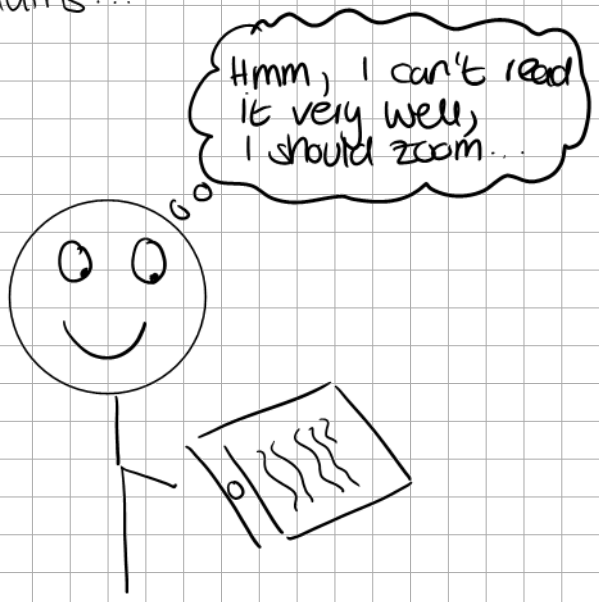
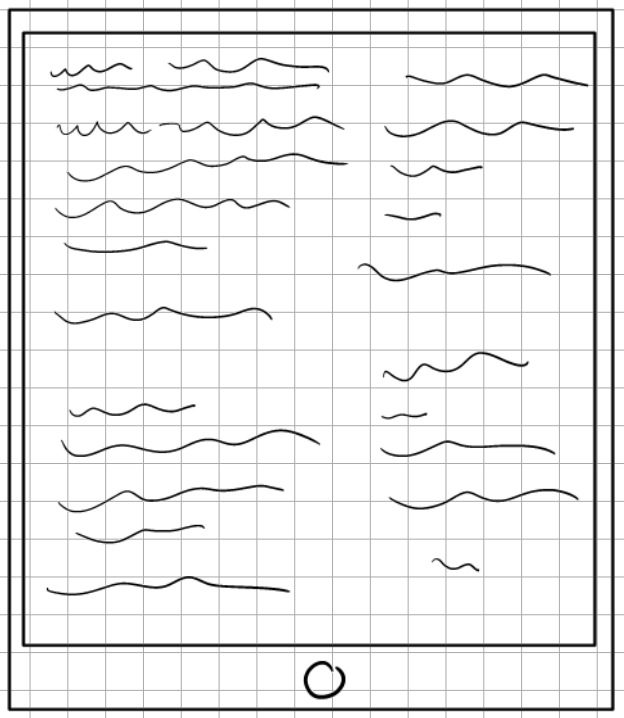
Reading a book about maths...





Zooming in...

reward function is:

$$J(\theta) = \sum_{s \in S} d^{\pi}(s)\, V^{\pi}($$

where $d^{\pi}(s)$ is the Markov chain for π_θ (o under (π). for simplicity, omitted for the policy written in full. Imagine the Markov chains

swipe to right to read right side of the text...

$$s) = \sum_{s \in S} d^{\pi}(s) \sum_{a \in A} \pi_\theta(a|$$

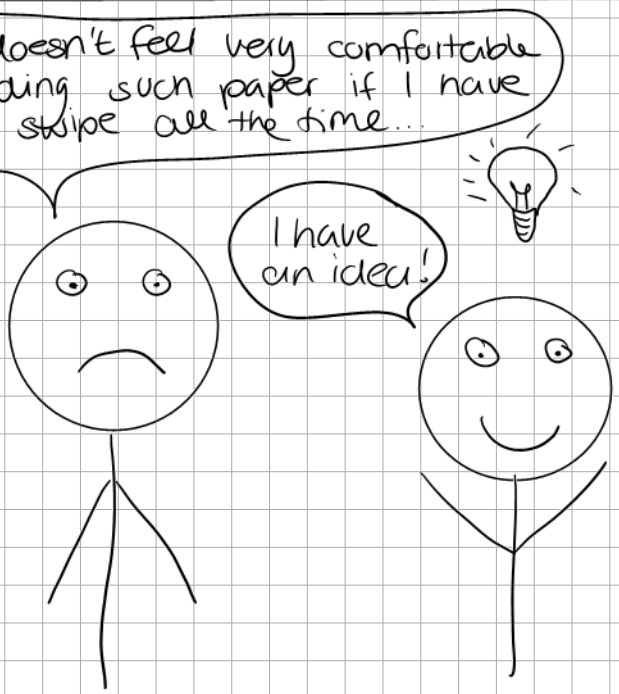
tionary distribution of

n-policy state distributio

the θ parameter would

t of other functions; if

that you can travel



Use the hint-views! It formats the text so you don't need to swipe!
And formulas don't get separated in an inappropriate way!

The reward function is:

$$J(\theta) = \sum_{s \in S} d^{\pi}(s)\, V^{\pi}(s)$$

$$= \sum_{s \in S} d^{\pi}(s) \sum_{a \in A} \pi_\theta(a|s)\, Q^{\pi}(s,a)$$

where $d^{\pi}(s)$ is the stationary distribution of Markov chain for π_θ.