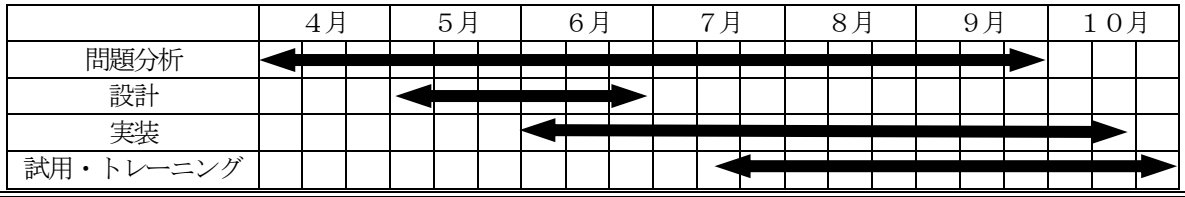
部　　門	競　技　部　門	No.1 登録番号	30009

No.2

1) 予定開発期間：6ヶ月
2) 予定開発人数：3人



No.3

実現方法

1) 公開フィールドへの事前対策

公開フィールドの攻略にあたり，機械学習を用いて事前に学習を行うことで対策とする．

深層学習と強化学習を組み合わせた DQN(Deep Q Network)を利用する．事前対策として公開されたフィールドの盤面上で配置されたエージェントが大会ルールに従い学習を進めることによって，学習データを蓄積し，これを利用して本番の競技対戦に利用する．特にエージェントが複数であるため，学習には相当な時間を要すると考えられる．

また，学習に時間がかかるため早急に実装し試用しなければならない．試用・トレーニング期間を多く用意することで調整する期間を多く設ける．そのためには学習のためのフィールドを作成し，フィールドが公開される前から調整に入れるよう準備をする必要がある．

2) 公開フィールドでの戦略アルゴリズム

公開フィールドに関しては，前述の事前対策を行い，学習データを利用して本番の競技対戦に利用する．

学習手法としては前述の通り深層学習と強化学習を組み合わせた DQN(Deep Q Network)を利用する．エージェントは自ターンである場合，移動できる範囲で動き，敵チームと対戦を行い指定のターン数経過後に勝ち負けを報酬として与える．行動についてはモンテカルロ木探索を用いて次の手をどのようにすれば最善か予測しながら学習を行っていく．

また，今回はチーム毎にエージェントが複数配置されている．特に大会ルールでは最大 8 人と多いため，エージェントが独立して動く場合自タイルを失ってしまう又は効率よく領域ポイントが獲得できない可能性がある．よって各エージェントと連携を取らなければならないために群知能を導入してチーム全体での利得を最大限にできるよう実装を行う．

3) 非公開フィールドでの戦略アルゴリズム

次の手を決定するにあたり，「そのタイルをとることによる利得」，「相手チームに与える損失」やエージェントが 1 か所に集まるのを防ぐため「エージェント間の距離」などを用いて複数の評価関数を作成する．さらに，これらの評価関数による評価に対し，遺伝的アルゴリズムを用いて最適化する．

最適化された評価関数を用いて，モンテカルロ木探索やビームサーチなど複数の探索方法を試し，最終的に勝率の高かった結果を採用して次の手を決定する．なお，今回の競技では各チームのエージェントが最大 8 人と多いため，深く探索すると組合せ爆発が起こりかねない．従って，探索を進めるごとに勝率の低い手に対する枝刈りや，探索の深さをある程度浅くするなどして計算量を大幅に削減する．

4) その他（独創的なところ）

公開フィールドに関しては，深層学習と強化学習を用いるため，長い時間をかけることによって実現できるものであり，公開フィールド攻略手段に対して十分に有効である．また，非公開フィールド攻略については，複数の探索方法を利用しながら枝切りや探索の深さを調整することによる短時間で質の高い攻略方法を取っている．

No.4

開発環境
OS：Windows10，Mac OS X，Arch Linux
IDE：VisualStudio 2017，VisualStudio code
言語：C++，Python