Regularization in RL

- DRA's KL
- [Todorov's KL](#)
  - Abstracts away from actions
  - KL on divergence term is of control dynamics from default dynamics
- [Linear RL](#)
  - RL version of Todorov's work where dynamics are replaced by policies
- [SAC](#)/maxEnt RL
- Moskovitz+Botvinick's regularization
  - unpublished
  - has two regularizing terms:
    1. the first term is the same as Linear RL, a KL between behavioral policy and default policy, which is learned here unlike linear RL/Todorov
    2. the second term is a complexity (max description length) penalty on the default policy
  - has an interesting information-theoretic story and easy implementation thanks to the legendary Kingma: both these terms are what come out of [variational dropout](#)
- [Noisy Nets](#)
  - Noisy weights in a DQN
  - maybe only related to DRA and none of the others

Ideas linked to DRA vs MaxEnt

- Read [lilian weng's blog on policy gradient algorithms](#)
  - DPG
  - DDPG
  - TRPO
  - PPO
  - SAC
- Read [noisynets](#) and [plappert et al 2017](#)