

# SISTEMA AUTÔNOMO INTELIGENTE BASEADO EM COMPUTAÇÃO EVOLUTIVA APLICADO À NAVEGAÇÃO DE ROBÔS MÓVEIS

RENATO REDER CAZANGI, MAURÍCIO FIGUEIREDO

*Departamento de Informática, Universidade Estadual de Maringá – UEM*

*Av. Colombo, 5790, CEP 87020-900, Maringá – PR*

*E-mails: rrcazang@din.uem.br, mauricio@din.uem.br*

**Resumo** — O sistema autônomo inteligente proposto é capaz de aprender uma estratégia geral de navegação para guiar um robô ao alvo, sem colisões contra obstáculos. O sistema é projetado de acordo com os conceitos básicos da teoria da computação evolutiva. Cada indivíduo (regra se-então) é associado a um cromossomo. A função objetivo é definida diferentemente para cada um dos três procedimentos de aprendizagem possíveis, que são disparados por eventos específicos: colisão, captura de alvos e monotonia (um evento virtual, tal qual uma percepção, que ocorre após um longo período de tempo sem que tenha sido detectado colisão ou captura). Cada um dos procedimentos de aprendizagem define um diferente tipo de mecanismo evolutivo. As várias contribuições apresentadas visam explorar o poder dos sistemas classificadores na aprendizagem de comportamentos conflitantes (busca de alvo e desvio de obstáculos). Resultados de simulação mostram que o sistema de navegação proposto é capaz de aprender estratégias de navegação reativas em ambientes desconhecidos, com reduzido custo computacional.

**Abstract** — The intelligent autonomous system proposed is able to learn a general navigation strategy for guiding a robot to the target, without collisions against obstacles. The system is designed according to basic concepts of the evolutionary theory. Each individual (if-then rule) is associated to a chromosome. The fitness function is defined differently for each of three possible learning procedures, triggered by specific events: collisions, target captures or monotony (a virtual event, like a perception, that happens after a very large time period without neither collision nor target capture detection). Each learning procedures define a distinct reproduction mechanism to generate a new population. Simulation results show that the robot initially fails in its task (it does not know any navigation strategy, either collision avoidance or target seeking). As the autonomous system guides the robot, it interacts with the environment (unknown environment) and acquires knowledge from it. After many trials, it generates a successful trajectory. The computational effort associated to the learning period is not so high as reported by many similar works that adopt the same approach (evolutionary computing).

**Keywords** — mobile robot; autonomous navigation system; evolutionary computing; learning classifier system.

## 1 Introdução

Sistemas autônomos inteligentes que detenham a habilidade de navegação são capazes de conduzir um robô móvel até seu alvo, através de uma trajetória isenta de colisões. Tais sistemas de navegação representam internamente estratégias de navegação que definem adequadamente as ações do robô para cada uma das inúmeras situações distintas que se apresentam ao longo da trajetória. Devido à variabilidade de situações, sendo algumas inéditas, o desempenho de um sistema de navegação depende de sua capacidade de aprendizagem. Ou seja, somente aperfeiçoando sua estratégia de navegação, através da aprendizagem, um sistema de navegação torna-se apto para guiar eficientemente o robô em ambientes diversos (associados a diferentes graus de complexidade) e sujeitos a perturbações variadas. Propostas de sistemas de navegação [1] [2] [3] incapazes de aprender suas estratégias a partir da interação com o ambiente têm sido criticadas [4].

Atualmente, a tendência dos esforços de pesquisa em sistemas de navegação é contribuir para o aperfeiçoamento de sua aprendizagem. Neste sentido, técnicas da inteligência computacional, e.g., redes neurais [5] [6] e computação evolutiva [7] têm sido adotadas para investigação e aprofundamento das respectivas teorias.

A proposta em [8] apresenta um sistema de navegação baseado em redes neurais nebulosas, capaz de aprender estratégias de navegação gerais a partir de um ambiente particular. Para superar possíveis dificuldades causadas por configurações em “U” (trajetórias cíclicas), a estrutura deste sistema é expandida com a adição de um repertório neural que modela uma memória de curta duração [9].

Uma das variantes da computação evolutiva, a saber, os sistemas classificadores [10][11], tem proporcionado confortáveis (modelagem) e poderosos (aprendizagem) modelos para sistemas de navegação [12] [13]. Entretanto, trabalhos relatam o excessivo custo computacional para a sua aprendizagem [14].

O sistema autônomo inteligente proposto tem como base teórica os sistemas classificadores. Cada indivíduo (regra se-então) é associado a um cromossomo com dois genes especiais, responsáveis por tipo de dominância e fator de dominância, respectivamente. A função objetivo é definida diferentemente para cada um dos três procedimentos de aprendizagem possíveis, que são disparados por eventos particulares: colisão, captura de alvos e monotonia (um evento virtual, tal qual uma percepção, que ocorre após um longo período de tempo sem que tenha sido detectado colisão ou captura). Cada um dos procedimentos de aprendizagem define um diferente tipo de mecanismo de reprodução. As várias contribuições apresentadas visam explorar o poder dos sistemas classificadores na aprendizagem de comportamentos conflitantes (busca de alvo e desvio de

obstáculos). Resultados de simulação mostram que o sistema de navegação proposto é capaz de aprender estratégias de navegação reativas em ambientes desconhecidos, com reduzido custo computacional.

O artigo tem continuidade com a seguinte seqüência de seções. A seção 2 descreve as características do robô móvel utilizado, bem como o ambiente de simulação. A arquitetura básica do sistema é tratada na seção 3. Na seção 4, os resultados obtidos em simulação são apresentados. A seção 5 apresenta as conclusões resultantes das simulações e os comentários finais.

## 2 Características do Robô

O robô é modelado de forma simplificada, não apresentando dinâmica interna. A velocidade é constante. O ajuste de direção do robô (incremental) pode assumir valores no intervalo de  $-15^\circ$  a  $+15^\circ$ .

Há duas classes de sensores para captar direção do alvo e distância aos obstáculos, respectivamente. Cada sensor de direção do alvo apresenta uma saída cuja intensidade depende do grau de alinhamento do alvo com o sensor. Cada um dos sensores de distância detecta a distância do obstáculo presente na direção que o sensor está alinhado. São considerados 37 sensores de cada tipo. A Figura 1 ilustra as leituras dos sensores correspondentes à situação de simulação também apresentada. O gráfico superior exibe as amplitudes dos estímulos capturados pelos sensores de obstáculos e o gráfico inferior apresenta resultados análogos para os sensores de alvo.

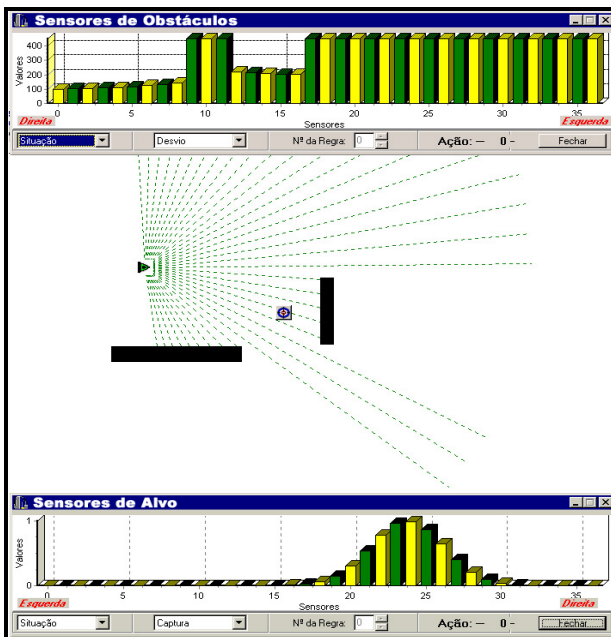


Figura 1: Leitura dos sensores em uma simulação.

Quando o robô captura um alvo, este é prontamente removido do ambiente e outro é inserido.

## 3 O Sistema de Navegação

O sistema de navegação é composto por quatro componentes principais: população de regras, módulo de sensoriamento, módulo de navegação e módulo de aprendizagem. Elas se relacionam através de um ciclo de execução, estabelecendo aprendizagem da estratégia de navegação e determinação do ajuste de direção do robô (Figura 2).

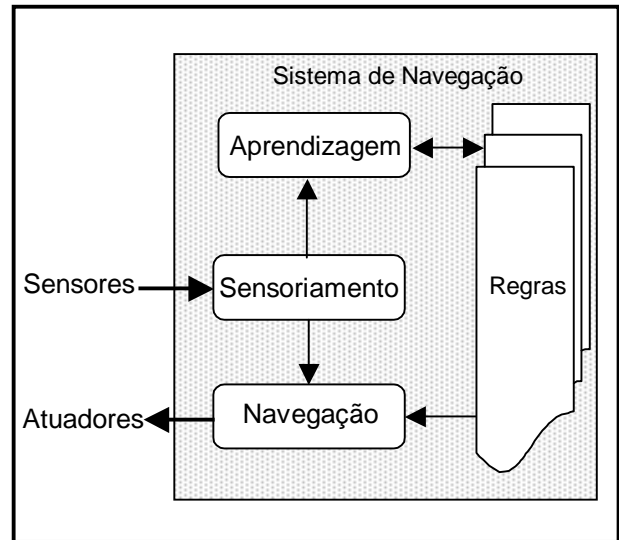


Figura 2: Diagrama do sistema de navegação.

### 3.1 População de Regras

Tal qual em um sistema classificador (SC), cada indivíduo da população é uma regra do tipo se-então. O antecedente corresponde a padrões de distância a obstáculos e direção do alvo, assim como aqueles captados pelos sensores; e o conseqüente, aos ajustes de direção definidos antes de cada deslocamento do robô.

A Figura 3 apresenta a representação tanto do antecedente quanto do conseqüente, na forma de vetores. O antecedente é representado por dois vetores, a saber, desvio e captura e seus componentes são valores inteiros e reais, respectivamente. O número de componentes de cada vetor é idêntico ao número de sensores. Cada componente representa o valor lido pelos sensores após cada deslocamento do robô. Os componentes do vetor de ação (correspondente ao conseqüente) assumem valores binários. A dimensão do vetor é definida pelo intervalo de variação dos ajustes de direção. Neste trabalho, o vetor de ação possui cinco bits, sendo que o bit mais significativo é o bit de sinal: 0 é positivo e 1 é negativo.

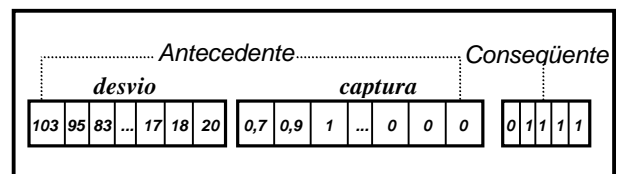


Figura 3: Estrutura das regras da população

Para que o robô inicie sua navegação sem conhecimento sofisticado incorporado, a população de regras é inicializada com valores aleatórios.

### 3.2 Módulos de Sensoriamento e Navegação

O módulo de sensoriamento é responsável por receber os sinais captados pelos sensores e enviá-los para o módulo de navegação. Este módulo também envia ao módulo de aprendizagem estímulos originados devido à captura de alvo, à colisão ou à monotonia (os respectivos sensores não são mostrados na figura).

O módulo de navegação determina o ajuste da direção de movimento do robô. Para tanto, os vetores do antecedente de cada regra são comparados com os respectivos vetores de estímulos (direção do alvo e distância). Neste trabalho, adota-se a distância Euclidiana para efeitos de comparação de vetores. A regra cujo antecedente apresente maior semelhança com os estímulos é selecionada. O valor representado pelo respectivo conseqüente é enviado para o atuador, determinando o ajuste de direção do robô.

### 3.3 Módulo de Aprendizagem

Os procedimentos de aprendizagem podem ser descritos segundo uma seqüência de três etapas básicas: avaliação dos indivíduos, aplicação de mecanismos evolutivos e atualização da população (Figura 4). Há três tipos de procedimentos de aprendizagem diferenciando-se pelo evento de disparo e objetivo.

Quando a aprendizagem é disparada, devido a uma colisão, esta tem o objetivo de melhorar a capacidade de desvio de obstáculos. No caso de captura, o objetivo é aprimorar a habilidade de captura de alvos. No caso de monotonia, estimula-se a busca a alvos.

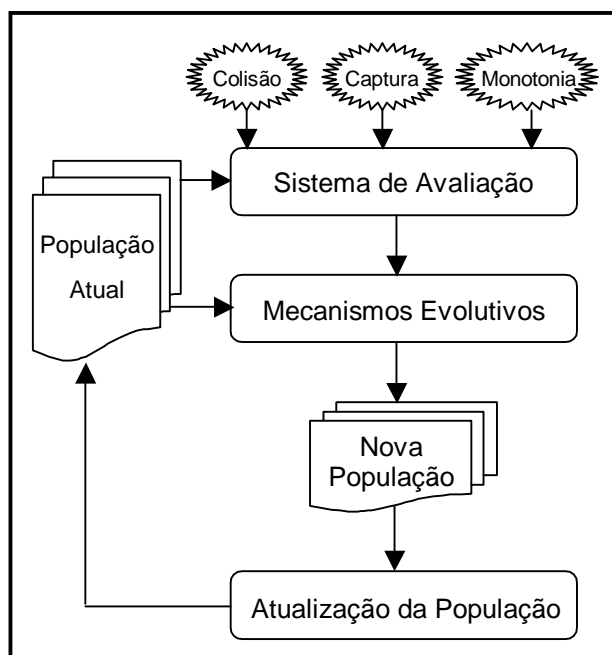


Figura 4: Módulo de aprendizagem.

Considerando que cada um dos tipos de aprendizagem tem objetivo distinto, fazem-se necessárias diferenciações internas nas etapas de aprendizagem, principalmente na etapa de avaliação. A avaliação possui processos e configurações peculiares a cada tipo de procedimento de aprendizagem. Na etapa de mecanismos evolutivos são realizadas modificações nos indivíduos através de técnicas baseadas na evolução natural das espécies, sendo elas: seleção, cruzamento e mutação. Concluindo a aprendizagem, a etapa de atualização substitui a população atual pela nova geração.

#### 3.3.1 Aprendizagem por Colisão

Na aprendizagem por colisão as novas regras são construídas gerando suas partes (antecedente e conseqüente) de maneira separada e distinta. Para cada novo indivíduo, primeiro é gerada a parte antecedente (condição) e só depois a parte conseqüente (ação). Os genitores dos antecedentes e conseqüentes são diferentes.

Inicialmente, há a avaliação de antecedentes segundo sua semelhança com os estímulos recebidos no momento da colisão. As regras de maior pontuação correspondem àquelas cujos antecedentes apresentam maior semelhança. Em seguida, os mecanismos evolutivos geram os antecedentes das regras da nova geração. Para preservar o conhecimento presente na população, apenas 1/3 da população é alterada (2/3 dos indivíduos permanecem sem alteração). Uma regra especial é introduzida, chamada ideal. Seu antecedente é formado pelos vetores recebidos dos sensores de distância e direção do alvo no instante da colisão, visando acelerar a convergência das regras. Em seguida são gerados os conseqüentes para as regras, a partir de uma nova avaliação. Para tanto, considera-se o lado do robô em que a colisão se sucedeu e o reflexo instintivo associado (colisões do lado esquerdo ajustam a direção do robô para o lado direito e vice-versa). Procede-se nova avaliação de regras, considerando-se a semelhança do reflexo instintivo e o conseqüente de cada regra. Os mecanismos evolutivos atuam sobre a população, originando os conseqüentes para as novas regras (já criadas com a geração de antecedentes). Por fim, atualiza-se a população: as novas regras (1/3 do total) substituem as regras antigas enquanto as demais regras são perpetuadas na nova geração.

#### 3.3.2 Aprendizagem por Captura

Na aprendizagem por captura as novas regras também são construídas gerando suas partes separadamente. Com relação à parte antecedente, são realizados procedimentos análogos à aprendizagem por colisão. Os antecedentes são avaliados de acordo com sua semelhança com os vetores de estímulos correspondentes ao momento da captura. Em seguida, são executados os mecanismos evolutivos também em

apenas 1/3 da população. Uma regra especial, cujo antecedente é igual ao vetor de estímulos no momento da captura, é inserida na nova população. Procede-se outra avaliação da população, classificando as regras segundo seu desempenho nas interações que precederam a captura. As regras que atuaram nestas interações são avaliadas de acordo com o desvio provocado entre a direção do alvo e a direção de movimento do robô. Maior pontuação corresponde à redução de desvio. Atuando sobre a população, os operadores de seleção, cruzamento e mutação geram os conseqüentes e definem os novos indivíduos que passam a compor a população.

### 3.3.3 Aprendizagem por Monotonia

A aprendizagem por monotonia é desencadeada quando o robô passa um longo tempo (definido *a priori*) sem colidir ou capturar alvos. Inicialmente o sistema de navegação não possui uma estratégia de navegação eficaz, portanto não busca alvos e eventos de captura ocorrem ao acaso. A aprendizagem por monotonia tem como objetivo gerar regras associadas ao comportamento de captura de alvos. Sendo assim, o procedimento de aprendizagem atua apenas no conseqüente das regras. Para a avaliação da população, todas as regras recebem uma mesma pontuação inicial. Considera-se a atuação das regras no período associado à respectiva monotonia. A pontuação de cada regra é alterada de acordo com sua contribuição para a redução ou aumento do desvio entre a direção de movimento do robô e a direção do alvo. Caso a regra diminua o desvio do robô com relação ao alvo, ela recebe uma bonificação em pontos, caso contrário, a regra é penalizada, perdendo pontos. Somente as regras que tiverem pontuação abaixo do valor inicial são selecionadas para alterações. Os mecanismos evolutivos que alteram as regras são distintos do cruzamento e mutação. As regras selecionadas têm seus conseqüentes ajustados diretamente de acordo com os vetores de captura dos respectivos antecedentes. Ou seja, se uma regra apresenta um padrão de captura de alvo à esquerda, seu conseqüente deve representar uma ação que gere ajuste de direção à esquerda. O vetor de desvio do antecedente de cada regra selecionada é reinicializado (cada componente recebe valores aleatórios) garantindo a consistência da regra. Finalizando o processo, as regras modificadas são colocadas na nova geração juntamente com as demais que foram mantidas.

### 3.3.4 Genes Especiais

Ainda há um fator importante no módulo de aprendizagem que é o sistema de dominância das regras. Ele foi criado para que regras evoluídas não sejam alteradas de forma destrutiva. A idéia de dominância tem relação com alguns conceitos de genes

dominantes e recessivos da genética, mas nesse caso com outros fins. Todas as regras possuem um gene associado com a dominância e outro com o tipo de dominância:

- tipo de dominância: associado à ação determinada pela regra para o robô, a saber, desvio ou captura;
- fator de dominância: grau de importância da regra para a população, variando entre 0 e 100.

Estes genes restringem o espaço de seleção de regras. Caso a aprendizagem seja causada por captura, todas as regras em que o tipo de dominância também seja captura são consideradas pelos mecanismos evolutivos. Entretanto, caso a aprendizagem seja causada por colisão, nem todas as regras em que o tipo de dominância seja captura são consideradas pelos mecanismos evolutivos, apenas aquelas cujo fator de dominância é menor que uma certa constante predefinida. A idéia é análoga para as regras cujo tipo de dominância é colisão. Não há restrições de seleção quando a aprendizagem é causada por monotonia. Neste caso, todas as regras resultantes assumem tipo de dominância captura e fator de dominância máximo.

O fator de aprendizagem das regras ideais (definidas a partir dos vetores de estímulos de distância e direção do alvo) assume valor máximo assim que são inseridas nas populações (são regras de conhecimento sofisticado).

A recombinação de regras estabelece alterações nos valores do fator de dominância, sendo, de uma forma geral, uma média dos valores dos respectivos genitores ponderada pela posição do corte. Por exemplo, se o genitor 1 e o genitor 2 são do tipo desvio, possuem respectivamente dominância 80 e 30 e o ponto de cruzamento é no gene 5 (de um total de 20) o resultado seria o seguinte: regra 1 e regra 2 seriam do tipo desvio e as dominâncias seriam, respectivamente, 42 e 68. São definidos diferentes ajustes caso os genitores não tenham o mesmo tipo de dominância.

## 4 Resultados

Em seguida apresentam-se experimentos desenvolvidos em ambiente de simulação, visando demonstrar as potencialidades do sistema de navegação proposto. Em todos os experimentos (exceto o experimento 2), apenas um alvo é colocado inicialmente no ambiente. Após cada evento de captura de alvo, um alvo adicional é inserido. A população de regras é constituída por 180 indivíduos.

A Figura 5 apresenta os resultados do experimento 1. O ambiente de navegação não possui obstáculos. A trajetória percorrida pelo robô indica que as regras inicialmente definidas representam algum conhecimento, mesmo sendo geradas aleatoriamente,

pois o sistema de navegação guiou o robô aos dois primeiros alvos inseridos no ambiente, sem qualquer aprendizagem. Ao buscar o terceiro alvo, um evento de monotonia ocorre (no ponto da trajetória indicado pela seta). Após a aprendizagem por monotonia, o robô demonstra um comportamento de captura mais definido. Outros alvos são capturados em seguida.

No experimento 2, o ambiente é configurado com um único obstáculo. Não são inseridos alvos. O robô colide algumas vezes. À medida que interage com o ambiente (estimulado pelas colisões), o sistema de navegação adquire conhecimento, gerando uma estratégia de desvio de obstáculos mais eficiente.

Os dois experimentos apresentados não exigem do sistema autônomo inteligente explorar completamente suas potencialidades, pois apenas um único comportamento foi aprendido, captura de alvos ou desvio de obstáculos (observe-se que cada experimento considera um sistema de navegação distinto).

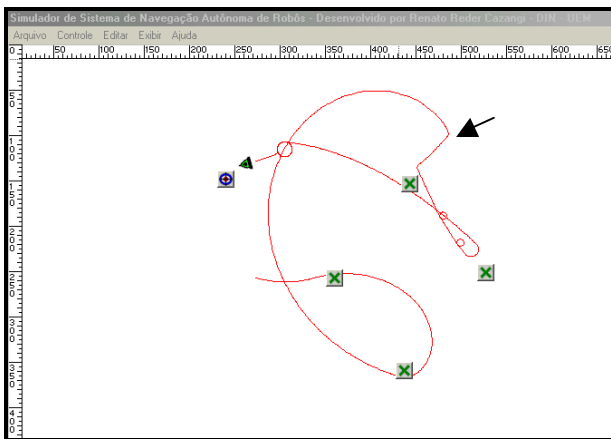


Figura 5: Aprendizagem de estratégia de captura de alvos (ambiente sem obstáculos).

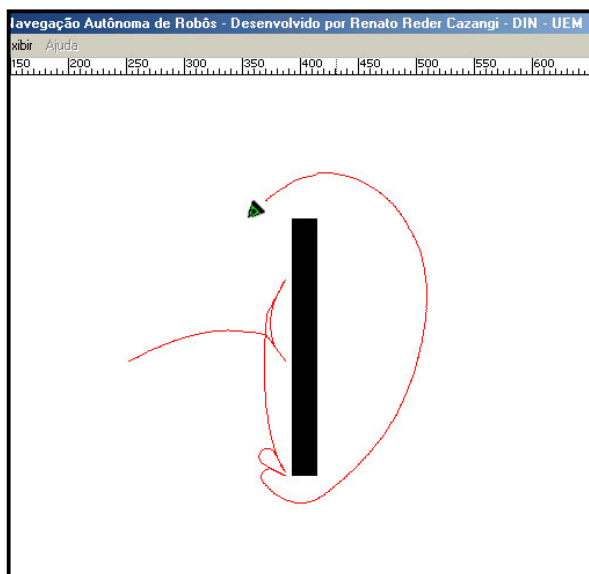


Figura 6: Aprendizagem de estratégia de desvio (ambiente sem alvos).

No experimento 3 obstáculos e alvos estão presentes no ambiente. Após colisão contra o obstáculo, o robô percorre uma trajetória até o primeiro alvo. em sua busca pelo segundo alvo, o robô evita uma colisão iminente. Observa-se que duas estratégias estão sendo geradas a partir da interação com o ambiente.

O comportamento de monotonia também interfere na aprendizagem de uma estratégia de navegação em ambiente com obstáculos e alvos. O experimento 4 ilustra este fato conforme ilustrado na Figura 8. Inicialmente o robô colide. Contornando o obstáculo, alcança o alvo e assume uma trajetória pouco eficiente. Um evento de monotonia ocorre após longo tempo. O sistema de navegação aprende e passa a definir uma trajetória mais eficiente.

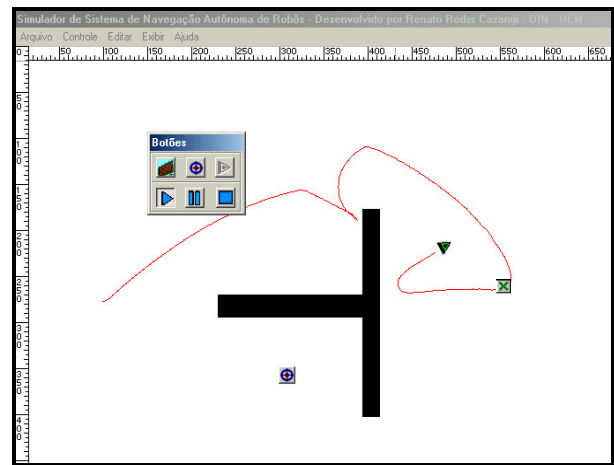


Figura 7: Aprendizagem de estratégia de navegação (captura de alvos e desvio de obstáculos).

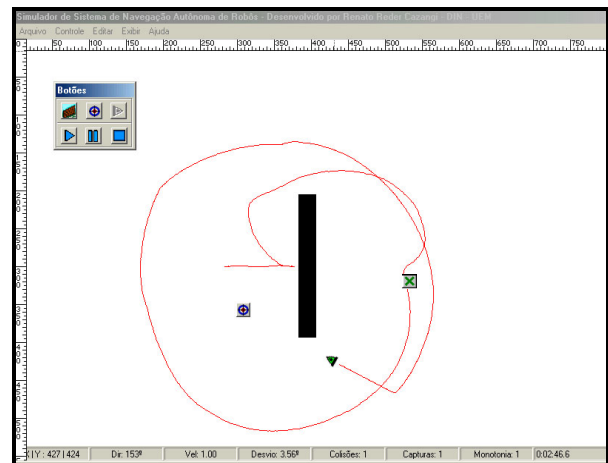


Figura 8: Evento de monotonia disparando procedimento de aprendizagem.

## 5 Conclusão

A navegação autônoma de robôs proporciona desafios bastante propícios para a investigação da aprendizagem em sistemas autônomos inteligentes. Uma das tendências da pesquisa consiste em desenvolver sistemas de navegação capazes de gerar estratégias a

partir da interação com o ambiente. Em geral, as propostas apresentadas têm mostrado uma simbiose de teoria e técnicas provenientes de diferentes áreas de pesquisa (inteligência artificial, biologia, psicologia, neurofisiologia, engenharia, etc.).

Muitas propostas, mesmo as baseadas em redes neurais ou computação evolutiva, consideram que o sistema de navegação seja dotado, a priori, de comportamentos instintivos, e.g. desvio de obstáculos ou busca de alvos. Nestes casos o sistema deve aprender a coordená-los.

O resultado de pesquisa apresentado tem como motivação explorar a grande capacidade de aprendizagem da computação evolutiva (especificamente dos sistemas classificadores) inclusive reduzindo o custo computacional. O sistema autônomo inteligente apresenta várias contribuições: funções objetivo, mecanismos de avaliação, mecanismos de reprodução, conceitos de dominância, entre outras.

Resultados de simulação demonstram que o sistema autônomo inicialmente desconhece qualquer estratégia de navegação (também não há comportamentos instintivos inseridos no robô). Eventuais eventos de colisão e captura disparam os respectivos procedimentos de aprendizagem. No entanto, a princípio o evento de monotonia (originalmente sugerido neste trabalho) parece ser o evento de maior importância na aquisição de conhecimento e, conseqüentemente, para desenvolvimento de estratégia de navegação.

A expectativa para trabalhos futuros, considera a concentração de esforços no aperfeiçoamento dos mecanismos evolutivos, particularmente na investigação do evento de monotonia. Experimentos preliminares mostram que este evento possa causar degradação do conhecimento presente na população de regras (estratégias eficientes), pois sua ocorrência também depende do tipo de configuração do ambiente e não só da qualidade da estratégia adquirida.

### Agradecimentos

Os autores agradecem ao CNPq pelo suporte financeiro ao projeto 300982/97-3.

### Referências

[1]Saffiotti, A., Ruspini, E. e Konolige, K., "Using Fuzzy Logic For Mobile Robot Control", D. Dubois, H. Prade and H. Zimmermann (eds.), International Handbook of Fuzzy Sets, Kluwer Academic Publisher, 1999.

[2]Ben-Shahar, O. e Rivlin, E., "To push or not to push: On the rearrangement of movable objects by a mobile robot", IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, v. 28, nº 5, pp. 667-679, outubro, 1998.

[3]Rosenblatt, J., "DAMN: A Distributed Architecture for Mobile Navigation", Journal of Experimental and Theoretical Artificial Intelligence, v. 9, nº 2-3, pp.339-360, abril-setembro, 1997.

[4]Verschure, P.F.M.J., "A Bottom Up Approach Towards the Acquisition and Expression of Sequential Representations Applied to a Behaving Real-World Device: Distributed Adaptive Control III, Neural Networks, v. 11, pp. 1531-1549, 1998.

[5]Millan, J., "Rapid, Safe, and Incremental Learning of Navigation Strategies", IEEE transactions on System, Man, and Cybernetics, Part B, pp. 408-420, 1996.

[6]Janabi-Sharifi, F., "A fuzzy-neural navigation system for a mobile robot", Proceedings of CSME Forum 1998, Toronto, Canada, pp. 58-62, maio, 1998.

[7] Wilson, M., King, C. e Hunt, J., "Evolving hierarchical robot behaviours", Robotics and Autonomous Systems, v. 22, nº 3-4, pp. 215-230, 1997.

[8]Figueiredo, M. "Redes Neurais Nebulosas Aplicadas a Problemas de Modelagem e Controle Autônomo", Tese de Doutorado, Departamento de Engenharia de Computação e Automação Industrial, FEEC, UNICAMP, 1997.

[9] Melo, W. e Figueiredo, M., "Memória de Curta Duração em Sistema Autônomo Aplicado à Navegação de Robôs", Anais do V Congresso Brasileiro de Redes Neurais, pp. 493-498, Rio de Janeiro, abril, 2001.

[10]Holland, J.H., "Escaping Brittleness: The Possibilities of General-Purpose Learning Algorithms Applied to Parallel Rule-Based Systems", em Michalsky, R.S., Carbonell, J.G. e Mitchell, T.M., (eds.), Machine Learning, Morgan Kaufmann, San Mateo, EUA, 1986.

[11]Bonarini, A., "Learning fuzzy classifier systems", em P. L. Lanzi, W. Stolzmann, e S. W. Wilson (eds.), Learning Classifier System: new directions and concepts, Springer Verlag, Berlim, 2000.

[12] Iwakoshi, Y., Furuhashi, T. e Uchikawa, Y. "A Fuzzy Classifier System for evolutionary learning of robot behaviors", Applied Mathematics and Computation, v. 91, nº 1, pp. 73-81, abril, 1998.

[13] Pipe, A. e Carse, B. "Autonomous Acquisition of Fuzzy Rules for Mobile Robot Control: First Results from two Evolutionary Computation Approaches", em D. Whitley, D. Goldberg, E. Cantu-Paz, L. Spector, I. Parmee & H-G. Beyer (eds), GECCO-2000: Proceedings of the Genetic and Evolutionary Computation Conference, pp849-864, Morgan Kaufmann, New York, 2000.

[14]Colombetti, M., Dorigo, M. e Borghi, G., "Behavior Analysis and Training –A Methodology for Behavior Engineering", IEEE transactions on System, Man, and Cybernetics, Part B, pp. 365-380, 1996.