

On Estimating Similarity of Artificial and Real Organisms

Maciej Komosinski¹, Grzegorz Koczyk² and Marek Kubiak¹

¹Institute of Computing Science, Poznan University of Technology, Poznan, Poland

²Faculty of Biology, Adam Mickiewicz University, Poznan, Poland

Address for correspondence: Maciej Komosiński, Institute of Computing Science, Poznan University of Technology, Piotrowo 3A, 60-965 Poznan, Poland, Fax: +48 (61) 8771525; E-mail: Maciej.Komosinski@cs.put.poznan.pl

Summary: In this paper, a dissimilarity measure for artificial organisms is proposed. The organisms are simulated in the Framsticks system [10]. Properties of agents are enumerated formally, and the heuristic algorithm for estimating overall phenetic dissimilarity of two agents is described. An example of performance is shown on two selected organisms. Two clustering experiments with interesting results are presented using the UPGMA method. The properties of the measure are then discussed. Computer simulations of complex systems and their characteristics are compared to biological systems, which may bring up ideas for further experiments related to biology.

Key words: similarity measure, simulation, artificial life, autonomous agents, dendrograms

1. Introduction

Simulating evolving systems, apart from practical applications, is interesting because it allows for studies and analyses of evolutionary processes. The investigated system can be designed specifically to follow the desired properties, and then it may be attractive from a biological point of view. However, as the complexity of a simulated system increases, it is difficult to understand the behavior of existing agents in detail [3,5,11]. The only way is to observe them carefully and use human intelligence to draw conclusions. Usually, the behavior of such agents is non-deterministic, and their control systems are sophisticated, often coupled with morphology and very strongly connected functionally [12].

Thus for the purposes of studying behaviors and populations of individuals, one needs high-level, intelligent support tools [9]. In the real world, there are human biologists who investigate living organisms. In artificial worlds, such work is very tedious and time-consuming, and it is not likely that automatic tools will soon be able to produce understandable, non-trivial explanations of sophisticated artificial agents.

One of the simplest support tools is a measure of similarity. It can be addressed in many ways, including aspects of morphology (body), brain, size, function, behavior, performance, fitness, etc. Even though similarity measure is a simple tool, its reliable definition for complex agents may be quite sophisticated. Having such a measure, it is possible to

- introduce artificial niches by modification of fitness values in optimization [6,13],
- study evolutionary processes and the structure of populations of individuals,
- study function/behavior of agents,

- reduce the number of agents to a small subset of interesting, diverse, unique individuals,
- infer dendrograms (and hopefully, phylogenetic trees) based on distances between organisms.

Thus construction and exploitation of similarity measure is interesting both for computer science and biology. For our experiments we used *Framsticks* system, which models realistic organisms [10,9,1]. Although the model is simpler than biological creatures, it is general enough to observe various properties and difficulties.

The paper is organized as follows: section 2 describes the model of simulation, as well as the formal definition of an agent. In section 3 we focus on the dissimilarity measure we proposed, and present the heuristic algorithm. Section 4 contains two uses of the measure: an illustrative example on two individuals, and two clustering experiments. Section 5 speculates about various approaches to measuring similarity, relation of biological and artificial-world methods, discusses properties of our measure, and compares characteristics of real life and artificial life systems in context of biological research. Section 6 summarizes the work and outlines future goals.

2. The Framsticks model

Artificial creatures in *Framsticks* are built of body and brain. Body is composed of material points (called *parts* in this paper) connected by elastic *joints*. Brain is made from *neurons* and their elements called *neuroitems* (these are receptors, effectors, and neural connections). For more detailed description of the model mechanics and neural network refer to [7,10,1].

It is possible to use the system for simulation of various processes, including local optimization, evolutionary optimization, coevolution of populations, spontaneous evolution, multi-criteria optimization, etc. Such universality is obtained by the use of a scripting language, which controls the overall architecture of the system as well as more specific issues. Results used in this paper come from the directed model of evolution (called evolutionary optimization in computer science [6,13]). In this simplified setup, a constant number of genotypes was subject to fitness-proportional selection, mutation and recombination, and evaluation. Fitness criteria were defined explicitly (average height or speed of organisms during their life span).

Framsticks supports a few genetic representations of organisms (i.e. functions mapping from genotype to phenotype), including direct and developmental [9]. It is possible to easily “plug in” a new encoding and genetic operators. The fact of diverse genotype “languages” may influence the characteristics of the organisms described by them, but each encoding has to produce an organism. Thus it is possible to compare phenotypes which are “compatible” because they are build of a set of standardized components (parts, joints, neurons, and neuroitems).

Elements of a framstick are characterized by many properties. Morphological elements have mass, friction, stiffness, etc., and neurons have weights, sigmoid coefficients, etc. Body and brain are connected. Thus the model is quite complex. Its most important elements are described formally below.

Let P denote a non-empty set of parts (material points) of a creature’s body. We define several functions which describe physical and biological attributes of parts (\mathbf{R} denotes the set of real numbers):

- *position*: $P \rightarrow \mathbf{R}^3$
- *orientation*: $P \rightarrow [0, 2\pi)^3$
- *mass, volume, friction, ingestion, assimilation*: $P \rightarrow \mathbf{R}_{>0}$

Furthermore, we introduce J as a non-empty set of joints, and a function:

- *joint_connects*: $J \rightarrow P^2$

which represents connections (arcs) between pairs of parts (vertices). The graph corresponding to morphology must be connected, which means that the creature’s body cannot have isolated parts. No arcs joining the same pair of vertices may exist. The properties of joints include:

- *stiffness, rot_stiffness, stamina*: $J \rightarrow \mathbf{R}_{>0}$

For brain, we define a set of neurons N , which may be empty. Properties of a neuron can be defined as functions:

- *force, inertia*: $N \rightarrow [0, 1]$
- *sigmoid, initial_state*: $N \rightarrow \mathbf{R}$
- *placement*: $N \rightarrow P$

The set C (which may be empty) concerns neuroitems: elements related to connections between neurons and interfacing parts of the body (including receptors of equilibrium, touch, smell, and muscles). The structure of this set and related functions are described in detail in [8].

Thus a creature can be described by four sets: P , J , N , and C , and relations between their elements. It can be clearly seen that the structure of an agent is quite complicated – it has different aspects: structural (one graph defining body, another one defining neural network which is located on the body's graph), geometrical (position and orientation in a 3-D space), and also several parameters describing elements of body and brain.

3. Similarity measure: the heuristic algorithm

The problem of similarity estimation in *Framsticks* is closely related to the problem of isomorphism of graphs. Taking only morphology into consideration, the task would be to find the matching between parts (P sets) of the two agents. However, such approach is not feasible. An exact algorithm working on such sophisticated graph representations would have an unacceptably high computational complexity (the problem of finding the maximal common sub-graph would be NP-hard).

Following these considerations and another aspects discussed in section 5, we constructed a heuristic method. This method tries to match the body structure of two individuals based on the degrees of parts as the main piece of information. Sets of parts are sorted by *degree*, then by *nitem_count* and *neuron_count* (defined below). Then the algorithm is matching parts starting from the highest degree groups and trying to find parts that have similar *nitem_count* and *neuron_count*. Starting from the highest degree is likely to preserve the most important parts from being unmatched. In case of ambiguity (when points have the same *degree*, *nitem_count* and *neuron_count*), the remaining properties of parts are used as discriminating features (except of geometric ones). When two organisms have different number of parts, each missing part produces dissimilarity ingredient as big as would be produced by two maximally dissimilar parts.

Let $G_1=(P_1, J_1, joint_connects_1)$ and $G_2=(P_2, J_2, joint_connects_2)$ represent the graphs of bodies of two compared agents. Let us assume without loss of generality, that $|P_2| \geq |P_1|$.

The goal of the matching algorithm is to find a mutually single-valued function *matching*: $P_1 \rightarrow M$, where $M \subseteq P_2$ (not all parts can be matched when organisms differ in size). This function is feasible for computing dissimilarity.

For all vertices in G_1 and G_2 , we define:

- *degree*(p) – degree of a vertex $p \in P$ in graph,
- *neuron_count*(p) = $|\{n \in N: placement(n)=p\}|$,
- *nitem_count*(p) = *neuron_count*(p) + $|\{c \in C: placement(parent_neuron(c))=p\}|$.

The heuristic algorithm that finds the *matching* function is outlined below:

SORT(P_1 , descending, by *degree*)

SORT(P_2 , descending, by *degree*)

SORT(groups in P_1 with the same *degree*, descending, by *nitem_count*)

SORT(groups in P_2 with the same *degree*, descending, by *nitem_count*)

REPEAT

FIND U, V – groups of parts in P_1, P_2 with the same, highest possible, degree and with parts not matched yet. Assume below that $|U| \leq |V|$

LET $d=|degree(u \in U)-degree(v \in V)|$

FIND the first part $u \in U$ which is yet unmatched

```

FIND  $V_1 \subseteq V$  so that  $ni = |nitem\_count(v \in V) - nitem\_count(v_1 \in V_1)|$  is minimal
IF ( $|V_1|=1$ ) THEN match-parts( $v_1 \in V_1, u$ )
ELSE
BEGIN
  FIND  $V_2 \subseteq V_1$  so that  $ne = |neuro\_count(v_1 \in V_1) - neuro\_count(v_2 \in V_2)|$  is minimal
  IF ( $|V_2|=1$ ) THEN match-parts( $v_2 \in V_2, u$ )
  ELSE
  BEGIN
    IF ( $d=0$  and  $ni=0$  and  $ne=0$ ) THEN
    BEGIN
      FIND  $v_3 \in V_2$  that has all properties (except geometrical) same as  $u$ 
      IF (exists  $v_3$ ) THEN match-parts( $v_3, u$ )
      ELSE match-parts( $v_3, u$ ), where  $v_3$  is the last in  $V_2$ 
    END
    ELSE match-parts( $v_3, u$ ), where  $v_3$  is the last in  $V_2$ 
  END
END
END
UNTIL all parts in  $P_1$  are matched.

```

Having found the matching, two components of dissimilarity can be computed:

$$dDeg(matching) = \sum_{p_1 \in P_1, p_2 \in M: p_2 = matching(p_1)} |degree(p_1) - degree(p_2)| + \sum_{p_2 \notin M} |degree(p_2)|$$

The second criterion of dissimilarity considers the numbers of neurons on already matched parts:

$$dNeu(matching) = \sum_{p_1 \in P_1, p_2 \in M: p_2 = matching(p_1)} |neuron_count(p_1) - neuron_count(p_2)| + \sum_{p_2 \notin M} |neuron_count(p_2)|$$

Finally, the overall dissimilarity between two organisms is evaluated using the weighted sum model:

$$dissimilarity(O_1, O_2) = w_{Deg} \cdot dDeg + w_{Neu} \cdot dNeu$$

where w_{Deg} and w_{Neu} are weights of these two criteria. The discussion of properties of this function is provided in section 5.

4. Computational experiments

Illustrative example of similarity estimation

To illustrate the algorithm, a simple example is presented below. Consider two individuals A and B (nicknamed ‘‘Scorpion’’ and ‘‘Rolling blender’’ shown on fig. 1.

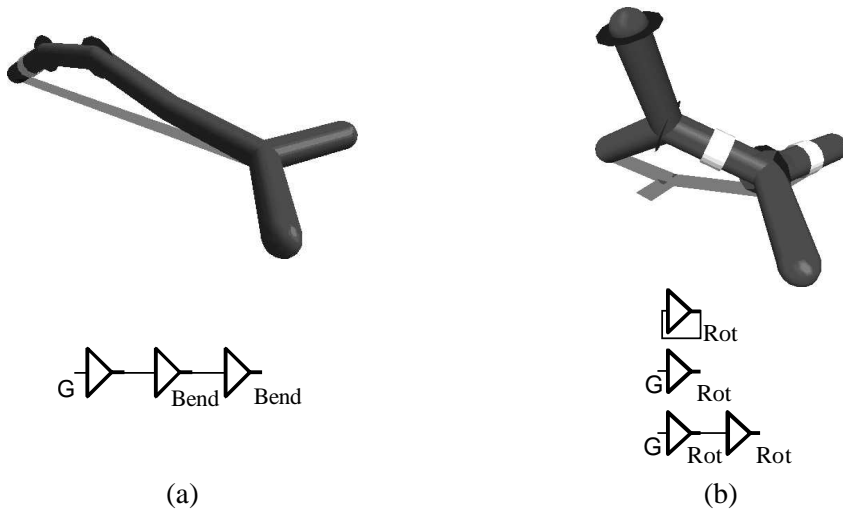


Fig. 1. Two individuals. Morphology and neural networks are shown. White rings in morphology are equilibrium receptors (named “G” in neural network diagrams). “Rot” and “Bend” mean that a neuron output controls a muscle (rotating or bending). The muscles can also be seen in the morphology view (dark rings).

The structure of these creatures (both body and brain) is shown on fig. 2.

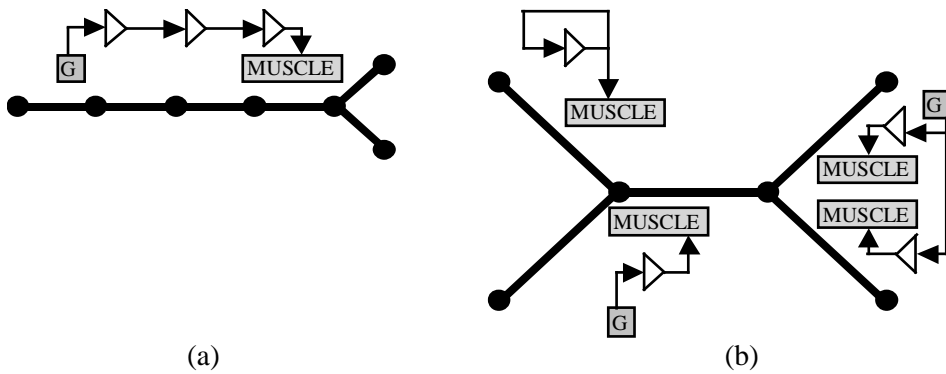


Fig. 2. The structure of body and neural network of the two individuals. “G” is the equilibrium receptor.

Fig. 3 shows properties of parts of the body.

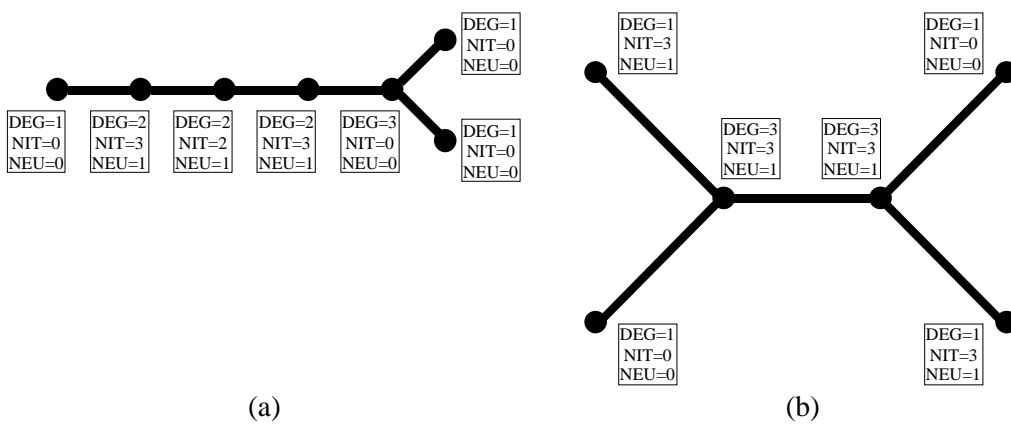


Fig. 3. The properties of body parts: DEG (degree of a body part), NIT (number of neuroitems on a part), and NEU (number of neurons on a part).

The algorithm uses data shown on fig. 3 to match the two individuals. Organism A has one part more, so one part remains unmatched. The matching process is shown on fig. 4: parts are grouped by degree, and then best-fitting parts from the other creature are found.

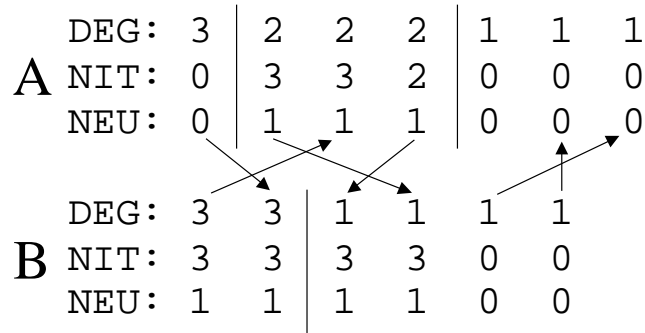


Fig. 4. The matching process of creatures A and B. Arrows show matched parts. Vertical lines separate groups of parts with equal degree.

Thus, the dissimilarity components are:

- Difference in degrees within matching $dDeg = 4$ (unmatched parts are compared to null)
- Difference in the number of neurons within matching: $dNeu = 1$

The overall dissimilarity value can be computed as some aggregation of these two values. For example, using weighted sum with weights $w_{Deg}=1$ and $w_{Neu}=0.5$, the overall dissimilarity is equal to $1 * 4 + 0.5 * 1 = 4.5$.

Example of clustering

The dissimilarity measure was finally applied to a small example of a clustering problem. We performed 10 independent optimization processes for each of three tasks. Each task used a different criterion:

- Maximize average height of creatures, neural network turned off (“hp” task)
- Maximize average height of creatures, neural network turned on (“h” task)
- Maximize average speed of creatures, neural network turned on (“s” task)

Then we have chosen 30 best individuals from these 30 independent evolutionary processes, and computed dissimilarity matrix (dissimilarity for each pair of individuals). Developmental genotype encoding was used [9] to represent individuals during optimization. The values for weights of dissimilarity criteria were adjusted experimentally ($w_{Deg}=1$ and $w_{Neu}=0.5$), constituting an acceptable tradeoff between importance of differences in body and brain. Fig. 5 presents the clustering hierarchical tree. UPGMA method was used (unweighted pair group method with arithmetic averages) [15], a simple and intuitively appealing example of the phenetic approach to data summary. UPGMA groups objects that differ least according to the similarity measure without other points of consideration.

It can be observed that “h” and “hp” individuals are mixed. These tasks were very similar, and obtained individuals were alike. However, it is very interesting that all “s” individuals (evolved to obtain high speed of movement) are clustered together. Thus the criterion of fitness during evolution (height or speed) caused different characteristics of individuals, and the dissimilarity measure was able to capture this difference very well.

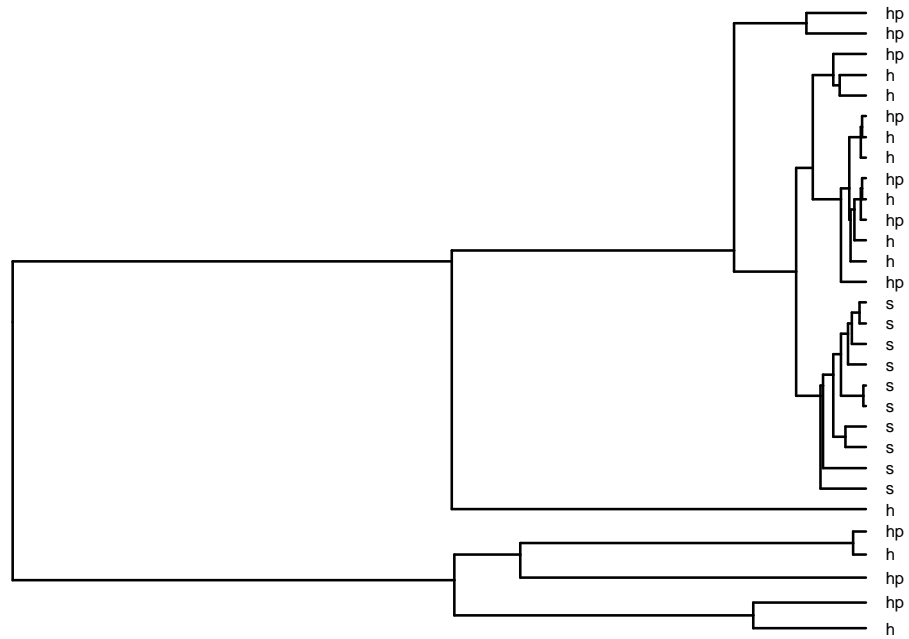


Fig. 5. The clustering tree for 30 best individuals from 3 optimization tasks

Discussions of the actual behavior of these best individuals is outside the scope of this paper. However, to be able to observe the actual operation of the measure, we have applied UPGMA to 10 individuals from the “hp” task. Figure 6 shows their morphologies together with the clustering tree.

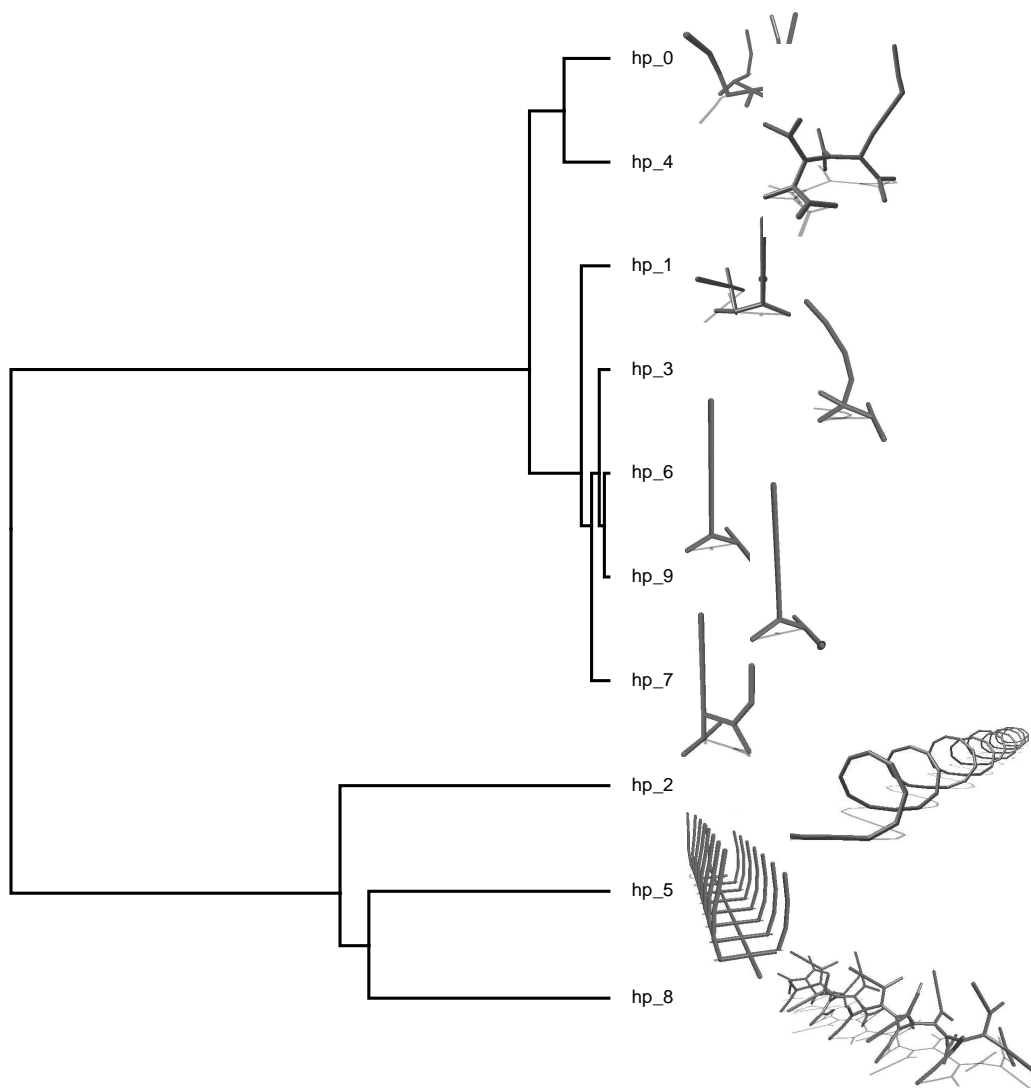


Fig. 6. The clustering tree for 10 best individuals from the “hp” optimization task

It can be seen that three big organisms are in a single, distinct cluster. They are similar in size, but different in structure, so the distance in-between them is high. Moreover, the measure captured also functional similarity (hp_1, 3, 6, 9, 7) – all these agents have a single stick upwards and a similar base. The agents hp_0 and hp_4 are of medium size, but certainly closer to the small organisms group than to the big ones. They are also similar in structure; that is why they constitute a separate cluster.

In summary, the dissimilarity measure is very helpful for study and analysis of groups of individuals. The manual work of classification of the agents shown on fig. 6 yielded similar results, but it was a very mundane and time-consuming process. It also lacked objectivism and accuracy, which are properties of the automatic measure.

5. Approaches to constructing the measure and their consequences

Cladistic vs. phenetic; molecular vs. morphological

The measure we constructed was patterned after the methods of numerical taxonomy (see also section 6). However, there were several problems to be addressed at this stage. We shall discuss them in turn.

The first issue that emerged was the approach to describing organism morphology. In our case, it is the graph theory approach. As we have already seen (section 2), all available information about the organism structure is present at this level. The problem remained how to treat such structural data.

Each of the “building blocks” (parts) could be paired with the corresponding parts of another organism – in an attempt to align them according to homology. This approach would give a measure directly reflecting the patristic component of similarity (i.e. the component due to shared ancestry). In biology this is known as the cladistic approach. The authors decided to use an overall similarity measure, akin to an overall phenetic resemblance coefficient [15]. The computed value takes into account both patristic and homoplastic components of similarity. In biology, searching for the patristic component of similarity is necessary but difficult on the morphological level – especially when we compare evolutionarily distant groups of organisms. It is important to notice that the measure need not ideally mirror the patristic component – it must only change monotonically in accordance with its degree. Then and only then will the topology of the phenogram be the same as the topology of the cladogram. That is why our all-inclusive approach is similar to Jardin's approach (cited for example in [15]), which was to compare entire organisms based on a relational system describing relative positions of organs. We also compare entire representations of body structure (graphs), trying to find the optimal matching (1:1, if it is possible) of the graph vertices. The measure reflects the homology, when corresponding parts are matched minimizing the dissimilarity coefficient.

During construction of the measure, we decided not to use the molecular approach (matching genotypes) [2]. There were some reasons for our decision:

- organisms genomes vary greatly in size and generally are very small compared to real life organisms (which would cause unreliability of matching outcome, as referenced to in [14]),
- in the current Framsticks genotypes, it is hard to identify redundant information, for which additional probabilistic assumptions about accumulation of changes could be formulated (in molecular biology such sequences-introns, NTSs (non-transcribed spacers) allow for Maximum Likelihood analyses [2], as changes in the sequence accumulate randomly),
- graphs of the body and neural network (morphology) are directly and explicitly derived from genotype; characters are fully hereditary. As there is no bias, comparing genotypes would not gain any advantages over the morphological approach.

Molecular methods – as specialized as sequence matching, restriction site analysis or the comparison of specific isozymes – would not help in this case. They are too well adjusted to the known facts in molecular biology.

Properties of the proposed measure

The goal was to compare entire organisms with the least possible loss of generality of the measure. It allows us to construct a metric space (for entire sets of organisms) with distance function d satisfying the following constraints (for each i, j, k):

- (1) $d(i,i) = 0$,
- (2) $d(i,j) = d(j,i)$,
- (3) $d(i,j) \leq d(i,k) + d(k,j)$.

Constraint (3), when strengthened, yields the ultrametricity criterion:

- (3*) $d(i,j) \leq \max\{d(i,k), d(k,j)\}$,

which ensures that the pair-wise dissimilarity function shown on the dendrogram (phenogram) monotonically rises and that all compared organisms are equidistant from the root. The latter is true, if the organisms evolve with equal mutation and recombination rates and come from the same computational experiment.

We were able to show that one of the components of the measure ($dDeg$) satisfies (3). Basing on the computational experiments, we can partially confirm it for the measure as a whole. To this day we didn't formally prove (3*).

Artificial vs. real life. Towards possible uses of the Framsticks system

Artificial life systems supply all the needed information for research, and every aspect can be designed to fit the goal of performing studies. However, some aspects of reality cannot be simulated because they are not yet enough understood, or are too complex. For this reason, it is worthwhile to discuss similarities and differences between *Framsticks* model and the real life. Such comparison may help identify possible applications of artificial life systems (and, specifically, their similarity measures) in biology, and show where computer models cannot be used.

	ALife – Framsticks	Real life
Organism		
Genotype coding	Various formats with desired properties, differing in level of complexity, but equivalent in meaning	Single format of description – DNA sequence
Complexity/size of genome	Small-sized in comparison to real life; usually dependent on the size of phenotype	Huge; of unmapped consequences for regulation and expression of genotype (there is no all-inclusive model on the macro level)
Genotype-phenotype associations	Defined association given explicitly	Association known rarely; micro and macro levels are distinct and many genes are epistatic; so far no common language describing various phenotypes was found
Character's heritability	Full	Varies; environmental conditions influence many traits, which complicates genotype-phenotype associations
Assumptions regarding morphology (when inferring processes)	Known explicitly; can be referenced to model parameters (e.g. formats, available elements, etc.)	Given implicitly, which often causes trouble in constructing taxonomies or phylogenies for complex sets of OTUs (Operational Taxonomical Units [15]); incorrectly chosen similarities or mistakenly identified/weighted traits
Evolution		
Mutation	Can be modeled so that most mutations are individually insignificant, but can accumulate	Small, local changes of DNA sequence, most of which are neutral, but a few can disrupt entire metabolic processes/organism structure
Recombination	Usually rearranges parts of parental genotypes producing an offspring containing both	Greater changes, rearrangements of entire sequences or their parts
Selection	In most cases, based on explicitly given parameters, which define the adaptive value of an individual organism	Criteria often unknown or ill-understood; assessed <i>post factum</i> basing on calculated adaptive value of factions of organisms
Records	Full information about populations at any time	Earlier states not available, often only final state of a system is recorded; earlier states are deduced under some assumptions
Starting conditions	Can be changed; given explicitly	Unknown; cannot be experimented upon; deduced indirectly under some assumptions

6. Conclusions and future work

In this paper, we described the algorithm for estimating dissimilarity of *Framsticks* artificial organisms. This dissimilarity measure captures structural information contained in the agents. We have illustrated the performance of the algorithm, and applied it in two clustering experiments.

There are strong complexity constraints imposed on such measure, because it is usually used many times. For artificial speciation, the measure has to be computed for all pairs of individuals in each generation. In analysis of structure of populations, it has to be computed once, but the population size is usually much greater than in the case of a typical evolutionary algorithm. Thus only the heuristic approach is possible; exact measures would be too slow to compute for all pairs of individuals.

Numerical taxonomy and phylogenetic inference tools try to fulfill three criteria [15], which are presented below in relation to the *Framsticks* model:

- objective – our measure uses an extensive graph description containing all information relevant to organism's morphology; all underlying assumptions are defined in accordance with the *Framsticks* model,
- explicit – genotype-phenotype relations, as well as modeled environmental constraints are given explicitly and can be changed freely. This also applies to the weights and characters included in dissimilarity measure,
- repeatable – simulation of evolutionary processes can be recorded at any moment; parameters and constraints allow for many repeated computational experiments.

Thus it is our conclusion that the dissimilarity measure presented in this work does fulfill these criteria.

As shown on fig. 5 and fig. 6, the constructed similarity measure was used to compute distances and perform cluster analysis using UPGMA method [2] on organisms undergoing directional selection maximizing a particular trade (height, speed). Properties of the dissimilarity function allowed us to construct dendrograms. We expect that the function used to compare evolving framsticks does indeed reflect the degree of (dis)similarity due to shared ancestry (patristic similarity). Thus, potentially, topology of the phenogram is identical to the topology of the cladogram [15]. Answering the question, to what degree this conclusion applies, requires further research. Partial confirmation comes from the fact that organisms evolved under different environmental constraints (in this case: evolution oriented toward height and toward speed, respectively) are indeed put into distinctive clusters (see fig. 5).

The next step is to use the constructed measure to discriminate between organisms in evolutionary horizon, in order to construct groups of OTUs (in section 4 we have shown a clustering, but it was the clustering of individuals, not populations). These groups, in large simulations, would be used as a basis in construction of dendrograms and, hopefully, cladograms, phylogenetic trees. Until now *Framsticks* was used to simulate individuals, whereas in simulating evolutionary processes the goal is to simulate populations. Mutation and recombination in our system are performed on the individual level – same as in real life. But in real life selection is a driving force, working on factions of or even entire populations, as well as on individual members of these populations – and thereby forming the core of all evolutionary processes.

The ill-thought perspective, unfortunately an often-encountered one, is to perceive these associations as both purposeful and singular entities (as noticed, for example, by R. Dawkins [4]). In nature these associations follow the cause-and-effect rule and happen continuously, at the same time – not one at the time. The mathematical models (such as Markov chain models, Monte-Carlo models [16]) do not alleviate the problem on a global basis (i.e. for entire data sets), either being based on specific assumptions about the molecular level of evolution (and aim to accurately divine the course of evolution of a given number of sequences), or approaching one or several phenomena (on population level). The latter applies for example to birth-death models of population genetics.

The measure will be further developed to take into account all the relevant features of an agent, including geometrical and biological properties. Considering geometry creates additional problems, which are also present in biological analyses, namely affine transformations (scaling, rotation, etc.). However, geometry has a great impact on the behavior of agent, functionality, and dynamics of its movement, so it cannot be neglected.

In an artificial life system, adjusting its architecture and parameters allows for simulation of aggregations of evolutionary processes working at the same time. Then, the complete record and a well-founded similarity measure allow simulating and examining these processes in detail. The *Framsticks* model will be enriched with additional features to allow for various experiments. Some of

the possible experiments related to biology include introducing geographical constraints, and then looking at the differences in clusters obtained after a given period of time, or studying two or more populations of much differing sizes. The latter, under geographical constraints, could be used to simulate and understand speciation.

The strength of the *Framsticks* approach lies not in the capability to simulate any single factor or process, but rather in the capability to simulate complex processes on a simplified model. The dissimilarity measure proposed in this work is an interpretive tool for grouping individuals, and can be used to discriminate between adaptations and populations. It can also be employed to diversify populations of individuals during optimization and in another applications, being a versatile tool for a versatile model.

Acknowledgements

This work has been supported by the State Committee for Scientific Research, from KBN research grant no. 8T11F 006 19, and by the Foundation for Polish Science, from subsidy no. 11/2001.

References

1. Adamatzky A., Komosinski M. and Ulatowski S. (2000) Software review: Framsticks, *Kybernetes: The International Journal of Systems & Cybernetics*, **29** (9/10), 1344-1351.
2. Avise J.C. (1994) *Molecular Markers, Natural History and Evolution*. Chapman & Hall, New York.
3. Bentley P. (1999) *Evolutionary design by computers*. Morgan Kaufmann.
4. Dawkins R. (1995) *A River Out of Eden. A Darwinian View of Life*. Orion Publishing Ltd.
5. Funes P. and Pollack J.B. (1998) Evolutionary body building: adaptive physical designs for robots, *Artificial Life*, **4** (4, Autumn), 337-357.
6. Goldberg D.E. (1989) *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Publishing Co.
7. Komosinski M. (2000) The World of Framsticks: Simulation, Evolution, Interaction. In: *Virtual Worlds. Lecture Notes in Artificial Intelligence 1834* (ed. by J.-C. Heudin), pp. 214-224. Springer-Verlag.
8. Komosinski M. and Kubiak M. (2001) Taxonomy in Alife. Measures of Similarity for Complex Artificial Organisms. In: *Proceedings of 6th European Conf. on Artificial Life* (ed. by J. Kelemen and P. Sosik). Prague, Czech Republic. Springer-Verlag.
9. Komosinski M. and Rotaru-Varga A. (2000) From Directed to Open-Ended Evolution in a Complex Simulation Model. In: *Artificial Life VII* (ed. by M.A. Bedau, J.S. McCaskill, N.H. Packard and S. Rasmussen), pp. 293-299. MIT Press.
10. Komosinski, M. and Ulatowski, Sz. Framsticks Internet site, <http://www.frams.alife.pl/>
11. Lipson H. and Pollack J.B. (2000) Automatic design and manufacture of robotic lifeforms, *Nature*, **406** (6799), 974-978.
12. Lund H.H., Hallam J. and Lee W.-P. (1997) Evolving Robot Morphology. In: *Proceedings of IEEE 4th International Conference on Evolutionary Computation*. NJ. IEEE Press. Invited paper.
13. Michalewicz Z. (1996) *Genetic Algorithms + Data Structures = Evolution Programs*. Springer-Verlag.
14. Setubal J. and Meidanis J. (1997) *Introduction to Computational Molecular Biology*. PWS Publishing Company.
15. Sneath P.H. and Sokal R.R. (1973) *Numerical Taxonomy*. Freeman & Co, San Francisco.
16. Suchard, M.A., Weiss, R.E. and Sinsheimer, J.S. Bayesian Selection of Continuous-Time Markov Chain Evolutionary Models (2001) *Mol Biol Evol* **18**, 1001-1013.