

## Modelo para Previsão de Soluções a Problemas Pouco Estruturados.

Robinson Vida Noronha<sup>1</sup>, Clovis Torres Fernandes<sup>2</sup>

<sup>1</sup>Universidade Tecnológica Federal do Paraná

<sup>2</sup>Instituto Tecnológico de Aeronáutica

vida@utfpr.edu.br, clovis@ita.br

*Abstract: How to anticipate which relevant elements of domain could be selected by students to compose solutions to ill-structured problems is the main core of this paper. Markov Chain was used to solve this research question. This solution was simulated in a computer environment. The results of this simulation was contrasted with a real example.*

*Resumo: Como prever quais elementos do domínio poderiam selecionados pelos aprendizes para a composição de soluções a problemas pouco estruturados é o tema deste relato de pesquisa. Para solucionar essa questão, empregou-se Cadeias de Markov para identificar os elementos mais prováveis de serem selecionados. A resultado da simulação da solução apresentada é comparada com um exemplo criado em 1976 que não empregou a técnica descrita neste artigo.*

### 1- Introdução.

Ao encerrar uma atividade instrucional ou aula no ambiente virtual, por exemplo, um determinado professor fornece aos aprendizes como tarefa complementar:

- i) um documento texto introdutório que complementa o assunto da aula;
- ii) um conjunto de problemas que devem ser solucionados pelos aprendizes. Destaca-se que esses problemas possuem características bem específicas. Eles são problemas do tipo Pouco Estruturado;
- iii) um conjunto de elementos relevantes do domínio selecionados pelo professor. Esses elementos são organizados no formato de matriz ou tabela. Essa matriz recebe o nome de Matriz de Resposta ou simplesmente Matriz Resposta [Egan 1976].

Cada aprendiz, separadamente, deve ler o documento texto. Após isso, solucionar os problemas propostos. A solução para cada problema é formada pela seleção de alguns elementos da matriz e a justificativa dessa seleção.

Haveria uma forma desse professor se preparar previamente para essa atividade? Como esse professor poderia antecipar quais elementos da matriz podem ser selecionados ou omitidos das soluções?

Com a finalidade de preencher essa lacuna, este artigo apresenta um modelo procedimental para auxiliar o professor ou o autor da atividade instrucional na tarefa de identificar quais são esses elementos. O modelo descrito neste artigo utiliza Cadeias de

Markov [Feller 1968] para antecipar quais seriam os elementos da matriz mais prováveis de serem selecionados para a composição de uma solução.

Este artigo está dividido em 6 seções. A Seção 2 destaca que os elementos da matriz podem ser representado por vetores. A partir dessa representação, a Seção 3 descreve uma forma de adaptação dos conceitos da Cadeia de Markov. Na Seção 4, esse modelo é simulado computacionalmente a partir de um exercício extraído de Egan (1976). Finalmente a Seção 5 encerra o artigo tecendo conclusões.

## 2- Representação Estruturada dos Elementos da Matriz Resposta.

Um elemento qualquer da matriz  $F_n$  pode também ser representado por um vetor que indica a presença ou não de um conjunto de elementos relevantes do domínio  $KWK_j$  [Noronha 2007]. De maneira geral, considerando  $m$  KWKs de um domínio, um elemento  $n$  qualquer da matriz pode ser representado por

$$F_n = \alpha_1 * KWK_1 + \alpha_2 * KWK_2 + \alpha_3 * KWK_3 + \dots + \alpha_m * KWK_m$$

O fator  $\alpha_j$  corresponde a um índice que pode assumir os valores 0 para a ausência de ocorrência do elemento  $KWK_j$  ou 1 para uma simples ocorrência desse elemento.

## 3- Modelo de Cálculo da Probabilidade de Seleção dos Elementos da Matriz.

Este trabalho considera que a seleção dos elementos da Matriz Resposta é realizada de maneira sequencial pelo aprendiz. Ou seja, o aprendiz seleciona um primeiro elemento e depois outros elementos são sucessivamente selecionados e adicionados para compor a solução do aprendiz.

A similaridade dos conceitos entre dois elementos da matriz é representada por meio do Cálculo da Distância Euclideana -  $D_{ij}$ . Quanto maior for o valor de  $D_{ij}$ , maior será a similaridade conceitual entre esses dois elementos quaisquer  $F_i$  e  $F_j$  matriz, por exemplo. Logo, a probabilidade de um elemento  $F_j$  qualquer ser selecionado logo em seguida à seleção de um elemento  $F_i$  é definida neste trabalho por:

$$p_{i,j} = P(F_{i,j}) = \frac{D_{ij}}{\sum_{\substack{k=1 \\ k \neq i}}^m D_{ik}} \quad (Eq. 1)$$

Esse valor de probabilidade é calculada a partir da divisão do valor da distância euclideana entre as duas expressões de conceitos consideradas e o somatório de todas as distâncias euclidianas.

Qual seria a probabilidade do aprendiz selecionar um terceiro, um quarto ou um  $n$ -ésimo elemento para compor a sua solução? Para responder a essa questão, fez-se uso das Cadeias de Markov [Feller 1968].

O encadeamento de Markov permite calcular a probabilidade de uma sequência de eventos  $\{S_1, S_2, S_3, \dots, S_n\}$  ocorrerem. A probabilidade de um determinado encadeamento de eventos ocorrer pode ser determinada pela Eq. 2 a seguir.

$$p(S_1, S_2, S_3, \dots, S_n) = a_1 * p_{1,2} * p_{2,3} * \dots * p_{n-1,n} \quad (Eq. 2)$$

Os elementos  $p_{i,j}$  dessa expressão correspondem à probabilidade condicional de um evento  $S_j$  ocorrer em seguida ao evento  $S_i$ . No contexto deste trabalho, os eventos significam a seleção de um determinado elemento da matriz.

O índice  $a_i$  na expressão anterior representa a probabilidade do evento  $S_i$  ocorrer por primeiro. Como esse primeiro índice poderia ser determinado? O índice  $a_i$  pode ser calculado de duas maneiras: em relação ao documento texto e ao enunciado do texto do problema. As informações contidas no enunciado do problema e no texto introdutório podem influenciar na escolha desse primeiro elemento.

Para a definição do valor de  $a_i$ , um índice é definido, conforme apresentado pelas Eqs. 3 e 4 a seguir.

$$a_{i,Apresentação} = P(F_{i,Apresentação}) = \frac{n_i}{B_{Apresentação}} \quad (\text{Eqs. 3 e 4})$$

$$a_{i,Desafio} = P(F_{i,Desafio}) = \frac{n_i}{B_{Desafio}}$$

Onde  $n_i$  é a quantidade de ocorrências das KWKs pertencentes ao elemento da matriz e que também podem ser encontradas no documento texto e no enunciado do problema.  $B_{Apresentação}$  e  $B_{Desafio}$  correspondem respectivamente ao total de KWKs pertencentes ao documento texto e ao enunciado do problema.

Caso o aprendiz selecione o primeiro elemento da matriz de forma aleatória, o valor de  $a_i$  é dado por  $a_i = 1/m$ , onde  $m$  é a quantidade de elementos da matriz. Esse caso particular não foi considerado para a simulação computacional relatada na próxima seção deste artigo.

#### 4- Análise do Modelo Proposto

Nesta seção analisa-se alguns dados obtidos por meio de uma simulação computacional do modelo proposto no exemplo analisado a partir de Egan (1976). As KWKs definidas e utilizadas na simulação podem ser encontradas em [Noronha 2007].

Para essa simulação, desejou-se saber quais seriam os dois ( $w=2$ ), três ( $w=3$ ) e quatro ( $w=4$ ) elementos da matriz mais prováveis de serem selecionados para compor uma solução. A Tabela 1 resume os dados obtidos por essa simulação. Esses dados fornecem indicações de quais elementos da matriz seriam os mais prováveis de serem selecionados. Dessas indicações, pôde-se observar que:

- O elemento da matriz mais provável para a seleção é o terceiro elemento da matriz, ou seja o F3.
- A seqüência dos dois elementos mais prováveis para a seleção são: F3, F20.
- A seqüência dos três elementos mais prováveis para a seleção são { F3, F20} e qualquer elemento do conjunto { F9, F11, F12, F13, F18, F19}.

- A seqüência dos quatro elementos mais prováveis são: os elementos apresentados no ítem anterior acrescido de qualquer um dos elementos apresentados na última coluna da Tabela 1.

**Tabela 1 – Possíveis Sequências a Serem Utilizadas pelos Aprendizes.**

W=1	W=2	W=3	W=4
F3	F20	F9	F11, F12, F18, F19
		F11	F9, F18, F19
		F12	F9, F18, F19
		F13	F2, F4, F5, F6, F7, F8, F14, F15, 17
		F18	F9, F11, F12, F19
		F19	F9, F12, F18

Ao compararmos os elementos listados na Tabela 1 com os elementos da matriz indicados por Egan, como sendo os mais prováveis de serem selecionados, pelos aprendizes pode-se observar uma certa similaridade. Todos os elementos da Tabela 1 também estão antecipados por Egan (1976). Isso pode ser considerado um forte indício da validade deste modelo de previsão de soluções.

## 5- Conclusão e Trabalhos Futuros

Na abordagem de solução investigada utilizou-se Cadeia de Markov e a representação estruturada do domínio através de vetores.

Os resultados obtidos pela simulação computacional do processo descrito neste artigo, quando comparados com a atividade instrucional criada por Egan apresentam um elevado grau de similaridade. Egan registrou quais elementos da matriz seriam mais prováveis de serem selecionados para a composição de várias soluções. Esses elementos destacados por Egan também foram obtidos como resultado da simulação.

Apesar do relato dessa investigação preliminar ser positivo, é prematuro qualificar ou não a Cadeia de Markov como método de previsão de soluções. Destaca-se aqui a necessidade da continuidade dessa investigação para outros exercícios do mesmo tipo.

## Agradecimento

Aproveita-se para agradecer à Fapesp pelo apoio a esse pesquisa (FAPESP 05/50621-2).

## 6- Referências

- Egan, K. (1976). *Structural Communication*. 6 Davis Drive, Belmont, California 94002: Fearon Publishers.
- Feller, W. (1968). *An Introduction to Probability Theory and Its Applications*. Terceira Edição, vol.I, John Wiley & Sons, Inc.
- Jonassen, D. H., (2004). *Learning to Solve Problems - An Instructional Design Guide*. 989 Market Street, San Francisco, CA .Instructional Technology & Training Series.
- Noronha, R., V., 2007. *Conceitos e Modelos de Execução de Exercícios de Comunicação Estrutural*. Tese de Doutorado, Instituto Tecnológico de Aeronáutica, SP, Brasil.