

**调用顺序详解**

1. **主模块启动 (logic\_search.py)**
   * 接收原始矩形数据 (fixed\_area) 和格式信息 (format\_info)
   * 调用布局分析模块进行预处理
2. **布局分析 (layout\_analysis.py)**
   * **步骤 1**：按垂直位置分组矩形 (sort\_fixed\_area\_into\_lines)
     + 输入：原始矩形字典 {key: (x0, y0, x1, y1)}
     + 输出：按行分组的矩形列表 {y0: [(key, rect), ...]}
   * **步骤 2**：过滤无效间隔 (filter\_fences)
     + 移除稀疏行和行内异常间隔
   * **步骤 3**：计算平均距离 (calculate\_average\_distance)
     + 为后续关系推断提供距离基准
3. **表格结构分析 (table\_structure.py)**
   * **步骤 1**：检测网格线 (detect\_grid\_lines)
     + 从矩形边界提取可能的表格网格线
   * **步骤 2**：划分单元格 (partition\_cells)
     + 基于网格线将布局划分为逻辑单元格
   * **步骤 3**：识别表头和数据 (identify\_header\_and\_data)
     + 通过格式特征（加粗、字体大小）区分表头和数据
   * **步骤 4**：检测合并单元格 (detect\_merged\_cells)
     + 识别表格中的合并单元格区域
4. **关系推断 (relationship\_inference.py)**
   * **多策略分析**：
     + **位置关系**：右侧、包含、垂直重叠规则
     + **格式特征**：字体大小、颜色匹配度
     + **距离度量**：水平距离归一化
     + **表格结构**：同一单元格、同行列关系
   * **加权综合**：
     + 按策略权重（位置 0.6、格式 0.3、距离 0.2、表格 0.3）计算总置信度
     + 过滤低置信度关系（阈值 0.3）

数据流向如下

