Warping Search: A New Metaheuristic Applied to the Protein Structure Prediction

Richard A. Gonçalves
CPGEI UTFPR
Paraná, Brazil
richard@unicentro.br

Marco C. Goldbarg
DIMAP UFRN
Rio Grande do Norte, Brazil
gold@dimap.ufrn.br

Elizabeth F. Goldbarg
DIMAP UFRN
Rio Grande do Norte, Brazil
beth@dimap.ufrn.br

Myriam R. Delgado
CPGEI UTFPR
Paraná, Brazil
myriamdelg@utfpr.edu.br

ABSTRACT

This paper presents a new technique to search the solution of difficult optimization problems. The Warping Search is inspired on the movements of celestial bodies in the universe and their interactions. Based on some concepts of Einstein's General Relativity, the search operators proposed in this approach are inspired on gravitational wave sources such as mass transfer, spin, collapse, explosion and collision and defined based on gravitational fields. The technique has been tested in the Protein Structure Prediction Problem in lattice models that is known to be NP-hard. The presented approach is compared with recent heuristics proposed for the problem and the computational results attest the efficiency of the method.

Categories and Subject Descriptors

I.2.8 [**Computing Methodologies**]: Artificial Intelligence—*Problem Solving, Control Methods, and Search*; J.3 [**Computer Applications**]: Life and Medical Sciences

General Terms

Algorithms

Keywords

Metaheuristics, Protein Structure Prediction, Warping Search

1. INTRODUCTION

In 1916, Albert Einstein formulated a set of general laws known as General Relativity [1] where he described gravity as the warping of space-time, i.e. a property of the universe's fabric, not a force acting at a distance as Newton has previously proposed. This paper presents an optimization method inspired on the movements and interactions between objects of the universe and on the warping of space-time called Warping Search (WS). In WS, the solutions of a combinatorial problem are thought as celestial bodies in movement. The bodies' movements are guided by computational adaptations chosen based on the principia of the General Relativity. The proposed heuristic is used to solve the Protein Structure Prediction Problem (PSPP) in the 3D HP model.

GECCO'08, July 12–16, 2008, Atlanta, Georgia, USA.
ACM 978-1-60558-130-9/08/07.

2. WARPING SEARCH

In the Warping Search candidate solutions (or part of them), called celestial bodies, evolve by means of operations inspired on gravitational waves sources (such as spin, mass transfer, collapse, explosion and collision). Gravitational waves are waves that warp the space-time. The evolution of celestial bodies also happens due to their movements caused by the distortion in space-time. A celestial body has the following characteristics: mass, velocity, a position in space-time and a gravitational field. The mass corresponds to the adequacy (fitness) of the solution represented by the body. The mass, in relativity, is distorted (modified) by the velocity of the body. The velocity of a body gives an estimation of the history of the evolution or the evolution rate of the celestial body. If it has a high velocity, it means that it is evolving fast (improving its solution's quality with few movements). The position of the celestial body in the space-time (search space) encodes the solution represented by it. It's important to note that not all the celestial bodies encode complete solutions, some bodies represent partial solutions. Complete solutions are considered denser than the partial ones, i.e. they have more information encoded on them. The gravitational field of a celestial body determines how attractive a celestial body will be and how intense will be the gravitational wave sources applied to it. A celestial body can represent a star, a planet, a satellite, an asteroid, etc. In this paper only stars are considered. There is a special kind of celestial object: a black hole. A black hole, physically speaking, is a very dense celestial object that greatly warps the space-time. In WS a black hole represents an optimum. So the objective of the Warping Search is to find out black holes in the space-time. The general form of the Warping Search algorithm is as follows. The algorithm starts with a huge explosion, the Big Bang. In Physics, the Big Bang is a theory that explains the origin of the universe. It states that about 15 billion years ago a tremendous explosion started the expansion of the universe. In WS this stage corresponds to the creation of various galaxies and the initialization of the celestial bodies (complete and partial solutions). The initialization of a celestial body corresponds to the calculation of the initial position, initial velocity and initial mass (inertial and distorted, various formulations for the distortion can be applied).

Different galaxies may be created in the universe - so different sets of celestial bodies may exist and evolve in different rates and forms, interacting during the whole process. Galaxies permit specialization of the celestial bodies (a kind of niching) and parallel implementations (similar to multiple populations). Afterwards, the warping of the space-time and the gravitational field of each celestial body are calculated. In Physics, there are several possible for-

Table 1: **Results obtained by the proposed algorithm to the 3D HP model compared to methods of the literature.**

	Warping Search				IA			F-EA		
Seq.	Best ($-\eta$)	Mean	SD	Worst	Best ($-\eta$)	Mean	SD	Best ($-\eta$)	Mean	SD
1	-11	-11	0	**-11**	-11	-11	0	-11	-10.32	0.61
2	-13	-13	0	**-13**	-13	-13	0	-13	-10.90	0.98
3	-9	-9	0	**-9**	-9	-9	0	-9	-7.98	0.71
4	-18	**-18**	0	**-18**	-18	-16.76	1.02	-18	-14.38	1.26
5	**-31**	**-29.43**	0.97	-27	-29	-25.16	0.45	-25	-20.80	1.61
6	**-31**	**-28.87**	1.35	**-24**	-23	-22.60	0.40	-23	-20.20	1.50
7	**-52**	**-49.80**	1.29	**-47**	-41	-39.28	0.24	-39	-34.18	2.31
8	**-54**	**-51.23**	2.63	**-43**	-42	-39.08	0.95	-39	-33.01	2.49

mulas for calculating the warp. Although any formula can be used, all the formulas share the characteristic that the warping due to a celestial body is located in its vicinity and is proportional to the distorted mass of the body. The gravitational field of a celestial body is the average value of the gravitational force between this celestial body and all other bodies of the universe. The next step in the algorithm is the movement of the celestial bodies. The movement, due to Einstein's theory, must be guided by the warping in space-time. The movement of the bodies can be implemented by a path relinking procedure, a differential operator or any other technique that modifies one solution toward another. Thereafter, the gravitational wave sources are applied. Gravitational wave sources include spin, mass transfer, collapse, explosion and collision. The intensity of a wave source is proportional to the gravitational field of the celestial body to which it is applied. In the computational context spin can be thought as any operation on a single solution (a dense celestial body), such as a local search, perturbations, solution of subparts of the problem by an exact method, etc. Mass transfer is the process of transferring mass from one body to another denser body. From the computational point of view this can be thought as a process where, given two celestial bodies the denser receives "material" (parts) of the other. The mass transfer can occur between complete or incomplete solutions, where the less dense body donates mass to the more dense one. The natural collapse happens when the inward gravitational pull of an object is no longer resisted by outward pressure. In the natural context, a star exhausts its nuclear fuel and then it collapses under its own mass. In the computational context the collapse of a solution can be thought as its destruction after it has exhausted its internal energy (it has not improved for a certain number of iterations, for example). The destruction of the body is partial and, once a solution is "destructed" it is immediately rebuilt. It can be randomly rebuilt or a constructive heuristic can be used. The explosion of a celestial body may cause its pieces (parts of the solution) to reach other bodies. In the algorithmic context, the expelled pieces can be absorbed by other bodies. At this point, the termination criterion is tested. If it is satisfied the algorithm ends, otherwise the parameters are updated and the search proceeds. Parameters updating corresponds to a self-adaptation stage of WS. Examples of parameters that can be adapted include the number of celestial bodies, galaxies, the probability of applying each gravitational wave source, the type of movement applied to the solutions, the neighborhood used in a local search, and so on. With these ideas in mind we can think in the search process of WS given by the evolution of one or more galaxies. In this evolutionary process, the explosion (Big Bang) that follows the universe creation is so strong that the universe expands exploring different regions of the search space. Thereafter, the harmonization takes place with matter being organized into different galaxies formed by celestial bodies. The bodies' movement and the gravitational wave sources are mechanisms that balance diversification and intensification in the search process.

3. PROTEIN STRUCTURE PREDICTION

Deciding the structure of a protein is fundamental to find out its biological function. Due to the inherited complexity of Protein Structure Prediction Problem (PSPP), simplified models have become very popular, specially the HP model. Conformations of an HP sequence are restricted to self avoiding walks on a lattice. For the 3D HP model a three-dimensional cubic lattice is adopted. The PSPP in this model is known to be NP-hard.

4. EXPERIMENTS AND RESULTS

Some experiments were carried out to test the efficiency of the methodology while solving the PSPP in the 3D HP model. The benchmark used in this experiment is composed of eight protein instances with different length, taken from [3]. For a fair comparison the termination criterion was set as 10^5 function evaluation, the same criterion adopted in [3]. The results obtained with the Warping Search and the results presented in [2] (F-EA) and in [3] (IA) are compared in Table 2, demonstrating that the proposed approach is capable of finding best energy values to the last four instances and better mean energy values for all but the first three instances.

5. CONCLUSION

In this paper we proposed a new evolutionary metaheuristic based on cosmological principles for combinatorial problems called Warping Search. The algorithm mimics various aspects from nature, such as the warping of the space-time, the movement of stars, spin, collapse and explosion of celestial bodies and mass transfer between two celestial bodies. Comparing the results obtained by the WS implementation for the PSPP in the 3D HP model with those reported in the literature it is possible to observe that WS exhibits an overall better performance, finding energy minima that other algorithms are not capable of.

6. REFERENCES

[1] A. Einstein, "Relativity: The Special and General Theory". New York: Crown, 1961.

[2] C. Cotta, "Protein Structure Prediction Using Evolutionary Algorithms Hybridized with Backtracking," Lecture Notes in Computer Science, 2687, pp. 321–328, 2003.

[3] V. Cutello, G. Nicosia, M. Pavone, J. Timmis, "An Immune Algorithm for Protein Structure Prediction on Lattice Models", IEEE Trans. on Evol. Comp., 11(1):101-117, 2007.