Ambigious vs. Unambigious Grammars

Definitions. In some context-free grammars, the same word can be generated in two or more different ways. Such grammars are called *ambiguous.* If every word has at most one derivation, the grammar is called *unambiguous.*

Why this is important. How can we describe arithmetic expressions with numbers, such as 5, or $2 + 3 \cdot 4$? At first glance, this is straightforward:

- every digit is an arithmetic expression, so we have 10 rules

$$E \to 0, \quad E \to 1, \ldots, E \to 9$$

- if we have two arithmetic expressions and we combine them with the plus sign, we still get an arithmetic expression; this corresponds to the rule

$$E \to E + E;$$

- if we have two arithmetic expressions and we combine them with the multiplication sign, we still get an arithmetic expression; this corresponds to the rule

$$E \to E \cdot E;$$

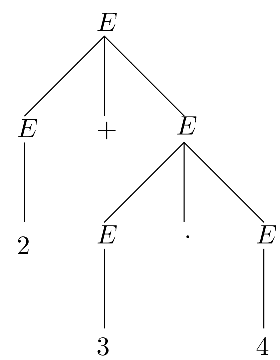
- if E is an arithmetic expression, then (E) is also an arithmetic expression; this corresponds to the rule

$$E \to (E).$$

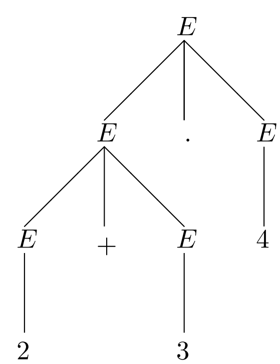
How can we derive the expression $2 + 3 \cdot 4$ from these rules? Our understanding of this expression is that it is the sum of two expressions: the number 2 and the product $3 \cdot 4$, so we should use the rule $E \to E + E$ and then derive 2 from the first E and the expression $3 \cdot 4$ from the second E.

Deriving 2 from E is easy: we have a rule $E \to 2$.

The expression $3 \cdot 4$ can be represented as a product of 3 and 4, which are also expressions. So, we have the following natural derivation of the expression $2 + 3 \cdot 4$:



The problem is that with this grammar, we can derive this same expression differently: view it as the product of $2+3$ and number 4:

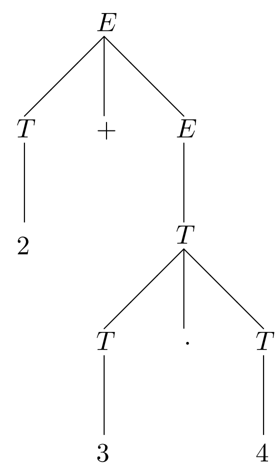


Since one of our main purposes is compiling, this is *not* what we want: this derivation will lead to computing a different value $(2+3)\cdot 4$. We would like to have only one derivation – i.e., we want our grammar to be unambiguous.

How to correct this situation: example. In this example, instead of just considering expressions (E), we can also introduce the notion of a *term* T – which is either a number, or a product, or an expression in parentheses. In this case, we have the following rules:

$$E \to T; \quad E \to T + E; \quad T \to (E); \quad T \to T \cdot T; T \to 0; \ldots T \to 9.$$

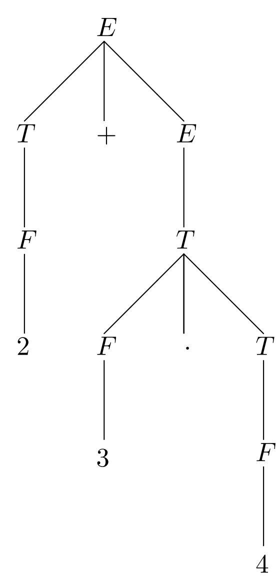
In this grammar, the expression $2+3\cdot 4$ has only one derivation:



This grammar, by the way, is still ambiguous: e.g., the expression $1 \cdot 2 \cdot 3$ can be interpreted as $(1 \cdot 2) \cdot 3$ (as Java does), or as $1 \cdot (2 \cdot 3)$. So, to get a truly unambiguous grammar, we need more arrangement. For example, we can introduce an additional concept of a factor F and get the following rules:

$$E \to T; \quad E \to T + E; \quad T \to F; \quad F \to (E); \quad T \to F \cdot T; F \to 0; \ldots F \to 9.$$

Then, the expression $2 + 3 \cdot 4$ can be derived as follows:



How can we derive the word $2 + 3 \cdot 4$ in this grammar: a detailed explanation. We start with the starting variable E. To this variable, we can apply one of the two rules: $E \to T$ and $E \to T + E$.

If we use the rule $E \to T$, then next, we should have a derivation $T \to 2{+}3{\cdot}4$. However, T indicates a term i.e., a number, or a product, or an expression in parentheses, and $2+3\cdot4$ is not a number, not a product, and not an expression in parentheses.

Thus, the only rule that we can apply is the rule

$$\underline{E} \to T + E.$$

In the expression $2+3\cdot4$, there is only one plus sign. So, the only way to match this expression with the expression $T+E$ is to associate T with what is before the plus sign, and E with what is after the plus sign. In other words, we need to have the derivations $T \to 2$ and $E \to 3\cdot4$.

Let us see how we can derive 2 from T. There are two rules that replace the variable T: the rule $T \to F$ and the rule $T \to F\cdot T$. If we apply the second rule, we will get an expression with a multiplication sign, and the desired expression 2 does not have a multiplication sign. Thus, the only rule we can apply is the rule $T \to F$. By applying this rule, we get the following derivation:

$$\underline{E} \to \underline{T} + E \to F + E.$$

Now, we need to derive 2 from F. The derivation $F \to 2$ is actually one of the rules. So, we have the derivation

$$\underline{E} \to \underline{T} + E \to \underline{F} + E \to 2 + E.$$

Let us see how we can derive $3\cdot4$ from the variable E. To this variable, we can apply one of the two rules: $E \to T$ and $E \to T + E$. If we apply the second rule, we get an expression containing the plus sign, and in the expression $3\cdot4$ that we want to derive there is no plus sign. Thus, we cannot apply the second rule, and the only rule we can apply is the rule $E \to T$. By adding the application of this rule, we have the following derivation:

$$\underline{E} \to \underline{T} + E \to \underline{F} + E \to 2 + \underline{E} \to 2 + T.$$

We already derived the first part 2+ of the desired expression. To derive the whole expression $2+3\cdot4$, we need to derive the remaining part of the expression $3\cdot4$ from the variable T.

There are two rules that replace the variable T: the rule $T \to F$ and the rule $T \to F\cdot T$. If we apply the first rule, then we will need to derive the expression $3\cdot4$ from the variable F. To the variable F, we can apply either the rule $F \to (E)$, or one of the rules $F \to 0, \ldots, F \to 9$. In the first case, we get an expression with parentheses, and $3\cdot4$ does not have parentheses. In the second case, we get a digit, and $3\cdot4$ is not a digit. In both cases, we do not get the expression $3\cdot4$.

Thus, to get this expression, we cannot apply the rule $T \to F$. So, the only rule that we can apply is the rule $T \to F\cdot T$. By adding the application of this rule, we have the following derivation:

$$\underline{E} \to \underline{T} + E \to \underline{F} + E \to 2 + \underline{E} \to 2 + \underline{T} \to 2 + F\cdot T.$$

Our expression has only one multiplication symbol, so the only way to match our expression $3 \cdot 4$ with $F \cdot T$ is to match 3 with F and 4 with T. The transition from F to 3 is straightforward, we have a rule $F \to 3$:

$$\underline{E} \to \underline{T} + E \to 2 + \underline{E} \to 2 + \underline{T} \to \underline{F} + E \to 2 + \underline{F} \cdot T \to 2 + 3 \cdot T.$$

What is left is to derive 4 from T. There are two rules that replace the variable T: the rule $T \to F$ and the rule $T \to F \cdot T$. If we apply the second rule, we will get an expression with a multiplication sign, and 4 does not have any sign. So, our only choice is to apply the first rule $T \to F$. By adding the application of this rule, we have the following derivation:

$$\underline{E} \to \underline{T} + E \to \underline{F} + E \to 2 + \underline{E} \to 2 + \underline{T} \to 2 + \underline{F} \cdot T \to 2 + 3 \cdot \underline{T} \to 2 + 3 \cdot F.$$

Finally, to match, we need to derive 4 from the variable F. This can be done by applying the rule $F \to 4$, so the final derivation takes the following form:

$$\underline{E} \to \underline{T}{+}E \to \underline{F}{+}E \to 2{+}\underline{E} \to 2{+}\underline{T} \to 2{+}\underline{F}{\cdot}T \to 2{+}3{\cdot}\underline{T} \to 2{+}3{\cdot}\underline{F} \to 2{+}3{\cdot}4.$$

This is exactly the derivation that we described in the tree form.