Evolutionary Algorithms as Fitness Function Debuggers

Proceeding of ISMIS'99, LNAI 1609, Springer Verlag, pp 639–647, 1999.

F. Mansanne[1], F. Carrère[2], A . Ehinger[2], and M. Schoenauer[3]

Abstract. All Evolutionary Algorithms experienced practitioners emphasize the need for a careful design of the fitness function. It is commonly heard, for instance, that *"If there is a bug in your fitness function, the EA will find it"*. This paper presents a case study of such a situation in the domain of geophysical underground identification: some weird solutions are found by the Evolutionary Algorithm, obviously physically absurd, but fulfilling almost perfectly the geophysical criterion.

1 Introduction

Among the crucial steps of the design of an Evolutionary Algorithm, practitioners mostly focus on parameter tuning, with work ranging from good-old systematic trial-and-error [16] to meta-optimization [7] and to more recent self-adaptive strategies (initiated in [20], see [8] for a recent survey). The design of an efficient representation is also recognized as a crucial step, as witnessed by the never-ending debate binary vs real for real-parameter search spaces [6, 1, 9], or more recent attempts to actually separate representation from operators [22]. The choice of operators, somehow dual of the choice of representation [2], also raised a lot of discussions [5, 11]. And, to a less extend, some works addressed the initialization issue [23, 12]. But it is amazing to see that the design of the fitness function is generally omitted from the discussions. One exception is given by the field of constraint handling, for which a range of penalty methods (see [14] for a survey) can be seen as addressing the issue, though somehow indirectly: the fitness function is modified in order to take into account the constraints, but the main objective function is itself hardly discussed.

Of course, only real-world problems make the issue of the objective function worth investigating: when the function to optimize is a given analytic function, little room is left for improvement. But in complex domains, a blind use of "state-of-the-art" objective function within an Evolutionary Algorithm may sometimes lead to unexpected results: The optimization algorithm lacks the "common sense" that implicitly prevents the experts to even think of looking for solutions in some yet unexplored regions of the search space. On the other hand, such unexpected strange results might allow for new insights into some yet unquestioned long-used criterion, yielding to a better understanding of the problem at hand.

[1] Laboratoire de Mathématiques Appliquées, Université de Pau, France
[2] Geophysics and Instrumentation, IFP, Pau, France
[3] Centre de Mathématiques Apliquées, UMR CNRS 7641, Ecole Polytechnique, Palaiseau, France

This paper presents a case study of "fitness debugging" using Evolutionary Computation in the domain of geophysical inverse problem solving: The goal is to identify the velocity distribution in the underground from surface records of reflection profiles of elastic waves. The problem is presented in section 2, together with the fitness function derived from the "aligned iso-X" criterion that geophysicists have been using for a long time. The evolutionary algorithm, using the variable length representation based on Voronoï diagram introduced in [17] and used on the same geophysical problem, but with a different fitness function, in [18], is described in section 3. Results on simple synthetic 2D simulated data of section 4 illustrate the weaknesses of the iso-X criterion: the algorithms exhibits weird solutions that are better than the known synthetic solution with respect to the iso-X criterion – and that would clearly be immediately eliminated by human eyes. But is it that easy to introduce "common sense" in evolutionary algorithms? The conclusion section discusses this issue.

2 The geophysical problem

The goal of the inverse problem in geophysical prospection is to identify the velocity distribution in the underground from recorded reflection profiles of elastic waves. A seismic experiment starts with an artificial explosion (a *shot*) at some point of the surface. The acoustic waves propagate through the underground medium, eventually being reflected by multiple interfaces between different media (see Figure 1). The reflected waves are measured by some receptors on the surface, actually recording pressure variations along time, termed *seismograms*. The aim is to identify the repartition of the velocities in the underground domain from the seismograms obtained from a sequence of shots.

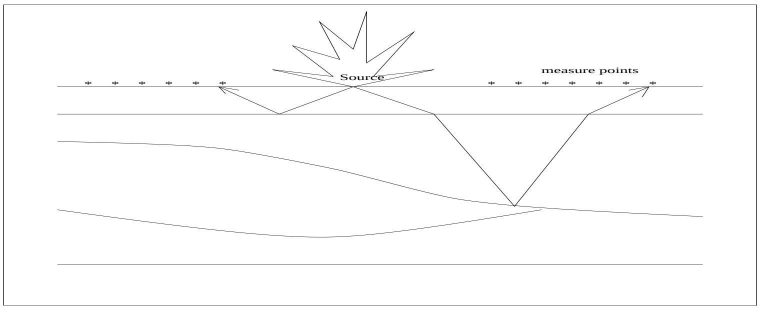


Fig. 1. *A two-dimensional view of a typical seismic-reflection experiment.*

Several methods for the determination of underground velocities have been proposed. The most successful [4] makes a clever use of migration velocity analysis, a method for determining long wavelength velocity structure, which is successfully applied to 2D data from North Sea. However, a prerequisite to the efficient description of a model in this method is the identification of regions of smooth variation.

Another way to handle the problem is to directly compute the difference between observed and simulated wave fields, as has been done in [21] (as well as in [18], using the same evolutionary approach based on Voronoï diagrams described in section 3). But such approach requires solving the wave equation in the whole domain, which is highly resource consuming when accuracy matters. Moreover, the velocity information can be derived solely from travel times. Hence, it has been decided here to work with migrated data rather than with original data: only the “useful” part of the wave field has to be computed, for each shot-receptor pair.

2.1 Migrated images

The principle of a prestack shot record migration is the following (see [24] for a more detailed presentation): Two wave fields are considered: the wave propagated from the seismic source (the explosion of figure 1), and the virtual wave obtained by backward propagation from the receptors down to the diffracting obstacles of the upward traveling diffracted wave.

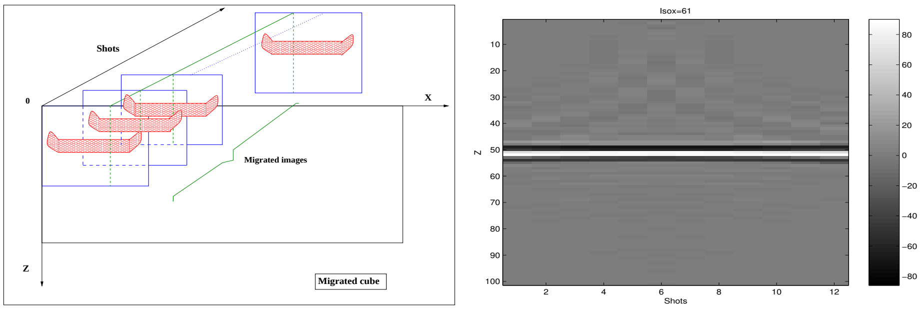
The key idea is that diffracting points are the only points where these two waves coincide in time. Hence, the so-called *Clearbout imaging principle* can be applied. It answers the question of, telling whether, at a certain point in space, there is time coincidence of the source wave and the backward-propagated diffracted wave.

Thus, a migrated image can be considered as a map of coefficients of reflection, the highest coefficients assessing the presence of reflectors. The imaging step finally concludes the migration process, yielding an image of the subsurface. The inverse problem of underground identification can then be addressed through as a maximization problem based on the *semblance* of the migrated images.

2.2 The objective function

Most practical underground velocity determinations are based on image gathers analysis (for example, migration velocity analysis in [4]). In an image gather each trace represents a migrated image of the subsurface at the same horizontal position. The objective function relies on the fact that reflection events in an image gather are horizontally aligned if the underground velocity model is correct (Figure 2). To measure the horizontal alignment of the reflection events in an image gather, the criterion first proposed in [24], and applied with success in [10] to 1D seismic profile from the North Sea, will be used hereafter, yielding the following expression, where ξ represents the set of receptors:

$$Fitness = \sum_{X \in \xi} \frac{|| \sum_{i=1}^{i=nshots} trace_i(X,Z)||}{\sum_{i=1}^{i=nshots} ||trace_i^2(X,Z)||} \tag{1}$$



(a) The whole Shots × X × Z cube *(b) Cross-section at a given X (receptor 61)*

Fig. 2. *The aligned events in the plane Shots × Z (plot b) correspond to the simple test case of section 4.*

If the underground velocity repartition is correct, traces for different positions at the surface location are aligned in the $Shot \times X$ cross-sections (Figure 2-b) and the fitness is hence maximal. Next section introduces the the evolutionary algorithm used to optimize the above fitness.

3 Evolutionary underground velocity identification

A number of authors recently addressed seismic processing problems with EA techniques. In [21], a genetic algorithm with multiple crossover locations is developed for the inversion of plane-wave seismograms; in this work, the underground is represented by a series of layers. In [3], the 'pseudo-subspace' method for the inversion of geophysical data is proposed: the complexity of models is progressively increased, but following a user-defined evolution. In [4], a classical GA algorithm is used with spline coefficients encoded into a binary string.

The central issue when addressing such inverse problem is the trade-off one has to make between the complexity of the representation and the accuracy of the models in the resulting search space. Whatever underground velocity model is used, it has to be later translated into a regular mesh which is the input data to the finite difference wave propagation algorithm. As an example, a grid of 400x400=160,000 points can be represented as a full 160,000-long vector in the full parametric approach (one velocity per element of the mesh), packed into 40x20=800 spline coefficients if spline approximation is preferred, or into a few dozens of Voronoï sites, each involving 3 real coefficients in the approach chosen here.

3.1 Voronoï representation

The underlying hypothesis in this work is that the underground domain can be partitioned into regions of homogeneous velocity, i.e. the velocity is supposed

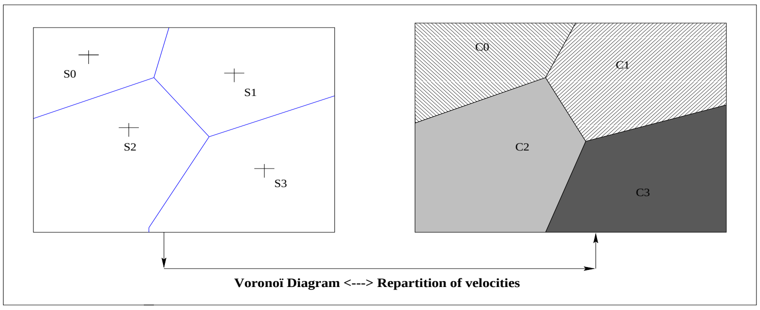


Fig. 3. *The Voronoï representation: Voronoï sites S_i define a partition of the domain into polygonal subsets in (a). Associate a velocity v_i to each cell C_i to obtain the piecewise constant function shown in (b).*

piecewise constant. The idea is to evolve both the partition and the real-valued velocities.

A possible way of representing partitions of a given domain comes from computational geometry, more precisely from the Voronoï diagram theory (see [15] for a detailed introduction to Voronoï diagrams). The representation of partitions by Voronoï diagrams has been introduced to tackle Optimum Design problems [17], identification problems in Structural Mechanics [19], and, recently, to address the geophysical inverse problem of section 2 [18].

Voronoï diagrams Consider a finite number of points $S_0, \ldots, S_N$ (the *Voronoï sites*) of a given subset of $\mathbb{R}^n$ (the design domain). To each site S_i is associated the set C_i of all points of the design domain for which the closest Voronoï site is S_i, termed *Voronoï cell.* The *Voronoï diagram* is the partition of the design domain defined by the Voronoï cells. Each cell C_i is a convex polyhedral subset of the design domain, and any partition of a domain of $\mathbb{R}^n$ into convex polyhedral subsets is the Voronoï diagram of at least one set of Voronoï sites.

The genotype Consider now a (variable length) list of Voronoï sites, in which each site S_i is given a real valued coefficient v_i. The corresponding Voronoï diagram represents real-valued piecewise constant function if all points in each Voronoï cell are given the velocity value of the associated site (here the Voronoï diagram is supposed regular, i.e. to each cell corresponds exactly one site). Example of Voronoï representations can be seen in Figure 3, where the Voronoï sites are the dots in the center of the cells.

Note that the complexity of this representation does not depend in any way on the discretization that will be used to compute the response of the model to wave propagation. Moreover, as the number of Voronoï sites is adjusted by the algorithm itself, the complexity is in fact *self-adaptive.*

3.2 Evolution operators

In order to apply evolutionary algorithms to the Voronoï representation, one needs to define some crossover and mutation operators. All details can be found in [17], [19] or [18].

- The crossover operator exchange Voronoï sites on the basis of geometrically-based choice. The same (randomly chosen) straight line divides both parents into two regions, and Voronoï sites on both sides of that line are exchanged.
- The main mutation operator performs a Gaussian mutation on the coordinates and/or on the velocities of the sites. As in Evolution Strategies, the standard deviation is self-adapted, i.e. undergoes a log-normally distributed random mutation, as proposed by Schwefel [20].
- Mutation operators for variable-length representations must include random addition and destruction of some items in the list.

3.3 The algorithm

The evolution paradigm used in all experiments is similar to "traditional" Genetic Algorithm: population size of 30; binary tournament selection and generational replacement (with elitism); crossover rate of 0.7 and mutation rate of 0.6. The algorithm stops after a fixed number of evaluation without improvement of the best fitness.

The value of the mutation rate may seem high in the light of GA standards. But first this is the *individual* mutation rate, as opposed to usual probability to flip each bit; Second, it should be related to the self-adaptation of standard deviations: The underlying hypothesis of the self-adaptation mechanism is that there exists some correlation between the value of the standard deviations of an individual and the location of that individual on the fitness landscape. Such correlation is obviously lost if mutations happen too rarely. Several experiments (see [13]) witness for that – and also enforce the usefulness of crossover in that framework.

4 Results on a synthetic problem

4.1 Experimental conditions

The chosen synthetic test starts from an homogeneous repartition of velocity ($2000ms^{-1}$) in a 2D area of depth $1000m$ and width $3100m$, with one single horizontal reflector at depth $500m$ across the whole domain. For all 12 shots that are considered, the acquisition system consists in 101 receptors. Shot sources and receptors are equi-distributed on the surface with respective spacings of $50m$ and $25m$. Sampling rate is $4ms$ and the record length is $2s$. Velocities are searched between $1000ms^{-1}$ and $3000ms^{-1}$.

The direct problem is solved using the IFP dedicated solver. The cost of one fitness evaluation is $70s$ on a Dec Alpha 500 workstation, which will be reduced to

35s on a Nec vector calculator (most of the computing time is obviously devoted in the evaluation of the fitnesses). Hence the maximum number of evaluations has to be severely limited to 2500 in all runs of section 4 and a single run still takes around 24h.

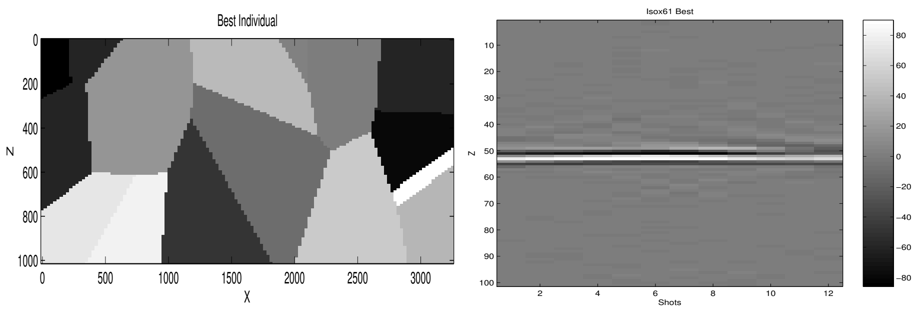
4.2 Validation results

In the first experiments, the number of Voronoï sites was severely limited (2 or 5 sites at most). The reflector was perfectly identified, and the homogeneous velocity approached within a few percent.

The migrated images are horizontal at the depth of $500ms^{-1}$, exactly at the position of the only reflector, and the iso-X plots are strictly horizontal (see figure 2-b). The conclusion of these early experiments was that the chosen algorithmic approach seemed to work quite right.

4.3 A bug in the fitness?

To further illustrate the actual effectiveness of the proposed approach, the constraint on the maximum number of Voronoï sites was relaxed. And strange results began to emerge, as the one plotted in Figure 4-a. Indeed this individual cannot be thought as a solution by anyone having even the smallest experience in seismic data. But it did give optimal values of the fitness function defined in section 2, as witnessed by a sample Iso-X diagram of Figure 4-b, almost "as horizontal as" the one of the perfect solution of Figure 2-b.



(a) *Black*=1300 ms^{-1}, *white*=2700 ms^{-1} (b) *Iso-X plot for receptor 61*

Fig. 4. *The absurd underground velocity repartition (a) is miles away from the exact solution (homogeneous* $2000ms^{-1}$ *velocity), and nevertheless gives almost perfect iso-X plot – compare (b) with Figure 2-b!*

At this stage, it is not possible to say if the model itself is too simple compared to the degrees of freedom of the individuals, or if the formula of equation (1) used to test the horizontality is too raw, or, more dramatically, if the horizontality criterion itself should be questioned. But these results did raise many issues among the expert geophysicists, who are how trying to understand the case.

Conclusion

The above results clearly enlighten the need for some guidelines when transposing human-designed criterion to computer optimization: human experts in fact use a lot of implicit criteria when tackling complex problems. Those criteria usually don't need to be explicitly written down as long as other experts with the same scientific background are addressing the same kind of problem. But computer programs don't have any scientific background – unless explicitly programmed.

The question thus becomes how to incorporate "common sense" in our favorite algorithms? As for the seismic inversion problem addressed in this paper, some straightforward partial answers can be given to avoid weird underground velocities like the one of Figure 4: the idea would be to forbid velocity distribution that look too chaotic – but roughness is not easy to measure. A first could be to forbid too high discontinuities between adjacent Voronoï cells. But the implementation of such constraint is not straightforward (see [14]): should one penalize high jumps in velocities – with fixed or dynamic penalties? – or repair bad individuals by modifying their genotype, ... ? Moreover, whatever the constraints on the genotype, there is absolutely no guarantee at all that other types of weirdos will not emerge, meeting these artificial constraints, but finding other unguarded common sense holes.

On the other hand, such additional constraints, though difficult to design and to incorporate efficiently are reliably in the algorithm, might prove beneficial for the whole search process, by limiting the size of the space to be searched. As a matter of fact, even if complexity was not its main drawback, the mesh-based representation, where the unknown parameters to identify are made of one real-valued parameter (the velocity) in each element of a given mesh, suffers from the high uncorrelation of adjacent velocities. For instance, pure random choice of the one velocity per element results in highly non-physical underground models (see also [12] for a related discussion on initialization procedures).

To conclude, let us say that, even though experts are sometimes doubtful about the ability of Evolutionary Algorithms to efficiently optimize the function that took them years to begin to understand, they are indeed convinced by the discovery of unbelievable cases such as the one presented in that paper.

References

1. J. Antonisse. A new interpretation of schema notation that overturns the binary encoding constraint. In J. D. Schaffer, editor, *Proceedings of the* 3^{rd} *International Conference on Genetic Algorithms*, pages 86–91. Morgan Kaufmann, June 1989.
2. D.L. Battle and M.D. Vose. Isomorphisms of genetic algorithms. In G. J. E. Rawlins, editor, *Foundations of Genetic Algorithms*, pages 242–251. Morgan Kaufmann, 1991.
3. F. Boschetti. *Application of genetic algorithms to the inversion of geophysical data*. PhD thesis, University of Western Australia, 1995.
4. P. Docherty, R. Silva, S. Singh, Z. Song, and M. Wood. Migration velocity analysis using a genetic algorithm. *Geophysical Prospecting*, 45:865–878, 1997.

5. D.B. Fogel and L.C. Stayton. On the effectiveness of crossover in simulated evolutionary optimization. *BioSystems*, 32:171–182, 1994.
6. D. E. Goldberg. A theory of virtual alphabets. In Hans-Paul Schwefel and Reinhard Männer, editors, *Proceedings of the* 1^{st} *Parallel Problem Solving from Nature*, pages 13–22, 1991.
7. J. J. Grefenstette. Optimization of control parameters for genetic algorithms. *IEEE Trans. on Systems, Man and Cybernetics*, SMC-16, 1986.
8. R. Hinterding, Z. Michalewicz, and A. E. Eiben. Adaptation in evolutionary computation: A survey. In *Proceedings of the* 4^{th} *IEEE International Conference on Evolutionary Computation*, pages 65–69. IEEE Press, 1997.
9. C. Z. Janikow and Z. Michalewicz. An experimental comparison of binary and floating point representations in genetic algorithms. In R. K. Belew and L. B. Booker, editors, *Proceedings of the* 4^{th} *International Conference on Genetic Algorithms*, pages 31–36. Morgan Kaufmann, 1991.
10. S. Jin and P. Madariaga. Background velocity inversion with a genetic algorithm. *Geophysical research letters*, 20(2):93–96, 1993.
11. T. Jones. Crossover, macromutation and population-based search. In L. J. Eshelman, editor, *Proceedings of the* 6^{th} *International Conference on Genetic Algorithms*, pages 73–80. Morgan Kaufmann, 1995.
12. L. Kallel and M. Schoenauer. Alternative random initialization in genetic algorithms. In Th. Bäck, editor, *Proceedings of the* 7^{th} *International Conference on Genetic Algorithms*, pages 268–275. Morgan Kaufmann, 1997.
13. F. Mansanne. . PhD thesis, Université de Pau, In preparation.
14. Z. Michalewicz and M. Schoenauer. Evolutionary Algorithms for Constrained Parameter Optimization Problems. *Evolutionary Computation*, 4(1):1–32, 1996.
15. F. P. Preparata and M. I. Shamos. *Computational Geometry: an introduction*. Springer Verlag, 1985.
16. J. D. Schaffer, R. A. Caruana, L. Eshelman, and R. Das. A study of control parameters affecting on-line performance of genetic algorithms for function optimization. In J. D. Schaffer, editor, *Proceedings of the* 3^{rd} *International Conference on Genetic Algorithms*, pages 51–60. Morgan Kaufmann, 1989.
17. M. Schoenauer. Representations for evolutionary optimization and identification in structural mechanics. In J. Périaux and G. Winter, editors, *Genetic Algorithms in Engineering and Computer Sciences*, pages 443–464. John Wiley, 1995.
18. M. Schoenauer, A. Ehinger, and B. Braunschweig. Non-parametric identification of geological models. In *Proceedings of the* 5^{th} *IEEE International Conference on Evolutionary Computation*. IEEE Press, 1998.
19. M. Schoenauer, L. Kallel, and F. Jouve. Mechanical inclusions identification by evolutionary computation. *European J. of Finite Elements*, 5(5-6):619–648, 1996.
20. H.-P. Schwefel. *Numerical Optimization of Computer Models*. John Wiley & Sons, New-York, 1981. 1995 – 2^{nd} edition.
21. P.L. Stoffa and M.K. Sen. Nonlinear multiparameter optimization using genetic algorithms : inversion of plane-wave seismograms. *Geophysics*, 56, 1991.
22. P.D. Surry and N.J. Radcliffe. Formal algorithms + formal representations = search strategies. In H.-M. Voigt et al., editors, *Proceedings of the* 4^{th} *Parallel Problems Solving from Nature*, pages 366–375. Springer Verlag, 1996.
23. P.D. Surry and N.J. Radcliffe. Inoculation to initialise evolutionary search. In *AISB Evolutionary Computing*, LNCS 1141, pages 269–285. Springer Verlag, 1996.
24. M.T. Taner and F. Koehler. Velocity spectra-digital computer derivation and applications to velocity functions. *Geophysics*, 56:1794–1810, 1969.