Associative Pattern Matching and Inference Using Associative Graph Data Structures

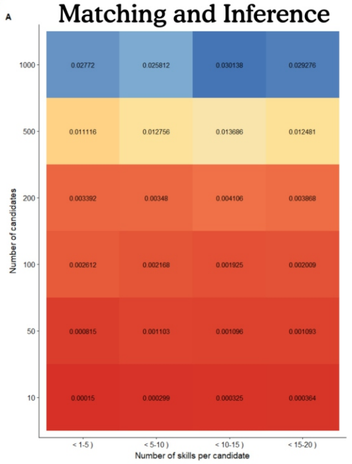
Adrian Horzyk AGH University of Science and Technology, Department of Biocybernetics and Biomedical Engineering, Mickiewicza Av. 30, 30-059 Krakow, Poland

Agata Czajkowska Warsaw University of Technology, Koszykowa 75, 00-662 Warsaw, Poland

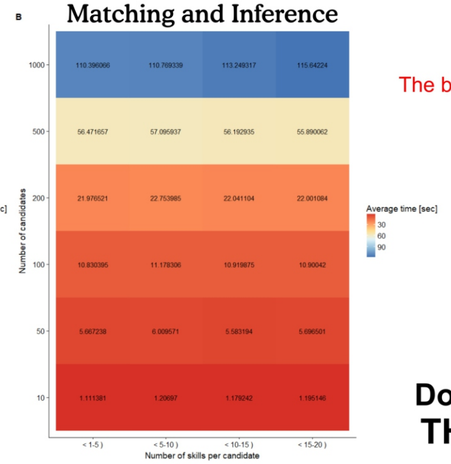
	Technical Skills			PersonalSkills			LanguageSkills			Education Level	Sallary	Time Work Type	Field Importance
	Name	Years	Weight	Name	Level	Weight	Name	Level	Weight				
Job Offer	C#	≥ 3	(10/10)	Communication skills	60%-80%	(8/10)	English	≥ C1	(8/10)	Bachelor	7500	FullTime	TechnicalSkills (9/10)
	Entity Framwork	≥ 3	(9/10)										PersonalSkills (8/10)
				Ability to work under pressure	80%-100%	(10/10)							LanguageSkills (5/10)
	T-SQL	≥ 2	(8/10)				Polish	≥ Native	(7/10)				Education (10/10)
				Decision making	60%-80%	(7/10)							Sallary(10/10)
	Cloud dev	≥ 1	(6/10)										TimeWorkType (10/10)
Candidate 1	C#	2		Ability to work under pressure	70%								
	T-SQL	3	X			X	Polish	Native	X	Bachelor	7000	HalfTime	X
	Cloud dev	2		Communication skills	80%								
Candidate 2	Entity Framwork	3		Communication skills	70%								
	C#	1	X	Decision making	90%	X	English	B2	X	Master	6500	FullTime	X
	T-SQL	1		Ability to work under pressure	90%								
Candidate 3	Cloud dev	2	X	Communication skills	80%	X	Polish	Native	X	Master	6500	HalfTime	X
	C#	2		Decision making	90%		English	C1					
Candidate 4	Entity Framwork	1		Ability to work under pressure	70%								
	C#	1	X			X	English	Native	X	Bachelor	8000	FullTime	X
	T-SQL	2		Decision making	50%								
	Cloud dev	2											
Candidate 5	Cloud dev	1		Communication skills	70%								
	Entity Framwork	2	X	Ability to work under pressure	80%	X	Polish	C1	X	Bachelor	7000	HalfTime	X
	T-SQL	3		Decision making	90%								

EFFICIENCY COMPARISON

Associative Pattern Matching and Inference



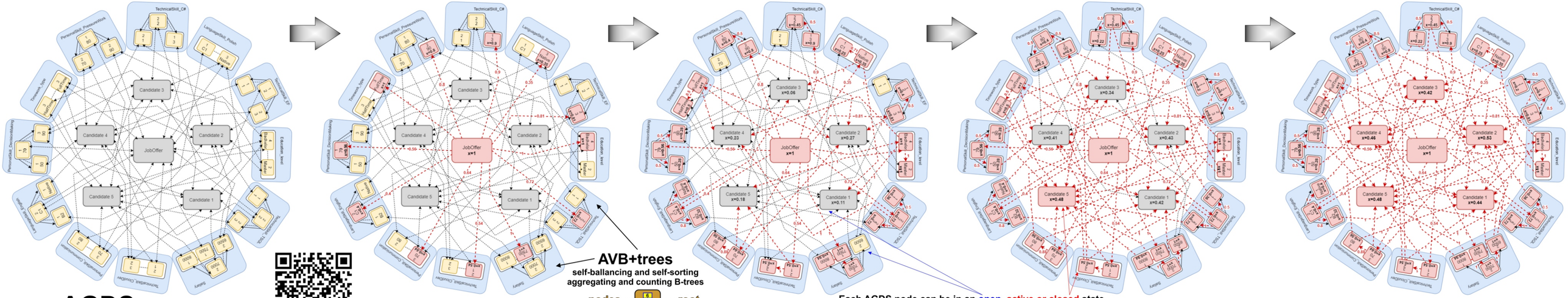
KNN Pattern Matching and Inference



The best matching

Object node	Stimulus strength	Nodes sending signal
Candidate1	$\frac{1}{12} \cdot (1 + 0.35) \approx 0.11$	*Bachelor (Eduacation_level)* *Native (LanguageSkill_Polish)*
Candidate2	$\frac{1}{12} \cdot (1 + 0.8 + 0.81 + 0.64) \approx 0.27$	*FullTime (Timework_type)* *90 (PersonalSkill_PressureWork)* *3 (TechnicalSkill_EF)* *70 (PersonalSkill_Communication)*
Candidate3	$\frac{1}{12} \cdot (0.4 + 0.35) \approx 0.06$	*C1 (LanguageSkill_English)* *Native (LanguageSkill_Polish)*
Candidate4	$\frac{1}{12} \cdot (1 + 1 + 0.72) \approx 0.23$	*FullTime (Timework_type)* *Bachelor (Eduacation_level)* *2 (TechnicalSkill_TSQL)*
Candidate5	$\frac{1}{12} \cdot (0.64 + 0.54 + 1) \approx 0.18$	*70 (PersonalSkill_Communication)* *1 (TechnicalSkill_CloudDev)* *Bachelor (Eduacation_level)*

Don't connect network elements using all-to-all schema!
THE STRUCTURE OF CONNECTIONS MATTERS!

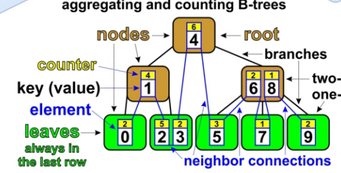


AGDS

Associative Graph Data Structure accelerated by AVB+trees inspired by the brain associative processes

A. Horzyk and K. Gołdon, Associative Graph Data Structures Used for Acceleration of K Nearest Neighbor Classifiers, In: 27th International Conference on Artificial Neural Networks, ICANN 2018, Springer-Verlag, LNCS 11139, pp. 648-658, 2018.

AVB+trees
self-ballancing and self-sorting aggregating and counting B-trees



Each AGDS node can be in an open, active or closed state.
At the beginning, all nodes are in the open state.
When the node is stimulated the right number of times, it becomes active, sends its charging level to all connected nodes which are in the open state, and switches to the closed state.
We stimulate the network until the destination nodes are open.

A. Horzyk, Associative Graph Data Structures with an Efficient Access via AVB+trees, In: 2018 11th International Conference on Human System Interaction (HSI), IEEE Xplore, pp. 169 - 175, 2018, DOI: 10.1109/HSI.2018.8430973.

The BFS tree stimulation order is used to charge AGDS network nodes and find similar nodes to inputs according to the associative strength of connections.

THE WEIGHT BETWEEN NODES REPRESENTING NEIGHBOR VALUES OF THE SAME ATTRIBUTE

$$w_{v_i^{a^k}, v_j^{a^k}} = 1 - \frac{\left| v_i^{a^k} - v_j^{a^k} \right|}{r^{a^k}}$$

CHARGING LEVEL OF THE Nth NODE COMPUTED OVER TIME

$$x_n = \sum_{k=1}^{S_n} x_k \cdot w_k$$

THE WEIGHTS FROM THE NODES REPRESENTING VALUES TO THE NODES REPRESENTING OBJECTS

$$w_{v_i^{a^k}, o_n} = \frac{1}{\theta_n}$$

THE WEIGHTS FROM THE NODES REPRESENTING OBJECTS TO THE NODES REPRESENTING VALUES

$$w_{o_n, v_i^{a^k}} = 1$$

THE WEIGHTS FROM THE NODES REPRESENTING DEFINING OBJECTS TO THE NODES REPRESENTING DEFINED OBJECTS

$$w_{o_m, o_n} = \frac{1}{\theta_n}$$

THE WEIGHTS FROM THE NODES REPRESENTING DEFINED OBJECTS TO THE NODES REPRESENTING DEFINING OBJECTS

$$w_{o_n, o_m} = 1$$

Associative Pattern Matching and Inference:

The associative inference process starts from the question node that is stimulated by one, immediately changing its state to active and stimulating the connected source nodes. The input criteria can be defined with different priorities which define the weights between the question node and the source nodes. In effect, all source nodes become active, stimulate connected open nodes and switch to the closed state. The source nodes are connected to various nodes defining other similar values and associated objects, so the inference process propagates stimuli over the AGDS graph in the BFS order until all or a demanded number of the destination nodes become active. Each activated destination node sends its charging level (association strength) out of the graph as its matching level to the input criteria. The mostly matching patterns (represented by the strongest associated nodes) are the final answer of this network.

ICAISC 2019 Zakopane
18-th International Conference on Artificial Intelligence and Soft Computing
JUNE 16-20, 2019

In: Rutkowski L., Korytkowski M., Scherer R., Tadeusiewicz R., Zadeh L., Zurada J. (eds) Artificial Intelligence and Soft Computing, Springer-Verlag, LNCS, Vol. 10509, pp. 371-383, 2019 Proceedings of ICAISC 2019, DOI: 10.1007/978-3-030-20915-5_34