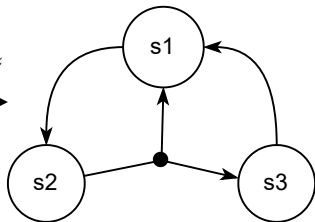
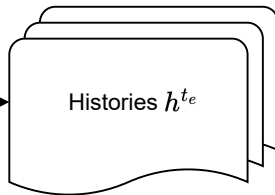


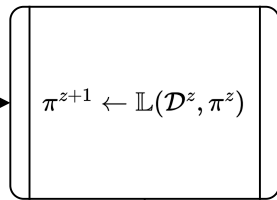
Game model



Data  $\mathcal{D}^z$



Learning algorithm



Joint policy  $\pi^z$

Updated joint policy  $\pi^{z+1}$