# 项目文档

## 摘要

本项目实现了AlphaZero算法，并在五子棋上验证了算法的有效性。同时本项目在原有AlphaZero的基础上，尝试为价值网络引入Attention Mechanism，因为其拥有全局关系建模能力，较传统CNN应该更有效。通过对基于Attention的价值网络和基于传统CNN价值网络的性能对比，我们发现利用基于Attention价值网络的AlphaZerro算法能够更快地收敛。实验结果验证了我的猜想。

## 项目背景

AlphaZero算法于2017年由DeepMind提出，在完美信息零和博弈的棋类游戏上能够击败顶尖人类玩家。研究该项目一是为了体会此绝妙算法的种种，二是为了探索将其与前沿的研究成果相结合。由此我构建了AlphaZero算法，并且在五子棋上进行训练验证。

同时我思考到传统的CNN网络只具备局部建模的能力，而全局建模能力更加适合棋类游戏，发表在Nature上的算法是基于CNN网络实现的，缺少全局关系建模的能力。目前学界火热的Attention Mechanism拥有全局建模的能力，如果将Attention和全局建模能力和AlphaZero结合，AlphaZero应该能更加强大。

## 已有方法

自从2017Nature AlphaZero现世以来，知名度较高且成熟的开源复现唯有Facebook AI 研究院的ELF OpenGo。不过该复现着重于算法的复现和平台的构建，尚未有相关报道说明了ELF OpenGo结合了相关的Attention技术。

## 我的方法

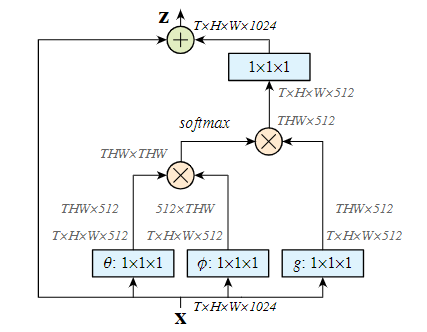
### 传统的方法

传统的方法是采用卷积层来提取棋面特征，之所以使用卷积在于其平移不变性，参数共享以及像素级建模能力。这些特征都使得卷积比其他神经网络模型比如RNN，FCNN等更加适合作为特征提取器物。在我的实现中，为了做对比，我使用了基于CNN的价值网络。

### 新的方法

尽管卷积已经较为适合作为棋盘的特征提取器了，但是卷积缺少全局建模能力。对于棋类游戏这种需要考虑全局以此来下决定的环境来说，全局建模能力显得尤为重要。为此，我引入Attention Mechanism相关组件加入原来的价值网络，为模型补充全局关系建模能力。具体而言，我使用了基于Non Local neural network（以下简称NL）作为我的Attention模块。

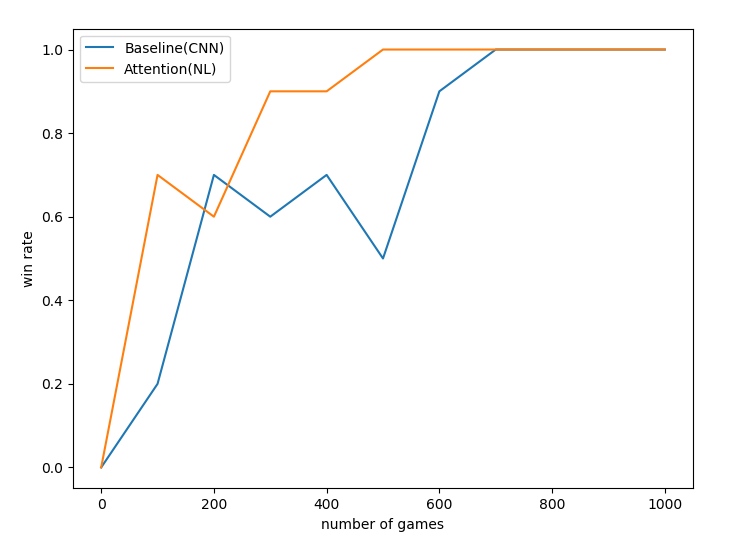
NL是出自CMU王小龙博士，是一种实现较为简单同时效果很好的Space Attention Mechanism。具体而言如下图所示



这里的theta,phi,g可以理解为NLP Transformer中的Q,K,V。由上图可知，这是一种基于Key-Value查询的Attention机制。对于输入X，首先经过三个模块的矩阵变换，之后theta和phi矩阵相乘&softmax生成权值。这里的权值的每一行表示对应像素相对于其他像素的响应，若是响应值较大，表示该像素相对于其他像素拥有更重要的特征，反之则表示该像素相对其他像素携带的特征信息较少。由此经过整个模块的变换，那些对全局特征来说重要的信息会被加强，因此更加容易被激活，更加容易被“注意”，反之则容易被抑制，即“不注意”。因为特征中元素是相对于全局来说的，所以Attention具备全局建模能力。

## 四．结果对比分析

我们在五子棋上验证了AlphaZero算法的有效性。由于算力限制，我选择的是6\*6的棋盘，任意方向连成4子即可获胜。由于测试环境是五子棋，所以我将此算法命名为AlphaBang。评价算法的标准是每隔一段时间，就将训练中的AlphaBang与1000次蒙特卡洛树搜索进行对弈，每次对弈10局，得出胜率。



上图中，横轴代表AlphaBang训练的自对弈局数，纵轴是AlphaBang和1000蒙特卡洛树搜索对弈的胜率。橘黄线代表的是添加了Attention组件的AlphaBang的表现，蓝色线代表的是传统CNN的AlphaBang的表现。可以看到基于Attention的AlphaBang更快地收敛，性能更好。

## 总结

本项目我实现了AlphaZero算法，并通过分析传统的基于CNN的AlphaZero的局限性，提出加入Attention来补充模型的全局建模能力，并且在五子棋上进行了算法验证和对比分析，证实了拥有全局建模能力的AlphaZero相较传统AlphaZero性能更好，验证了我的想法。