its rule and thus it will be fired. The Mobile ArabChat differentiates between matches to select the best match (the rule that has the highest strength) that represent the conversation's topic. The Mobile ArabChat has the facility to navigate among topics through the scripted scripts. The applied domain in this paper was simple and for evaluation and testing purposes. However, the Mobile ArabChat shows a good accuracy from the both evaluation approaches sides; objective and subjective. From the objective side, using the RMUT and logs manual analysis showed that it handled well 78.64% of the conversations. This figure might reflect the general user's satisfaction and the Mobile ArabChat's performance. From the other side (the subjective), 73.68% of users who filled the questionnaire agreed that the Mobile ArabChat was excellent. Also, 96.5% of users found that using the Mobile ArabChat on their smartphones better than using the same system through their personal computers. This due to the flexibility of the mobile platform.

ACKNOWLEDGMENT

The authors are grateful to the Applied Science Private University, Amman, Jordan, for the full financial support granted to this research.

REFERENCES

[1] TuringFIT Facebook group, A., Computing machinery and intelligence. Mind, 1950: p. pp 433-60.

[2] Weizenbaum, J., ELIZA: A computer program for the study of natural language communication between man and machine. Communications of the ACM, 1966. 10: p. 36-45.

[3] ConvAgent. Business Rule Automation Knowledge Management Conversational Agent. 2005 [cited; Available from: www.convagent.com/what-we-do.html.

[4] ConvAgent. ConvAgent Foundation- ADAM Conversational Agent. 2015 [cited; Available from: www.ConvAgent.com.

[5] Sammut, C., Managing Context in a Conversational Agent. Electronic Transactions on Artificial Intelligence, 2001.

[6] Sammut, C. and D. Michie, InfochatTM Scripter's Manual, Convagent Ltd. . 2001: Manchester.

[7] Wallace, R. ALICE: Artificial Intelligence Foundation Inc. . 2008 [cited; Available from: http://www.alicebot.org.

[8] Kadous, M. and C. Sammut, InCA: A Mobile Conversational Agent. Trends in Artificial Intelligence, 2004. Volume 3157: p. pp 644-653.

[9] Hijjawi, M., et al. ArabChat: An Arabic Conversational Agent. in Computer Science and Information Technology (CSIT), 2014 6th International Conference on. 2014. Amman, Jordan.

[10] Weizenbaum, J., ELIZA: A computer program for the study of natural language communication between man and machine. Communications of the ACM., 1966. Vol 10.: p. PP 36-45.

[11] Block, N., The mind as the software of the brain, in An Invitation to Cognitive Science, D.N. Osherson, et al., Editors. 1995, MIT Press.

[12] Wallace, R.S. ALICE: Artificial Intelligence Foundation Inc. 2008 [cited; Available from: http://www.alicebot.org.

[13] Abu Shawar, B. and E. Atwell, A Comparison Between Alice and Elizabeth Chatbot Systems. 2002. p. 21.

[14] Goh, O., A framework and evaluation of Conversational Agents, in Information Technology. 2008, Murdoch University.

[15] Timothy, B. and G. Toni, Health dialog systems for patients and consumers. J. of Biomedical Informatics, 2006. 39(5): p. 556-571.

[16] Maragoudakisa, M., et al., Natural Language in Dialogue Systems, a case study on a medical application. , in Proceedings of Panhellenic

Conference with International Participation in Human–Computer Interaction. 2001: Greece. . p. 197–201.

[17] University of Alberta: Using Online Help Desk Tools to Enhance Client Service and Department Operations in ECAR-EDUCAUSE CENTER FOR APPLIED RESEARCH. 2007.

[18] LLP, D., Arab States Mobile Observatory 2013, p.b.D.L.f.t.G.A. (GSMA), Editor. 2013. p. 68.

[19] The Arab Social Media, G.a.I.P., Mohammed bin Rashid School of Government, The Arab World Online 2014: Trends in Internet and Mobile Usage in the Arab Region. 2014.

[20] Zuehlke, E., In Arab Countries, Mobile Internet and Social Media Are Dominant, but Disparities in Access Remain, in Population Reference Bureau. 2012.

[21] Go-Gulf, Smartphone Usage in the Middle East- Statistics and Trends, in Go-Gulf. 2013.