

RoboCup サッカーにおける共有知を用いたプランニング推測に基づく行動選択

Action Selection based on Planning Estimation using Knowledge Sharing in RoboCup Soccer

鈴木 雄大†, 中島 智晴†, 秋山 英久‡

Yudai SUZUKI†, Tomoharu NAKASHIMA†, Hidehisa AKIYAMA‡

大阪府立大学†, 福岡大学‡

Osaka Prefecture University†, Fukuoka University‡

sya01138@edu.osakafu-u.ac.jp

tomoharu.nakashima@kis.osakafu-u.ac.jp

akym@fukuoka-u.ac.jp

Abstract

The aim of this research is to achieve kicker's planning using the idea of knowledge sharing between players. When a kicker makes a decision he plans action sequences. However, our team opuSCOM has a low success rate of the planning action sequences. It is caused by no consideration of other players' movements in the kicker's planning, which is a difficult task. We employ knowledge sharing to consider the kicker's planning and to select an action. In the experiment, we show effectiveness of the knowledge sharing.

1 はじめに

ロボット工学と人工知能の領域横断型研究プロジェクトとしてRoboCup [1] が知られている。RoboCup には様々なリーグが存在しており、それぞれのリーグにおいて活発に研究、開発が行われている。チーム開発においては、ただ単に勝利するだけでなく、賢く安定して勝利することが望まれている。ランダムや思いつきで作られた戦術を使って勝利することよりも、緻密なデータ分析や機械学習によるモデル化を活用して勝利につながる戦術を生成することがRoboCupの理念と合致する。ドリブル等の行動を洗練する [2], 相手の行動を予測して意思決定を行う [3], 相手チームに合わせた戦術をとる [4], といった研究が様々な観点から進められている。

RoboCup サッカーリーグの一つであるサッカーシミュレーション 2D リーグ [5] では、近年戦術が多様化しており、相手戦術との相性が試合結果に大きな影響を与えている。このような背景から、勝利のために、相手に合わせた戦術に切り替える研究が行われている。しかし、これには相手戦術の見極めや、複数の対抗戦術を用意する必要が

ある。そこで、相手戦術に依存せず、自身の戦術の完成度を高めていく実装に取り組んでいる。

本論文では、オフENS時のプレイヤーの行動選択を研究対象とする。オフENS時、ボールを保持したプレイヤー (Kicker) は、より有効な戦術を実現するために、パスやドリブルなどのプランニングを数手先まで行い、行動候補の中からより評価値の高い行動を選択する。ボールを保持していないプレイヤー (nonKicker) は、フォーメーションに従うことのみを目的としたポジショニングを行い、そこにはKickerのプランニングが考慮されていない。本研究室で開発しているチーム opuSCOM は、Kickerがプランニングを行っているにも関わらず、そのプランニング通りに行動できていない場面が多い。この問題を解決するために、本論文では nonKicker の行動に着目し、共有知を用いたプランニング推測に基づく行動選択を提案する。全プレイヤーが推測したプランニングに基づいて行動することで、プランニング通りの行動を実現させることを目的とする。この目的を達成することで、opuSCOM の性能の向上を図る。数値実験により、共有知の実装前後における opuSCOM の性能の変化を調査する。

2 研究背景

RoboCup サッカーシミュレーション 2D リーグでは、戦術の多様化が進んでおり、各チームは様々な戦術を用いる。それに伴い相手戦術との相性が試合結果に大きな影響を与える。この問題に対して、相手戦術に合わせた戦術への切り替え、または、相手戦術に依存しない戦術の実装が対応策として挙げられる。戦術切り替えでは、Nakashimaら [6] が提案した勝敗予測などから相手戦術の識別を行うことに加え、多様な相手戦術の定義や、自身のチームに複数の戦術の実装が必要である。よって、相手の全戦術に適切な対応を行うことは困難と考えられる。そのため、相手戦術に依存しない有効な戦術を実装することが望ましい。全ての戦術に対応するには、プレイヤーがあらゆる状況が

ら動的にプランニングを行い、有効な行動を選択することが重要である。本研究室で開発しているチーム opuSCOM は、nonKicker や敵の移動など未来の状態を正確に考慮できておらず、プランニングが効果を発揮していない。未来の状態を Kicker が考慮する手法として、Akiyama ら [7] の、Delaunay Triangulation を用いたボール位置に基づくポジショニングを用いて味方プレイヤーの位置を予測する手法が考えられる。しかし、この手法の場合、フォーメーションは予め設定されたものであるため動的なプランニングができていないとも言えない。そのため、敵のあらゆる戦術に対応できないと考えられる。また、敵プレイヤーの位置を予測する研究が山下ら [8] によって提案されたが、精度は十分なものではない。これらのことから、Kicker のプランニングを Kicker による予測のみで有効にするのは困難である。そこで本論文では、Kicker のプランニングではなく、nonKicker の行動選択に着目した。未来の状態を予測するなど Kicker のプランニングの精度を高めるのではなく、nonKicker が Kicker のプランニングを考慮して行動することで、問題の解決を図った。Kicker のプランニングを考慮する方法として、評価関数と行動候補を共有知として用いる。従来では nonKicker はフォーメーションに従い、Kicker のプランニングを考慮していなかった。本論文では、共有知を用いたプランニング推測により Kicker のプランニングを考慮した行動選択を提案する。これにより、Kicker のプランニング通りに行動できるようになり、opuSCOM の性能を向上させることができると考えた。

3 共有知を用いたプランニング推測

3.1 プランニング

プランニングとは、パスやドリブルなどの行動候補から、より評価関数の評価値が高い行動、もしくは行動列を探索することである。Kicker は行動決定の際、より評価値の高い行動を選択するためにプランニングを行う。行動列において、直近の行動を深さ 1、その次の行動を深さ 2 と呼ぶ。本論文では、行動意思決定のモデルとして Akiyama ら [9] の提案した木探索による協調行動プランニングを用いる。このモデルでは、ボールキック時において探索木を生成し探索することにより、行動プランを作成している。本論文では、行動プランを次にプレイヤーが行うべき一定数の長さを持つ行動列と定義する。また、探索木の走査アルゴリズムとして、最良優先探索を用いる。

以下の手順により行動プランを作成する。まず、ルートノードに現在の状態を格納する。そして、ルートノードから行動候補を生成する。このとき、プレイヤーが観測した現在の状態から、自分と他者を含めた複数のプレイヤーによって実行される行動候補（パスやドリブル、シュートなど）を生成する。生成された行動候補の評価値を評価関

数により計算し、行動候補と評価値を探索木へ子ノードとして格納する。すべてのノードが追加された後、評価値が最大のノードを選択し、そのノードからさらに行動候補を生成する。これを繰り返すことで、探索ノード数があらかじめ設定された最大値に達するまで探索木を成長させる。ただし、木の深さがあらかじめ設定した値を越える場合や、行動列の終了条件に設定されている行動（シュート）が生成された場合は、その葉ノードでの子ノード生成は行わないものとする。構築された木構造の中から評価値の高いノードをつなげると、行動列が得られる。探索終了後、生成された行動列の中で評価値が最大の行動列を行動プランとする。

プランニングにより、プレイヤーは、より戦術的価値が高い行動を選択することが可能となる。最良優先探索に基づく行動列の探索例を図 1 に示す。図 1 では、簡略化のためノードには行動の評価値のみを記し、エッジ上に行動を記す。図 1 では、初期状態からパスが 2 つ、ドリブルが 1 つ生成されている。そして、それぞれに対し評価関数により、評価値を計算している。その結果、パスにはそれぞれ評価値 30、評価値 20、ドリブルには評価値 15 が計算されている。この中で最も評価値の高い行動である、評価値 30 のパスから、行動候補をさらに生成し評価値を計算している。深さ 1 でパスを受け取るとプランニングされた nonKicker を nonKickerB、深さ 2 でパスを受け取るとプランニングされた nonKicker を nonKickerA とする。仮に、図 1 の状態で探索が終了されたならば、この探索木では、Kicker が nonKickerB へパスを行い、その後さらに nonKickerB が nonKickerA へパスを行う行動プランが生成される。現在 opuSCOM は、深さ 2 時点でほとんどの場合、行動プラン通りに行動できていない。これは、Kicker のプランニングを nonKicker が知らないことに起因している。

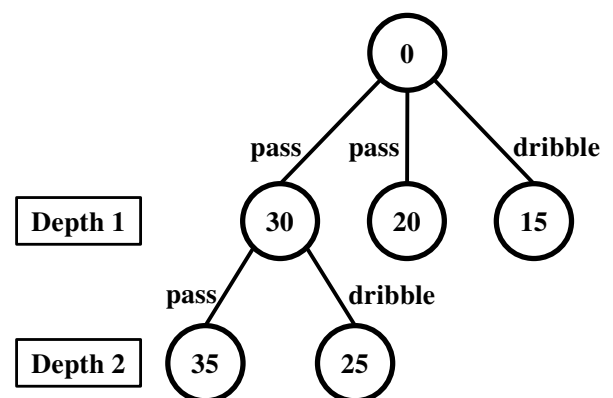


図 1: 最良優先探索に基づく行動列の探索例
Example of best first search in action planning

3.2 共有知

共有知とは、複数のプレイヤーで共有される知識のことである。従来では Kicker のみが持つ評価関数と生成する行動候補を、本論文では全プレイヤーの共有知とし、nonKicker が Kicker のプランニングを推測できるようにする。non-Kicker が Kicker と同様の行動候補を生成できれば、全プレイヤーで同じ行動プランが得られる。従って、Kicker のプランニング推測結果を使って、プレイヤー間の協調行動を獲得できる。評価関数と行動候補を共有知とすることで、試合中に生成される動的な行動プランを複数のプレイヤーが共有することを実現する。nonKicker に Kicker と同様の行動候補生成を可能にする共有知の実装を行い、数値実験によりその効果を検証する。

4 数値実験

4.1 実験設定

opuSCOM に共有知を実装した場合と実装しなかった場合に対して、opuSCOM と実力が近い 10 チームとそれぞれ 300 試合実施する。対戦チームは、2017 年の JapanOpen に出場した Esperanza, Toyosu-Galaxy, WIT の 3 チーム、世界大会に出場した HillStone, Fifty-Storms, ITAndroids, Persiangulf2017, Ri-one2017 の 5 チーム、agent2d[10] と共有知を実装前の opuSCOM の計 10 チームとした。本論文では、Kicker が nonKickerB へパスを行い、さらに nonKickerB が nonKickerA へパスをする、といった行動プラン“ Kicker (深さ 1)nonKickerB (深さ 2) nonKickerA ”を Kicker が意思決定した場合を対象とする。つまり、深さ 1、深さ 2 が共にパスの行動プランを研究対象とする。深さ 1 をパスとした提案理由として、ドリブルが実現されたかどうかの判断が難しいためである。ドリブルに対してパスは、ボールを保持するプレイヤーが移り変わるため、実現されたかの判断が容易であり、本実験を行う上で適していると考えた。

図 2 に実験の概要図を示す。図 2 では、破線がプランニングを、実線が nonKickerA の移動を示す。共有知により、nonKickerA が推測したプランニングが図 2 に示すようなプランニングの場合、矢印の先が示すボール受け取り予測位置へ nonKickerA が移動する。これによりプレイヤーが行動プラン通りに行動できるようになると考えた。本実験では、Kicker と nonKickerA のプランニング共有率、Kicker のプランニング実現率、試合結果から opuSCOM の性能の変化を調査する。プランニング共有率は、パスのみで構成された行動プランのうち、Kicker がプランニングを行った前後 10 サイクル以内に深さ 2 で受け取る nonKickerA が同様のプランニングを行った割合と定義する。実現率は、パスのみで構成された行動プランのうち、Kicker がプランニングを行った後、深さ 1 では 10 サイクル以内、深さ 2 では 20 サイクル以内に行動プランに基づ

くパス先のプレイヤーにボールが渡った割合と定義する。座標は考慮せず背番号のみで一致しているか判断する。また、nonKicker は Kicker の位置座標から行動プランの推測を行うため、ノイズを含む情報の認識ではプランニングの推測は困難である。実験では、共有知の有効性を調査するために、ノイズがなくプレイヤーが他のプレイヤーの位置座標を正確に認識できる fullstate を用いて実験を実施する。

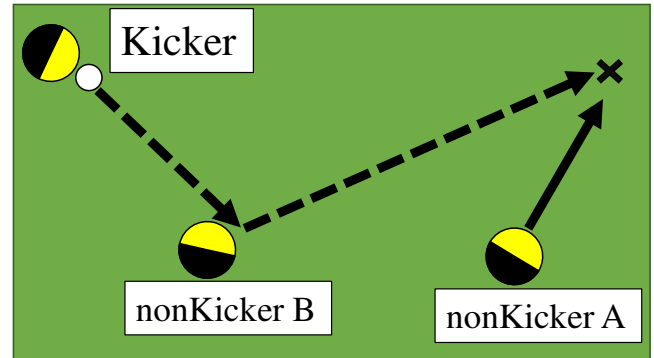


図 2: 実験の概要図
Experimental situation

4.2 実験結果

プランニング共有率、プランニング実現率、試合結果の調査結果を以下に示す。

4.2.1 プランニング共有率

共有知を実装後の opuSCOM と 300 試合行った各チームに対する、1 試合の平均プランニング回数 (#Planning), Kicker と nonKicker のプランニング共有率 (Sharing Rate) を表 1 に示す。表 1 より、どのチームに対しても 90% 程度の高い共有率を示した。このことから、共有知を用いることで nonKicker は Kicker のプランニングを推測できていた。

表 1: プランニング共有率

Planning sharing rates		
Opponents	#Planning	Sharing Rate(%)
agent2d	95	89.40
HillStone	89	89.64
Esperanza	84	90.16
Toyosu-Galaxy	30	92.20
WIT	51	89.49
Fifty-Storms	100	90.48
ITAndroids	84	89.81
PersianGulf2017	73	87.45
Ri-one2017	65	89.72
opuSCOM (Before)	114	90.29

4.2.2 プランニング実現率

共有知を実装後の opuSCOM と 300 試合行った各チームに対する、1 試合の平均プランニング回数 (#Planning) と全プランニングのうち深さ 1 の実現率 (D1)、深さ 1 が実現したプランニングのうちの深さ 2 のプランニングが実現した割合 (D2) を表 2 に示す。表 2 より、深さ 1 が実現したプランニングのうちの深さ 2 のプランニングが実現した割合は実装前後で 2 % ~ 11 % 上昇した。比率の差の検定により、有意水準 5 % で、共有知を実装前後のプランニング実現率に差があると言えた。このことから、共有知を用いたプランニング推測によって、深さ 2 時点でプランニング通りに行動することができた割合は上昇した。

表 2: プランニング実現率
Planning achievement rates

Opponents		#Planning	D1 (%)	D2 (%)
agent2d	Before	101	73.37	57.40
	After	95	73.39	64.25
HillStone	Before	94	71.92	54.52
	After	89	70.20	65.23
Esperanza	Before	89	72.49	60.05
	After	84	73.34	66.70
Toyosu-Galaxy	Before	29	74.46	40.96
	After	30	75.41	49.22
WIT	Before	51	74.06	59.90
	After	51	73.57	64.65
FiftyStorms	Before	104	67.80	55.34
	After	100	68.73	62.58
ITAndroids	Before	91	73.63	48.36
	After	84	73.95	54.97
PersianGulf2017	Before	71	70.29	56.54
	After	73	68.67	58.61
Ri-one2017	Before	67	73.71	48.84
	After	65	73.06	50.68
opuSCOM (Before)	Before	120	71.13	57.70
	After	114	72.43	65.10

4.2.3 試合結果

試合結果を表 3 に示す。項目は、勝ち (Win)、負け (Lose)、引き分け (Draw) の割合をそれぞれ示す。表 3 より、勝率に着目すると勝率が向上したチームもあるが、比率の差の検定により、有意水準 5 % で共有知を実装前後で勝率に差があるとは言えなかった。つまり、有意に勝率が向上したとは言えなかった。

4.3 考察

プランニング実現率が向上したにも関わらず、試合結果で勝率が向上しなかった。この原因として、敵陣では味方プレイヤーの数が少なくパス パスのプランニングが生成されにくいことが原因として挙げられる。図 3 は、agent2d との 1 試合の行動プランを全てプロットした図である。座

表 3: 試合結果

Game results

Opponents		Win (%)	Lose (%)	Draw (%)
agent2d	Before	59.33	29.67	11.00
	After	56.33	30.67	13.00
HillStone	Before	24.67	60.67	14.67
	After	24.00	64.67	11.33
Esperanza	Before	33.00	52.33	14.67
	After	33.67	47.33	19.00
Toyosu-Galaxy	Before	83.33	2.33	14.33
	After	81.33	5.67	13.00
WIT	Before	74.33	14.33	11.33
	After	79.33	10.33	10.33
FiftyStorms	Before	24.00	65.67	10.33
	After	23.67	64.33	12.00
ITAndroids	Before	5.00	86.33	8.67
	After	6.33	87.00	6.67
PersianGulf2017	Before	40.00	46.67	13.33
	After	41.67	48.00	10.33
Ri-one2017	Before	21.33	61.33	17.33
	After	20.67	65.67	13.67
opuSCOM (Before)	Before	37.33	48.00	14.67
	After	41.14	39.13	19.73

標はサッカーフィールドを表し、右側が敵陣、左側が自陣を示す。赤線が成功した行動プラン、青線が失敗した行動プランを示す。図 3 より、敵陣内での行動プランが少ないことが分かる。よって、研究対象とした行動プラン自体が敵陣内で少なく、プランニング実現率が上昇しても勝率への影響が少なかったことが考えられる。また、プランニング実現率は上昇したものの、未だ深さ 2 時点での値は高くないことも原因として考えられる。

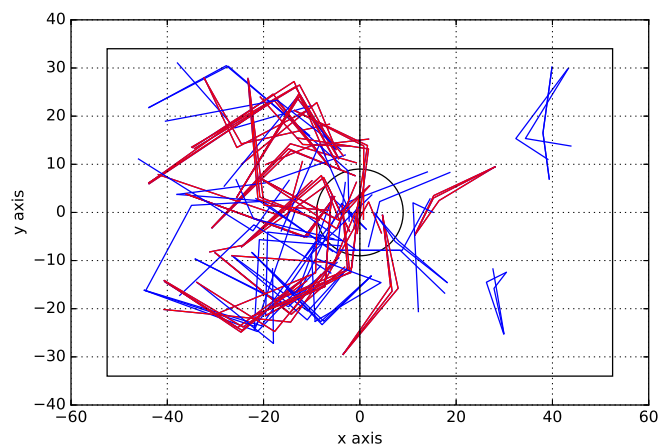


図 3: 1 試合の全行動プラン

All planning during the game

5 おわりに

本論文では、共有知を用いたプランニング推測に基づく行動選択を提案した。共有知によって nonKicker が Kicker のプランニングを推測することができた。これにより、non-Kicker がプランニングに従った行動を選択することができた。実験結果として、プランニング実現率は上昇したが、勝率の向上は見られなかった。今後の課題として、敵陣での攻撃的な行動生成の実装、コーナーキックなど特定の場面への限定化など共有知を活かした実装、または、プランニング実現率がより上昇するような工夫への取り組みが挙げられる。

参考文献

- [1] Hitoaki Kitano, Minoru Asada, Yasuo Kuniyoshi, Itsuki Noda, Eiichi Osawa and Hitoshi Matsubara, “RoboCup: A Challenge Problem for AI,” *AI Magazine*, Vol. 18, No. 1, pp. 73-85, 1997.
- [2] Martin Riedmiller, Roland Hafner, Sascha Lange and Marti Lauer, “Learning to dribble on a real robot by success and failure,” *Proc. of IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2207-2208, 2008.
- [3] Tomonari Nakade, Tomoharu Nakashima, Hidehisa Akiyama and Seki Hirosato, “Position Prediction of Opponent Players by SIRMs Fuzzy Models for RoboCup Soccer 2D Simulation,” *Proc. of 8th IEEE International Conference on Soft Computing and Intelligent Systems (SCIS) and 17th International Symposium on Advanced Intelligent Systems*, pp. 224-228, 2016.
- [4] 三船 哲史, 中島 智晴, 秋山 英久, Tomonari Nakade, Tomoharu Nakashima, Hidehisa Akiyama and Seki Hirosato, “RoboCup サッカーにおけるキック分布を用いたゲーム戦術の切り替え,” *Proc. of the JSAI Technical Report SIG-Challenge*, pp.1-6, 2016.
- [5] 秋山 英久, “ロボカップサッカー シミュレーション 2D リーグ必勝ガイド,” 秀和システム, 2006.
- [6] Tomoharu Nakashima, Satoshi Mifune, Jordan Henrio, Oliver Obst, Peter Wang and Mikhail Prokopenko, “Kick Extraction for Reducing Uncertainty in RoboCup Logs,” *Proc. of 17th International Conference on Human Interface and the Management of Information*, pp. 622-633, 2015.
- [7] Hidehisa Akiyama and Itsuki Noda, “Multi-Agent Positioning Mechanism in the Dynamic Environment,” *Proc. of the 12th RoboCup Symposium*, pp. 377-384, 2008.
- [8] 山下 雄大, 三船 哲史, 中島 智晴, 秋山 英久, “RoboCup に対する SIRMs を用いた位置予測モデルの開発,” 第 25 回ソフトサイエンス・ワークショップ講演論文集, 下関市立大学, 山口, 4 ページ, 3 月 10-11 日, 2015.
- [9] Hidehisa Akiyama, Shigeto Aramaki and Tomoharu Nakashima, “Online Cooperative Behavior Planning using a Tree Search Method in the RoboCup Soccer Simulation,” *Proc. of 4th IEEE International Conference on Intelligent Networking and Collaborative Systems(INCoS)*, pp. 170-177, 2012.
- [10] Hidehisa Akiyama and Tomoharu Nakashima, “Helios Base: An Open Source Package for the RoboCup Soccer 2D Simulation,” *Proc. of Robot Soccer World Cup XVII*, pp. 528-535, 2013.