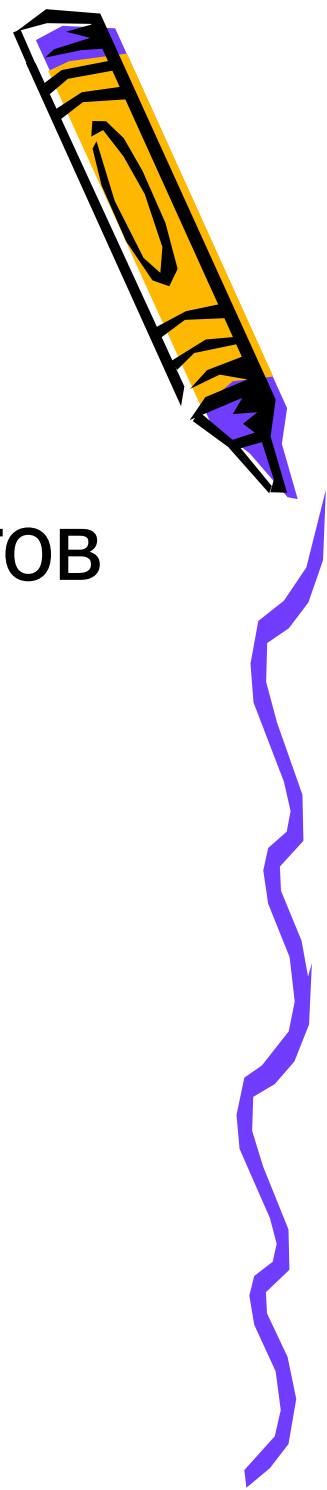


**АЛГОРИТМ ПРЕОБРАЗОВАНИЙ  
XML ДОКУМЕНТА С  
СОХРАНЕНИЕМ ДОПУСТИМЫХ  
ИЕРАРХИЧЕСКИХ ОТНОШЕНИЙ**



Сысоев Т. М.  
МСП РАН



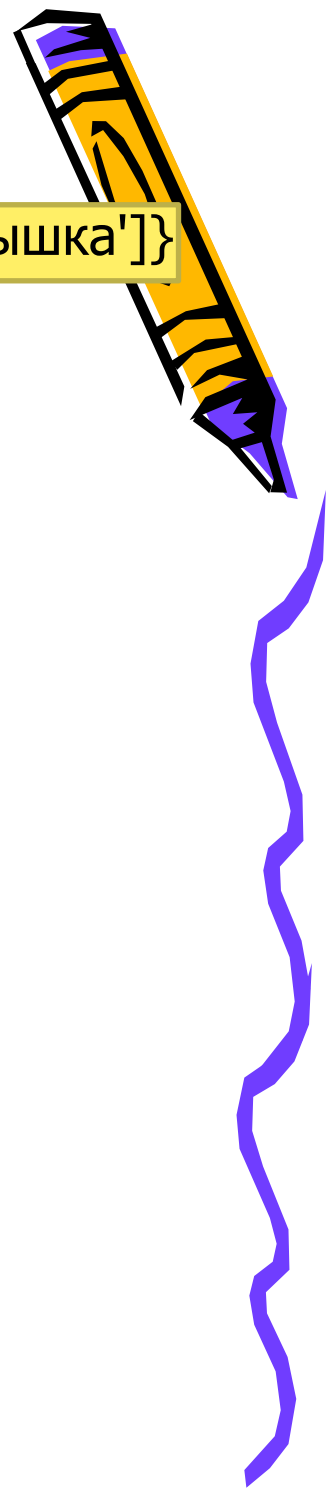
## Задача построения отчетов:

- генерация произвольных документов с заданным форматированием на основе шаблонов и информации, содержащейся в базе данных
- простое изменение шаблонов



# Язык шаблонов

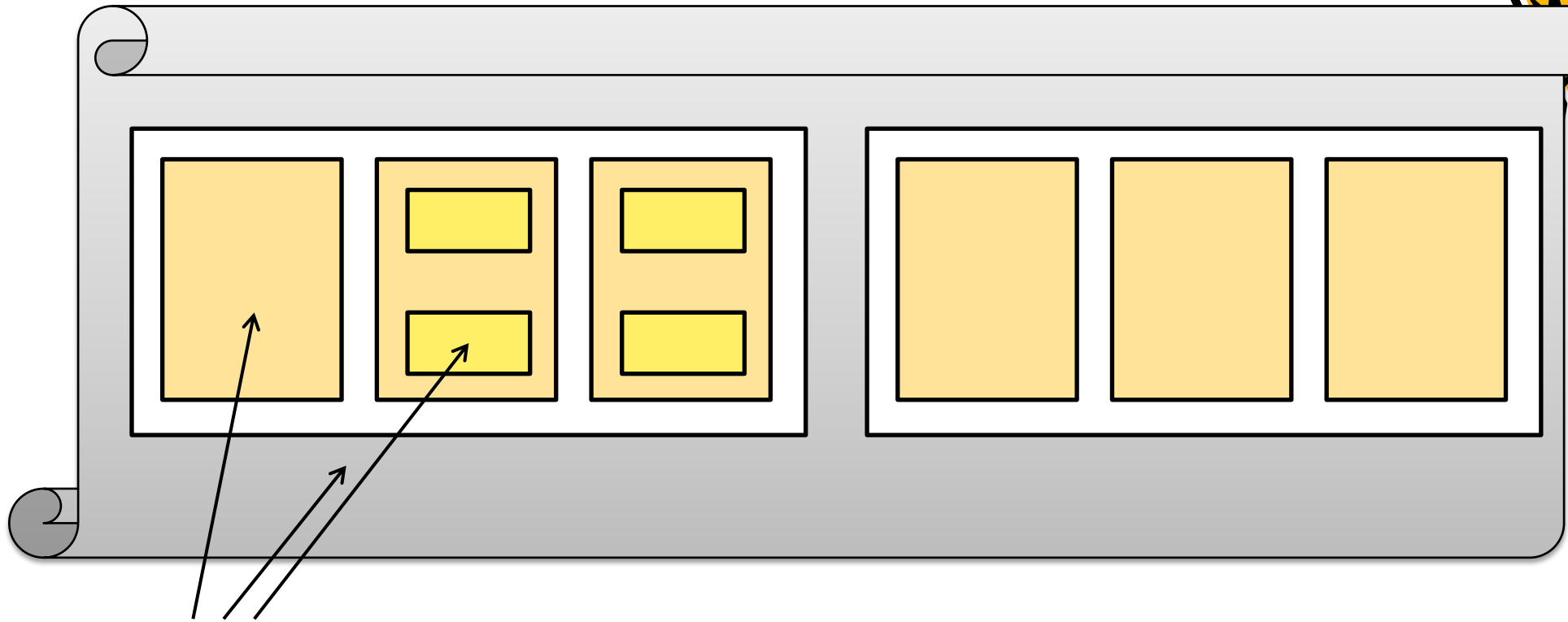
```
{set participants ['репка', 'дедка', 'бабка', 'внучка', 'жучка', 'кошка', 'мышка']}  
{for $i in 0..length($participants)}  
  {if $i > 0}  
    {for $j 1..$i}  
      ${participants[$j]} за ${genitive(participants[$j-1])}  
    {end}  
  {end}  
{if $i = length($participants) - 1}  
  - вытащили репку  
{end}  
{if $i < length($participants) - 1}  
  - не вытащили репку  
{end}  
{end}
```



# Варианты взаимодействия с офисными пакетами



# Обработка больших текстов



области действия  
переменных

цикл обработки:

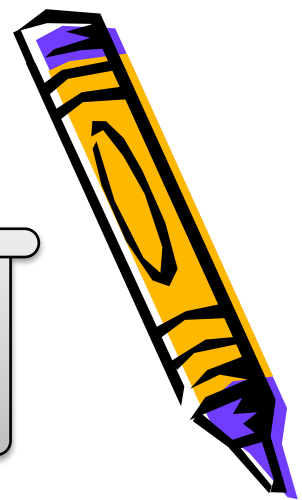
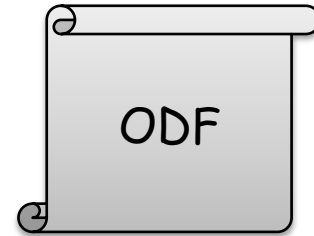
- поиск первой директивы
- выполнение директивы

сложность очередного шага  $\sim$  количеству областей  
большинство шагов порождают новые области



# Гибридный подход

выполнение основных операций



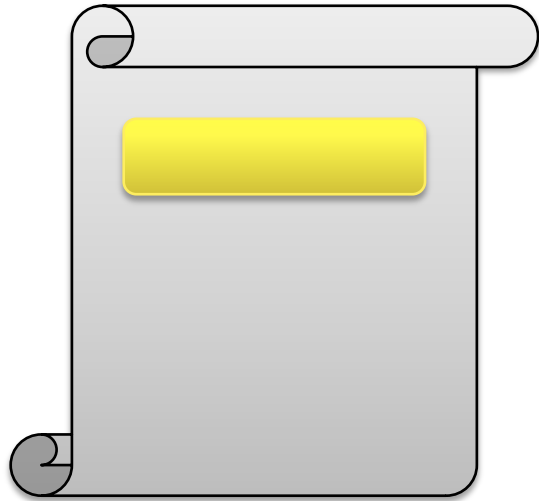
Программа генерации отчетов

операция установки оптимальной ширины колонки таблицы

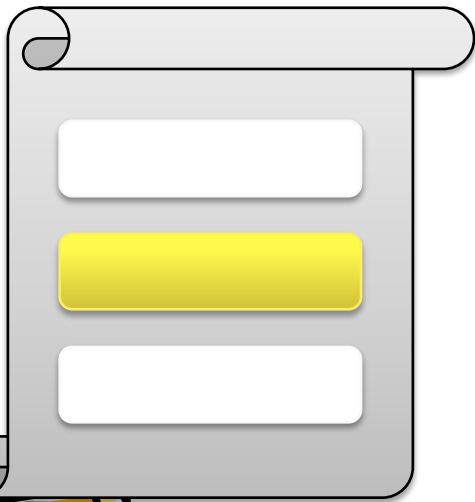
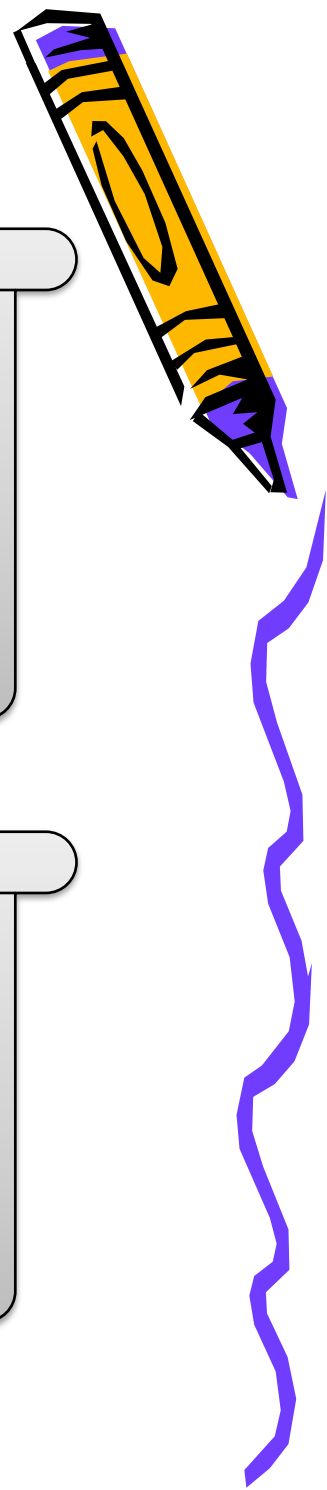
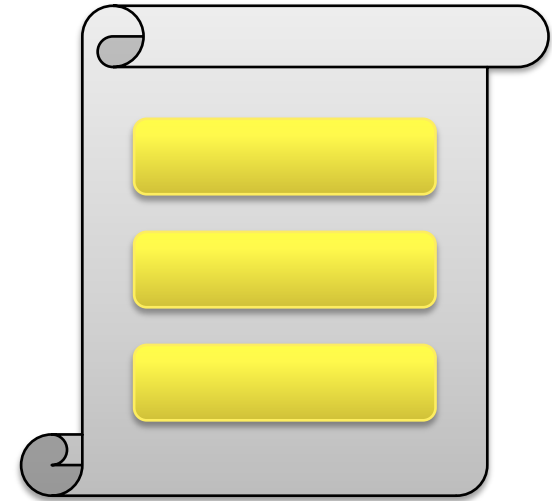
OpenOffice.org



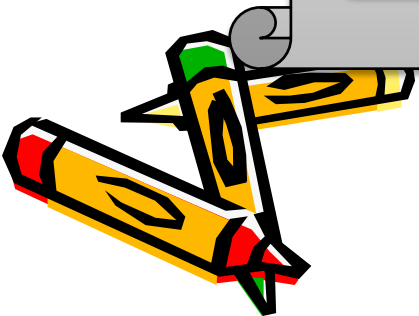
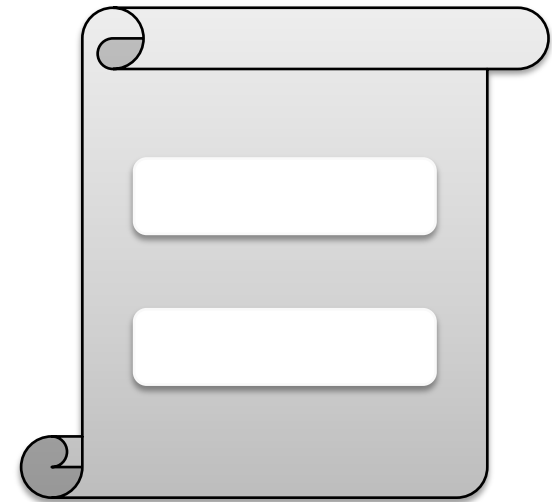
# Основные операции



копирование



удаление



# Проблемы выполнения операций

## Колобок

### Ингредиенты:

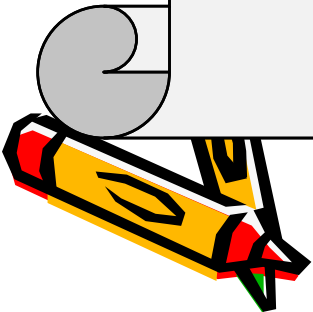
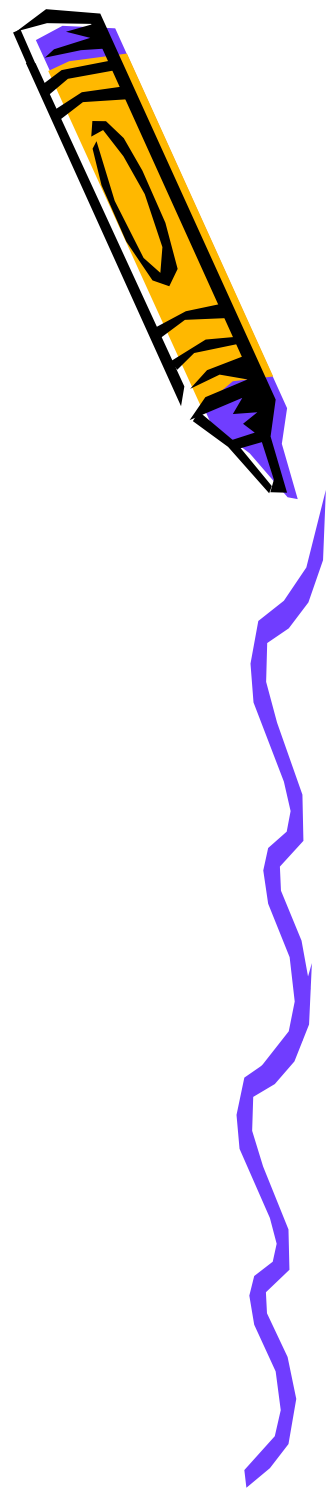
1. мука
2. сметана
3. сахар

### Приготовление:

- замесить тесто
- испечь в масле
- поставить на окошко

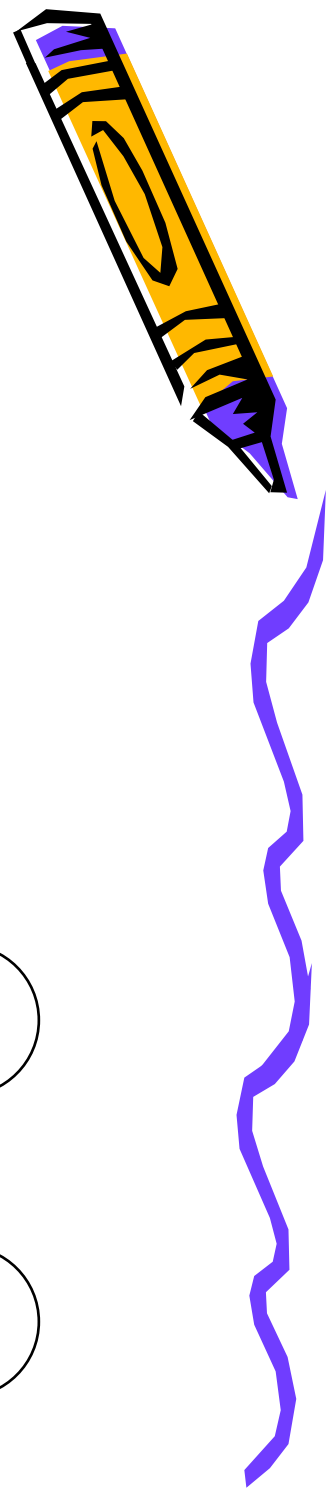


удаление

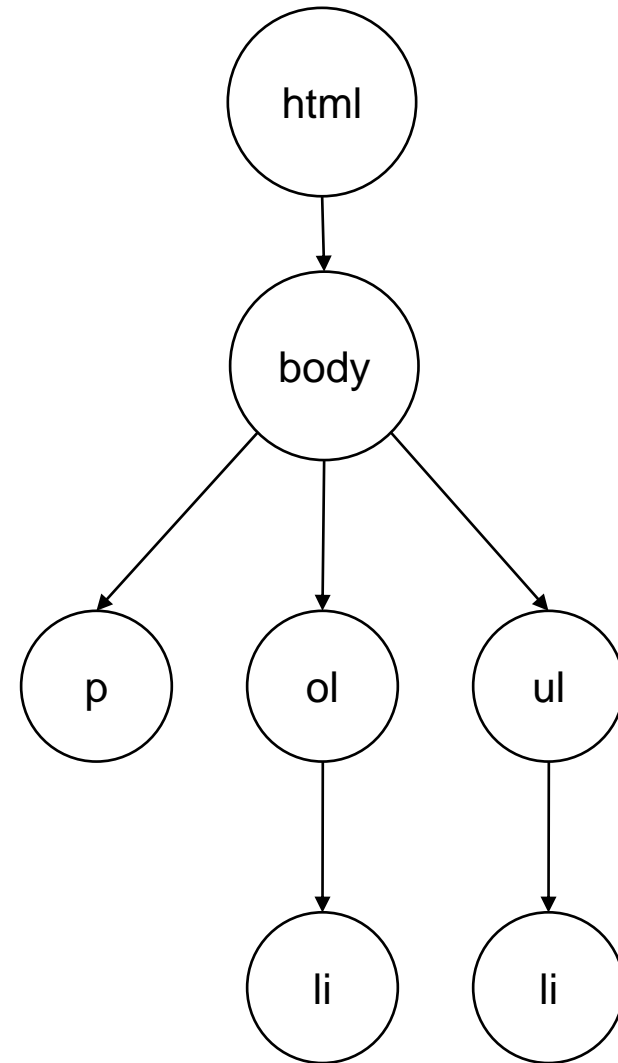


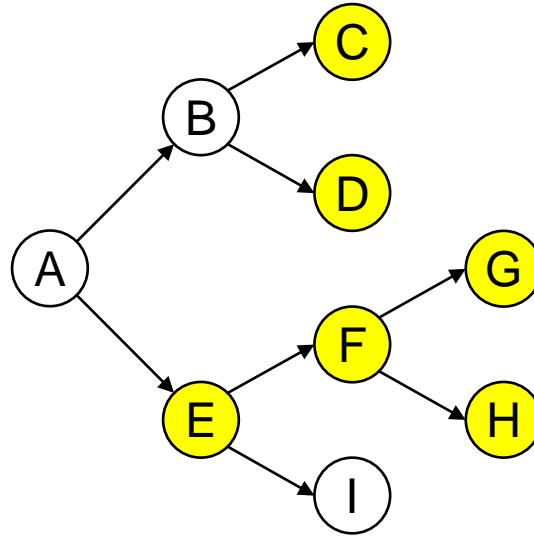
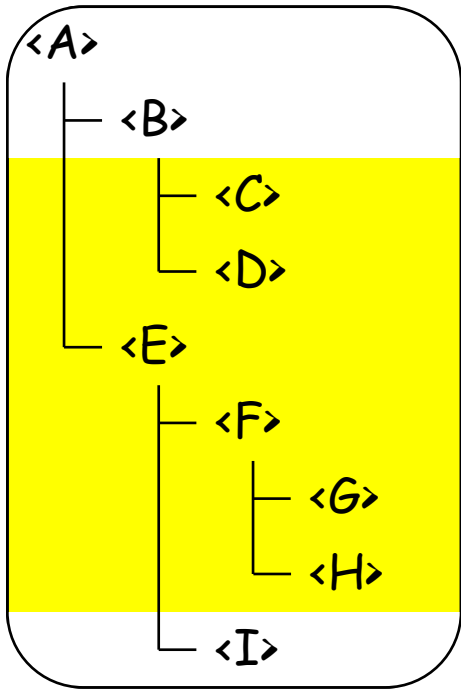


# Допустимая иерархия

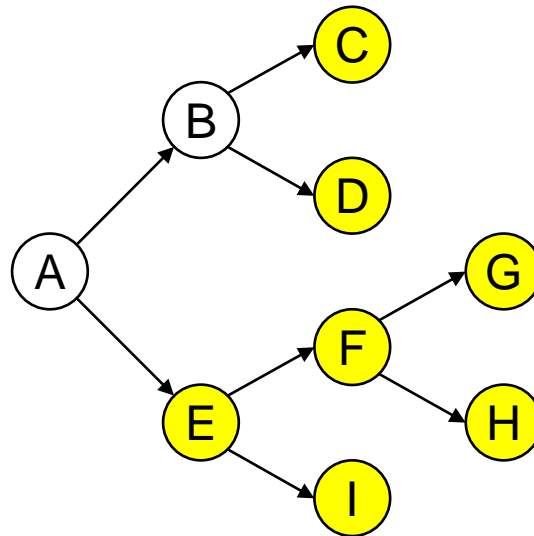
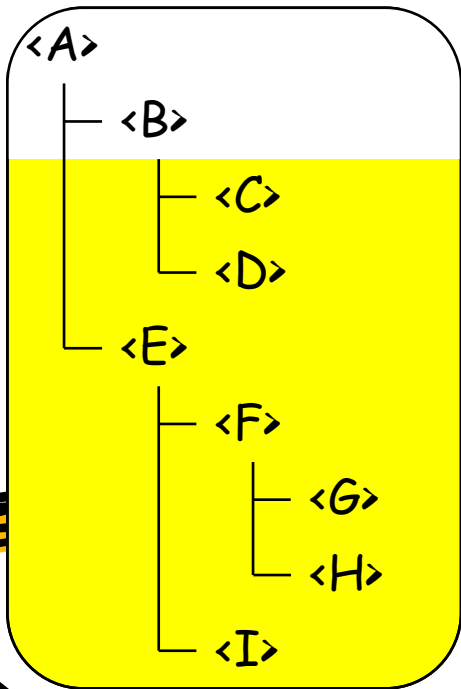
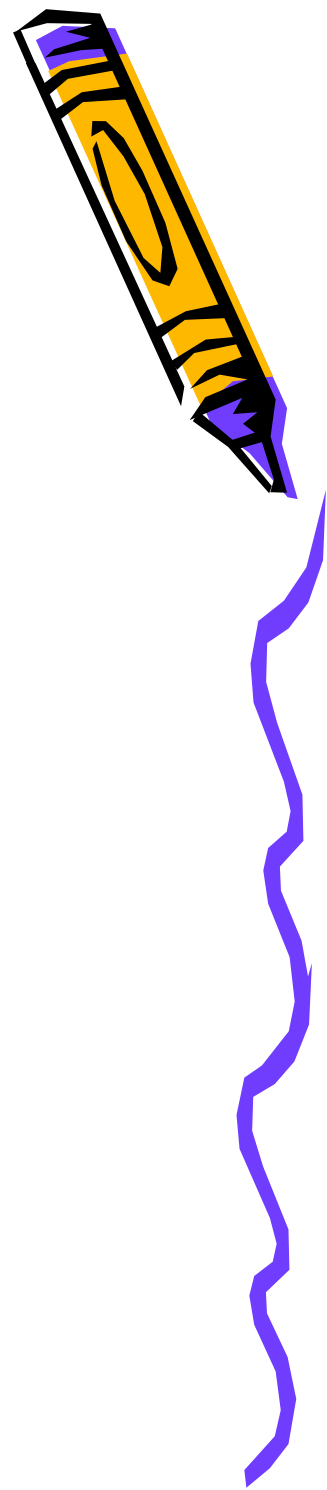


```
<html>
<body>
  <h1>Колобок</h1>
  <p>Ингредиенты</p>
  <ol>
    <li>мука</li>
    <li>сметана</li>
    <li>сахар</li>
  </ol>
  <p>Приготовление</p>
  <ul>
    <li>замесить тесто</li>
    <li>испечь в масле</li>
    <li>поставить на окошко</li>
  </ul>
</body>
</html>
```



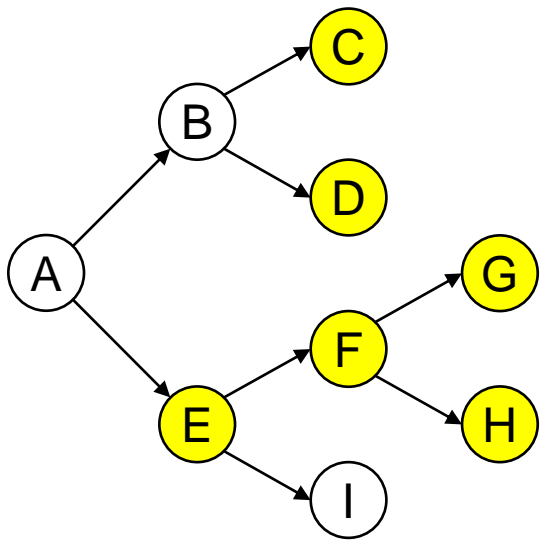


$(B, H]$

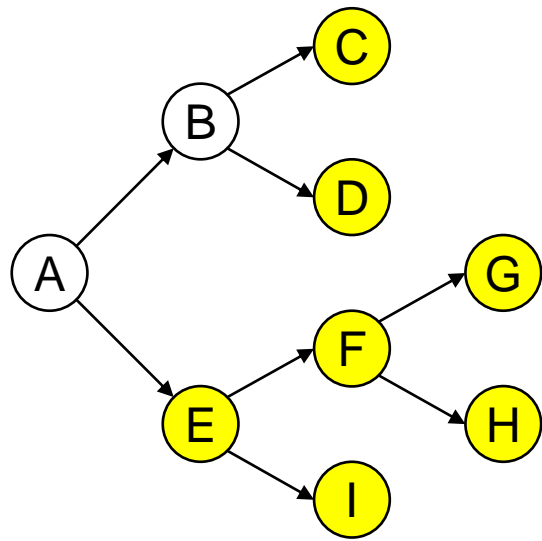


$\{A, B\}$

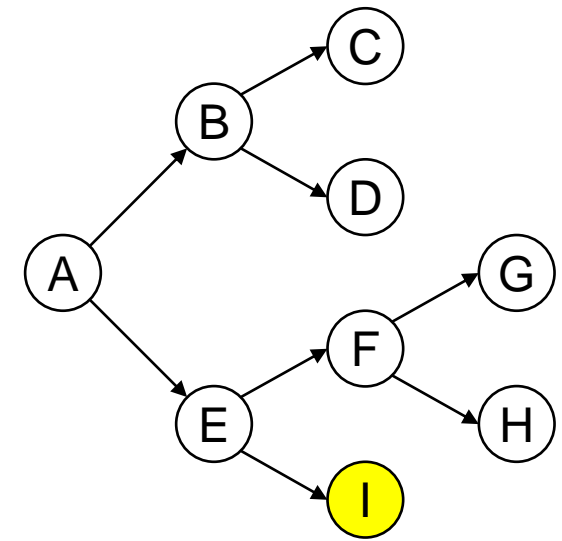




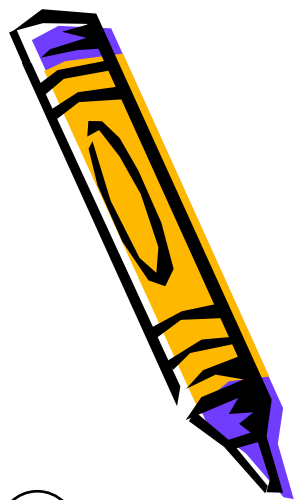
=



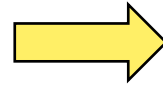
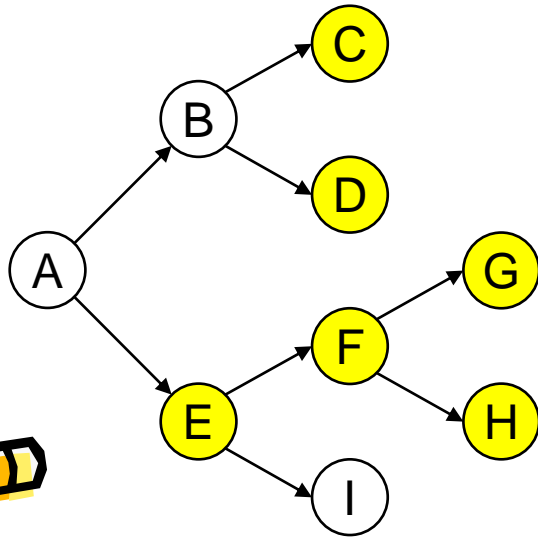
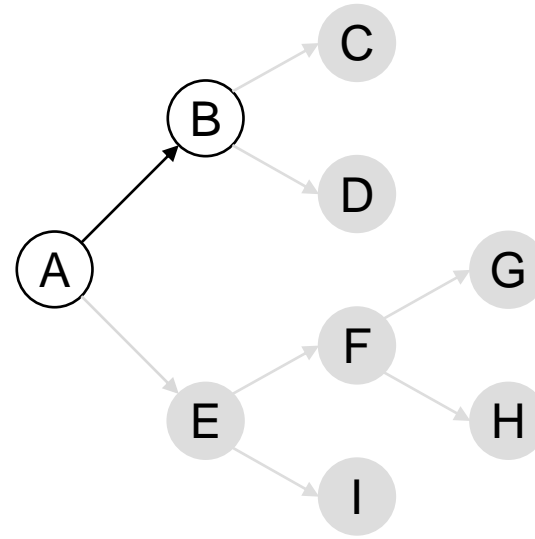
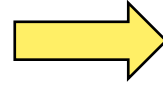
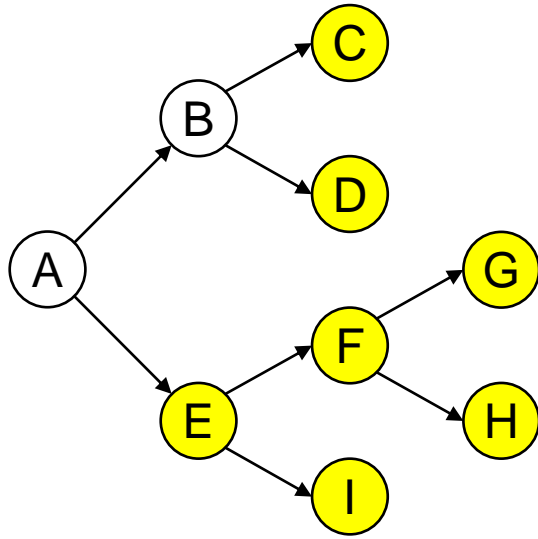
-



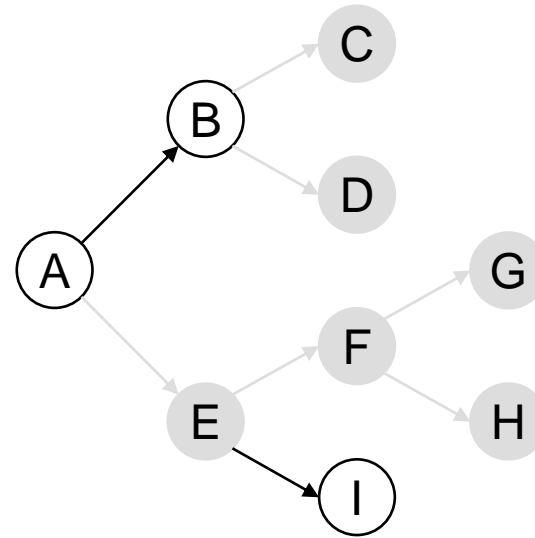
$$(B, H] = \{A, B\} - \{A, H\}$$



# Удаление фрагмента

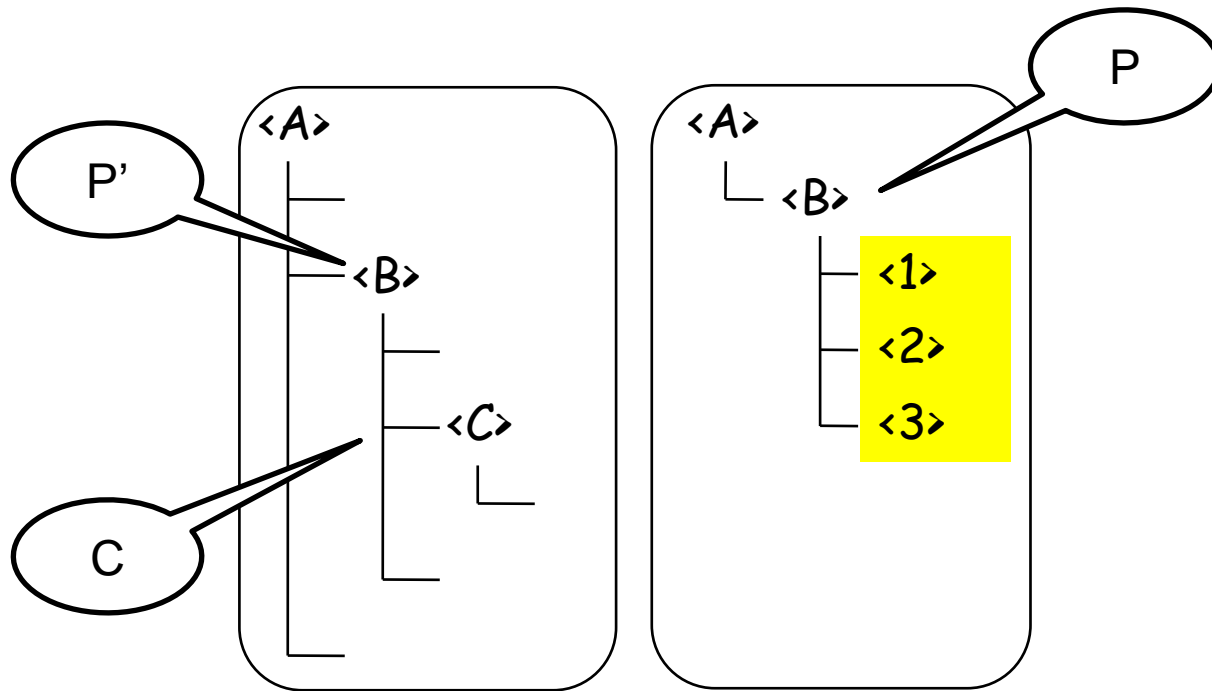


?



# Операция поиска элемента для копирования

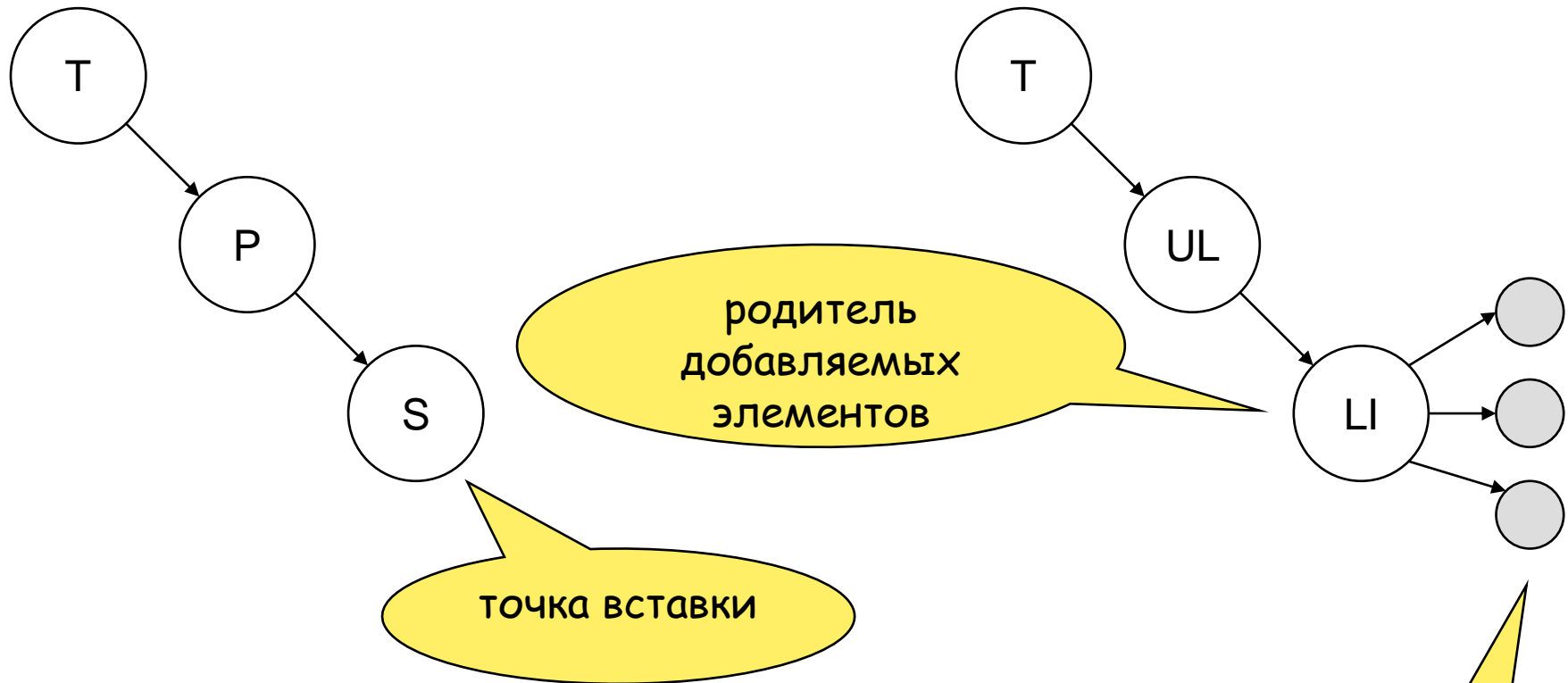
Для заданной вершины документа  $P$  найти соответствующую вершину  $P'$  в области точки для вставки  $C$  такую, что добавление потомков  $P$  к вершине  $P'$  не приведёт к нарушению допустимой иерархии



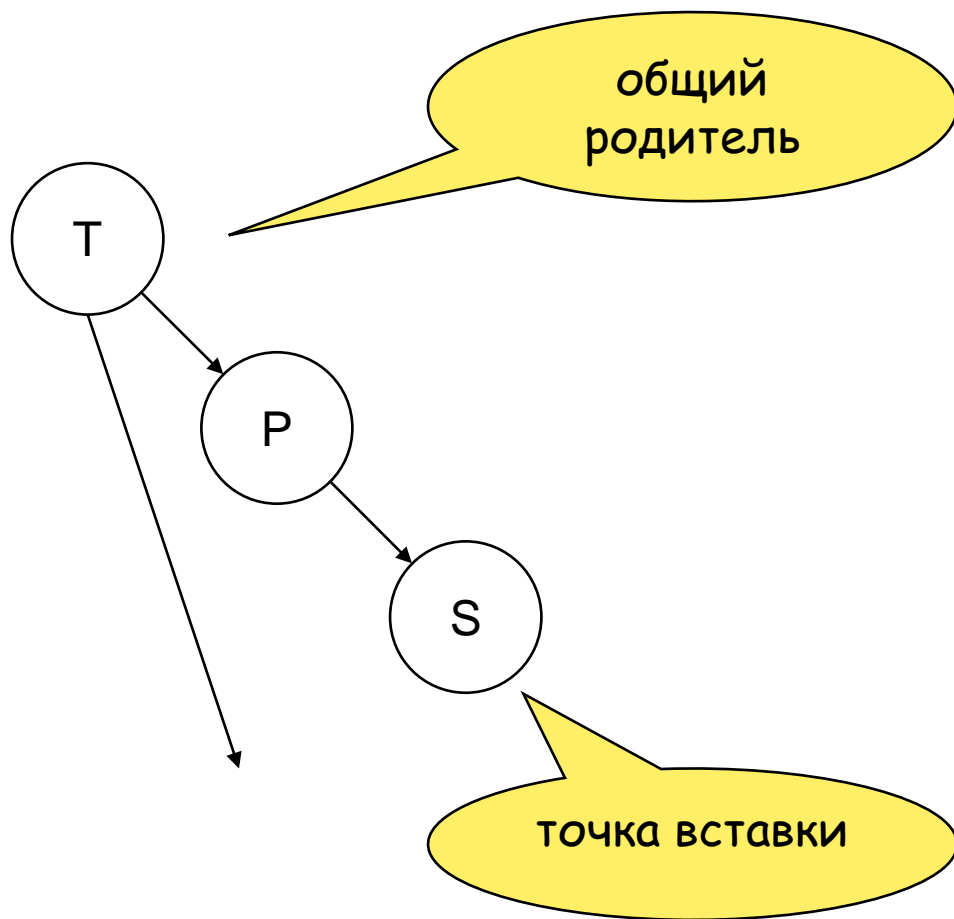
Если  $P'$  - родитель  $C$ , то можно найти элемент среди дочерних для  $P'$ , после которого следует начинать копирование



# Операция поиска элемента для копирования

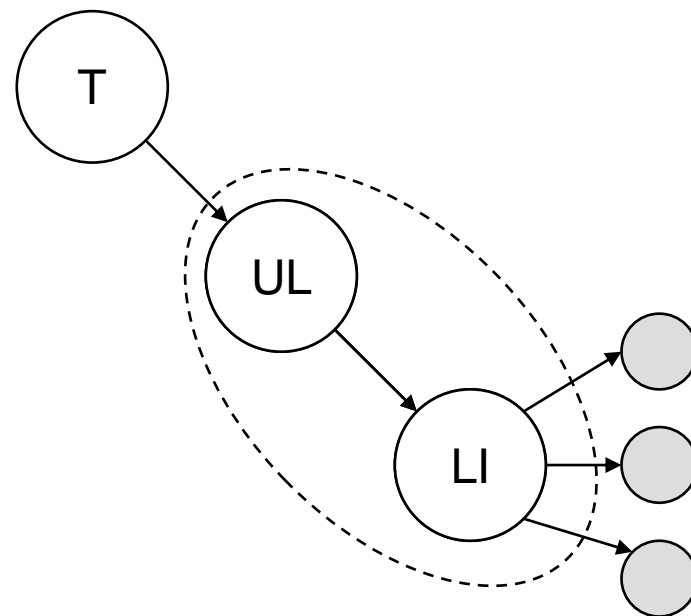


# Операция поиска элемента для копирования



точка вставки

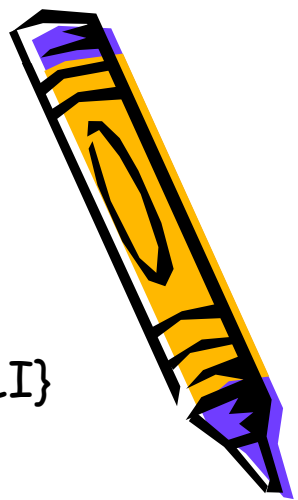
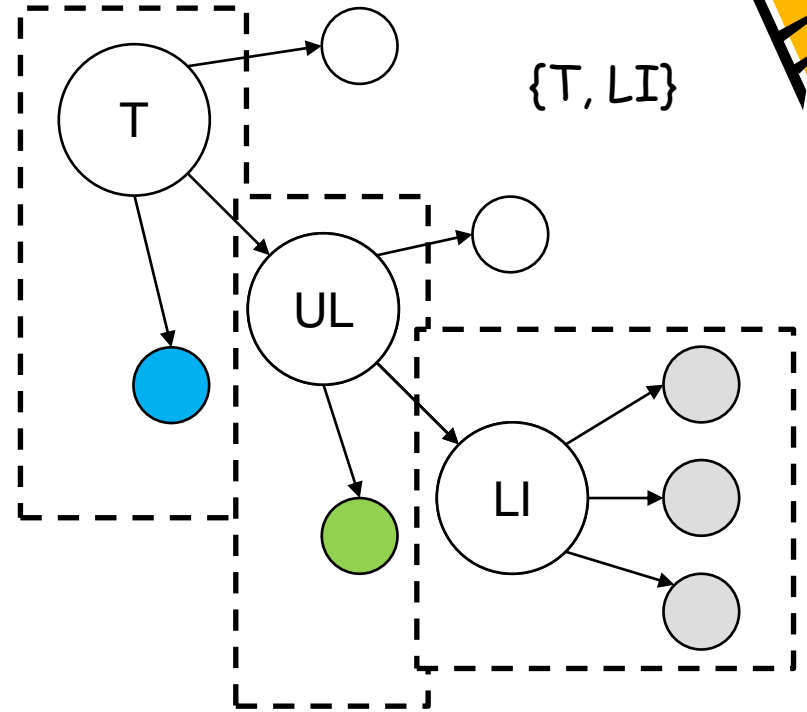
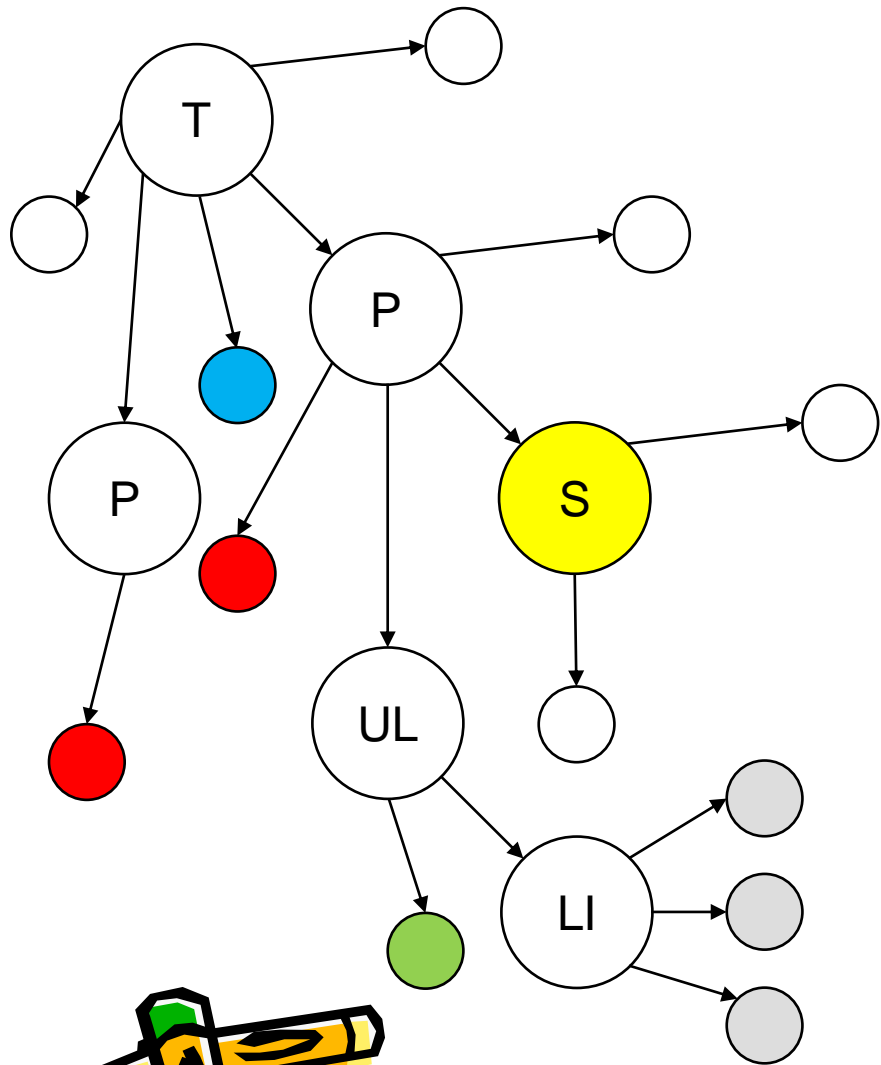
новая точка вставки



путь от общего родителя до добавляемых элементов



# Операция копирования фрагмента





# Операции копирования и удаления

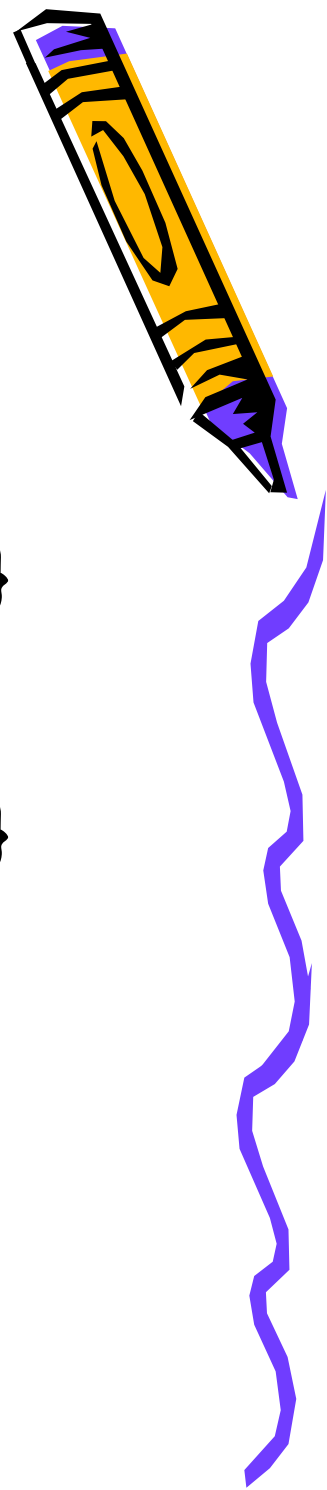
$$(B, H] = \{A, B\} - \{A, H\}$$

копирование  $(B, H]$ :

копирование  $\{A, B\}$ , удаление  $\{A, H\}$

удаление  $(B, H]$ :

удаление  $\{A, B\}$ , копирование  $\{A, H\}$





**СПАСИБО  
ЗА ВНИМАНИЕ!**

