Alpha: .75; gamma: .5

```
############################
## - - - X - - - - - - - ##
## - - - X X X - - - - - ##
## - - X T X X X X - - - ##
## - - - X X - - - B - - ##
## - - X X X - - - - - - ##
## - - X X X X P - P - - ##
## - - X X X X - - - - - ##
## - - - X X - - - - - - ##
## - - - - - - - - - - - ##
## - - - - - - - - - - - ##
############################
Burgler score: 241; Pony report: 66% of
ponies saved!
```

Alpha: .1; gamma: .5

```
############################
## - - - - - - - - - - - ##
## - - - - - - - - - - - ##
## - - - T P - - - X X ##
## - - - - - - - - - B - ##
## - - - - P - - - - - - ##
## - - - - - - P P - P - ##
## - - - P - - - - - - ##
## - - - - - - - - - - - ##
## - - - - - - - - - - - ##
## - - - - - - - - - - - ##
############################
Burgler score: 17; Pony report: 0% of
ponies saved!
```

Alpha: .5; gamma: .5

```
############################
## - - - - - - - - - - - ##
## - - - - - - - - - - - ##
## - - - T P - - - - - - ##
## - - - - - - - - - B - ##
## - - - - X X - X - - - ##
## X X - X X X X - P - - ##
## X - - X X X - - - - - ##
## X - - X X - - - - - - ##
## X X X - X - - - - - - ##
## - X - - - - - - - - - ##
############################
Burgler score: 157; Pony report: 66% of
ponies saved!
```

Alpha: .9; gamma: .5

```
############################
## - - - - - - - - - - - ##
## - - - - - - - - - - - ##
## - - - T X - - - - - - ##
## - - - - X X - X B - - ##
## - - - X X X X - - - ##
## - - X - X X X - P - - ##
## - X X X X X - - - - - ##
## - X X X X - - - - - - ##
## - X - X - - - - - - - ##
## - - - - - - - - - - - ##
############################
Burgler score: 249; Pony report: 83% of ponies
saved!
```

Alpha: .6; gamma: .1

```
############################
## - - - - - - - - - - - ##
## - - - - - - - - - - - ##
## - - - T P - - - - - - ##
## - - - - - - - - - B - ##
## - - - X X X - X - - - ##
## - - - X - X X X X X ##
## - - X X - - X X - X ##
## - X X - - - - - X - ##
## X - - - - - - - - - ##
## - X X - - - - - - - ##
############################
Burgler score: 307; Pony report: 83% of ponies
saved!
```

Alpha: .6; gamma: .5

```
############################
## - - - - - - - - - - - ##
## - - X X X X - X - - ##
## - - X T X X X - X - ##
## - - - X X - - - B - ##
## - - - X X - - - - - ##
## - - - X X X P - P - ##
## - - - X X - - - - - ##
## - - X X X X - - - - ##
## - - - X - X - - - - ##
## - - - - X - - - - - ##
############################
Burgler score: 387; Pony report: 66% of ponies
saved!
```

Alpha: .6; gamma: .9

```
############################
## - - - - - - - - - - - ##
## - - B - - - - - - - ##
## - - - T P - - - - - ##
## - - - - - - - - - E - ##
## - - - - P - - - - - ##
## - - - - - P P - P - ##
## - - - P - - - - - - ##
## - - - - - - - - - - - ##
## - - - - - - - - - - - ##
## - - - - - - - - - - - ##
############################
Burgler score: 0; Pony report: 0% of ponies
saved!
```

The biggest way I can see that the parameters affected the burglar's final policy is very simple: some converged consistently, and some resulted in infinite loops in the final exploitation (never in exploration). The original parameter setup converged almost every time. When alpha was .1, I ran it a lot and only one ever converged, which is the one included here. It seems to be a lucky one as well, since the burglar is so close to the start. It seems to be the case that when the learning rate is low, the resultant Q values are arranged in such a way that the burglar moves in circles, or more likely just back-and-forth forever. This probably happened because having a low learning rate means the Q values aren't as dramatically affected, and the agent simply wasn't learning enough from any given move.

When gamma was .1, I experienced a similar frequency of infinite loops. However, it did end up getting the one I included here. I can't say whether or not this was a fluke, but I'm going to assume it's meaningful. Under this assumption, I would say that the low discount rate hindered the final Q values in a similar way to the low learning rate. But since the learning rate was higher, the agent still managed to learn well enough. Somehow, the low discount rate was causing the final Q values to be better overall, but to still be riddled with infinite loop traps.

When gamma was .9, I couldn't even get it to converge once. I'm really not quite sure why this would happen, but it seems that having a nice medium discount rate is important.

The others converged frequently (not always, however), and I can't really differentiate how effectively each parameter combination navigated, even when looking at multiple sets of result paths. It seems that each parameter combination was equally effective at collecting ponies, getting to the exit, and avoiding trolls as long as it didn't loop infinitely. Even one that looped infinitely most of the time (low gamma) eventually produced a path that rivals any of the other, more stable combinations' paths.