

Evolving hominins in HomininSpace — Genetic algorithms and the search for the ‘perfect’ Neanderthal

F. Scherjon

Man’s longing for perfection finds expression in the theory of optimization. It studies how to describe and attain what is Best, once one knows how to measure and alter what is Good or Bad.
(Beightler et al., 1979, p. 1)

Introduction

Optimisation of agent-based models is generally seen as a complex problem. Computation of a solution involves the execution of a simulation in all but the most trivial models. The search for optimal solutions is therefore computationally expensive, especially when the level of realism is high and the number of parameter combinations is large. This research applies Genetic Algorithm based search and optimisation techniques in an attempt to characterise Neanderthal mobility using a realistic large scale agent based modelling system. The methodology and software developed in this study are part of the doctoral research of the author.

Genetic Algorithms

Genetic Algorithms (GAs) form, together with Evolutionary Strategies and Genetic Programming, the Evolutionary Algorithm programming paradigm (Coello Coello, 2002). Differences between these techniques are fading (Michalewicz, 1992, p. 132), and while the chosen terminology no longer reflects our current understanding of genetics, they all aim to simulate an evolutionary process in the computer (Coello Coello, 2002). Basic elements in evolutionary algorithms are a population of individuals to work with, a string with values that define an individual which can be manipulated (referred to as the genes or the chromosome), a fitness function that calculates how well adapted an individual is within the modelled environment (Michalewicz, 1992), and an optimisation technique targeting optimal solutions.

The term Genetic Algorithm was coined by John H. Holland (1975) but the general principle was already recognized by Alan Mathison Turing in the 1948 essay ‘Intelligent Machinery’: ‘There is the genetical or evolutionary search by which a combination of genes is looked for, the criterion being the survival value’ (Turing, 1948, p.18). The idea is that an individual represents a single solution to the problem at hand; that individuals vary from one to the other; that an evaluation function represents the environment in calculating the fitness of each individual; and that the computer searches for the individual with the highest fitness value (the optimal solution) in the problem space, without actually programming that solution, but instead by manipulating the genes of the individuals (Coello Coello, 2002).

GAs are very well adapted for nonlinear search spaces, with multiple locations present in those spaces that can yield good solutions. A non-traditional methodology is required when traversing these resulting parameter spaces searching for an optimal solution, where exhaustive search algorithms are replaced by a heuristic procedure. A GA differs from more traditional optimisation and search algorithms in at least four ways (Goldberg, 1989):

1. A GA searches in a population of points, not just a single point;
2. GAs operate with probabilistic, unbiased transition rules, not deterministic ones;
3. A GA operates on a set of parameter values, not the parameters themselves, and does not require any explicit knowledge of the actual structure of the solution space (what these parameters are about);
4. A direct, explicit fitness function is used.

For the structure of a GA, a probabilistic selection is made from the population based on some measurement of the individual's adaptation to the environment or its fitness within it. The design of the fitness function is at the discretion of the modeller. After the selection of one or more individuals, operators are applied on the selection. These operators are inspired by the early perception of how the genomes of offspring were created in nature. They include Mutation, where random changes are made to the genes of an individual, and Crossover (a special kind of Recombination), where the genes of two individuals are mixed in sequence. Then the fitness function is calculated for the generated offspring, and the new individuals are inserted into the population, generally replacing other individuals. The process is then restarted, until some stop criteria are met. Generally the stop criterion is a certain value for the fitness function being attained, or no measurable improvement of the calculated fitness values after a certain number of generations.

Agent Based Modelling

An Agent-based Model (ABM) is a complex system aiming to reproduce the dynamics of the real world, and one that cannot be solved mathematically. Therefore GAs form a perfect exploration and tuning method for such models (Calvez and Hutzler, 2006). Emerging properties, high parameter sensitivity, and non-linear solutions often characterise the solution space for such models. There is no analytical description, and simulation results cannot easily be reduced to the input data. There is development and emergence in a non-trivial way. As such they are well suited to modelling complex systems and to explore this complexity. Archaeological research questions connected to such models often incorporate the resultant output of the (inter)actions of many individuals through time, and as such can be targeted by ABM techniques (Lake, 2014). However, there have been few attempts to apply GAs to archaeological simulations, although they are implicitly used in the more widely applied Genetic Algorithm for Rule-Set Prediction tool (GARP) for biogeographical niche modelling (Banks *et al.*, 2008).

HomininSpace

HomininSpace is an agent based modelling and simulation environment where a fluctuating carrying capacity in a reconstructed paleoenvironment is the key attractor for hominin dispersal (Scherjon, 2015a). Simulations executed in the HomininSpace modelling system are used to assess the character of past hominin dispersal, taking the patterns of presence and absence of Neanderthals in North-west Europe in the Middle Palaeolithic (130kya–50kya) as a case study. Mobility, the sum of small scale movements through larger geographic and temporal scales, enables hunter gatherers to survive (Kuhn *et al.*, 2016).

Energy in the landscape, in the form of herds of medium to large ungulates, will support groups of Neanderthals moving through the reconstructed environment. That environment includes a topography influenced by fluctuating sea levels. The question is whether the dispersal and movement of Neanderthals was based on tracking preferred habitats. The tracking of favourable habitats has also been described as the 'ebb and flow' of populations

(e.g. Hublin and Roebroeks, 2009), and involves individuals or groups of individuals moving in the area where the most favourable circumstances are found. Today the best known example of such behaviour is displayed by migratory birds, which fly south towards the Mediterranean or beyond in autumn and return when spring sets in.

The 'ebb and flow' of moving populations has often been opposed to a 'sources and sinks' model, where local populations must adapt behaviourally and/or genetically to cope with the changing climate or subsequently become extinct when conditions become less favourable (Pulliam, 1988, 1996). They are replenished from more productive areas when the situation improves (MacDonald *et al.*, 2012). Obvious examples of this 'sources and sinks' model are most species of flora. Since individuals of this kind cannot move by themselves, they invariably die when the climate deteriorates too far. For that species to live there again, the area must be re-colonized from other areas where local reproduction more than balances mortality. The question is whether Neanderthal dispersal patterns more closely resemble ebb and flow movement or a sources and sinks model.

To address the mobility of Neanderthals, the two opposing types have been implemented in the parameterised model (16 parameters in total) underlying the HomininSpace simulation system. These types are termed Static and Dynamic. Simulations are executed for both types and a fitness function, and have been constructed to compare simulation results with the archaeological data. This fitness function counts and totals how often an archaeologically determined presence period is matched by a presence in the simulation at the same point in time. A period of presence is defined as the radiometrically determined date (with standard deviation) for an artefact or for sediment associated with the archaeology. For instance, it has been shown that Neanderthals were present in Pech de l'Azé IV with a thermoluminescence (TL) measurement on heated flint (Richter *et al.*, 2013). This artefact is dated to 68.5 kya, with a standard deviation on this TL date of 6.6 ky. If in any simulation a Neanderthal group is within the area of Pech de l'Azé IV during the period 75.1-61.9 kya, each year of that presence is counted and added to the total simulation score. This sum is referred to as the MatchingVisits, and is the quantitative fitness value for that specific simulation (Calvez and Hutzler, 2006, p. 47). Initial results presented in earlier research suggested that for manually selected parameter values, the Static mobility type produced the best solution, or in other words that simulation results for Static hominins matched the Neanderthal archaeology best (Scherjon, 2015a).

Each simulation is characterised by a combination of parameter values and a fitness value that is the result from running the simulation with those parameter values. This combination of parameter values is referred to as the *chromosome* for that individual simulation. GAs are then used to search for the parameter value combination that matches the archaeology best. An initial population is constructed by varying parameter values randomly, and then running the simulations with these parameter sets. Then, individual simulation results are used to select promising parameter combinations, running new simulations for the generated offspring in a search of the optimal solution.

This rest of this paper is structured as follows: the next section will describe the implemented GA and the underlying data set. Then, results for the experiments are provided and discussed. The conclusions are presented in the last section, together with some directions for further research and suggestions about future use of genetic algorithms in archaeological ABMs.

Materials and Methods

In this research the following steps were implemented:

1. The construction of a list of one thousand randomly created parameter sets that can be used as input to the HomininSpace modelling system. Default values for the parameters were selected from the literature, and the random values for the parameters were generated within the interval of 10%-200% of that default.
2. The creation of simulation results from two initial populations; one of Static and a second of Dynamic hominins. This involved running all 1000 parameters sets constructed in the first step for both mobility settings, and calculating the fitness value for each combination for both settings.
3. Run GAs for both initial populations, using as selection criterion the match with the archaeology. As two variants were applied to the Dynamic hominins, the result is three data sets.
4. Analyse these three data sets to find the most optimal solution, or the 'Perfect Neanderthal'.
5. Explain the results, using correlation analysis on the initial populations and computation of the coefficient of variation for the results from the GAs.

The parameters that are used in the model underlying the HomininSpace (HS) modelling system are described in Table 1. There are 16 parameters in total, distributed in three groups: those for Demographics, for Energetics and one for Group Dynamics. Parameters are based on ethnographic data (default) with wide plausible extended value ranges, from which instance values are randomly selected. Employing these 16 parameters, if you need to systematically explore the total parameter space, this can exponentially expand the number of parameters. For example, if you take only three values per parameter (a minimum, a maximum and a middle value), this would require the simulation of a 3^{16} parameter set (or almost 300 million simulations). Therefore the combinatorial explosion with so many parameters requires a non-exhaustive exploration. Nevertheless, due to the non-linear character of the solution space, an automatic and systematic exploration of parameter space is needed (Calvez and Hutzler, 2006).

Table 1 - The parameters and default values for the modelled hominins in HomininSpace.

The implemented GA in the HS system selects individuals from the population using a tournament selection procedure (Miller and Goldberg, 1995). For the selection of each individual, a tournament is organised in which n random individuals are chosen from the population. From this subset the highest ranking (most optimal) solution is declared winner. This procedure ensures that even mediocre solutions can produce offspring. The selected individuals participate in the creation of the next generation of seven new individuals through the application of several operators. HS implements a real-coded GA where each individual is represented by a string of 16 integer values, one for each model parameter (Herrera *et al.*, 1998). Then additional simulations for the newly generated parameter combinations are executed, the fitness value computed, and the offspring is added to the general population. This process is repeated until the end of each experiment (see Figure 1).

Figure 1 - Application of Genetic Algorithms in the HomininSpace modelling system.

The operators that are applied are crossover and mutation. Mutation is implemented as the random modification of one single parameter value by plus or minus 10%, and thus implements a low rate mutation mechanism to increase coverage of the search space and to prevent convergence to a local optimum (Yao, 1993). An additional advantage of the chosen mutation operator is that when needed, the search space can be expanded even beyond the original chosen random value domain. In other words, values can be created that are not present in the (initial) population (Djurišić *et al.*, 1997). Crossover aims to recombine two good parent solutions into potentially even better offspring solutions. Implemented is a uniform multi-point crossover, where the offspring is a single individual that is a stochastic mix of the parameter values from both parents (Syswerda, 1989). This combines very well with tournament selection (Djurišić *et al.*, 1997, p. 7860). Seven individuals (four crossover results, three mutations) are created in each generation, because the hardware running the experiments has seven parallel processors available for executing simulations.

Figure 2 - Mutation operator (left) and uniform crossover operator (right) producing offspring individuals (denoted by C).

Since offspring is added to the population without any replacement of parents or other less performing individuals, no lineage is terminated prematurely, and each individual competes until the very end of each experiment. This also ensures that the best performing individuals survive intact within each generation. However, successful lineages tend to dominate the tournaments, which can be used to create an informed stop criterion for the GA.

Results

Two experimental setups were created: one for Static hominins (those that stay in the same area even when the climate deteriorates) and one for Dynamic hominins (that constantly move to that area that has the most resources). In each experiment, initial populations were created by running 1000 simulations with randomly created parameter values (the same values for both experiments). The fitness value (referred to as MatchingVisits) is shown for both experiments in Figure 3. Static results are first plotted and then overlain by Dynamic results. Most obvious are the overall green peaks, suggesting the more optimal solutions are for Static hominins. Of interest are the simulations with better Dynamic results, which in Figure 3 are the results in red peaks, with some more promising results highlighted in blue in the figure. These two sets are used as input for the GAs, for which the results are presented below.

Figure 3 - Results for the simulations of the initial populations for both experimental setups. On the vertical axis the score for MatchingVisits, and on the horizontal axis is the simulation number. For each simulation two results are plotted; for Static in green and for Dynamic in red. In the circles there are three Dynamic peaks, meaning that for those simulations Dynamic scores higher than Static.

However, first a characterisation of the parameter influence is undertaken by calculating a Pearson product-moment correlation coefficient (PCC) for each parameter against the fitness value, in both experimental starting populations. The PCC gives the degree of linear dependence between two ratio-scale variables, and can be a positive, a zero or a negative correlation (Fletcher and Lock, 2005, p. 117). These results are presented in Table 2. For Static hominins the Birthrate, Max_ForagingRange and to a lesser extent the CohortSize_Fertile are positively correlated. In addition, negative correlations are especially clear in DeathRate_PostFertileCohort, and also to a lesser extent in DeathRate_PreFertileCohort, DeathRate_FertileCohort and GroupSizeFertile_BeforeMerge. For Dynamic hominins, a positive correlation was found for Max_ForagingRange, Birthrate and GroupSize_BeforeSplit. There is a negative correlation between fitness value and DeathRate_FertileCohort, GroupSizeFertile_BeforeMerge, and Temperature_Tolerance.

Table 2 - The Pearson correlation coefficient for each parameter against the MatchingVisits fitness value. ** denotes a significant correlation, which for easy reference is also colored red. Parameter names are taken from the source code.

GAs were executed for both experimental setups. For the Static hominins, 731 new individuals were generated. The stop criterion was an experiment duration of two weeks, and no further improvement after 13 generations (>100 individuals). The maximum value of 3,945,109 was obtained in simulation number 1638. For Dynamic hominins, 540 extra simulations were run. Here the maximum of 3,948,133 was reached in simulation 1495, and after an additional 45 simulations with no improvement the experiment was ended (end result higher than Static max). In total 3271 simulations were executed, with an average execution time of 20 minutes per simulation.

Figure 4 - GA results for both Dynamic (red) and Static hominins (green).

The top 5 results for both experimental setups are presented in Table 3. To illustrate the effect of the optimisation effort that GA can achieve, the best results from both initial populations are also included. It shows that improvement is especially spectacular for Dynamic hominins; changing from a maximum score of 2.4m to 3.9m (where Static hominins achieve an improvement from 3.5m to 3.9m). These results can be compared against results from previous research (Scherjon, 2015b) where manually constructed parameter combinations were executed for both Static and Dynamic hominins (Figure 5). Parameter values here were derived from the ethnographic literature and modified slightly to accommodate expected Neanderthal deviation (for instance, energy usage was increased, see Verpoorte, 2006). Although the implemented system was subsequently further developed, the general trend was clear: for the manually constructed parameter sets, Static simulations always scored higher than Dynamic (comparable to most of the non-GA optimized results presented in Figure 3). In that research, no higher scoring Dynamic simulations were found.

Table 3 - Top 5 results for both Dynamic and Static hominins, results with and without GA optimisation.

Figure 5 - Taken from Scherjon (2015b). For each simulation the simulation scores for Static hominins are better than those from the corresponding (same parameter value) Dynamic simulations. The simulation score quantifies the match with the actual archaeology.

Discussion

The results from previous research in Scherjon (2015b) that are reproduced in Figure 5 were explained to a large extent by the more intense resource competition that would result from Dynamic behaviour. Areas with many resources would attract all hominins from the surrounding area, who would then deplete the resource patches, with resulting food shortages for all. Apparently, though, there are hominin types that follow a Dynamic strategy and that do show a good (and even best) match to the archaeology (Table 3). To characterise these dynamic hominins, the values for all significant parameters for the 30 most optimal solutions implementing the Dynamic mobility pattern are presented in Figure 6.

Figure 6 shows that the most successful Dynamic hominins all have a relatively high birth-rate, around 46% (dark blue line). It also shows that the difference between number 1 and number 2 is only a slight decrease in the variable *GroupSize_BeforeSplit*, the variable that indicates when groups become unstable (note that there can be other differences in the non-significant variables). This variable is also the most variable one. Also observe that the value suggested in the literature is 25 (Sørensen, 2009), which here is only sported by the 25th ranked Dynamic simulation.

To understand these results further, it is important to realise that the fitness value is constructed by matching with *presence* data in the archaeological record. These data points are spread through the simulation area, both in space and time. The theoretical best match would be attained by a hominin that is present everywhere all the time (!). Such a hominin cannot be sustained by the limited amount of resources produced by the environment. So the system, implemented in a GA, searches for a sub-optimal solution, and detects a family of solutions that represents a hominin type which is not, as such, recognised in the literature. This is a hominin that constantly travels through the landscape within very small groups (around ten individuals) and with a very high birth-rate (close to the physical limit a female modern human body would be able to sustain, around 46%).

Figure 6 - The top 30 Dynamic hominins, for all significantly correlated parameters.

When inspecting the non-significant parameters for the same set of solutions (Figure 7), it becomes clear that values for these vary more than for the significant parameters. This makes sense since they influence the final result less. Contra-intuitive energy related variables are also non-significant for scoring against the archaeology (not included in the figure).

Figure 7 - The non-significant parameters for the 30 best ranking Dynamic simulations.

Conclusions

A Genetic Algorithm (GA) has been implemented in the HomininSpace simulation system (HS) to systematically explore the parameter space that results from the chosen parameter set in the underlying model. Simulation results are compared against an archaeological record of actual presence data, resulting in a quantitative fitness value per simulation. It is shown that the implemented GA is capable of finding more optimal fitting parameter value combinations that result in a higher fitness value than informed manually selected parameter values. When applied to the research question on Neanderthal mobility, it must be concluded that the results for both strategies are very comparable. The fitness values for improved individuals are within the same order of magnitude, and there is no statistically significant difference between Static and Dynamic hominins (contrasting previous research). However, it is interesting that the best matching simulations were those which, by a narrow margin, have hominins that are implementing a Dynamic mobility strategy.

The model not only implements the mobility type, but also many other parameters involved. This results in the following characterisation of the most optimal fitting solution in the model underlying HS. The 'perfect' Neanderthal:

- implements a Dynamic mobility strategy;
- has a (very) high birth-rate;
- sports low death rates for pre-fertile and fertile segments;
- has high death rates for the post-fertile segment;
- has a low energy intake;
- can resist cold fairly well;
- has a short childhood;
- operates in relatively small groups;
- hunts and collects from a large foraging range.

From these results it can be concluded that the implemented Genetic Algorithm works. It improves upon randomly constructed initial population results and falsifies previous research: the evolved Dynamic Neanderthals have an equal or better fit than all Static ones (evolved or manually constructed). The search for optimal solutions is generic, systematic and produces better results than informed manual selection of parameter values. The character of the parameter values for the set of most optimal solutions confirms the statistical analysis on the significance of certain parameters on the fitness value. This information can be used in future parameter reduction efforts on the model, but care must be taken; some parameters which are unimportant in some simulations, might be important in others, and vice versa (for example see the `Temperature_Tolerance` and the death rate for the post-fertile cohort parameter).

Genetic Algorithm techniques applied in ABM are, unfortunately, computationally expensive, since the calculation of the fitness value is the actual simulation run with the evolved parameter set. Therefore the following elements must be considered carefully when constructing a GA for an archaeological ABM: the total computational costs, stochasticity in the genetic algorithm, choice of GA operators, the stop criterion and the chosen fitness function. And as with all stochastic modelling: nothing is guaranteed!

It is acknowledged that the simulation results are matched against presence data only, and this has a major impact on the direction of the search for optimal solutions. The best match with the archaeology would be achieved by a simulated hominin that is present everywhere all the time. Such is a general issue when using archaeological data, since 'absence of evidence is not evidence of absence' (Phillips *et al.*, 2006). However, archaeologists are quite convinced that Neanderthals were not present in certain areas at certain points in time during the simulated period (Ashton, 2002; Wragg Sykes, 2017). Further research should take into account the tendency to fill the landscape when presence only information is used and should investigate the effects of incorporating evidenced absence.

Acknowledgements

Research by Fulco Scherjon was supported by the Netherlands Organisation for Scientific Research (N.W.O. Spinoza program). I thank three anonymous reviewers for their useful suggestions which significantly improved the paper and the editors for their efforts in checking the details.