

Rede Neural Modular e Aprendizagem por Reforço para Sistema de Navegação Autônoma

Rodrigo Calvo, Maurício Figueiredo

Departamento de Informática / Universidade Estadual de Maringá – Av. Colombo 5790, Maringá – PR, CEP 87020-900

E-mails: rcalvo@din.uem.br, mauricio@din.uem.br

Abstract

This work describes an autonomous navigation system based on a modular neural network. The environment is unknown and the system is not able for navigation. As the robot collides and takes targets, it improves its navigation strategy and efficiently guides the robot to targets. A reinforcement learning mechanism adjusts parameters of the neural networks at target capture and collision moments. Simulation experiments show performance comparisons. Only the proposed system reaches targets if the environment presents a high risk (dangerous) configuration (targets are very close to obstacles).

1. Introdução

Sistemas de navegação devem ser capazes de guiar robôs sem nenhuma interferência externa (sistemas independentes). Algumas abordagens de sistemas de navegação consideram que o ambiente (topologia e eventos) é totalmente conhecido, ou seja, o sistema de navegação tem o conhecimento do mapa do ambiente e todos os possíveis eventos pertencem a um número finito de classes conhecidas. Neste caso, o sistema de navegação pode ser programado com antecedência, com poucos ou nenhum parâmetro adaptável. Este cenário é encontrado em ambientes especiais, como por exemplo, sistemas de robôs industriais ou estacionamentos.

Porém, há grandes expectativas para aplicações da robótica onde o ambiente é totalmente desconhecido. Pelo menos, neste caso, a autonomia é uma característica essencial para sistemas de navegação, em particular, e para sistemas de controle em geral [1] e [2].

A aprendizagem de sistemas de navegação autônoma inteligente é baseada nas iterações com os ambientes de navegação. A aprendizagem ocorre se o sistema de navegação não exibe um desempenho aceitável ou se novas experiências ocorrem. Essas premissas são investigadas em uma simples classe de problemas de navegação: o ambiente é desconhecido, há somente um alvo no ambiente e o robô precisa alcançar o alvo evitando colisões contra obstáculos [3] e [4]. Muitos sistemas autônomos inteligentes mostram resultados considerando esse problema de navegação. Redes neurais são adotadas para projetar um sistema

autônomo que aprende controlar variáveis de controle de navegação: velocidade e direção [1]. O sistema proporciona duas classes de campos sensoriais: direção do alvo e distância aos obstáculos. Um controlador mais complexo de navegação é baseado em técnicas evolucionárias (sistema de classificação). Resultados de simulação mostram que o sistema aprende, simultaneamente, a desviar de obstáculos, capturar alvos e coordenar esses comportamentos quando são conflitantes [5].

Este trabalho considera o mesmo problema de navegação definido anteriormente e propõe um controlador de navegação autônoma baseado em redes neurais modular. O sistema aprende de acordo com o novo mecanismo de aprendizagem por reforço para ajustar camadas em diferentes módulos. As ocorrências de captura de alvos e colisões determinam os instantes em que o processo de aprendizagem é realizado. Resultados de simulação mostram que, após um período de aprendizagem, o sistema autônomo coordena adequadamente os comportamentos inatos (busca a alvo e desvio de obstáculos). Avalia-se o sistema de navegação autônomo comparando-o com uma proposta distinta.

Após a fase de aprendizagem, quando os sistemas alcançam os alvos distantes dos obstáculos, alvos são posicionados próximos a estes. Somente o sistema proposto é capaz de ajustar sua estratégia de navegação e alcançar os alvos em situações de risco.

Este artigo é organizado como segue. Seção 2 define os modelos do robô e do ambiente de navegação. Também são apresentados os fundamentos de técnicas de redes neurais. O sistema de navegação autônomo é descrito na Seção 3. Seção 4 mostra resultados de simulação, inclusive uma comparação para propósitos de avaliação de desempenho. Uma breve discussão sobre os resultados obtidos e possibilidades para trabalhos futuros são apresentados na Seção 5.

2. Modelos e fundamentos

2.1. Modelo do robô

O modelo do robô não apresenta dinâmica interna (Figura 1). Sua velocidade é constante (1 unidade de distância por iteração) e sua direção pode assumir valores entre -15° e 15° .

O robô utilizado nas simulações do sistema proposto contém três classes de campos sensoriais com o objetivo de obter informações provenientes do ambiente. Tais classes são: 1) Os sensores de obstáculo estão dispostos na parte frontal do robô (de -90° a 90°) em um número de 50 sensores com a finalidade de obterem a distância entre o robô e o obstáculo mais próximo que estiver alinhado em sua direção. A saída é um valor pertencente ao intervalo $[0, 1]$; 2) Os sensores de direção ao alvo estão distribuídos ao longo do perímetro do robô em uma quantidade de 100 sensores. A saída de cada sensor está relacionada com o ângulo entre sua direção e o alinhamento do robô com o alvo, assumindo valores no intervalo $[0, 1]$; 3) Os sensores de distância ao alvo fornecem ao sistema a distância entre o robô e o alvo. Um único sensor captura a distância desejada. Tal distância é mapeada, para um valor dentro do intervalo $[0, 1]$, que é, então, convertido para todos os 51 sensores de distância ao alvo através de uma função gaussiana.

Há dois conjuntos de sensores para detectar colisões contra obstáculos e capturas de alvos denominados sensor de colisão e de captura, respectivamente. A saída desses sensores é binária, 1 se o evento (colisão e/ou captura) ocorrer e 0 caso contrário.

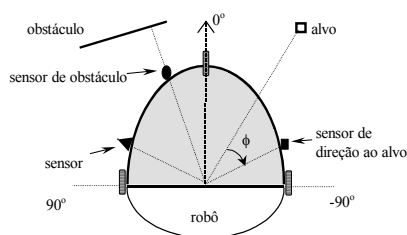


Figura 1: Um esboço do modelo do robô

2.2. Redes neurais

Redes neurais são inspiradas no sistema neural biológico e reproduz qualitativamente muitas de suas características, tais como: paralelismo, aprendizagem, adaptação, generalização e aproximação [6]. As redes neurais são umas das mais interessantes técnicas para aplicações onde a capacidade de aprendizagem é essencial. Este é o caso considerado, já que o ambiente é desconhecido e o sistema de navegação, inicialmente, não tem o controle eficiente de estratégias para guiar o robô ao alvo.

Há três principais estratégias de aprendizagem para redes neurais: não-supervisionada, supervisionada e por reforço. A estratégia de aprendizagem não-supervisionada diz respeito à somente entradas, ou seja, redes neurais se auto-organizam de acordo com as características das entradas. Desde que aplicações de controle são associadas com um mapeamento entrada-saída, a estratégia aprendizagem não-supervisionada não é adequada, a menos que juntamente com esta seja considerada outra estratégia. A aprendizagem supervisionada depende do armazenamento de

conhecimento do problema, dado que é necessário um conjunto de pares de entrada-saída corretos para efetuar o processo de aprendizagem. Esta estratégia pode ser adequada para diversas aplicações de controle, mas não para o controle autônomo. Este é o caso em que o ambiente modifica-se tornando inválidos os pares de entrada-saída usados durante o período de aprendizagem.

A estratégia de aprendizagem por reforço fornece mecanismos para a aprendizagem e adaptação de uma rede neural utilizada em aplicações de controle autônomo. De acordo com esta abordagem, os sistemas autônomos aprendem enquanto interagem com o ambiente. Em todos os instantes em que o sistema depara-se com eventos discrepantes (situações desconhecidas, situações de sucesso ou falha); o sistema recebe um sinal de avaliação (reforço) associado ao seu desempenho. O processo de aprendizagem é baseado neste sinal de avaliação.

A aprendizagem por reforço em sistemas artificiais pode ser entendido como uma metáfora para o processo de condicionamento de aprendizagem em sistemas biológicos [7]. Brevemente, o condicionamento operante é descrito em termos de estímulos não-condicionados (sinais de entrada), e reflexos não-condicionados e condicionados (sinais de saída ou ação): 1) o sistema de aprendizagem, seleciona um reflexo condicionado que modifica o ambiente; 2) o ambiente gera um estímulo não-condicionado obtido por um campo sensorial do sistema; 3) o sistema gera um reflexo não-condicionado associado com um estímulo não-condicionado (sinal de saída) e 4) os parâmetros do sistema são ajustados para aumentar ou diminuir a probabilidade do reflexo condicionado, selecionado no passo 1 acima.

Este trabalho adota o princípio proposto por Edelman: a melhor maneira para adquirir conhecimento sobre sistemas autônomos é investigar modelos artificiais inspirados em sistemas biológicos [8]. Este princípio é o argumento mais importante para justificar a abordagem adotada: redes neurais e aprendizagem por reforço. Ambas as técnicas de inteligência artificial possuem uma fundamentação biológica e são baseadas na neurofisiologia e na psicologia do comportamento.

3. Sistema de navegação autônoma

O sistema de navegação autônoma consiste de três módulos principais, conectados a um neurônio de saída. Dois destes módulos, Busca ao Alvo (BA) e Desvio de Obstáculos (DO) geram comportamentos instintivos. Um módulo de coordenação estabelece (após o processo de aprendizagem) pesos adequados para ponderar os comportamentos gerados pelos módulos BA e DO e combiná-los no neurônio de saída (Figura 2).

3.1. Módulos neurais DO e BA

Os módulos BA e DO são redes neurais com o conhecimento do respectivo comportamento já armazenado, desprovido totalmente de aprendizagem. Estes módulos geram comportamentos inatos de busca de alvos e de desvio de obstáculos.

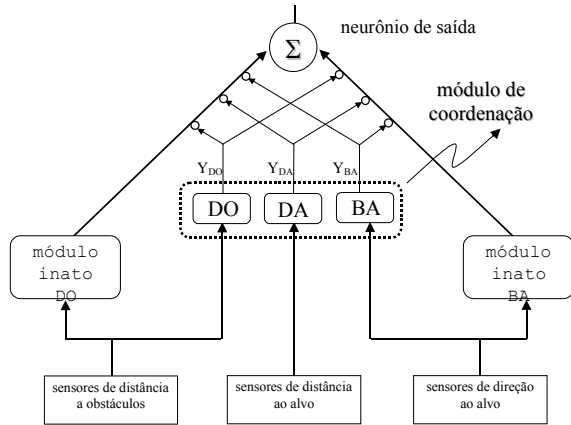


Figura 2: Sistema de navegação autônoma.

O módulo DO gera o comportamento inato de desvio de obstáculos. As entradas para este módulo são os sinais emitidos pelos sensores de obstáculos e suas correspondentes direções. Assim, o módulo tem a capacidade de ajustar a direção do robô adequadamente para evitar colisões. Isoladamente, este módulo garante o comportamento de desvio de obstáculos, não possuindo a capacidade de capturar alvos. O módulo BA gera o comportamento inato de busca de alvos. Para tanto, o módulo precisa dos sinais provenientes dos sensores de direção ao alvo, e de acordo com a direção do sensor que estiver mais alinhado com ao alvo, o ajuste da direção do robô é efetuado. Se este módulo for executado isoladamente, somente o comportamento de busca ao alvo é desenvolvido. Caso haja um obstáculo entre o robô e o alvo, a colisão se torna inevitável. Para manter a analogia com os sistemas biológicos, as respectivas redes neurais destes módulos são configuradas de acordo com uma abordagem evolucionária [9].

3.2. Módulo de coordenação neural

Se o sistema de navegação consistisse somente dos módulos inatos e se eles operassem juntos sem nenhuma coordenação, haveria muitas situações conflitantes e o desempenho da navegação seria certamente insatisfatória. A função do módulo de coordenação é coordenar os comportamentos instintivos gerados pelos módulos inatos DO e BA.

O Módulo de Coordenação é formado por três redes neurais nebulosas não-supervisionadas: Desvio de Obstáculos (DO), Busca a Alvo (BA) e Distância ao Alvo (DA). As redes são conectadas a diferentes tipos de sensores: sensores de distância de obstáculos,

direção de alvos e distância de alvos, respectivamente (Figura 2).

Após o processo de aprendizagem, o módulo de coordenação pondera os sinais provenientes dos módulos inatos, balanceando os comportamentos instintivos. As saídas das redes do módulo de coordenação estabelecem os pesos sinápticos para o neurônio de saída. O método de aprendizagem por reforço é adotado para estas redes.

Rede Neural de Desvio de Obstáculo (DO). A arquitetura da rede consiste de duas camadas de neurônios nebulosos (neurônios modelados de acordo com a Teoria Nebulosa [10]). A primeira camada é construtiva, ou seja, nos instantes que o ciclo de aprendizagem é acionado, alguns neurônios são inseridos na primeira camada. Cada neurônio da primeira conecta com todos os sensores de distância de obstáculos. Há somente um neurônio na segunda camada, que é conectado com todos os neurônios da primeira camada (Figura 3).

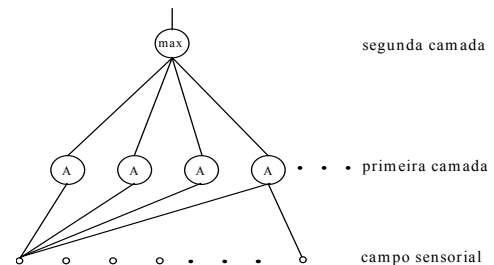


Figura 3: Arquitetura da rede de coordenação.

De acordo com o modelo neurônio definido para a primeira camada, as entradas x_i e os pesos sinápticos w_i , $i = 1, 2, \dots, N$, são valores modais para os conjuntos nebulosos (modelados de acordo com funções gaussianas) Ω_{x_i} e Φ_{w_i} , respectivamente (dispersões das funções são constantes).

Durante o processamento neural, estes conjuntos são manipulados usando o operador medida de possibilidade $\Delta(\cdot)$. Então a t-norma produto $T(\cdot)$ é aplicada sobre seus argumentos: o resultado de $\Delta(\cdot)$ e o peso s_i (há dois pesos sinápticos para cada entrada do neurônio: w_i e s_i). Após esse processo, o operador $A(\cdot)$ encontra o máximo sinal sináptico e a função identidade $f(x)=x$ determina a saída. A eq. (1) mostra a saída y_k do k-ésimo neurônio da primeira camada [9].

$$y_k = f \left(\bigwedge_{i=1}^N \left(T \left(\Delta \left(\Omega_{x_i}, \Phi_{w_i} \right), s_i \right) \right) \right) \quad (1)$$

em que N é o número de sensores conectados ao neurônio.

As entradas do neurônio da segunda camada provêm de todos os neurônios da primeira camada. A saída desse neurônio, y_{DO} , é também a saída da rede DO e

corresponde ao valor máximo dentre todas as suas entradas (eq. (2)).

$$y = \underset{k=1}{\overset{Q}{MAX}}(y_k) \quad (2)$$

em que Q é o número de neurônios da primeira camada e MAX é o operador máximo.

A rede DO influencia no balanceamento de cada comportamento instintivo do robô (busca de alvo e desvio de obstáculo), baseado-se no mapa do ambiente (vetor de distância do(s) obstáculo(s) fornecido pelos sensores de distância ao alvo). Assim, a rede aprende a associar classes de mapas do ambiente com situações de eminência de colisão.

As colisões definem instantes t_c quando o processo de aprendizagem é realizado, ou seja, quando os pesos sinápticos w_i e s_i são ajustados de acordo com a eq. (3).

$$\begin{aligned} w_{ij}(t+1) &= w_{ij}(t) + (z(t) - w_{ij}(t)) / (1 + s_{ij}(t)) \\ s_{ij}(t+1) &= s_{ij}(t) + (\mu - s_{ij}(t)) / (\alpha + 1) \end{aligned} \quad (3)$$

em que o termo $z(t)$ é a entrada no instante t , $z(t_c) \in \{x_i(t) / t = t_c, i = 1, 2, \dots, N\}$, tal que $\mu = \Delta(\Phi_{z(t)}, \Phi_{w_i}) \geq \Delta(\Omega_{x_i(t)}, \Phi_{w_i}), i = 1, 2, \dots, N$, $\Delta(\cdot)$ é o operador de medida de possibilidade, α é a taxa de aprendizagem e $x_i(t)$ é o i -ésimo componente do vetor de distâncias de obstáculos.

Rede Neural de Busca ao Alvo (BA). As idéias adotadas para conceber a rede BA são similares ou análogas àquelas mencionadas para a rede DO. As diferenças entre as redes são mostradas a seguir. A arquitetura (modelo dos neurônios, camadas, e conexões) é idêntica a rede DO, mas para a rede BA, cada neurônio da primeira camada é conectado com o campo sensorial de direção ao alvo

A rede BA aprende a associar as classes de sinais de direção ao alvo com o comportamento adequado de busca ao alvo. O processo de aprendizagem é ativado a todo o momento t_a quando o robô captura um alvo (o sensor de captura detecta este tipo de evento). Nestes instantes, os pesos sinápticos w_i e s_i são ajustados de acordo com a eq. (3), considerando que o termo $z(t)$ é uma entrada no instante $t = t_a$, $z(t_a) \in \{x_i(t) / t = t_a, i = 1, 2, \dots, N\}$, tal que $\mu = \Delta(\Phi_{z(t)}, \Phi_{w_i}) \geq \Delta(\Omega_{x_i(t)}, \Phi_{w_i}), i = 1, 2, \dots, N$, e $x_i(t)$ é o i -ésimo componente do vetor de direção de alvos.

Rede Neural de Distância de Alvos (DA). A arquitetura da rede DA (modelo dos neurônios, camadas, e conexões) e seu processamento difere das rede DO e BA somente em relação às entradas da rede. Cada neurônio da primeira camada conecta-se com o campo sensorial de distância ao alvo.

A rede neural DA aprende a associar classes de sinais de distância ao alvo com momentos em que o robô deve adotar o comportamento de busca ao alvo. O processo de aprendizagem é ativado a todo o momento t_a quando o robô captura um alvo. Nestes instantes, os pesos sinápticos w_i e s_i são ajustados de acordo com a eq. (3), considerando que o termo $z(t)$ é uma entrada no instante $t = t_a$, $z(t_a) \in \{x_i(t) / t = t_a, i = 1, 2, \dots, N\}$, tal que $\mu = \Delta(\Phi_{z(t)}, \Phi_{w_i}) \geq \Delta(\Omega_{x_i(t)}, \Phi_{w_i}), i = 1, 2, \dots, N$, e $x_i(t)$ é o i -ésimo componente do vetor de distância ao alvo.

Neurônio de Saída. O modelo do neurônio de saída é definido de acordo com a expressão geral dada pela eq. (4).

$$y = f\left(\sum_{i=1}^N w_i^* x_i\right), \quad w_i^* = w_i / \sum_{j=1}^N w_j \quad (4)$$

em que x_i e $w_i, i = 1, 2, \dots, N$, são as entradas e os pesos sinápticos, respectivamente, e $f(\cdot)$ é a função identidade.

O neurônio de saída recebe duas entradas ($N = 2$), $x_1 = U_{DO}$ e $x_2 = U_{BA}$, dos módulos inatos DO e BA, respectivamente (veja Figura 2). O neurônio de saída agrega as entradas e, conseqüentemente, os comportamentos instintivos. O módulo de coordenação determina os pesos sinápticos para cada entrada do neurônio, como é mostrado a seguir (eq. (5)).

$$\begin{aligned} w_1 = w_{DO} &= \varepsilon + w_{DO} y_{DO} - (\bar{w}_{BA} y_{BA}) (\bar{w}_{DA} y_{DA}) \\ w_2 = w_{BA} &= \varepsilon - \bar{w}_{DO} y_{DO} + (w_{BA} y_{BA}) (w_{DA} y_{DA}) \end{aligned} \quad (5)$$

em que ε é uma constante ($\varepsilon = 0.05$); \bar{w}_{DO} , w_{BA} , e w_{DA} são pesos sinápticos associados com as conexões axônio-axônio entre as redes neurais DO, BA e DA e o módulo inato BA, respectivamente; w_{DO} , \bar{w}_{BA} , e \bar{w}_{DA} são de forma análoga os pesos sinápticos referentes ao módulo inato DO (Figura 2).

Os pesos sinápticos na eq. (5) são ajustados de acordo com a eq. (6a) e eq. (6b) nos instantes $t = t_c$ e $t = t_a$, respectivamente.

$$\begin{cases} w_{DO}(t+1) = \zeta(1 - w_{DO}(t))f(y_{DO})g(U_{DO}) \\ \bar{w}_{DO}(t+1) = \zeta(1 - \bar{w}_{DO}(t))f(y_{DO})g(U_{BA}) \end{cases} \quad (6a)$$

$$\begin{cases} w_{BA}(t+1) = \eta(1 - w_{BA}(t))f(y_{BA})g(U_{BA}) \\ \bar{w}_{BA}(t+1) = \eta(1 - \bar{w}_{BA}(t))f(y_{BA})g(U_{DO}) \\ w_{DA}(t+1) = \eta(1 - w_{DA}(t))f(y_{DA})g(U_{BA}) \\ \bar{w}_{DA}(t+1) = \eta(1 - \bar{w}_{DA}(t))f(y_{DA})g(U_{DO}) \end{cases} \quad (6b)$$

em que $\zeta = 0.015$ e $\eta = 0.03$ são fatores de aprendizagem.

Além dos ajustes dos pesos sinápticos, em todo o momento que ocorre uma colisão ou uma captura (t_c ou

ta), a primeira camada das redes DO, BA e DA pode ter suas arquiteturas alteradas. Portanto, são redes neurais construtivas. Um sistema nebuloso [11] determina, a partir de três parâmetros, entre três possibilidades o estado da arquitetura das redes. Tais estados são: sem alteração, inserção de neurônio ou remoção de neurônio. Devido às limitações de espaço, o sistema nebuloso não é descrito neste artigo. Quando um neurônio é inserido, seus pesos sinápticos recebem os sinais do respectivo campo sensorial da rede, ou seja, $w_i = x_i(t_c)$ e $s_i = 0.5$ (rede DO); e $w_i = x_i(t_a)$ e $s_i = 0.5$ (rede BA e DA).

Embora os modelos axônio-dendrito sejam geralmente considerados em redes neurais, a conexão axônio-axônio é adotada para o modelo das sinapses do neurônio de saída. Ambos os tipos de conexão têm uma contrapartida em sistemas neurais biológicos. Observe na eq. (5) que as redes neurais DO, BA e DA cooperam/competem para estabelecer a correta influência em cada comportamento instintivo sobre o desempenho do robô.

4. Resultados

Esta seção mostra resultados de comparação entre o sistema de navegação descrito anteriormente e um sistema de navegação distinto proposto em [9]. Estes sistemas são chamados, a partir daqui, de CS e FS, respectivamente. O sistema FS é também um sistema neural e adota os mesmos comportamentos inatos descritos para CS, denominado de desvio de obstáculos e busca a alvos [9].

A Figura 4 apresenta o ambiente definido para a avaliação. O retângulo preenchido representa um obstáculo, os círculos são os alvos e o triângulo o robô. As posições dos alvos são definidas em duas classes: alvos próximos a obstáculos (classe P) e alvos que estão distante dos obstáculos (classe D). Durante a simulação, somente um alvo está presente no ambiente. Após a captura, o alvo (capturado) é eliminado e um novo alvo é inserido no ambiente.

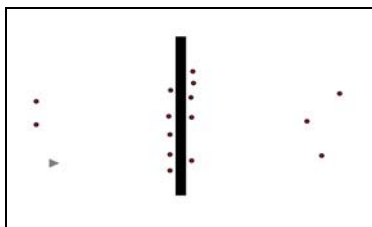


Figura 4: Ambiente de navegação para avaliação.

Os experimentos são realizados em duas fases. A primeira é a fase de aprendizagem, em que somente os alvos da classe D são considerados. A segunda é a fase de avaliação, nesta fase os alvos considerados são os pertencentes à classe P.

Inicialmente o sistema de navegação não conhece a estratégia de navegação. Durante a primeira fase, o robô interage com o ambiente recebendo estímulos de colisão e captura (Figura 5). O sistema de navegação aprende o conhecimento básico para exercer suas habilidades, mas somente para situações onde os alvos estão distantes do obstáculo.

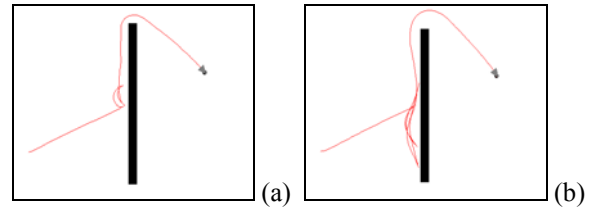


Figura 5: A captura do primeiro alvo após algumas colisões: (a) sistema FS e (b) sistema CS.

A Figura 6 mostra a evolução da aprendizagem para os dois sistemas. Observe que o número de colisões e o número de iterações necessárias para o sistema CS alcançar todos os alvos (classe D) são maiores que aqueles mensurados para o sistema FS.

Durante a segunda fase, somente alvos muito próximos do obstáculo são inseridos no ambiente. Conseqüentemente as situações são completamente diferentes daquelas mostradas na fase anterior. A Figura 7 mostra que somente o sistema CS foi capaz de se adaptar (aprender) às novas situações. As colisões sofridas por este sistema foram mais numerosas, mas todos os alvos (classe P) foram capturados. Ao final da segunda fase o robô estava apto para alcançar os alvos sem sofrer colisões. Por outro lado, o sistema FS, não alcançou alvos próximos ao obstáculo.

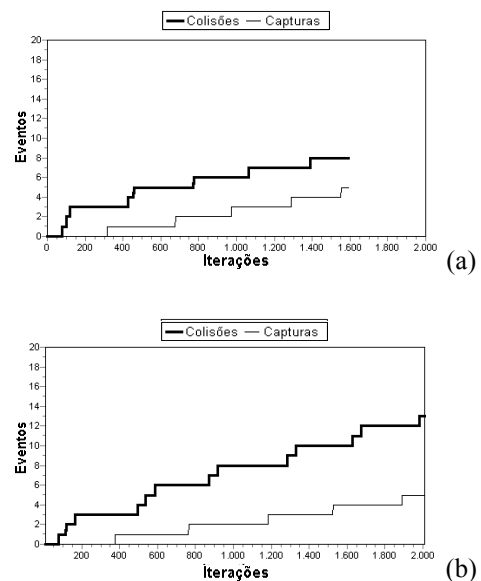


Figura 6: Evolução da fase de aprendizagem: (a) sistema FS e (b) sistema CS

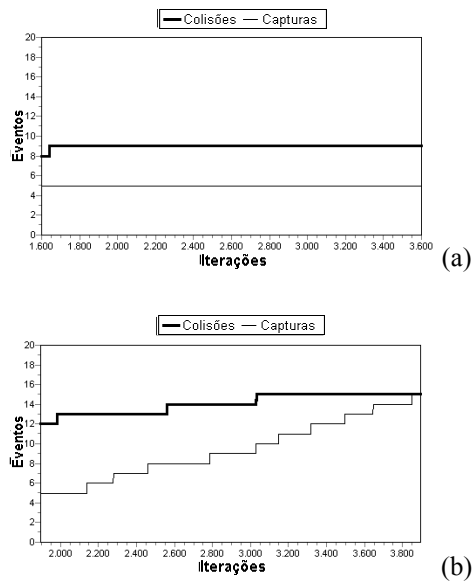


Figura 7: Evolução da fase de avaliação: (a) sistema FS e (b) sistema CS

4. Conclusão

A expectativa é muito grande para aplicações com robôs móveis em que o ambiente é desconhecido. Nestes casos, autonomia é uma característica essencial para sistemas de navegação de robôs. Este trabalho descreve um sistema de navegação autônomo baseado em redes neurais modulares. O sistema possui dois módulos que geram comportamentos inatos: desvio de obstáculo e busca ao alvo. Inicialmente o sistema não tem a capacidade de balancear estes comportamentos instintivos e conflitantes. A partir de sua aprendizagem, o sistema torna-se capaz de coordenar os comportamentos inatos. A estratégia de aprendizagem por reforço é adotada, proporcionando condições para que o sistema aprenda a partir da interação com o ambiente, inicialmente desconhecido. Simulações computacionais mostram que o sistema aprimora sua estratégia de navegação à medida que eventos de colisão e de captura ocorrem. O sistema de navegação generaliza o conhecimento adquirido em um ambiente particular tornando o robô capaz de navegar em diferentes ambientes com um desempenho idêntico ou similar. O sistema proposto é avaliado em um ambiente que força a mudança de estratégia de navegação. Um segundo sistema é considerado para efeitos de comparação. Inicialmente os dois sistemas aprendem suas respectivas estratégias de navegação para alvos distantes dos obstáculos. Tais estratégias tornam-se inadequadas para alvos próximos a obstáculos, correspondentes a situações de risco. Somente o sistema proposto se tornou capaz de ajustar a estratégia de navegação e capturar alvos próximos ao obstáculo.

5. Agradecimentos

Os autores agradecem ao CNPq pelo apoio financeiro ao Projeto PIBIC/CNPq 2001/2002.

Referências

- [1] Crestani, P., Figueiredo, M., e Von Zuben, F. A Hierarchical Neuro-Fuzzy Approach to Autonomous Navigation. *Proceedings of 2002 International Joint Conference on Neural Networks*, (cd-rom), Honolulu, USA, 2002.
- [2] Antsaklis, P. J. Intelligent Learning Control. *IEEE Control Systems*, Vol. 15, (3), pp. 5-7, 1995.
- [3] Colombetti, M., Dorigo, M. e Borghi, G. Behavior Analysis and Training – A Methodology for Behavior Engineering, *IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics*, Vol. 26, (3), pp. 365-380, 1996.
- [4] Vershure, P. 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*, Vol. 11, pp. 1531-1549, 1998.
- [5] Cazangi, R. e Figueiredo, M. Simultaneous Emergence of Conflicting Basic Behaviors and Their Coordination in an Evolutionary Autonomous Navigation System. *Proceedings of 2002 Congress on Evolutionary Computation*, (cd-rom), Honolulu, USA, 2002.
- [6] Haykin, S. *Neural Networks*, Prentice Hall, New York, USA, 1994.
- [7] Donahoe, J. W. e Palmer, D. C. *Learning and complex behavior*. Allyn and Bacon, Needham Heights, USA, 1994.
- [8] Edelman, G. *Neural Darwinism: The Theory of neuronal group selection*. Basic books, New York, USA, 1987.
- [9] Figueiredo, M. e Gomide, F. Evolving neurofuzzy networks for basic behaviors and a recategorization approach for their coordination. In: *Genetic Algorithms and soft Computing* (Herrera, F, and Verdegay, J. (Ed)), pp 533-552, Springer Verlag, USA, 1996.
- [10] Gomide, F. e Pedrycz, W. An introduction to fuzzy sets: analysis and design. *The MIT Press*, Cambridge, USA, 1998.
- [11] Calvo, R. e Figueiredo M. Fuzzy System for Architecture Learning in a Constructive Neurofuzzy Network, unpublished, 2003.