ULM C Compiler
12 November 2019

Michael C. Lehn

Contents

Chapter 1

Syntax

1.1 Structure of a C program

⟨translation-unit⟩ ⟶
⟶ ⟨external-declaration-list⟩
⟨external-declaration-list⟩ ⟶ ⟨external-declaration⟩
⟶ ⟨external-declaration-list⟩ ⟨external-declaration⟩
⟨external-declaration⟩ ⟶ ⟨function-definition⟩
⟶ ⟨declaration⟩
⟨declaration⟩ ⟶ ⟨declaration-specifiers⟩ ⟨init-declarator-list⟩ **;**
⟶ ⟨declaration-specifiers⟩ **;**
⟨function-definition⟩ ⟶ ⟨declaration-specifiers⟩ ⟨declarator⟩ ⟨compound-statement⟩
⟨declaration-specifiers⟩ ⟶ ⟨type-specifier⟩
⟶ ⟨storage-class-specifier⟩ ⟨type-specifier⟩

1.2 Storage classes and types

⟨storage-class-specifier⟩ ⟶ **static**
⟶ **extern**

⟨type-specifier⟩ ⟶ ⟨integer-type⟩
⟶ ⟨void-type⟩
⟶ ⟨struct-specifier⟩

⟨integer-type⟩ ⟶ **int8_t**
⟶ **int16_t**
⟶ **int32_t**
⟶ **int64_t**
⟶ **uint8_t**
⟶ **uint16_t**
⟶ **uint32_t**
⟶ **uint64_t**

In the lexical analysis **char** gets replaced with **int8_t** , and **int** with **int16_t** .

⟨void-type⟩ ⟶ **void**

⟨struct-specifier⟩ ⟶ **struct** ⟨identifier⟩
⟶ **struct** ⟨identifier⟩ **{** ⟨struct-declaration-list⟩ **}**
⟶ **struct {** ⟨struct-declaration-list⟩ **}**
⟨struct-declaration-list⟩ ⟶
⟶ ⟨struct-declaration-sequence⟩
⟨struct-declaration-sequence⟩ ⟶ ⟨struct-declaration⟩
⟶ ⟨struct-declaration-sequence⟩ ⟨struct-declaration⟩
⟨struct-declaration⟩ ⟶ ⟨type-specifier⟩ ⟨declarator-list⟩ **;**

⟨init-declarator-list⟩ ⟶ ⟨init-declarator⟩
⟶ ⟨init-declarator-list⟩ **;** ⟨init-declarator⟩
⟨init-declarator⟩ ⟶ ⟨declarator⟩
⟶ ⟨declarator⟩ **=** ⟨initializer⟩
⟨initializer⟩ ⟶ ⟨expression⟩
⟶ **{ }**
⟶ **{** ⟨initializer-list⟩ **}**
⟨initializer-list⟩ ⟶ ⟨initializer-list-items⟩
⟨initializer-list-items⟩ ⟶ ⟨initializer⟩
⟶ ⟨initializer-list-items⟩ **,** ⟨initializer⟩

⟨declarator⟩ ⟶ ⟨embedded-declarator⟩
⟨embedded-declarator⟩ ⟶ ⟨direct-declarator⟩
⟶ ⟨pointer-declarator⟩

⟨pointer-declarator⟩ ⟶ * ⟨embedded-declarator⟩

⟨direct-declartor⟩ ⟶ **(** ⟨embedded-declarator⟩ **)**
⟶ ⟨identidier⟩
⟶ ⟨direct-declarator⟩ **[** ⟨constant-expression⟩ **]**
⟶ ⟨direct-declarator⟩ **(** ⟨parameter-list⟩ **)**
⟶ ⟨direct-declarator⟩ **( )**

⟨type-name⟩ ⟶ ⟨type-specifier⟩
⟶ ⟨type-specifier⟩ ⟨abstract-declarator⟩
⟨abstract-declarator⟩ ⟶ ⟨embedded-abstract-declarator⟩
⟨embedded-abstract-declarator⟩ ⟶ ⟨abstract-direct-declarator⟩
⟶ ⟨abstract-pointer-declarator⟩
⟨abstract-pointer-declarator⟩ ⟶ * ⟨embedded-abstract-declarator⟩
⟶ *

⟨abstract-direct-declarator⟩ ⟶ **(** ⟨embedded-abstract-declarator⟩ **)**
⟶ **[** ⟨constant-expression⟩ **]**
⟶ ⟨abstract-direct-declarator⟩ **[** ⟨constant-expression⟩ **]**
⟶ **(** ⟨parameter-list⟩ **)**
⟶ ⟨abstract-direct-declarator⟩ **(** ⟨parameter-list⟩ **)**
⟶ ⟨abstract-direct-declarator⟩ **( )**
⟶ **( )**

⟨declarator-list⟩ ⟶ ⟨declarator⟩
⟶ ⟨declarator-list⟩ `,` ⟨declarator⟩

⟨parameter-list⟩ ⟶ ⟨parameter-declaration⟩
⟶ ⟨parameter-list⟩ `,` ⟨parameter-declaration⟩
⟨parameter-declaration⟩ ⟶ ⟨type-specifier⟩ ⟨declarator⟩

⟨constant-expression⟩ ⟶ ⟨logical-or-expression⟩

⟨condition⟩ ⟶ ⟨assignment-expression⟩
⟨expression⟩ ⟶ ⟨assignment-expression⟩

⟨assignment-expression⟩ ⟶ ⟨logical-or-expression⟩
⟶ ⟨unary-expression⟩ `=` ⟨assignment-expression⟩
⟶ ⟨unary-expression⟩ `+=` ⟨assignment-expression⟩
⟶ ⟨unary-expression⟩ `-=` ⟨assignment-expression⟩

⟨logical-or-expression⟩ ⟶ ⟨logical-and-expression⟩
⟶ ⟨logical-or-expression⟩ **| |** ⟨logical-and-expression⟩
⟨logical-and-expression⟩ ⟶ ⟨equality-expression⟩
⟶ ⟨logical-and-expression⟩ **&&** ⟨equality-expression⟩
⟨equality-expression⟩ ⟶ ⟨relational-expression⟩
⟶ ⟨equality-expression⟩ **==** ⟨relational-expression⟩
⟶ ⟨equality-expression⟩ **!=** ⟨relational-expression⟩
⟨relational-expression⟩ ⟶ ⟨additive-expression⟩
⟶ ⟨relational-expression⟩ **<** ⟨additive-expression⟩
⟶ ⟨relational-expression⟩ **<=** ⟨additive-expression⟩
⟶ ⟨relational-expression⟩ **>=** ⟨additive-expression⟩
⟶ ⟨relational-expression⟩ **>** ⟨additive-expression⟩
⟨additive-expression⟩ ⟶ ⟨multiplicative-expression⟩
⟶ ⟨additive-expression⟩ **+** ⟨multiplicative-expression⟩
⟶ ⟨additive-expression⟩ **-** ⟨multiplicative-expression⟩
⟨multiplicative-expression⟩ ⟶ ⟨unary-expression⟩
⟶ ⟨multiplicative-expression⟩ ***** ⟨unary-expression⟩
⟶ ⟨multiplicative-expression⟩ **/** ⟨unary-expression⟩
⟶ ⟨multiplicative-expression⟩ **%** ⟨unary-expression⟩

⟨unary-expression⟩ ⟶ ⟨postfix-expression⟩
⟶ ⟨address-of⟩
⟶ ⟨pointer-dereference⟩
⟶ ⟨unary-plus⟩
⟶ ⟨unary-minus⟩
⟶ ⟨prefix-plusplus⟩
⟶ ⟨prefix-minusminus⟩
⟶ **!** ⟨unary-expression⟩
⟶ **sizeof** ⟨unary-expression⟩
⟶ **sizeof (** ⟨type-name⟩ **)**
⟨address-of⟩ ⟶ **&** ⟨postfix-expression⟩
⟨pointer-dereference⟩ ⟶ ***** ⟨unary-expression⟩
⟨unary-plus⟩ ⟶ **+** ⟨unary-expression⟩
⟨unary-minus⟩ ⟶ **-** ⟨unary-expression⟩
⟨prefix-plusplus⟩ ⟶ **++** ⟨unary-expression⟩
⟨prefix-minusminus⟩ ⟶ **--** ⟨unary-expression⟩

⟨postfix-expression⟩ ⟶ ⟨primary-expression⟩
⟶ ⟨function-call⟩
⟶ ⟨postfix-expression⟩ **[** ⟨expression⟩ **]**
⟶ ⟨postfix-expression⟩ **.** ⟨identifier⟩
⟶ ⟨postfix-expression⟩ **->** ⟨identifier⟩
⟨function-call⟩ ⟶ ⟨postfix-expression⟩ **( )**
⟶ ⟨postfix-expression⟩ **(** ⟨argument-expression-list⟩ **)**
⟨argument-expression-list⟩ ⟶ ⟨expression⟩
⟶ ⟨argument-expression-list⟩ **,** ⟨expression⟩

⟨primary-expression⟩ ⟶ ⟨identifier⟩
⟶ ⟨constant⟩
⟶ ⟨string-literal⟩
⟶ **(** ⟨assignment-expression⟩ **)**

⟨compound-statement⟩ ⟶ **{ }**
⟶ **{** ⟨block-item-list⟩ **}**
⟨block-item-list⟩ ⟶ ⟨block-item⟩
⟶ ⟨block-item-list⟩ ⟨block-item⟩
⟨block-item⟩ ⟶ ⟨local-declaration⟩
⟶ ⟨statement⟩
⟨local-declaration⟩ ⟶ ⟨declaration⟩

⟨statement⟩ ⟶ ⟨compound-statement⟩
⟶ ⟨expression-statement⟩
⟶ ⟨if-statement⟩
⟶ ⟨while-statement⟩
⟶ ⟨for-statement⟩
⟶ ⟨return-statement⟩
⟶ ⟨break-statement⟩
⟶ ⟨continue-statement⟩

⟨expression-statement⟩ ⟶ ⟨expression⟩ **;**

⟨if-statement⟩ ⟶ **if (** ⟨condition⟩ **)** ⟨compound-statement⟩
⟶ **if (** ⟨condition⟩ **)** ⟨compound-statement⟩ **else** ⟨compound-statement⟩
⟶ **if (** ⟨condition⟩ **)** ⟨compound-statement⟩ ⟨else-if-list⟩ **else** ⟨compound-statement⟩
⟨else-if-list⟩ ⟶ ⟨else-if⟩
⟶ ⟨else-if-list⟩ ⟨else-if⟩
⟨else-if⟩ ⟶ **else if (** ⟨condition⟩ **)** ⟨compound-statement⟩

⟨while-statement⟩ ⟶ **while (** ⟨condition⟩ **)** ⟨compound-statement⟩

⟨for-statement⟩ ⟶ **for (** ⟨for-expressions⟩ **)** ⟨compound-statement⟩
⟨for-expressions⟩ ⟶ ⟨initial-clause⟩ ⟨for-condition⟩ **;** ⟨for-increment⟩
⟨initial-clause⟩ ⟶ ⟨for-declaration⟩
⟶ ⟨expression⟩ **;**
⟶ **;**
⟨for-declaration⟩ ⟶ ⟨declaration⟩
⟨for-increment⟩ ⟶ ⟨expression⟩
⟶
⟨for-condition⟩ ⟶ ⟨condition⟩
⟶

⟨return-statement⟩ ⟶ **return ;**

⟨return-statement⟩ ⟶ **return ;**
⟶ **return** ⟨expression⟩ **;**

⟨break-statement⟩ ⟶ **break ;**

⟨continue-statement⟩ ⟶ **continue ;**

⟨constant⟩ ⟶ ⟨integer-constant⟩
⟶ ⟨character-constant⟩
⟶ ⟨bool-constant⟩

⟨integer-constant⟩ ⟶ **decimal-constant**
⟶ **octal-constant**
⟶ **hexadecimal-constant**

⟨character-constant⟩ ⟶ **char-constant**

⟨bool-constant⟩ ⟶ **true**
false

⟨string-literal⟩ ⟶ ⟨string-literal-items⟩
⟨string-literal-items⟩ ⟶ **string-literal**
⟶ ⟨string-literal-items⟩ **string-literal**