Pasta Experiment, Friday 12th October 2007

Training Data:

Classes	Conchigle, Farfalle, Fusilli, Penne, Volanti
Training Set	135 shapes.
Test Set	150 shapes
Image Resolution: 640x480	
Camera	Logitech Quickcam Pro 9000

	Training %	Test %	Time (ms)
SGGP	100.0	90.7	4745
My GP System (DT)	96.3	86.7	114062
My GP System (DRS)	97.8	89.3	6860
ECJ (DT)	91.1	82.7	109000
KNN	NA	87.3	NA

Something which occurred to me:

Solution space increases in proportion to the problem space

And the fitness of an individual is necessarily independent from its implementation

Therefore within any population, there may exist more than one candidate solution

Which may be totally different, and therefore attack different parts of the training data

Standard GP searches for only one solution at a time

So the transient existence of other, useful solutions is wasted

SGGP aims to make full use of all potential solutions in a generation at once.

Which is why it is faster.

Size

174

Pasta Experiment 2, Thursday 8th November 2007

Training Data:

Classes	Conchigle, Farfalle, Fusilli, Penne, Volanti
Training Set	135 shapes.
Test Set	300 shapes
Image Resolution: 640x480	
Camera	Logitech Quickcam Pro 9000

	Training %	Test %	Time (ms)
SGGP	100.0	89.0	3872
My GP System (DT)	100.0	90.3	71022
My GP System (DRS)	96.3	89.0	2568
ECJ (DT)	97.8	83.3	69362
KNN	NA	86.7	NA

Size

26 partial solutions

37

5

372

NA

Big ANPR, Thursday 9th November 2007

Training Data:

Classes	32 classes – numbers and letters
Training Set	175 shapes.
Test Set	455 shapes
Image Resolution: Variable	
Camera	Nikon D70

	Training %	Test %	Time (ms)
SGGP	100.0	82.2	18456
My GP System (DT)	45.1	35.6	317342
My GP System (DRS)	66.9	52.3	167606
ECJ (DT)	45.7	40.0	313349
KNN	-	85.0	-

Size

80 classifiers

184

nr

449

-

Posture Experiment, Thursday 18th October 2007

Training Data:

Classes	10: Open Palm, Fist, Thumbs Up, Live Long, Pinch, Closed Palm,
Training Set	50 shapes
Test Set	50 shapes
Image Resolution: 640x480	
Camera	Logitech Quickcam Pro 9000

	Training %	**Test %**	**Time (ms)**
SGGP	100.0	*82.0*	*2309*
My GP System (DT)	84.0	64.0	145000
My GP System (DRS)	94.0	30.0	81000
ECJ (DT)	88.0	64.0	341000
KNN	NA	38	NA

NOTE: SGGP results udated since last sent to Adrian (mon 22 oct)

Point1, Point2, Okay, Bang

Colour Segmentation Experiment, 25th October 2007

Training Data:

Classes	7: Red, Green, Black ,White, Silver, Yellow, Blue
Training Set	213 training samples with 10 colour-based features
Test Set	10471 training samples
Further info	Each sample represents a single pixel in an image
	Colours are taken from images of cars, so red samples
	Come from parts of a red car
Camera	Nikon D70

	Training %	Test %	Time (ms)	Tree Size
SGGP	100.0	*96.8*	*12698*	
My GP System (DT)	90.6	85.1	53652	
My GP System (DRS)	87.7	85.0	199115	
ECJ (DT)	87.7	84.7	125291	292
KNN	NA	99.2	NA	NA

Note: ECJ often produces vast trees

Statlog SatImage Data Set

http://www.is.umk.pl/projects/datasets-stat.html#Satellite

Training: 4435 samples

Testing: 2000 samples

Features: 36

Classes: 6

	Training %	Test %	Time (ms)	Error
SGGP	100% of 250	80.0	7742	0.200
My GP System (DT)	69.1	67.8	454021	0.322
My GP System (DRS)	82.7	80.8	79869	0.192
ECJ (DT)	73.9	73.2	1132980	0.268
KNN	NA	89.4	NA	0.106

Not very happy with this result at the moment

It isn't quite competitive with the results using other ML techniques

I   If I could have some more time to improve this one, that would be good.

DRS improves significantly (15% extra on training) when using normalised data

Size
NR
NR
8
315
NA

a.

BE SGGP			DRS SGGP		
Hits	**Time**	**Learners**	**Hits**	**Time**	**Learners**
31	5173	36	30	4956	25
31	2351	32	42	4842	25
33	3731	35	33	4798	23
41	2309	30	36	4705	26
33	3706	32	35	4819	22
35	2535	34	39	4679	20
34	2486	29	40	48.32	24
28	2634	40	38	4703	20
29	2346	36	34	4887	21
34	3978	35	37	4800	24
Average: **32.9**	**3124.9**	**33.9**	**36.4**	**4323.73**	**23**
Max: **41**	**5173**	**40**	**42**	**4956**	**26**

Comparison of SGGP using binary classifiers that are either small boolean ex
With SGGP using binary classifiers that use dynamic thresholding (DRS)
This is on the postures training data

:pressions (BE)

Without Normalisation

Optimising 10%			Optimising None		
Hits	Time	Hits/s	Hits	Time	Hits/s
3126	116	26.95	2631	85	30.95
2418	111	20.84	2544	107	23.78
2597	106	22.39	2262	90	25.13
2468	112	21.28	2163	98	22.07
2234	89	19.26	2423	99	24.47
2568.6	106.8	22.14	2404.6	95.8	25.28

With Normalisation