


**FIRST STEPS****Contents**

1.	General Information .....	2
1.1.	About .....	2
1.2.	Authors and references .....	2
2.	How to Start.....	3
2.1.	Defining a solution pattern .....	5
2.2.	Creating a population of solution candidates .....	7
2.3.	Creating a genetic algorithm instance .....	8
2.4.	Organizing an evolving loop.....	10
2.5.	Composing a fitness function.....	12
2.6.	Inspecting results.....	15
3.	Samples.....	17
3.1.	Evo. DiagonalProblem sample .....	17
3.2.	Evo. AckleyFunction sample.....	19
3.3.	Evo. TravellingSalesmanProblem sample.....	21



## 1. General Information


### 1.1. About

**Evo** is a custom open-source package designed to be used in Dynamo environment (version 1.3.3. or newer). So far includes C#-coded nodes responsible for solving single-objective optimization problems using Genetic Algorithms (GA). The package is suitable for optimizing multi-variable number-based problems as well as combinatorial ones, e.g. [Travelling Salesman Problem](#) .

The document presents basic information about preparing visual scripts using the package. It also brings elementary terms and methodologies used in the solution as well as describes samples enclosed to the package<sup>1</sup>. Integral part of the specification are separate documents: *Evo Specification. Crossovers.pdf*, *Evo Specification. Mutations.pdf* and *Evo Specification. Selections.pdf* including theory background regarding basic GA operators: crossover, mutation and selection methodologies.

### 1.2. Authors and references

The package was created by [Marcin Jasiński](#) , a PhD candidate at the Silesian University of Technology, Faculty of Civil Engineering, Department of Mechanics and Bridges, Gliwice, Poland. The source of the package is available at [Evo](#)  GitHub repository. Contact: [marcin.jasinski@polsl.pl](mailto:marcin.jasinski@polsl.pl).



The main part of the solution is based on [GeneticSharp](#)  by **Diego Giacomelli** – a fast, extensible, multi-platform and multithreading C# Genetic Algorithm library that simplifies the development of applications using Genetic Algorithms (GA). License: MIT.





---

<sup>1</sup> Can be found in: %APPDATA%\Dynamo\Dynamo Core [or Dynamo Revit]\1.3 [or 2.0]\packages\Evo\extra

## 2. How to Start

All the package nodes are included in **Evo** → **GeneticAlgorithms** category visible in Dynamo libraries panel. The category is divided into several subcategories:

Category	Description
<p>Basic</p> 	<p>A group of nodes responsible for crucial application of the genetic algorithm solver. In general, each Dynamo script utilizing the package should contain at least one of each: <b>Basic.CreateGeneticAlgorithm</b>, <b>Evo.Init</b>, <b>Evo.ContinueWhile</b> and <b>Evo.LoopBody</b> nodes. It should also contain at least one <b>LoopWhile</b> node which is package-independent Dynamo node available under BuiltIn category.</p>
<p>Chromosomes</p> 	<p>A group of nodes responsible for definition of a solution candidate pattern, i.e. what type of variables are used, how many of them is defined in the optimization problem, what are their minimum and maximum values, what is their required precision and so on.</p>
<p>Crossovers</p> 	<p>Crossover, besides selection and mutation, is one of the basic GA operators. Crossover is responsible for mating the best parents in the input generation.. Crossover nodes define objects that are send to the <b>Basic.CreateGeneticAlgorithm</b> node.</p>
<p>Mutations</p> 	<p>Mutation, besides selection and crossover, is one of the basic GA operators. Mutation is responsible for slight modification of randomly chosen offsprings produced by the crossover. Mutation reintroduces genetic diversity back into the population and prevents from catching local optima, right after converging pressure imposed by crossover. Mutation nodes define objects that are send to the <b>Basic.CreateGeneticAlgorithm</b> node.</p>

Category	Description
<b>PopulationStrategies</b> 	<p>A small group of nodes that define the way in which generations are stored within a genetic algorithm instance. The instance is created by <b>Basic.CreateGeneticAlgorithm</b> node and iterated by <b>LoopWhile</b> built-in Dynamo node. Each iteration produces new generation. In the case of a huge problem, a significant number of generations may be produced. The strategy nodes define if all generations should be stored in the solution or only selected ones.</p>
<b>Populations</b> 	<p>A group of nodes responsible for operations on populations and their generations. A population is a set of generations. A generation is a set of candidate solutions. At the beginning, a population is generated using <b>Populations.GeneratePopulation</b> node, basing on the defined chromosome pattern. The newly generated population includes initial generation of randomly obtained solution candidates. The solver creates new generations in the population that can be then accessed by other nodes from this category.</p>
<b>Selections</b> 	<p>Selection, besides crossover and mutation, is one of the basic GA operators. Selection is responsible for picking individuals from the most recent generation to so called parents' set. The parents are then mated by one of the crossover operators and produce offspring. Selection methodologies try to select the best candidates but also try to maintain a certain diversity level in the parents' set. Selection nodes define objects that are send to the <b>Basic.CreateGeneticAlgorithm</b> node.</p>
<b>Terminations</b> 	<p>A group of nodes defining specific conditions that are required to stop a running evolving processes. Termination nodes define objects that are send to the <b>Basic.CreateGeneticAlgorithm</b> node.</p>

## 2.1. Defining a solution pattern

First of all, the algorithm should know a kind of the problem to solve. In general, so far three kinds of problems may be defined – expressed by either by binary chromosomes, combinatorial chromosomes or double chromosomes. If the problem is a single-objective function of multiple variables, that can be expressed mathematically, **Chromosomes.CreateBinaryChromosome** or **Chromosomes.CreateDoubleChromosome** nodes should be used. The first one generates a chromosome in form of a binary string representation that can be then decoded by **Chromosomes.DecodeChromosome** node (see: Fig. 1). It is recommended for optimization problems based on non-negative integer variables. Negative variables or the ones with a significant amount of fraction digits generate long binary string representations that can reduce convergence rate of the algorithm. In general, in case of numerical optimization tasks, based on real numbers with no limitation to sign and precision, **Chromosomes.CreateDoubleChromosome** is recommended (see: Fig. 3).

If the problem is a single-variable optimization, minimum and maximum values of the variable should be given by connecting Number nodes to the chromosomes nodes:

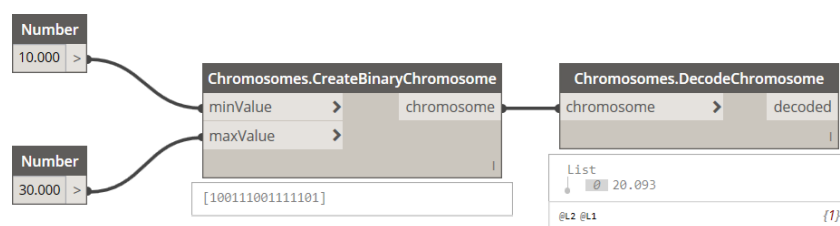


Fig. 1. Creating a binary chromosome for a single-variable function.

As seen above, output of the **Chromosomes.CreateBinaryChromosome** node has a form of binary string representation that, after decoding, is equal to 20.093. The encoded chromosome is an exemplary solution candidate that will be used to generate a bigger population of similar candidates when passed to the **Populations.GeneratePopulation** node (see: chapter 0). One can also provide required precision of the binary variable (fraction digits, 3 by default) and, if desired, number of bits (length of the binary string).

If the problem is a multi-variable optimization, instead of single values, lists of the values should be passed to the **Chromosomes.CreateBinaryChromosome** node, i.e. list of minimum values, list of maximum values and, if needed, list of fraction digits and list of binary representation lengths, as seen in the Fig. 2.

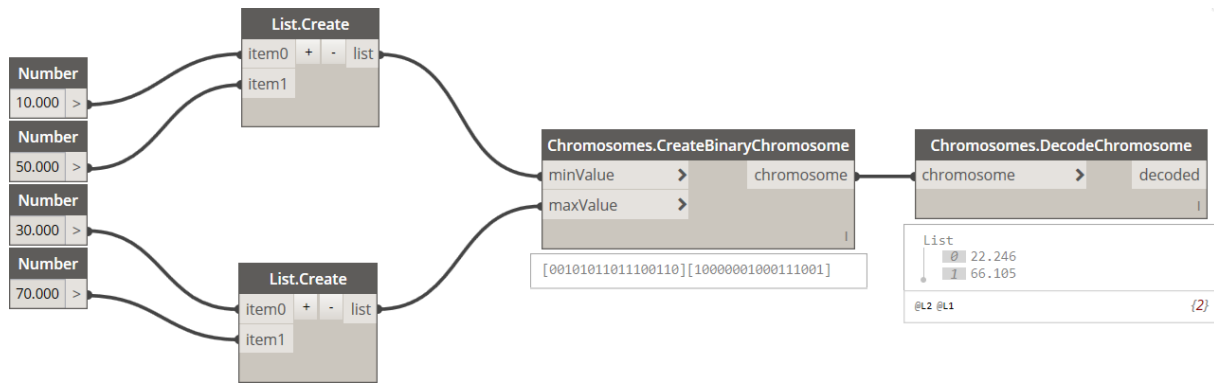


Fig. 2. Creating a binary chromosome for a multi-variable function.

Example of a double (real-variable) chromosome of a two-variable optimization problem:

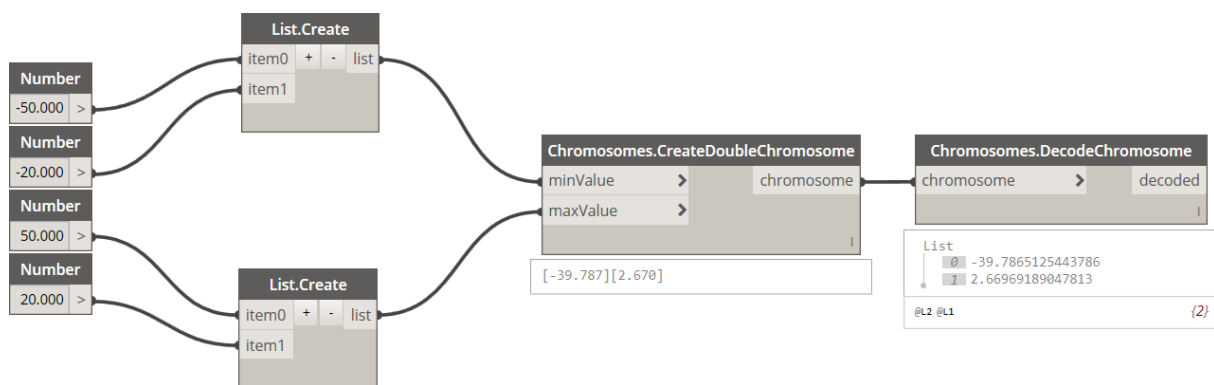


Fig. 3. Creating a double chromosome for a multi-variable function.

If the problem is a combinatorial optimization, e.g. Travelling Salesman Problem (TSP), it is required to obtain a chromosome in a form of an array of unique elements (numbers). In such the case the dedicated **Chromosomes.CreateCombinatorialChromosome** node should be used. The node takes a number of elements as an input. The output has a form a numerical string translated to a list of unique numbers when decoded (see: sample 3.3).

## 2.2. Creating a population of solution candidates

The exemplary solution candidate created in chapter 2.1. can be then used to generate the whole population of other candidates based on the same constraints (minimum and maximum values, required precisions and so on). To do so, one of the `Populations.CreatePopulation` nodes should be used (see: Fig. 4). In general, a population is a set of generations. Each generation is a set of chromosomes.

To create a population, the required inputs are: size of the population (number of solution candidates per generation) and the exemplary chromosome. The created population can be also supplemented with a strategy defined by one of the **PopulationStrategies** nodes. So far there are two population strategies available: `TrackingStrategy` that records all individuals of all generations created during optimum search (in case of complex problems number of generations may be significant); and `PerformanceStrategy` that records only a given number of most recent generations (older generations, including initial generations are overwritten). If no strategy is specified, `PerformanceStrategy` is taken by default with 10 most recent generations recorded.

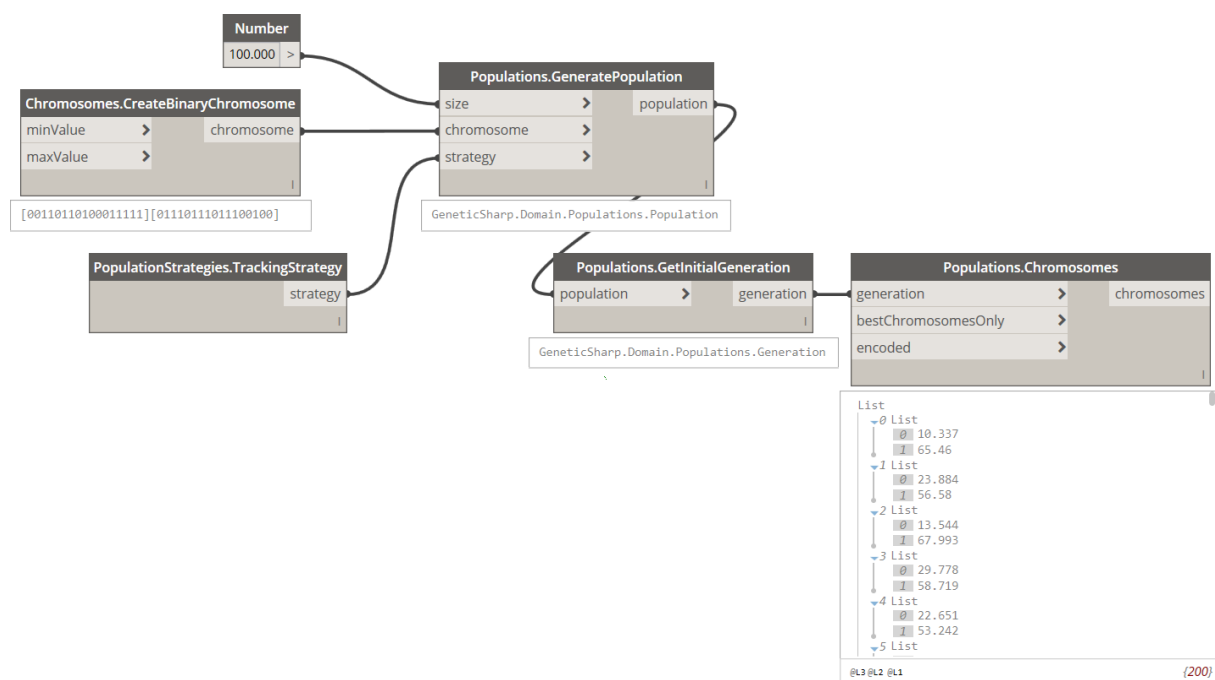


Fig. 4. Creating a new population.

Creating a population generates also an initial generation. If not yet overwritten by the strategy, it can be inspected by the **Population.GetInitialGeneration** node. Chromosomes that belong to the initial generations can be checked using the **Populations.Chromosomes** node. In the Fig. 4, a new population of size 100 was created. It utilizes the chromosome defined in the Fig. 2, hence 200 numbers are encoded in the population (100 solution candidates – 100 sets of two variables per each set).

### 2.3. Creating a genetic algorithm instance

The population from the previous point, so far including the initial generation only, should be passed to the **Basic.CreateGeneticAlgorithm** node. The node is used to define what population should be evolved and which operators should be used.

There are three main GA operators, in turn: selection, crossover and mutation. Selection is used for inspecting the input generation and selecting a portion of most fitting individuals, called parents. Usually, selection methodologies are also designed to ensure that a certain diversity level in the selected set is maintained so the algorithm is less likely to converge to local optima. Both the selection method and size of the selection set are inputs of the **Basic.CreateGeneticAlgorithm** node. The selection size can be lower or equal to the population size defined in the previous point. The selection method is an object created by one of the nodes included in the **Selections** category. So far four selection methodologies are implemented in the solution, described in *Evo Specification. Selections.pdf*.

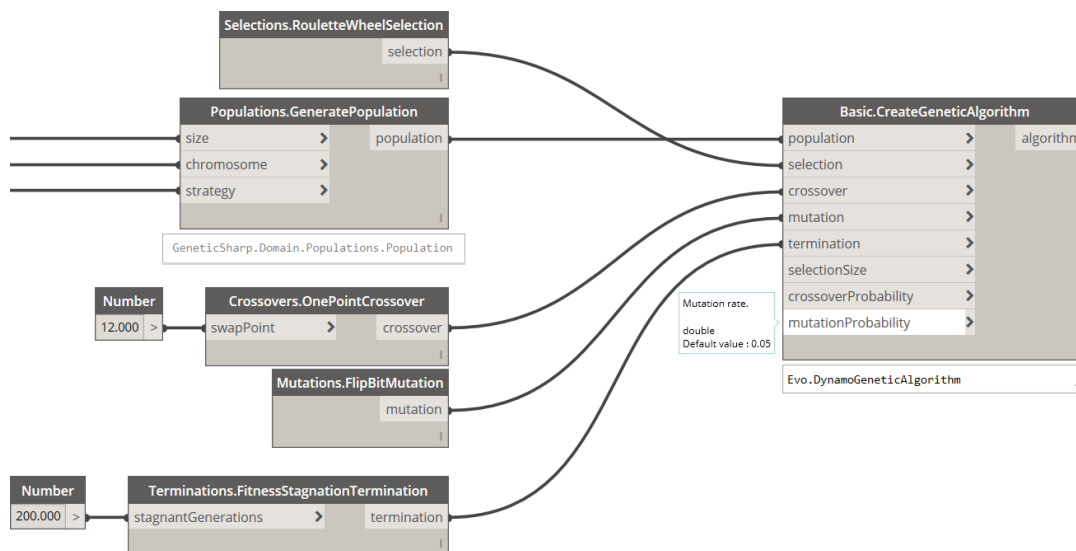


Fig. 5. **Basic.CreateGeneticAlgorithm** node with several inputs connected.

Selected parents are then passed to the crossover. In the crossover, parents are matched with each other, producing new individuals called offsprings. Usually one pair of the parents, randomly selected from the parents' set, produce two new offsprings. It should be however noted that there are several crossover methods that are exceptions to this principle. Structure of the offspring consists of its parents' structures, combined in a specific way. Still, not all parents take part in the crossover, even if passed to the parents' set. Some portion of the parents is directly taken to the offsprings' set with no change in their structures. The crossover probability is one of the GA parameters, set to 0.75 by default. So far thirteen crossover methodologies are implemented, described in *Evo Specification. Crossovers.pdf*. These are objects returned by nodes of the **Crossovers** category.



Finally, the offsprings' set is affected by the mutation operator. In (Konak et al., 2006), it is said that “[...] *crossover leads the population to converge by making the chromosomes in the population alike. Mutation reintroduces genetic diversity back into the population and assists the search escape from local optima.*”. Indeed, newer generations in the population are become less diverse. The point is to make sure that the fitness function values evaluated from such the generation is the global optimum, not a local one. The mutation operator is responsible for slight modification of an offspring, usually at gene level. There are nine mutation operators implemented so far, each one working in a specific way, described in details in *Evo Specification. Mutations.pdf*. Not all offsprings are however mutated – the mutation probability is one of GA parameters and is relatively low compared to the crossover probability, usually between 0.01 and 0.05. The value of 0.05 is used by default. The desired mutation method should be passed to the **Basic.CreateGeneticAlgorithm** node, returned by one of the nodes grouped in the **Mutations** category.

Please note that some of crossover and mutation operators, included in **Crossovers** and **Mutations** Evo libraries are designed especially for binary, combinatorial or real-value (double) problems. I.e. if one is designed for binary chromosomes, it may fail on combinatorial or double chromosomes. The type of chromosomes which the operator was designed for can be checked in description of an operator node or in documentation: *Evo Specification. Crossovers.pdf*, *Evo Specification. Mutations.pdf*.

Offsprings, included mutated ones, are then passed to the new generation. It is ensured size of an input generation and size of the new generation are equal. Free slots, if remaining, are filled with individuals taken from the input generation, without crossover and mutation operators imposed. So far elitist reinsertion method is implemented, meaning that  $n$  best individuals from the input population are passed to the new generation, where  $n$  – number of free slots in the new generation.

Finally, the termination can be defined. So far there are four different termination conditions that define the break point in the evolving loop. These are included in the **Terminations** category and can be combined by two additional operator nodes: *AndTerminations* and *OrTerminations*. E.g. it is possible to break the solution search process if a given number of generations was produced, when a given number of newly produced generations do not produce any better individuals, when evolving time is exceeded or several conditions combined.

## 2.4. Organizing an evolving loop

Dynamo specificity requires organizing a set of nodes that will iterate through the whole evolving process. A single iteration includes selection, crossover and mutation of individuals and these are the role of both **Basic.ProduceOffsprings** and **Basic.ProduceNewGeneration** nodes. Basing on the sequence, a new generation is produced that is, by nature, fitter than the previous one. Because during each iteration a fitness function results for new individuals must be evaluated, the function body must be passed somewhere as a parameter. In textural programming, procedures or functions taking other functions as parameters are called higher-order methods and these are hard to implement due to the specific way in which functions are encoded in Dynamo.

Therefore, a Dynamo built-in **LoopWhile** node will be utilized. The node is not a part of the package but can be found in the Dynamo **BuiltIn** category. If you never used the **LoopWhile** node yet, some basic information can be found [here](#). As you can read, the node takes three inputs: *init*, *continueWhile* and *loopBody*. The last two are the function-type inputs. In Dynamo, it is easy to distinguish between nodes that return functions and nodes that return results of these functions. If a node does not receive all required inputs, its title bar has a light grey color (see Fig. 6). Such the node returns a function object that can be used by some other nodes in Dynamo library, e.g. **List.Map**. If all required inputs are passed to the node, its title bar is brown, hence returning function results. Note that some nodes inputs may stay unconnected yet still the node title bar is dark, what indicates that unhandled inputs take default values, imposed programmatically. It is important to connect light greyed nodes to *continueWhile* and *loopBody* inputs of the **LoopWhile** node. Still, a browned node should be passed to the **LoopWhile** as the *init* input.



Fig. 6. Nodes returning results of functions (left) and function objects (right).

Integral part of the Evo package are three custom nodes that handle all the three **LoopWhile** inputs and make using the **LoopWhile** node easier. These nodes are included in the **Basic** Evo category: **Evo.Init**, **Evo.ContinueWhile** and **Evo.LoopBody** for *init*, *conitnueWhile* and *loopBody* **LoopWhile** node inputs, respectively. The nodes are clusters of other nodes, including standard Dynamo ones. Such the structures are saved as separated DYF files and are indicated by slightly modified outfit among other nodes in the network. To inspect them, simply put one of the nodes somewhere in the work area, click right mouse button on the node and then, from the context menu: *Edit Custom Node*. It may be noticed

that the **Evo.LoopBody** custom node includes both `Evo.ProduceOffsprings` and `Evo.ProduceNew Generation` nodes that is responsible for creating and modifying new generations in the evolving search process, hence you do not need to handle them on your own.

An exemplary arrangement of the `LoopWhile` and `Evo` nodes is presented in the figure below. Make sure there is nothing passed to *init* inputs of `Evo.ContinueWhile` and `Evo.LoopBody` nodes. The *keepEvolving* input of `Evo.init` node may stay unhandled as it takes *true* by default. If you want to disable evolving process at start, just pass *false* value to the *keepEvolving* input and change it when desired.

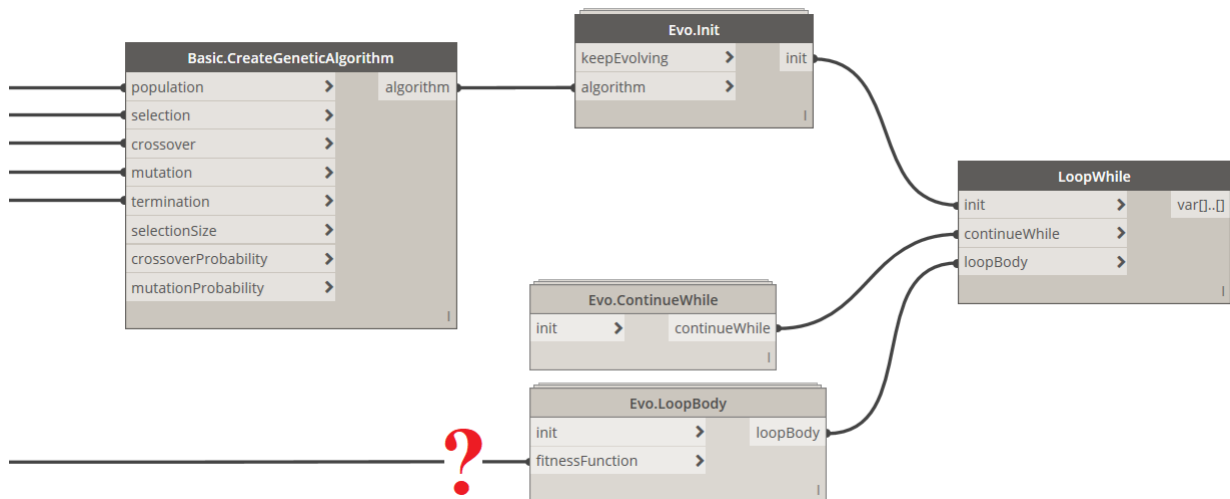


Fig. 7. Exemplary arrangement of `LoopWhile` and `Evo` nodes.

Last but not least is the *fitnessFunction* input of the **Evo.LoopBody** node. This is actually the key part of the optimization problem, that defines the function to optimize. Such the function is called a fitness function. Basic rules regarding composing the fitness function are given in the next chapter.

## 2.5. Composing a fitness function

The fitness function defines a problem to resolve. There are no special limits as for formulas and operations performed inside the functions. These may be mathematical or logical operations, using numbers, strings, geometries and any other custom types or utilizing outer environments and datasets, e.g. text files, excel sheets, Revit or even other platforms, e.g. Robot, if dedicated packages handling them are used. Note that in case of complex problems, utilizing external sources and environments, evolving time may be significant.

There are several rules that should be obeyed when composing a fitness function:

1. All of the operations inside the function should be composed inside a custom node (saved as external DYF file): *File* → *New* → *Custom Node*,
2. Keep in mind the fitness function takes a list of input variables, operates on them and returns a list of results. If a single-variable optimization is the goal, there is no special concerns regarding organizing the initial operations performed by the function – just a list of  $n$  variables is taken and a list of  $n$  results is evaluated.

However, if the optimization is a multi-variable problem, the fitness function should take a list of  $n$  **sets** of variables and return a list of  $n$  results – one numerical result per one set of variables. Each set can contain numerous single variables, hence prior to process the variables, their extraction out of the set is required. It can be noticed in one of the samples files **Evo. Diagonal Problem.dyn** and its custom DYF node called **DistanceBetweenPoints.dyf**<sup>2</sup>:

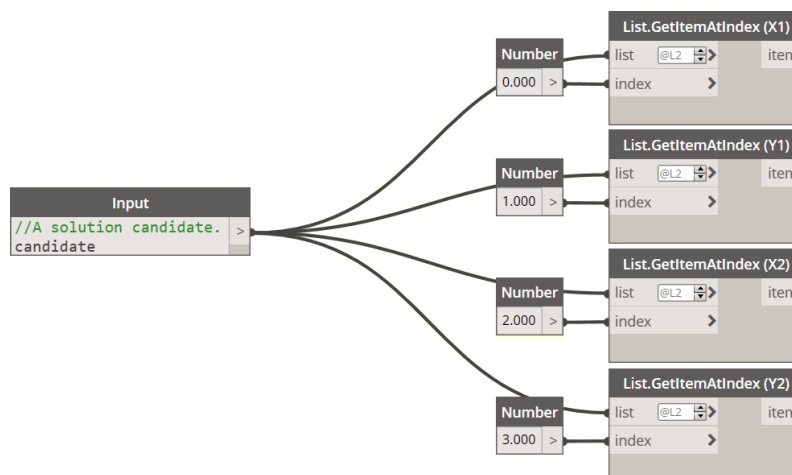


Fig. 8. Extraction of variables from a set of variables passed to the input of a fitness function custom node.

<sup>2</sup> Can be found in: %APPDATA%\Dynamo\Dynamo Core [or Dynamo Revit]\1.3 [or 2.0]\packages\Evo\dyf

In this particular case, an Euclidean distance between two points is calculated. Each point is defined by two numbers: X and Y coordinates for the first point and X and Y coordinates for the second point. Therefore, a four-variable problem is analyzed.

The input *candidate* may be a set of four variables or just a list of such the sets – in fact it does not matter as looping through the list of sets is handled automatically by Dynamo. Our concern is to take care of the four variables inside a single set passed from the *candidate* input node. The input and output nodes are available from standard Dynamo library, under **Input** category. The custom node we create can have several inputs and several outputs however **the whole generation of solution candidates** (in this case the whole list of four-variable sets) **should be passed to one and only one input**. Other inputs, if introduced, can be used to pass constant values to the fitness function that are not the optimization subject but are used to calculate the fitness function (e.g. a list of coordinates of cities in TravellingSalesmanProblem.dyf node).

As you can see in the Fig. 8, four coordinates are extracted out of a *candidate* using built-in *List.GetItemAtIndex* nodes. Note the levelling feature was used for their *list* inputs, meaning that if a list of sets is passed to the fitness function, *List.GetItemAtIndex* does **not** return first, second, third and fourth set, but first, second, third and fourth element from **inside** of the set. To do so, just click on the > icon to the right of the input title and tick the *Use Levels* option:

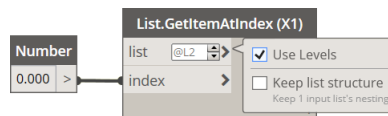


Fig. 9. Nodes levelling.

3. The custom node, for each set of variables should return one numerical value that is the fitness of the set. The final value obtained by the function operations should be passed to the **Output** node which is a built-in Dynamo node. The custom node can also calculate and return auxiliary results however only one output should be passed to the *fitnessFunction* input of the **Evo.** **LoopBody** node. The *fitnessFunction* input accepts only a function object so make sure nothing is connected to your custom node when passing it to the **Evo.** **LoopBody**:

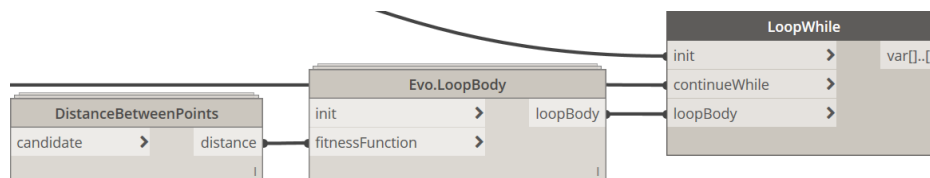


Fig. 10. A custom node defining a fitness function passed to the *Evo.LoopBody* node.

4. The searched global optimum is the **maximized fitness function** – the greater the result of the fitness function, the better candidate solution. Therefore, if your objective consist of function minimization, it is recommended to multiply final function result by  $-1$  right before passing it to the Output in the custom node, or – even better – using a function normalizing the fitness.
5. Errors in fitness function evaluation can lead to incorrect optima found or even termination of the whole evolving process. Therefore, it is recommended to test the fitness custom node before starting the genetic algorithm. Examples of such of the tests are presented in most of sample files, e.g. calculating distance between coordinates:

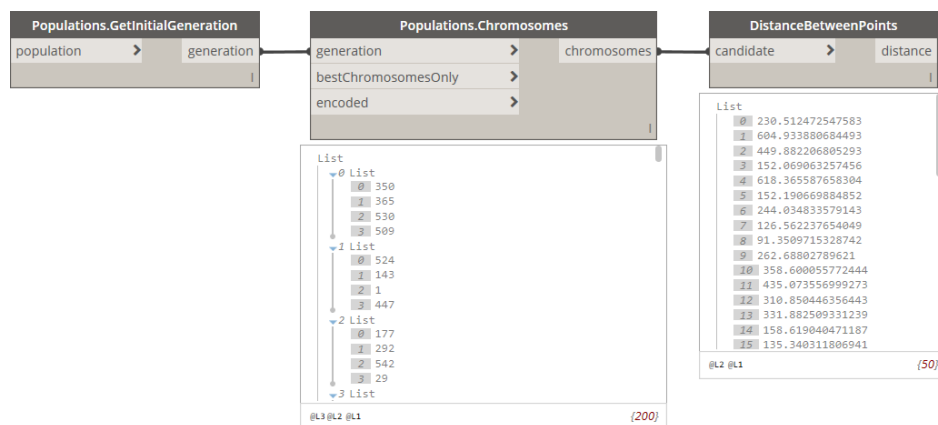


Fig. 11. Inspecting values returned by a fitness function custom node.

## 2.6. Inspecting results

Once the evolving process was finished, its results can be inspected. Note the **LoopWhile** node returns a list of two elements. First element is the *false* value meaning that the processing has finished and no further generations are created.

The second element is the genetic algorithm instance – the same we created in chapter 2.3, but now supplemented with new generations. Note that if **PopulationStrategies.PerformanceStrategy** was used as the population strategy, older populations, including the initial one may be overwritten and, from now on, inaccessible. If you want to store all generations that appear during the whole evolving process, just use **PopulationStrategies.TrackingStrategy**.

First of all the genetic algorithm instance must be taken from the **LoopWhile** output, preferably using **List.LastItem** available in the standard Dynamo library. Then, the population should be retrieved from the genetic algorithm instance, using the **Populations.GetPopulation** node. From the population, inner generations may be extracted. It is possible to retrieve elements from the most recent generation, *n*th generation or all elements of progressive generations (i.e. producing better and better fitness values).

In the example below, the **Populations.GetCurrentGeneration** node was used. The node returns an instance of the generation. To take chromosomes out of it, you can use the **Populations.Chromosomes** node. To extract the best solution (best chromosome), use the **Populations.BestChromosome** node from the **Populations** category:

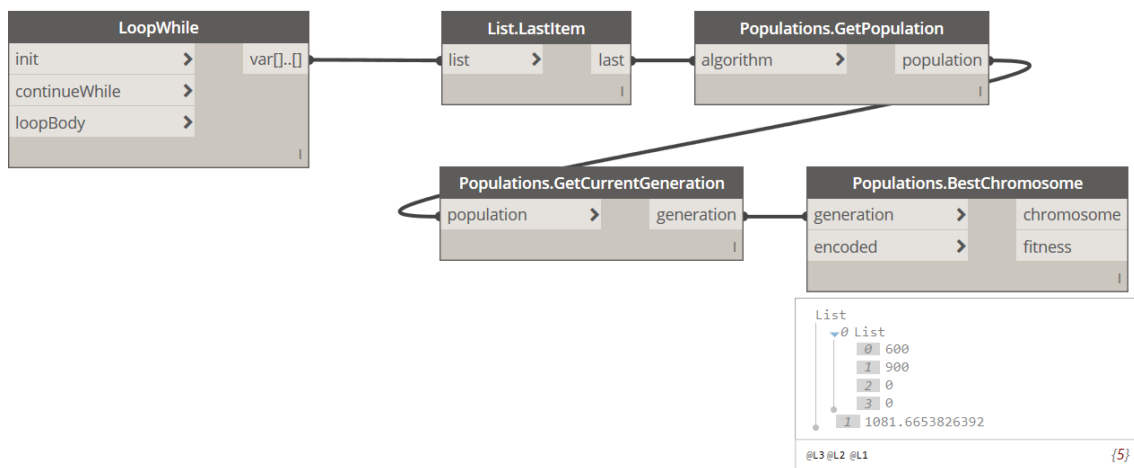


Fig. 12. Inspecting best solution of the optimization problem.

Please note that genetic algorithms are classified as metaheuristic techniques. Therefore, by definition, results obtained by such the techniques do not guarantee to be optimal, but instead sufficient for reaching an immediate goal. The obtained result is as close to the real global optimum as possible, however, due to stochastic methods, can slightly differ from results noted in previous or further runs. In most cases optima obtained from genetic algorithms are sufficient from the engineering point of view, as long as correct parameters and operators are used.

A large number of genetic algorithm evaluations were carried out in the research field. A lot of them include assessment of operators and parameters subjected to the given type of problem. The other ones suggest new selection, crossover or mutation methods to be introduced. The numerous implementations in this package cover these proposals. In further versions of the package, additional operators and features may be expected, including multi-objective optimization techniques.



### 3. Samples

This chapter presents sample Dynamo scripts prepared to demonstrate possibilities of the Evo package as well as the way in which the scripts may be built. All the samples are \*.dyn files and auxiliary custom nodes (\*.dyf files), saved in Dynamo, v. 1.3.3.

Samples are available in the package directory. For **Dynamo Core**:

```
%APPDATA%\Dynamo\Dynamo Core\1.3\packages\Evo\extra
or
%APPDATA%\Dynamo\Dynamo Core\2.0\packages\Evo\extra
```

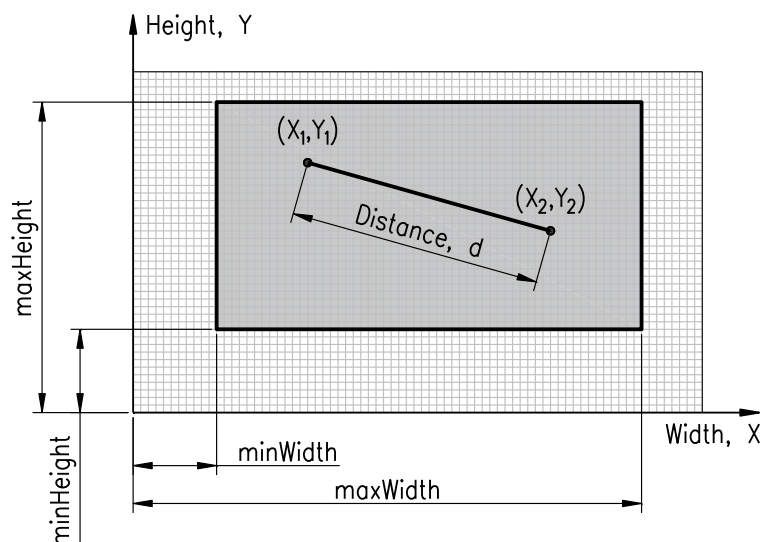
For **Dynamo Revit**:

```
%APPDATA%\Dynamo\Dynamo Revit\1.3\packages\Evo\extra
or
%APPDATA%\Dynamo\Dynamo Revit\2.0\packages\Evo\extra
```

#### 3.1. Evo. DiagonalProblem sample

##### Problem Description

The purpose of this sample is to find the start and end points of the longest line that can fit in a rectangle of a given height and width. Such the trivial problem has a natural solution in the form of a diagonal of the rectangle, however the genetic algorithm itself has no idea about the geometrical background of the problem.



*Fig. 13. Evo. DiagonalProblem. Visualization of the optimization task.*

The idea is to generate random solution candidates. Each candidate consists of four numbers:  $X_1$  and  $Y_1$  for the start point and  $X_2$  and  $Y_2$  for the end point. For each set a fitness function value is evaluated. The fitness function is an Euclidean distance function that should be maximized. An exemplary solution candidate and its boundaries are presented in the Fig. 13.

The fitness function – Euclidean distance:

$$d = \sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2}$$

Constraints:

$$\text{minWidth} \leq X_1 \leq \text{maxWidth}, \text{minHeight} \leq Y_1 \leq \text{maxHeight}$$

$$\text{minWidth} \leq X_2 \leq \text{maxWidth}, \text{minHeight} \leq Y_2 \leq \text{maxHeight}$$

The genetic algorithm will try to find the solution by operating on the sets and function results only, manipulating them iteratively. In fact, it has no even access to the fitness function body. Instead, it utilizes raw results the function returns for provided solution candidates. In the example, minimum width and height are set as 0.0, maximum width is 900.0 and maximum height is 600.0.

The optimization task is a four-variable problem. At the beginning of the script, a binary chromosome is created using the **Chromosomes.CreateBinaryChromosome** node. Its second variant is used, requiring following inputs: list of minimum values per each variable, list of maximum values per each variable and list of precisions (numbers of fraction digits) per each variable. All the three lists consists of four elements – one element per each variable.

Then a population of 50 solution candidates is generated. The population is passed to the **Basic.CreateGeneticAlgorithm** node. Default operators are used: Elite Selection, Uniform Crossover, Flip Bit Mutation, termination after 100 stagnant generations, selection size equal to the half of the population size, crossover and mutation probabilities: 0.75 and 0.05, respectively.

The Euclidean distance fitness function is included in the **DistanceBetweenPoints** custom DYF node. To inspect the node, click right mouse button on the node and, from the context menu, select: *Edit Custom Node*. Its structure is a nodal version of the formula shown above.

The total evolving time is about 3 seconds with 179 generations produced. The best solution is a combination of four values: 0, 0, 600 and 900 of the total length between points:  $d = 1081,7$  units. Note that there are two diagonals in the rectangle so the output numbers can have different order. Evaluation time and number of produced generations may also differ when compared to other runs.

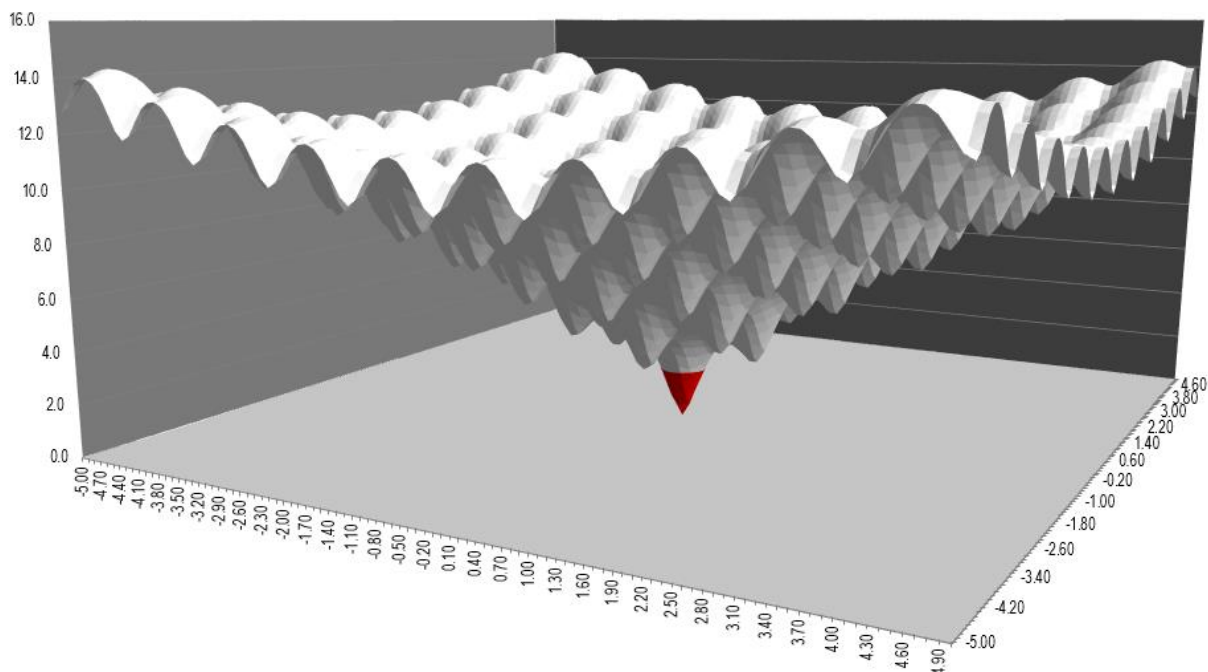
### 3.2. Evo. AckleyFunction sample

#### Problem Description

The Ackley function is a non-convex, two-variable function used as a performance test for optimization algorithms, originally proposed by David Ackley in 1987. The function can be expressed mathematically using the following formula:

$$f(x, y) = -20 \cdot e^{-\frac{1}{5} \sqrt{\frac{1}{2}(x^2 + y^2)}} - e^{\frac{1}{2}(\cos 2\pi \cdot x + \cos 2\pi \cdot y)} + e + 20$$

The function is characterized by numerous local optima and only one global minimum:  $f(0, 0) = 0$ . 3D plot for the Ackley function is presented below:



*Fig. 14. Evo. AckleyFunction. Visualization of the optimization task.*

The optimization task is to find the function global minimum using the visual programming environment. The chromosome pattern of the problem is a bit simpler than the one used in the sample 3.1. as two variables are considered, instead of four. Other, most common test functions are listed [here](#).

For real variables, including negative numbers, it is recommended to use double chromosomes. At the beginning of the script, the **Chromosomes.CreateDoubleChromosome** node is used. The most fitting variables –  $x$  and  $y$  – are looked in a numerical range between -100.0 and 100.0. Then a population of 100 solution candidates is generated. The population is passed to the **Basic.CreateGeneticAlgorithm** node. Following operators are used: Tournament Selection, Voting Recombination Crossover, Flip Bit

Mutation (by default), termination after 100 stagnant generations (by default), selection size equal to population size, crossover and mutation probabilities: 0.75 and 0.01, respectively.

The Ackley function is included in the **AckleyFunction** custom DYF node. To inspect the node, click right mouse button on the node and, from the context menu, select: *Edit Custom Node*. Its structure is a nodal version of the modified formula shown above.

The total evolving time is about 20 seconds, producing 430 generations. The best solution varies around the point  $x = 0.0$  and  $y = 0.0$ :

Populations.BestChromosome			
generation	>	chromosome	
encoded	>	fitness	
List			
0 List			
0		0.00152122311140126	
1		-3.35232036864946e-05	
1		-0.00614755678488166	
@L3 @L2 @L1			
{3}			

In the sample, **Populations.GetProgressiveGenerations** is also used, listing all the generations that have been improving the found optimum consecutively. One can notice that the latter generations are characterized by slightly improved fitness. To reduce the computation time, other kind of termination can be considered.

Populations.GetProgressiveGenerations		Populations.BestChromosome		Object.Identity	
population	>	generation	>	obj	>
		encoded	>		var[...]
1 -3.35268630658671e-05		1 -3.35268630658671e-05		52 -0.00614755678532219	
1 -0.00614755678489942		1 -0.00614755678489942		53 -0.00614755678529022	
66 List		66 List		54 -0.00614755678523338	
0 List		0 List		55 -0.00614755678521206	
0 0.0015212230339737		0 0.0015212230339737		56 -0.00614755678499535	
1 -3.35267171917036e-05		1 -3.35267171917036e-05		57 -0.00614755678499179	
1 -0.00614755678489587		1 -0.00614755678489587		58 -0.00614755678498469	
67 List		67 List		59 -0.00614755678497758	
0 List		0 List		60 -0.00614755678497048	
0 0.00152122305738415		0 0.00152122305738415		61 -0.00614755678496337	
1 -3.35256548708813e-05		1 -3.35256548708813e-05		62 -0.00614755678495982	
1 -0.00614755678488876		1 -0.00614755678488876		63 -0.00614755678491719	
68 List		68 List		64 -0.00614755678490653	
0 List		0 List		65 -0.00614755678489942	
0 0.00152122310922519		0 0.00152122310922519		66 -0.00614755678489587	
1 -3.35233024321443e-05		1 -3.35233024321443e-05		67 -0.00614755678488876	
1 -0.00614755678488166		1 -0.00614755678488166		68 -0.00614755678488166	
@L4 @L3 @L2 @L1		{207}		@L2 @L1	
				{69}	

### 3.3. Evo. TravellingSalesmanProblem sample

#### Problem Description

The Travelling Salesman Problem (TSP) is a classical NP-hard combinatorial problem. Its input is a list of cities, in our case expressed in the form of points list (Fig. 15). The optimization task is to find the closest path between the points with no repetitions. The problem is visualized by a salesman who travels between the cities but during the route he cannot enter the same city twice and his path should form a closed curve.

The presented sample utilizes operations on graphical Dynamo objects. The first step in the solution search is formulating an adequate type of a chromosome pattern. In case of such the problems, the dedicated chromosome can be created using **Chromosome.CreateCombinatorialChromosome**. The only input of the node is number of unique indices the chromosome should store.

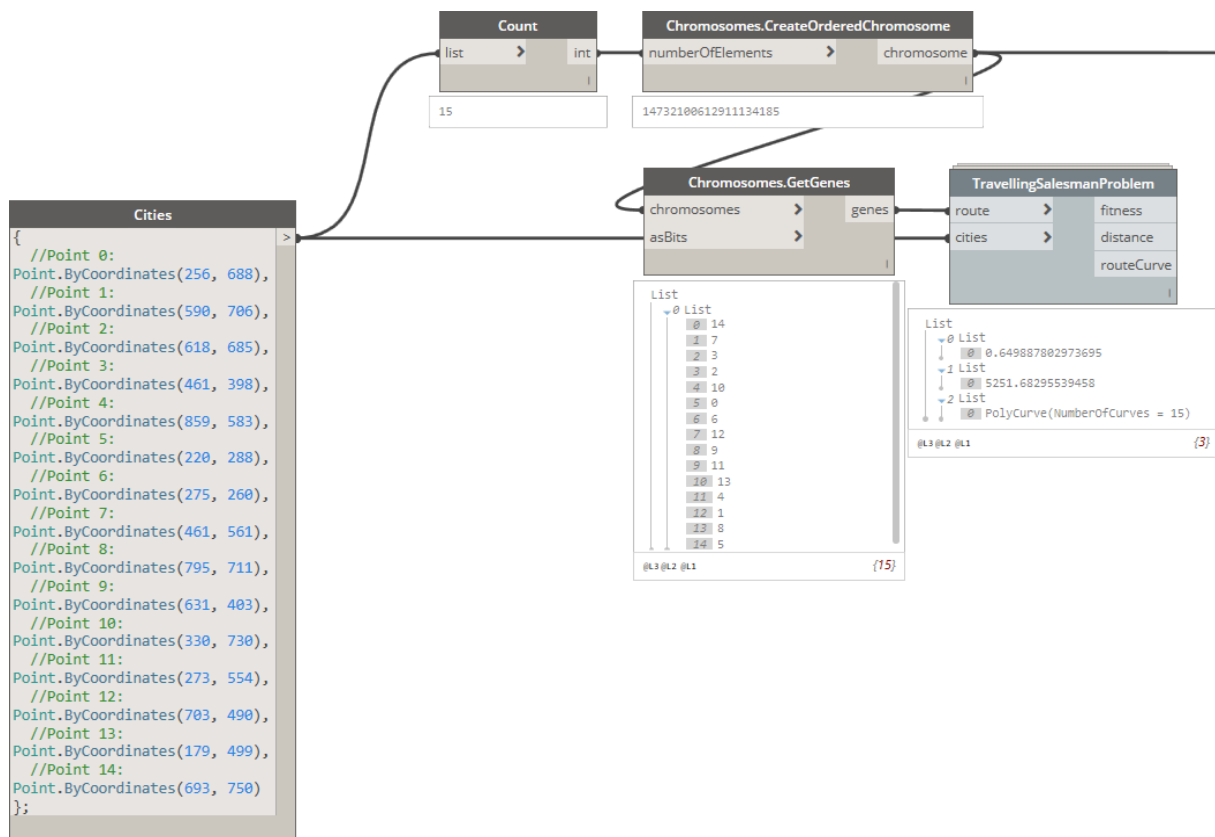
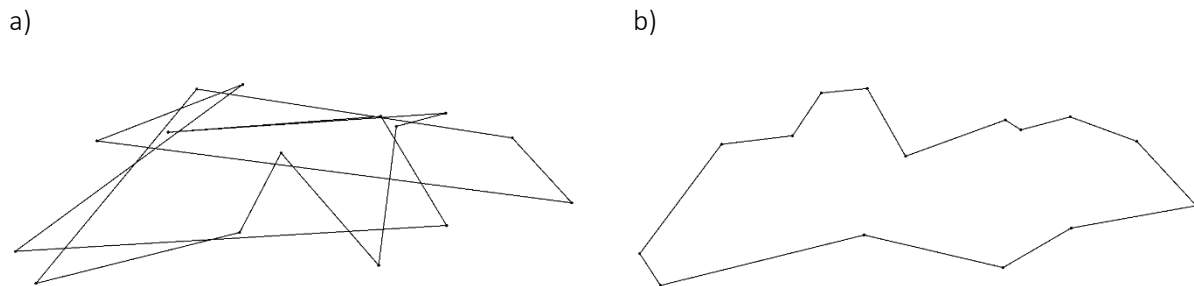


Fig. 15. Definition of an ordered chromosome, genes extracted out of the chromosome and the route calculated.

The **TravellingSalesmanProblem** is a custom DYF node responsible for calculating normalized fitness function value, real travel distance between the cities in turn and visualizing the route in the form of a polycurve. The travel distance is calculated by a **Curve.Length** node that gets the polycurve as an input, created by a **PolyCurve.ByPoints** node. All the operation can be checked within the DYF node.

Then a population of 100 random routes is generated. The population is passed to the **Basic.Create GeneticAlgorithm** node. Following operators are used: Roulette Wheel Selection, Ordered Crossover, Reverse Sequence Mutation, termination after 100 stagnant generations (by default), selection size equal to population size and default probabilities for both crossover and mutation: 0.75 and 0.05, respectively. Please keep in mind that some of the crossover methods, as well as some mutation methods, are designed especially for combinatorial types of optimization problems. The other ones operate on floating point chromosomes encoded to the binary strings (non-combinatorial problems).

The total evolving time is 15 to 30 seconds. Iterations were terminated after generation no. 201. The figure below shows the initial proposed route obtained by a random sequence between the cities. The total travel distance for the proposal is 5118.2 units. After search process the best route of total distance 2096.4 units was obtained, presented below to the right:



*Fig. 16. One of the initial route proposal (left) and the best solution obtained by the algorithm (right).*

## Scientific publications

Deb, K.: *An Efficient Constraint Handling Method for Genetic Algorithms*. Computer Methods in Applied Mechanics and Engineering, vol. 186, issue 2, 4, June 2000, pp. 311 – 338.

Deb, K., Deb, D.: *Analyzing Mutation Schemes for Real-Parameter Genetic Algorithms*. International Journal of Artificial Intelligence and Soft Computing, vol. 4, issue 1, February 2014, pp. 1 – 28.

Hussain, A., Muhammad, Y. S., Sajid, M. N., Hussian, I., Shoukry, A. M., Gani, S.: *Genetic Algorithm for Traveling Salesman Problem with Modified Cycle Crossover Operator*. Computation Intelligence and Neuroscience, vol. 2017, August 2017, pp. 1 – 7.

Konak, A., Coit, D. W., Smith, A. E.: *Multi-Objective Optimization Using Genetic Algorithms: A Tutorial*. Reliability Engineering & System Safety, vol. 91, issue 9, September 2006, pp. 992 – 1007.

Larrañaga, P., Kuijpers, C. M. H., Murga, R. H., Inza, I., Dizdarevic, S.: *Genetic Algorithms for the Travelling Salesman Problem: A Review of Representations and Operators*. Artificial Intelligence Review, vol. 13, issue 2, April 1999, pp. 129 – 170.

Umbarkar, A. J., Sheth, P. D.: *Crossover Operators in Genetic Algorithms: A Review*. ICTACT Journal on Soft Computing, vol. 6, issue 1, October 2015, pp. 1083 – 1092.