Towards a Quantum-Inspired Multi-Gene Linear Genetic Programming Model

Guilherme Cesário Strachan, Adriano Soares Koshiyama, Douglas Mota Dias, Marley M. B. R. Vellasco, & Marco Aurélio Cavalcanti Pacheco
Pontifical Catholic University of Rio de Janeiro, Dept. of Electrical Engineering, Rio de Janeiro, Brazil
[strachan,adriano,douglasm,marley,marco]@ele.puc-rio.br

ABSTRACT

This paper presents a new model for regression problems based on Multi-Gene and Quantum Inspired Linear Genetic Programming. We discuss theoretical aspects, operators, representation, and experimental results.

1. INTRODUCTION

The development of new GP regression models is relevant to provide more accurate results through fewer evaluations, and two distinct approaches are possible: (i) modify the GP basic structure: finding new ways to codify a solution and thus providing new recombination operators. There is some work in this area, such as Linear GP [1, 2]; (ii) allowing more outputs per GP individual: a simple approach is to enable more functions per individual and to combine their outputs, like Multi-Gene (or Multi-Tree) GP (MGGP) [3].

Thus, new GP based models that can operate in both senses could generate better results within fewer evaluations. This paper proposes a Quantum-Inspired Multi-Gene Linear GP model (QIMuLGP) for regression tasks, a generalization of Quantum-Inspired Linear GP (QILGP) [2]. QIMuLGP modifies canonical GP structures, explores new recombination operators and enables several outputs per individual that can be combined applying the least squares method. We evaluated this approach on 11 datasets, comparing its results with GP, MGGP and QILGP.

2. QUANTUM-INSPIRED MULTI-GENE GP

We propose a novel GP model in this paper: Quantum-Inspired Multi-Gene Linear GP (QIMuLGP). The main difference between the original QILGP [2] and this generalization is that each individual has more than one chromosome. Therefore the fitness of an individual results from a linear combination of each chromosome output, where the weights are adjusted using least squares method – like MGGP.

Figure 1 illustrates QIMuLGP structure and its basic operation. It has a quantum population with N quantum individuals (QIs) with M chromosomes each (e.g. $N = 3$ and $M = 4$ in Figure 1). QIMuLGP has two classical individuals (CIs): one to store an observed individual and other for the best individual found. Through CIs, x86 machine code programs are generated. Figure 1 also enumerates four basic steps that repeat N times to complete a generation: 1^{st}. a QI is observed generating a CI (*Observed Individual*); 2^{nd}. the M chromosomes of *Observed Individual* are linearly combined to calculate its fitness; 3^{rd}. if its fitness is better, it is copied to *Best Classical Individual*; 4^{th}. an operator P is applied to the QI observed in step 1, taking as reference *Best Classical Individual*, increasing the probability that future observations of the QI results in CI more similar to the best found.

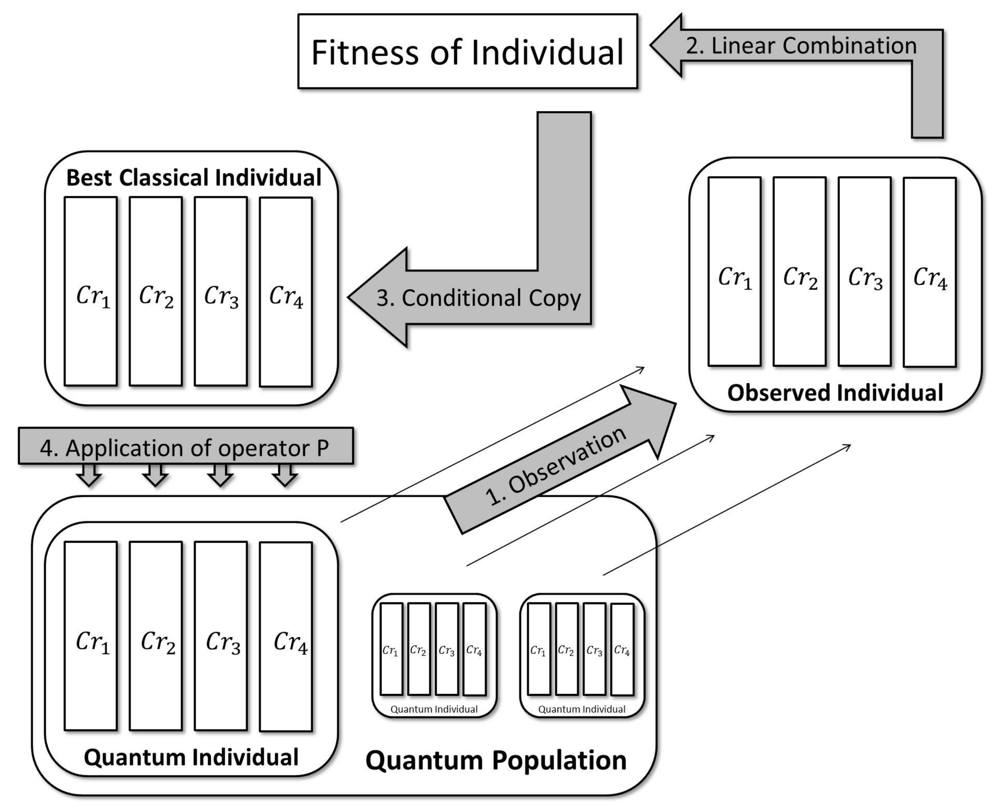

Figure 1: Basic diagram of QIMuLGP model.

The observation of a QI comprises observing each of its chromosomes, which defines the chromosomes of the resulting CI. The Figure 2 shows the observation process. The evolution continues the same way as QILGP.

3. RESULTS AND DISCUSSIONS

Tables 1 and 2 present the main results (RMSE) and standard deviation (σ) for the test set as well as the time spent for performing an evaluation (milliseconds). We varied the number of evaluations according to the number of variables of each dataset (parenthesis in the first columns) multiplied by some default values (3,000, 5,000, 7,000, 11,000).In general, two patterns can be identified: (i) performing more evaluations can benefit almost all evolutionary algorithms; (ii) Multi-Gene approaches (MGGP and QIMuLGP) compare favorably with their canonical counterparts.

QIMuLGP enhanced the average RMSE of QILGP about 54%, with a dispersion reduction of 10%, but the computational cost was 24 times higher. The comparison with MGGP shows that QIMuLGP RMSE values were 19% higher on average. However, the proposed model had an speedup factor of 8.

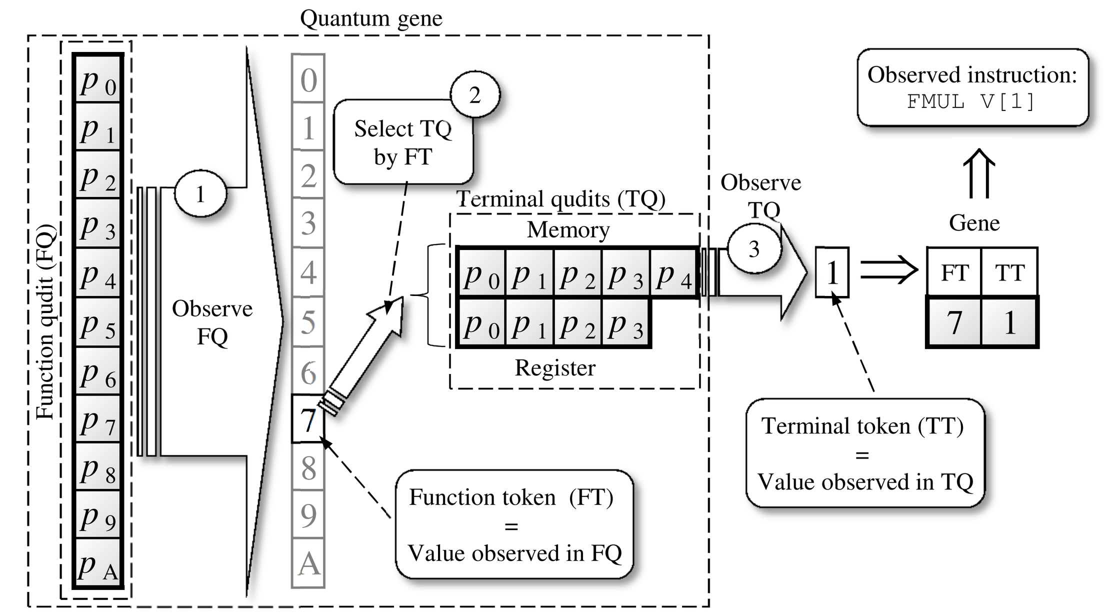

Figure 2: Creation of a gene by the observation of a quantum gene.

Table 1: Main results of GP and MGGP for test set.

	GP				
Evals/vars	3,000	5,000	7,000	11,000	ms/eval
ABA(8)	2.32489	2.34160	2.27052	2.25400	
σ	0.01600	0.02610	0.05027	0.03850	3.37462
MPG6 (5)	5.85305	6.42550	4.90581	4.65530	
σ	0.34086	0.16460	0.27672	0.27380	0.23375
MPG8 (7)	4.84849	4.93290	4.45391	4.30520	
σ	0.26502	0.09670	0.12933	0.16150	0.48711
FRD (5)	2.65968	2.66270	2.36383	2.54460	
σ	0.20180	0.28990	0.28181	0.30970	1.50323
LAS (4)	29.2293	29.8512	28.8769	25.6498	
σ	1.25435	0.97750	1.46215	2.12210	1.52929
CPU (6)	80.6985	94.9208	78.1299	74.3228	
σ	6.13706	7.82560	8.83915	6.10730	0.24385
PLA (2)	3.11145	3.59640	2.83740	2.66930	
σ	0.12580	0.41730	0.20408	0.10280	3.63541
QUA (3)	0.20042	0.19010	0.18941	0.18940	
σ	0.05219	0.00130	0.00015	0.00020	1.18020
ELE (18)	0.00505	0.00497	0.00486	0.00479	
σ	0.00007	0.00021	0.00013	0.00017	7.82954
TRE (15)	0.40564	0.38500	0.36876	0.39916	
σ	0.10046	0.01660	0.02252	0.17031	1.07788
BAS (16)	714.178	709.767	728.983	910.453	
σ	24.5729	16.2693	18.7587	55.0068	0.36862
	MGGP				
Evals/vars	3000	5000	7000	11000	ms/eval
ABA(8)	2.09448	2.09005	2.15677	2.09146	
σ	0.01536	0.00939	0.27578	0.01209	68.9318
MPG6 (5)	2.82620	2.79916	2.76916	2.78887	
σ	0.03991	0.05446	0.04314	0.08218	19.0408
MPG8 (7)	2.77983	2.79353	2.74484	2.79120	
σ	0.05648	0.06879	0.04656	0.06131	21.7090
FRD (5)	1.19303	1.10039	1.16372	1.11400	
σ	0.08375	0.03939	0.07745	0.03803	56.9523
LAS (4)	7.18401	7.27544	6.62436	6.91258	
σ	0.43145	0.31578	0.37549	0.36189	49.4215
CPU (6)	78.3732	81.5594	75.6617	68.5437	
σ	50.6581	41.5579	6.05547	32.4190	42.4149
PLA (2)	1.36655	1.32200	1.30754	1.31548	
σ	0.02762	0.01607	0.01867	0.01511	179.986
QUA (3)	0.19185	0.19101	0.19196	0.19195	
σ	0.00143	0.00160	0.00097	0.00222	82.7172
ELE (18)	0.00245	0.00235	0.00685	0.00504	
σ	0.00021	0.00008	0.01827	0.00676	107.327
TRE (15)	0.21615	0.21601	0.23694	0.21140	
σ	0.00837	0.00309	0.09619	0.00361	100.430
BAS (16)	1437.11	941.210	887.302	962.363	
σ	3232.07	487.200	66.2986	392.300	3.87404

4. CONCLUSIONS

We applied QILMuGP in set of 11 regression benchmarks, and it was found that QILMuGP greatly improve the accuracy when comparing to its simplified version (QILGP);

Table 2: Main results of QILGP and QIMuLGP for test set.

	QILGP				
Evals/vars	3000	5000	7000	11000	ms/eval
ABA(8)	2.49880	2.42167	2.33994	2.30975	
σ	0.10909	0.07016	0.06592	0.08731	0.97413
MPG6 (5)	5.65584	4.84559	4.30004	3.82211	
σ	0.66805	0.35753	0.54516	0.60763	0.12580
MPG8 (7)	5.02799	4.48504	4.19458	3.72854	
σ	0.52746	0.58040	0.64397	0.71112	0.13271
FRD (5)	3.83546	3.40657	3.20121	2.80673	
σ	0.16641	0.18097	0.13782	0.08024	0.30340
LAS (4)	28.9468	28.1398	26.0149	22.9493	
σ	1.91894	2.34428	2.53016	2.14484	0.24450
CPU (6)	74.1313	71.0203	69.7652	57.5957	
σ	15.5803	14.0342	16.3019	8.33990	0.08650
PLA (2)	3.17037	2.82697	2.68247	2.39313	
σ	0.16257	0.20336	0.10046	0.12619	0.33500
QUA (3)	0.18977	0.18946	0.18948	0.18952	
σ	0.00643	0.00635	0.00635	0.00631	0.46000
ELE (18)	0.58902	1.16181	1.16172	2.31833	
σ	1.30531	2.58649	2.58656	5.17306	3.54139
TRE (15)	0.33768	0.31326	0.30514	0.29105	
σ	0.03757	0.04039	0.04268	0.03810	0.32467
BAS (16)	844.192	757.278	748.243	732.011	
σ	26.4487	68.6481	35.7884	86.1586	0.13319
	QIMuLGP				
Evals/vars	3000	5000	7000	11000	ms/eval
ABA(8)	2.15181	2.12817	2.14707	2.13025	
σ	0.08167	0.08105	0.10566	0.08106	11.6134
MPG6 (5)	2.87843	2.81456	2.77730	2.74637	
σ	0.24769	0.20421	0.22262	0.25929	3.04560
MPG8 (7)	2.88234	2.81248	2.81635	2.76679	
σ	0.38581	0.36920	0.35363	0.35807	3.85343
FRD (5)	2.06512	1.88007	1.70563	1.52325	
σ	0.13197	0.18702	0.10826	0.09174	6.93420
LAS (4)	10.5812	9.48600	9.17668	7.82286	
σ	2.01984	1.62191	1.90272	1.60768	6.38325
CPU (6)	94.6812	61.6998	178.479	148.720	
σ	107.974	33.4788	248.655	218.108	2.90633
PLA (2)	1.51029	1.50002	1.49180	1.48186	
σ	0.04283	0.04069	0.04147	0.05259	8.90900
QUA (3)	0.19564	0.19032	0.35561	0.19337	
σ	0.01749	0.00785	0.37354	0.00917	8.25233
ELE (18)	0.00695	0.00266	0.00259	0.00390	
σ	0.00923	0.00009	0.00006	0.00302	63.2082
TRE (15)	0.23948	0.23853	0.23182	0.22692	
σ	0.04567	0.04297	0.04221	0.04040	52.4000
BAS (16)	737.169	747.094	984.207	780.099	
σ	87.4307	39.7709	543.770	79.8790	1.85350

and MGGP obtained slightly better results than QILMuGP, however using twice the computational effort.

5. ACKNOWLEDGMENTS

The authors would like to thank CNPq, CAPES, FAPERJ and PUC-Rio for supporting this work.

6. REFERENCES

[1] M. F. Brameier and W. Banzhaf. *Linear genetic programming*. Heidelberg, Springer, 2007.

[2] D. M. Dias and M. A. C. Pacheco. Quantum-inspired linear genetic programming as a knowledge management system. *The Computer Journal*, 56(9):1043–1062, 2013.

[3] D. Searson, M. Willis, and G. Montague. Coevolution of nonlinear pls model components. *Journal of Chemometrics*, 21(12):592–603, 2007.