A COMPUTATIONAL MODEL OF REINFORCEMENT LEARNING AT BASAL GANGLIA

D. Hasan, Ö. Karabacak, N. S. Şengör
Istanbul Technical University, Faculty of Electrical Electronics Eng. Istanbul Turkey
daghanhasan@hotmail.com, ozkan2917@yahoo.com, neslihan@ehb.itu.edu.tr

AIM
To regenerate the responses of action network
To model ‘stimulus-response’ association learning
To use reinforcement learning during this process

MOTIVATION

Taylor & Taylor [1,2] – Action Network
- A dynamical system for BG-TH-C loop
- Temporal sequence storage and generation

Doya [3]
- Continuous time and space form of reinforcement learning
- Continuous Actor-Critic
- Radial Basis Neural Network
  - Value function
  - Policy function
- Minimize temporal difference(TD) errors
  expectation=reward

REFERENCES


PROPOSED MODEL

\[ p(t+1) = w^T f(p(t)) + w^T g(b(t)) \]

\[ b(t+1) = w^T f(p(t)) + w^T g(b(t)) \]

\[ m(t+1) = w^T f(p(t)) + w^T g(b(t)) \]

SIMULATION RESULTS

ACTION NETWORK

TRIAL 1
(r(t) is applied one time step after GO signal for ten time steps.)

TRIAL 2
(r(t) is applied simultaneously with GO signal for ten time steps.)

CONCLUSION

- Even though we could not succeed to obtain signals as action network generated, we managed to apply reinforcement signal to get similar signals.
- The signals related with movement and premovement neurons are almost same, but the signals related with simple memory are not as expected.
- If we had known more about the experiment, we could have better managed to apply reinforcement signal.
- Instead of using gradient based method it could have been better to use Hebb rule.