Parallel Genetic Programming

Francisco Fernández de Vega
Grupo de Evolución Artificial
University of Extremadura (Spain)

http://gea.unex.es

Summary:

- Introduction.
- History.
- Parallel GP.
- The Island Model.
- Successful Applications.
- Future.
Some considerations

- Genetic Programming can be considered a Machine Learning system:
  - “[machine learning] is the study of computer algorithms that improve automatically through experience [Mitchell, 1996].
- The emphasis is on learning (instead of on knowledge).
- The dream of computers that program themselves [Samuel, 1963] could be reached soon.
What’s Genetic Programming?

- According to Banzhaf et al*, GP is a system that induce computer programs by evolutionary means.
- GP (Koza, 1992) is a kind of Evolutionary Algorithm (EA). (Genetic Algorithms, Genetic Programming, Evolutionary Strategies, Evolutionary Programming).
- EAs can be seen as search techniques
  (stochastic search technique).


Different Search Techniques

Search Techniques

- Enumeratives
- Calculus Based
- Stochastic

- Evolutionary Algorithms
- Simulated Annealing
- Neural Networks
- Beam Search
- Genetic Programming
- Genetic Algorithms

As classified by Banzhaf et al
How does an EA work

- A summary:
  - T=0;
  - Initialize and evaluate \[ P(t) \]
  - While not stop_condition do
    - \[ P'(t) = \text{variation}[P(t)] \]
    - Evaluate \[ P'(t) \]
    - \[ P(t+1) = \text{select}[P'(t), P(t)] \]
    - \[ T = t+1 \]
  - end while
How does GP work?

Genetic Operators:
• Crossover.
• Mutation.
• Selection.
• Reproduction.

Fitness Function

Genetic programming
Genetic Programming

Mutation Point

After the operation

\[ b + a \times (b - c) \]

Subtree to be deleted

Before the operation

\[ b + a \times (b - c) \]

New subtree randomly generated

Genetic programming

Selection

Fitness Function

Fitness values

Selection

Genetic programming
Summary:

- Introduction.
- History.
- Parallel GP.
- The Island Model.
- Successful Applications.
- Future.
There are two main ideas behind Parallel EAs:

- Increase performances:
  - in principle, by adding processors, memory and interconnection networks and putting them to work together on a given problem.
  - Modifying the underlying algorithm can also help in the finding of solutions.

Ideas involving both EAs and Parallel Computing can be traced back to Holland, 1976.
But the field had to wait until early 1980s when parallel implementations appear.
Grefenstete, 1981, was one of the first in examining some issues concerning parallel implementations of Gas.
History

- Other researchers began more systematic studies: Gross, Cohoon, Tanese, Pettey, Georges-Schleter, Mühlenbein, and Manderick*.
- They studied Hypercubes parallel architecture, distributed models, theoretic models, island models, and cellular model.


Parallelism & EAs

- Flynn model is still widely accepted for classifying computer architectures.
- The taxonomy is based on the notion of instruction and data stream:
  - SIMD: Single Instruction, Multiple Data stream (the preferred model).
  - MISD: Multiple instruction, single data stream.
  - MIMD: Multiple instruction, multiple data stream.
- Shared or Distributed Memory.
Taxonomy

- Parallel EAs can be classified attending to different features.
  - M. Nowostawski and R. Poli 1999:
    - Master/Worker: A single population and the fitness evaluation of multiple individuals in parallel.
    - Static subpopulations with migration.
    - Static overlapping subpopulations without migration.
    - Massively Parallel genetic algorithms.
    - Dynamic demes.
    - Parallel Steady-state Ga
    - Parallel Messy Ga
    - Hybrid methods.

---

Taxonomy

- M. Tomassini, 1999:
  - Global parallel evolutionary algorithm (parallelization at the fitness level) also called by the author master/worker model (coarse-grained model).
  - Island distributed evolutionary algorithms (population based approach).
  - Cellular evolutionary algorithm (fine-grained model).
Structured EAs

- Structured populations has been used for improving EAs.
- Two main types of algorithms:
  - Distributed EAs.
  - Cellular EAs.
- On the opposite side, panmictic EA is the classic model.
- None of the models require a parallel implementation.
Nonstandard Structured EAs

- We could use different parameters/representations in different subpopulations (Tanese, Lin & Punch & Goodman, Herrera & Lozano & Moraga).

- These algorithms are sometimes called heterogenous.

Summary:

- Introduction.
- History.
- Parallel GP.
- The Island Model.
- Successful Applications.
- Future.
Why should we use a parallel model?
• We want to increase performances.

How could we parallelise?
• At the individual level.
• At the population level.
• At the fitness evaluation level

Parallelising at the fitness level.

This model is also called “global model”.

Parallelising at the population level (also called island model or coarse-grain model).

Fine-grained model (also called Grid or Cellular model)
Parallel GP: A review

- Juillé and Pollack 1995 presented one of the first attempts to parallelize GP.
- They implemented a global model, although also presented some result using sub-populations.
- The proposal was somehow specific for the SIMD model they were using.
- One of their aims was to reduce interprocessor communication.

Parallel GP: A review

- P. Tuffs, 1995, presented the same year another master/worker parallel version of GP.
- He approached a classification problem by means of GP. -the development of a system to do data mining on a fairly large (multi-gigabyte) database of credit-card transactions. The task was to classify customers and predict their future behavior-
- Probably it was the first attempt to parallelize GP (the book correspond to year 1993, although was published on 1995).
Parallel GP: A review

- Andre and Koza 1996 presented another Parallel GP Implementation using a network of 66 transputers (VLSI device containing 32 bit on-chip processor, memory and links)
- They employed the island-model.
- An appropriate migration rate showed improvement in the computational effort required for the Boolean 5-parity function. (64 demes, SubPop_size=500, Migration_rate=5%)
- The main conclusion was that Parallel GP could achieve super-linear speedups.

Parallel GP: A review

- Several processors generate independently its own segment of the next generation.
- They implemented crossover in parallel: Individuals from different processors may undergo crossover.
Parallel GP: A review

- Oussaidène et al 1997, presented a parallel implementation of GP for trading model induction.
- They employed the global model architecture (parallelization at the fitness level), employing a master/worker model, where each node from the network is in charge of evaluating individuals coming from a master node.
- The master node is in charge of the main GP algorithm.
- The model may undergo a load imbalance problem.

Parallel GP: A review

- Results offered by Koza were questioned later by W. F. Punch 1998. His experiments on the Royal Tree problem were not so optimistics. His main conclusions were that multiple-solution problems would be more amenable to multiple populations than single-solution problems.
- On the other hand, non-deceptive problems would be more amenable to multiple populations than deceptive problems.
- He only tried a set of parameters for the parallel model.
Parallel GP: A review

- Fernandez et al. 1999, presented experimental results on an island-model implementation of GP.
- Although results were preliminary, this is the first time some important parameters of the island-model are tested (communication topology).

- Folino et al. 2000, presented a new implementation of GP using the cellular model.
- Fernández et al. 2000, studied more deeply the relationship among several important parameters for the island-model (subpop size, number of subpop, communication topology).
Parallel GP : A review

- The latest results on both the island and cellular GP models have been presented very recently:
  - F. Fernández et al., 2003, described latest results with the island-model, while G. Folino et al, 2003, presented a Cellular Scalable implementation for GP. Their results are compared with previously described results using parallel GP.

Parallel GP tools

- Many GP tools allow the use of “demes” but simulated in a sequential fashion.
- There have been several parallel implementations during the last few years.
- Several languages (C, java, C++) and communication frameworks (sockets, java rpc, pvm, mpi …) have been employed.
Communication Tools

- **GLOBUS.**
- **Others (Sockets, Java-RMI…)**

Parallel GP tools

- **Chong, 1998**, presented DGP, a java based distributed approach to genetic programming on the Internet.
- **F. Fernández et al., 1999**, developed a parallel GP tool implementing the island-model, and communicating subpopulations by means of PVM. This tool was later improved by means of MPI (Fernández et al., 2000).
- **Spezzano et al., 2001**, presented CAGE: A tool for parallel genetic programming applications. They implemented the cellular model.
- Classic LiiGP software (**Punch**) also has a couple of parallel version implemented using PVM and MPI (see Fernandez parallelilgp)
Parallel GP tools

- **DREAM** project: It is aimed at providing a framework for evolutionary computation.
- It allows distributed computing.
- Any Evolutionary Algorithm could be used, by adjusting some parameters, within DREAM.
- Founded by European Union.
- See: [http://www.dcs.napier.ac.uk/~benp/dream/dream.html](http://www.dcs.napier.ac.uk/~benp/dream/dream.html)

---

Parallel GP tools

- **Paradiseo**: Parallel and Distributed Evolving Objects.
- It is based on **EO** (Evolutionary Computation Framework).
- Includes tools for:
  - Population Based Metaheuristics.
  - Single Solution Based Metaheuristics.
  - Multi-Objective Metaheuristics.
  - Parallel and Distributed…
Parallel GP tools

- **ECJ**: A Java-based Evolutionary Computation Research System.
- Includes asynchronous Island Model over TCP/IP.
- Multiobjective Optimization.

Summary:

- Introduction.
- History.
- Parallel GP.
- The Island Model.
- Successful Applications.
- Future.
Island Model

- Traditionally experimental results are shown comparing Fitness/Generation.
- Two reasons for avoiding this kind of comparisons:
  - The Bloat phenomenon in GP.
  - Populations with different size require different time to evaluate a generation.
Proposal: Evaluate Fitness/Effort (for convergence) or Fitness/Time (for speedup)
(Fernández, Galeano, Gómez).

Computational effort:
The total number of nodes GP has evaluated for a given number of generations.

$$p = \text{the number of populations}$$
$$i = \text{the number of individuals per population}$$
$$\text{avg}_\text{length}_g = \text{the average length of individuals in all the populations in generation } g.$$ 

The computational effort $E_g$ at a generation $g$ is:

$$E_g = PE_g + PE_{g-1} + \ldots + PE_1 + PE_0$$
Island Model – Comparing Results

Island Model - Topology


The Island Model
Island Model - Topology

- **Punch, 1998**, used a typical Island model with ring topology.

  He employed that topology for his experiments but no comparisons with different topologies were provided in the paper.

Fernandez et al. 2000 introduces a random topology and compare it with grid and ring topology.

- Random Topology: It changes dynamically.

The main conclusion is that if the remaining parameters stay fix, there are no significant differences when changing topology.
Island Model - Topology

![Graph showing the Island Model topology with different grid, circle, and random patterns.](image)

The Island Model 51

Island Model - Topology

![Graph showing another aspect of the Island Model topology with grid, circle, and random patterns.](image)

The Island Model 52
Island Model – Migration Rate

- How many individuals should migrate each migration step?
- Depending on the number of individuals, results are different. The limits are:
  - 0 individuals migrating (isolated populations)
  - All the individuals migrating.
- Different migration rates applied in literature:
  - Juillé & Pollack: 1 migrating individual per subpopulation.
  - Andre and Koza: 0%-8% migrating individuals per subpop.
  - Punch 1998: 2 individuals.

Fernandez et al 2003: best migration rate is between 5% and 10% in 4 test problems (2 classic and 2 real-life problems):
- Even Parity 5.
- Ant Problem.
- Routing and Placing circuits on FPGAs.
- Medical Diagnosing.
Island Model – Migration Rate

The Island Model
Island Model – Migration Frequency

- Both Juillé & Pollack and also Andre and Koza employ migration every generation.
- In Punch 1998, subpop. wait for 10 generations before the migration step.
- In Fernandez et al 2003, a wider study have been carry on, comparing different frequencies.
- Best convergence results appear when about 10% of individuals from each subpopulation are sent every 5-10 generations.
Island Model – Migration summary

[Graph showing migration rates across different populations and periods.]

FPGA Problem

The Island Model

Island Model – Migration Rate

[Graph showing migration rates across different populations and periods.]

ANT Problem

The Island Model
Island Model – Migration Summary

- For large values of the grain, exchange individuals less frequently.
- For low values of the grain, exchange more frequently.
- Recommendation: Exchange 10% of the population every 10 generations.

Island Model – Subpop. Size

- Experiments presented by Andre & Koza, make use of the large computational capability of the network they employed.
- 32000 individuals are distributed among 64 demes, each with 500 individuals.
- Punch 1998, employed 5 populations, 200 individuals each, and also 7 populations, 700 individuals each.
F. Fernández et al., 2003, presents a set of trials.

**Conclusions:**
- There is a number of individuals with which best results are obtained (regardless of the number of subpops).
- We must carefully select the number of subpops, not any number of populations obtain the same results.

---

**Island Model – Subpop. Size**

---

The Island Model

63
Island Model – Synchronisation

- Synchronism is an important issue when using Parallel GP: different individuals may require different processing time for their evaluation.
- Two models:
  - Synchronous: Exchange step takes place at a given generation.
  - Asynchronous: Populations send individuals when they are ready, and check every generation if new incoming individuals are awaiting.
Island Model – Synchronisation

- Andre and Koza worked with an asynchronous model: each generation is typically working on different generations after a few ones.
- Dracopoulos and Kent, employed the synchronous model in both the global and island models.
- Fernández et al, 2002, presented a study comparing synchronous and asynchronous models in monoprocessor systems.

---

Synchronous - Asynchronous

Results obtained using 1 processor

Synchronous model better for monoprocessor system.
Asynchronous model better for parallel systems.
Synchronous - Asynchronous

- Tongchim and Chongstitvatana, presented a comparison among models using a restricted migration policy.
- Their results only focus on a problem, and shows better performance with the asynchronous model.

---

Island Model - Bloat

- The Island Model seems to prevent the bloat phenomenon.
Some comments on diversity

- Genotypic diversity decreases.
- Phenotypic diversity improves.

What about fault tolerance?

- Fault tolerance is an important issue.
- Different techniques have been employed:
  - Check pointing.
  - Redundancy.
  - Others…
- Is GP Fault tolerant?
What about fault tolerance?

The Cellular Model

- Folino et al, 2003, have presented a comparison with panmitic and island model approach.
- The method provides results of similar quality than the island model (an small error in the comparisons seems to favor the cellular model in the paper, but a detailed revision shows that results are similar).
- They apply the model to induce decision trees.
The Cellular Model

- Each individual is located in a grid position.
- Individuals interact only with their neighboring ones.

Summary:

- Introduction.
- History.
- Parallel GP.
- The Island Model.
- Successful Applications.
- Future.
Some Applications

- C. Miccio et al, 1995, described an implementation on a T3D computer for inducing binary decision diagrams.
- M. Oussaidène et al, presented an application to trading model induction.
- F. Fernández, 2001, described a proposal for solving the problem of Placement and Routing of circuits on FPGAs.
- Folino apply the cellular model to generate decision trees.

---

Some Applications

- Koza et al 2000, presents a list of “Human-Competitive results obtained by Means of Genetic Programming”, including:
  - Synthesis of Analog Circuits.
  - Synthesis of PID controllers.
  - Applications to biomedicine (protein detection).
  - Previously patented inventions, reinvented.
  - Some patented invention.
Summary:

- Introduction.
- History.
- Parallel GP.
- The Island Model.
- Successful Applications.
- Future.

Future

- Some topics for future research:
  - Theoretical models.
  - Heterogeneous models.
  - Improvements by means of better scheduling policies.
  - Bloat phenomenon.
References


References

References


References

References


References

References


References

- **W. Punch**, iligp: [http://garage.cse.msu.edu/software/lil-gp/lilgp-index.html](http://garage.cse.msu.edu/software/lil-gp/lilgp-index.html)
References


References