

—論文題目—

# 負の報酬を考慮したIL-RBMの強化学習 タスクへの応用

Application of IL-RBM Considering Penalty to Enforcement  
Learning

指導教授

萩原 将文 教授

学習指導副主任

今井 倫太 教授

慶應義塾大学 理工学部 情報工学科

平成 27 年度

学籍番号 61214423

中村 晃貴

# 目 次

あらまし	1
参考文献	3
付録	4
付録 A 異なる学習率を用いた予備実験の結果	4

## あらまし

本論文ではRBMを改良し，強化学習タスクへの応用を行った．従来のRBMは学習済みの状態から新たなデータセットを追加学習する場合に，既学習情報を失ってしまうという問題点が存在した．そこで大澤らは追加学習時に素子を追加するという手法を用いたが，このRBMを強化学習タスクへ応用する際に，報酬の正負をネットワーク内で記憶できないという欠点が存在した．そこで提案された追加学習可能なRBMにさらに二つのネットワークを持たせることで，それぞれのネットワークが正の報酬へ関連付けられるデータセットと，負の報酬へ関連付けられるデータセットを分けて記憶することを可能にした．また，今回改良したRBMを用いて，強化学習タスクへの応用を行った．評価実験では三目並べを行い，既存手法に比べ，勝率の増加と，負の報酬を取り扱えることに起因する大きな敗北率の低下を示した．[1][2][?]

## 謝辞

本研究を行うに当たって、丁寧な御指導をして頂いた萩原将文教授に深く感謝致します。

また、共同研究行って下さった大澤正彦先輩には、研究にあたり多くの助言や示唆を与えて下さり、また、システムの制作を共に行っていただきました。ここに感謝の意を示します。

また、システムの構築，論文の執筆など多くの場面で技術的，精神的に支えて頂いた研究室の皆様へ感謝を示したいと思います。

## 参考文献

- [1] Karol Gregor, Ivo Danihelka, Alex Graves, Danilo Rezende, and Daan Wierstra. Draw: A recurrent neural network for image generation. In *Proceedings of The 32nd International Conference on Machine Learning*, pp. 1462–1471, 2015.
- [2] Alec Radford, Luke Metz, and Soumith Chintala. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*, 2015.
- [3] David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. Mastering the game of go with deep neural networks and tree search. *Nature*, Vol. 529, No. 7587, pp. 484–489, 2016.

## 付録 A

### 異なる学習率を用いた予備実験の結果

??章にて行った予備実験にて，異なる学習率を用いた場合の結果のグラフを以下に示す．

まず，RBMの学習時におけるエポック数とクロスエントロピーの関係について示した後，出力層の学習におけるエポック数と誤差の関係について示す．

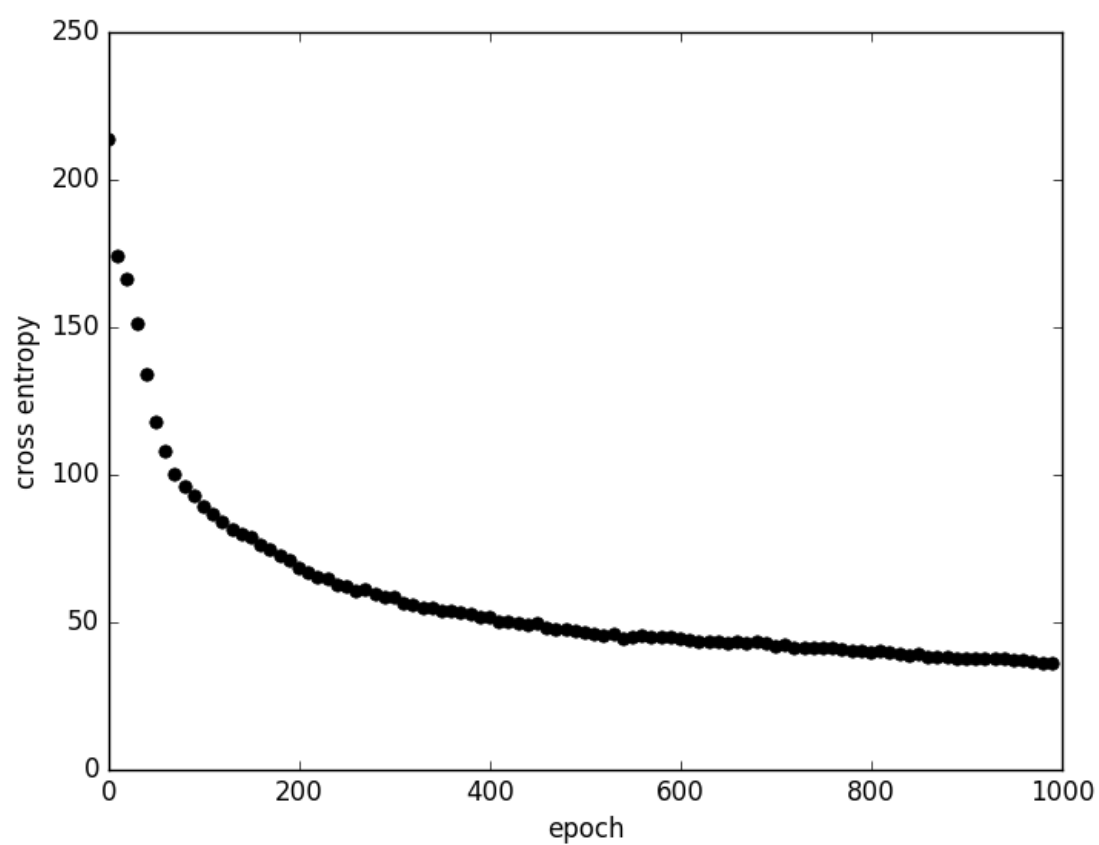


図 A.1 RBM の学習の様子 (学習率 0.01)

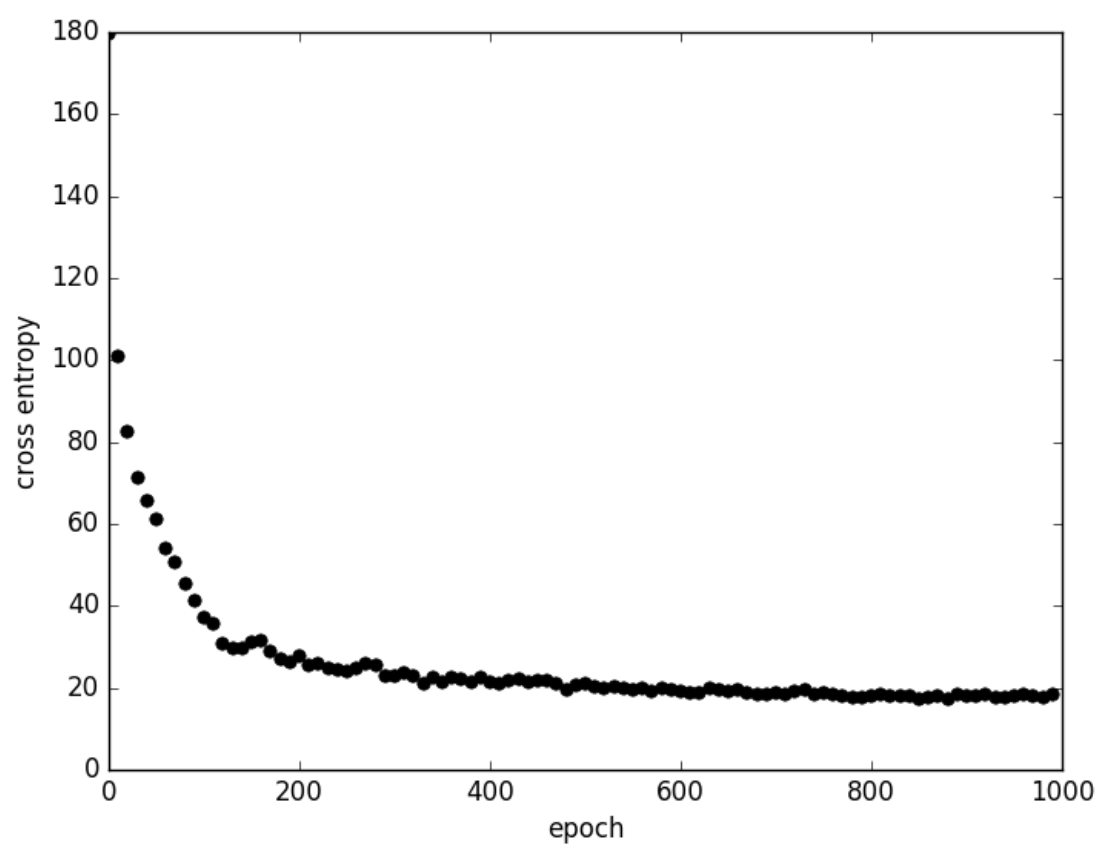


図 A.2 RBM の学習の様子 (学習率 0.05)



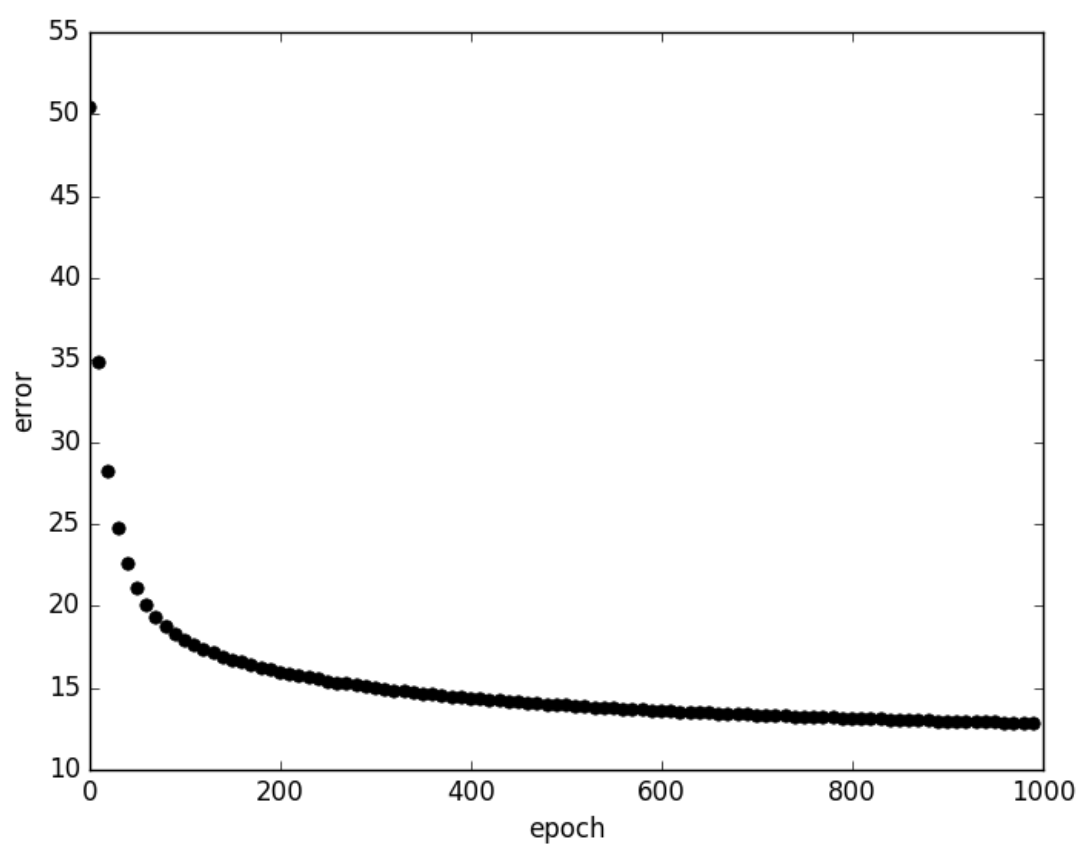


図 A.3 出力層の学習の様子 (学習率 0.01)

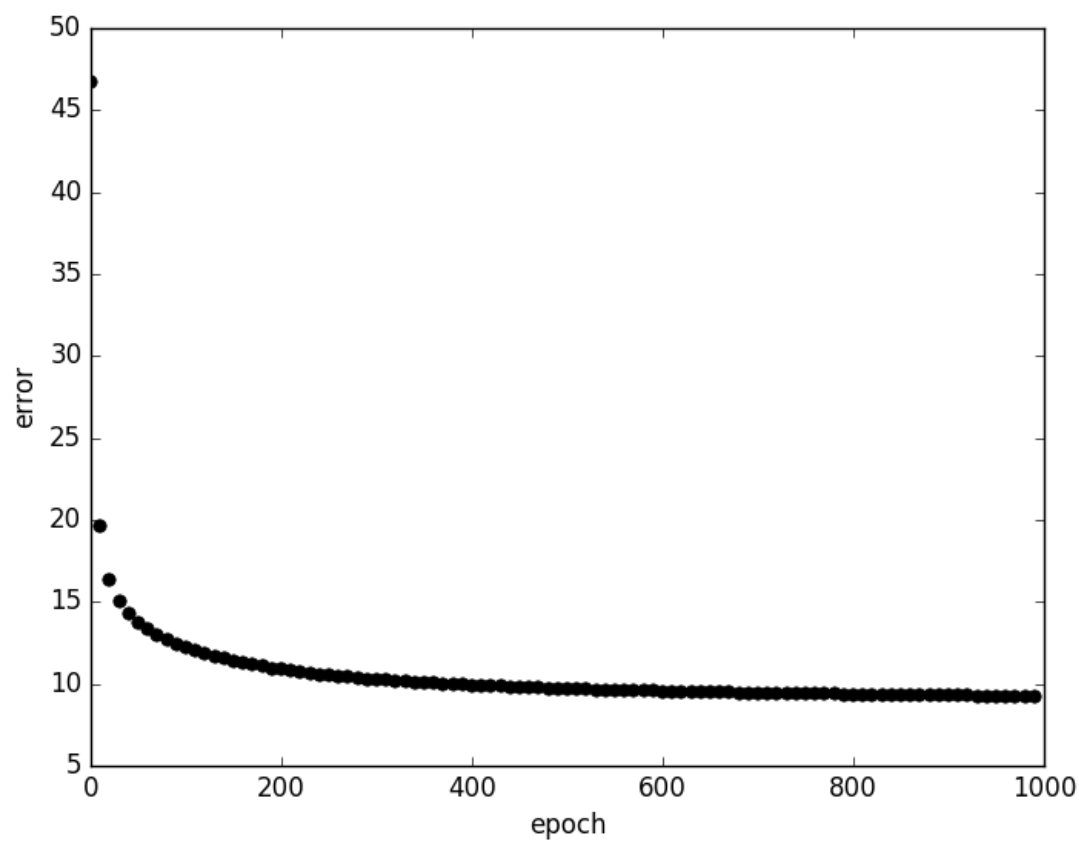


図 A.4 出力層の学習の様子 (学習率 0.05)