

UMA ABORDAGEM MULTIAGENTES PARA SISTEMA DEFENSIVO EM UM TIME DE FUTEBOL DE ROBÔS BÍPEDES

Camila Laranjeira
mila.laranjeira@gmail.com

Alan dos Santos Soares
alandossantossoares@gmail.com

Emmanuel Marmund Argollo
enmanueru3553@gmail.com

Diego Frias, Marco A. C. Simões
diegofriass@gmail.com

Marco A C Simões
marco.c.simoess@gmail.com

Josemar Rodrigues de Souza
josemarsbr@gmail.com

Núcleo de Arquitetura de Computadores e Sistemas Operacionais, Universidade do Estado da Bahia, Brasil

Abstract: This article describes the defense module developed for an autonomous robot soccer team. Believing artificial intelligence is capable of overcoming limitations of movement such as slow walking, a new behaviour state was created and added to the decision making logic of our agents: a player marker focused on the opponent with current possession of the ball, defensively supporting the ally currently active. This module was developed during the 2014 edition of RoboCup, where it was initially validated and, later on, automatic tests were performed showing an improvement not only on the team's defense, but also in its overall performance. A few variables, such as ball possession, game concentration zone, among others were measured to estimate the impacts of adopting this module.

Keywords: Robot soccer; Artificial intelligence; Multi-agent systems; Cooperation.

Resumo: Este artigo descreve o módulo de defesa cooperativa desenvolvido para um time autônomo de futebol de robôs. Acreditando que a inteligência artificial é capaz de superar limitações da movimentação, foi criado um novo estado comportamental para compor a tomada de decisão dos agentes: uma marcação focada no oponente com atual posse da bola, dando suporte de defesa ao jogador aliado em estado ativo. Este módulo foi desenvolvido durante a edição de 2014 da RoboCup, tendo sido inicialmente validado ainda em competição, e posteriormente através de testes automatizados, através dos quais foi possível perceber uma melhora não só na defesa do time, mas também em seu desempenho geral. Uma série de indicadores como posse de bola, zona de concentração de jogo entre outros foram mensurados para aferir os impactos da adoção deste módulo.

Palavras-chave: Futebol de robôs; Inteligência artificial; Sistemas multiagente; Cooperação.

I. INTRODUÇÃO

A RoboCup [6], fundada em 1997, é uma iniciativa para fomentar a pesquisa em robótica e inteligência artificial. Desde o seu surgimento, ela evoluiu de muitas formas, mas seu objetivo principal ainda se mantém: em meados do século XXI formar um time de robôs totalmente autônomos capaz de derrotar a mais recente seleção campeã da FIFA. Na busca pelo seu objetivo, ela se dividiu em diferentes categorias e ligas, continuando a crescer a cada ano. O desafio de futebol de robôs apresenta alto nível de complexidade reunindo grande parte dos desafios presentes em problemas reais. Decisões complexas precisam ser tomadas pelos agentes jogadores de futebol em tempo real de forma a atingir um objetivo coletivo maior. Para explorar as diversas disciplinas que compõem este desafio, a RoboCup o dividiu em cinco ligas. A liga de robôs de pequeno porte (*Small Size*) é formada por times com cinco robôs cilíndricos movidos por rodas omnidirecionais. A maior parte dos times utiliza uma visão global através de uma câmera suspensa sobre o campo que captura e informa a posição da bola, das traves e de todos os robôs em campo. Outra possibilidade é o uso da visão parcial com uma câmera embarcada no robô. Neste caso apenas os objetos no ângulo de visão da câmera terão sua posição conhecida. A visão parcial também é mais ruidosa uma vez que é móvel e captura objetos com distância e ângulos variados. Nesta liga o uso da visão parcial é opcional. Os robôs utilizam uma bola de golfe e possuem tamanho limitado a 180 mm. Esta liga desenvolve características da eletromecânica dos robôs

permitindo o desenvolvimento de veículos móveis rápidos e omnidirecionais. Também são pesquisados desafios na área de algoritmos, inteligência artificial e sistemas multiagentes. Trabalhos na área de visão computacional também são desenvolvidos e testados neste ambiente. A liga de robôs de médio porte (*Middle Size*) é baseada em robôs maiores que demandam maiores estudos em eletromecânica, baterias, etc. Neta liga a visão parcial é obrigatória e o campo e bola têm dimensões próximas a de um campo de futebol de salão. O ambiente desta liga é o que mais se aproxima do ambiente real de um jogo de futebol. Também estão presentes os desafios de algoritmos e inteligência artificial. Na liga de robôs humanoides, os robôs com rodas são substituídos por robôs bípedes com anatomia semelhante à de um ser humano. Existem competições com robôs pequenos (*kid size*), médios (*teen size*) e grandes (*adult size*). A visão é parcial e a construção dos robôs em si já é um grande desafio devido à necessidade de uso de vários motores independentes para suportar os diversos graus de liberdade presentes. O controle também é muito mais complexo que nos robôs com rodas. Apesar dos times serem menores (1 a 5 robôs – a depender do tamanho dos mesmos) os desafios em algoritmos e inteligência artificial também estão presentes. A liga plataforma padrão (*Standard Platform*) também utiliza robôs humanoides mas utiliza um modelo de robô de mercado, de forma que todos os times têm o mesmo hardware e o desafio fica a cargo apenas do software de controle e inteligência dos robôs. Por fim, na liga de simulação é possível realizar experimentos com robôs simulados, abstraindo a complexidade presente na parte eletromecânica, de controle e visão computacional, o foco principal da simulação é a inteligência artificial. A Simulação se divide em duas competições: 2D e 3D. Na 2D, o ambiente simulado assemelha-se ao de robôs de pequeno ou médio porte. A simulação considera robôs que deslizam pelo campo de forma omnidirecional, abstraindo totalmente o problema da locomoção. A cooperação entre os agentes e inteligência de alto nível são o grande desafio nesta liga. Abordagens avançadas de aprendizagem por reforço e redes neurais são usadas pelos times para desenvolver a inteligência coletiva avançada. O conhecimento produzido nesta liga ainda não consegue – em sua maioria – ser utilizado nas ligas com robôs físicos devido às limitações do hardware e do controle. Na liga de simulação 3D, a física do mundo real é simulada em um ambiente tridimensional, representando o cenário de um jogo de futebol onze contra onze de agentes humanoides. Essa liga tenta minimizar as dificuldades do mundo real, mas mantém o desafio de desenvolver movimentos similares aos humanos, bem como estratégias de inteligência artificial cooperativas. Na 3D, os agentes se aproximam mais dos robôs reais, pois, apesar de não haver desafios de hardware, as dificuldades do controle dinâmico de diversos graus de liberdade estão presentes, sendo assim, muito do que é produzido de conhecimento nesta liga será aplicado nos robôs reais do futuro. A liga de simulação se encontra à frente das

ligas físicas em termos de inteligência artificial e software.

A equipe BahiaRT é o berço de testes das pesquisas do Núcleo de Arquitetura de Computadores e Sistemas Operacionais (ACSO) da Universidade do Estado da Bahia. Através das competições científicas, os pesquisadores do ACSO validam seus projetos em robótica e inteligência artificial. No desafio de futebol de robôs, a equipe BahiaRT vem se destacando nos últimos anos na liga de simulação 3D com foco principal em sistemas multiagentes e inteligência artificial.

A inteligência artificial cooperativa é um desafio moderno encontrado nos mais diferentes cenários, e especificamente no contexto do jogo de futebol a cooperação se torna um forte diferencial competitivo. Enquanto a maioria dos times na liga 3D foca no desenvolvimento de movimentos mais eficazes e robustos, outras equipes acreditam que a inteligência artificial é capaz de superar limitações de movimentação. Como consequência de manter o foco na inteligência, o movimento de andar do BahiaRT é lento se comparado aos campeões dos últimos anos, fazendo com que recuperar a posse da bola durante o jogo seja uma tarefa muito mais difícil. Por este motivo, este trabalho foi desenvolvido com o objetivo de desenvolver um módulo de cooperação multiagentes visando recuperar a posse de bola quando a equipe BahiaRT não a possui.

Vale salientar que, mesmo após uma extensa revisão da literatura existente, não foi encontrado nenhum trabalho voltado para o desenvolvimento de mecanismos de defesa cooperativa em um jogo de futebol, cuja aplicação estivesse focada em robôs bípedes. Os autores acreditam que este estudo é de suma importância para incentivar futuros trabalhos dessa natureza, destacando a importância de aumentar o nível das estratégias implementadas para agentes humanoides.

Entre as consequências esperadas do uso desta estratégia, estão a redução dos gols concedidos, maximização dos gols marcados, e melhoria do desempenho global do time. O sucesso deste trabalho representa a possibilidade de robôs autônomos tomarem decisões complexas, de forma cooperativa, em um espaço muito curto de tempo. Isto pode ser aplicado em uma gama de aplicações que demandam este comportamento, como veículos autônomos, controle de tráfego urbano e aéreo, robôs de serviços, entre outros.

Para tratar esse tema, o trabalho foi dividido em 8 seções. A seção II apresenta uma breve revisão da literatura referenciando os principais trabalhos relacionados. Na seção III é descrita a inteligência comportamental responsável pela definição de papéis dos agentes do BahiaRT. A seção IV conta com uma descrição da lógica por trás da solução proposta. A seção V detalha a metodologia de validação da solução. Na seção VI são apresentados os resultados do método aplicado, seguida da seção VII, onde são

apresentadas as conclusões do estudo, e por fim na seção VIII são apresentadas as propostas de trabalhos futuros.

II. TRABALHOS RELACIONADOS

Existem muitas abordagens para definição de funções ou papéis em sistemas multi-agentes [5], [6], [9]. Estas definições permitem que atividades possam ser desempenhadas por um ou mais agentes para alcançar os resultados esperados.

Na abordagem proposta por [8], um agente com papel de coordenador é definido para coordenar uma defesa de forma adaptativa. Esta coordenação visa identificar possíveis ameaças e problemas que possam comprometer a segurança da sociedade. De acordo com os problemas, o coordenador verifica quais ações são necessárias para reparar ou proteger a sociedade de possíveis danos.

No futebol de robôs, agentes individuais tendem a encontrar políticas adequadas às funções específicas dos agentes, como apresentado em [7]. Ao contrário do agente coordenador, agentes individuais tomam suas próprias decisões baseando-se na sua experiência e no comportamento de seus aliados. Como demonstrado em [11], um agente que escolhe suas próprias estratégias baseando-se nas expectativas de outros agentes resulta em um processo cooperativo baseado na antecipação de ações. Uma expectativa é definida por uma função que mapeia o comportamento de um agente em função de suas interações.

Várias abordagens foram propostas para a área de jogos. A coordenação de um único estágio tem sido usada como um modelo no estudo de sistemas multi-agente [3] [4]. Este tipo de abordagem se baseia na definição de um conjunto de ações individuais que são avaliadas em função do custo de execução.

Outras abordagens utilizam aprendizagem por reforço [4] para garantir que os agentes aprendam a escolher as melhores ações de acordo com a situação [10]. Estas abordagens não utilizam um agente coordenador, portanto agentes individuais tomam suas próprias decisões. No entanto, este tipo de abordagem requer que o agente execute as n ações até encontrar a melhor ação que resolva o problema, sendo um processo que requer tempo.

Os artigos [3] e [2] fazem uma discussão sobre papéis sociais em sistemas multi-agente. Em [4] é feita uma revisão geral sobre questões e abordagens do problema de alocação de papéis em sistemas multi-agente.

III. INTELIGÊNCIA COMPORTAMENTAL DOS AGENTES DO TIME BAHIA RT

No início da Robocup 2014, realizada em João Pessoa, PB, Brasil, os agentes do time BahiaRT possuíam 3 estados comportamentais ou papéis: (1) limitado, (2) cooperativo e (3) ativo. Cada agente a partir da análise do estado, direção e posição dele, da posição dos colegas, dos oponentes e da bola, no campo, determina seu rol. Cada papel tem um objetivo a curto prazo definido.

O agente se coloca no papel limitado quando não está totalmente em pé devido a uma queda ou se encontra fora do campo. Neste estado o objetivo a curto prazo é recuperar sua capacidade para participar do jogo, seja se levantando ou voltando ao campo.

Se não estiver limitado, o agente assume o papel ativo quando é o jogador com a maior chance de tomar ou manter o controle da bola. Na tática atual, apenas um jogador pode assumir o papel ativo, aquele com o menor tempo para chegar até a bola ($T2B$). Para isto, cada agente calcula seu $T2B$ levando em conta sua distância da bola, sua velocidade, e o tempo necessário para girar caso não se encontre de frente para a bola. O tempo de giro é estimado levando em conta o ângulo de giro necessário e a velocidade angular do movimento de girar. O cálculo do $T2B$ é repetido para os 3 agentes aliados mais próximos da bola, mas sem levar em conta o tempo de giro e assumindo a velocidade máxima dos agentes do time. Se o $T2B$ do agente for o menor de todos, então o agente assume (ou mantém) o papel ativo. O fato de não considerar nem a direção nem a velocidade real dos aliados mais próximos à bola, dá uma vantagem aos agentes aliados, evitando auto-superestimação do agente, isto é, o agente assume o papel ativo apenas no caso em que ele é realmente o mais habilitado para isso. Desta forma se evitam situações em que dois agentes assumem o papel ativo causando interferência indesejada na jogada. Vale destacar, que este procedimento, por outro lado, cria situações em que nenhum jogador toma o papel ativo ficando o time todo inativo. Mesmo que essa situação seja momentânea devido ao dinamismo do jogo, uma nova estratégia está em desenvolvimento para evita-la. Veja mais detalhes na seção V.

Por último, o agente assume o papel cooperativo quando não está limitado nem é o agente ativo. Neste papel o agente tem a missão de se posicionar no campo seguindo o fluxo determinado pela formação dinâmica do time. A posição de referência de cada agente no campo depende da posição que ocupa na formação inicial e da posição atual da bola [9].

Na primeira fase da Robocup 2014 confirmamos um fato observado nos testes pré-competição, de que a mobilidade de nossos jogadores era inferior à dos melhores times adversários. Muitas vezes nosso jogador ativo acabava sendo deixado para trás e passava a perseguir o oponente com posse da bola, sem nenhuma chance de alcançá-lo. Para resolver essa vulnerabilidade decidimos reforçar a IA de defesa,

criando um novo papel, o defensivo. O jogador no papel defensivo tem a missão de marcar efetivamente o oponente com posse de bola, se aproximando frontalmente, sempre que possível, dando suporte ao jogador ativo durante todo o tempo de jogo. A escolha do marcador se realiza entre os jogadores que antigamente teriam o papel cooperativo, não interferindo na escolha do papel ativo. Desta forma acrescentamos uma nova posição de suporte mais ativa que a dos aliados cooperativos, que participa diretamente na jogada. O principal desafio foi estabelecer os critérios de seleção do defensor, definir a movimentação mais efetiva e as regras necessárias para evitar interferir no jogador ativo, além de implementar as mudanças no agente e fazer todos os testes de aceitação em aproximadamente 12 horas. Estes assuntos serão tratados na seção 3.

IV. MÓDULO COOPERATIVO DE DEFESA

O modo de defesa se inicia quando se cumprem as seguintes condições:

1) A bola se encontra mais perto do nosso gol do que do gol adversário.

2) O time adversário está com posse da bola (pelo menos um jogador oponente a menos de d_p metros da bola) ou com maior probabilidade de tomar posse dela, do que nosso time. A probabilidade de posse é calculada com base no menor tempo estimado $T2B$ para chegar a uma distância prefixada d_p (em metros) da bola.

No modelo cinemático atual o tempo mínimo necessário para um agente j chegar de um ponto dado até uma distância d_p de um objeto vem dado por duas componentes:

- Linear: Dada pela divisão de distância d_j até o objeto e a velocidade máxima do agente v_{Mj} , da forma:

$$T2BL_j = \frac{d_j - d_p}{v_{Mj}}$$

- Angular: Dada pela divisão do ângulo a_j em relação ao objeto e a velocidade de giro máxima do agente w_{Mj} , da forma:

$$T2BG_j = \frac{a_j}{w_{Mj}}$$

Denotando por d_{m_o} a distância do oponente mais próximo da bola e por v_{M_o} a velocidade máxima esperada dos agentes do time oponente (lida de arquivo de configuração ao inicializar os agentes em cada período da partida) podemos calcular o tempo mínimo esperado para a posse pelo oponente, da forma

$$T2B_o = \frac{d_{m_o} - d_p}{v_{M_o}}$$

onde os subscritos m e M indicam *mínimo* e *máximo*, respectivamente. No resto do texto os segundos subscritos o e a indicam agentes *oponentes* e

aliados, respectivamente. Observe que o tempo de giro foi desconsiderado para o oponente, dando assim vantagem ao adversário em relação à posse de bola ao tempo que evita o trabalho de manter atualizadas estimativas das velocidades de giro dos times adversários no arquivo de configuração. Contudo, as velocidades lineares devem ser atualizadas sistematicamente com base na análise dos jogos mais recentes dos adversários. No arquivo de configuração existe uma estimativa de velocidade máxima do oponente para ser usada no caso de um time desconhecido. No jogo, a identificação do time adversário é feita pelo nome do time retornado pelo servidor.

No estágio atual do BahiaRT cada agente aliado i_a conhece sua velocidade máxima de giro w_{Mia} e sua orientação a_{ia} em relação à bola, mas não dispõe dessa informação dos aliados. Neste contexto, cada agente aliado, não limitado, calcula seu $T2B$ da forma:

$$T2B_{i_a} = \frac{d_{i_a} - d_p}{v_{Mia}} + \frac{a_{i_a}}{w_{Mia}}$$

e assume $T2BG_j = 0$ para todos os outros agentes aliados. Desta forma se dá também vantagem ao aliado em relação à posse de bola evitando-se interferências.

A escolha do agente aliado a assumir o papel ativo é feita pelo critério de menor $T2B$. Já a escolha do agente a assumir o novo papel defensivo satisfaz o objetivo principal do time na defesa: proteger o gol. De acordo com isto, se seleciona entre os agentes com papel cooperativo aquele que tenha melhores probabilidades de interceptar o atacante oponente durante sua aproximação a nosso gol. Para simplificar a escolha, fizemos uma decomposição do domínio espacial, dividindo o campo de 30 metros de comprimento (C) x 20 metros de largura (L) em três zonas transversais, uma central de 6.1 metros, isto é, de aproximadamente três vezes o comprimento do gol, e duas laterais de 6.95 metros, como pode ser visto na fig. 1.

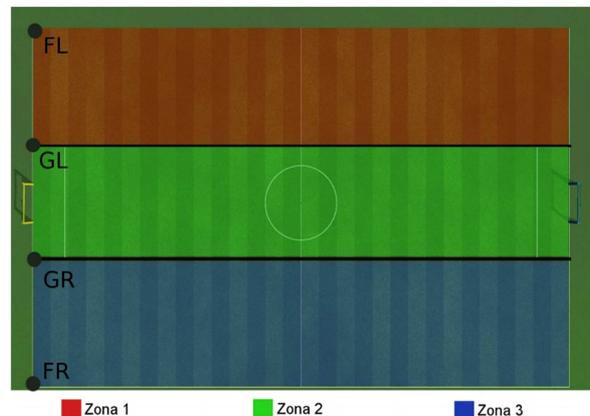


Figura 1 - Divisão do campo em zonas de controle com nosso gol à esquerda do campo.

Definimos 4 pontos de referência localizados na linha de fundo do campo nas fronteiras das três zonas, como também apresentado na fig. 1. Denotamos os pontos de referência por *GL*, *GR* para os marcadores ao lado do gol, e por *FL*, *FR* para os vértices do campo. Na notação utilizada a segunda letra representa a posição do ponto de referência, à direita (*R*, $y < 0$) ou à esquerda (*L*, $y > 0$), em relação a nosso goleiro, quando está de frente para o campo.

Para determinar em qual zona se encontra o atacante oponente, é verificado entre quais marcadores ele está localizado e, numa segunda fase de decomposição do domínio espacial, cada agente define em tempo real uma Região Triangular de Perigo (*RTP*), que é única, dado que um vértice da *RTP* é a posição do oponente e os outros dois vértices (na base do triângulo) são pontos de referência, escolhidos de forma tal que nosso gol sempre esteja contido na *RTP*. Isto é, quando o atacante se encontra em uma das zonas laterais, esquerda ou direita, um dos vértices da base da *RTP* é o ponto de referência externo nessa zona, *FL* ou *FR*, e o outro é o interno na zona oposta, *GR* ou *GL*, respectivamente, como apresentado nas figuras 2 e 4. No caso do atacante se encontrar na zona central, os dois vértices da base da *RTP* são os dois marcadores internos, *GL* e *GR* (fig. 3). Para melhor entendimento das figuras considere de vermelho o ponto de localização do oponente com posse da bola, bem como a *RTP* definida, e representados em azul estão jogadores aliados.

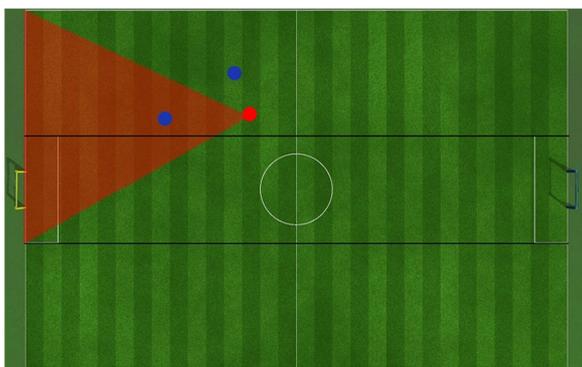


Figura 2 - Triangulação da região de perigo na zona 1.

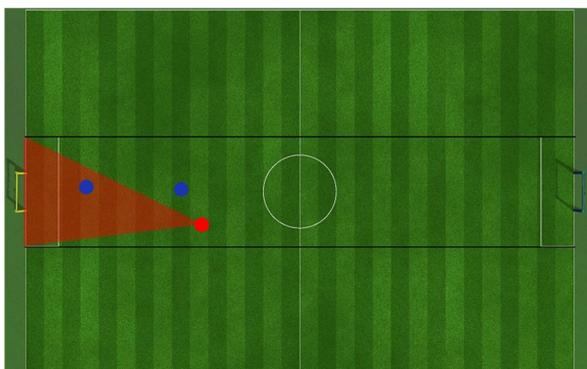


Figura 3 - Triangulação da região de perigo na zona 2.

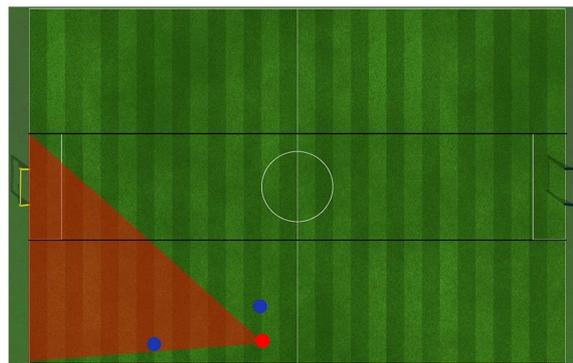


Figura 4 - Triangulação da região de perigo na zona 3.

Uma vez estabelecida a *RTP*, todos os agentes com papel cooperativo dentro dela determinam se devem ou não assumir o papel defensivo. Para isto, se executa um procedimento bastante similar ao descrito para a auto-escolha do agente ativo, trocando apenas o objeto de referência, i.e, a bola pelo atacante. De forma análoga a *T2B* definimos *T2A* como o tempo necessário para um agente aliado chegar a uma distância d_p do atacante definido por:

$$T2A_{i_a} = \frac{d_{i_a} - d_p}{v_{M_{i_a}}} + \frac{\alpha_{i_a}}{w_{M_{i_a}}}$$

onde i_a são todos os agentes aliados dentro da *RTP*; d_{i_a} é a distância linear do agente i_a até o atacante oponente com posse de bola e α_{i_a} o ângulo entre a orientação do agente i_a e o atacante oponente com posse de bola.

Da mesma forma que no cálculo do *T2B*, para os demais aliados na *RTP*, será considerado $T2A_j = 0$.

A escolha do agente aliado dentro da *RTP* a assumir o papel defensivo é feita pelo critério de menor *T2A*. As figuras 2, 3 e 4 também ilustram o papel diferenciado da marcação, pois ainda que haja outro aliado mais próximo do atacante, o aliado marcador será o mais apto dentro da *RTP* selecionada, pois este terá maiores chances de interceptar o oponente durante seu provável deslocamento na direção do gol. Esta forma de definição do papel defensivo, representa uma das principais vantagens deste modelo, pois o mais usual entre os times que disputam a liga de simulação 3D é sempre considerar o agente mais próximo ao alvo. Este modelo considera de forma estratégica o agente mais apto a interceptar o adversário ainda que este não seja o mais próximo.

Uma vez que o papel defensivo é atribuído, este agente passará a ter o comportamento de tentar interceptar a bola recuperando a posse da mesma para o time BahiaRT. É importante observar que neste momento, dois agentes terão o mesmo objetivo: o ativo e o defensivo. Os dois estarão tentando recuperar a posse de bola. Entretanto, o agente defensivo terá o benefício de estar posicionado de forma frontal ao

oponente, estando no caminho entre este e o gol a ser protegido.

O agente ativo – devido às já relatadas limitações na velocidade dos agentes da equipe BahiaRT – na maioria dos casos estará atrás ou ao lado desta trajetória. Entretanto, seu comportamento é fundamental para a efetiva recuperação da posse de bola. No atual ambiente de simulação não há um modelo de faltas implementado. Apesar deste modelo já está em discussão há alguns anos no comitê técnico da liga de simulação 3D, até a RoboCup 2015 ainda não há um modelo implementado.

A consequência disto é que os agentes ainda podem colidir uns com os outros sem que seja registrada uma infração (falta). Desta forma, o agente defensivo – ao tentar recuperar a bola – possivelmente irá colidir com o atacante oponente derrubando-o mas também indo o próprio agente defensivo ao chão. O resultado desta colisão deve ser a sobra da bola para o agente ativo que estará próximo à bola justamente devido ao seu comportamento anterior de tentar interceptá-la.

Então o agente que de fato terá maior chance de retomar a posse de bola é o agente com papel ativo, mas isto só é possível devido a ação do agente defensivo em impedir o progresso do atacante oponente.

O modelo aqui proposto adapta-se facilmente tanto ao cenário atual – onde não há marcação de faltas – quanto o possível cenário futuro com um modelo de faltas definido. Isto é possível pela definição do valor da variável d_p . No cenário atual, adotamos $d_p=0$, fazendo com que o agente defensivo tente de fato interceptar a bola, ao invés de apenas marcar a uma certa distância o jogador adversário. Com o modelo de faltas, este comportamento frequentemente iria gerar uma falta contra o BahiaRT, o que é indesejável. Então neste caso bastará ajustar o valor de $d_p > 0$ impedindo a colisão com o adversário e a possível falta.

Esta abordagem de defesa também pode ser ajustada para marcação de outros jogadores adversários além do que está com a posse de bola. Para isto, basta mudar o alvo de todo o cálculo descrito para o jogador em foco. Isto permitirá criar um mecanismo de marcação homem a homem, impedindo que o time adversário tenha sucesso em realizar um passe por exemplo. Ajustando o valor de d_p pode-se obter um sistema de marcação mais ou menos apertado, dependendo do cenário de jogo. Esta característica pode ser ajustada dinamicamente de acordo com o adversário, com o placar atual de jogo, com o tempo de jogo e uma série de outros fatores.

V. PROCEDIMENTO EXPERIMENTAL

Para validação do módulo desenvolvido, foram realizadas 150 partidas contra 3 dos 4 melhores times da edição 2014 da Liga de Simulação 3D, de acordo com o ranking oficial encontrado na página da RoboCup 2014 [21]. As partidas foram igualmente

distribuídas, totalizando 50 jogos contra cada adversário, sendo metade desses jogos sem o módulo de defesa aqui proposto, e metade o incluindo. Os adversários selecionados foram o UT Austin Vila (campeão), Robocanes (vice-campeão) e FCPortugal (4º colocado). A equipe BahiaRT ficou na 5ª colocação nesta edição da competição. A equipe 3ª colocada (magmaOffenburg) não foi utilizada por ter apresentando problemas nos scripts de início do time fornecidos após a RoboCup 2014.

Estando o trabalho inserido no contexto da Liga de Simulação 3D, o ambiente de testes foi configurado de acordo com as regras da competição, onde são necessárias quatro máquinas dedicadas. Uma das máquinas roda apenas o servidor de simulação Simspark [1], responsável por simular o ambiente de um jogo de futebol, bem como a física do mundo real. Duas máquinas alocadas para executar os times que irão se enfrentar, e uma quarta máquina para rodar o monitor Roboviz [10], ferramenta utilizada na competição oficial cuja função durante os jogos é limitada apenas à exibição da partida.

Para que não houvesse interferência humana nos testes de validação, foi utilizada uma ferramenta desenvolvida no ACSO para automação de testes, chamada Trainer 3D, através da qual é possível rodar uma série de partidas de forma automatizada e coletar dados de cada partida.

Foram determinadas três medidas de desempenho do time no contexto analisado, baseado em três pontos de vista diferentes:

1) *Efetividade no controle da bola*: O objetivo desse indicador é mensurar o Tempo de Posse de Bola (TPB) de cada time, para isso consideramos três possibilidades: posse de bola pertencendo aos aliados, aos oponentes, ou indefinida. Para melhor entender esse critério, considere que, segundo a inteligência do BahiaRT, um jogador tem posse da bola quando a distância entre ele e a bola é inferior a 0.5 metros. No caso de nenhum jogador atender à condição, a posse é considerada indefinida. A comparação do tempo de posse de bola antes e depois da implementação do módulo da defesa, permite-nos confirmar a hipótese de que essa estratégia aumenta a posse de bola da equipe BahiaRT.

a) Em primeira instância se considera como critério de efetividade do controle da bola, isto é, que um time é mais efetivo, na medida que aumenta o Tempo de Posse de Bola pelo time aliado (TPB_T) em relação ao TPB pelo time adversário (TPB_A). Para este critério, pode-se definir a métrica $E_{C,A}$ dada por:

$$E_{C,A} = \frac{TPB_T}{TPB_T + TPB_A}$$

onde o subscrito C refere-se a controle e o subscrito

A ao adversário. Note que $E_{C,A} \in [0, 1]$. O mínimo 0 se atinge quando $TPB_T = 0$ e o máximo 1 quando $TPB_A = 0$. No caso $TPB_A = TPB_T$ se tem $E_{C,A} = 0.5$, por isso $E_{C,A} > 0.5$ indica controle do nosso time e $E_{C,A} < 0.5$ indica controle do time adversário.

b) Em segunda instância se considera como critério de efetividade do controle da bola, que um time é mais efetivo, no controle da bola, na medida que aumenta o Tempo de Posse de Bola (TPB) pelo time TPB_T em relação ao Tempo de Posse Indefinida da Bola (TPIB) na partida $TPIB_p$. Para este critério, pode-se definir a métrica $E_{C,B}$ dada por:

$$E_{C,B} = \frac{TPB_T}{TPB_T + TPIB_p}$$

onde o subscrito C refere-se a controle e o subscrito P à partida. Note que $E_{C,B} \in [0, 1]$. O mínimo 0 se atinge quando $TPB_T = 0$ e o máximo 1 quando $TPIB_p = 0$.

c) Combinação de Critérios: Os dois critérios acima podem ser combinados de diversas formas. A escolha de qual critério corresponde melhor às avaliações feitas por peritos humanos, é um assunto a ser pesquisado.

- Combinação linear: Se pode dar peso diferente a $E_{C,A}$ e $E_{C,B}$.

$$E_{C,L}(\partial) = \partial E_{C,A} + (1 - \partial) E_{C,B}$$

Com $\partial \in [0, 1]$

Note que esta métrica também pertence ao intervalo unitário dado que é a combinação linear de dois critérios também unitários. Note também que depende de um parâmetro.

- Integração formal: O máximo ocorre quando ambos critérios são implicitamente satisfeitos:

$$E_{C,I} = \frac{TPB_T}{TPB_T + TPB_A + TPIB_p}$$

Note que $E_{C,I} \in [0, 1]$. O mínimo 0 se atinge quando $TPB_T = 0$ e o máximo 1

quando $TPIB_p = TPIB_A = 0$. Note também que o denominador é igual ao Tempo da Partida (TDP), isto é:

$$TDP = TPB_T + TPB_A + TPIB_p.$$

2) *Efetividade do posicionamento*: Se considera como critério de efetividade do posicionamento, isto é, que um time é mais efetivo que outro, na medida que aumenta seu Tempo na Área de Ataque (TAA), TAA_T , o qual foi mensurado coletando o tempo da bola na área de ataque durante toda a partida, já que o posicionamento do BahiaRT é atraído pela posição da bola (a coordenada X da bola define em que área ela se encontra). Considera-se que estando na área de ataque, diminui a probabilidade de receber um gol, ao tempo que aumenta a probabilidade de fazer um gol. Para este critério, pode-se definir a métrica E_p dada por:

$$E_p = \frac{TAA_T}{TDP}$$

Note que o Tempo na Área de Ataque (TAA) do adversário TAA_A vem dado por:

$$TAA_A = TDP - TAA_T = TDP(1 - E_p)$$

3) *Efetividade real*: O propósito da defesa é essencialmente interceptar oponentes se deslocando na direção do gol, tentando assim minimizar o número de gols sofridos, portanto como critério de efetividade real foi escolhido o saldo de gols relativo, definido da forma:

$$E_R = \frac{1}{2} \left(1 + \frac{G_T - G_A}{G_T + G_A + \varepsilon} \right)$$

onde G_T e G_A são os gols do nosso time e do time adversário, respectivamente, e ε é uma constante que tende a zero. Em caso de empate $G_T = G_A$ atribui-se uma efetividade real de 50%. Num placar positivo (1-0) temos aproximadamente 75% de efetividade real, e num placar negativo (0-1), temos efetividade aproximada de -75%. No último jogo da seleção contra Alemanha na Copa do Mundo 2014 a efetividade real de nosso time seria:

$$E_R = \frac{1}{2} \left(1 + \frac{G_T - G_A}{G_T + G_A + \varepsilon} \right) = \frac{1}{2} \left(1 + \frac{1 - 7}{1 + 7 + \varepsilon} \right) \cong \frac{1}{2} \left(\frac{2}{8} \right) = 12.5\%$$

Se Brasil não tivesse feito o gol de honra daria:

$$E_R = \frac{1}{2} \left(1 + \frac{-7}{7 + \varepsilon} \right) \cong \frac{1}{2} (0) = 0\%$$

Note que o resultado seria o mesmo ainda que o saldo de gols fosse apenas 1 a favor do adversário. A escolha de uma métrica que não diferencie tanto perder de 1 a 0 do que perder de 5 a 0 se deve ao fato de que o jogo se perde com apenas um gol de saldo negativo. Portanto o foco deve ser voltado para não negatar o saldo de gols e não tanto para evitar que ele aumente demais.

Os três critérios aqui definidos foram coletados durante as partidas e armazenado em uma planilha eletrônica para facilitar a compilação e manipulação dos resultados.

VI. DISCUSSÃO DE RESULTADOS

O módulo de defesa foi desenvolvido durante a edição de 2014 da RoboCup, entre a primeira e a segunda rodada, onde já foi possível vivenciar na prática a sua eficácia. Na primeira rodada, jogando contra 5 times dos quais apenas 1 ficou entre os cinco primeiros colocados no ranking final, o BahiaRT permitiu 5 gols. Já na segunda rodada com a nova defesa implementada, jogando também contra 5 times, mas dentre eles o primeiro e o segundo colocados no ranking final, o BahiaRT permitiu apenas 1 gol. Foi possível perceber também durante as partidas uma melhora sensível no desempenho geral do time. Os resultados em detalhe da liga de Simulação 3D podem ser encontrados na wiki da RoboCup [20].

Os resultados dos testes automatizados com e sem o módulo de defesa foram agrupados na tabela 1. Computamos em cada caso a média de gols sofridos, feitos e a média do saldo de gols. Da esquerda para a direita temos os times 1, 2 e 3, correspondendo ao UT Austin Villa, Robocanes e FC Portugal, respectivamente. Na última coluna se encontram as médias de todas as partidas com e sem defesa.

Tabela 1: Compilação dos placares das partidas jogadas contra cada time com e sem defesa ativada.

Defesa	Gols/Jogo	Time 1	Time 2	Time 3	Total
Sem	Sofridos	0.88	0.48	0.24	0.53
	Feitos	0.04	0.36	0.64	0.34
	Saldo	-0.84	-0.12	0.40	-0.19
Com	Sofridos	0.80	0.64	0.24	0.56
	Feitos	0.04	1,08	0.72	0.61
	Saldo	-0.76	0.44	0.48	0.05

A primeira constatação nos resultados é que mesmo sem o módulo de defesa ativado o número de gols sofridos foi relativamente baixo, totalizando 0.53 gol por partida, em jogos realizados contra os times top do ranking mundial. Já o desempenho ofensivo do BahiaRT sem o módulo de defesa ativado foi de 0.34 gol por jogo.

Quando analisamos os resultados com o módulo de defesa ativado no agente observamos que a média de gols sofridos no geral se manteve aproximadamente igual. Entretanto houve um crescimento significativo da quantidade de gols marcados com uma média de 0.61 gols por partida. Isto representa um crescimento de 79.4% na eficácia ofensiva do time. O saldo de gols que foi negativo sem a defesa ativa (-0.19 gols por partida) tornou-se positivo (0.05 gols por partida), demonstrando claramente a evolução proporcionada pelo módulo de defesa no desempenho global do time.

As três métricas descritas na seção V foram calculadas com e sem a ativação do módulo de defesa. A avaliação quantificada da mudança de desempenho foi estimada como a diferença percentual relativa, dada pela forma:

$$Df = \frac{Com - Sem}{Sem} 100\%$$

As tabelas 2, 3 e 4 apresentam a compilação dos resultados coletados relativos às três métricas, com e sem defesa, e na última coluna se encontra a diferença percentual entre os dois cenários.

Tabela 2: Diferença percentual relativa contra o Time 1 (UT Austin Villa)

Indicadores		Sem	Com	Dif
Controle	ECL	24.11%	25.92%	7.50%
	ECI	11.55%	11.64%	0.81%
Posicionamento	EP	34.16%	36.43%	6.64%
Resultado	ER	4.35%	4.76%	9.52%

Tabela 3: Diferença percentual relativa contra o Time 2 (Robocanes)

Indicadores		Sem	Com	Dif
Controle	ECL	27.84%	30.01%	7.80%
	ECI	11.82%	12.61%	6.66%
Posicionamento	EP	59.35%	59.16%	-0.31%
Resultado	ER	42.86%	62.79%	46.51%

Tabela 4: Diferença percentual relativa contra o Time 3 (FC Portugal)

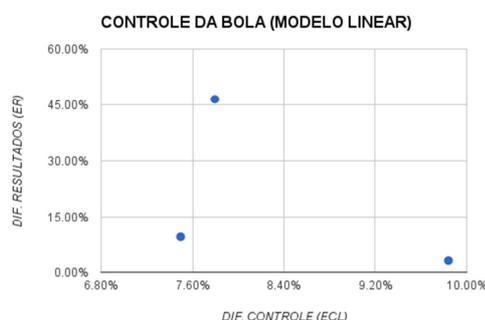
Indicadores		Sem	Com	Dif
Controle	ECL	28.31%	31.10%	9.84%
	ECI	12.34%	13.47%	9.14%
Posicionamento	EP	63.34%	62.99%	-0.56%
Resultado	ER	72.73%	75.00%	3.13%

Contra os três times houve melhoria do indicador de resultados: 9.52% contra o Time 1, 46.51% contra o Time 2 e 3.13% contra o Time 3, contudo o indicador de posicionamento piorou ligeiramente (Time 2: -0.31%, Time 3: -0.56%) exceto contra o UT Austin Villa, onde houve uma melhoria de 6.64%. Os dois indicadores de controle da bola também melhoraram, oscilando entre 0.81% e 9.84%.

Foi estudada a existência de correlação entre as variações dos indicadores de controle e posicionamento com as variações do indicador de resultados.

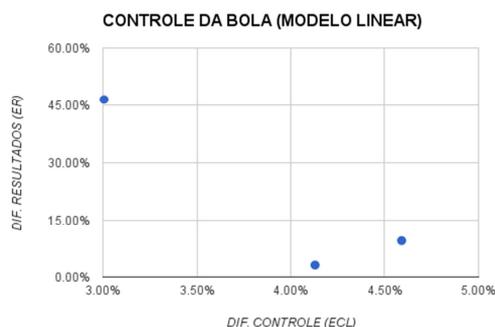
1) Indicador $E_{C,L}$ vs E_R

a) $\partial = 0.3$ (30% de peso para $E_{C,A}$ e 70% para $E_{C,B}$)



Uma pequena melhoria no indicador produz uma grande melhoria nos resultados. Depois, uma grande variação do indicador produz séria perda de performance.

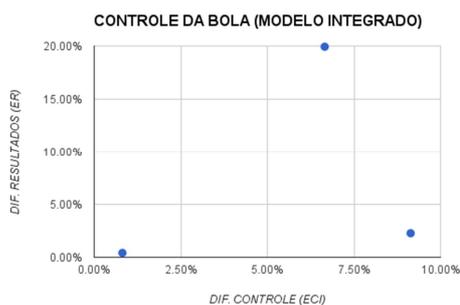
b) $\partial = 0.6$ (60% de peso para $E_{C,A}$ e 40% para $E_{C,B}$)



Uma grande variação do indicador produz séria perda de performance. Depois, uma pequena melhoria no indicador produz melhoria nos resultados.

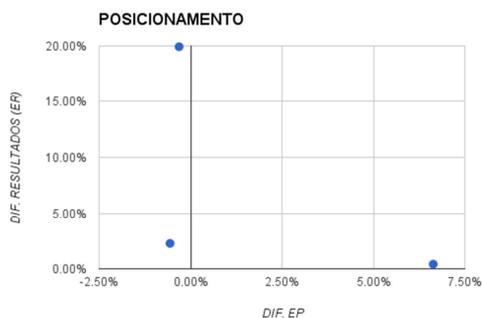
Os resultados mostram que o comportamento dos resultados em função do indicador $E_{C,L}$ varia significativamente em função do peso ∂ . Foi observado que para $\partial < 0.32$ obtém-se o comportamento da figura (a) e que para $\partial > 0.32$ o comportamento da figura (b). Ambos os resultados são contraditórios entre si, o que sugere a exclusão desse indicador enquanto não seja realizado um maior número de testes com mais times visando um melhor entendimento do comportamento do mesmo.

c) Indicador $E_{C,I}$ vs E_R



Este indicador, que não depende de parâmetro, possui uma correlação global positiva com os resultados. Os resultados para os dois maiores indicadores são maiores que o resultado para o menor valor. Contudo, a melhoria dos resultados no ponto intermediário (Robocanes) parece desproporcionada.

d) Indicador E_p vs E_R



Os resultados mostram a inexistência de dependência entre o indicador de posicionamento e o de resultados. Mesmo ficando menor tempo na área de ataque o time conseguiu melhores resultados no saldo de gol contra o FC Portugal e o Robocanes. Já contra o UT Austin Villa um aumento considerável da permanência na área de ataque após ativar o módulo de defesa não conseguiu uma melhoria significativa dos resultados.

VII. CONCLUSÕES

A modificação na IA defensiva, acrescentando um novo comportamento na Inteligência Comportamental do agente, teve resultados significativos sem precisar melhorar as habilidades físicas dos agentes no campo: inteligência vs músculo.

Foi refutada a hipótese de que aumentando o tempo de posse de bola na área de ataque deve melhorar o saldo de gols (recebendo menos gols e fazendo mais gols), embora o universo da pesquisa seja muito limitado. A análise preliminar realizada permite afirmar que o resultado vai depender muito, entre outras coisas, da qualidade da defesa e da potência de contra-ataque adversária. A melhora expressiva dos resultados contra o Robocanes, mesmo perdendo em permanência na zona de ataque e apenas

ganhando um pouco em controle da bola, demonstram que essa hipótese não é verdadeira, atribui-se então a melhora dos resultados a uma provável deficiência na formação defensiva desse time.

A análise dos resultados nos permite concluir que é muito pouco provável prever os resultados derivados de melhorias que aumentem os indicadores de controle da bola e de posicionamento. Por enquanto, o indicador de resultados (saldo de gols) é o único que deve ser utilizado para guiar a otimização do time. A melhoria na IA possibilitou obter um aumento de 79.4%, em média, nos gols marcados, o que implica num aumento considerável da efetividade do time BahiaRT, resultante da implementação do módulo de defesa descrito neste artigo.

A definição de papéis em sistemas multi-agente para desempenhar funções cooperativas específicas demonstrou-se efetiva na Liga de Simulação 3D. No entanto, de acordo com o número de papéis, sistemas multi-agente cooperativos podem obter resultados diferentes, portanto, tais sistemas devem ser analisados cuidadosamente para identificar o número de papéis adequado que maximize os resultados.

Por último, o critério utilizado para determinação da posse de bola não descreve corretamente a realidade do jogo, pois na maior parte do tempo de jogo (aproximadamente 79%) a posse aparece como indefinida. Será necessário reavaliar o critério de posse, o que exigirá a repetição do estudo e a reanálise de hipóteses.

VIII. TRABALHOS FUTUROS

Nas seções anteriores comentamos que a auto-escolha dos papéis de agente ativo e defensor era feita considerando apenas uma velocidade para os aliados e para os oponentes, o que não corresponde com a realidade atual, onde existe uma composição heterogênea de jogadores em cada time. As velocidades de cada agente dependem da morfologia e características do robô físico simulado, existindo 6 tipos diferentes no momento atual. Como é escolhida do técnico do time, associar a cada um dos 11 jogadores escalados um tipo específico, esta informação poderia ser:

- 1) Lida na inicialização dos agentes desde um arquivo de configuração
- 2) Comunicada verbalmente pelos agentes usando o canal auditivo no início de cada tempo de jogo.

Em qualquer caso, para considerar a diferença das velocidades dos aliados mais próximos na hora da auto-escolha do papel pelos agentes, seria necessário realizar modificações importantes no agente.

Já a direção ou o ângulo com relação à bola de cada agente é uma informação dinâmica que precisa

ser estimada em tempo real. Cada agente calcula a partir da visão o ângulo até a bola, usando o ângulo da bola no seu campo de visão e a posição angular da junta do pescoço. A alternativa de menor custo computacional é o compartilhamento dessa informação pelo canal auditivo, contudo, este canal que tem uma largura muito pequena e está totalmente saturado com as informações transmitidas por ele. Neste cenário resulta interessante criar um método, que embora simples, possa estimar a direção do corpo dos agentes aliados e oponentes a partir dos dados recebidos da visão.

Em relação à coleta de dados experimentais, o Trainer 3D deverá ser capaz de identificar e contabilizar as vezes que o agente defensor conseguiu, além de evitar um ataque efetivo do oponente, possibilitar um contra-ataque através do agente que era marcador e que passou a ter um papel ativo.

Atualmente, os principais times que disputam anualmente a liga de simulação 3D possuem pesquisas em andamento para tentar criar jogadas efetivas envolvendo o passe – um dos fundamentos essenciais ao futebol humano e ainda pouco usado no futebol de robôs bípedes – de forma que a adaptação do modelo de defesa para este cenário tornar-se necessária. O modelo global pode ser mantido, mas será necessário introduzir uma fase preliminar de seleção do alvo. Assim, ao invés de ter como alvo fixo a ser marcado o oponente com posse de bola, outros oponentes podem ser tomados por alvo e um esquema de marcação adaptativo ser gerado a partir do mesmo modelo.

Outros trabalhos futuros, não menos importantes, estão relacionados com a melhoria das habilidades físicas dos agentes: aumento da velocidade e estabilidade, aumento da precisão angular e da regulação do alcance do chute, aumento do alcance máximo do chute, para os vários tipos específicos de agentes. Esta otimização poderá melhorar significativamente, em complemento com o modelo de defesa apresentado, os resultados positivos da defesa em relação aos números de gols sofridos.

AGRADECIMENTOS

Agradecemos aos programas e instituições que financiam parcialmente este projeto: CNPq/PIBIC, FAPESB/IC, UNEB/PICIN. Agradecemos também a equipe FC Portugal pela cooperação em parceria com as Universidades de Aveiro, Minho e Porto em Portugal.

REFERENCES

[1] Boedecker, Joschka, and Minoru Asada. "Simsparckoncepts and application in the robocup 3d soccer simulation league." *Autonomous Robots* (2008): 174-181.

[2] Boella, Guido, and Leendert Van Der Torre. "The ontological properties of social roles in multi-agent

systems: Definitional dependence, powers and roles playing roles." *Artificial Intelligence and Law* 15.3 (2007): 201-221.

[3] The dynamics of reinforcement learning in cooperative multiagent systems. In *Proceedings of the fifteenth national/tenth conference on Artificial intelligence/Innovative applications of artificial intelligence*, pages 746-752. American Association for Artificial Intelligence, 1998.

[4] L.P. Kaelbling, M.L. Littman, and Andrew Moore. Reinforcement learning: A survey. *Journal of Artificial Intelligence Research*, 4:237-285, 1996.

[5] Zhang, Xiaoqin, Haiping Xu, and Bhavesh Shrestha. "An integrated role-based approach for modeling, designing and implementing multi-agent systems." *Journal of the Brazilian Computer Society* 13.2 (2007): 45-60.

[6] Kim, In-Cheol. "Dynamic role assignment for multi-agent cooperation." *Computer and Information Sciences-ISCIS 2006*. Springer Berlin Heidelberg, 2006. 221-229.

[7] Marsella, Stacy, et al. "On being a teammate: Experiences acquired in the design of RoboCup teams." *Proceedings of the third annual conference on Autonomous Agents*. ACM, 1999.

[8] Wells, David, et al. "Adaptive defense coordination for multi-agent systems." *Multi-Agent Security and Survivability, 2004 IEEE First Symposium on*. IEEE, 2004.

[9] Neruda, Roman, and Ondřej Kazík. "Role-based design of computational intelligence multi-agent system." *Proceedings of the International Conference on Management of Emergent Digital EcoSystems*. ACM, 2010.

[10] Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Prentice-Hall, Englewood Cliffs, NJ, 2nd edition edition, 2003. Vo, Quoc Bao, and Minyi Li. "On the role of expectations in multi-agent reasoning and decision making." *Web Intelligence and Intelligent Agent Technology (WI-IAT), 2012 IEEE/WIC/ACM International Conferences on*. Vol. 2. IEEE, 2012.

[11] Wells, David, et al. "Adaptive defense coordination for multi-agent systems." *Multi-Agent Security and Survivability, 2004 IEEE First Symposium on*. IEEE, 2004.

[12] Zhang, Xiaoqin, Haiping Xu, and Bhavesh Shrestha. "An integrated role-based approach for modeling, designing and implementing multi-agent systems." *Journal of the Brazilian Computer Society* 13.2 (2007): 45-60.

[13] J. Clerk Maxwell, *A Treatise on Electricity and Magnetism*, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68-73.

[14] I.S. Jacobs and C.P. Bean, "Fine particles, thin films and exchange anisotropy," in *Magnetism*, vol. III, G.T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271-350.

- [15] K. Elissa, "Title of paper if known," unpublished.
- [16] R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.
- [17] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," IEEE Transl. J. Magn. Japan, vol. 2, pp. 740-741, August 1987 [Digests 9th Annual Conf. Magnetism Japan, p. 301, 1982].
- [18] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.
- [19] "Soccer Simulation League/RoboCup2014." - RoboCup Federation Wiki. Web. 27 May 2015. <http://wiki.robocup.org/wiki/Soccer_Simulation_League/RoboCup2014#3D>.
- [20] "RoboCup Soccer 3D Simulation League." RoboCup 2014. Web. 27 May 2015. <http://www.robocup2014.org/?page_id=3383>.