

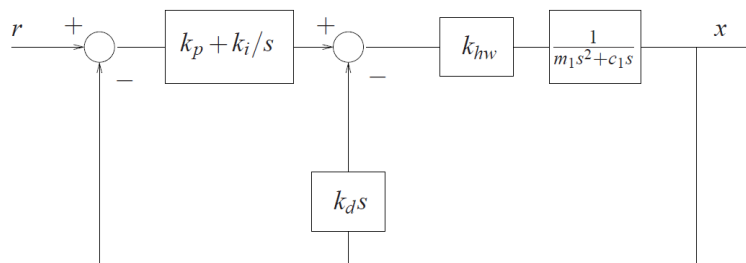
EA072 –Exercícios de Fixação de Conceitos 2 (EFC2)

Algoritmos Evolutivos + Navegação Autônoma + Sistemas de Recomendação

Peso da Lista: 4 || Data de entrega: 19/11/2015

1 Síntese de controle PID

Nos laboratórios de controle de processos da FEEC/Unicamp, como EA619 e EA722, estudam-se projetos de controle PID (controle proporcional-integral-derivativo em malha fechada), onde é necessário definir os valores de k_P , k_D e k_I para se maximizar um determinado critério de desempenho. No controle PID, a ação de controle vai ser proporcional ao erro de saída da planta, à derivada deste mesmo erro e à integral deste mesmo erro. São vários critérios de desempenho que podem ser considerados no projeto, como tempo de subida, tempo de acomodação, sobressinal, erro de regime, margem de ganho e margem de fase. Não é simples definir simultaneamente os valores de k_P , k_D e k_I , pois há interferência entre eles. Mesmo assim, há formas sistemáticas de projeto, que são estudadas em EA721 (curso teórico associado ao laboratório) e em EA722. Nesta atividade do curso EA072, iremos supor que o projetista não domina conceitos de projeto de controle PID, mas tem a tarefa de projetar um controlador PID visando maximizar dois critérios de desempenho. Para tanto, o projetista está convidado a empregar uma versão de algoritmo genético em Matlab, fornecida pelo professor, sendo que supõe-se que os parâmetros k_P , k_D e k_I excursionam no intervalo $[0, +5]$ e que a planta é o Sistema Retilíneo presente no Laboratório de EA722 (é um dos 4 tipos de plantas presentes nas bancadas do laboratório), na configuração das Experiências 3 e 4 e com parâmetros conhecidos e previamente fornecidos. A configuração de controle mostrada na figura abaixo, na verdade, é a denominada PI&D, pois o ganho k_D multiplica a derivada do sinal de saída da planta, e não do erro.



Atividade prática:

- (a) O programa emprega os operadores genéticos de mutação não-uniforme, crossover aritmético e crossover uniforme, além do operador de seleção por torneio de 3 indivíduos. Explique como funcionam esses 4 operadores.

- (b) Qual é o tamanho da população? Qual é o número máximo de gerações? Como é criada a população inicial? Quais são as taxas de mutação e crossover? Com que frequência cada operador de crossover é escolhido?
- (c) Nas linhas 26 e 111 do programa [prog_PID.m] fornecido pelo professor, insira o comando [fitness(i,1) = 100/(10*(S.RiseTime+1));], indicando que o objetivo é encontrar valores para k_P , k_D e k_I que minimizam apenas o tempo de subida da resposta ao degrau do sistema de controle PID. Execute o programa pelo menos 5 vezes, com a parametrização fornecida, reporte e comente os resultados. Procure novas configurações de parâmetros (os parâmetros do algoritmo são os mencionados no item (b) acima) que possam conduzir a melhores resultados.
- (d) Nas linhas 26 e 111 do programa [prog_PID.m] fornecido pelo professor, insira o comando [fitness(i,1) = 100/(10*(S.RiseTime+1)+((Pm-60)^2+1));], indicando que o objetivo é encontrar valores para k_P , k_D e k_I que minimizam o tempo de subida da resposta ao degrau do sistema de controle PID e que levem a margem de fase a valores no entorno de 60 graus. Repita as atividades do item (c), incluindo uma comparação entre os resultados dos itens (c) e (d).

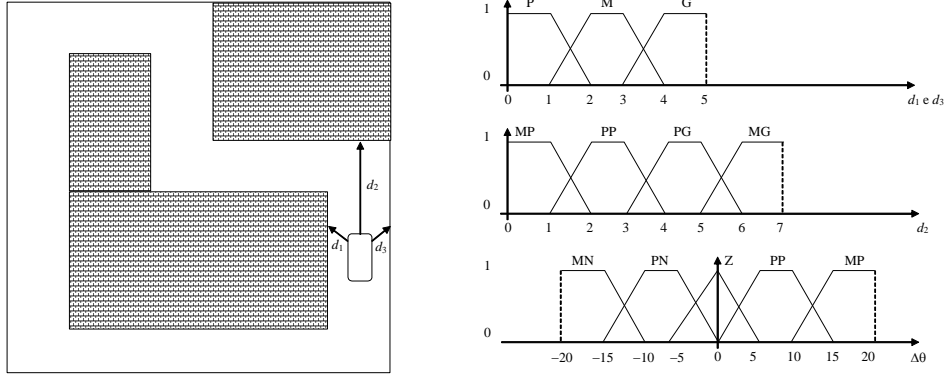
Nota: Nas análises e comentários dos itens (c) e (d), vai ser necessário empregar alguns conceitos de controle de processos.

2 Aproximação de funções multidimensionais

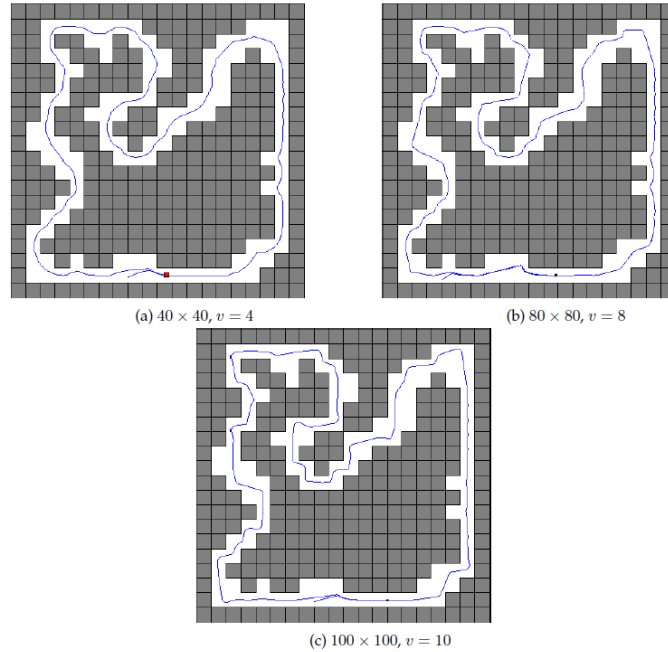
Gere um mapeamento $\mathcal{R}^1 \rightarrow \mathcal{R}^1$, ou seja, $y = f(x)$, amostre dados para um intervalo restrito de entrada e empregue o software Eureqa, disponível em <http://www.nutonian.com/products/eureqa/>, para propor alternativas de funções de aproximação para a sua $f(\cdot)$. Faça o mesmo para um mapeamento $\mathcal{R}^3 \rightarrow \mathcal{R}^1$, ou seja, $y = f(x_1, x_2, x_3)$. Controle o intervalo de cada variável de entrada, produza mapeamentos desafiadores (mas evite muita composição de funções, ou seja, funções que são argumentos de outras funções), trabalhe com um bom número de amostras e insira ruído junto às amostras. Escolha adequadamente o conjunto de funções-base que serão utilizadas na busca. Apresente a fronteira de Pareto obtida, com soluções de compromisso entre acurácia e simplicidade da solução e escolha uma solução com boa relação de custo-benefício entre esses dois critérios (acurácia e simplicidade).

3 Controle nebuloso e robótica evolutiva

Um veículo autônomo deve ser controlado por um sistema nebuloso, ou seja, um sistema baseado em regras nebulosas, e, para tanto, ele deve receber como entrada as informações de 3 sensores de distância (que produzem valores para d_1 , d_2 e d_3) e fornecer na saída a variação no ângulo do veículo ($\Delta\theta$). As funções de pertinência associadas aos termos linguísticos que definem as partições das variáveis linguísticas d_1 , d_2 , d_3 e $\Delta\theta$ são fornecidas a seguir, à direita. Proponha um consequente adequado para cada uma das 36 regras $\langle \underline{SE} \ d_1 \text{ é } ? \ \underline{E} \ d_2 \text{ é } ? \ \underline{E} \ d_3 \text{ é } ? \ \underline{ENTÃO} \ \Delta\theta \text{ é } ? \rangle$.



Escolha um labirinto com dimensões compatíveis com a escala apresentada para d_1 , d_2 e d_3 , tome o robô como um ponto e considere que os sensores laterais estão a 45° do sensor central. Apresente simulações e analise os resultados obtidos. Os resultados de atuação do controle nebuloso devem produzir trajetórias como as ilustradas a seguir.



Em seguida, utilize o mesmo ambiente e o mesmo robô, mas substitua o controlador do robô por uma rede neural MLP com uma camada intermediária. Utilize um algoritmo evolutivo para obter os pesos da rede neural que levem o robô a navegar sem colisão. Para obter o *fitness* de cada controlador candidato, utilize um elenco de condições iniciais que impliquem desafios distintos para a navegação e defina o *fitness* como o inverso do (número de colisões mais um). O veículo tem velocidade constante (defina o valor de $v \cdot \Delta t$) e a atualização de posição do veículo deve se dar na forma:

$$\begin{cases} x_{t+1} = x_t + v \cdot \cos(\theta_{t+1}) \cdot \Delta t \\ y_{t+1} = y_t + v \cdot \sin(\theta_{t+1}) \cdot \Delta t \end{cases}, \text{ com } \theta_{t+1} = \theta_t + \Delta\theta_t.$$

4 Sistema de recomendação empregando k -NN

Em muitos contextos de aplicação, é comum a existência de uma matriz cujas linhas representam clientes de uma empresa prestadora de serviço (ou fornecedora de produtos) e cujas colunas representam serviços (produtos). Cada elemento da matriz é um campo que admite uma nota que o cliente dá ao serviço (produto) da empresa. A tendência é que essas matrizes sejam esparsas, particularmente em casos em que existem muitos clientes e muitos serviços (produtos). Ou seja, cada cliente deve ter utilizado (adquirido) e avaliado um subconjunto pequeno dentre todos os serviços (produtos) disponíveis. Os campos associados a pares (cliente, serviço) (ou (cliente, produto)) que ainda não foram preenchidos representam serviços (produtos) candidatos à recomendação para o cliente. No entanto, a recomendação só deve ser feita caso se estime que a avaliação que o cliente possa vir a atribuir ao serviço (produto) for alta. Dispondo de tal matriz, contendo algumas poucas avaliações de serviços (produtos) por parte dos clientes e muitos campos com conteúdo faltante, um algoritmo capaz de realizar a imputação de qualquer dado faltante da matriz é o k -NN (k nearest neighbour), o qual se encontra entre os 10 algoritmos mais utilizados em aprendizado de máquina. Serviços (Produtos) que receberem imputações de valor elevado para um certo cliente, devem ser recomendados para aquele cliente. Uma empresa pioneira em sistemas de recomendação foi a Amazon, mas hoje praticamente todas as empresas de médio a grande porte dispõem de sistemas de recomendação, por exemplo, visando personalizar o conteúdo das páginas acessadas, buscando assim maximizar a satisfação de seus clientes. Com base no paper “State-of-the-art Recommender Systems”, disponível na página do curso, explique como funciona o algoritmo k -NN em suas versões *user-based* e *item-based*. Para poder recomendar o item j ao usuário i , é necessário encontrar uma forma de estimar o grau de satisfação do usuário i para com o item j . Empregando o toolbox em Matlab fornecido juntamente com o enunciado da questão, estude como opera o programa [exemplo.m] e o adapte para realizar a imputação de dados faltantes na matriz de dados [MovieLens]. **Nota: O professor agradece Rosana Veroneze pela autoria e cessão do toolbox para uso nesta atividade.**

5 Sistema de recomendação – Paradigma Alternativo

Uma forma distinta de proposição de um sistema de recomendação se dá quando, além da matriz de clientes por serviços (produtos), existe também um vetor de atributos associado a cada cliente e um vetor de atributos associado a cada serviço (produto). Como a matriz tende a ser esparsa, o uso dessa informação adicional que permite caracterizar clientes e serviços (produtos) pode conduzir a ganho de desempenho. Uma proposta nesta linha é o algoritmo SCOAL (Simultaneous Co-Clustering and Learning), proposto em Deodhar & Ghosh (2010) SCOAL: A framework for simultaneous co-clustering and learning from complex data. ACM Transactions on Knowledge Discovery from Data (TKDD), vol. 4, no. 3, paper no. 11. (1) Descreva o princípio de funcionamento do SCOAL; (2) Apresente as principais propriedades do arquivo MovieLens; (3) Após ler o arquivo [Readme.txt] do toolbox fornecido, execute o programa para o arquivo MovieLens, de modo a produzir 16 co-clusters. (4) Forneça os modelos lineares obtidos para cada um dos 16 co-clusters. (5) Explique o significado de MSE, Precision e Recall obtidos. **Nota: O professor agradece André Luiz Vizine Pereira pela autoria e cessão do toolbox para uso nesta atividade.**

6 Pesquisadores de apoio

Além do professor da disciplina, os alunos da disciplina podem solicitar apoio técnico e tirar dúvidas com o doutorando da FEEC:

- Marcos Medeiros Raimundo [marcosmrai(_at_)gmail.com]