
**

*

2008/1/23:

2008/5/8:

B+

B+

.

.

:

.

Abstract

Usually, the natural language grammar is represented as a rules inside the program of natural language processing (NLP), but with a language such as the Arabic language that have a flexible syntax, using this representation will produce a large number of these rules that lead to increase the search time required and to increase the size of the program, so in this paper we proposed a new representation method. We represent the Arabic language grammar as logical terms in the knowledge base, and we use the B+tree to store these terms since it is suitable with large data, so we reach the required syntactical term in a little time by using the suitable searching method that b+tree provide it.

In this paper we choose the Augmented Transition Network(ATN) to represent the language grammar since it is the most suitable parsing technique for a language such as the Arabic language.

.[1]

.

Context Free)

(Grammar

.1

:[3]

الجملة ← فعل + فاعل + مفعول به
 فعل ← كتب
 فاعل ← اسم
 مفعول به ← اسم
 اسم ← أداة تعريف + اسم
 اسم ← ولد
 اسم ← درس

.2

.3

(Natural Language Processing)

(States)

.(Transitions)

()

.()

" "

." "

(Phonology)

(Syntax)

(Morphology)

(Semantics)

.[2](Lexicon)

(Pragmatics)

- RTN)

(Recursive Transition Network

-ATN)

(Augmented Transition Network

1

(Grammar)

ATN

.[4]1970 William A.Woods

¹ القواعد (Grammar) للغة هي المفاهيم التي تحدد الجمل المسموحة في تلك اللغة.

(Registers)

)

(...) :ATN

(...) Saad builds houses in Iraq.

: ATN

[[NP Saad][VP builds [NP houses][PP in [NP Iraq]]]]

. Path :Int_list (Saad)

) (builds houses in Iraq)

.(... :Roles (Houses) (builds)

:(in Iraq)

.(noun phrase) :NP

.(verb phrase) :VP

preposition) :PP

.(phrase

B+ (Terms)

(Logical Terms)

.(B+ tree)

B

list String

nested list

[2] B+ .(Compound term

B+

(data structure)

(Predicate Calculus)

B+

index to)

(database

.(indices)

.[5]

"bnf"

:

:New States

B+)

(leaves) (key) .(ATN

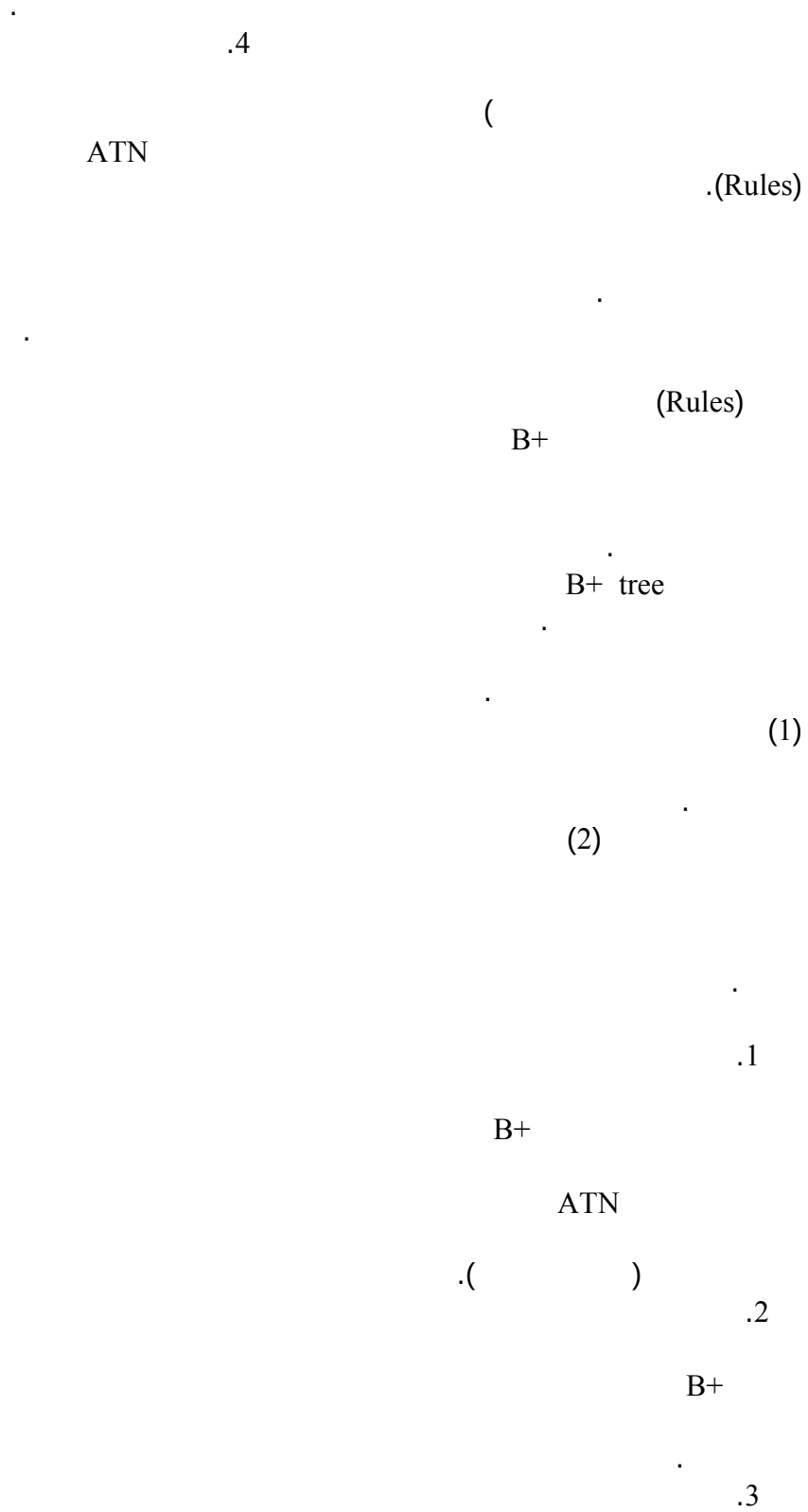
:Path

(reference number)

" " B+
 . " "
 .
 B+
)
 (
 (1) .B+
 B+
 (leaves)
 :
 B+ .
 (binary tree)
 B+
 (node)
 B+
 "s,[[2,1]]"
 :
 ")). bnf([np1],[[8]],[9],[
 : "np1,[[6,5]]"
 " bnf([m5],[[6,5]],[13]
 .
 " m5,[[2,1],[6,5]]"
 :
 . bnf([np1], [19], :
 "np1,[[2,1],[6,5]]"
 :
 . bnf([np1], [19]
) :
 (Type

² يتم تحويل المفتاح إلى سلسلة حرفية لأن مفتاح
 الهيكل الشجري هو سلسلة حرفية.

"	:	2-8	(Count)
.		"	
.(19)		3-8	
:		.9	(Path)
bnf(NewStateList,List,IntList,R		oles)	(ATN)
		.10	
			.(StateList)
			StateList .1
:		.11	:
	:	1-11	1-1
		.(12)	(s) : 2-1
"	:	2-11	.StateList
.		"	.2
.(19)		3-11	.State StateList
		.12	.3
.Roles		.13	:)
.IntList		.14	: (
NextState		.15	.(4) : 1-3
.NewStateList		.16	LastType : 2-3
NextState		.17	.
.StateList		.18	.(5) 3-3
.List SubType		.19	Path .4
Path			:
.Count SubType			LastType : 1-4
			.Path
			.(5) 2-4
			.(7) : 3-4
			.5
			LastType Type
			.State
			.6
			:
. ATN			.(7) : 1-6
			.(9) : 2-6
			.7
(Rules)			.State Type
			.8
			:
)			.(9) : 1-8

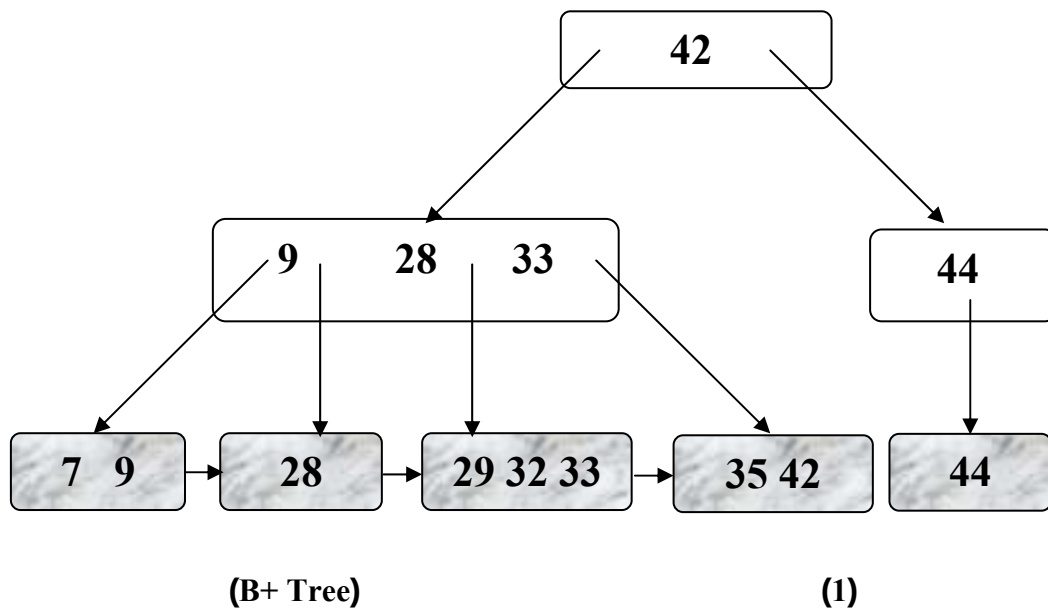


بيانات قواعد اللغة العربية	الحالة	نوع الكلمة أو المركب الحالي	الحالات المتوقعة التالية	رموز الحالة الحالية المتوقعة	قائمة Path	الأدوار الوظيفية المتوقعة للحالة الحالية	المفتاح	مثال
bnf([s_1],[[1]],1,["فعل"])	s	فعل لازم	s_1	[1]	[]	فعل	s,[1,0]	يذهب
bnf([s_2],[[2]],7,["منفذ"])	s_1	اسم فاعل	s_2	[2]	[[1]]	منفذ	s_1,[2,7]	طالب
bnf([s_2],[[3]],11,["مضاف اليه"])	s_2	اسم علم	s_2	[]	[[2],[1]]	مضاف اليه	s_2,[2,1]	العلم
bnf([s_2],[[4]],5,["صفة"])	s_2	صفة	s_2	[]	[[2],[1]]	صفة	s_2,[3,1]	المجتهد
bnf([s_2],[[5]],5,["صفة"])	s_2	صفة	s_2	[]	[[2],[1]]	صفة	s_2,[3,1]	النشيط
bnf([n_2],[[6,4]],3,["حرف جر"])	s_2	حرف جر	n_2	[6,4]	[[2],[1]]	حرف جر	s_2,[6,4]	إلى
bnf([s_2],[[18,2]],["شبه جملة"])	n_2	اسم مكان	s_2	[]	[[2],[1],[6,4]]	شبه مجرور	n_2,[2,5],[6,4]	المدرسة
bnf([n_2],[[6,4]],3,["حرف جر"])	s_2	حرف جر	n_2	[6,4]	[[2],[1]]	حرف جر	s_2,[6,4]	في
bnf([s_2],[[18,2]],["شبه جملة"])	n_2	اسم علم	s_2	[]	[[2],[1],[6,4]]	شبه مجرور	n_2,[2,1],[6,4]	الصباح
bnf([s_2],[[5]],5,["صفة"])	s_2	صفة	s_2	[]	[[2],[1]]	صفة	s_2,[3,1]	الباكر
bnf([s_2],[[18,2]],["شبه جملة"])	s_2	حرف جر + اسم آلة	s_2	[]	[[2],[1]]	شبه مجرور	s_2,[2,5],[6,4]	بالسيارة
bnf([final],[[]],[[]])	s_2	فعل، منفذ	final	[]	[[2],[1]]	-	s_2,[2],[1]	جملة مقبولة

جدول (1) يمثل خط سير المعرب مع الجملة العربية (يذهب طالب العلم المجتهد النشيط إلى المدرسة في الصباح الباكر بالسيارة)

.	.
)	)
.(	.(
.	B+.

(2)



Scatter plot showing the relationship between the year of the first publication (X-axis) and the number of citations (Y-axis). The X-axis ranges from 1987 to 2004, and the Y-axis ranges from 0 to 1.5. Data points are represented by black dots, and a linear regression line is shown as a solid black line. The regression line starts at approximately (1987, 0.1) and ends at approximately (2004, 0.5). The data points are scattered around this line, with some points showing higher citation counts than others for the same year.