Indonesian Journal of Electrical Engineering and Computer Science
Vol. 28, No. 1, October 2022, pp. 611~624
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v28.i1.pp611-624

Inheritance issues' features extraction using Arabic text analyzer (IFAA)

Abeer Khalid Al-Mashhadany[1], Marwan B. Mohammed[2], Mawlood Mukhlis Alrawi[3]
[1]Department of Computer Science, College of Science, Al-Nahrain University, Baghdad, Iraq
[2]Department of Computer Science, Al-Nahrain University, Baghdad, Iraq
[3]Department of Jurisprudence and Its Fundamentals, Imam Al-Azam University College, Baghdad, Iraq

Article Info

Article history:

Received Feb 20, 2022
Revised Aug 1, 2022
Accepted Aug 10, 2022

Keywords:

Arabic NLP
Features extraction
Heirs table
Text analytics
Text mining

ABSTRACT

Inheritance issue is part of our life. Daily, many persons may die. Person is gone, while his money stays for others. Islamic law took care of the issue of inheritance. Al-Quran has verses dedicated in inheritance issue. Al-Quran gives every person its rights, so; the share for each heir is determined. Islamic law jurists are asked frequently to solve inheritance issues. This work; inheritance issues' features extraction using Arabic text analyzer (IFAA) hopes to analyze inheritance issue. It receives the issue as Arabic unstructured characterized text. It applies Arabic analyzer system to extract all features. Many commercial applications are constructed to solve inheritance issue; they receive the features manually, while this work is an attempt to computerize features' extracting. This work needs a good experience in analyzing Arabic text. So, this research attempts developing Arabic analyzer system dedicated in inheritance issues, which has the ability to analyze inheritance issue and extract its features. It will be shown that Arabic analyzer system is useful in converting Arabic text into data that are understandable by programming languages, and those data could be used to perform arithmetic calculations, and achieve high accuracy reaches to 100%.

Corresponding Author:

Abeer Khalid Al-Mashhadany
Depertment of Computer Science, College of Science, Al-Nahrain University
Baghdad, Iraq
Email: aabeeeeraa@yahoo.com

1. INTRODUCTION

In Islamic law inheritance science is called Al-faraid science علم الفرائض. In Arabic, the word Faraid فرائض is plural of the word Frida فريضة, which is a share determined by Islamic law for each one of the heirs. Al-faraid science refers to everything related to inheritance. It is the jurisprudence of inheritance and mathematics which is necessary to know the right of each one of the heirs. Islamic law took care of the issue of inheritance and considered it as important part of Islamic jurisprudence. The reason for interesting inheritance is that it deals with the affairs of life that related to money, which is the nerve of life [1]-[6]. Many references arranged tables to simplify understanding the science of Al-faraid [1], [2], [5], [7]. Such tables must include all possible heirs. Mostly the share for each heir is found. Some tables show the verse for each case. Really all researchers on this domain depend on verses of Al-Quran. All heirs and their relationships are found in Al-Quran; verses; 7, 11, 12 and 176 in Sura Al-Nisa [7].

Al-faraid science deals with inheritance issue which may be found as a written text. Usually, people describe the case of inheritance by short text, which is unstructured text. Natural language processing techniques analyze unstructured text to extract data and information from its content. Text mining, in other words; text data mining is the intelligent text-analysis. It is an example of Natural language processing

techniques. It analyzes unstructured text and extracts information from it. So it is one of the most complex analysis tasks. Inheritance issues texts mostly written in Arabic language. Arabic unstructured text analysis is faced with many challenges because of the complexity of this language. It is morphologically rich language [8]-[14]. Implementing such job needs high skills in domain of Arabic text analysis. Many text analysis systems are proved in previous works. Arabic diagnosing expert system shell (ADESS), Knowledge Acquisition and hybrid inference with a stem-based approach (KISB), and root-stem approach in general analyzer system for Arabic language (RSGAS) provide analysis systems to prevent; typos, duplication, and inconsistent rules [15]-[17]. They were successful in discovering the wrong words and the similarity in meaning for phrases. Other previous work [18] merged statistics and many linguistic approaches to perform keyword extraction. Other one [19] performed text summarization. The previous work [20] had built high accurate method for testing text similarity. The previous work extracting numerical data from unstructured Arabic texts (ENAT) extracts numeric data from Arabic unstructured text [21].

Any inheritance problem includes specific features. The main feature that must be found is heirs and their relationship with died person, gender of each one is very important to calculate the exact share. Some problems may include other features such as gender of died person, marital status, and information about the inheritance [6], [22]. To clarify these features 60 examples are collected, see Figure 1. Heirs are marked with double under line, while a dead is marked with single under line. The inheritance, when found is marked with bold under line.

1. ماتت امرأة وتركت زوجا والأب وأخ ولها خمسة عشر ملايين
2. توفي رجل عن والدين وأخ وزوجة وله تركة
3. توفي عن أخ لأبوه وأخت لأب وعم وله مليون دينار
4. توفت امرأة عن بنت شقيقات والزوج وأب وجد
5. توفي رجل عن بنت وبنت الإبن وزوجة وجد وشقيق
6. ميت عن إبنه الأول وإبنه الثاني وبناته الأربعة وله مليار
7. ماتت امرأة وتركت زوجاً وأباً
8. توفي رجل عن بنت وأخ
9. توفي رجل عن زوجة و أخ
10. توفي عن زوجة وابن
11. توفي رجل عن بنتين وأخ
12. توفي عن بنتين وأخ
13. توفي شخص عن أم و أخ
14. توفي عن أم وأخ
15. توفي عن أخ لأم و أخت لأم وعم
16. مات امرأة عن زوج وأم وأخ شقيق
17. مات امرأة عن زوج وأخ شقيق وأخ لأب
18. مات شخص عن أخت وعم
19. مات شخص عن أختين شقيقتين وعم
20. مات رجل عن زوجة وأخت شقيقة وعم
21. مات عن زوجة وأختين شقيقتين وعم
22. توفي شخص عن أب وأم وبنت وبنت إبن
23. توفي شخص عن أخت شقيقة وأخت لأب وأخت لأم
24. توفي شخص عن زوجة و بنت وأخت شقيقة
25. توفي شخص عن أب وأخ
26. توفي رجل عن زوجة وبنت وأم وأخ شقيق
27. مات رجل عن زوجة وأخت شقيقة وأخ لام وأخ لأب
28. مات رجل عن بنت وأم وأب
29. توفي شخص عن بنتيه ووالديه
30. توفت امرأة عن زوج وست شقيقات
31. توفت عن زوج وشقيقتان
32. ماتت امرأة عن زوج وثلاث بنات إبن وأخ شقيق
33. مات رجل عن زوجتين واربع إخوة أشقاء
34. مات عن زوجتين واربع إخوة
35. مات شخص عن اربع بنات وثلاث أخوات شقيقات
36. توفت امرأة عن زوج واين وبنت واحدة، وتركت مبلغاً قدره ستمئة الف دينار
37. توفي رجل عن زوجة وأم واثنين وبنتاً واحدة وتركة قدرها ست مليون دينار فما نصيب كل منهم من التركة
38. توفي شخص عن جد وأخ شقيق
39. توفي شخص عن جد وأخ شقيق وأخت شقيقة
40. توفي شخص عن جد وأخوين
41. توفي شخص عن جد وثلاث إخوة
42. توفي شخص عن أم وجد وأخ
43. توفي شخص عن أم وجد وثلاث إخوة
44. توفت امرأة عن زوج وبنتين وجد وأخ
45. مات رجل وخلف زوجة وأخا شقيقاً
46. ماتت امرأة وتركت زوجاً وابناً وأباً
47. مات رجل وترك ابناً وأباً
48. مات رجل وترك أباً وبنتاً
49. ماتت امرأة وتركت زوجاً وبنتاً وأباً
50. توفي رجل عن زوجة وابن وأب وجد
51. توفي رجل عن زوجة وابن وبنت
52. ماتت. امرأة عن زوج وأم وأب
53. توفي رجل عن بنت وبنت إبن وزوجة وجد
54. توفي شخص عن بنت وجد وأم الأم.
55. توفي شخص عن أخ لأم وأخت لأم وعم
56. مات عن ثلاث أخوات شقيقات، وأختين لأم وأم
57. ماتت عن زوج وأختين شقيقتين وأم
58. مات عن زوجة وبنتين وأم وأخ شقيق
59. رجل عن أم وزوجة وأخ شقيق وأخت شقيقة
60. متوفى وله شقيق وأخت لأم وأخت لأب

Figure 1. Examples of inheritance issues

Many novelty articles interested with inheritance in Islamic law [23], [24] but not at the domain of applying artificial intelligence. Researchers' search on websites proved that there are no novelty articles treats the inheritance issue as Arabic unstructured text. Instead, there are many software applications dedicated to solve inheritance issue [7], [25]-[29]. But some of them described as testing version, incomplete and available of error such as [26], while the others looked as complete and has no error. All mentioned applications include software constructed to solve inheritance issue. The researcher studied them and found that they are receiving features manually. All of them provide empty fields that specified to issues. There are notes to tell user for which issue the field is specified. After receiving all required data, the verses that contain Islamic rule which is dedicated to current case must be displayed. Then Aza calculate and display the share for each heir. The users of such applications; people asking about inheritance issue, they have to break up the problem into fields and have to know each field where to be input, to achieve right answer. They have to understand How to run the application.

This work applies a new idea, which is represented by developing new algorithm; inheritance issues' features extraction using Arabic text analyzer (IFAA), it provides the ability of receiving inheritance problem as a characterized text written in Arabic language which is unstructured text. In this case, text mining task is necessary in order to extract all features. Task of text mining requires applying Arabic analyzer system. So the aim of this work is to solve the inheritance issue that received as unstructured characterized text. The plan is to construct a dedicated Arabic analyzing system to perform the tasks; extracting heirs, recognizing the kinship with a dead, extracting amount of legacy if found.

2. INHERITANCE ISSUES' FEATURES EXTRACTION USING ARABIC TEXT ANALYZER ARCHITECTURE

The proposed method IFAA is stands for inheritance issues' features extraction using Arabic text analyzer. Its architecture is decomposed into four main components; central unit, analyzer unit, the heir-table,

and ENAT method, see Figure 2. Following sections describe IFAA main components and subcomponents in details.

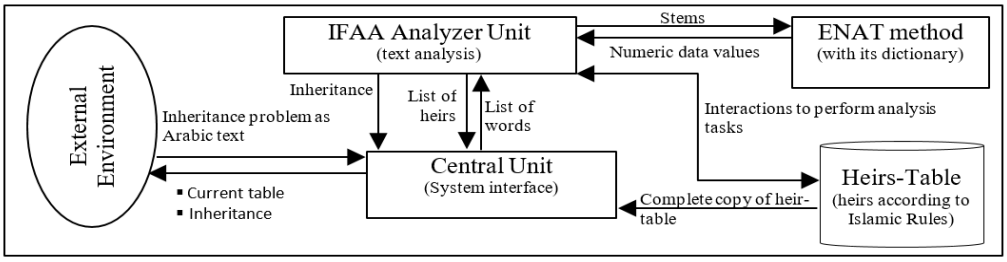


Figure 2. IFAA architecture

2.1. IFAA heirs-table

As mentioned before; there are many and many types of heirs. Each one of theme has a different rule to apply the inheritance operation. And each one of them has a specific share that differs from others. IIFFA heirs-table is constructed depending on Al-Khatrawi which divided heirs into fourteen males, and ten females [30]. It is important to talk about the two terms branches and origins. Branches term is used to indicate that the heir is a branch from the inherited person, such as; son and son's son. Origins term is used to indicate that the heir is an origin for the inherited person, such as; father and grandfather. IFAA build a good structure to solve the problems; by saving all possible heirs. The structure looks like a table, in which each type of heirs has a specific row, so it is called heirs-table, see Table 1. Each type of heirs is described with five fields, as listed below.

Table 1. IFAA heirs-table

g-sym	h-r	r-sym	h-ds	co-sym
G6a	Son	S	integer	Sco
G6a	son's son	Ss	list of integer	ScoL
G5a	Father	F	boolean	null
G9a	Grandfather	Gf	list of boolean	null
G7a	Brother	B	integer	Bco
G7a, G8a	brother by father	Bf	integer	Bfco
G7a, G8b	brother by mother	Bm	integer	Bmco
G6a, G7a	brother's son	Bs	list of integer	BscoL
G6a, G7a, G8a	brother's by father son	Bfs	list of integer	BfscoL
G10	Uncle	U	integer	Uco
G10, G8a	uncle by father	Uf	integer	Ufco
G6a, G10	Cousin	C	list of integer	CcoL
G6a, G10, G8a	cousin by father	Cf	list of integer	CfcoL
G4a	husband	H	boolean	null
G6b	daughter	D	integer	Dco
G6b, G6a	son's daughter	Ds	list of integer	DscoL
G5b	mother	M	boolean	null
G9b, G8b	mother's mother	Mm	boolean	Co
G9b, G8a	father's mother	Mf	boolean	Co
G5b, G5a, G5a	father's father's mother	Mff	boolean	Co
G7b	sister	Sis	integer	Sisco
G7b, G8a	sister by father	Sisf	integer	Sisfco
G7b, G8b	sister by mother	Sism	integer	Sismco
G4b	wife	W	integer	Wco

a. Group symbol field let ***g-sym***; it is the symbol that refers to the group of words at IFAA dictionary, to which the heir belongs.
b. Heir relationship field let ***h-r***; it is just for describing the relationship that relates the heir with died person.
c. Relationship symbol field let ***r-sym***; just for providing a symbol to be used for referring to the heir.
d. Heir data structure field let ***h-ds***; it determines a suitable structure to contain the heir. Because of differences among hairs, different types of structures may be used, as follows:
   - The heir may be a single person such as; ***wife*** and ***husband***, so a ***boolean*** is suitable to tell whether the heir is existing or not.

- The heir may be a list of single person such as; ***grandfather*** –different levels of grandfathers-, so ***list-of-boolean*** is suitable.
- The heir may be a number of persons such as; ***brothers*** and ***sisters***, so an ***integer*** is suitable to save the number.
- The heir may be a list of persons such as; ***sons of son*** –sons in different levels-, so the suitable structure is ***list-of-integer***.

e. Counter symbol field let ***co-sym***; it is just a symbol that refers to the counter responsible for counting the counted heir, such as ***sisters*** and ***brothers***. While uncounted heir such as ***husband*** has no such symbol, so null is used.

2.2. ENAT method

ENAT stands for Extract Numerical data from Arabic Texts. ENAT component is a complete proved system. It is taken from the previous work [21], which is passed in manipulating numeric words written in Arabic language. Current work needs ENAT to recognize numeric words and extract numerical data, here; that is very necessary because numerical data comes frequently in inheritance problems, as shown in examples at Figure 1. Through implementing jobs at current article, ENAT will be called; which tests the stems, if finds numeric data, then returns its value.

2.3. IFAA central unit

The central unit works as system interface. It receives inheritance problem from external environment. The problem is an Arabic text that includes set of heirs and may be the amount of inheritance. Central unit interacts with IFAA analyzer unit, which performs analysis and then extracts the problem issues. Central unit interacts with IFAA heirs-table to construct a table of heirs for the current problem.

2.3.1. Central unit connections with analyzer

A stream of characters is received from external environment. It is waiting for analysis process, but before that; some pre-processes must be done; firstly performing tokenization which means converting characters into list of words, and secondly removing stop-words. So, two lists have provided from the previous work RSGAS (Abeer *et al.* 2014); list of delimiters to perform tokenization task, and list of stop-words to perform recognizing stop-words. Now the central unit connects with the analyzer to transmit the list of words. After performing analysis, two issues will be extracted; firstly is list-of-heirs, and secondly is amount of inheritance. The central unit receives all extracted issues from the analyzer. List-of-heirs is a list of compound object h-co-obj which is decomposed as follows: h-co-obj (r-sym, level, count), where:

- r-sym is the symbol that specified to a specific heir, as found in heirs-table.
- Level is an integer number to recognize depth-degree of the relationship. For a relationship such as wife or husband; level value is -1 to indicate that there is no level. For a direct relationship such as son or father; level value is zero. For an indirect relationship such as son's son or grandfather; level value is more than zero. Figure 3 illustrates some examples about the assignment of values to level.
- Count is an integer which represents the count of this heir. As an example the word *بنتان* includes count =2 for daughter heir.

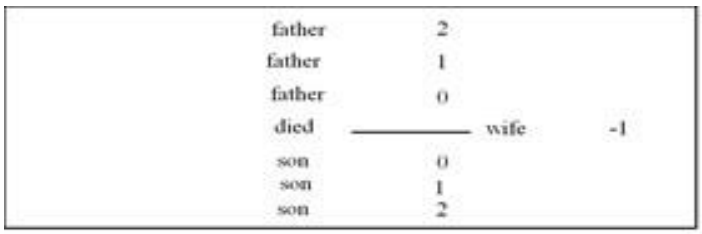


Figure 3. Assigning values to level

2.3.2. Central unit connection with heirs-table

Now list of heirs stayed at central unit, which starts interaction with the heirs-table to get a complete copy of it. Central unit must improve that copy; to produce a new version dedicated for the current problem. The new version is called current-table. Central unit analyses the list of heirs that were received from analyzer unit. It studies the effecting of each heir on the current-table. Flowchart 1 is responsible for performing the task of current-table construction, see Figure 4.

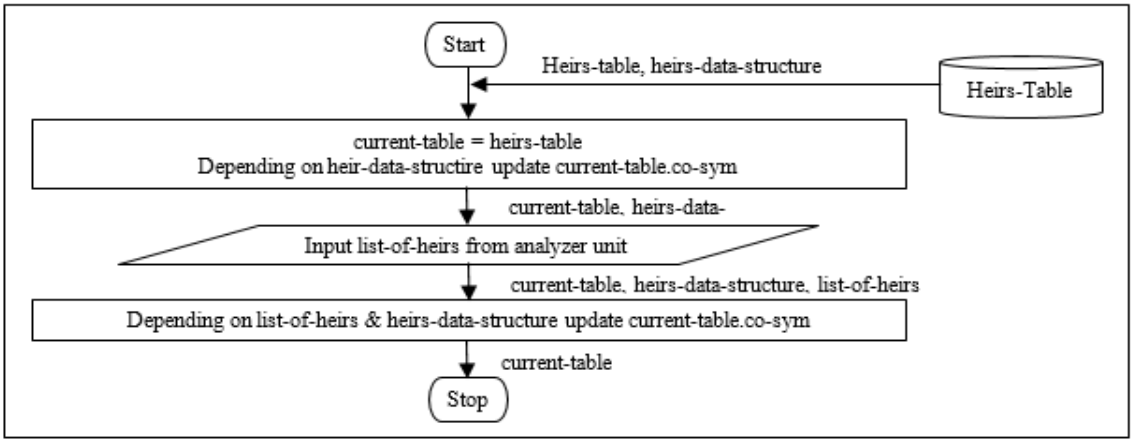


Figure 4. Flowchart 1: current-table construction

2.4. IFAA analyzer

IFAA analyzer is a stem-based analyzer dedicated for inheritance problems. Its tasks are; extracting list of heirs and extracting amount of inheritance. Analyzer unit interacts with central unit and the heir-table to perform its tasks. IFAA analyzer is decomposed into four subunits; analyzer center, dictionary, stemmer, and sections analyzer, see Figure 5.

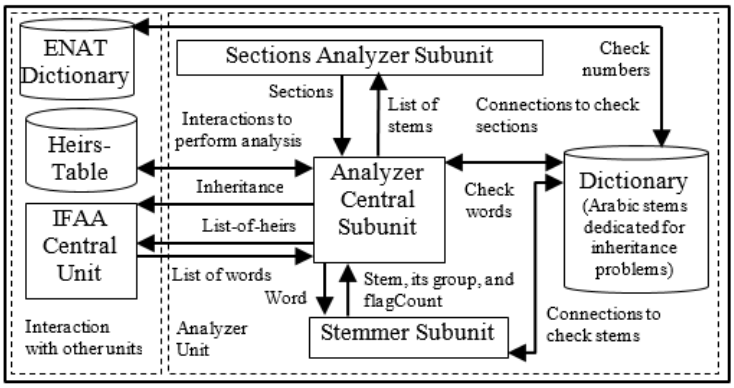


Figure 5. IFAA analyzer unit

2.4.1. Analyzer dictionary

IFAA analyzer is provided with dictionary, in which all Arabic words that may be used in inheritance problems are stored. Analyzer dictionary is a stem-based, which means that stems of words are stored. Because of the limited range of words in this work, stem-based approach is preferable than root-based approach. Many examples of inheritance issues were collected as showed in Figure 1. The examples are studied and all possible words have been collected. Words in the dictionary were classified according to their purposes.

All previous works' dictionaries divided words into two main classes; verbs and nouns. While, for current work; the studying of all possible examples shows that; verb-chunks are not important, because all heirs are noun-chunks. As shown in examples, verb-chunks either for telling that a person is died, or telling that there are heirs. The task here is to search about heirs, and they are all found in noun-chunks. So, IFAA dictionary does not be classified into verbs and nouns. IFAA dictionary classified words into ten groups, see Table 2. These groups are classified into two main classes:

- Keyword class includes all stems that distinguish the standard structure of the inheritance problem.
- Heir class includes all stems that refer to the heirs.

Table 2. IFAA dictionary table

g-sym	g-title	g-purpose	Stem	Stem-co-obj Flag count	Stem	Flag count
G1	Died group	referring to the died person	ميت	-1	متوفي	-1
			مات	-1	توفي	-1
			توفى	-1	توفت	-1
			رحل	-1	فارق	-1
			هلك	-1	راحل	-1
			مفارق	-1	هالك	-1
G2	Heir keyword	Keywords refer to the heirs	ترك	-1	خلّف	-1
			عن	-1	له	-1
			لها	-1		
G3	Inheritance keyword	Keywords refer to inheritance	ترك	-1	خلّف	-1
			ورّث	-1	يملك	-1
			تملك	-1	تمتلك	-1
			يمتلك	-1	يتملك	-1
			تتملك	-1	تركة	-1
			رزق	-1	ارزاق	-1
			مبلغا	-1	ارث	-1
			ثروة	-1	ممتلكات	-1
			له	-1	لها	-1
G4	Husband group	Stems refer to heir (describe husband or wife relationship)		Sub-group a		
			زوج	-1	زوجاً	-1
				Sub-group b		
			زوجة	1	زوجتين	2
			زوجات	-2	زوجتان	2
			زوجتا	2	زوجتي	2
G5	Parents group	Stems refer to heir (describe parents relationship)		Sub-group a		
			أب	-1	أباً	-1
			والد	-1	أبي	-1
			والداً	-1		
				Sub-group b		
			أم	-1	أماً	-1
			والدة	-1		
				c		
			والدي	2	والدين	2
			والدان	2		
G7	Brother1 group	Stems refer to heir (describe brother or sister relationship)	Sub-group a			
			أخ	1	أخا	1
			أخو	1	شقيق	1
			شقيقان	2	شقيقين	2
			أشقّاء	-2	أخوين	2
			أخين	2	أخان	2
			إخوة	-2		
G8	Brother2 group	Stems tell about brother or sister relationship; from one parent, or from two parents	Sub-group a			
			لأب	-1	لوالد	-1
			لأبو	-1	لأبي	-1
			Sub-group b			
			لوالدة	-1	لأم	-1
			Sub-group c			
			شقيق	1	شقيقة	1
			شقيقان	2	شقيقين	2
			شقيقتين	2	شقيقتان	2
			أشقّاء	-2	شقيقات	-2
			لوالدان	2	لوالدين	2
			لوالدا	2	لوالدي	2
G9	Grand group	Stems refer to heir (describe grandfather relationship)	Sub-group a			
			جد	1	جدان	2
			جدين	2		
			Sub-group b			
			جدة	1	جدتان	2
			جدتين	2		
G10	Uncle group	Stems refer to heir (describe uncle relationship)	عم	1	عمان	2
			عمين	2	أعمام	-2
			عما	2	عمي	2
			عموم	-2		
G6	Son group	Stems refer to heir (describe Son or daughter relationship)	Sub-group a			
			إبن	1	إبناً	1
			إبنان	2	إبنين	2
			إبني	2	أبناء	-2
			Sub-group b			
			بنت	1	بنتاً	1
			بنتين	2	بنتان	2
			بنتي	2	بنات	-2

Table 2 is IFAA dictionary table. For each row in the table there is a group of words, for example, the fifth raw has many words that may refer to the father, such as (أب،والد). The table contains row for each one of possible heirs. In addition to heirs-words groups, there are several other groups. The other groups deal with words used to one of the following purposes:

- They tell about died person such as *مات، توفي*.
- Start to list the heirs, such as *عن، ترك*.
- May be word to describe the brothering relationship, such as *شقيق، لأم*.
- Or may tell about the inheritance, such as *ارث، ثروة*.

As shown in Table 2, the dictionary has four fields. First three fields are; group symbol ***g-sym***, group title ***g-title***, and group purpose ***g-purpose***. They are used to describe about the classification. Such fields are necessary to facilitate the dictionary manipulation during this work. The fourth field contains all stems. The stem is presented by compound object ***stem-co-obj***, it has two parts; one for the stem-word, and another for counting.

stem-co-obj(stem, flagCount) where:

stem is a stem-word in Arabic language, which is used in inheritance issues.

flagCount (***fco)*** is an integer value used as a flag to indicate whether the stem is a countable word or not. It is also used as a flag of the prefix *و*, when its value is a fractional.

Really the ***flagCount*** is according to the usage in inheritance issues, not according to Arabic language rules. For example, the word '*زوج*' could be countable word according to Arabic language rules, by using the stems ('*زوجين*' ,'*زوجان*', and '*ازواج*'). While in inheritance issues it is impossible to be countable. The reason is that if a wife is died, then there is only one -on live- husband who must inherit a dead wife. So the ***flagCount*** may be one of the following values:

- -1 or -0.1 for all uncountable stems, for example 'زوج' and 'ام'.
- 1 or 0.1 for all stems that refer to single, for example, 'بنت' and 'اخ'.
- 2 or 0.2for all stems that refer to two heirs, for example, 'بنتان' and 'زوجتان'.
- -2 or -0.2 for all stems that refer to more than two heirs, for example, 'بنات' and 'زوجات'. Such heirs require numeric word to count them; it may be before or after the stem for example, 'ثلاث زوجات'.

2.4.2. Stemmer subunit

Stemmer analyzer receives an Arabic word from analyzer central subunit. After performing stemming analysis; it returns the stem compound object stemG-co-obj. That contains the stem with its specifications.

stemG-co-obj(stem, g-sym, fco) where:

- stem is the Arabic stem as stored in IFAA dictionary.
- g-sym is the group symbol of the stem as stored in IFAA dictionary.
- fco is the flagCount of the stem as stored in IFAA dictionary.

Stem analyzer searches the dictionary trying to seek the word with its specifications. A word may be a stem, so it is found directly, or a word may be a stem with affixes, so it applies Arabic morphological rules to perform word analysis and then seek the stem with its specifications. In stem-based approach any word is decomposed into (prefix, stem, and suffix). Flowchart 2 performs the stemmer tasks, see Figure 6.

All prefixes and suffixes in Arabic language were collected at the previous work KISB [16]. Test was done to determine a sub-list of them that are necessary for current work. Each group of dictionary's words is tested alone. Table 3 shows the IFAA-prefixes pre and IFAA-suffixes suf grouped according to IFAA dictionary groups. Current work affixes is limited because of the limitation of used words, in addition to the limitation of tenses. Only past tense is suitable with such issues, because a person cannot inherit from any person that is on live now, so all collected examples were in past tense.

Table 3. IFAA prefixes and suffixes

g-sym	pre	suf
G1	a: و، ف b: ال	a: ت b: ت
G2	a: و، ف b: ال، و، ف، وال، فال c: و	a: ت b: ت
g-sym	pre	suf
G4, G5, G6, G7, G9, G10	ال، و، ف، لل، وللل، فال، فلل، وال	ه، ها
G8		ه، ها، وه، وها، يه، يها
G3	a: ف، و b: ال، و، ف، وال، فال c: و	a: ت

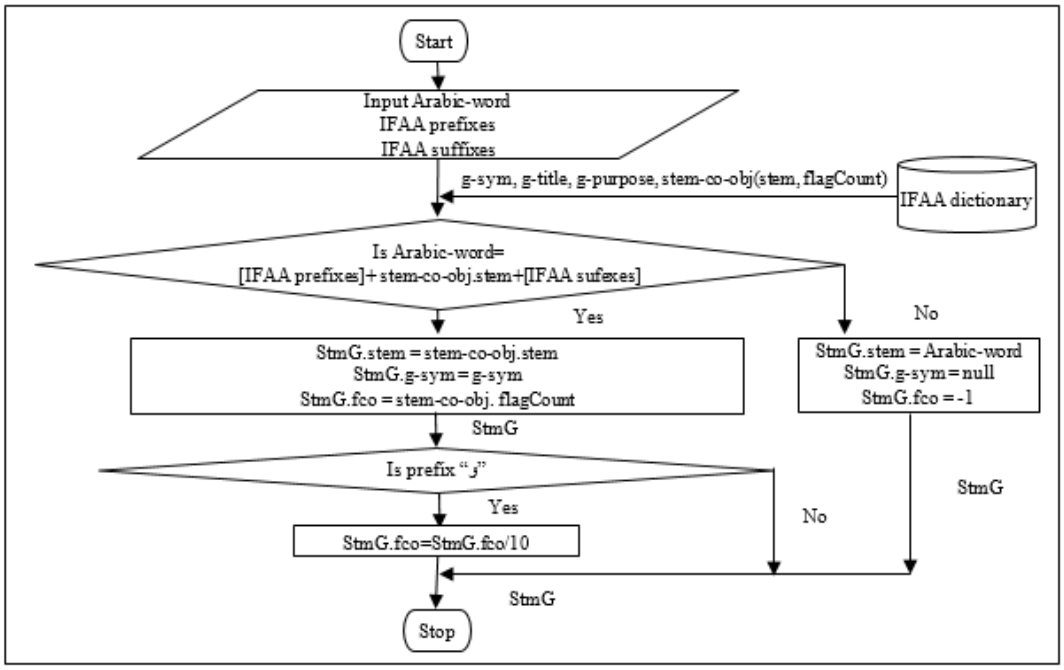
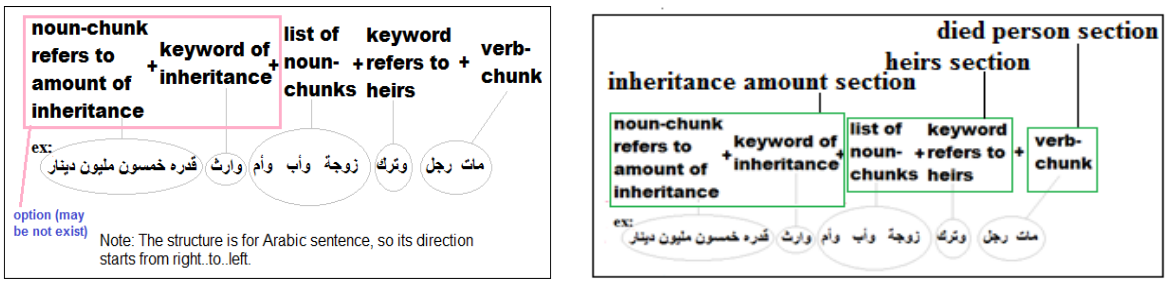


Figure 6. Flowchart 2: stemmer tasks

2.4.3. Sections analyzer subunit

All collected examples that showed at Figure 1 has been studied. It was important to build a suitable structure that could be applied on the examples. This work suggests a suitable standard structure; which is decomposed into three main sections; died person section, heirs section, and inheritance amount section. Sections analyzer subunit puts the list of received stems; list of ***stemG-co-obj***, in the standard structure by recognizing the three sections, as shown in Figure 7 that shows an inheritance issue example in Arabic language, Figure 7(a) specifies the group for each word in the example, while Figure 7(b) shows the three standard sections. All collected examples are started with verb-chunk. It is just to tell that a person has been died. So it could be neglected, because it has no effect on heirs and their shares. Next to the verb-chunk, mostly there is a word refers to the heirs. So this part of text is important and must be saved in a suitable form; ***list-of-heirs***, as mentioned before in this paper. Next to heirs, information about amount of inheritance may be found. So a buffer is needed to save the amount of inheritance.

Flowchart 3 implements sections analyzer jobs, see Figure 8. The input is the list of ***stemG-co-obj***, which is received from the analyzer central subunit. The output is the three standard sections; died person section ***Sec1***, heirs section ***Sec2***, and inheritance amount section ***Legacy***. First job is to extract died person section. Second job is to test existing of heirs. So there is a sub-section for each heir, which must be checked to extract number and extract relationship specification if found. Numeric data is tested by using ENAT method. Heir sub-section is one compound object of type ***h-co-obj***, but it will be two heirs in case of the parents. Third job is to extract value of legacy which is calculated by using ENAT method.



(a) (b)

Figure 7. Inheritance issue example in Arabic language; (a) shows the group of each word and (b) recognizes the three standard sections

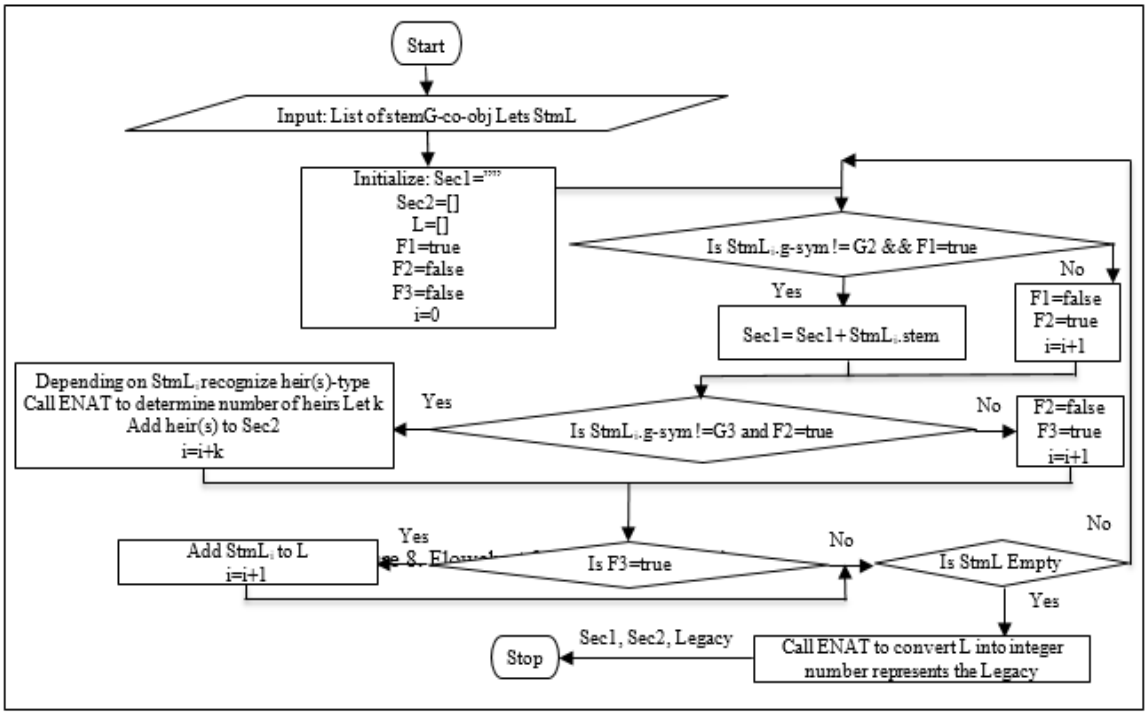


Figure 8. Flowchart 3: section construction

3.4.4. Analyzer central subunit

Analyzer central sub-unit receives list of Arabic words from IFAA central unit. It provides interactions among dictionary, stemmer, sections analyzer, and other main components to perform following:

- Extract stems from Arabic words.
- Put the current problem in the standard form and recognize its sections.
- So, it will be easy to extract list-of-heirs and the amount of inheritance that must be returned into IFAA central unit.

3. RESULTS AND DISCUSSION

This work is an attempt to build a dedicated Arabic analyzer for inheritance problem that is received as unstructured text. Manipulating unstructured text and performing mining tasks is a big challenge. A previous work ENAT passed in extracting numeric data from unstructured Arabic text. Because of inheritance problem, as well as any unstructured Arabic text; may include numeric word within it, so IFAA make useful of the proved work ENAT. The behavior of IFAA will be shown by follow the analysis of examples 1..6 that shown in Figure 1. Each example represents Arabic unstructured text. IFAA receives the text by its central unit which performs preprocessing and results list of words, as shown in Table 4.

Table 4. Preprocessing job on received text

Ex.No.	List of words									
	w.10	w.9	w.8	w.7	w.6	w.5	w.4	w.3	w.2	w.1
1	ملايين	عشر	خمسة	ولها	وأخ	والأب	زوجا	وتركت	امرأة	ماتت
2			تركة	وله	وزوجة	وأخ	والدين	عن	رجل	توفي
3	دينار	مليون	وله	وعم	لأب	وأخت	لأبوه	أخ	عن	توفي
4			وجد	وأب	والزوج	شقيقات	ست	عن	امرأة	توفت
5		وشقيق	وجد	وزوجة	الإبن	وبنت	بنت	عن	رجل	توفي
6	مليار	وله	الاربعة	وبناته	الثاني	وإبنه	الأول	إبنه	عن	ميت

List of words is transmitted into the analyzer unit, at which the first job performed is stemming analysis as shown in Table 5. The output of this job is list of ***stemG-co-obj***. That contains the stem of the word with its specifications; stem, group, flagCount.

Table 5. Stemming analysis of Arabic words

Ex.No.		List of words									
		w.10	w.9	w.8	w.7	w.6	w.5	w.4	w.3	w.2	w.1
1	word	ملايين	عشر	خمسة	ولها	وأخ	والأب	زوجا	وتركت	امرأة	ماتت
	stem	ملايين	عشر	خمسة	لها	أخ	أب	زوجا	ترك	امرأة	مات
	group	null	null	null	G3	G7a	G5a	G4a	G2	null	G1
	fco	-1	-1	-1	-0.1	0.1	-0.1	-1	-0.1	-1	-1
2	word			تركة	وله	وزوجة	وأخ	والدين	عن	رجل	توفي
	stem			تركة	له	زوجة	أخ	والدين	عن	رجل	توفي
	group			G3	G3	G4b	G7a	G5c	G2	null	G1
	fco			-1	-0.1	0.1	0.1	2	-1	-1	-1
3	word	دينار	مليون	وله	وعم	لأب	وأخت	لأبوه	أخ	عن	توفي
	stem	دينار	مليون	له	عم	لأب	أخت	لأب	أخ	عن	توفي
	group	null	null	G3	G10	G8a	G7b	G8a	G7a	G2	G1
	fco	-1	-1	-0.1	0.1	-1	0.1	-1	1	-1	-1
4	word			وجد	وأب	والزوج	شقيقات	ست	عن	امرأة	توفت
	stem			جد	أب	زوج	شقيقات	ست	عن	امرأة	توفت
	group			G9a	G5a	G4a	G8c	null	G2	null	G1
	fco			0.1	-0.1	-0.1	-2	-1	-1	-1	-1
5	word		وشقيق	وجد	وزوجة	الإبن	وبنت	بنت	عن	رجل	توفي
	stem		شقيق	جد	زوجة	إبن	بنت	بنت	عن	رجل	توفي
	group		G7a	G9a	G4b	G6a	G6b	G6b	G2	null	G1
	fco		0.1	0.1	0.1	1	0.1	1	-1	-1	-1
6	word	مليار	وله	الاربعة	وبناته	الثاني	وإبنه	الأول	إبنه	عن	ميت
	stem	مليار	له	الاربعة	بنات	الثاني	إبن	الأول	إبن	عن	ميت
	group	null	G3	null	G6b	null	G6a	null	G6a	G2	G1
	fco	-1	-0.1	-1	-0.2	-1	0.1	-1	1	-1	-1

List of ***stemG-co-obj*** is transmitted into section analyzer subunit, at which first job is to extract first section, died person section. This job is performed by implementing flowchart 3. Table 6 shows the first section for each example.

Table 6. Extraction of first section

Ex.No	Sec.1	Rest of text								
1	مات امرأة	ملايين	عشر	خمسة	لها	أخ	أب	زوجا	ترك	
		null	null	null	G3	G7a	G5a	G4a	G2	
		-1	-1	-1	-0.1	0.1	-0.1	-1	-0.1	
2	توفي رجل			تركة	له	زوجة	أخ	والدين	عن	
				G3	G3	G4b	G7a	G5c	G2	
				-1	-0.1	0.1	0.1	2	-1	
3	توفي	دينار	مليون	له	عم	لأب	أخت	لأب	أخ	عن
		null	null	G3	G10	G8a	G7b	G8a	G7a	G2
		-1	-1	-0.1	0.1	-1	0.1	-1	1	-1
4	توفت امرأة			جد	أب	زوج	شقيقات	ست	عن	
				G9a	G5a	G4a	G8c	null	G2	
				0.1	-0.1	-0.1	-2	-1	-1	
5	توفي رجل		شقيق	جد	زوجة	إبن	بنت	بنت	عن	
			G7a	G9a	G4b	G6a	G6b	G6b	G2	
			0.1	0.1	0.1	1	0.1	1	-1	
6	ميت	مليار	له	الاربعة	بنات	الثاني	إبن	الأول	إبن	عن
		null	G3	null	G6b	null	G6a	null	G6a	G2
		-1	-0.1	-1	-0.2	-1	0.1	-1	1	-1

The next job is to extract list of heirs. This job needs many steps. Firstly is the extracting of sub-section, which represents just one heir, as shown in Table 7. Secondly ENAT is called to extract numeric data from the subsection. Thirdly is determining the relationship specification from the subsection. Note that numeric data (N) and specification (S) are not necessary exist, as shown in Table 8. The value (0) of N means there is no numeric data. The value (0) of S means that there is no specification. The value 1 at example 3 means that the relationship is by father. The rest of heir sub-section found at field ***Rh-sub***.

Fourthly more checking must be done to decide the actual details of ***h-co-obj***. Then the result is either one heir as shown at examples 1, 3, 4, 5, and 6. Or it may be two heirs as shown at example 2, see Table 9. Field ***Heir*** contains ***h-co-objs*** of first heir sub-section.

Loop at flowchart 3 is repeated extracting all heirs and the rest of text may contain the amount of inheritance, so ENAT method is called to extract the numeric data if found and then saved at the third section, as shown in Table 10. Section 2 detailed into heirs with their ***h-co-obgs***.

Know the central unit receives list of ***h-co-objs*** and the amount of inheritance. Next job is to construct current-table by implementing flowchart 1. The current table is established by copy heirs-table, at which all flags false and all counters equal to 0. Then, updating must be done according to list of h-co-objs, as shown in Table 11. That was the last job; the output of IFAA has been achieved.

Table 7. Extract a sub-section represents heir

Ex.No	first heir sub-sec	Sec.1	Rest of text					
1	زوجا	مات امرأة	ملايين	عشر	خمسة	لها	أخ	أب
	G4a		null	null	null	G3	G7a	G5a
	-1		-1	-1	-1	-0.1	0.1	-0.1
2	والدين	توفي			تركة	له	زوجة	أخ
	G5c	رجل			G3	G3	G4b	G7a
	2				-1	-0.1	0.1	0.1
3	أخ لأب	توفي	دينار	مليون	له	عم	لأب	أخت
	G8a G7a		null	null	G3	G10	G8a	G7b
	-1 1		-1	-1	-0.1	0.1	-1	0.1
4	ست شقيقات	توفت امرأة			جد	أب	زوج	
	G8c	null				G9a	G5a	G4a
	-2	-1				0.1	-0.1	-0.1
5	بنت	توفي رجل	شقيق	جد	زوجة	إبن	بنت	
	G6b		G7a	G9a	G4b	G6a	G6b	
	1		0.1	0.1	0.1	1	0.1	
6	إبن الأول	ميت	مليار	له	الاربعة	بنات	الثاني	إبن		
	null	G6a		null	G3	null	G6b	null	G6a
	-1	1		-1	-0.1	-1	-0.2	-1	0.1

Table 8. Extract numeric data and specification from the heir subsection

Ex.No	S	N	Rh-sub	Sec.1	Rest of text					
1			زوجا	مات امرأة	ملايين	عشر	خمسة	لها	أخ	أب
	0	0	G4a		null	null	null	G3	G7a	G5a
			-1		-1	-1	-1	-0.1	0.1	-0.1
2			والدين	توفي			تركة	له	زوجة	أخ
	0	0	G5c	رجل			G3	G3	G4b	G7a
			2				-1	-0.1	1	0.1
3			أخ	توفي	دينار	مليون	له	عم	لأب	أخت
	1	0	G7a		null	null	G3	G10	G8a	G7b
			1		-1	-1	-0.1	0.1	-1	0.1
4			شقيقات	توفت امرأة				جد	أب	زوج
	0	6	G8c					G9a	G5a	G4a
			-2					0.1	-0.1	-0.1
5			بنت	توفي رجل		شقيق	جد	زوجة	إبن	بنت
	0	0	G6b			G7a	G9a	G4b	G6a	G6b
			1			0.1	0.1	0.1	1	0.1
6	0	1	إبن الأول	ميت	مليار	له	الاربعة	بنات	الثاني	إبن
			null	G6a	G3	null	G6b	null	G6a	G6a
			-1	1	-0.1	-1	-0.2	-1	0.1	0.1

Table 9. H-co-obj of first subsection

Ex.no.		Heir		Sec.1	Rest of text				
	sym	lev	co						
1	H	-1	-1	مات امرأة	دينار	مليونان	لها	أخ	أب
					null	null	G3	G7a	G5a
					-1	-1	-0.1	0.1	-0.1
2	F	0	1	توفي رجل				زوجة	أخ
								G4b	G7a
	M	0	1					0.1	0.1
3	Bf	-1	1	توفي	دينار	مليون	له	عم	لأب	أخت
								G10	G8a	G7b
								0.1	-1	0.1
4	Sis	-1	6	توفت امرأة				جد	أب	زوج
								G9a	G5a	G4a
								0.1	-0.1	-0.1
5	D	0	1	توفي رجل		شقيق	جد	زوجة	إبن	بنت
						G7a	G9a	G4b	G6a	G6b
						0.1	0.1	0.1	1	0.1
6	S	0	1	ميت	مليار	له	الاربعة	بنات	الثاني	إبن
					null	G3	null	G6b	null	G6a
					-1	-0.1	-1	-0.2	-1	0.1

Table 10. Results of flowchart 3

Ex.No	Sec.3	Sec.2				Sec.1
		heir	sym	lev	co	
1	15000000	زوجا	H	-1	-1	مات امرأة
		أب	F	0	-1	
		أخ	B	-1	1	
2	0	والدين	F	0	-1	توفي رجل
			M	0	-1	
		أخ	B	-1	1	
		زوجة	W	-1	1	
3	1000000	أخ لأب	Bf	-1	1	توفي
		أخت لأب	Sisf	-1	1	
		عم	U	-1	1	
4	0	ست شقيقات	Sis	-1	6	توفت امرأة
		زوج	H	-1	-1	
		أب	F	0	-1	
		جد	Gf	1	1	
5	0	بنت	D	0	1	توفي رجل
		بنت إبن	Ds	1	1	
		زوجة	W	-1	1	
		جد	Gf	1	1	
		شقيق	B	-1	1	
6	1000000000	إبن الأول	S	0	1	ميت
		إبن الثاني	S	0	1	
		بنات الاربعة	D	0	4	

Table 11. Updating on heirs-table to construct current-table

Ex.no	Updating on heirs-table
1	The flag of Father is true
	The counter of Brother is 1
	The flag of Husband is true
2	The flag of Father is true
	The counter of Brother is 1
	The flag of Mother is true
	The counter of Wife is 1
3	The counter of Brother by Father is 1
	The counter of Uncle is 1
	The counter of Sister by Father is 1
4	The flag of Father is true
	The flag of first Grandfather is true
	The flag of Husband is true
	The counter of Sister is 6
5	The flag of first Grandfather is true
	The counter of Brother is 1
	The counter of Daughter is 1
	The counter of Son's Daughter is 1
	The counter of Wife is 1
6	The counter of Son is 2
	The counter of Daughter is 4

The presented implementation of current system includes sis examples of Arabic text. They covered different cases of heirs. Current system; IFAA succeeded in applying the suggested standard structure and convert the unstructured Arabic text into three sections. IFAA succeeded in analyzing heir section and extracts all heirs and determine their numbers and types. Also IFAA will be succeeded with any inheritance issue text written in Arabic, because of the flexibility of the suggested standard structure. IFAA test a word, if it is a key for a specific section then trying to extract the section.

So there is no condition for order of sections. IFAA collects all possible Arabic stems in field of inheritance, and organize them in related groups. The presented implementation proves that IFAA is a perfect analyzer with high accurate reaches to 100%. The limitation of such system just appeared with wrong words that may be resulted from typos. The wrong word may be a key, so the section will not be recognized. Or it may be heir, so he will not be extracted. Or it may be a number, so this number will be lost. It is good idea to merge the spell checker method; that is proved in a previous work, with IFAA analyzer to treat the limitation of wrong words for future.

4. CONCLUSION

Extracting features from unstructured text is a challenge. Suggesting a standard structure to convert unstructured text into number of structured sections is good process to treat unstructured text. This work passed in developing algorithm called IFAA which is Arabic analyzer system that extracts features from unstructured text and it is dedicated for texts talking about inheritance issues. IFAA passed in developing dedicated dictionary for inheritance issues which collects all Arabic stems related to such problems and groups them according to their purposes. IFAA collected all possible heirs and suggested a perfect structure to accomplish all possible cases related to heirs. IFAA suggested the standard structure and applied it to divide the unstructured text into three sections with full flexibility. IFAA treated cases of numeric words by invoking ENAT method. IFAA developed perfect algorithm to recognize heirs and determine their numbers and types. IFAA treated the difficulties of using allowed inheritance applications, those that require difficult steps to be used. Because of all that; IFAA is a perfect analyzer system for inheritance issues feature extraction, with accurate ratio reaches to 100%. There is only one point of limitation with wrong words that may be caused from typos.

REFERENCES

[1] M. B. Jumah, "Ahkam Almyrath fi Alsharyah Alislameh," *Dar Al-Fikr*, Ammaan, Jordan 1981.
[2] Mawlood M. Al-Rawy, "Ilm Alfaraid wAlmawareeth*," Mktebh Kelimh*, Baghdad, Iraq, 2021.
[3] M. Al-Sjawendy, "Sharh Alsrajyh*," Dar Alkutub Alilmyh,* Beirut, Lebanon, 2018.
[4] M. Al-Dimashqy, "Sharh Alrahbyh fi Ilm Alfaraidm*," Dar Alkutub Alilmyh,* Beirut, Lebanon, 2015.
[5] Ahmed S. Al-Rasmoki, "Ilm Almyrath-Idah Alasrar Almsonh fi Aljawahir Almaknonh fi Sudf Alfraid Almesnonh," *Ministry of Awqaf and Islamic Affairs*, Kingdom of Morocco, 2004.
[6] Mehfod A. Al-Kaldony, "Altahtheeb fi Ilm Alfaraid wAlwasaya," *Dar Al-Fikr*, Ammaan, Jordan, 1995.
[7] N. H. H. Mohammed, "Web based expert system for i-faraid," *University Utara Malaysia*, 2009.
[8] F. H. E. Al-Qawasmeh, and P. Pit, "Improving arabic text categorization using decision trees," *First International Conference on Networked Digital Technologies, IEEE*, pp. 110-115, 2009.
[9] S. Kamenka, "Natural language generation and semantic technologies," *Cybernetics And Information Technologies*, vol. 14, no. 2, pp. 3-23, 2014, doi: 10.2478/cait-2014-0015.
[10] M. G. H. A. Zamil and Q. Al-Radaideh, "Automatic extraction of ontological relations from Arabic text," *Journal of King Saud University – Computer and Information Sciences*, vol. 26, no. 4, pp. 462-472, 2014, doi: 10.1016/j.jksuci.2014.06.007.
[11] M. I. Rahman, N. A. Samsudin, A. Mustapha, and A. Abdullahi, "Comparative analysis for topic classification in Juz Al-Baqarah," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 12, no. 1, pp. 406-411, 2018, doi: 10.11591/ijeecs.v12.i1.pp406-411.
[12] S. A. Salloum, Q. A. Ahmad, M. Al-Emran, and S. Khaled, "A survey of arabic text mining. in: (eds) intelligent natural language processing: trends and applications," *Studies in Computational Intelligence Springer, Cham*, vol. 740, pp. 417-431, 2018, doi: 10.1007/978-3-319-67056-0_20.
[13] E. Reda, B. Gamal, and A. Al-Katib, "Extractive arabic text summarization using modified pagerank algorithm," *Egyptian Informatics Journal*, vol. 21, no. 2, pp. 73-81, 2020, doi: 10.1016/j.eij.2019.11.001.
[14] A. Alsayat and N. Elmitwally, "A comprehensive study for arabic sentiment analysis (challenges and applications)," *Egyptian Informatics Journal*, vol. 21, pp. 7-12, 2020, doi: 10.1016/j.eij.2019.06.001.
[15] A. K. Al-Mashhadany, B. Taha and S. Venus, "Arabic diagnosing expert system shell (ADESS)," *2010 The Proceedings of 2nd Conference for Information Technology: Applications and Horizons, University of Technology, Baghdad, Iraq*, 2010.
[16] A. K. Al-Mashhsdany, "Knowledge acquisition & hybrid inference with a stem-based approach (KISB)," *Journal of Al-Nahrain University*, vol. 15, no. 2, pp. 169-183, 2012, doi: 10.22401/JNUS.15.2.25.
[17] A. K. Al-Mashhsdany, *et al.*, "Root-stem approach in general analyzer system for arabic language (RSGAS)," *Journal of University of Anbar for Pure Science*, vol. 8, no. 1, pp. 41-49, 2016, doi: 10.37652/juaps.2016.135154.
[18] A. T. Sadiq, A.-Obaidi, and A. K. Al-Mashhadany, "Using english morphological analyzer to decrease the dictionary size in knowledge extraction techniques," *International Journal of Research and Reviews in Soft and Intelligent Computing (IJRRSIC), Science Academy Publisher, United Kingdom*, vol. 2, no. 1, pp. 114-118, 2012.
[19] A. K. Al-Mashhadany, "Text summarization based on several natural language techniques," *Eng. & Tec. Journal, University of Technology, Baghdad/ Iraq*, vol. 32, pp. 379-385, 2014.
[20] A. K. Al-Mashhadany, M. A. Sura, and K. T. Sawsan, "Text similarity based on english morphological analyzer approach," *Journal of Al Nahrain University- Sience, Baghdad/ Iraq*, vol. 17, no. 4, pp. 203-212, 2014, doi: 10.22401/JNUS.27.
[21] A. K. AL-Mashhadany, N. H. Dalal, T. S. Ahmed, W. K. Al-Mashhadany, "Extracting numerical data from unstructured Arabic texts (ENAT)," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 21, no. 3, pp. 1759-1770, 2021, doi: 10.11591/ijeecs.v21.i3.pp1759-1770.
[22] M. Mawlood, "Alasaleeb alhisabh fi hel almasail alirthyh- qadym whadeeth*," The Great Imam University College*, Iraq, 2014.
[23] B. Afroza, "Implication of islamic law of inheritance: ultimate solution to family conflict*," Asian Journal of Applied Science and Engineering*, vol. 2, no. 2, pp. 118-126, 2013.
[24] R. A. Fipy and A. Haswida, "The Implementation of inheritance based on the tribe of kaili ledo and islamic inheritance laws*," Jurnal Dinamika Hukum*, vol. 19, no. 2, pp. 218-232, 2019, doi: 10.20884/1.jdh.2019.19.2.2525.
[25] M. Mawlood, "Alqassam Alshariy Program*," Iraqi Ministry of Justice*, Baghdad, Iraq, 2006.
[26] H. Aydemir, "Program of inheritance division," *Bursa/ Uludag University, Turkey*, 2010.
[27] A. M. Thamir, "Derash fiqhih fanih lbarnamig alqassam alshariy*," Faculty of Sharia, Department of Jurisprudence*, 2013.
[28] H. F. AL-Shahad and A. Zeina, "Computation of inheritance share in islamic law by an expert system using decision tables," *Quarterly Adjudicated Journal for Natural and Engineering Research and Studies*, vol. 1, pp. 105-114, 2018.
[29] S. Zouaoui and K. Rezeg, " Islamic inheritance calculation system based on Arabic ontology (AraFamOnto)," *Journal of King Saud University – Computer and Information Sciences*, vol. 33, pp. 68–76, 2021, doi: 10.1016/j.jksuci.2018.11.015
[30] M. A. Al-Khatrawy, "Alraid fi Ilm Alfaraid," *Mktebh Dar Alturath*, Baghdad, Iraq, 1988.

BIOGRAPHIES OF AUTHORS

Abeer Khalid Al-Mashhadany received her B.Sc. in Computer Science in 2002 and M.Sc. in Computer Science (Artificial Intelligence) in 2005 from Al-Nahrain University, Baghdad, Iraq. She is Assistant Prof. in A.I. since 2015. Her research interests include Artificial intelligence, Natural language processing, and analysis of Arabic and English texts. She can be contacted at email: aabeeeeraa@yahoo.com.

Marwan B. Mohammed received her B.Sc. in software engineering in 2008 from Baghdad College of Economic Sciences University and M.Sc. in Computer Science in 2015 from Baghdad University, Baghdad, Iraq. He is an Assistant teacher. in A.I. since 2015. Her research interests include Artificial intelligence, Natural language processing, security, and English texts. He can be contacted at email: marwan.badran@nahrainuniv.edu.iq.

Mawlood Mukhlis Alrawi received his B.Sc. in electrical engineering from the University of Technology-Baghdad in 1979 & B.Sc. in Islamic sciences from the College of Sharia and Law at the Iraqi University in 2010 & M.Sc. in (Jurisprudence and its Foundations) in 2014, & Ph.D. in (Comparative Jurisprudence) in 2017, from Imam Al-Azam University College, Baghdad, Iraq. His research intrests in solving inheritance and commandments problems. He can be contacted at email: mawlood@gmail.com.