

## 10. Идея синтаксического анализа. Два способа анализа. Использование МП-автомата для нисходящего анализа.

---

**Синтаксический анализ** (или разбор) — это процесс определения того, принадлежит ли входная цепочка символов  $w$  заданному формальному языку  $L(G)$ .

Если цепочка принадлежит языку, анализатор строит её структурное представление (например, дерево вывода), а если нет — выдает сообщение об ошибке.

В компиляторах этот этап обычно следует за лексическим анализом, работая с последовательностью токенов.

### 1. Два способа синтаксического анализа

В зависимости от того, как строится дерево вывода, выделяют два основных подхода:

- **Нисходящий анализ (сверху вниз, top-down):** Дерево строится от корня к листьям. Процесс начинается с аксиомы грамматики, которая последовательно разворачивается в терминальную цепочку, соответствующую входным данным.
- **Восходящий анализ (снизу вверх, bottom-up):** Дерево строится от листьев (входных символов) к корню. Входная цепочка «сворачивается» к аксиоме путем поиска основ и их замены на соответствующие нетерминалы.

### 2. Использование МП-автомата для нисходящего анализа

**Автомат с магазинной памятью (МП-автомат)** является естественной моделью для реализации нисходящего анализа, так как он позволяет моделировать **левосторонний вывод** цепочки.

**Механизм работы:**

1. **Инициализация:** В стек помещается начальный символ грамматики (аксиома).
2. **Моделирование вывода:**
  - Если на вершине стека находится **нетерминал**, автомат выбирает подходящее правило вывода и заменяет этот нетерминал правой частью правила (записывая её в стек в обратном или прямом порядке в зависимости от реализации).
  - Если на вершине стека находится **терминал**, он сравнивается с текущим символом входной цепочки. При совпадении терминал удаляется из стека, а указатель входа сдвигается на следующий символ.
3. **Завершение:** Цепочка считается распознанной, если после прочтения всей цепочки стек оказался пуст.

Для **LL(1)-грамматик** такой выбор правила всегда **детерминирован** и делается на основе одного текущего символа входа и нетерминала на вершине стека.