

# 表达式：UrlGenerator

Beyebe项目 · 小蜜蜂

## 表达式：UrlGenerator

单url

最简语法

多url

最简语法

自定义操作符与增量

POST参数

语法

自定义属性

语法

事务组

语法

UrlGenerator表达式用于生成一个、多个或者一组url，多个表达式以换行符分隔。

## 单url

### 最简语法

```
=url
```

在一个完整url的前面添加一个等号，表示为任务种子的起始地址增加该url。

**Demo：**

```
1.  输入 :
2.      =http://z.cn
3.
4.  输出 :
5.      http://z.cn
```

## 多url

### 最简语法

```
~url...{from,to}...
```

多url模式用于为任务种子的起始地址添加多个url。

#### Demo :

```
1.  输入 :
2.      ~http://z.cn?page={2,5}
3.
4.  输出 :
5.      http://z.cn?page=2
6.      http://z.cn?page=3
7.      http://z.cn?page=4
8.      http://z.cn?page=5
```

### 自定义操作符与增量

多url模式默认从“from”为起点，按每次“+1”的方式执行“to”次，生成“to”个url。

这里“+1”动作的操作符“+”和增量“1”都可以自定义。

#### 语法

```
~url...{from,to | [+*/]增量}...
```

## Demo :

```
1.  输入 :
2.      ~http://z.cn?page={0,3 | +20}
3.
4.  输出 :
5.      http://z.cn?page=0
6.      http://z.cn?page=20
7.      http://z.cn?page=40
8.      http://z.cn?page=60
9.
10. -----
11.
12. 输入 :
13.      ~http://z.cn?page={0,3 | *3}
14.
15. 输出 :
16.      http://z.cn?page=0
17.      http://z.cn?page=3
18.      http://z.cn?page=6
19.      http://z.cn?page=9
```

## POST参数

单url模式和多url模式都能添加post参数列表，拥有post参数列表的url会被以post方式发送请求。

## 语法

```
= [key1:value,key2:value2...]url
~ [key1:value,key2:value2...]url
```

## 示例 :

```
1.  = [page:3,keyword:"手机"]http://z.cn
```

# 自定义属性

单url模式和多url模式都能添加自定义属性，添加的自定义属性会包含在该url的采集结果中。

POST参数列表和自定义属性可以在一个表达式中共存，自定义属性需要放置在POST参数前面。

## 语法

```
={key1:value,key2:value2...}url  
={key1:value,key2:value2...}url
```

### 示例：

1. `={city:"深圳",category:"护肤"}http://z.cn`
- 2.
3. `={city:"深圳",category:"电子产品"}[page:3,keyword:"手机"]http://z.cn`

# 事务组

有些网站需要使用精确的采集路径进入页面才能正常打开，使用事务组，可以让采集器在单个Browser对象内依次访问指定的url。

## 语法

```
(UrlGenerator),(UrlGenerator),(UrlGenerator)...(UrlGenerator)
```

### 示例：

1. `(=http://waimai.baidu.com/waimai?  
qt=shoplist&lat=4828483.6823517&lng=12933686.676485&address=address),(  
={category:"超市购",point:"苹果园地铁站",city:"北京  
市"}http://waimai.baidu.com/waimai/shoplist/d246740c572f1145?  
taste=10&display=json&page=1&count=40)`

