

Evolving Ensembles: What Can We Learn from Biological Mutualisms?

Michael A. Lones¹(✉), Stuart E. Lacy², and Stephen L. Smith²

¹ School of Mathematical and Computer Sciences,
Heriot-Watt University, Edinburgh EH14 4AS, UK
M.Lones@hw.ac.uk

² Department of Electronics, University of York,
York YO10 5DD, UK
{stuart.lacy,stephen.smith}@york.ac.uk

Abstract. Ensembles are groups of classifiers which cooperate in order to reach a decision. Conventionally, the members of an ensemble are trained sequentially, and typically independently, and are not brought together until the final stages of ensemble generation. In this paper, we discuss the potential benefits of training classifiers together, so that they learn to interact at an early stage of their development. As a potential mechanism for achieving this, we consider the biological concept of mutualism, whereby cooperation emerges over the course of biological evolution. We also discuss potential mechanisms for implementing this approach within an evolutionary algorithm context.

1 Introduction

Data mining is the process of finding meaningful patterns embedded within data. It typically comprises several stages, including feature extraction, feature selection, classification, and knowledge extraction. Here we focus on classification, which is the process of correctly identifying which category a previously unseen data point belongs to, based on the values of different explanatory variables (or *features*). Classification is the basis of computer-aided decision making, and has become an important tool in many spheres of knowledge, including biomedical diagnosis [9]. A classifier can be considered a function that maps features to classes. This function may be implemented in many different ways, including linear models, probabilistic models, decision trees, artificial neural networks, and support vector machines, to name but a few [5].

Traditionally, classification uses a single model instance, trained on a single sample of the data. Recently, however, there has been a shift towards ensemble models, which train multiple model instances (or base classifiers) and then combine their outputs [6]. By using a diverse selection of base classifiers, it is often possible to reach more accurate and robust decisions than those produced by individual classifiers. Diversity amongst the base classifiers is typically achieved by training them on different samples of the data and/or by restricting them to using different subsets of the features. Widely used examples include bagging, boosting and random forests [5,6].

A common element of ensemble generation algorithms is that the base classifiers are generated sequentially. With the notable exception of boosting, where the performance of each classifier influences the training of the next one, base classifiers are also typically trained independently. This is in stark opposition to the patterns of interaction seen within biological systems, where the elements of cooperative systems emerge in parallel and with considerable inter-dependency. This begs the question: is there some advantage to the biological approach? Moreover, can we learn how to build better ensembles by studying the organisation and evolution of biological systems? To answer this question, we begin by surveying some of the previous work on using models of evolution, namely evolutionary algorithms, to generate ensembles. We also survey previous work on introducing ideas of biological mutualism to evolutionary algorithms. We then discuss how mutualistic models might be practically applied to the problem of ensemble generation, outlining potentially useful research directions.

2 Evolving Ensembles

Evolutionary algorithms (EAs) are population-based optimisation algorithms modelled upon the process of biological evolution. From an ensemble classification perspective, an evolutionary algorithm's population is a potentially valuable resource. During the course of evolution it often contains multiple solutions with diverse behaviours, which is precisely what we are looking for in an ensemble. This has promoted interest in using EAs to train ensemble classifiers.

However, in general, an EA's population loses diversity as evolution progresses. This means that it is often necessary to use a diversity preservation mechanism to maintain behavioural diversity. In our work, we have been looking at whether niching methods can be used to maintain a diverse population of base classifiers throughout an EA run, and whether we can then form effective classifiers from the final population [7, 8, 13]. Niching techniques in EAs are motivated by observations from biology, notably the appearance and preservation of ecological niches within evolving biological populations. Common examples are crowding, fitness sharing and geographical segregation. Our results are promising, showing that the resulting ensembles are significantly more predictive than individually trained classifiers [8]. More recently, we have applied this approach to a medical diagnosis problem, finding that evolved ensembles can compensate for the small differences in clinical practice that often occur between different sites in medical studies. Because evolved ensembles comprise a relatively small number of individually strong classifiers, rather than many weak classifiers (as in random forests), there is also the potential to extract knowledge, which is important for diagnostic classifiers used within a medical context [10].

In biology, niching often reflects the loss of competition between organisms that have no need to compete, for instance because they use different energy and nutrient sources. Species in different niches can then cooperate [16], producing mutualistic interactions that can be beneficial to the ecosystem as a whole [5]. Interestingly, a classifier ensemble is in some ways similar to an ecosystem. It

comprises a group of different classifiers (akin to species) that reach their decisions (their behaviour) as a result of processing different data sources. In general, it makes no sense for classifiers in different niches to compete. However, it does make sense for them to cooperate. We might therefore expect that biological mutualism can guide us in finding more effective ways of evolving ensembles.

3 Biological Mutualism

A mutualism is a symbiotic relationship that is beneficial to all parties [19,20]. These are very common in biology. For example, over 80% of plant species form mycorrhizal associations with fungi, in which the fungus absorbs sugars from the plant's roots and the plant receives phosphates and other hard to absorb nutrients that the fungus scavenges from the surrounding area [21]. This allows the plant to grow in nutrient-poor environments and the fungus to grow in energy-poor environments. Other well known examples include mutualisms between flowering plants and pollinating insects, and between mammals and their colonies of gut bacteria. In many cases, evolution has driven the process to the stage where the individual members of the mutualism can not survive without their mutualistic partners. Mutualistic interactions can involve more than two species [2]. Furthermore, wider interactions involving mutualistic partners lead to the formation of mutualistic networks, which may involve hundreds of species, and which have been described as the 'architecture of biodiversity' [1]. It has also been suggested that selection inevitably pushes populations in the direction of ever increasing cooperation and integration, so the scope and connectivity of these mutualistic networks is likely to increase over evolutionary time [19].

An important issue, both for biology and for any system that seeks to emulate biology, is understanding the conditions required for mutualisms to occur. In particular, what are the mechanisms that prevent the emergence of selfish behaviour in a population of cooperating species—or, to put it another way, which mechanisms support collective action? This question is also studied in the context of economic and social interactions, where it is framed as dealing with asymmetric or hidden information, i.e. where there is a lack of trustworthy information regarding a potential partner's intentions or abilities. Two broad mechanisms for dealing with this situation are signalling and screening [20]. Signalling involves an up-front expenditure of energy or resources (a *strategic waste*) in order to attract cooperation from a potential partner. Examples are the investment in costly phenotypes (e.g. large tail feathers) in animal mating displays, and the distribution of dividends by companies looking for investors. Screening is similar, but is initiated by the less well-informed partner; for example, the expectation that potential suitors should fight for mating rights in order to demonstrate their fitness, or requiring an employee to undergo a probationary period. In symmetric situations, where neither party has sufficient knowledge of the other, combinations of both signalling and screening are often used.

4 Mutualistic Evolutionary Algorithms

EAs are based upon relatively simple models of biological evolution. Ideas of biological mutualism, which have become increasingly influential in the understanding of biological systems, have so far had little impact upon their development. Nevertheless, there are notable exceptions, and several forms of EA do exhibit behaviours that resemble those seen in biological mutualisms.

Co-evolutionary algorithms [14], in particular, model the co-evolution of multiple species within one or more populations. Although mutualistic relationships are co-evolutionary, the two terms are not synonymous, and co-evolution can also lead to symbiotic, predator-prey, constructive or destructive relationships between species. Co-evolutionary algorithms are mainly restricted to two-species systems, and model either competitive (e.g. predator-prey) or cooperative co-evolutionary systems. The latter approach has been particularly successful in recent years, with cooperative co-evolutionary algorithms being successfully used to solve high-dimensional optimisation problems [22]. We have also had some success using these within the context of ensemble classification. For instance, in [13], this mechanism was used to evolve pattern recognisers (i.e. classifiers) in DNA sequences, using a second population of pattern combinatorics (i.e. ensembles) to promote the discovery of meaningful groups of patterns. This was only possible because the base classifiers were evolved in parallel, using information about their ensemble behaviour to focus the search towards the most useful population of base classifiers.

Although not directly motivated by biological mutualism or co-evolution, two other forms of EA that display behaviours akin to mutualism are Michigan-style learning classifier systems and cultural algorithms. Learning classifier systems (LCS) [3] evolve distributed rule sets that can be used to determine responses, for example class predictions in classification problems. In a Michigan-style LCS, each member of the population is a simple rule; however, the overall behaviour of the LCS is a result of cooperation between the rules. Cultural algorithms [15] use a shared belief space where population members can read and write information. This provides a mechanism for population members to interact with one another, and this in turn can be used to support mutualism, amongst many other kinds of behaviour. Both systems use mechanisms to discourage selfish behaviour. In the case of the LCS, rewards are divided between groups of interacting rules using credit assignment techniques such as the *bucket brigade*. Selfish behaviour by a rule will reduce the overall reward for the group, reducing the likelihood of it being propagated to the next generation. In the case of cultural algorithms, only the fittest population members are typically allowed to write to the belief space, with the expectation that fitter members are more likely to be part of a beneficial mutualism. A notable feature of LCS and certain cultural algorithms [18] is that the population *is* the solution, since individual members of the population may not have a meaningful function when separated from other population members. In this respect, they bear significant resemblance to mutualistic biological systems. However, the mechanisms used to promote mutualisms are quite different to those seen within biological systems.

5 Evolving Mutualistic Ensembles

So, might biological ideas of mutualism be useful for evolving ensembles? The combination of specialism and co-operation seen within biological populations is precisely the behaviour we desire in classifier ensembles. We want individual base classifiers to focus on particular sub-tasks, such as processing different subsets of the data (comparable to different nutrient and energy sources). We then want them to co-operate, combining their behaviours in a way that increases their mutual fitness, i.e. we want them to be well-adapted to one another. With current ensemble methods, the behaviours of base classifiers are combined in a variety of simple ways, such as voting or averaging of their outputs (Exceptions include our work on evolving non-linear output combinators [7]). In addition, the ensembles produced are strictly hierarchical, with only the raw features delivered to the inputs, and only the final outputs delivered to the combinator.

It would be interesting to explore forms of interaction more akin to those we see in biology. For example, the output of one classifier could be delivered to an intermediate decision node within another classifier. This is more comparable to a plant/fungus mutualism, with a classifier that is specialised at processing a difficult part of the decision space (analogous to sparse nutrients) feeding into a classifier that makes broader judgements. Using suitable classifier models, we could even remove the hierarchical organisation of classifiers entirely, allowing cyclic interactions to occur (see Fig. 1). Recurrent dynamical systems models, such as recurrent ANNs, and artificial biochemical networks [11], would be appropriate for this. Our work on artificial signalling networks [4] is particularly relevant, showing how complex decision making networks can be formed from interactions between simpler recurrent pathways. However, these are just some of the possible models of representation and interaction that could be

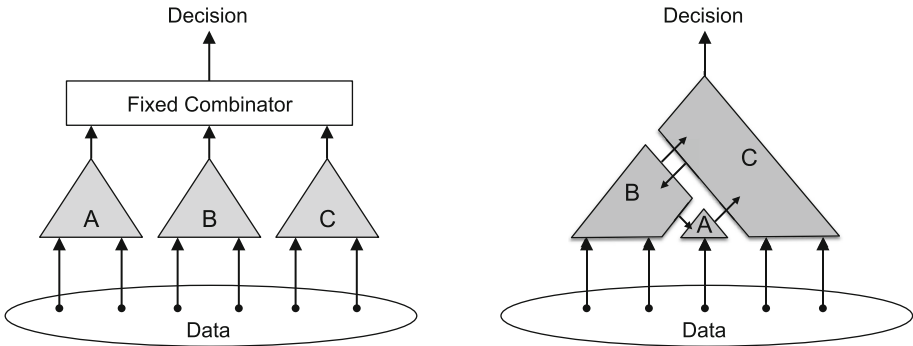


Fig. 1. Left — the standard form for an ensemble classifier, with information feeding hierarchically upwards from the data set to the individual classifiers to the combinator. Right — a mutualistic arrangement in which information flows between classifiers. Classifier A is not able to reach a decision by itself, but generates an important sub-decision that feeds into the rest of the system. An appropriate form of credit assignment would then cause it to be promoted in the next generation.

used within a mutualistic classifier ensemble, and there is considerable scope for adventure in this area. Nevertheless, all approaches would have to address the issue of how to promote the emergence of mutuality in classifier populations.

5.1 Promoting Mutuality

Biological organisms enter into mutualisms because it directly improves their fitness. A similar motivation is required for classifiers, otherwise there is no incentive for them to enter into mutualistic partnerships. In general, we expect to achieve this by redistributing the fitness of the ensemble amongst its component base classifiers, with the expectation that ensembles will outperform individual classifiers and therefore exert selective pressure. Reinforcement learning techniques, such as those used in LCS, could be used for this redistribution. However, there are potential complexities. For example, it would not be unreasonable for a single classifier to be involved in multiple mutualistic ensembles. In such a circumstance, should it receive a reward from all its partnerships, or just the most successful one? The former might promote overly-promiscuous classifiers, yet the latter might discourage wider interactions. Also, should a classifier involved in an ensemble also receive its standalone fitness? This would promote the propagation of individually strong classifiers, but would disadvantage a classifier that plays an important role in supporting an ensemble but which has little standalone worth.

Credit allocation acts as a carrot for the formation of mutualistic interactions, but a stick is also needed to prevent selfish or inefficient behaviours akin to parasitic or symbiotic relationships in biology. Based on biological understanding, signalling and screening have the potential to be effective mechanisms for achieving this. For instance, classifiers could be required to pay a fitness penalty in order to be involved in an ensemble. This form of signalling would be expected to discourage selfish behaviour, since parasitic classifiers would lose individual fitness at the same time as reducing the potential fitness payback from joining an ensemble. There is also the potential for using such a system to study biological hypotheses of how mutualisms are formed. Unlike many of the game theoretic models typically used for this, an evolving ensemble has its own complex behavioural outcomes. In this respect, the mutualistic evolution of classifier ensembles can be seen as an opportunity to test, refine, and potentially even discover biological theories of mutualism.

5.2 Identifying Partners

Biological mutualisms involve different forms of interaction. Mycorrhizal associations between plants and fungi, for instance, usually involve direct contact, with the fungus identifying and then invading the cells of its interacting plant species. Other kinds of mutualism are based upon geographical proximity. For example, a tree provides shelter for birds, birds excrete beside the tree, and the tree absorbs the nutrients. This is also an example of using the environment (the soil in this case) as an intermediary for sharing nutrients, energy, or information.

The classifiers involved in an ensemble must also be able to identify one another, and must be able to do this in a way that is robust to evolutionary change. For instance, a classifier could reference another classifier by a pointer. However, pointers are not robust to evolution, since the reproduction process usually involves cloning, and often leads to multiple offspring. An alternative is to use the environment as a go-between. LCS, for example, do this, with the behaviour of a rule changing the environment, and this change of context then leading to the firing of another rule. It is also the basis of interaction in cultural algorithms, with information being transferred via the belief space environment. However, environmental interaction can lead to unintended interactions, and can increase the scope for selfish behaviour.

Another alternative is to use some form of functional referencing in which the reference describes the form or behaviour of the intended partner. Assuming that no major change in form or behaviour takes place during reproduction, this form of referencing would also be fairly robust to evolution. An example of a functional reference is our work on implicit context representation in genetic programming systems [12].

6 Conclusions

Biological systems are a potentially useful source of information and motivation for the design of computer algorithms. However, there is a danger of taking inappropriate motivation from biology, or of unthinkingly applying bio-inspired algorithms to inappropriate problems. This has led to a degree of criticism in the literature [17]. In this paper, we propose that inspiration from biological mutualisms could be applied to the design and generation of classifier ensembles. To us, this is an example of where biological systems display a clear example of how to achieve complex behaviour in a system of distributed decision makers, and where useful lessons could be learnt and applied to a problem which shows parallels to these biological systems. We have discussed some of the issues that would be involved in doing this. However, we have not tried to suggest one particular way of doing this, because there are many paths that could be taken. Rather, we wish to begin a discussion of the ways in which this could be done, and the potential benefits to which it could lead.

Whilst we have restricted our attention to the particular problem of generating classifier ensembles, it seems likely that mutualism could also provide solutions to other kinds of problem addressed using EAs. For example, genetic programming is concerned with evolving executable structures, such as computer programs. In this context, mutualistic interactions could potentially be used to address the scalability barrier associated with evolving single large programs. Mutualisms could even develop over a series of evolutionary runs, with subsequent runs interacting with solutions found in earlier runs, and possibly bringing them back into the evolutionary process. We would argue that mutualism is an important area of biology that is often overlooked by the evolutionary computation community, but which could bring significant benefits to this community

if modelled and used appropriately. Doing so would also provide an opportunity to test, and potentially generate, biological theories of the ways in which mutualisms evolve.

Acknowledgements. This work was supported by the EPSRC [grant ref. EP/M013677/1].

References

1. Bascompte, J., Jordano, P.: Plant-animal mutualistic networks: the architecture of biodiversity. *Annu. Rev. Ecol. Evol. Syst.* **38**(1), 567–593 (2007)
2. Biere, A., Bennett, A.E.: Three-way interactions between plants, microbes and insects. *Funct. Ecol.* **27**(3), 567–573 (2013)
3. Bull, L.: *Learning Classifier Systems: A Brief Introduction. Applications of Learning Classifier Systems*, pp. 1–12. Springer, Berlin (2004)
4. Fuente, L.A., Lones, M.A., Turner, A.P., Stepney, S., Caves, L.S., Tyrrell, A.M.: Computational models of signalling networks for non-linear control. *BioSystems* **112**(2), 122–130 (2013)
5. Kuhn, M., Johnson, K.: *Applied Predictive Modeling*. Springer, Berlin (2013)
6. Kuncheva, L.I.: *Combining Pattern Classifiers*. Wiley-Interscience, Chichester (2004)
7. Lacy, S., Lones, M.A., Smith, S.L.: A comparison of evolved linear and non-linear ensemble vote aggregators. In: *Proceedings of 2015 Congress on Evolutionary Computation, CEC 2015*. IEEE Press, May 2015
8. Lacy, S., Lones, M.A., Smith, S.L.: Forming classifier ensembles with multimodal evolutionary algorithms. In: *Proceeding of 2015 Congress on Evolutionary Computation, CEC 2015*. IEEE Press, May 2015
9. Lones, M.A., Smith, S.L., Alty, J.E., Lacy, S.E., Possin, K.L., Jamieson, D.R.S., Tyrrell, A.M.: Evolving classifiers to recognize the movement characteristics of parkinson’s disease patients. *IEEE Trans. Evol. Comput.* **18**(4), 559–576 (2014)
10. Lones, M., Alty, J.E., Lacy, S.E., Jamieson, D., Possin, K.L., Schuff, N., Smith, S.L., et al.: Evolving classifiers to inform clinical assessment of parkinson’s disease. In: *2013 IEEE Symposium on Computational Intelligence in Healthcare and e-health (CICARE)*, pp. 76–82. IEEE (2013)
11. Lones, M.A., Turner, A.P., Fuente, L.A., Stepney, S., Caves, L.S., Tyrrell, A.M.: Biochemical connectionism. *Nat. Comput.* **12**(4), 453–472 (2013)
12. Lones, M.A., Tyrrell, A.M.: Modelling biological evolvability: implicit context and variation filtering in enzyme genetic programming. *BioSystems* **76**(13), 229–238 (2004)
13. Lones, M.A., Tyrrell, A.M.: A co-evolutionary framework for regulatory motif discovery. In: *IEEE Conference on Evolutionary Computation, CEC 2007*, pp. 3894–3901. IEEE (2007)
14. Popovici, E., Bucci, A., Wiegand, R.P., De Jong, E.D.: *Coevolutionary Principles. Handbook of Natural Computing*, pp. 987–1033. Springer, Berlin (2012)
15. Reynolds, R.G.: An introduction to cultural algorithms. In: *Proceedings of the Third Annual Conference on Evolutionary Programming*, pp. 131–139 (1994)
16. Santos, F.C., Pinheiro, F.L., Lenaerts, T., Pacheco, J.M.: The role of diversity in the evolution of cooperation. *J. Theor. Biol.* **299**, 88–96 (2012)

17. Sörensen, K.: Metaheuristics—the metaphor exposed. *Int. Trans. Oper. Res.* **22**(1), 3–18 (2015)
18. Spector, L., Luke, S.: Cultural transmission of information in genetic programming. In: *Proceedings of the First Annual Conference on Genetic Programming*, pp. 209–214. MIT Press (1996)
19. Stewart, J.E.: The direction of evolution: the rise of cooperative organization. *Biosystems* **123**, 27–36 (2014)
20. Turcotte, M.M., Corrin, M.S.C., Johnson, M.T.J.: Adaptive evolution in ecological communities. *PLoS Biol.* **10**(5), e1001332 (2012)
21. Wang, B., Qiu, Y.L.: Phylogenetic distribution and evolution of mycorrhizas in land plants. *Mycorrhiza* **16**(5), 299–363 (2006)
22. Yang, Z., Tang, K., Yao, X.: Large scale evolutionary optimization using cooperative coevolution. *Inf. Sci.* **178**(15), 2985–2999 (2008)