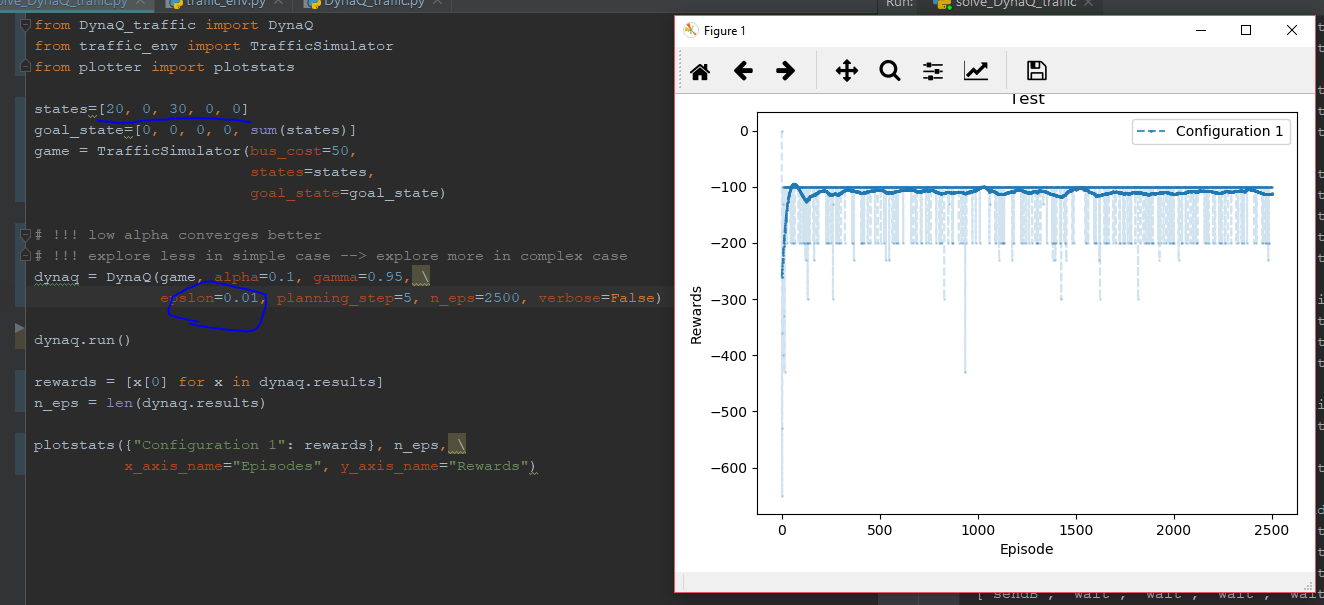
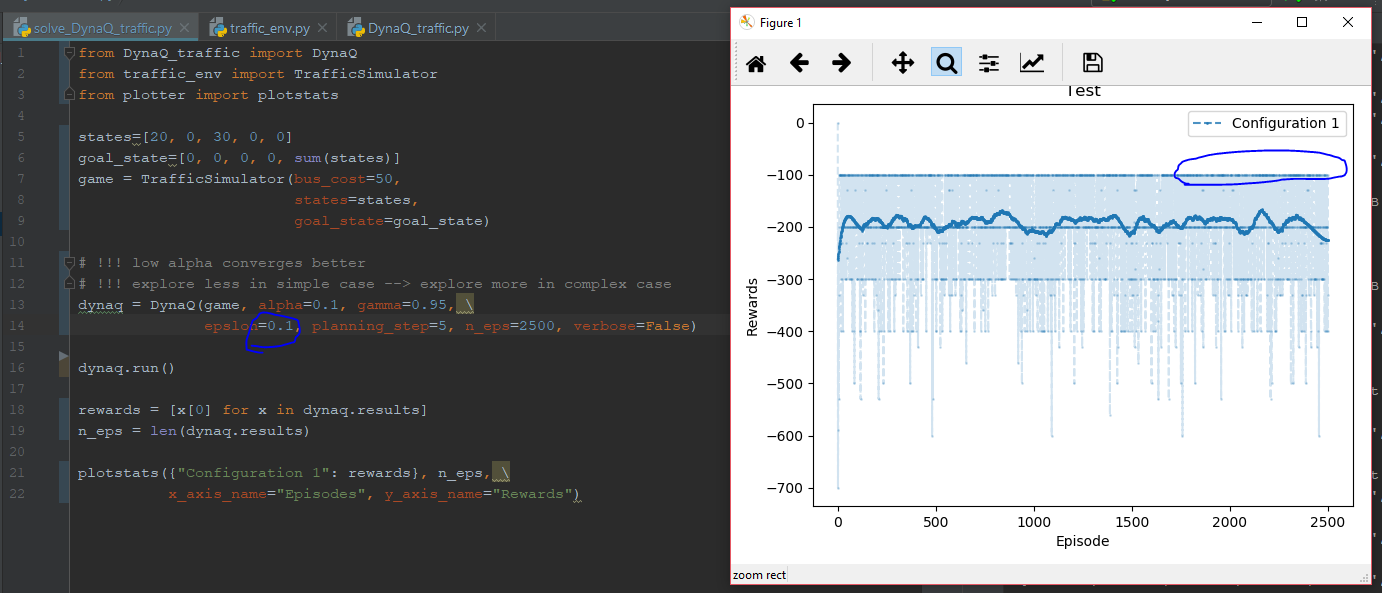
## Simple case

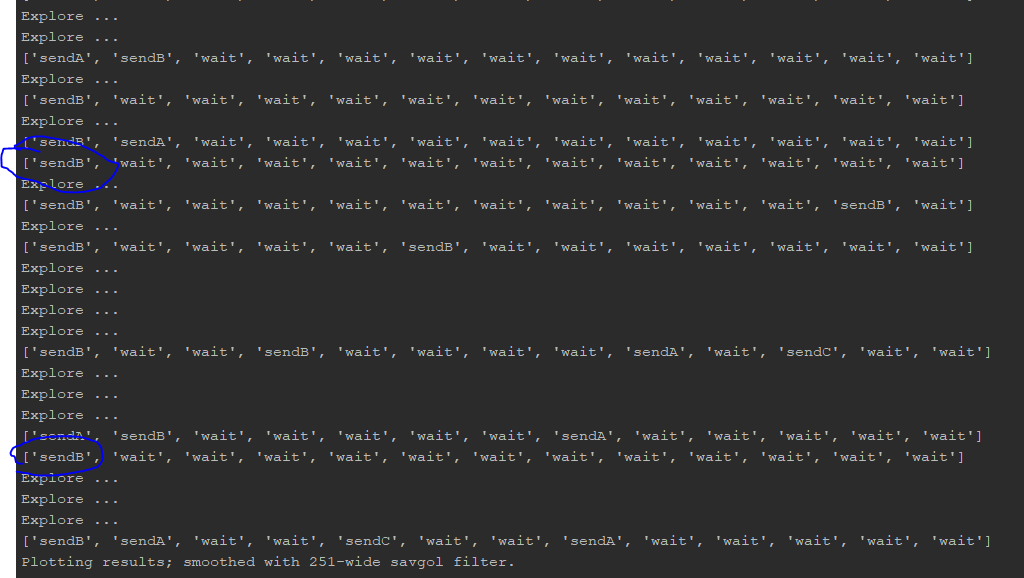
### Low explore



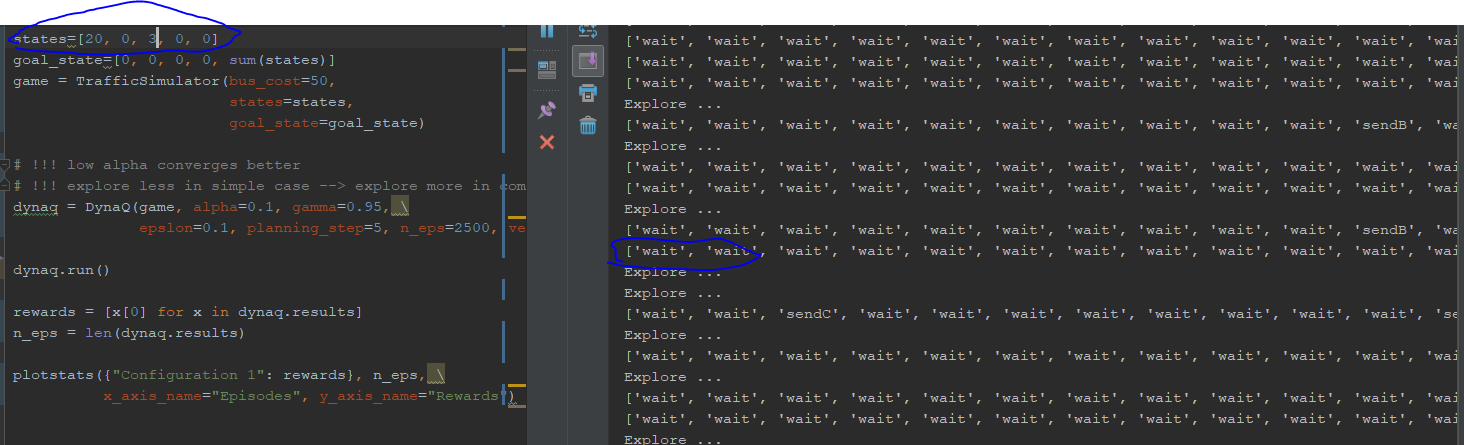
### High explore



Optimal policy still found:



## Another simple case



It knows to not sending bus when there are less passengers waiting

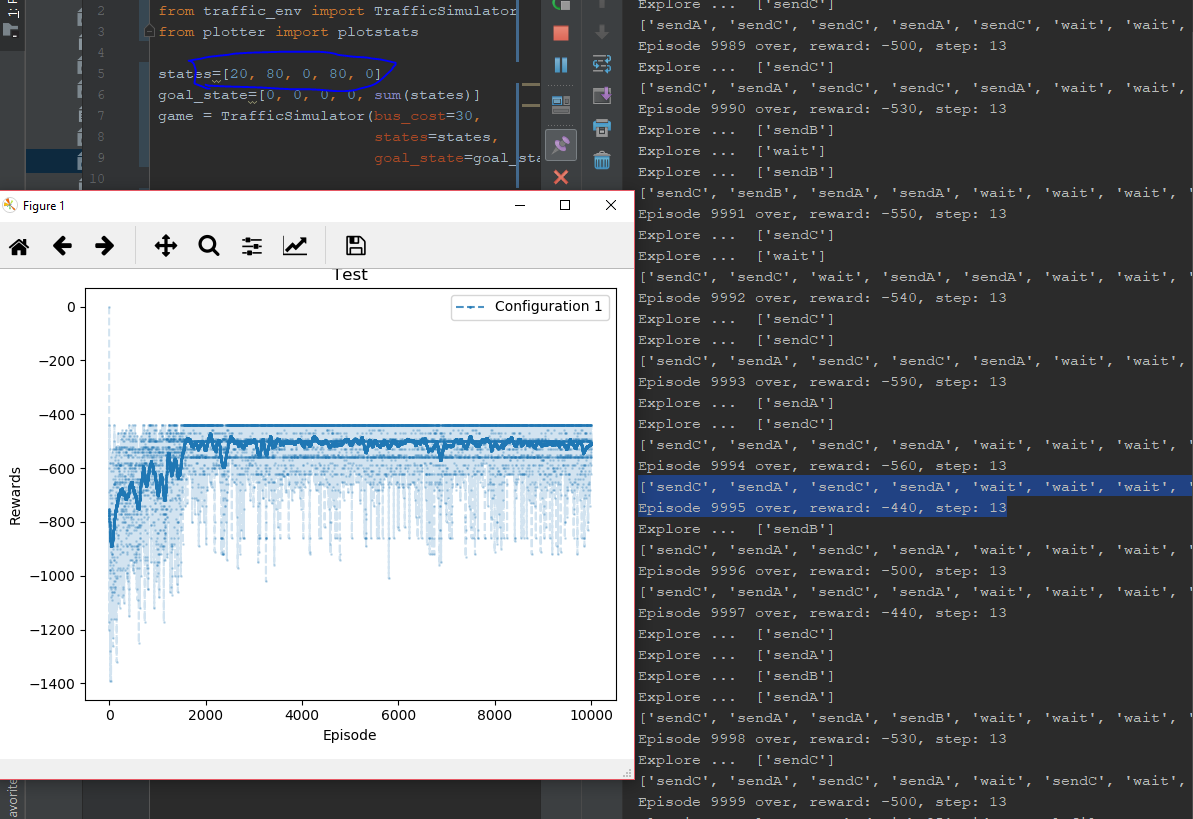
## Complex case

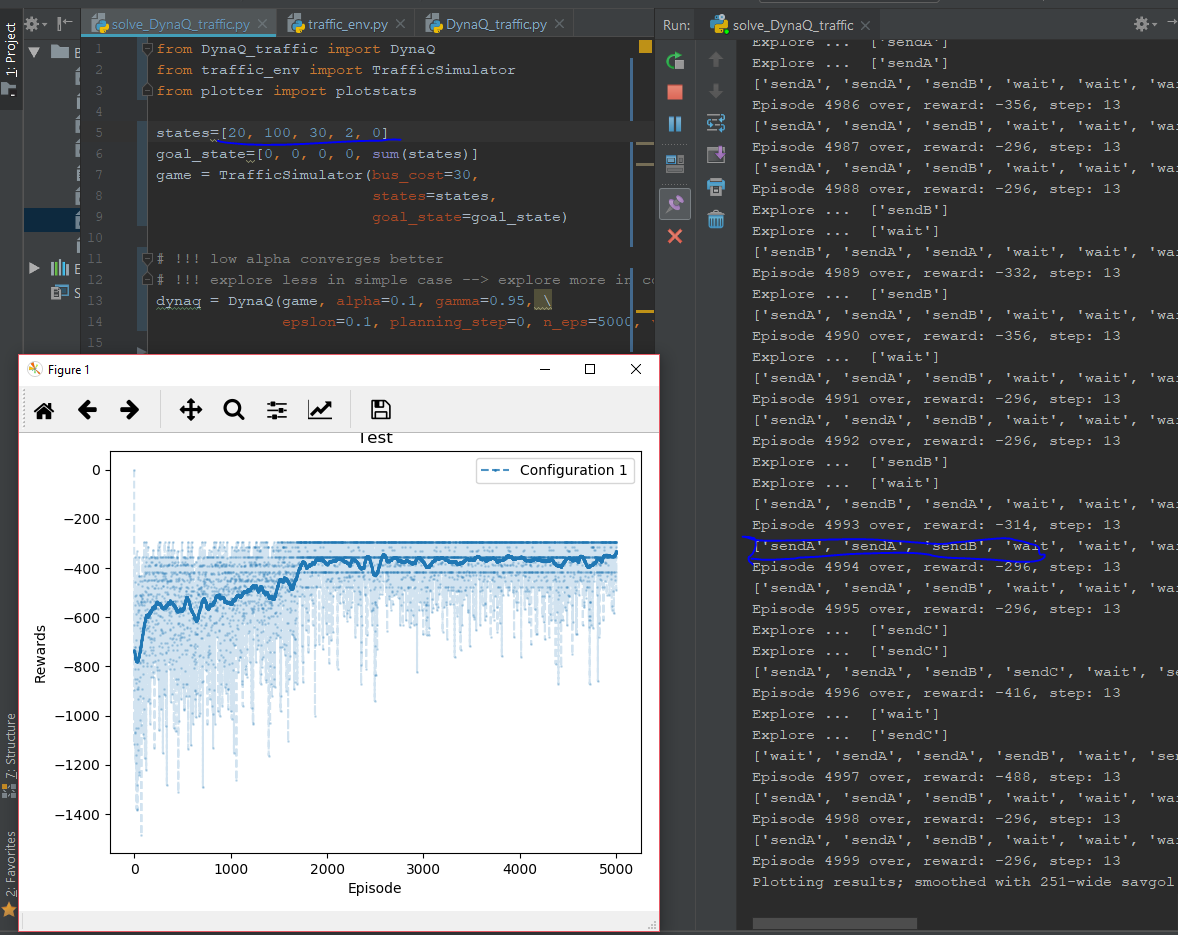
Leaning from experience and exploring become much more important.

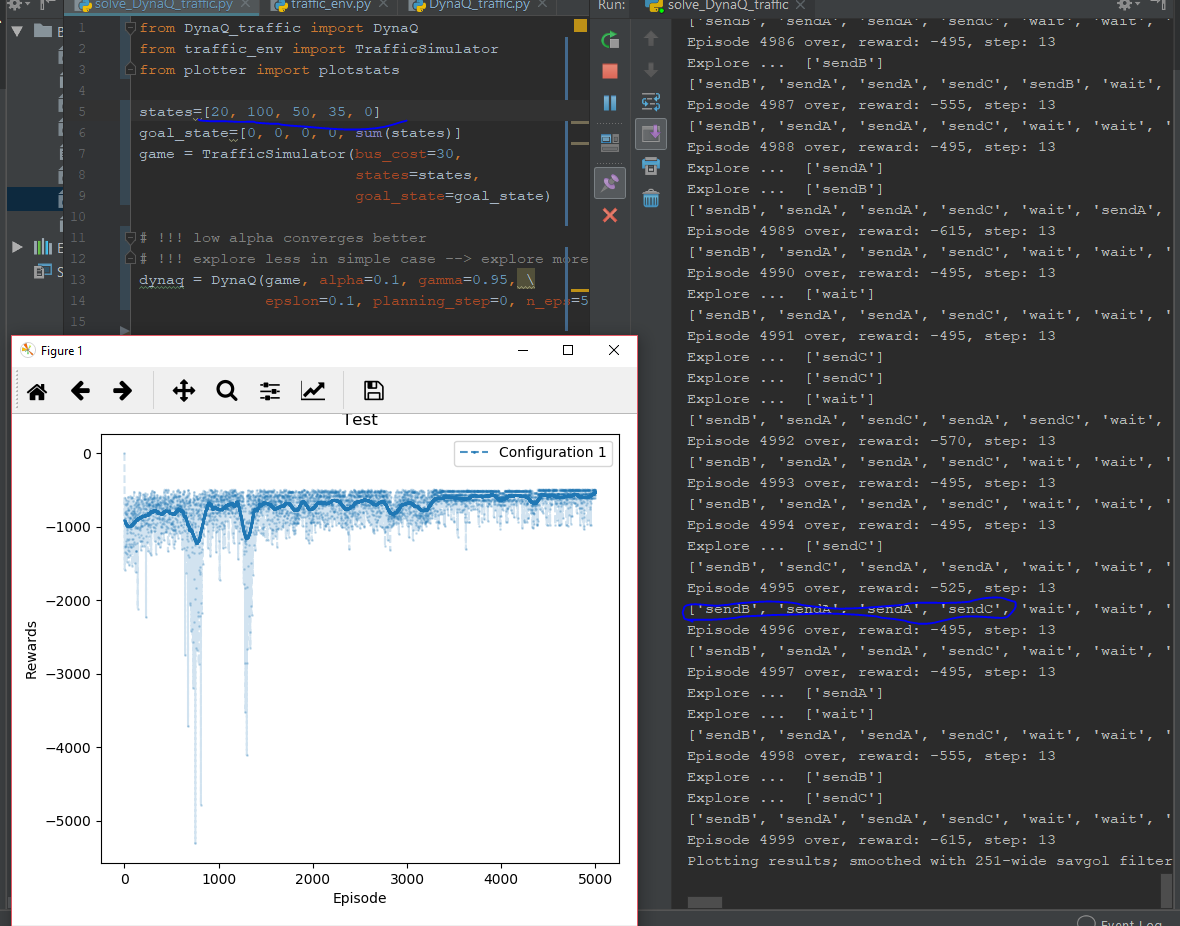
When exploring rate is 0.1, the result plot doesn’t converge because it explores to much instead of using optimal policy

To fix this, run optimal policy at the end of dynaq (no exploring)

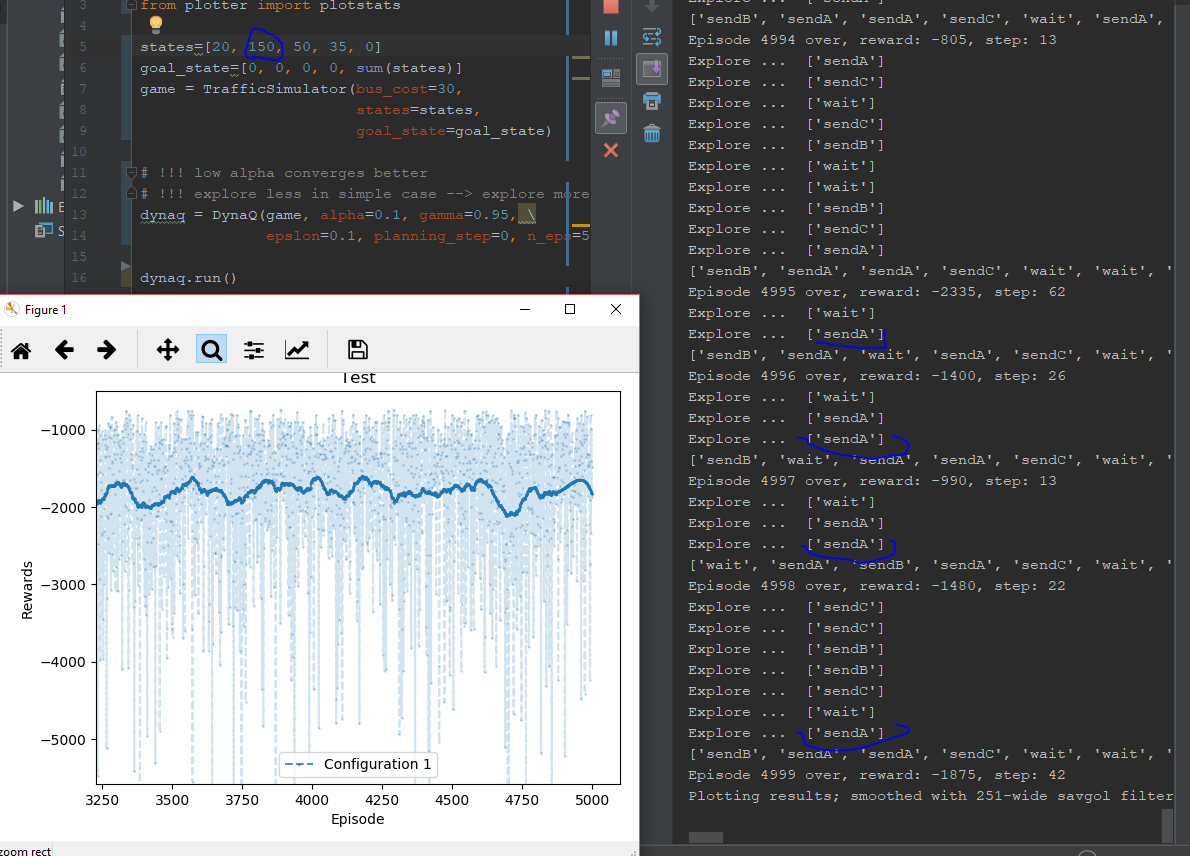
### Case that works







Fail to converge:



Dynaq is not able to handle dynamic environment now. Need to upgrade!