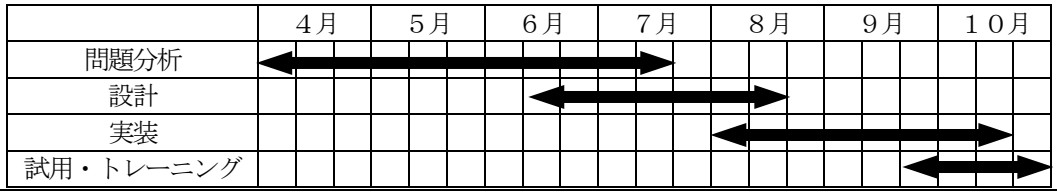
部　　門	競　技　部　門	No.1 登録番号	30021

No.2

1) 予定開発期間：6ヶ月
2) 予定開発人数：4人



No.3

実現方法

1)　司令塔の指示決定アルゴリズム

指示決定アルゴリズムは大きく分けて探索部と評価関数部の2つに分かれる。

・探索部

主な探索にはモンテカルロ木探索を使用する。モンテカルロ木探索によりランダムに手を打ち、勝てる可能性の高い手を探索する。今回は、機械学習により評価関数を練り上げて行くが、敗北に向かう手を打つ可能性は否定できない。そのため、モンテカルロ木探索を利用してランダムに回数を重ねて探索しながら明確な敗北手を防ぐことを考えている。しかし、最大120ターンという膨大な探索空間を対象に本探索を行うことは現実的ではない。従って、ある程度探索が進むごとに勝敗確率を求める関数を用いて形勢を評価する。これにより、全探索と比較して大幅な計算量の削減が期待できると考えている。

・評価関数部

評価関数には主に機械学習を用いる。最初は人間同士の対戦データを学習ケースとして使用し、ある程度の精度が得られ始めた後に機械同士のセルフプレイに移行する。セルフプレイによって学習に必要な膨大なデータを補いたいと考えている。しかし、単純なセルフプレイでは人間に勝てるほどの精度は得られないと考えている。そのため、ある程度学習データを蓄積した後、精度が低いと考えられるデータを削減して評価関数の全体的な性能向上を図る。

機械学習は、評価過程が自動化できる一方で、手動での微調整は困難である。そこで、補助として我々が作成した補助評価関数を併用する。これら2つの評価値それぞれに重み付けをし、評価の調整を柔軟に行うことによりアルゴリズムの安定化を図る。

2)　エージェントへの指示伝達方法

指示伝達はトランプとハンドサインを併用して行う。トランプでは、指示の対象であるエージェントをトランプのマークで判別し、トランプの絵柄を用いて指示が移動とタイル除去のいずれかをエージェントが判別する。ハンドサインでは、指示の方向を識別する。方向は腕を伸ばして約45度ずつ傾けて見せる。また、エージェントが指示を受け取ったらサインを返してもらい伝達を確認する。これらにより、伝達ミスを減らしたいと考えている。

3)　その他（独創的なところ）

指示伝達において司令塔側が指示のポーズを間違えた場合、たとえ伝達ミスがなかったとしても想定とは異なる動きになってしまう。そのため、あらかじめ GUI に指示のポーズを鏡合わせで表示しておき司令塔が同様にポーズを取ることで指示ミスを軽減させることを考えている。

機械学習では、入力と出力を結びつける都合上、入力に明確な特徴が現れている方が有利となる。今回のように評価関数に機械学習を用いる際には、盤面のデータを機械学習のフレームワークに入力できる形式にする必要がある。現時点では、人間同士の対戦データをその入力形式に変換する際の盤面特徴がより明確に現れるような形式を模索中である。

No.4

開発環境

【OS】Windows7/10
【IDE】Visual Studio 2017 Community
【Language】C++17
【Library】OpenSiv3D, Caffe2, ZBar