

Detecção Automática e Dinâmica de Memória de Trabalho utilizando *Q-Learning* e Média Móvel Exponencialmente Ponderada

Alessandro Vivas¹, Luciana Assis¹, Cristiano Pitangui²

¹Departamento de Computação – UFVJM
Diamantina – MG

²DTECH – UFSJ
Ouro Branco, MG.

{alessandro.vivas, lupassis, pitangui.cristiano}@gmail.com

Abstract. *Intelligent Tutoring Systems work to customize Virtual Learning Environments according to the learner's cognitive profile. In order to customize these environments, it needs to apply Artificial Intelligence techniques to detect Affective Traits, Working Memory Capacity, and Learning Styles. This work proposes the application of Q-Learning and Exponentially Weighted Moving Average techniques for Working Memory Capacity detection through the use of the apprentice's navigation traces. Experimental results show the potential of both methods to detect the Working Memory Capacity and point out the superiority of the Exponentially Weighted Moving Average technique considering the scenarios evaluated.*

Resumo. *Sistemas de Tutoria Inteligentes trabalham para customizar Ambientes Virtuais de Aprendizagem de acordo com o perfil cognitivo do aprendiz. Para customizar estes ambientes, faz-se necessário a aplicação de técnicas de Inteligência Artificial para detecção de Traços Afetivos, Memória de Trabalho, e Estilos de Aprendizagem. Este trabalho propõe a aplicação das técnicas Q-Learning e Média Móvel Exponencialmente Ponderada para detecção de Memória de Trabalho por meio dos traços de navegação do aprendiz. Resultados experimentais demonstram o potencial de ambas as técnicas para detecção da Memória de Trabalho e apontam que a técnica de Média Móvel Exponencialmente Ponderada é superior considerando os cenários avaliados.*

1. Introdução

A customização de Ambientes Virtuais de Aprendizagem (AVA) por meio do mapeamento do Modelo Cognitivo do aprendiz é um dos principais objetivos de um Sistema de Tutoria Inteligente (STI). Com esta finalidade, existem na literatura diversas abordagens que utilizam diversos traços cognitivos como Estilos de Aprendizagem, Computação Afetiva e mais recentemente Memória de Trabalho (*Working Memory Capacity* - WMC). O objetivo é se opor ao paradigma *one size fits all* que predomina na maior parte dos sistemas de gestão de conteúdo, e customizar o ambiente de acordo com as necessidades do aprendiz.

Os STIs buscam mapear o perfil cognitivo do estudante utilizando técnicas de Aprendizado de Máquina e Inteligência Artificial para adaptar a interface e a

navegação, e oferecer um conteúdo individualizado para melhorar a experiência de ensino-aprendizagem. Memória de Trabalho ou Memória de Curto Prazo (WMC) é uma característica do ser humano que permite manter uma quantidade limitada de informação (aproximadamente 7 ± 2 itens) [Miller 1994, Bernard et al. 2016] por um curto período de tempo [Miller 1994].

Os primeiros trabalhos sobre o uso de WMC para customização de AVAs foram propostos no ano de 2008. Apesar de possuir uma vasta literatura de investigação no campo da psicologia, ainda são escassos os trabalhos que discutem algoritmos para detecção automática e dinâmica da WMC e sua possível aplicação em STIs. Neste sentido, o presente trabalho apresenta e discute duas abordagens automáticas e dinâmicas para a detecção da Memória de Trabalho de alunos.

A metodologia adotada consiste no estudo e análise dos trabalhos relacionados, e desenvolvimento de duas abordagens para classificação da WMC utilizando os algoritmos *Q-Learning* e Média Móvel Exponencialmente Ponderada (MMEP), para uma detecção *online* da capacidade de memória de trabalho do aluno. A abordagem *online* se mostra útil uma vez que a capacidade de memorização pode ser afetada por diferentes fatores inerentes ao ser humano e, por isso, o processo de detecção requer ajustes constantes. Por meio de testes realizados foi possível avaliar condições de ajuste de parâmetros das abordagens propostas para redução dos erros de classificação do tipo da WMC.

Este trabalho se organiza como segue. A seção 2 discute brevemente o conceito de WMC. A seção 3 apresenta os principais trabalhos relacionados a esta pesquisa. A seção 4 apresenta a relação existente entre perfis de navegação e WMC. A seção 5 apresenta as estratégias gerais de personalização de AVA considerando a WMC. A seção 6 discute as técnicas *Q-Learning* e Média Móvel Exponencialmente Ponderada. A seção 7 apresenta os experimentos computacionais realizados para a avaliação das abordagens propostas. Finalmente, a seção 8 conclui o trabalho e aponta algumas pesquisas futuras.

2. Memória de Trabalho (WMC)

Memória de Trabalho (WMC) corresponde ao número de itens que uma pessoa pode armazenar na memória de curto prazo. É importante ressaltar que interfaces de ambientes computacionais precisam levar em conta a WMC de forma a evitar a sobrecarga cognitiva. WMC pode ser definida como a capacidade de memória de um trato cognitivo que influencia o processo de aprendizagem, em termos da velocidade de aprendizagem, memorização de conceitos aprendidos e capacidade de aquisição de competência [Graf et al. 2008, Bernard et al. 2016].

Exceder tal capacidade pode reduzir o desempenho do estudante, diminuir a transferência de conhecimento e aumentar a quantidade de tempo necessário para aprender [Kirschner 2002, Teigen 1994, Bernard et al. 2016]. A WMC afeta diretamente as habilidades do estudante, tais como, a forma de compreensão de textos, tomada de decisões, e resolução de problemas [Tortorella et al. 2015, Ford and Chen 2000, Broadway and Engle 2011].

Autores classificam os aprendizes de acordo com sua WMC em duas categorias: baixa e alta. Neste sentido, por meio da identificação da WMC, pode-se individualizar a carga de trabalho de maneira a facilitar o processo de ensino-aprendizagem [Bernard et al. 2016].

A detecção estática da WMC utiliza um teste denominado *Operation Span Task* (OSPAN) [Turner and Engle 1989, Bernard et al. 2016]. O problema de abordagens estáticas é que muitas vezes questionários são imprecisos devido a falta de interesse de preenchimento, pois muitas vezes estes são muito longos e tediosos. Uma melhor maneira é detectar a WMC por meio de algoritmos durante o processo de navegação do estudante em um AVA. Assim, utiliza-se uma maneira na qual o aprendiz não sabe que está sendo avaliado e assim não influencia intencionalmente o processo de detecção do seu perfil cognitivo.

3. Trabalhos Relacionados

As primeiras abordagens relacionadas ao uso da WMC objetivaram correlacionar a teoria sobre Estilos de Aprendizagem e a Memória de Trabalho [Graf and Kinshuk 2008, Graf et al. 2008]. A correlação entre as teorias foi comprovada a partir da aplicação de dois questionários para 297 alunos dos cursos de Computação e Sistemas de Informação. Um dos questionários (*Index of Learning Styles questionnaire*) identificava o Estilo de Aprendizagem e o outro a WMC (OSPAN) [Graf et al. 2009].

Em 2013, Chang et al. (2013) propôs um mecanismo para identificação da WMC utilizando o perfil de navegação dos aprendizes. Este trabalho correlacionou vários perfis de navegação (linear, revistar conteúdo, lembrar conteúdos de sessões de aprendizagem anteriores, tempo gasto em objetos de aprendizagem) à WMC [Chang et al. 2013a, Chang et al. 2013b].

Em 2014, Chang et al. (2014) trabalhou com estratégias e recomendações, direcionadas aos docentes, de como lidar com alunos de diferentes capacidades da WMC [Chang et al. 2014]. Estas recomendações podem ser utilizadas para a construção de AVAs dinâmicos para lidar com alunos com diferentes perfis cognitivos.

Em 2015, Tortella et al. (2015) relaciona Sistemas Colaborativos de Aprendizagem a modelos cognitivos e *Learning Analytics* [Tortorella et al. 2015]. No mesmo ano, M. M. El-Bishouty (2015) apresenta um *framework* para cálculo da WMC e discute três metodologias para analisar características cognitivas e Estilos de Aprendizagem [El-bishouty et al. 2015].

Em 2016, Bernard et al. (2016) compara os resultados de um Algoritmo Genético (AG) com os dados levantados por meio de questionários para detecção da WMC [Bernard et al. 2016]. Resultados apontam que o AG proposto detecta automaticamente a WMC com uma precisão de 80%.

4. Perfis de Navegação e Memória de Trabalho

Estudos correlacionam 6 padrões (ou características) de navegação que podem auxiliar o processo de cálculo da WMC. O presente trabalho considera apenas 5 padrões, pois descarta o padrão que trata da correlação com a teoria de Estilos de Aprendizagem, uma vez que tal teoria é alvo, recentemente, de severas críticas [Kirschner 2017].

A primeira característica de navegação informa que estudantes com alta WMC tendem a utilizar navegação linear ou sequencial, isto é, seguir os conteúdos do AVA sem retornar constantemente aos conteúdos anteriores. Por outro lado, quando um estudante utiliza navegação não-linear, exhibe uma tendência de baixa WMC [Huai 2000, Chang et al. 2013a].

Um padrão de navegação reversa, onde o estudante revisita o conteúdo, é a segunda característica relacionada a baixa WMC. Tal padrão pode ser detectado pela média do número de vezes que o aprendiz revisita o conteúdo apresentado [Chang et al. 2013a].

A terceira característica é a capacidade de realização de tarefas simultâneas. Quando um estudante, por exemplo, que está assistindo a uma vídeo-aula não termina o seu conteúdo, visita um fórum, e só depois retorna a tarefa anterior e, mesmo assim, obtém um bom desempenho, tal comportamento indica uma alta WMC [Woehrle and Magliano 2012, Engle 2010, Chang et al. 2013a].

A quarta característica relaciona WMC e memória de longo prazo. Uma maneira de detectar esta característica é monitorar a navegação do estudante. Se um aprendiz estuda um conteúdo e finaliza a sessão de aprendizagem, e, ao retornar, realiza o exame sem revisar o material, tal comportamento é um indicativo que o estudante está recuperando eficientemente informações da memória de longo prazo, indicando uma alta WMC. Por outro lado, se o estudante revisa os objetos de aprendizagem após terminar uma sessão de aprendizagem, tal comportamento pode ser um indicativo de baixa WMC [Engle 2010, Unsworth et al. 2012].

A quinta característica relaciona o tempo que um aprendiz desprende em um recurso de aprendizagem. Chang et al. (2013) sugere que estudar durante muito tempo um objeto de aprendizagem particular indica baixa WMC [Chang et al. 2013a]. Este parâmetro pode ser inferido através da média do tempo gasto em cada objeto de aprendizagem. Alunos que despendem mais tempo que a média geral tendem a ter baixa WMC.

5. Estratégias de Personalização de Ambientes Virtuais de Aprendizagem

De maneira geral, e sob o ponto de vista da WMC, os AVAs podem ser customizados à duas grandes classes, a saber: a categoria que abrange alunos com baixa WMC, e a categoria que abrange alunos com alta WMC. Tais categorias se justificam uma vez que alunos com baixa WMC podem sofrer sobrecarga cognitiva quando submetidos a AVAs com grande volume de conteúdos. Por outro lado, alunos com alta WMC possuem tendência a um baixo desempenho quando submetidos a AVAs com pequeno volume de conteúdos.

Neste sentido, para estudantes com baixa WMC, o AVA deve ser customizado diminuindo-se *links* e conteúdos de maneira a evitar a sobrecarga cognitiva, uma vez que estes estudantes podem facilmente perder o foco ao serem submetidos a um grande volume de informações. Outra estratégia é que os temas estudados devem ser revistos para auxiliar o processo de memorização e retenção.

Por outro lado, [Broadway and Engle 2011] propõe o aumento de recursos multimídia para melhorar o desempenho de estudantes que possuem alta WMC. Adicionalmente, deve-se estender a oferta de conteúdos e encorajar os estudantes a aplicarem o que foi aprendido a problemas do mundo real [Sweller 2005, Chang et al. 2015]. Alguns autores sugerem, ainda, a utilização de mapas mentais para correlacionar os temas estudados e o desenvolvimento de habilidades meta-cognitivas.

6. Q-Learning e Média Móvel Exponencialmente Ponderada

Este trabalho propõe duas abordagens para detecção automática e dinâmica de Memória de Trabalho. Adicionalmente, realiza-se um estudo sobre a influência dos parâmetros das

técnicas utilizadas no comportamento das abordagens propostas.

Existem diversas técnicas para se trabalhar com a predição e classificação em tempo real, dentre elas, destacam-se: Redes Bayesianas, *Q-Learning*, *Dynamic Scripting* e Média Móvel Exponencial. Tais métodos já foram aplicados para detecção automática de Estilos de Aprendizagem, e apresentaram resultados promissores quanto a capacidade de convergência, bem como alta potencialidade a se adaptarem às mudanças no ambiente [Salazar et al. 2017, Dorça 2012, Silva et al. 2017, Ribeiro et al. 2017].

6.1. *Q-Learning*

De forma geral, algoritmos baseados em Aprendizado por Reforço, tal como o *Q-Learning* [Mitchell 1997], utilizam funções de ganho que definem se a política de ações que foi aplicada é boa em relação ao contexto onde o agente atua [Iglesias et al. 2009]. O objetivo é que o agente reforce comportamentos que impliquem em ganho e penalize ações que impactam negativamente no seu desempenho. Tal motivação pode ser matematicamente expressa pela Equação 1.

$$Q(s, a) \leftarrow Q(s, a) + \alpha * (R(s) + \gamma * \max_{a'} Q[s', a'] - Q(s, a)) \quad (1)$$

Onde:

- $Q(s, a)$: qualidade da ação do estado s ;
- $\alpha \in [0, 1]$: taxa de aprendizagem;
- $R(s)$: reforço aplicado ao estado s ;
- γ : parâmetro de controle sobre futuras recompensas;
- $Q[s', a']$: qualidade da ação futura a' no estado futuro s' .

Dorça (2012) apresenta uma proposta de detecção de Estilos de Aprendizagem por meio do uso de Aprendizagem por Reforço e Cadeias de Markov. Uma das vantagens desta abordagem é adaptação do conteúdo apresentado em função do modelo cognitivo do aluno [Dorça 2012].

6.2. Média Móvel Exponencialmente Ponderada (MMEP)

Média Móvel Exponencialmente Ponderada (MMEP) surgiu em aplicações empresariais na tentativa de estimar a produção nas próximas semanas levando-se em conta dados passados [Box et al. 2005]. É aplicada também em outros contextos, como por exemplo, na previsão do próximo ciclo de CPU em Sistemas Operacionais, e para estimar o atraso no algoritmo de controle de congestionamento do protocolo *Transmission Control Protocol* (TCP).

Diferentemente do cálculo da média ponderada, MMEP utiliza um fator para suavizar a influência dos valores passados para o cálculo dos valores futuros, conforme Equação 2.

$$\hat{y}_{t+1} = \lambda * y_t + (1 - \lambda) * \hat{y}_t \quad (2)$$

Onde:

- \hat{y}_{t+1} : é a previsão que se deseja fazer;
- y_t : é o valor atual;
- \hat{y}_t : é a previsão feita anteriormente;
- $\lambda \in [0, 1]$: é o fator de suavização.

Este recurso tem a capacidade de filtrar os *outliers* encontrados em previsões erradas que podem levar a um erro de estimativa. Essa capacidade é vista como uma maneira de impedir que resultados ruins oriundos de fatores externos destruam uma tendência que a algum tempo era construída.

Ribeiro et al. (2017) utilizou MMEP no processo de detecção de Estilos de Aprendizagem. Ela foi aplicada para evitar que possíveis erros de estimativas no processo de detecção alterassem as probabilidades de predição, isto é, em vez de utilizar o valor atual de predição, utiliza-se o valor médio e, assim, evitam-se flutuações indesejáveis no algoritmo de predição [Ribeiro et al. 2017].

7. Experimentos Computacionais

7.1. Metodologia Experimental

Para avaliar as abordagens propostas foram geradas (simuladas artificialmente) 10.000 instâncias de sessões de aprendizagem considerando os 5 fatores descritos na seção 4. Assim, foram gerados 10.000 vetores com 5 posições cada. Cada posição de cada vetor foi gerada utilizando-se distribuição normal com $\mu = 0.4$ e $\sigma = 0.3$.

A cada vetor é atribuído um valor médio (considerando-se cada um dos 5 fatores) obtido por uma interação (sessão de aprendizagem) do aprendiz no AVA. Este valor é calculado somando-se os valores de cada uma das 5 componentes do vetor e dividindo-se este resultado por 5.

Geram-se valores médios de interação em duas categorias, a saber:

- i) um conjunto de valores que indica um valor correto da WMC;
- ii) um conjunto de valores que indica um valor incorreto da WMC.

O objetivo de gerar valores segundo ii) é induzir erroneamente um valor da WMC e verificar o comportamento do sistema nesta situação. Neste sentido, caso o estudante possua alta WMC, gera-se um conjunto de valores que induza a detecção de uma baixa WMC. De maneira similar, caso o estudante possua baixa WMC, gera-se um conjunto de valores que induza a detecção de uma alta WMC. Geraram-se, para tanto, aproximadamente 400 exemplos (dos 10.000 utilizados) que induzam a erros de classificação. Assim, avalia-se o comportamento do algoritmo quando lhe são apresentados traços de navegação que não correspondem ao perfil cognitivo do aprendiz.

Todos os experimentos foram replicados 30 vezes para aumentar a significância estatística dos resultados apresentados. Variaram-se os valores de α e λ de 0.1 a 0.9 (com passos de 0.1) para verificar suas influências no comportamento das técnicas. Em todos os testes definiu-se a WMC inicial no polo oposto das geradas nos padrões de testes.

7.2. Determinação dos Parâmetros α e λ

Este experimento tem como objetivo avaliar a relação entre os valores de taxa de aprendizagem (α) do *Q-Learning* e do fator de suavização (λ) da MMEP. Os parâmetros de

ambos as técnicas foram escolhidos em um mesmo padrão de testes conforme descrito na seção anterior. Além de avaliar a relação entre os valores de α e λ , objetiva-se determinar seus valores de modo que o *Q-Learning* e a MMEP incorram na menor quantidade de erros de detecção da WMC.

A Figura 1 apresenta os resultados da simulação para os diversos valores de α e λ , enquanto a Tabela 1 apresenta os resultados do valor médio e do desvio padrão dos erros de predição encontrados na simulação.

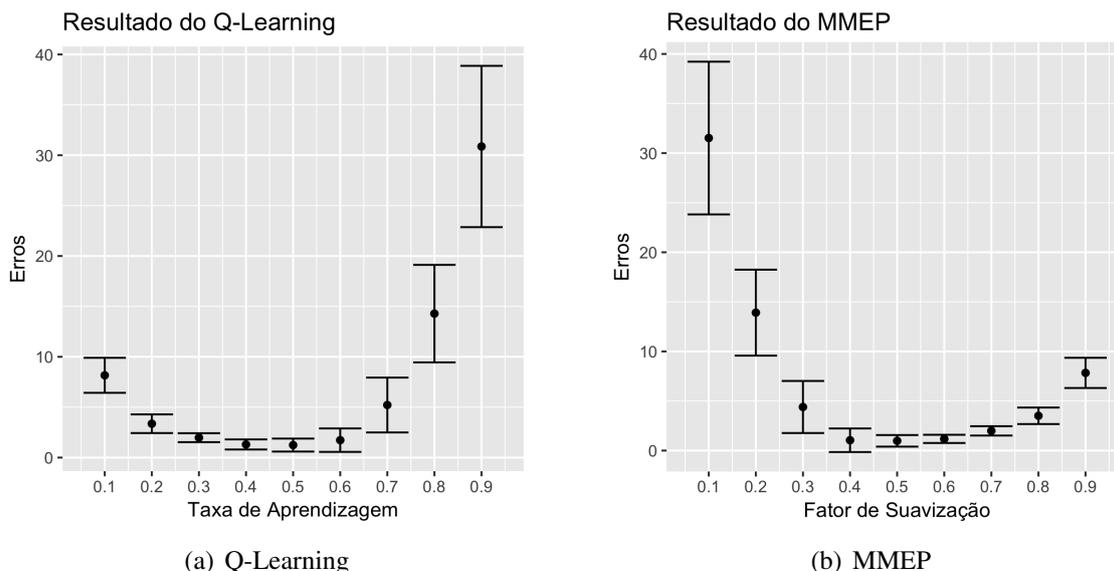


Figura 1. Calibragem dos Parâmetros

<i>Q-Learning</i>									
α	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
média	8.16	3.36	1.98	1.30	1.24	1.73	5.22	14.28	30.87
desvio padrão	1.73	0.93	0.44	0.50	0.64	1.16	2.72	4.84	8.00
MMEP									
λ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
média	32.62	14.36	5.59	1.72	1.11	1.24	2.19	3.61	8.00
desvio padrão	7.59	4.90	2.37	1.33	0.96	0.46	0.72	0.74	1.52

Tabela 1. Determinação dos Parâmetros Iniciais de Simulação

Resultados obtidos permitem concluir que o *Q-Learning* e a MMEP demonstram melhor comportamento quando $\alpha = 0.5$ e $\lambda = 0.6$, respectivamente.

7.3. Erros de Convergência

O objetivo deste experimento é avaliar quantos dos 400 exemplos gerados de forma a induzir erros de classificação são de fato classificados incorretamente.

A Figura 2 apresenta o número de erros para cada um dos 30 testes realizados para ambas as técnicas propostas. A média de erros do algoritmo *Q-Learning* foi de 25.27 com desvio padrão de 8.67 (25.27 ± 8.67) e, para o MMEP, a média foi de 5.53 com desvio

padrão de 3.73 (5.53 ± 3.73). Neste sentido, os resultados encontrados apontam que, para os cenários considerados, a MMEP tem melhor desempenho quando comparada ao *Q-Learning*.

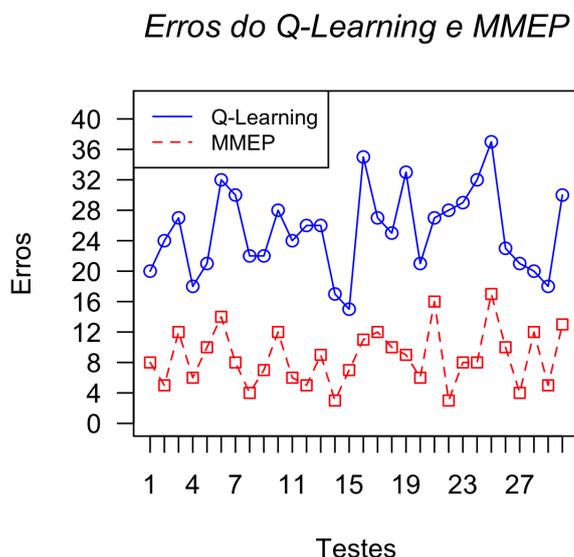


Figura 2. Comparativo dos Dois Algoritmos

Para averiguar se há diferenças significativas entre o desempenho dos algoritmos: *Q-Learning* e *MMEP*, foi feito um planejamento experimental e uma análise estatística dos dados com os resultados obtidos por ambos métodos analisados. Para cada algoritmo foram realizadas 30 replicações. Os testes foram executados de maneira independente e foi avaliada a taxa de erros de ambos os algoritmos implementados.

A hipótese nula indica ausência de diferença entre os algoritmos avaliados. Na análise dos experimentos foi considerado um nível de significância de 95%. Após a averiguação da normalidade dos dados foi utilizado o teste t. O resultado do teste de hipótese aponta para a rejeição da hipótese nula uma vez que o p-valor foi inferior a 0.05 ($p\text{-value} < 2.2e^{-16}$). O valor do intervalo de confiança obtido foi de 14.83 e 18.97 demonstrando que o algoritmo *Q-Learning* apresenta uma taxa de erros maior, em média 16.37 erros a mais do que o *MMEP*.

8. Conclusões e Trabalhos Futuros

Este trabalho apresentou duas propostas para detecção automática e dinâmica de Memória de Trabalho utilizando traços de navegação em Ambientes Virtuais de Aprendizagem. Uma das propostas utilizou o algoritmo *Q-Learning* e a outra fez uso da técnica de Média Móvel Exponencialmente Ponderada. Ambas as propostas foram devidamente avaliadas em bases de dados geradas de maneira artificial.

Resultados experimentais apontam que a proposta utilizando *MMEP* supera a proposta baseada no algoritmo *Q-Learning*. Mais precisamente, considerando erros de classificação, verificou-se que o algoritmo *Q-Learning* apresentou erro médio de 6.3%,

enquanto a MMEP de apenas 2.2%. Além da menor quantidade de erros, a MMEP obteve menor variabilidade em seus resultados.

Referências

- Bernard, J., Chang, T.-W., Popescu, E., and Graf, S. (2016). Optimizing Pattern Weights with a Genetic Algorithm to Improve Automatic Working Memory Capacity Identification. In *ITS 2016*, pages 334–340. Springer International Publishing.
- Box, G. E. P., Hunter, J. S., and Hunter, W. G. (2005). *Statistics for Experimenters: Design, Innovation and Discovery*. John Wiley & Sons, Inc., Hoboken, New Jersey, 2nd ed. edition.
- Broadway, J. M. and Engle, R. W. (2011). Lapsed attention to elapsed time? Individual differences in working memory capacity and temporal reproduction. *Acta Psychologica*, 137(1):115–126.
- Chang, T.-W., El-Bishouty, M. M., Graf, S., and Kinshuk (2013a). An Approach for Detecting Students' Working Memory Capacity from Their Behavior in Learning Systems. *2013 IEEE 13th International Conference on Advanced Learning Technologies*, pages 82–86.
- Chang, T. W., El-Bishouty, M. M., Graf, S., and Kinshuk (2013b). Recommendation mechanism based on students' working memory capacity in learning systems. *Proceedings - 2013 IEEE 13th International Conference on Advanced Learning Technologies, ICAIT 2013*, pages 333–335.
- Chang, T.-w., Kurcz, J., and El-bishouty, M. M. (2015). Ubiquitous Learning Environments and Technologies.
- Chang, T.-W., Kurcz, J., El-Bishouty, M. M., Graf, S., and Kinshuk (2014). Adaptive Recommendations to Students Based on Working Memory Capacity. In *2014 IEEE 14th International Conference on Advanced Learning Technologies*, pages 57–61. IEEE.
- Dorça, F. A. (2012). *Fabiano Azevedo Dorça*. PhD thesis, Universidade Federal de Uberlândia.
- El-bishouty, M. M., Chang, T.-w., and Lima, R. (2015). Analyzing Learner Characteristics and Courses Based on Cognitive Abilities, Learning Styles, and Context. In Chang, M. and Li, Y., editors, *Smart Learning Environments*, Lecture Notes in Educational Technology, pages 3–25. Springer Berlin Heidelberg, Berlin, Heidelberg.
- Engle, R. W. (2010). Role of Working-Memory Capacity in Cognitive Control. *Current Anthropology*, 51(S1):S17–S26.
- Ford, N. and Chen, S. Y. (2000). Individual Differences, Hypermedia Navigation, and Learning: An Empirical Study. *J. Educ. Multimedia Hypermedia*, 9(4):281–311.
- Graf, S. and Kinshuk (2008). Learner modelling through analyzing cognitive skills and learning styles. *Handbook on Information Technologies for Education and Training*, pages 179–194.
- Graf, S., Lin, T., and Kinshuk (2008). Improving student modeling: the relationship between learning styles and cognitive traits. *Computers in Human Behavior*, pages 122–137.

- Graf, S., Liu, T. C., Kinshuk, Chen, N. S., and Yang, S. J. H. (2009). Learning styles and cognitive traits - Their relationship and its benefits in web-based educational systems. *Computers in Human Behavior*, 25(6):1280–1289.
- Huai, H. (2000). *Cognitive style and memory capacity: effects of concept mapping as a learning method*. Tese de doutorado, Twente University, The Netherlands.
- Iglesias, A., Martinez, P., Aler, R., and Fernandez, F. (2009). Learning teaching strategies in an adaptive and intelligent educational system through reinforcement learning. *Applied Intelligence*, 31(1):89–106.
- Kirschner, P. (2002). Cognitive load theory: implications of cognitive load theory on the design of learning. *Learning and Instruction*, 12(1):1–10.
- Kirschner, P. A. (2017). Stop propagating the learning styles myth. *Computers & Education*, 106:166–171.
- Miller, G. (1994). The magic number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 101(2):343–352.
- Mitchell, T. M. (1997). Machine learning. *MIT Press*, pages 1–5.
- Ribeiro, P., Assis, L., Vivas, A., and Pitangui, C. (2017). Detecção de Estilos de Aprendizagem utilizando Média Móvel Exponencialmente Ponderada. In *Simpósio Brasileiro de Informática na Educação*, page 1247, Recife.
- Salazar, L., Assis, L., Vivas, A., Pitangui, C., and Falci, S. (2017). Detecção de Estilos de Aprendizagem em Ambientes Virtuais de Aprendizagem Utilizando Redes Bayesianas. In *Simpósio Brasileiro de Informática na Educação*, page 1317, Recife.
- Silva, J. C., Pitangui, C., Assis, L., and Vivas, A. (2017). Detecção Automática e Dinâmica de Estilos de Aprendizagem em Sistemas Adaptativos e Inteligentes utilizando Dynamic Scripting. In *Simpósio Brasileiro de Informática na Educação*, number S, page 1327, Recife.
- Sweller, J. (2005). Implications of Cognitive Load Theory for Multimedia Learning. *The Cambridge Handbook of Multimedia Learning*, pages 19–48.
- Teigen, K. H. (1994). Yerkes-Dodson: A Law for all Seasons. *Theory & Psychology*, 4(4):525–547.
- Tortorella, R. A. W., Hobbs, D., Kurcz, J., Bernard, J., Baldiris, S., Chang, T.-W., and Graf, S. (2015). Improving Learning Based on the Identification of Working Memory Capacity, Adaptive Context Systems, Collaborative Learning and Learning Analytics. *Proceedings of Science and Technology Innovations*, pages 39–55.
- Turner, M. L. and Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, 28(2):127–154.
- Unsworth, N., Redick, T. S., Spillers, G. J., and Brewer, G. A. (2012). Variation in working memory capacity and cognitive control: Goal maintenance and microadjustments of control. *Quarterly Journal of Experimental Psychology*, 65(2):326–355.
- Woehrle, J. L. and Magliano, J. P. (2012). Time flies faster if a person has a high working-memory capacity. *Acta Psychologica*, 139(2):314–319.