

اطلاعات اعضای تیم			
نام و نام‌خوادی	شماره دانشجویی	نام استاد	رشته
امیر احسان ملانی	۹۹۱۲۵۵۱۲	دکتر پوره‌هاشم	مهندسی کامپیوتر – گرایش نرم‌افزار
الهه کشاورزی خوزانی	۹۹۱۲۵۴۳۱		

مقدمه

تحلیلگر لغوی تابعی است که در ورودی خود متن برنامه مورد کامپایل را به عنوان یک فایل متن باز گشوده و کرکتر به کرکتر می‌خواند و با رسیدن به یک کرکتر جدا کننده لغت را تشخیص و تفکیک مینماید. جدا کننده‌ها مثل blank و tab یا new line فقط به منظور جدا سازی لغات به کار می‌روند. دسته دیگر از جدا کننده‌ها مثل semicolon یا + ارزش لغوی دارند. تحلیلگر لغوی در خروجی خود اطلاعات مربوط به لغت را در یک رکورد یا ساختار به نام بسته لغت قرار می‌دهد. بسته لغت شامل اطلاعاتی در مورد لغت تشخیص داده شده توسط تحلیلگر لغوی مانند سطر یا ستون و نوع لغت است که آن را در اصطلاح token گویند.

فاز اول پروژه:

شناسایی توکن‌ها و عبارات منظم مربوطه

- زبان مبدا انتخابی: C++
- توکن‌های شناسایی شده و عبارات منظم مربوطه:

float:	
num	{0 - 9}
point	{.}
exponent	{e, E}
sign	{+, -}
Regular expression:	
(sign λ)(num* point num num* num num* point num*)(λ exponent num num*)	

integer:	
num	{0 - 9}
exponent	{e, E}
sign	{+, -}
Regular expression:	
(sign λ) num num* (λ exponent num num*)	

integer_hex	
num	{0 - 9} ∪ {a - f} ∪ {A - F}
hex	{X, x}
Regular expression:	
0 hex num*	

integer_binary	
num	{0 , 1}
bin	{B, b}
Regular expression:	
0 bin num*	

identifier	
num	{0 - 9}
ch	{ _ } ∪ {a-z} ∪ {A-Z}
Regular expression:	
ch (ch num)*	

string	
ch	{ ' , ! , ... , ~ } - { " }
Regular expression:	
" ch* "	

character	
ch	{ ' , ! , ... , ~ } - { ' }
Regular expression:	
' ch '	

pointer	
point	{ * }
num	{ 0 - 9 }
ch	ch = { _ } ∪ { a-z } ∪ { A-Z }
Regular expression:	
ch (ch num)* point	

keywords								
auto	double	int	struct	break	else	long	switch	case
enum	register	typedef	char	extern	return	union	const	float
short	short	unsigned	continue	for	signed	void	default	goto
sizeof	volatile	do	if	static	while	asm	bool	catch
class	const_cast	delete	dynamic_cast	explicit	export	false	friend	inline
mutable	namespace	new	operator	private	protected	public	reinterpret_cast	static_cast
template	this	throw	true	try	typeid	typename	using	virtual

operators								
+	-	*	/	%	=	==	>	>=
<	<=	!=	&&		!	&		^
<<	>>	~	+=	-=	*=	/=		