

Empowerment as a Generic Utility Function for Agents in a Simple Team Sport Simulation

Marcus Clements(✉) and Daniel Polani

Adaptive Systems Research Group, School of Computer Science,
University of Hertfordshire, Hatfield, UK
m.clements@mxco.g.com, d.polani@herts.ac.uk

Abstract. Players in team sports cooperate in a coordinated manner to achieve common goals. Automated players in academic and commercial team sports simulations have traditionally been driven by complex externally motivated value functions with heuristics based on knowledge of game tactics and strategy. Empowerment is an information-theoretic measure of an agent’s potential to influence its environment, which has been shown to provide a useful intrinsic value function, without the need for external goals and motivation, for agents in single agent models. In this paper we expand on the concept of empowerment to propose the concept of *team empowerment* as an intrinsic, generic utility function for cooperating agents. We show that agents motivated by team empowerment exhibit recognizable team behaviors in a simple team sports simulation based on Ultimate Frisbee.

Keywords: Empowerment · Intrinsic Motivation · Artificial Intelligence · Information Theory

1 Introduction

Player behavior in team sports is a complex interplay of tactics and strategies that remains unpredictable and opaque despite the finely grained analysis of scientists, television pundits and millions of sports fans every weekend worldwide. Commercial computer sports simulations and RoboCup soccer teams adopt a “top-down” approach to analysis of the game, breaking down the strategic and tactical behaviors observed in team sports to construct highly complex ontologies of action and reaction[10]. Such ontologies can be used to devise tactics and strategies for human players, and with considerable effort, to devise heuristics for robot player actions. Humans can draw parallels with other sports, but a sophisticated model requires increased cognitive load[19], consuming a precious resource when decisions need to be made quickly. In the absence of explicit goals, or before an intricate system of goals and sub-goals has been developed through experience, humans and animals are nonetheless able take decisions and act, motivated by some internal value function. To mimic this phenomenon, Schmidhuber introduced “artificial curiosity”[20,21], a reward function based on the interestingness of the environment. Steels describes the “autotelic principle”[24], that proposes that agents may become self-motivated by balancing skills with challenges based on earlier work on “flow theory” by Csikszentmihalyi[7]. From dynamical system theory, *homeokinesis*[8] uses an adaptive model of a dynamic environment to predict actions. Subsequent work by Ay et al.[2] utilizes information theory to predict a future world state from a known history where, as with homeokinesis, the value of the predicted information is dependent on the depth of historical information. By adopting a “ground-up” perspective, in contrast to considering a complex set of game knowledge driven heuristics, we consider an *intrinsic motivation* measure based on *Empowerment*[16] for decision making in team game play that does not require prior tactical or strategic knowledge. Empowerment can provide a motivation for action in environments where little or nothing is known about the past, but it assumes that future options can be predicted from the current perceivable world state[18].

1.1 Empowerment

Empowerment is an information-theoretic measure of the *potential* of an agent to affect the world, using its actuators as detected by its sensors, and can be used as a generalized intrinsic utility function for individual agent behavior[16,14]. Salge et al. (2014)[18] suggest that “Empowerment provides a task-independent motivation that generates AI behavior which is beneficial for a range of goal-oriented behavior”. A colloquial expression for maximizing empowerment could be “keeping your options open”. If we consider the agent’s interaction with the world as a causal Bayesian network, we can discretize the agent’s actions as a series of iterations of the perception-action loop, in which the agent perceives the world, and acts on the received percepts during each iteration. Empowerment is formalized as the channel capacity[23] of the information flow between the actions of an agent, and the effect of those actions some time later[18]. We calculate the empowerment from the world state a given number of iterations in the future, referred to as *n-step* empowerment[18]. Previous studies have demonstrated the utility of empowerment in various settings, including in a simple maze[16], a grid world with a pushable box[16], and as a method of stabilizing a dynamic system in the form of a pole balancing on a cart[15].

1.2 Team Empowerment

Kelly (1994) defines the *superorganism* as “a collection of agents which can act in concert to produce phenomena governed by the collective”[13]. Duarte et al. proposed the concept of a sports team as a superorganism[9]. The team is considered to be a single entity, with its own intrinsic motivation, formed by the collective implicit exchange of information between team members, through movement and spatial positioning. Much of the existing work on empowerment has focused on individuals, however for multiple agents in the same environment, the agents actions each influence the world, and thus their empowerment is dependent on the actions of others. Capdepuy et al. demonstrated that structure can emerge in scenarios with multiple agents, all independently motivated by empowerment[5,4,6]. In this paper we propose an approach to the intrinsic motivation of a collective of cooperating agents, introducing the concept of *team empowerment* as a generalized utility function for the team superorganism, in a simplified team sport simulation, based on Ultimate Frisbee. We show that player agents, in teams motivated solely by empowerment, exhibit recognizable team sport behavior.

2 Model

2.1 Ultimate Frisbee

Ultimate Frisbee[26] is a field game with players in two teams of four to eight players throwing and catching a flying disc. The game is played on a rectangular field (see Fig. 1) with an area at each end, known as the end zones. For each point, play begins with the teams each in their own end zone. The team selects a player to start the point by throwing the disc downfield towards the opposition. As soon as a member of the opposing team takes possession of the disc, by catching or picking it up once landed, normal mid-game play begins. The player with the disc may not move, but will attempt to throw the disc (pass) to a teammate. Opposition players will attempt to block or intercept passes. If the disc is caught by an opponent, or touches the ground the opposing team gains possession. The aim of the game is to catch the disc in the opposing end zone, which scores one goal for the team in possession. Once a goal has been scored, the teams swap end zones and play begins again. The first team to reach an agreed number of points wins the game.

2.2 Basic Tactics

Players of Ultimate Frisbee adopt basic tactics of movement and possession of the disc that are common to other team games, notably Association Football (Soccer). Players try to gain possession by intercepting an opponents pass. Once in possession of the disc they try to pass to a teammate. Players on the same team as the player in possession “cut” (run) to try to find space on the field away from other players so that they are more likely to successfully receive a pass. Players on the opposing team try to intercept the pass, or make it more difficult for the player in possession by positioning themselves nearby and reducing the available throwing angles.

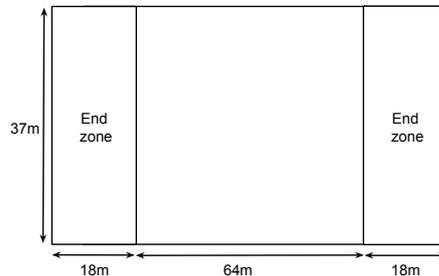


Fig. 1. Ultimate Frisbee field

2.3 Intrinsic Motivation

Our model considers a drastically simplified model of Ultimate, where nonetheless essential aspects of game play can be observed, requiring only a minimum set of rules. The beginning of the game, the scoring of a point, and the restarting of the game after a point is scored, involve special states and rules. For simplicity, just the mid-game is considered, where play has begun, and the disc is on the ground, or already held by a player. Scoring goals, and the subsequent restart of the game, are not simulated in our current model. The reason for this is that we wish to illustrate the operation of empowerment as an intrinsic motivation per se, while scoring goals introduces an external “motivation” in the form of a traditional utility. While the combination of the latter with the utility-less empowerment formalism is the topic of current research, in the present paper we focus on how far the pure, task-less empowerment formalism can already shape team behavior. As we will see, empowerment as an intrinsic motivation leads to the emergence of recognizable team sports behaviors including passing and interceptions, but in this experimental model, it will not lead to goals. There are many ways in which goals could be added to the model, but this would blur the contribution of empowerment itself, which can achieve several desirable behaviors without the concept of explicit goals.

2.4 Representation

The model consists of a 20 x 20 grid of locations with two teams of two players per team, compared to five or more players per team in the full setup. The grid represents the central part of the playing field, and does not include the end-zones. Only the mid-game is modeled, goals will not be scored, so end-zones and goal detection is not required. The model runs in turns, with each agent having an opportunity to act in each turn. Each action takes one turn to complete. The disc has four possible mutually exclusive states: Landed, Thrown, Flying, Held. Players have four possible states: Static, Moving, Catching or picking up the disc (cannot move), Throwing (can move). The available actions for each agent each turn are:

1. Move one square up, down, left or right, or stop.
2. Pick up the disc if it is on the ground in the same square.
3. Catch the disc if the grid square in which the agent is located is crossed by the trajectory of the flying disc.
4. Throw the disc to any square on the field within the configurable maximum throwing distance.

The model has the following constraints for the players based on the rules of the game:

1. Cannot move outside the field.
2. Cannot move onto the same square as another player.
3. Cannot move when holding the disc (as per Ultimate Frisbee rules).
4. Cannot take possession of the disc again after throwing until another player has thrown (as per Ultimate Frisbee rules).

Once the disc has been Thrown, in the next turn it will be Flying. The disc has a velocity configurable as the number of squares it will travel per turn once thrown. Once the disc has been flying for sufficient turns to reach its destination, it will be Landed unless it is caught by a Player while Flying in which case it is Held. As this is a two-dimensional model, the trajectory of the disc has no height, so it may be caught by a player at any point on its path. Fig. 2. shows the state diagram for players throwing and catching the disc. Each player (with number n) has a location vector (x_n, y_n) denoting a cell on the grid. The disc itself has a location vector (x_d, y_d) , and a destination vector (x_{dest}, y_{dest}) which is only relevant when it is in the Thrown or Flying state. The state of the model (S_m) in any given turn can be represented by the locations and state (s_p) of the players, and the location, state (s_d), and destination of the disc.

$$S_m := D[x_d, y_d, s_d, x_{dest}, y_{dest}], P_0[x_0, y_0, s_0], \dots, P_n[x_n, y_n, s_n] \quad (1)$$

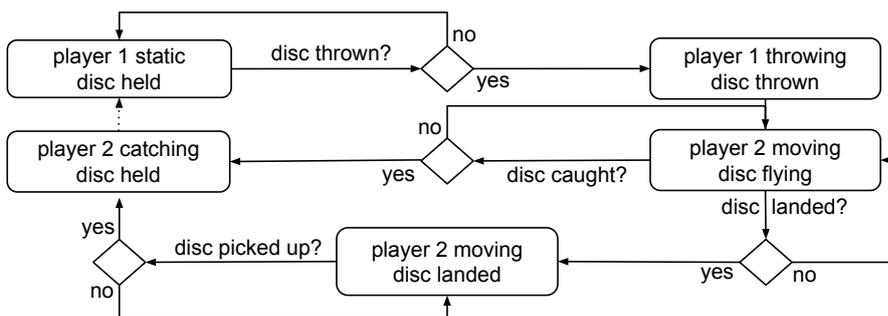


Fig. 2. State transitions for players and disc during throwing and catching action sequences.

2.5 Player precedence for catching and collisions

Players may not occupy the same square in the grid. During each turn the list of players is reordered according to a pre-determined, pseudo-random sequence where player positions in the list are evenly distributed over many iterations. Potential collisions are avoided because players take action in list order. If more than one player is located on the trajectory of the disc in a given turn, the closest to the disc origin may catch the disc.

3 Calculating Empowerment

In the general case, empowerment is computed via the channel capacity of the action-perception loop. This case captures partial information, stochastic and both continuous and discrete-state models. In our case, however, we will assume a discrete, deterministic, full-information model or

convert it into such, where necessary, drastically simplifying the computation of empowerment, while retaining the expressivity of the phenomena. This simplifies the empowerment calculation to computing the logarithm of the number of distinct states reachable in n steps, which will be measured in bits, since we will use the binary logarithm. For the derivation of the n -step empowerment calculation from channel capacity please refer to [16]. As empowerment is essentially reduced to counting the end states in our deterministic scenario, the logarithm is not strictly necessary here; however, we retain it in the formula, permitting expression of empowerment in bits and keeping the present results conceptually and quantitatively consistent with the formalism generalized to stochastic and partial information scenarios.

$$\mathfrak{E} = \log_2 |S_n| \text{ bits} \quad (2)$$

3.1 Empowerment from Movement

If the agent is in the middle of the field, far from the disc or any other players, the agent may move up, down, left or right, or stay still. For each of those actions there are the same five possible actions. For 1 step in the middle of the field, the Empowerment for each possible future action, as shown in Fig. 3., is given by $\mathfrak{E} = \log_2 5 = 2.32 \text{ bits}$. After 3 steps from a location in the middle of the field there are 25 possible end states, including not moving at all, as shown in Fig. 4. Considering three movement steps into the future, the empowerment of the agent is given by $\mathfrak{E} = \log_2 25 = 4.64 \text{ bits}$.

3.2 Empowerment from Throwing the Disc

An agent holding the disc may throw it to any other location within the maximum throwing distance. Given that the field has 20×20 locations, and a large maximum throwing distance, there are 399 potential resulting states for the disc as a result of the throw, thus the empowerment of the agent holding the disc is given by $\mathfrak{E} = \log_2 399 = 8.64 \text{ bits}$. Note that empowerment itself just assigns an intrinsic (pseudo-)utility value to the states under consideration. In the present model, this pseudo-utility induces behavior in the same way a traditional utility function would do; namely, in a given state, one chooses the action that changes the world state into a world state where the present agent has the largest value of empowerment (here: reachable states in the given lookahead horizon)¹. The deterministic, full-information case, allows us to even select this action directly from the lookahead tree propagating from the current state (see below for an illustration).

Example 1. Agent in cell adjacent to disc

As a more detailed example of determining the possible future states for an agent, let us consider a typical situation. The disc has landed in the adjacent square to an agent, and we are considering the empowerment of the possible next states given a three-step lookahead horizon. The agent has all the usual movement options at each step, which for two steps leads to 25 possible end states. If the agent steps onto the location of the disc, it has one additional option - to pick up the disc. Once the agent has the disc it may throw it to one of 399 possible locations. There are three groups of three-step action sequences available to the agent:

1. Move \rightarrow Move \rightarrow Move (25 end states)

¹ Note that this is not a utility in the conventional sense for reasons that require some subtle discussion which is beyond the scope of the present paper. For the present case, this distinction is immaterial.

2. Move → Move → Pick up disc (1 end state)
3. Move → Pick up → Throw disc (399 end states)

Moving onto the square with the disc in the first turn is the agent’s future state with the highest empowerment as none of the other possible future states can lead to throwing the disc, therefore that is the action the agent will take. Fig. 5 illustrates the possible end states for a player agent adjacent to the disc.

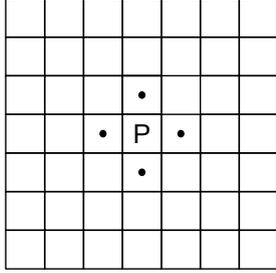


Fig. 3. Possible end states after one step. If the player at P does not move at all that is one possible end state. The dots show the reachable locations.

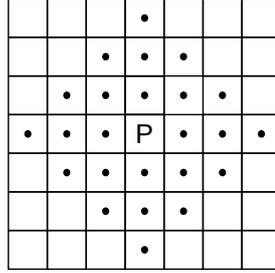


Fig. 4. Possible end states after three steps. If the player at P does not move at all that is one possible end state.

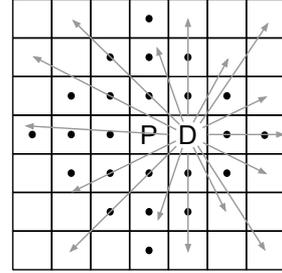


Fig. 5. Possible end states after three steps for player agent starting at P adjacent to disc D. The arrows show a subset of the possible disc trajectories.

3.3 Team Empowerment

All agents on the same team are considered as a single super agent (from superorganism) for calculating team empowerment. For Player 1 and Player 2, on the same team, when we calculate the empowerment for the resulting state of an action of Player 1, we include the states of Player 2 because the players are able to affect each other’s behavior, and so in catching and throwing the disc their states count towards the reachable states². The possible future states of agents on the opposing team however, are not considered in the reachable state calculation. Players in Team A, do not have a forward model of the behavior of players in Team B, they only have a model of the current state of the players (their locations). In the implementation of the model, when constructing the search tree of future states for a player in Team A, we leave the states of players in Team B unchanged from step to step. In the absence of information about the behavior of the opposition, the agents assume that opposition players take no action.

3.4 Implementation of the Model

The Java implementation of the model employs an object-oriented architecture to represent the model state as an acyclic directed graph. Each vertex in the graph represents the state of the simulation, and each edge in the graph represents an action. At each time step the graph is constructed by a recursive function. After n steps (for n -step empowerment) the number of end states is counted for each action. The choice of action to take from the current state is made

² It is possible to consider other friendly agents in a more general way, as still different from the egocentric agent, but nonetheless cooperative. This is research in progress.

simply according to which action (edge) will lead to the most future options (terminal vertices). The findings were captured into testable scenarios, providing initial conditions and assertable outcomes, which are repeatable because the model is deterministic. A graphical representation of the model is presented in the user interface, showing the changing position of the players and the disc (Figs. 6 and 7). The implementation provides the most basic environment for the expression of empowerment-driven behavior. Considerable further work is required to develop the team empowerment formalism to a more realistic stochastic, continuous environment before optimizing the model for real-time performance.

4 Experiments and Results

We employed the empowerment principle in a number of scenarios, typical of Ultimate Frisbee, to investigate how far it, by itself, and without further externally specified utilities, would drive the agents to take tactically plausible actions, based on the rules of the game[26]. After observation of these actions³, each result was confirmed by creating a repeatable test scenario inspecting the graph of states and actions generated during the empowerment calculation.

4.1 Finding Space

When not holding the disc or catching, an agent will move away from the sides of the field. At the sides the empowerment is lower because the future range of states available by move actions is restricted. An agent will also move away from another agent for the same reason. The distance the agents maintain from the sides, or each other, increases as the lookahead increases. “Finding space” is a common basic tactic in Ultimate and other team games, particularly Soccer.

4.2 Retrieving the Disc

An agent that is within the lookahead number of steps of a disc on the ground, will move directly to the disc and pick it up. Being in possession of the disc allows the throwing actions and so has a drastically higher empowerment than leaving the disc on the ground. Retrieving the disc, once it has hit the ground after a wayward pass, is an essential behavior of in Ultimate[26], and equivalent to collecting the ball after an inaccurate pass in Soccer.

4.3 Passing

With a three-step lookahead (or higher), an agent in possession of the disc will throw the disc to a location where a teammate may catch it. Throwing the disc to a teammate is the action leading to a successor state with the highest empowerment, because the future states of the teammate are taken into account in the calculation. In the subsequent turn, the disc will be in the *Flying* state with a trajectory crossing the cell containing the teammate, and thus catchable. For the teammate, being in possession of the disc has a higher empowerment than the disc landing on the ground, so they will catch it. After catching the disc, an agent must wait until the next turn to throw. While throwing the disc, an agent may also move, but while catching they may not. During the experiments, agents on the same team repeatedly throw (“pass”) the disc to each other, moving when not catching. Passing is an essential behavior in Ultimate[26] and other team sports including Soccer.

³ Video captures of the simulation can be viewed at: http://mxkog.com/research/team_empowerment

4.4 Interceptions

An agent located on the trajectory of the disc can catch the disc. Note that, to compute empowerment, agents need to have a forward model of what will ensue from taking particular actions, including the destination of the disc once thrown. During the empowerment calculation, the tree of possible future states will include catching the disc if the agent is on the trajectory of the disc. Since being in possession of the disc has a higher empowerment than not, the agent will catch the disc. Catching the disc, when a player on the opposing team has thrown it, is known as an “interception”, and is an important behavior in Ultimate[26] and other team sports. Figures 6 and 7 show an attempted pass and subsequent interception taking place in the model.

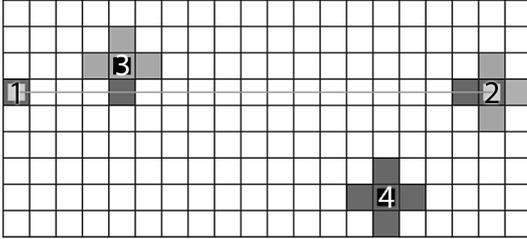


Fig. 6. Partial screen capture of the model running with a 3-step empowerment lookahead and a disc velocity of 20. Players 1 and 2 are on Team A, players 3 and 4 are on Team B. The grey line between the players shows the trajectory of the disc as thrown by Player 1 to Player 2. The darkness of the grey shading indicates the relative empowerment of that player's possible subsequent states.

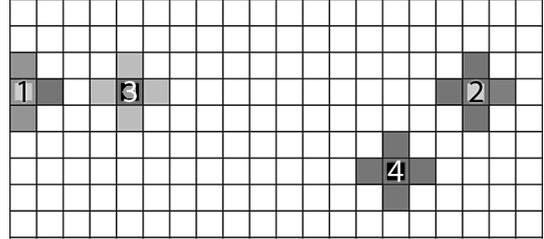


Fig. 7. Partial screen capture of the model one time step after the screen capture in Fig. 6. Player 1 can now move and will have a slightly higher empowerment moving right because the edge of the field constrains movement options. For Player 3 the highest empowerment state results from catching, so it has moved down and is now in possession of the disc.

5 Discussion

We have shown that *team empowerment* as a generic utility function for agents in a sports simulation leads to emergent team game behavior in the form of four basic tactics: *finding space*, *retrieving the disc*, *passing* and *interceptions*. By the rules of Ultimate[26], if no player on the same team catches a thrown disc, possession passes to the opposition. A player intercepting a pass between opponents gains possession of the disc, increasing their empowerment. Thus catching the disc is a better strategy than not catching, and consequently stepping into the trajectory of the disc is a better strategy than not doing so. Moving away from opposition players makes a successful pass more likely by reducing the opportunity for an opponent to intercept. A human playing Ultimate Frisbee, with no prior knowledge of tactics or strategy, could employ a pseudo-empowerment strategy by acting to maximize future options. Our results indicate that a player utilizing such a strategy, would be a much more useful member of the team from the outset than a player merely choosing actions at random from the options allowed by the game. The parameters of the model automatically favor capturing and subsequently throwing the disc because throwing options are more numerous than moving options. It therefore does not seem surprising that these action sequences emerge from the empowerment dynamics. However, this is emphatically intended — empowerment is expressly a quantity that emerges from the interaction of agents with their environment, and this behavior emerges from it “naturally”,

rather than being externally imposed on the agents. We remind the reader that the present experiment is inspired by the box-pushing scenario from [16] where a manipulable box on a grid, merely by virtue of its affordance, creates an empowerment “attractor” around the box. This can be seen as a drastic abstraction of a “ball” attracting an agent by pure virtue of being able to be kicked around; in a way, an agent detecting that “playing football” with a box is “fun” [22]. For this to work, it is unavoidable for generic measures for intrinsic motivation to make heavy use of the environment. However, previous studies show that empowerment produces plausible and human-like behaviors in a variety of scenarios, including balancing tasks, puzzles, obstacle avoidance and more [18], with its completely generic approach, not utilizing a task-specific utility (and emphatically not trying to be somehow “best” according to some specialized criterion, but rather “reasonable” in many generic scenarios). This indicates that the environment itself is often sufficient to induce naturally expected behaviors in many cases from the outset, and that utilities typically imposed on such scenarios may just capture properties which are implicitly “encoded” in the scenario itself. Consequently, one tentative hypothesis is that humans would utilize principles which are similar in spirit to cope with the large regions of their task space without the guidance of dedicated and detailed utility functions [18]. Thus, our proof-of-principle empowerment-based studies in the present paper suggest a concrete and operational approach of how to address the generic interaction of an agent team in a ball-like game, in an intrinsic and natural way; this would happen in advance of specifying a detailed, problem-specific utility function, endowing first the agents with a set of “common sense” baseline behaviors, that could later be refined by a more specific reinforcement learning approach. Extensions of the empowerment formalism to the team level could take into account hierarchization techniques such as *Layered Learning* [25], by which “ritualized action sequences” (which could be considered an empowerment-induced analogy of the Reinforcement Learning concept of options [3]) could be compacted, and utilized as meta-actions, to serve for the empowerment computation at a higher level. Basic empowerment, as implemented here, will select the states with most options and not take account of any future development. To look beyond the empowerment horizon, one could adopt the “soft horizon” method from [1] which augments the empowerment computation by incorporating a “halo” of futures corresponding to the states reached by the selected action sequences. However, we have shown that already with basic empowerment the resulting strategies are capturing many of the natural “first approximations” to tactically meaningful behavior.

6 Future Work

We hypothesize that with a more sophisticated simulation, further recognizable Ultimate Frisbee tactics would emerge such as blocking passes and marking (restricting options) of opponents. There are many similarities in game play and tactics between Ultimate Frisbee and Soccer, and we anticipate similar results using team empowerment in a comparable model based on Soccer. Furthermore we hypothesize that, when combined with reinforcement learning, empowerment as an underlying strategy could lead to the emergence of longer term strategies such as formations and movement in anticipation of expected actions by other players on the team. For strategies such as this to emerge, it is necessary to master the complexity of more general computations of empowerment for the continuous and stochastic case, and for the case of imperfect team members, and actively antagonistic players. While a number of empowerment estimation approaches for continuous/stochastic scenarios in the single-agent case have been developed [11,12,17] the team aspect sketched above requires considerable extension, improvement and scaling up of the formalism. This will be the focus of future work, to harness the opportunities that the empowerment formalism offers to produce “intuitive” baseline team behavior.

Acknowledgements DP would like to acknowledge support by the EC H2020-641321 socSMCs FET Proactive 713 project, the H2020-645141 WiMUST ICT-23-2014 Robotics project.

References

1. Anthony, T., Polani, D., Nehaniv, C.L.: General self-motivation and strategy identification: Case studies based on sokoban and pac-man. *IEEE Trans. Comput. Intellig. and AI in Games* pp. 1–17 (2014)
2. Ay, N., Bertschinger, N., Der, R., Güttler, F., Olbrich, E.: Predictive information and explorative behavior of autonomous robots. *The European Physical Journal B-Condensed Matter and Complex Systems* 63(3), 329–339 (2008)
3. Barto, A.G., Mahadevan, S.: Recent advances in hierarchical reinforcement learning. *Discrete Event Dynamic Systems* 13(4), 341–379 (2003)
4. Capdepuy, P.: Informational Principles of Perception-Action Loops and Collective behaviors. Ph.D. thesis, University of Hertfordshire (2010)
5. Capdepuy, P., Polani, D., Nehaniv, C.: Maximization of potential information flow as a universal utility for collective behavior. In: *Artificial Life, IEEE Symposium on*. pp. 207–213. IEEE (2007)
6. Capdepuy, P., Polani, D., Nehaniv, C.L.: Perception-action loops of multiple agents: informational aspects and the impact of coordination. *Theory in Biosciences* 131(3), 149–159 (2012)
7. Csikszentmihalyi, M.: *Beyond boredom and anxiety*. Jossey-Bass (2000)
8. Der, R., Steinmetz, U., Pasemann, F.: Homeokinesis: A new principle to back up evolution with learning. *Max-Planck-Inst. für Mathematik in den Naturwiss* (1999)
9. Duarte, R., Duarte, A., Correia, V., Davids, K.: Sports teams as superorganisms. *Sports Med* 42(8), 633–42 (2012)
10. Dylla, F., Ferrein, A., Lakemeyer, G., Murray, J., Obst, O., Röfer, T., Schiffer, S., Stolzenburg, F., Visser, U., Wagner, T.: Approaching a formal soccer theory from the behavior specification in robotic soccer. In: Dabnichki, P., Baca, A. (eds.) *Computers in Sport*, pp. 161–186. WIT Press, London (2008)
11. Jung, T., Polani, D., Stone, P.: Empowerment for continuous agent-environment systems. *Adaptive behavior* 19(1), 16–39 (January 2011)
12. Karl, M., Bayer, J., van der Smagt, P.: Efficient empowerment. preprint (2015)
13. Kelly, K.: *Out of control: the new biology of machines, social systems and the economic world*. Perseus, Cambridge, Mass. (1994)
14. Klyubin, A., Polani, D., Nehaniv, C.: Empowerment: A universal agent-centric measure of control. *Evolutionary Computation, The 2005 IEEE Congress on*, 1, 128–135 (2005)
15. Klyubin, A., Polani, D., Nehaniv, C.: Keep your options open: an information-based driving principle for sensorimotor systems. *PloS ONE* 3, 12 (2008)
16. Klyubin, A.S., Polani, D., Nehaniv, C.L.: All else being equal be empowered. In: *Advances in Artificial Life*. pp. 744–53. Springer (2005)
17. Mohamed, S., Rezende, D.J.: Variational Information Maximisation for Intrinsically Motivated Reinforcement Learning. In: *Neural Information Processing Systems (NIPS) 2015* (2015)
18. Salge, C., Glackin, C., Polani, D.: Empowerment – an introduction. In: Prokopenko, M. (ed.) *Guided Self-Organization: Inception*. Springer (2014)
19. Samuelson, L.: Analogies, adaptation, and anomalies. *J. of Economic Theory* 97(2), 320–366 (2001)
20. Schmidhuber, J.: Curious model-building control systems. In: International, I. (ed.) *Joint Conference on Neural Networks*. pp. 1458–1463. IEEE (1991)
21. Schmidhuber, J.: Exploring the predictable. *Advances in evolutionary computing* 6, 579–612 (2002)
22. Schmidhuber, J.: Formal theory of creativity, fun, and intrinsic motivation (1990-2010). *Autonomous Mental Development, IEEE Transactions on* 2(3), 230–247 (2010)
23. Shannon, C.E.: A mathematical theory of communication. *Bell Sys. Tech. Journal* 27, 623–656 (1948)
24. Steels, L.: The autotelic principle. In: *Embodied Artificial Intelligence: International Seminar, Dagstuhl Castle, Germany, July 7-11, 2003. Revised Papers*, pp. 231–242 (2004)
25. Stone, P.: Layered learning in multiagent systems: A winning approach to robotic soccer. MIT (1998)
26. World Flying Disc Federation Ultimate Rules Committee: *WFDF Rules of Ultimate 2017* (2017), http://www.wfdf.org/files/WFDF_Rules_of_Ultimate_2017_-_FINAL_-_31_Dec.pdf