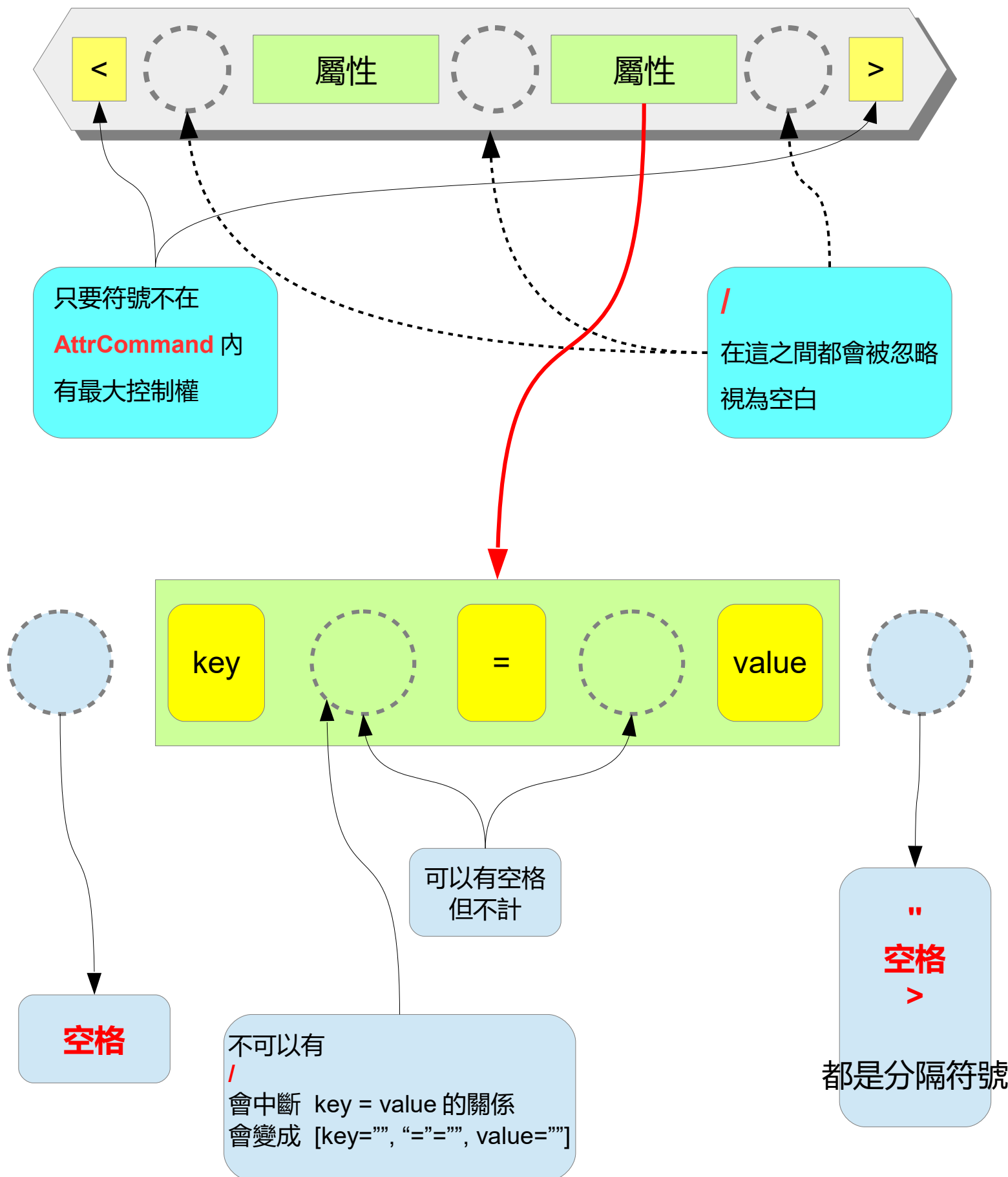
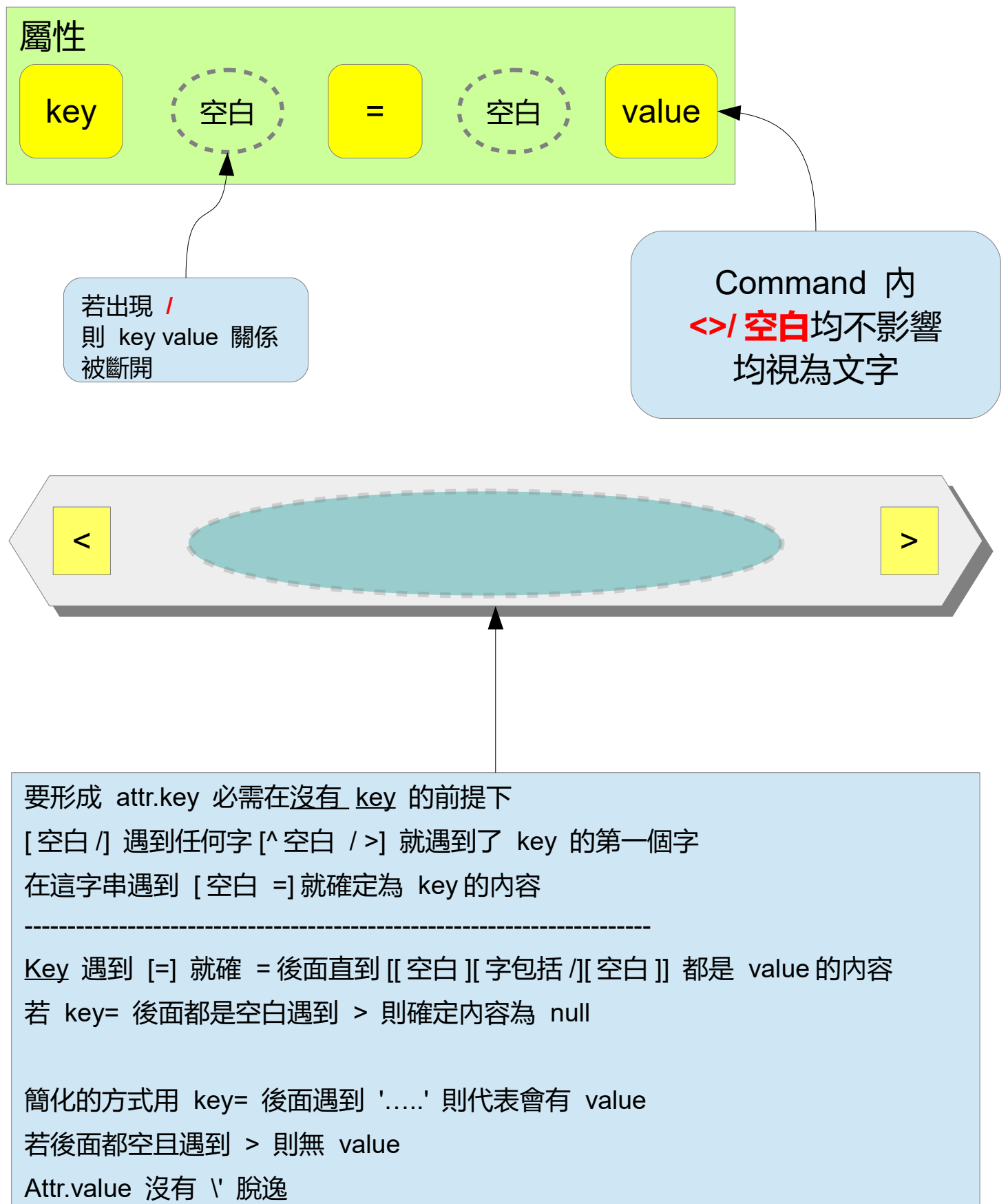


Html 標籤內的分隔符號

< > / 空白 是 4 個重要符號





文本解析方式

似乎是從頭開始解析 以 (<>) 以這兩為重

```
<div>  
<p class="</div>">  
</p>  
</div>
```

文本

文字

< 疑似標籤但其實是文本

文字

< 標籤 >

遇到 <
進入辨識狀態

找尋 >
遇到 x='___' 則跳過

若無法形成正常節點
則變成文字節點

若前方是文字節點與之合併

不管是 單標籤, 雙標籤的頭,
雙標籤的尾

都是一個獨立節點

最明顯例子

```
<div x=">"TODO write content</div>
```

這個 div 並沒有正確 close

一些標籤要注意的地方

end 標籤

不會有 attr

style 標籤

無套嵌

<style> 後面文字只要有遇到 </style>
就 close

script 標籤

無套嵌

<scrit> 文內 若有遇到
下面幾種文字組成 需跳過掃描
' '
//-----\n

domTree 的組成

tagNodeList

body

div

text

div

hr

/div

/div

/body

經過 一次 `<div>`
兩次 `</div>` 就能確定閉合點

