



Computational Quranic Linguistics @ Leeds University

Presented By

Abdul-Baquee Sharaf

The Team:

Eric Atwell, Kais Dukes, Majdi Swalha,

Language Research Group

I-AIBS: Institute for Artificial Intelligence and Biological
Systems

University of LEEDS, England



Brief Computation Corpus linguistic research on Modern Standard Arabic

Computation Corpus linguistic research on the Quran

- Qurany Subject Browser**
- Quranic Arabic Corpus**
- Tools on Text Mining the Quran**

A proposal for Future research: Quranic Knowledge Map

Future collaboration work

Tips for Text Mining Hadeeth Project

Arabic computing research at Leeds: Modern Arabic



UNIVERSITY OF LEEDS

Arabic corpus-trained chatbot

Corpus of Contemporary Arabic

aConCorde: concordance for Arabic texts

SALMA morphological analysis and tag-set

Arabic lexical resource from traditional Arabic dictionaries

Discourse Treebank for Modern Standard Arabic

180-Million-word Arabic Web Corpus, online concordance

<http://www.comp.leeds.ac.uk/arabic>

Arabic computing research at Leeds: Quran as Corpus



UNIVERSITY OF LEEDS

Quran chatbot: replies with verse from Quran

Qurany: browse Quran by concepts

Morphochallenge: Quran as Gold Standard for evaluation

Quranic Arabic Corpus: morphology and syntax annotations

Text mining the Quran: related verses; pronoun coreferences

“Qurany : Quran Subject Browser



UNIVERSITY OF LEEDS

<http://quranytopics.appspot.com/>

Qurany Concepts/Topics - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://quranytopics.appspot.com/

Print Save Verses Contact us About Qurany

Qurany قرآنك

Keyword Search

المشركون المنكرون لمره المعث
The Apostates Who Denies the Resurrection
وتعهد المصممين والمجرمين
والفاسقين
The Threatening of the Mischlevious, the Criminals and the Impious
الحمد (صلى الله عليه وسلم)
The Blessed Muhammad(PBUH)
أب النبيين معه
The Believers' Manners Towards Him(PBUH)
أخلاقه وصفاته
His Manners and Attributes
أزواجه وصفاته
His Wives and Daughters
أسرته وصفاته
His Midnight Journey to Jerusalem

أخلاقه وصفاته His Manners and Attributes

سورة آل عمران

فَإِذَا عَزَمْتَ فَتَوَكَّلْ عَلَى اللَّهِ إِنَّ اللَّهَ يُحِبُّ الْمُتَوَكِّلِينَ 159

Al-Emran

159

KHAN:- And by the Mercy of Allah, you dealt with them gently. And had you been severe and harsh-hearted, they would have broken away from about you; so pass over (their faults), and ask (Allah's) Forgiveness for them; and consult them in the affairs. Then when you have taken a decision, put your trust in Allah, certainly, Allah loves those who put their trust (in Him).

MAULANA:- Thus it is by Allah's mercy that thou art gentle to them. And hadst thou been rough, hard-hearted, they would certainly have dispersed from around thee. So pardon them and ask protection for them, and consult them in (important) matters. But when thou hast determined, put thy trust in Allah. Surely Allah loves those who trust (in Him).

<http://corpus.quran.com>

- A popular website (50,000 users per month)
- Used by researchers, scholars and students of the Quran and Arabic

Online Collaborative Annotation

- Anyone can view existing annotation
- Registered users can suggest corrections through a message board

Web-based Tools for Annotators and Researchers

- Concordance of the Quran (to see how related words have been tagged)
- Morphological search by root, lemma or stem
- 7 parallel translations into English for each verse
- Automatically generated phonetic transcription
- Natural language generation used to create grammatical summaries
- Audio recitation in Arabic

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Chapter (97) sūrat l-qadr (The Night of Decree)

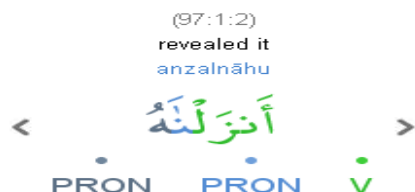
Translation	Arabic word	Syntax and morphology
(97:1:1) Indeed, We innā		<p>ACC – accusative particle</p> <p>PRON – 1st person plural object pronoun → Allah</p> <p>حرف نصب و «نا» ضمير متصل في محل نصب اسم «ان»</p>
(97:1:2) revealed it anzalnāhu		<p>V – 1st person masculine plural (form IV) perfect verb</p> <p>PRON – subject pronoun → Allah</p> <p>PRON – 3rd person masculine singular object pronoun → Quran</p> <p>فعل ماض و «نا» ضمير متصل في محل رفع فاعل والهاء ضمير متصل في محل نصب مفعول به</p>
(97:1:3) in fi		<p>P – preposition</p> <p>حرف جر</p>
(97:1:4) (the) Night laylati		<p>N – genitive feminine noun → Night of Decree</p> <p>اسم مجرور</p>
(97:1:5) (of) Power. l-qadri		<p>N – genitive masculine noun</p> <p>اسم مجرور</p>



Quranic Grammar - Word (97:1:2)

The second word of verse (97:1) is divided into 3 morphological segments. A verb, subject pronoun and object pronoun. The form IV perfect verb (فَعْلٌ مَاضٍ) is first person masculine plural. The verb's root is *nūn zāy lām* (ن ز ل). The suffix (هُ) is an attached subject pronoun. The attached object pronoun is third person masculine singular.

Chapter (97) sūrat I-qadr (The Night of Decree)



V – 1st person masculine plural (form IV) perfect verb
PRON – subject pronoun → Allah
PRON – 3rd person masculine singular object pronoun → Quran

فَعْلٌ مَاضٍ و«نَا» ضَمِيرٌ مُتَّصِلٌ فِي مَحَلِّ رَفْعِ فَاعِلٍ وَآلِهَاءِ
ضَمِيرٍ مُتَّصِلٌ فِي مَحَلِّ نَصْبِ مَفْعُولٍ بِهِ

See Also

- [Verbs, Subjects and Objects](#)
- [Dependency Graph](#) - visual syntax (*i'rāb*) for this verse
- [Concordance](#) - list occurrences of this word
- [Allah](#) - referred to by the subject pronoun
- [Quran](#) - referred to by the object pronoun

3 messages



Samir

11th January, 2010

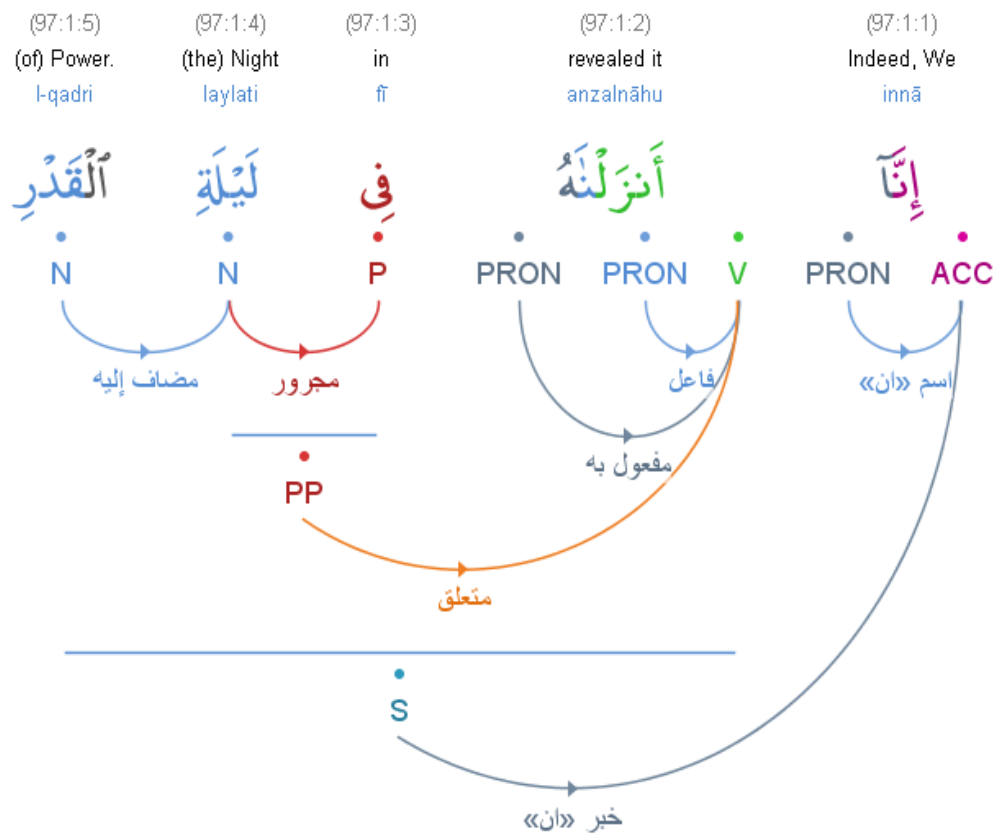
I suggest:

"V – 1st person plural (form IV) perfect verb"

It's not "masculine". there's no gender in the first person (either pl. or sing.).

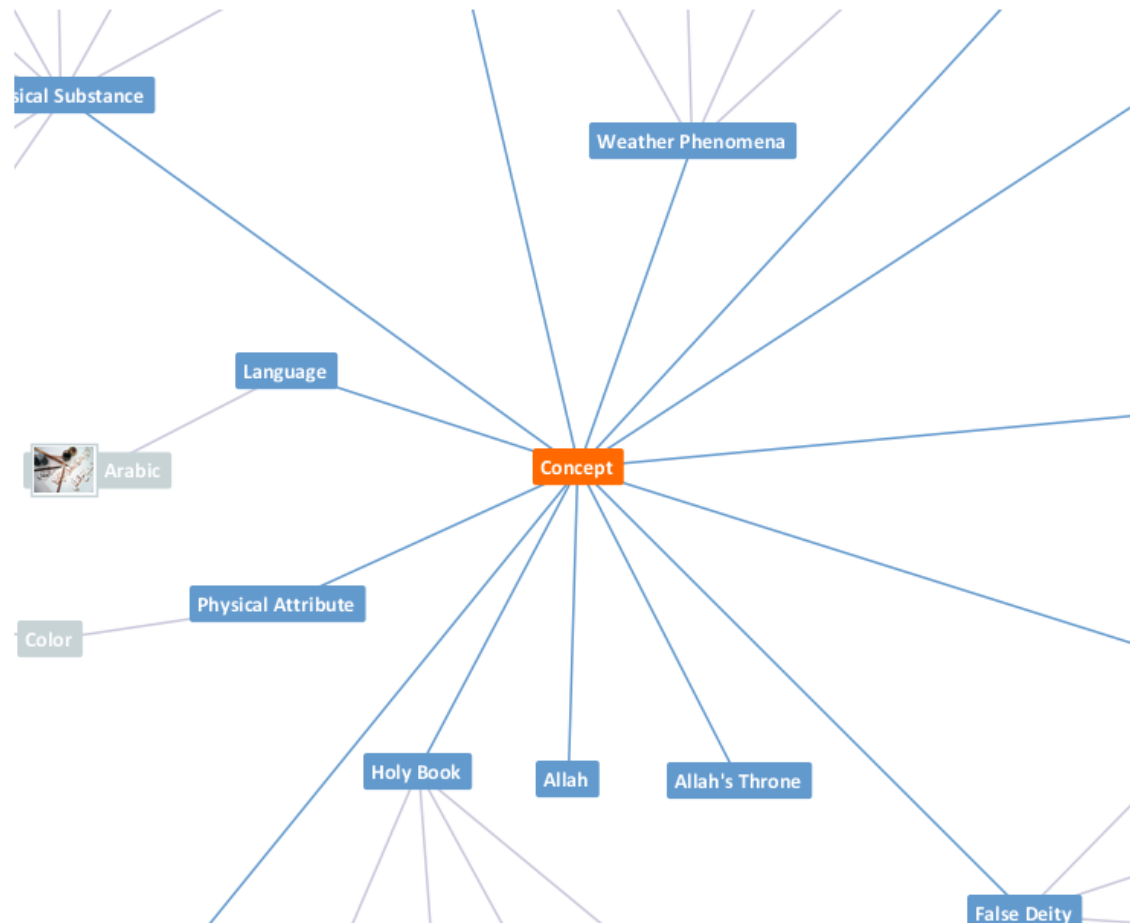
بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Chapter (97) sūrat l-qadr (The Night of Decree)



Visual Concept Map

The map below shows the part of the visual ontology for this concept. You can click on other concepts in the map for related information:



Quranic Analysis Tools : Concordance



UNIVERSITY OF LEEDS

2:52	وَأَذَاتِنَا مُوسَىٰ الْأَكْبَبُ وَالْفُرْقَانَ لَعَلَّكُمْ تَهْتَدُونَ	تَشْكُرُونَ	ثُمَّ عَفَوْنَا عَنْكُمْ مِّنْ بَعْدِ ذَلِكَ لَعَلَّكُمْ	1
2:56	وَوَضَّلْنَا عَلَيْكُمْ الْعُمَامَ وَأَنْزَلْنَا عَلَيْكُمُ الْمَنَّاءَ وَالسَّلْوَىٰ	تَشْكُرُونَ	تَنْظُرُونَ ثُمَّ بَعَثْنَاكُم مِّنْ بَعْدِ مَوْتِكُمْ لَعَلَّكُمْ	2
2:152	لِي وَآلَا تَكْفُرُونَ يَا أَيُّهَا الَّذِينَ آمَنُوا اسْتَعِينُوا	وَأَشْكُرُوا	وَيُعَلِّمُكُم مَّا لَمْ تَكُونُوا تَعْلَمُونَ فَأذْكُرُونِي أَذْكُرْكُمْ	3
2:158	عَلَيْمٌ إِنَّ الَّذِينَ يَكْفُرُونَ مَا أَنْزَلْنَا مِنْ	شَاكِرٌ	يَطُوفَ بِهِمَا وَمَنْ تَطَوَّعَ خَيْرًا فَإِنَّ اللَّهَ	4
2:172	لِلَّهِ إِنَّ كُنتُمْ إِيَّاهُ تَعْبُدُونَ إِنَّمَا حَرَّمَ	وَأَشْكُرُوا	الَّذِينَ آمَنُوا كُلُوا مِن طَيِّبَاتِ مَا رَزَقْنَاكُمْ	5
2:185	وَإِذَا سَأَلَكَ عِبَادِي عَنِّي فَإِنِّي قَرِيبٌ أُجِيبُ	تَشْكُرُونَ	الْعِدَّةَ وَلِتُكَبِّرُوا اللَّهَ عَلَىٰ مَا هَدَىٰكُمْ وَلَعَلَّكُمْ	6
2:243	وَقَاتِلُوا فِي سَبِيلِ اللَّهِ وَأَعْلَمُوا أَنَّ اللَّهَ	يَشْكُرُونَ	فَضْلٌ عَلَى النَّاسِ وَلَكِنَّ أَكْثَرَ النَّاسِ لَا	7
3:123	إِذْ تَقُولُ لِلْمُؤْمِنِينَ أَلَنْ يَكْفِيَكُمْ أَنْ يُمَدِّدْكُمْ	تَشْكُرُونَ	اللَّهُ بِدَرٍ وَأَنْتُمْ أَذِلَّةٌ فَاتَّقُوا اللَّهَ لَعَلَّكُمْ	8
3:144	وَمَا كَانَ لِنَفْسٍ أَنْ تَمُوتَ إِلَّا بِإِذْنِ	الشَّاكِرِينَ	عَقِبِهِ فَلَنْ يَضُرَّ اللَّهَ شَيْئًا وَسَيَجْزِي اللَّهُ	9

Word Cloud for the Quranic Chapters

Select Sura:

[Check All](#)

[Uncheck All](#)

[Madani](#)

[Makki](#)

102) Al-Takathur

103) Al-'Asr

104) Al-Humazah

105) Al-Fil

106) Quraysh

107) Al-Ma'un

108) Al-Kawthar

109) Al-Kafirun

110) An-Nasr

111) Al-Masad

112) Al-'Ikhlas

113) Al-Falaq

114) An-Nas

Word Cloud for Sura: 113,114,[**Stop words removed**]

[**Disclaimer:** The following order of words is **random** and does not reflect the original order of chapters of the Qur'an.]

حسد الذي حاسد الخناس العقد أعوذ
خلق الوسواس برب الفلق إله
والناس النفاتات **الناس** يوسوس
ملك **شر** غاسق الجنة قل
صدر

Quranic Analysis tools: Root co-occurrence



UNIVERSITY OF LEEDS

word : السماء

Submit

+2	+1	0	-1	-2
وما 6	ماء 20	السماء	من 51	أنزل 7
كانوا 3	والأرض 15	السماء	في 18	ولا 5
وإذا 3	وما 5	السماء	إلى 4	يوم 4
مداراً 3	ذات 3	السماء	خلقنا 2	برزقكم 3
به 3	من 3	السماء	وإذا 2	كسفا 3

<http://www.textminingthequran.com/apps/cooccur.php>

Very much abundant in the Quran

وَأَنَّهُمْ لَيَصُدُّونَهُمْ عَنِ السَّبِيلِ وَيَحْسَبُونَ أَنَّهُم مُّهْتَدُونَ

and We raised over you the mount [2:63]

الطُّورَ	فَوْقَكُمْ	وَرَفَعْنَا
at-Tour	over you	and we raised

children of Israel

Allah

And [mention] when We raised the mountain above them as
if it was a dark cloud [7:171]

فَوْقَهُمْ	الْجَبَلَ	نَتَقْنَا	وَإِذْ
other them	the mountain	we raised	and when

children of Israel

at-tour

Allah

gloss	Concept	Verse	Pronoun context	Antecedent #
Prophet Muhammad	محمد	3:3	إِلَّا هُوَ الْخَيُّ الْقَيُّومُ نَزَّلَ عَلَيْكَ الْكِتَابَ بِالْحَقِّ مُصَدِّقًا لِمَا بَيْنَ يَدَيْهِ	1
the Qur'an	القرآن	3:3	الْكِتَابَ بِالْحَقِّ مُصَدِّقًا لِمَا بَيْنَ يَدَيْهِ وَأَنزَلَ التَّوْرَةَ وَالْإِنْجِيلَ مِنْ قَبْلُ هُدًى	2 الْكُتُبِ

:Concordance lines for the concept Prophet Muhammad محمد

[78:36]	لَعُوًّا وَ لَأَكْفِبُنَا جَزَاءً مِّن رَّبِّكَ ۗ عَصَاً حِسَابًا رَبِّ آلِ سَمُوتَ وَ آلِ	1
[79:45]	إِلَىٰ رَبِّكَ ۗ كَافٌ مُّنتَهَىٰ ۗ هَا إِنَّ مَا أَنْتَ مُنذِرٌ مِّن يَحْضِيٰ هَا كَأَنَّ هُمْ يَوْمٌ	2
[79:44]	مَ أَنْتَ مِن ذِكْرِي ۗ هَا إِلَىٰ رَبِّكَ ۗ مُنْتَهَىٰ ۗ هَا إِنَّ مَا أَنْتَ مُنذِرٌ مِّن	3
[79:43]	آلِ سَاعَةَ أَيَّانٍ مُّرْسَىٰ ۗ هَا فِي مَ أَنْتَ مِن ذِكْرِي ۗ هَا إِلَىٰ رَبِّكَ ۗ كَافٌ مُّنتَهَىٰ	4
[79:42]	آلِ جِنَّةٍ هِيَ آلُ مَأْوَىٰ يُسْتَلُّ وَن ۗ عَن آلِ سَاعَةَ أَيَّانٍ مُّرْسَىٰ ۗ هَا فِي	5
[79:15]	إِذَا هُمْ بِآلِ سَاهِرَةٍ هَلْ أَرِيتِ ۗ حَبِيبَتِ مُوسَىٰ إِذْ نَادَىٰ هُوَ رَبُّ هُوَ	6
[80:2]	صَحِيٰ ۗ هَا عَيْسَىٰ وَ تَوَلَّىٰ أَن جَاءَ هُوَ ۗ آلِ أَحْمَرَ ۗ وَ مَا يُدْرِيكَ لَكَ لَعَلُّ	7
[80:3]	جَاءَ هُوَ ۗ آلِ أَحْمَرَ ۗ وَ مَا يُدْرِيكَ لَكَ لَعَلُّ هُوَ يُرْسِيٰ ۗ أَوْ يُدْعَرُ فَ تَنْفَعُ	8
[80:6]	هُوَ ۗ آلِ ذِكْرِي ۗ أَمَّا مَنِ اسْتَعْجَلَ فَ أَنْتَ لَ هُوَ مُصَدِّقٌ ۗ وَ مَا عَلِيٰ لَكَ	9
[80:7]	أَنْتَ لَ هُوَ مُصَدِّقٌ ۗ وَ مَا عَلِيٰ لَكَ ۗ أَلَا يُرْسِيٰ ۗ وَ أَمَّا مَنِ جَاءَ	10
[80:8]	ۗ أَلَا يُرْسِيٰ ۗ وَ أَمَّا مَنِ جَاءَ لَكَ ۗ يَسْعَىٰ ۗ وَ هُوَ يَحْضِيٰ فَ أَنْتَ عَن	11
[80:10]	جَاءَ لَكَ ۗ يَسْعَىٰ ۗ وَ هُوَ يَحْضِيٰ فَ أَنْتَ عَن هُوَ تَلْهَىٰ كَلِمًا إِنَّ هَا تُذَكِّرُهُ	12
[81:24]	بِآلِ أَهْلِ آلِ مُبِينٍ وَ مَا هُوَ عَلَىٰ آلِ غَيْبٍ بِ صَافِيٍّ وَ مَا	13
[85:12]	آلِ قَوْمٍ آلِ كَبِيرٍ إِنَّ يَحْضِيٰ رَبِّكَ ۗ لَ تَنْبِيءٌ إِنَّ هُوَ يُبْدِيٰ ۗ وَ	14
[85:17]	مَجِيدٌ فَعَلَّ لَ مَا يُرِيدُ هَلْ أَرِيتِ ۗ حَبِيبَتِ آلِ جُودٍ فِرْعَوْنَ وَ تَمُودَ بَل	15
[86:2]	سَمَاءَ وَ آلِ طَارِقٍ وَ مَا أُنزِيَٰ لَكَ ۗ مَا آلِ طَارِقٍ آلِ تُحْمِ آلِ تَابِ	16



Can we classify Quranic Chapters automatically into Makki and Madani?

Machine Learning Steps:

1. Identify feature set
2. Create training data
3. Run machine learning classifiers on training set
4. Evaluate performance of ML classifiers through test set
5. Improve feature set, ML classifiers, etc.

Makki vs Madani: features



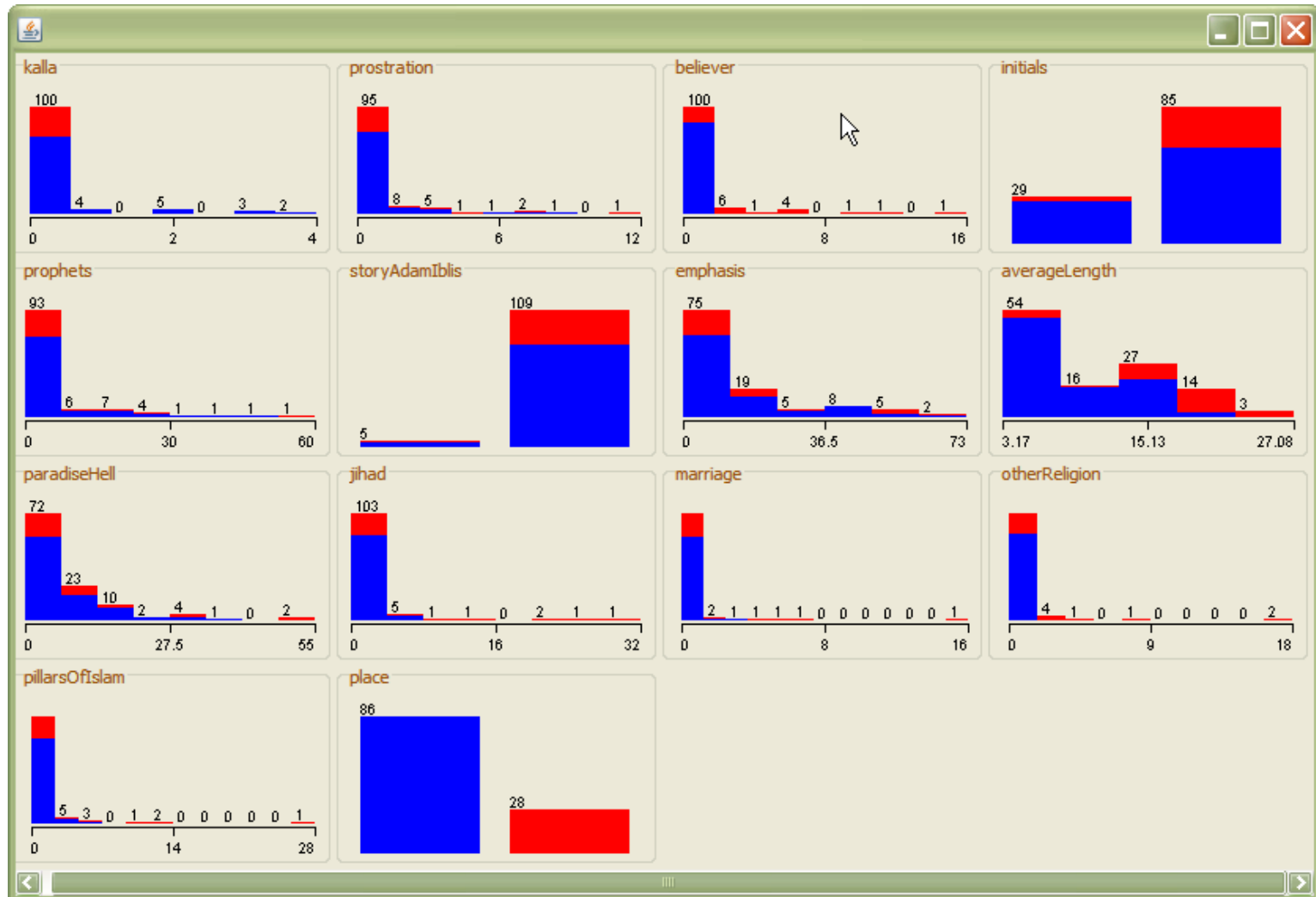
UNIVERSITY OF LEEDS

No	Feature	Category
1	Āyat of sajdah (Prostration)	Meccan
2	Use of aversion letter kalla	Meccan
3	The phrase: ya ayyiha alladhyna aamanu	Medinan
4	Qur'anic initials	Meccan
5	Story of Adam and Iblis	Meccan
6	Stories of previous prophets	Meccan
7	Instruments of denunciation and excoriation	Meccan
8	Average aya length	
9	Eschatological topics	Meccan
10	Reference to Jihad	Medinan
11	Marriage and Divorce	Medinan
12	Reference to the people of the Book	Medinan
13	Reference to the pillars of Islam	Medinan

Visual representation of features



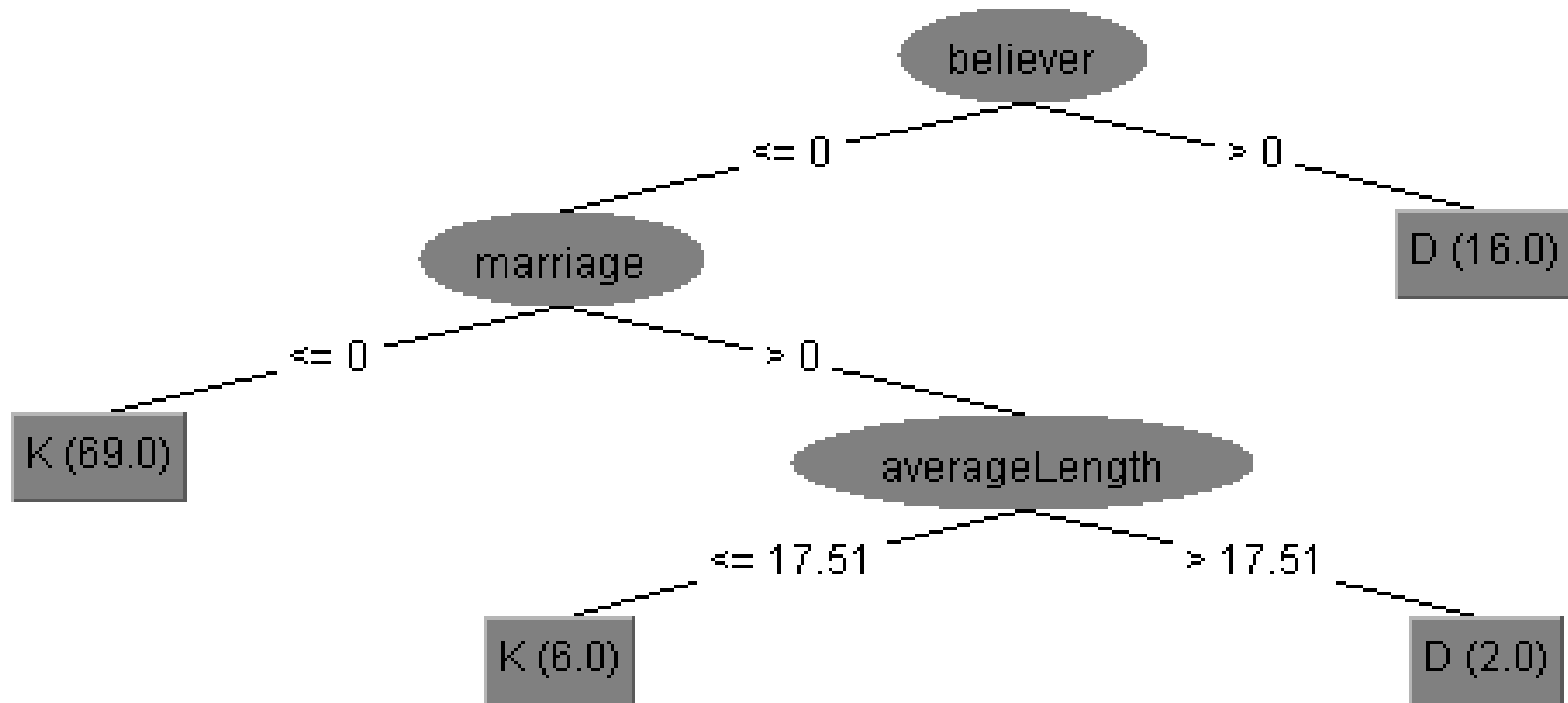
UNIVERSITY OF LEEDS



Decision trees



UNIVERSITY OF LEEDS



Automatic Classification



UNIVERSITY OF LEEDS

sura no.	kalla	prostrate	believer	initials	prophets	adam-iblis	emphasis	length	eschatology	jihad	marriage	other-religion	pillars	category	prediction
1	0	0	0	NO	0	NO	0	4.1	0	0	0	0	0	K	K
13	0	1	0	YES	0	NO	16	19.84	9	0	0	0	1	D	K
16	0	2	0	NO	4	NO	41	14.41	25	2	0	0	0	K	K
22	0	5	1	NO	6	NO	21	16.4	17	4	0	1	7	D	D
29	0	0	0	YES	13	NO	46	14.17	17	5	0	1	2	K	K
47	0	0	2	NO	0	NO	7	14.26	4	3	0	0	0	D	D
55	0	1	0	NO	0	NO	0	4.51	6	0	0	0	0	D	K
61	0	0	3	NO	3	NO	1	16.14	1	2	0	1	0	D	D
64	0	0	1	NO	1	NO	2	13.44	2	0	0	0	0	D	D
76	0	1	0	NO	0	NO	0	7.84	3	0	0	0	0	D	K
80	2	0	0	NO	0	NO	2	3.17	0	1	0	0	0	K	K
83	4	0	0	NO	0	NO	10	4.69	0	0	0	0	0	K	K
89	2	0	0	NO	0	NO	3	4.63	4	0	0	0	0	K	K
92	0	0	0	NO	0	NO	4	3.38	2	0	0	0	0	K	K
97	0	0	0	NO	0	NO	0	6	0	0	0	0	0	K	K
98	0	0	0	NO	0	NO	2	11.75	2	0	0	2	2	D	K
99	0	0	0	NO	0	NO	0	4.5	0	0	0	0	0	D	K
110	0	0	0	NO	0	NO	0	6.33	0	0	0	0	0	D	K
112	0	0	0	NO	0	NO	0	3.75	0	0	0	0	0	K	K
113	0	0	0	NO	0	NO	0	4.6	0	0	0	0	0	K	K
114	0	0	0	NO	0	NO	0	3.33	0	0	0	0	0	K	K

Automatic Clustering



UNIVERSITY OF LEEDS

1	2	3	4	5	6	7	8
2 D	6 K	15 K	74 K	8 D	1 K	31 K	3 D
7 K	10 K	19 K	83 K	24 D	50 K	35 K	4 D
	11 K	21 K	96 K	33 D	53 K	36 K	5 D
	12 K	23 K			55 D	44 K	9 D
	13 D	26 K			56 K	45 K	
	14 K	30 K			69 K	46 K	
	16 K	32 K			72 K	47 D	
	17 K	34 K			76 D	49 D	
	18 K	37 K			77 K	51 K	
	20 K	43 K			79 K	52 K	
	22 D	54 K			84 K	57 D	
	25 K	68 K			86 K	58 D	
	27 K				88 K	59 D	
	28 K				90 K	60 D	
	29 K				91 K	61 D	
	38 K				92 K	62 D	
	39 K				93 K	63 D	
	40 K				94 K	64 D	
	41 K				95 K	65 D	
	42 K				97 K	66 D	
	48 D				99 D	67 K	
					100 K	70 K	
					101 K	71 K	
					103 K	73 K	
					105 K	75 K	
					106 K	78 K	
					108 K	80 K	
					109 K	81 K	
					110 D	82 K	
					111 K	85 K	
					112 K	87 K	
					113 K	89 K	
					114 K	98 D	
						102 K	
						104 K	
						107 K	



Our resources are open-source rather than commercial; this is why they have been widely re-used, compared to resources kept “in-house” by other Arabic NLP research groups.

Our Quranic Arabic Corpus website <http://corpus.quran.com/> shows the advantages of making resources open-source : publications, press articles, Message Board for feedback, Google Analytics visualisation of global distribution of visitors to the website.

Understanding the Quran – a Grand Challenge for AI



UNIVERSITY OF LEEDS

Machine Learning research needs a “Gold Standard” – a corpus where each text is classified and marked up by experts, so ML can learn the classification.

The Quran is an excellent Gold Standard: many expert analyses exist (Tafsir), we can use these to train ML

Quranic scholarly work can ensure that Knowledge Based Systems based on the Quran are logically consistent and correct

Understanding the Quran – a Grand Challenge for AI



UNIVERSITY OF LEEDS

Huge worldwide interest in the Quran means we can harness volunteers for “crowd-sourcing” analysis

Quranic Arabic Corpus: initial automatic analysis, then proofreading and correction by many volunteers

A proposal for further research: the Quranic Knowledge Map



UNIVERSITY OF LEEDS

Understanding the Quran is a grand challenge for society, for western public education, for Muslim-world education, for knowledge representation and reasoning, for knowledge extraction from text, for systems robustness and correctness, and for online collaboration.

Understanding the Quran is a grand challenge for computer science and artificial intelligence

We propose a collaborative research effort to construct a Quranic Knowledge Map to address this challenge.

Three strands of research



UNIVERSITY OF LEEDS

Infrastructure. A set of tools used to develop the Quranic Knowledge Map: Arabic Natural Language Processing tools, tools for online collaborative annotation, and tools for knowledge engineering and automated reasoning.

Datasets. Tagging the Quran with morphology, syntax, semantics, pronoun and named entity references, concept ontology, other KR formalisms. Also, extending beyond Quran to linked Classical Arabic texts: Hadith etc. Each of these datasets is expected to be highly useful for further research and worthy in publication and distribution in itself.

End-user applications. These form the main contribution of the Quranic Knowledge Map to society, i.e. to interested researchers, students and public who will use the system.

Modules in the Quranic Knowledge Map



UNIVERSITY OF LEEDS

	INFRASTRUCTURE	DATASETS	APPLICATIONS
	Reusable Computational Tools	The Quranic Arabic Corpus	Online User Access
CORE TEXT ANNOTATION	Software for Classical Arabic NLP Arabic Morphological Analyzer Arabic Syntactic Parser Arabic Natural Language Toolkit Multi-lingual Word Alignment	Quranic NLP Datasets Quranic Arabic Text Morphological Tagging Syntactic Treebank Translations & Audio	Baseline Quranic Resource Online Tagged Quran Morphological Search Quranic Grammar Annotations Interlinear Translations
	Tools for Collaborative Annotation Linguistic Database Manual Annotation Tools Online Collaborative Annotation	Quranic Linguistic Datasets Pronoun Resolution Named Entity Resolution Quranic WordNet	Quranic & Arabic Linguistics Electronic Lexicon & Dictionary Word-sense Disambiguation Arabic Educational Resources
FURTHER LINGUISTICS	Knowledge Representation Semantic Annotation Framework Semantic Database QA & Information Retrieval Automated Reasoning and Inference	Quranic Knowledge Datasets Ontology of Quranic Concepts Quranic PropBank / FrameNet Related Texts: The Hadith Knowledge Representation	Quranic Knowledge Online Concept Topic Map & Search Verse similarly concordance Quran to Hadith Linkage Question Answering
QURANIC KNOWLEDGE			

Collaborators needed



UNIVERSITY OF LEEDS

Researchers with the following expertise:

Management and coordination: leading researchers acting in a supervisory capacity, ideally with skills in fund-raising and public relations!

Religious Studies experts, with experience in the Quran and exegesis (tafsir); and also experts in other religious texts

Full time annotators, familiar with Arabic and the Quran

NLP people, with good proven Arabic language computation skills

Software engineers for general infrastructure and web development, not necessarily with NLP skills

E-learning experts, ideally with a background in developing Arabic online language-learning or religious educational resources

In addition, it is expected that the project will leverage a large body of existing expert volunteers worldwide through collaborative annotation

Tips for Text Mining Hadeeth Project



UNIVERSITY OF LEEDS

- Stipulate a set of goals and objectives
- Setup a website (both Arabic and English) to make the world aware
- Participate in workshops and conferences
- Have clear understanding on the capability of computational approaches, and their performance in terms of precision and recall
- Based on objectives, create a Corpus and decide on the type of annotation needed
- Involve subject matter experts from Day-1
- Keep solutions open-sourced and exploit “crowd-sourcing” by publishing results online

Brief Computation Corpus linguistic research on Modern Standard Arabic

Computation Corpus linguistic research on the Quran

- Qurany Subject Browser**
- Quranic Arabic Corpus**
- Tools on Text Mining the Quran**

A proposal for Future research: Quranic Knowledge Map

Future collaboration work

Tips for Text Mining Hadeeth Project