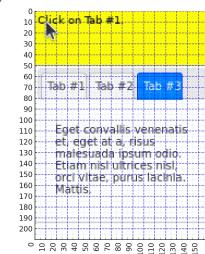


## Code generation (prior to training)



Task description



intents  
(e.g., click  
on a tab)

object template  
images

Tab #1 Tab #2 Tab #3



VLM

script  
</>

Critique  
VLM

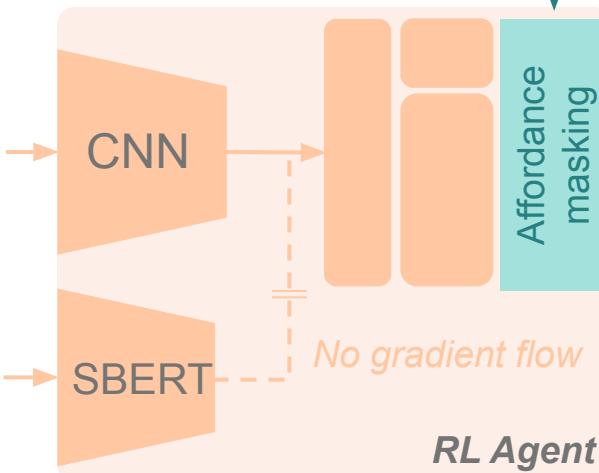
reviews +  
regenerates  
script  
</>  
(x3)

```
def determine_affordable_actions(observation)  
</>
```

{affordances}



“Click on  
Tab #1”



Action type  
(e.g., CLICK)

Pixel  
coordinates  
(X, Y)

observation

reward



action

Web environment