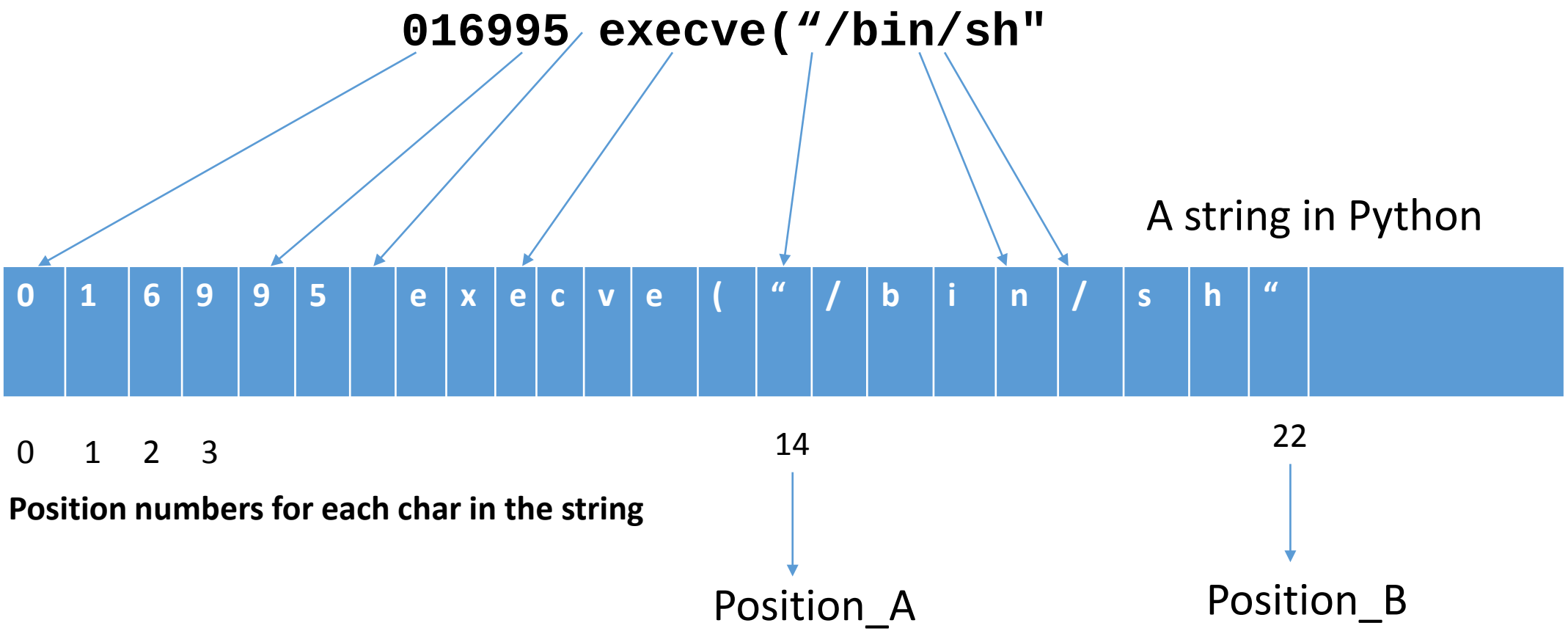
A lecture on how to extract program names

Key Observation

--- Program names are contained inside `execve` events

--- Every `execve` event log entry is a `string`

What is a string in Python?



The program name starts at position 15, which is Postion_A+1

Extract program names in 4 steps

- **Step 1: create a list to hold all the extracted program names**
- **Step 2: extract a program name from the current execve event log entry**
  - **Step 2A: find Position_A**
  - **Step 2B: find Position_B**
  - **Step 2C: extract the sub-string between Position_A and Position_B**
- **Step 3: append the extracted program name to the list created in Step 1**
- **Step 4: locate the next execve event log entry; go to Step 2**

Step 1

Here is the Python code:

```
program_names = []
```

Step 2A

Here is the Python code:

```
if ‘ execve(’ in line
    position_a = line.find(‘”’)
```

Here is the real value based on the example on slide 3:

```
position_a = 14
```

Step 2B

We do NOT want Python to find ‘”’ from the beginning of the string!

Here is the Python code:

```
position_b = line.find(‘”’, position_a+1)
```

Here is the real value based on the example on slide 3:

```
position_b = 22
```

Step 2C

Here is the Python code:

```
name = line[position_a+1 : position_b]
```

The char on this position will be extracted

The char on this position will NOT be extracted!

All the chars in between will be extracted!

Here is the real value based on the example on slide 3:

```
name = /bin/sh
```

Step 3

Here is the Python code:

`program_names` is a list of program names; `name` is a string which holds the program name extracted in Step 2; the `append` method will append `name` to the list

```
program_names.append(name)
```

Step 4

Here is the Python code:

```
for line in lines:
    if ‘ execve(‘ in line
```