<https://github.com/dennybritz/reinforcement-learning/issues/7>

<https://medium.com/@ranko.mosic/online-planning-agent-dyna-q-algorithm-and-dyna-maze-example-sutton-and-barto-2016-7ad84a6dc52b>

<http://incompleteideas.net/book/first/ebook/node96.html>

<http://intelligentonlinetools.com/blog/2018/10/28/reinforcement-learning-example-planning-using-q-learning-dyna/>

<https://stats.stackexchange.com/questions/428557/dyna-q-algorithm-reinforcement-learning>

<https://becominghuman.ai/lets-build-an-atari-ai-part-1-dqn-df57e8ff3b26>

<https://medium.com/@ranko.mosic/online-planning-agent-dyna-q-algorithm-and-dyna-maze-example-sutton-and-barto-2016-7ad84a6dc52b>

<https://arxiv.org/pdf/1703.01040.pdf>

<https://github.com/kkhetarpal/COMP767/blob/master/PlanningAndLearning/Dyna_Q.ipynb>

<https://arxiv.org/pdf/1805.07813.pdf>

<https://gym.openai.com/evaluations/eval_PnbXlpV3RqyzovY8xk1aQ/>

<https://github.com/openai/gym/blob/master/gym/envs/atari/atari_env.py>

search control

model query

value update

**Full model-based training**

Our full training pipeline involves alternating between collecting data using policy, training the world model and training the policy inside the model. It requires significantly more time (several days to a week, depending on your hardware and the model you use).

To train a deterministic model:

python -m tensor2tensor.rl.trainer\_model\_based \

--loop\_hparams\_set=rlmb\_base \

--loop\_hparams=game=pong \

--output\_dir ~/t2t\_train/mb\_det\_pong

To train a stochastic discrete model:

python -m tensor2tensor.rl.trainer\_model\_based \

--loop\_hparams\_set=rlmb\_base\_stochastic\_discrete \

--loop\_hparams=game=pong \

--output\_dir ~/t2t\_train/mb\_sd\_pong

Hyperparameter sets are defined in tensor2tensor/rl/trainer\_model\_based\_params.py. Hyperparameter sets for the world model and agent are nested within loop\_hparams by name. You can change them with:

--loop\_hparams=game=freeway,generative\_model=next\_frame\_basic\_deterministic,base\_algo\_params=ppo\_original\_params