

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/296223713>

# Finding Arabic Poem Meter using Context Free Grammar

Article · March 2013

CITATIONS

5

READS

76

3 authors:



**Mohammad Alnagdawi**  
Universiti Teknologi Malaysia  
4 PUBLICATIONS 6 CITATIONS

SEE PROFILE



**Hasan Rashaideh**  
Al-Balqa' Applied University  
18 PUBLICATIONS 40 CITATIONS

SEE PROFILE



**Ala F A Aburumman**  
University of South Australia  
7 PUBLICATIONS 34 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Implementing Voice for Mobile Ad hoc Networks [View project](#)

# Finding Arabic Poem Meter using Context Free Grammar

Mohammad A. Alnagdawi · Hasan Rashideh · Ala' Fahed Aburumman

Received: 24 January 2013/ Accepted: 10 March 2013

**Abstract** In this paper, we make a program that can find the poem meter name (called Bahar in Arabic) by Arud science, this science provides a methodology to classify Arabic poems into 16 meters, to help a user to find the meter name for any Arabic poem using context free grammar (CFG). And we discuss the solutions for problems, from the starting phase to the result, using regular expression and CFG. And the result 75% of the verse is found its meter, when input enters correctly.

**Keywords** Arud · Context Free Gramar · CFG · Regular expression · Arabic poem meter.

## 1 Introduction

The Arud science, collected and explained by Arabic scientist called Al-Farahidi, is concerned of finding the Arabic meter name by studying rhyming verses (verse called Bayt in Arabic), through a sequence of the vowel sound (called Haraka in Arabic) and the constant sound (called Sukoon in Arabic) in a specific verse, then finding the meter name that the verses belonging to. Finding the poem meter in a traditional way is hard work for a non-expert, this problem encourages us to build this program. To consider any verse as rhymed, it should belong to the one of the sixteen Arabic meters of the poetry rhythms. If not, this poem is expressed as deviate.

In this paper, a new computerized method is proposed to find the poetry meters using CFG, we generate program code using tool called ANTLR (Another Tool for Language Recognition [1]).

Mohammad A. Alnagdawi · Hasan Rashideh · Ala' Fahed Aburumman

Computer Science Department, Faculty of Graduate Studies, AL-Balqa' Applied University, Jordan.

E-mail: nagdawi@gmail.com, rashideh@gamil.com, alafahid@gmail.com

## 2 Related work

M. A. Ismail and others use the “Expert System for Testing the Harmony of Arabic Poetry”. The design of ESHT (Expert System harmony Test) is to check the name of the meter in three steps: First step: the user enters the poetry and heuristic rule in knowledge base converting text to Arud form, that means to keep any letter is spoken and discard any letter is not spoken, and the heuristic knowledge base is written by RBS (Rule-Based Systems) Second step: is converting from Arud form to binary form (0's and 1's) using deterministic knowledge, deterministic means same inputs give same outputs and by passing the same rules. Last step: this is the core of this system, consultation which is taken by human expert knowledge stored in a knowledge base. But the author doesn't mention the experiment result.[2]

Shalabi, Ryadh and others use the “Computing System for Analyzing Arabic Poems Meter” to find the Arabic poem meter name through two steps: First step, by using Arud rules which analyze the verse and define long and short sounds. Second step: using the generated string of short and long sounds, comparing between rhythms of each meter one by one with the rhythms of verses that are needed to check. Finally, if algorithm succeeds to find the correct rhyme for this verse, then it will find the correct Arabic meter name. This algorithm implemented using Turbo C under MS-DOS, and this tool can't be used in present operating systems.

This paper only contains algorithm, without implementation section or experiment results.[3]

## 3 Methodology

This tool (also called program) implementation using 3 phases: first phase, converting from poetry form to Arud form. Second phase, is segmentation phase. Third phase, is detecting the poem meter name of poem. Each phase prepares the input to the next phase and so on, until reaches to the final result (meter name). And now we will discuss the 3 phases in detail:

3.1 Converting from poetry form to Arud form

To convert any text to Arud form, we just need to write any letter is read and discard any letter is not

read, below are the rules to do this work, and we write these rule using Regular Expression[4] to make proper action:

Table 1. Arabic letter			
Diacritic	Letter in Paper	IPA	Definition
ـَ	dhammah	u	High back rounded vowel
ـِ	fathah	a	Low short vowel
ـِى	kasrah	i	High front vowel
ـْ	shaddah	:	Geminate: consonant is double in duration
ـ	sukoon	Ø	The letter not diacritized nor geminated
ـً	tanween_fathah	an	Low vowel + alveolar nasal
ـٌ	tanween_dhammah	un	High back rounded vowel + alveolar nasal
ـٍ	tanween_kasrah	in	High front vowel + alveolar nasal
آ	alif maddah	/ʔaː/	
ء	hamza		
ا	alif	/aː/	
ى	alif_maqsurah	/aː/	
ب	ba	/b/	
ت	ta'	/t/	
ث	tha	/ θ /	
ج	jeim	[ʒ]	
ح	ha	/h/	
خ	kha	/x/	
د	dal	/d/	
ذ	dhal	/ð/	
ر	ra	/r/	
ز	zayn	/z/	
س	sein	/s/	
ش	shin	/ʃ/	
ص	sad	/sˤ/	
ض	sad	/dˤ/	
ط	sa	/tˤ/	
ظ	sa	[ðˤ]	
ع	ayn	/ʕ/	
غ	ghayn	/ɣ/	
ف	fa	/f/	
ق	qaf	/q/	
ك	kaf	/k/	
ل	lam	/l/	
م	mim	/m/	
ن	noon	/n/	
هـ	ha	/h/	
و	waw	/w/	
ي	ya	/e/	



Prepositional: Remove Alif maqsurah (ى) or Ya (ي) from preposition Ila, Fi or Ala (في/على/الى), also if sukoon comes after Alif maqsurah and Ya then the letter should be removed.[6] Example: see Table 7.

Table 7. Prepositional examples			
Word in English	Pronounce	After modified	Original word
In	Fi	فـِ	في
to	Ila	إِلَ	إلى
On	Ala	عَلَ	على

Rule:  
Fi (في):  
Reg <- fa'h (fathah|dhammah|kasrah|sukoon)?  
Ya (fathah|dhammah|kasrah|sukoon)?

Ala and Ila (إلى على):

Reg <-(Ayn|Alif) (fathah|dhammah|kasrah|sukoon)?  
Lam (fathah|dhammah|kasrah|sukoon)?  
Alif\_maqsurah(fathah|dhammah|kasrah|sukoon)?

Special cases: Add the letter to these words as mentioned in the Table 8: [5]

Table 8. Add alif to some words			
Word in English	Pronunciation	Poetry text	Arud Text
This (masculine)	Hatha	هذا	هَذَا
This (feminine)	Hatheh	هذه	هَآذِهِ
This (use with double masculine)	Hathan	هذان	هَآذَانِ
This (use with double feminine)	Hatain	هتین	هَآتَینِ
These	ha'ula'	هؤلاء	هَآؤُلَآءِ
That	Thalek	ذلك	ذَآلِکَ
Allah	Allah	الله	اَللّٰهُ
But	laken	لكن	لَآکِنْ
David	dawuud	داود	دَاوُودَ
Peacock	tawuus	طاوس	طَاوُوسٍ
Sarcophagus	nawuus	ناوس	نَاوُوسٍ

One rule as example:

Haza (هذا):

Reg <- Ha (fathah|dhammah|kasrah|sukoon)?  
Dhal (fathah|dhammah|kasrah|sukoon)?  
Alif (fathah|dhammah|kasrah|sukoon)?

And with other words, same rules apply, with changed letters.

3.2 Segmentation Phase

After completing previous step we can start applying the next step. It is called segmentation phase. The Main idea of segmentation phase is to divide input text (Arud form) to: short sound (s), vowel (Harakh) or long sound (l), vowel (Haraka) flowed by constant (sukoon). In other words, if it is one letter so it is called short sound and if it is two letters, so it is called long sound[7].

After finishing this phase, we get sound string, and send the sound string to the next phase to find meter name, and this is the last phase.

Example 1:

سَتُبْدِي لَكَ الْيَوْمَ مَا كُنْتَ جَاهِلًا  
وَيَأْتِيكَ بِالْأَخْبَارِ مَنْ لَمْ تُزَوِّدِ

Show the segmentation example in Table9

Table 9. How segmentation work								
Arud	تَزَوِّدِ	رِ مَنْ لَمْ	لَكَ بِنِ اِخْبَارًا	وَيَأْتِي	تَجَاهِلُنْ	مُ مَا كُنْ	لَ كُلُّ أَيُّ يَا	سَنُتَبْدِي
Encode	1 s 1 s	11 s	111 s	11 s	1 s 1 s	11 s	111 s	11 s

### 3.3 Detecting the poem meter name

To find the poem meter name for this verse, compare its sound string (come from previous phase) with all grammars. When any of these grammars is valid, that means this verse, belongs to this meter. Now, we will see one meter (called Taweel meter) as example for showing how this phase work.

Taweel Meter (الطويل): This meter contains only 2 types of rhythms repeated 4 times for each, they are called: Faulun and Mphaeiln (مفاعيلن - فعولن), and this meters drives into 3 rhythms which are: Faulo, Mphaeil and Mphaeiln (مفاعيلن - فعولن - مفاعيلن), see Table10 .[1]

This is the fully pattern:

فَعُولُنْ مَفَاعِيلُنْ فَعُولُنْ مَفَاعِيلُنْ  
فَعُولُنْ مَفَاعِيلُنْ فَعُولُنْ مَفَاعِيلُنْ

Fulan Mphaeiln Faulun Mphaeiln  
Fulan Mphaeiln Faulun Mphaeiln

And this is a list of all cases may exist in this meter with the code of them (see Table10 ):

Table10 . Taweel meter sound string.

Twl_d2	Twl_d1
Mphaeiln (مَفَاعِيلُنْ): 1 1 1 s	Faulun (فَعُولُنْ): 1 1 s
Mphaeiln (مفاعيلن): s 1 1 s	Faulo (فَعُولْ): s 1 s
Mphaeiln (مفاعِلْ): s 1 s 1	

Rule:

This is rule for Al-Taweel meter. It is built and suitable from right to left Arabic direction text, and Twl\_d2 or Twl\_d1: means one of rhythm that exists in Table 10.

R -> Taweel | Baset | Maded  
| Wafer | Kamal | Hazaj  
| Mutadark | Mutakareb | Rajz  
| Raml | Mujtath | Kafef  
| Mudare | Muktadeb | Munsareh | Saree  
Taweel -> Twl\_d2 Twl\_d1 Twl\_d2 Twl\_d1 Twl\_d2  
Twl\_d1 Twl\_d2 Twl\_d1  
Twl\_d2 -> 111s | s11s | 1s1s

Twl\_d1 -> 11s | s1s

Example 2: show the process for finding meter for this Arabic versus (verse) poem:

This is original verse poem:

أَبَا مُنْذِرٍ أَفْنَيْتَ فَاسْتَنْقِ بَعْضَنَا

الشَّرُّ أَهْوَنُ مِنْ بَعْضِ حَنَائِكَ بَعْضَ

In Table 11, first row is the poem verse converted to Arud form, and second row is sound string generated from Arud form (first row).

Table 11. Example 2 segmentation phase

Arud	نُ مِنْ بَعْضِ	رَاهُو	كُ بَعْضِ شَرِّ	ح نَانِي	قِ بَعْضِ نَا	تَ فَهِنْ تِي	ذِرْنُ أَفْ تِي	أَبَا مُنْ
Sound string	1 1 1 s	s 1 s	1 s 1 s	1 1 s	1 s 1 s	1 1 s	1 1 1 s	1 1 s

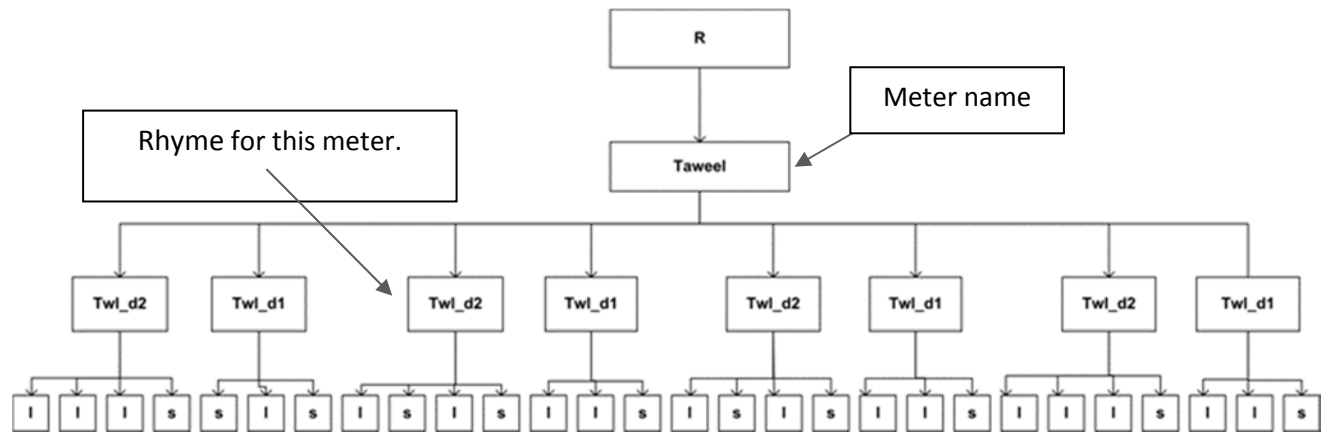


Fig. 1: Pares Tree for example 2

The parse tree in Fig. 1, shows us, how sound string finds the meter.

4 Conflict Between Meters

In some cases there exist verse belongs to 2 meters in the same time, and we can solve like this problem by defining sound string (long and sort string) as special case[8]. Now, discuss this problem using the following example:

أَلْمَوْتَ بَيْنَ الْخَلْقِ مُشْتَرِكٍ  
لَا سَوْقَهُ يَبْقَى وَ لَا مَلَكٌ

The sound string: “lsslslslslsslslslsl”.

The sound string mentioned above for this verse is accepted in 2 meters rule: Kamel and Saree meters, below find parse tree for 2 meters, see (Figure 2 and Fig. 3). Or you can use another verse form the same poem, to find correct meter name.

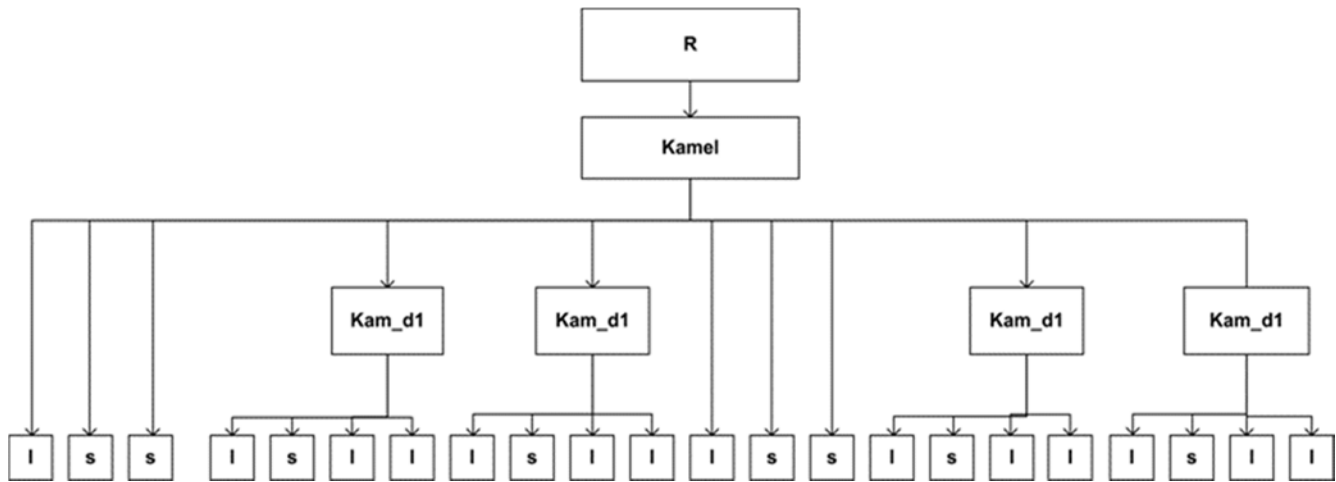


Figure 2. Kamel Meter Conflict





The analysis of the result shows 75% success ratio for a dataset with 128 verses of various Arabic poems. We plan to work on combining more algorithms and test it on a bigger dataset to reach an optimal solution with a high success ratio in the near future.

References

1. Mousa, Mahjoub. Al-Mezan (الميزان). s.l. : Madbole press, 1997. ISBN 997-208-198-9.
2. Expert System for Testing the harmony of Arabic poetry. M. A. Ismail and others. 2004.
3. Computing System for Analyzing Arabic Poems Meter. Shalabi, Ryadh, Kana'an, Gassan and AL-Jarah, Ahmad. Irbid : Yarmouk Research, 2003.
4. Regular Expressions. MSDN Library. [Online] Microsoft Corporation.  
<http://msdn.microsoft.com/en-us/library/2k3te2cs%28VS.80%29.aspx>.
5. Yamout, Phd. Gaze. Buhour Al-Sh'er Alaraby 2nd edition (بحور الشعر العربي). s.l. : Dar Al-Fikr Al-Lebnane, 1992.
6. Arud and Qwafi scince. [Online]  
<http://www.angelfire.com/biz/kha98/lifeskills/arabic/arabic5.htm>.
7. Younes, Ali. Poem Awzan and Qwafi (مدخل أوزان الشعر و قوافيه (ميسر لتتوقها و دراستها). Cairo : Dar Ghareeb, 2005. 977-215-855-8.
8. Ibrahim, Mahmoud. Al-Mujtathe and Al-Muktadeb meter (البحر المجتث و المقتضب). Language Jewels (جواهر اللغة). [Online]  
[http://zahra1.com/Edu\\_Letur/AROOD/AL3ROOD\\_Mgta th&Mqtadab.htm](http://zahra1.com/Edu_Letur/AROOD/AL3ROOD_Mgta th&Mqtadab.htm).
9. Alfred V. Aho and others,. Compilers : principles, techniques, and tool 2nd ed. s.l. : Pearson Education, Inc, 2007. ISBN 0-32 1-4868 1.
10. Meter Poetry- Al -Wafer (الوافر -البحور الشعرية). Arabic Network Subject (شبكة مادة العربية). [Online] [Cited: May 20, 2010.]  
<http://www.arabe.edunet.tn/aroudh/html/index12.html>.
11. Al-Wasel, Saad Abdullah. (موسوعة العروض و القافية). Encyclopedia of Arud and Rhyme. 2005.
12. Parr, Terence. ANTLR - Home. ANTLR. [Online] [Cited: December 6, 2010.] [www.antlr.org](http://www.antlr.org).
13. Liberty, Jesse and Xie, Donald. Programming C# 3.0. s.l. : O'Reilly Media, Inc, 2007. ISBN-0596527438.