|  |  |  |  |
| --- | --- | --- | --- |
| 构件ID |  | 构件名称 | CrawlerAlgorithmHandler ，  CrawlerConfig |
| 供给接口 | | | |
| 接口ID |  | 接口名称 | CrawlerConfig.updateAlgorithm |
| 语法 | boolean updateAlgorithm(Algorithm algorithm) | | |
| 前置条件 | 数据库管理员选择不同的爬虫策略算法时 | | |
| 后置条件 | 更新爬虫配置文件 | | |
| 功能描述 | 让爬虫采取不同的爬取策略 | | |
| 需求接口 | | | |
| 接口ID |  | 接口名称 | getAlgorithm |
| 提供构件 | CrawlerAlgorithmHandler | | |

|  |  |  |  |
| --- | --- | --- | --- |
| 构件ID |  | 构件名称 | CrawlerScheduleHandler，  CrawlerConfig |
| 供给接口 | | | |
| 接口ID |  | 接口名称 | CrawlerConfig.updateSchedule |
| 语法 | boolean updateSchedule(Schedule schedule) | | |
| 前置条件 | 数据库管理员重置爬虫爬取数据的时间和周期时 | | |
| 后置条件 | 更新爬虫配置文件 | | |
| 功能描述 | 让爬虫以不同的更新时间和周期运行 | | |
| 需求接口 | | | |
| 接口ID |  | 接口名称 | getSchedule |
| 提供构件 | CrawlerScheduleHandler | | |

|  |  |  |  |
| --- | --- | --- | --- |
| 构件ID |  | 构件名称 | CrawlerConfig，Crawler |
| 供给接口 | | | |
| 接口ID |  | 接口名称 | Crawler.restart |
| 语法 | boolean restart() | | |
| 前置条件 | 爬虫配置文件变更时，具体为爬虫算法和爬虫爬取数据的时间和周期 | | |
| 后置条件 | 重新启动爬虫 | | |
| 功能描述 | 能够重新启动爬虫，让它以不同的爬取策略运行 | | |
| 需求接口 | | | |
| 接口ID |  | 接口名称 | getConfig |
| 提供构件 | CrawlerConfig | | |

|  |  |  |  |
| --- | --- | --- | --- |
| 构件ID |  | 构件名称 | Crawler，CrawlerDataAnalyser |
| 供给接口 | | | |
| 接口ID |  | 接口名称 | Crawler |
| 语法 | void analyseData(html document) | | |
| 前置条件 | 爬虫从网上爬取了信息 | | |
| 后置条件 | 分析处理爬取的信息 | | |
| 功能描述 | 处理爬取的信息，使其符合统一的格式 | | |
| 需求接口 | | | |
| 接口ID |  | 接口名称 | getData |
| 提供构件 | CrawlerDataAnalyser | | |

|  |  |  |  |
| --- | --- | --- | --- |
| 构件ID |  | 构件名称 | CrawlerDataAnalyser，  CrawlerDataUpdater |
| 供给接口 | | | |
| 接口ID |  | 接口名称 | CrawlerDataAnalyser |
| 语法 | void updateData(html document) | | |
| 前置条件 | 分析处理出需要保存的信息 | | |
| 后置条件 | 传递给数据更新器 | | |
| 功能描述 | 将处理后的商品信息进一步筛选，找出相对于数据库中新的信息 | | |
| 需求接口 | | | |
| 接口ID |  | 接口名称 | getData |
| 提供构件 | CrawlerDataUpdator | | |

需要数据库的

|  |  |  |  |
| --- | --- | --- | --- |
| 构件ID |  | 构件名称 | CrawlerDataUpdater，数据库 |
| 供给接口 | | | |
| 接口ID |  | 接口名称 | CrawlerDataUp，dater |
| 语法 | boolean updateDB(Data data) | | |
| 前置条件 | 爬虫从网上爬取了信息 | | |
| 后置条件 | 分析处理爬取的信息 | | |
| 功能描述 | 将信息增量更新到数据库中 | | |
| 需求接口 | | | |
| 接口ID |  | 接口名称 |  |
| 提供构件 | 数据库 | | |

|  |  |  |  |
| --- | --- | --- | --- |
| 构件ID |  | 构件名称 | CrawlerDataAnalyser，  CompatibilityHandler |
| 供给接口 | | | |
| 接口ID |  | 接口名称 | CrawlerDataAnalyser |
| 语法 | Data compatible(html document) | | |
| 前置条件 | 处理不同内容的兼容性 | | |
| 后置条件 | 返回兼容的数据 | | |
| 功能描述 | 对爬取的内容进行兼容性处理，使其具有统一的格式 | | |
| 需求接口 | | | |
| 接口ID |  | 接口名称 | proceed |
| 提供构件 | CompatibilityHandler | | |

|  |  |  |  |
| --- | --- | --- | --- |
| 构件ID |  | 构件名称 | CompatibilityHandler，  PriceCompatibility |
| 供给接口 | | | |
| 接口ID |  | 接口名称 | PriceCompatibility |
| 语法 | Void update(Rule rule) | | |
| 前置条件 | 增加新的价格兼容方案 | | |
| 后置条件 | 更新兼容性设置 | | |
| 功能描述 | 采取不同的价格兼容方案 | | |
| 需求接口 | | | |
| 接口ID |  | 接口名称 | getPrice |
| 提供构件 | CompatibilityHandler | | |

|  |  |  |  |
| --- | --- | --- | --- |
| 构件ID |  | 构件名称 | CompatibilityHandler，  LanguageCompatibility |
| 供给接口 | | | |
| 接口ID |  | 接口名称 | LanguageCompatibility |
| 语法 | Void update(Rule rule) | | |
| 前置条件 | 增加新的语言兼容方案 | | |
| 后置条件 | 更新兼容性设置 | | |
| 功能描述 | 采取不同的语言兼容方案 | | |
| 需求接口 | | | |
| 接口ID |  | 接口名称 | getLanguage |
| 提供构件 | CompatibilityHandler | | |

|  |  |  |  |
| --- | --- | --- | --- |
| 构件ID |  | 构件名称 | CompatibilityHandler，  ScaleCompatibility |
| 供给接口 | | | |
| 接口ID |  | 接口名称 | ScaleCompatibility |
| 语法 | Void update(Rule rule) | | |
| 前置条件 | 增加新的计量单位兼容方案 | | |
| 后置条件 | 更新兼容性设置 | | |
| 功能描述 | 采取不同的计量单位兼容方案 | | |
| 需求接口 | | | |
| 接口ID |  | 接口名称 | getScale |
| 提供构件 | CompatibilityHandler | | |