A Corporate XCS

Andy Tomlinson
Faculty of Computer Studies and Matematics
Intelligent Computer System Centre
University of the West of England
Bristol BS16 1QY, U.K.
blob @ics.uwe.ac.uk

Larry Bull
Faculty of Computer Studies and Matematics
Intelligent Computer System Centre
University of the West of England
Bristol BS16 1QY, U.K.
larry@ics.uwe.ac.uk

Abstract

Previously we have applied rule linkage to ZCS and shown that the resultant system demonstrates performance improvements over ZCS in a series of sequential tasks, particularly tasks which present ambiguous stimuli to the system. In this paper we show that similar benefits can be gained by applying rule linkage to the more complex XCS. We then show that the benefits of rule-linkage can be increased by further XCS specific modifications to the system's rule-linkage mechanisms.

1 INTRODUCTION

Wilson and Goldberg (1986) originally proposed the theoretical possibility of forming rule-clusters or "corporations" within the rule-base of a Michigan-style classifier system (Holland, 1978); a theory that was further developed by Smith (1992).

In our previous work we have implemented a corporate classifier system (CCS) (Tomlinson and Bull, 1998, 1999) based on the ideas of Smith (1992) which demonstrates that rule-linkage can, for a certain class of problems, offer benefits to a system based on the zeroth-level classifier system (ZCS) (Wilson, 1994). Here we show that similar benefits can be gained when similar rule-linkage mechanisms are applied to XCS (Wilson, 1995).

The paper is arranged as follows, a brief description of XCS is given in the next section, followed by an overview of the rule-linkage mechanisms implemented in CCS. The application of rule-linkage in XCS is then discussed and performances of XCS and CXCS are compared in two contrasting classes of environments. The first, Woods2 (Wilson, 1995) tests the systems' abilities to form external associations, and in particular the capability to form accurately general hypotheses. The second class of environments is comprised of Delayed Reward Tasks (DRTs) (Tomlinson and Bull, 1998) and tests the systems' abilities to form predominantly internal associations. A number of modifications to the rule-linkage mechanisms are then proposed which are shown to improve performance of CXCS in this second class of environments.

2 XCS

The most significant differences between XCS and traditional Michigan-style systems are that XCS dispenses with the internal message list and perhaps more importantly that in XCS, rule fitness for the genetic algorithm (GA)(Holland, 1975) is based not on rule predictions (or strengths) but on the accuracy of these predictions (Frey and Slate, 1991). The intention is to steer the population to form a complete and accurate mapping of the search space rather than to simply focus on the higher payoff niches in the environment.

A further difference is that rather than executing the GA panmicticly, XCS restricts GA activity to within niches defined by the match sets (Booker, 1985). XCS has been shown to evolve rules that are maximally general, subject to an accuracy criterion. This encourages efficiency in knowledge representation within the rule-base. A brief overview of XCS functionality as implemented here is now given.

On each time-step, within the performance component, the system receives some binary-encoded sensory input and forms a match-set [M] consisting of all stimulus matching rules. A system prediction is then formed for each action represented in [M] according to a fitness-weighted average of the predictions of rules in [M] that advocate that action. The system action is selected either deterministically, probabilistically (roulette-wheel selection) or randomly from actions with non-zero predictions. Rules in [M] that advocate the selected action form an action-set [A]. The action is sent to the system effectors and a reward may or may not be received from the environment. If [M] is empty a covering operator is employed to create a new matching rule.

Reinforcement in XCS consists of updating three parameters, p, E and F for each qualifying rule.A rules fitness (F) is updated every time it belongs to $[A]_{-1}$ (or [A] in a single-step problem). The fitness is updated according to the relative accuracy of the rule within the set. There are three steps to the calculation:

1] Each rule's accuracy K_j is determined as follows:

$K_j = \exp[(\ln \alpha)(E_j - E_0)/E_0)] * 0.1$ for $E_j > E_0$ otherwise 1.

2] A relative accuracy K'_j is determined for each rule by dividing its accuracy by the total of the accuracies in the set.

3] The relative accuracy is used to adjust the classifier's fitness Fj using the moyenne adaptive modifee (MAM)(Wilson, 1995) procedure: If the fitness has been adjusted $1/\beta$ times, $F_j \leftarrow F_j + \beta(K'_j - F_j)$. Otherwise F_j is set to the average of the current and previous values of K'_j .

Next E_j is adjusted using P (see below) and the current value of p_j. The Widrow-Hoff technique is used as follows:

$E_j = E_j + \beta(|P - p_j| - E_j)$.

Finally p_j is adjusted. The maximum $P(a_i)$ of the system's prediction array is discounted by a factor γ ($0 < \gamma <= 1$) and added to any external reward from the previous time-step. This value is called P and is used to adjust the predictions of the rules in $[A]_{-1}$ using the Widrow-Hoff delta rule (Wilson, 1995) with learning rate β ($0 < \beta <=1$).

$pj = pj + \beta(P - p_j)$.

The GA acts on the match-set [M] . Two rules from the niche are selected for reproduction stochastically based on rule-fitness. The XCS population, [P] may be of fixed size or of variable size (and initialized to 0). A maximum size, N_p is defined. If [P] contains less than N_p members, the copies are inserted into the population and no compensating deletion occurs. Otherwise two rules are selected from [P] stochastically. Each rule keeps an estimate of the size of the match-sets in which it occurs. A rule's deletion probability is set proportional to this match-set size estimate; this tends to balance system resources across all presented niches.

The GA is activated within a match set if the number of time-steps since the last GA in that match-set exceeds a specified threshold. Each rule is time-stamped at birth with the value of a counter incremented each time-step. When a match-set is formed XCS computes the average time-stamp of its rules and executes the GA if the difference between the average and the current counter value exceed the threshold value (typically this threshold is set to 25).

An action selection strategy is randomly determined at the beginning of each task. There is a 50% likelihood that on each step the action will be selected randomly from those advocated within [M] by rules with non-zero predictions. Such trials are termed exploration trials. On the other trials, termed exploitation trials, a deterministic policy is adopted. During testing exploration is turned off for the last 1000 trials. The XCS GA is only active on exploratory trials, and as such is also turned off at the end of runs. The aim is to facilitate evaluation of the resultant rule-base's utility at the end of the learning period.

In this work, the system is equipped with a fixed size population (randomly initialised), and the concept of macroclassifiers (Wilson, 1995) is not employed. Wilson states that this is simply a coding issue and as such should not effect results.

3. CORPORATE CLASSIFIER SYSTEMS

Our previous CCS, based on Wilson's ZCS model (1994), employs linkage between rules in the population to form rule-chains or "corporations". Each rule in the population is equipped with two, initially inactive *link* parameters, a "link forward" and a "link back". When activated, either, or both of these links may reference another rule in the population. The result of such associations is a population of arbitrarily long rule-chains or corporations, whose members are treated as collective units, both by the discovery component and the performance component of the system.

Activities of the discovery component are based on a measure of *corporate fitness*. For a single rule, this value is the same as its strength parameter as determined by the performance component. For linked rules, corporate fitness is set to the mean strength of all rules in that particular corporation.

If a corporate rule is selected for deletion then the corporation is first disbanded (the linked-list is separated) and then the selected rule is deleted from the rule-base. If a corporate rule is selected for reproduction then the whole corporation is reproduced. The crossover mechanism is expanded to facilitate a form of *corporate crossover* which produces as offspring, a single hybrid corporation which inherits sections of both parent corporations (see figure 1).

Corporations are encouraged to encapsulate temporal chains of inference and so offer improved performance in tasks, such as the previously mentioned delayed reward tasks, which present arbitrarily ambiguous sensory stimuli. Rule-linkage, initially triggered with a fixed probability of 0.1 per time-step, acts across subsequent match sets ([M]) choosing a candidate to join probablistically (based on strength), from each niche. Unlike the original ZCS design, the GA is also restricted to acting in match set niches. These precautions ensure that corporations tend to represent viable sequences of previously presented stimuli, along with associated system responses.

In contrast to the original proposals of Wilson and Goldberg the performance component of CCS is made responsive to the presence of corporations. As in ZCS, on each time-step the system action is selected according to a roulette-wheel selection policy based on the strengths of rules within [M]. This rule is examined for corporate status (i.e. is it linked to another rule?). If a corporate rule has been selected then that corporation has absolute control over the system until either a reward is received (at which point a new task begins) or, the appropriate rule in the chain does not match the currently presented stimulus. In either event performance system functionality returns to standard ZCS behavior.

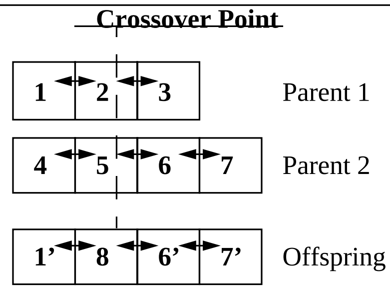

Fig. 1: Corporate Crossover

Rules 1', 6' and 7' are copies of the rules in the original corporations. Rule 8 is the product of crossover between rule 2 and rule 5.

So, if on some time-step *t*, a corporate rule takes control of the system, then on the next step, *t+1*, if the next rule in the corporation matches the new stimulus, control is held and the action of this rule automatically becomes the system action at time *t+1*. This characteristic of the performance component is referred to as persistence and has been shown to enable CCS to overcome sensory ambiguity during multiple-step tasks (Tomlinson and Bull, 1998).

Further encapsulation of corporations is achieved by preventing any rule which has an active "link back" (i.e. a follower) from entering [M], unless it links back to the rule that is currently in control of the system.

4. CXCS: CORPORATE XCS

The linkage mechanisms of CCS are now implemented in a version of Wilson's XCS. This version differs from the original design in that on each invocation, the GA (which acts in [M]) produces only a single offspring. As the XCS GA activation policy is based on the mean age of classifiers within the niche, it is anticipated that this difference will result in only minor performance variations. As the mean age within the niche will be less reduced by each GA action due to the addition of only one new rule, the GA should simply occur slightly more frequently within each niche. During comparisons, the GAs of both XCS and the corporate system, CXCS, produce a single offspring rule or corporation on each invocation.

Rules within the system are given the same linkage components as those in CCS and linkage occurs again between rules from subsequent match sets at a fixed rate (typically a 10% probability on each step). Like the GA, linkage occurs only on exploratory cycles and so is also turned off for the last 1000 trials of testing. In CCS, rule selection for linkage could be either random or probablistic, or deterministic, based on the relative strengths of rules within the niches. The equivalent parameter to ZCS/CCS rule strength in XCS is the prediction parameter. In XCS it is the accuracy of the prediction that is used to evaluate rules, and it is not in keeping with XCS philosophy to base discovery decisions on the prediction parameter alone. Accuracy and fitness are also discounted as possible weightings for rule selection for linkage. In CXCS it is possible that rules that appear to be inaccurate when evaluated alone are precisely the rules that could benefit from rule-linkage. If the inaccuracy is due to some sensory deception then the context of a corporate rule-chain may limit a rule's activation to instances in which it's action results in a more predictable consequence. In context the rule becomes more accurate. This is the main motivation for developing CXCS. With this in mind, selection for linkage in CXCS is determined randomly from rules (whose appropriate link is unattached) within the niche, imposing no bias based on the system's current perception of rule utilities.

Corporations are reproduced and evaluated collectively. As such rules within a corporation should share certain parameters used by the discovery component. These are fitness, which determines a rule's chance of selection for reproduction, and the estimate of mean match set size which determines a rule's chance of being selected for replacement; two parameters are introduced, "corporate fitness" and "corporate niche size estimate". For single rules these parameters are identical to their existing fitness and match set size estimates. For linked rules, these values can be determined in a number of ways. Each rule could be given the average fitness and match set size estimate of all rules within the corporation. Alternatively, corporate fitness could be based on the lowest exhibited fitness within the corporation. In this way, a corporation is considered only as accurate or fit as its weakest link. This approach certainly offers the theoretical advantage of a bias against unwanted parasites within corporations.In the initial design corporate fitness for each rule in a corporation will be set to the lowest exhibited fitness within the corporation. Corporate niche size estimates will be determined as the mean match set size estimate within that corporate unit.

As in CCS, corporations can, while they continue to match presented stimuli, maintain persistent control of the performance component. In CXCS corporations can only take control during exploitation cycles. To allow such behavior during exploration trials would represent a significant degradation of the system's discovery abilities. Again, as in CCS, *followers* (rules with an active "link-back" component) are given only limited access to [M].

When comparing systems that introduce different numbers of offspring per invocation of the GA it is important to consider the differences in relative rule replacement rates. Without such consideration it is possible to generate quite misleading comparisons of systems as rule replacement concerns tend to be amongst the more fragile aspects of classifier system design. To counteract this, a variable element is introduced into the CXCS GA activation. The system records the number of rules reproduced on each invocation of the GA (i.e. the size of offspring corporation, S_c). When the existent activation policy indicates that the GA should fire, a further mechanism will only allow the GA to fire with probability set according to the reciprocal of the mean offspring size parameter, S_m (initialized to 1). This estimate is adjusted on each invocation of the GA according to the standard Widrow-Hoff delta rule (Wilson, 1995) with the learning rate parameter β (typically 0.2), i.e. $S_m \leftarrow S_m + \beta(S_c - S_m)$. This modification to the GA activation mechanism ensures at least a more consistent rate of rule replacement throughout testing, however the drawback is that a corporate system, compared to a standard system will incur a relative reduction in crossover events. The more significant factor is perhaps the rule replacement rate and its effect on convergence within the rule-base, and so here, the variable GA activation policy is adopted for all tests.

An implementation of CXCS, as described above, is now compared to XCS initially in Wilson's Woods2, an environment that does not require the presence of internal associations within the system's rule-base, and then in a series of delayed reward tasks which can only be solved by the formation of internal associations between rules.

5. ANALYSIS OF PERFORMANCE

5.1 WOODS2

Woods2 (Wilson, 1995) is a two-dimensional, toroidal grid-world comprised of 30 x 15 cells. Each cell in the grid may be blank or occupied by one of four types of object, two of which are "food" and two are "rocks". The system is considered to be an artificial animal which traverses the rectilinear grid-world seeking food. It is capable of detecting the sensor codes of objects occupying its surrounding eight cells. These codes (each of which is three-bits long) comprise the system's stimulus at any time-step. 000 represents a blank cell, 110 and 111 represent food type objects, and 010 and 011 represent rocks. As such the stimulus length is 24-bits, with the left-hand three bits representing the cell due north of the current location, and the remainder corresponding to cells proceeding clockwise around it. On receipt of such a stimulus the system decides upon one of eight actions, which represent an attempt to move into one of the surrounding squares. If the cell is blank, the system moves into it, if the cell is occupied by a rock then the system is not able to make the move and if the cell contains food then the move is allowed and the system receives a reward (r_{imm}=1000), this is considered to be the termination of an individual trial. During testing, on the receipt of such a reward, the artificial animal, or animat (Wilson, 1985) is randomly relocated in some blank cell, again ready to seek some food object. A record is kept of the mean number of time-steps taken to reach food over a period of 4,000 successive trials and this is used as a measure of system performance. For further details of Woods2 see (Wilson,1995).

5.2 PERFORMANCE IN WOODS2

The XCS and CXCS models described above are now tested in Woods2. System parameters for these tests are as in (Wilson, 1995):
Rulebase Size: N_p = 800,
Probability of # at an allele position in the initial population: P# = 0.5,
Initial rule prediction = 10.0,
Learning Rate: β = 0.2,
Discount Factor: γ = 0.71,
Probability of crossover per invocation of the GA: χ = 0.8,
Probability of mutation per allele in the offspring: μ = 0.01,
If the total prediction of [M] is less than ϕ times the mean of the population ([P]), covering occurs: ϕ = 0.5.
GA activation threshold parameter = 25,
Number of single-rules or corporations produced by GA as offspring per invocation, 1.
Initial rule error: = 0.0
Initial rule fitness = 10.0
accuracy function parameter: e_0 = 0.01
accuracy function parameter: α = 0.1
Linkage Rate: = 0.1,

Plots of performance are presented (figure 2) which show that both designs reach near optimal performance. Performance plots here represent the average steps to food in the last 50 exploit problems, and the curves are averages of ten runs.

XCS achieves near optimal performance in Woods2 and so it was unlikely that CXCS would offer performance improvements in this Markovian environment which presents no sensory ambiguities to the system. The aim was simply to assess the extent of incurred overheads due to the more complex structures present in the new system. About 130 corporations of various sizes are present in the resultant rule-base and this figure breaks down approximately as follows:

Size	2	3	4	5+
No. of Corp.	80	30	10	10

This accounts for almost half of the rule-base, and so their presence can be considered significant. Their effect on performance however is less so. In this instance it appears that is an acceptable scenario. It should be stated that as the linkage rate is increased from 0.1 the presence of corporations does start to degrade CXCS performance in Woods2 more significantly (not shown).

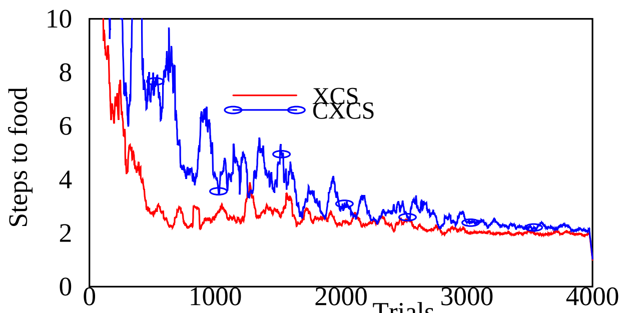

Fig. 2: Performance in Woods2

The systems are now tested in a series of environments in which it is hoped that corporations will in fact be able to improve system performance.

5.3 SIMPLE DELAYED REWARD TASKS (DRTs)

On each of *N* timesteps the system is presented with a stimulus and must select one of *A* actions, where *A* is a variable integer value which defines the breadth of a maze. *N* is the number of states or nodes to a reward and thus defines the maze depth. After *N* steps, the system receives a reward from the environment and a new task then begins. The size of the reward depends on which route the system chooses and so over time the system learns the optimum reward yielding route through the maze.

There is however more than one maze. There can be up to *Mz* different mazes. The system is informed which particular maze it is being presented with only on the first time-step of each trial. On all subsequent steps the stimulus is representative only of the current time-step in the trial. The maze is selected randomly at the start of each trial.

Figure 3 illustrates a simple task of this type with *A* set to 2, *N* set to 2 and *Mz* set to 2. The environmental stimulus at each time-step is also included. In this example, a reward of 1000 is awarded for one route on each map. All other routes receive a reward of 0.

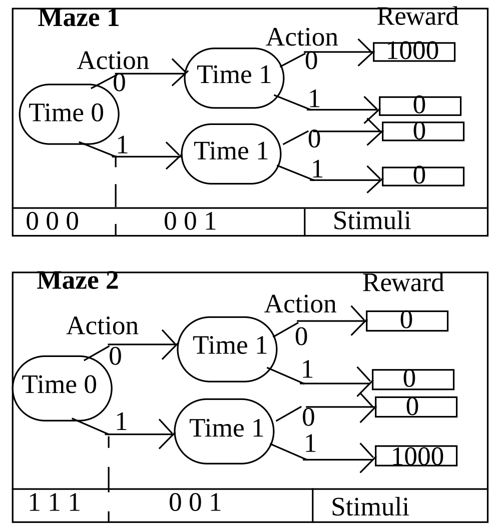

Fig. 3: Simple Delayed Reward Task - DRT 2:2

Throughout this work, as in figure 3, the message length *L* is set to 3. With *L* set to 3 there are 8 possible stimuli and so for a two map problem the maximum depth will be 7, as the first time-step (ts_0) takes two stimuli. Similarly with *L* set to 3 a four maze problem may have a maximum depth of 5.

In the task depicted in figure 3 rule-linkage will enable the system to determine the appropriate move on the second timestep (ts_1). In maze one the correct action is 0 and in maze two it is 1 (the stimulus however is the same - 001).

5.4 PERFORMANCE IN DRTs

Tests are now conducted in a series of four such delayed reward tasks. On each time-step the system must choose one of two actions (0 or 1). The first task, 2:2 consists of two mazes of length two the second task, 2:3 consists of two mazes of length three, etc. One path in each maze will yield a reward of 1000, all others return a reward of 0. The mazes are set up so that if the system selects the same action on each step through the maze it will be guaranteed a reward of 0. This precaution ensures that the successful solution of the mazes is not achievable by a single completely general rule, and will in fact require some form of cooperative behavior within the rule-base. Tests consist of a series of 10,000 trials and all curves are again averages of ten runs. The plots (figures 4, 5, 6 and 7) represent the average score over the last fifty exploitation trials. All parameters are as in the previous tests in Woods2 with the exception of the population size, (400), and the probability of a # at an allele position in the condition of a rule (0.33). These are more standard parameter settings, as used by Wilson when testing ZCS (Wilson, 1994), and also when testing XCS in the Boolean Multiplexor Problems (Wilson, 1995). Figures 4 to 7 also include plots of CXCS with linkage rate increased from 0.1 to 0.25. It can be seen that in these environments an increase in rule-linkage does provide some benefit.

According to the XCS action selection strategy, on a fixed proportion of trials selection will be random from all advocated actions. On such an exploratory trial all rules are likely to receive variations in profit according to the different contexts in which they fire (even if all rules are 100% specific). This will clearly result in low perceived accuracy for all firing rules. XCS accuracy is determined by a quite severe function with a sharp cutoff beyond the acceptable error margin. In such mazes it is possible that all firing rules on any time-step will exhibit a perceived accuracy of 0, leading to each rule having a resultant relative accuracy, and thus a fitness based on the reciprocal of the niche size. Although, during an exploitation cycle, rules representing the optimal system action may be present in the niche, with the bid based on the prediction scaled according to fitness, there will be no discernible bias towards the higher reward yielding action and so the system has certain difficulties mapping the maze which imposes rule co-dependencies that XCS is unable to facilitate.

The effects of the above problem are illustrated by the state of the rule-base at the end of testing and also by the continual activation of the cover operator, especially during the last 1000 exploitation trials. On a 2:2 task for instance, the rule-base will contain many rules of varying specificity that match the “second step stimulus” (001). All of these with the exception of the fully general ### rules have a fitness value of 0. The prevalent ### rules tend to increasingly occupy both niches and thus their prediction values will slowly fall to 0 (due to the previously mentioned precautions taken with the reward scheme designs), and over this period their fitness values will gradually rise to maximum fitness. Eventually [M] prediction on the second step falls below the threshold parameter value and the cover operator is invoked. This process continues throughout the test period.

CXCS does exhibit some ability to map these “internal association” tasks, however results are somewhat disappointing and certainly do not compare to the equivalent results produced by running the CCS model in the same tasks (Tomlinson and Bull, 1998, 1999). In a 2:2 task, for example, the resultant rule-base at the end of testing will typically contain about 100 corporations. These will be of indeterminate length and there will be a number of excessively long corporations. In the equivalent CCS test, the resultant rule-base would typically contain approximately 200 corporations, of which, just about all were of length two. In other words the entire population had linked to form corporations of the appropriate length to tackle the presented task.

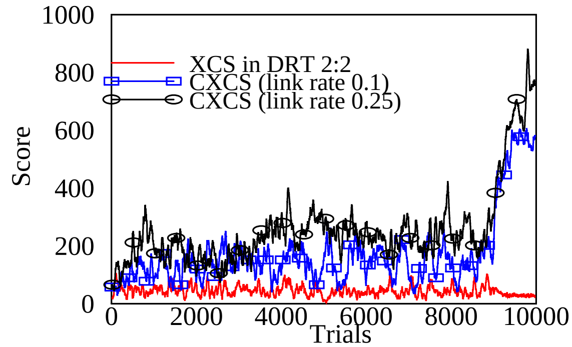

Fig 4. Performance in DRT 2:2

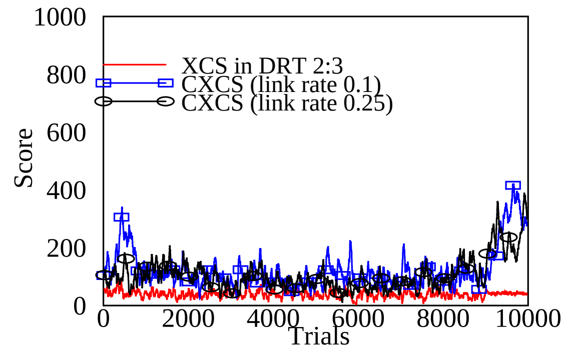

Fig 5. Performance in DRT 2:3

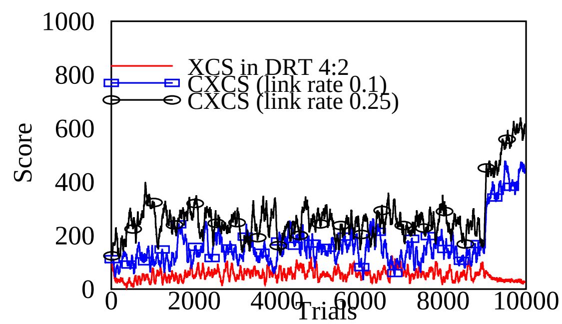

Fig 6. Performance in DRT 4:2

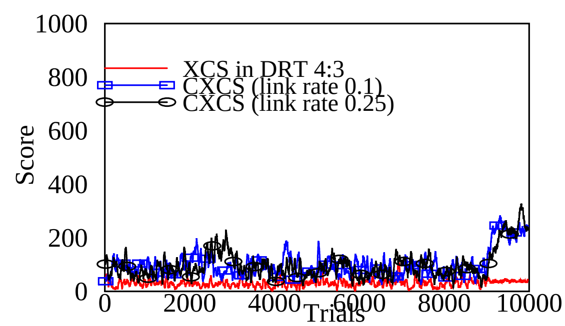

Fig 7. Performance in DRT 4:3

Finally, in the next section, we show that the CXCS mechanisms can be modified to provide significant performance improvements in the more complex delayed reward tasks.

6. MODIFICATIONS TO CXCS

6.1 A RE-EVALUATION OF CORPORATE FITNESS

Wilson (1995) suggests as a future avenue of research that the concept of fitness based on accuracy of prediction could be extended to classifiers with expectons. An expecton takes the form of an additional condition, but is a prediction of the resulting sensation or stimulus.

Such a system does not only match stimulus action pairs with prediction values, but with some anticipated stimulus. Fitness for such rules may not only be based on accuracy of prediction but may also, or separately, represent the accuracy of the expecton in predicting the next sensation. This principle can be adapted for use in CXCS.

Although rules in CXCS do not have an expecton, if their “link forward” is active, they are still associated with information about some anticipated stimulus. This is in the form of the condition of the rule that they are linked to. Actually such a rule may be associated with information regarding any number of anticipated future stimuli, dependent on the length of the corporation.

With this in mind, an alternative measure of fitness may be employed in CXCS for rules with active forward links, based not only on the accuracy of their prediction but also scaled according to the ratio of “control hits”(C_h) to the total number of times that control is received (C_c). If a rule takes control of the performance component and on the next time-step the following rule in the corporation matches the subsequent stimulus and therefore inherits control of the system successfully, then this is considered to be a “control hit”. If the rule fails to match the presented stimulus, then this is a “control miss”. As such, this ratio acts as a measure of how accurately the corporation is representative of some perceived aspect of the test environment. C_h and C_c are updated each time the rule takes control of the system (but on formation of $[M]_{+1}$). The control ratio parameter is adjusted as follows: $C_r = C_h / C_c$. This is determined prior to fitness calculations. On fitness adjustments, the relative accuracy parameter, K’ is scaled according to the control ratio before the fitness parameter is updated in the usual manner. As previously, corporate fitness is consistent for all rules in the corporation and is based on the lowest member fitness in the corporation. Corporate niche size estimate is again set to the mean match set size estimate of member rules.

Figure 8 shows performance plots of this modified version of CXCS in the same series of four delayed reward tasks as used in section 5. Rule linkage rate is again set to 0.25 and all other parameters are as in section 5.

The new fitness evaluation produces improved results in the harder delayed reward tasks. At the end of testing the system generally contains about 100 corporations, most of which are of the appropriate length for the particular task (i.e. of length two or three). On inspection of the plots it is clear that the excessive number of exploration trials are creating difficulties for the performance component. This is evident by the abrupt performance improvement when exploration is turned off for the last 1000 trials of testing.

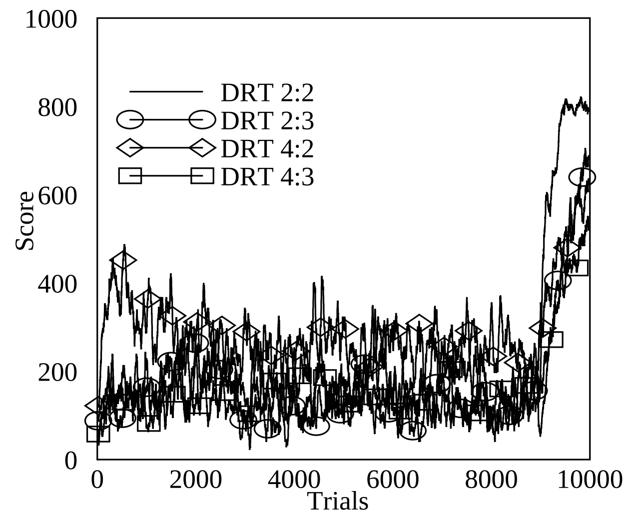

Fig 8: Performance of modified CXCS in DRTs.

6.2 A VARIABLE EXPLORATION RATE

An adjustment to the exploration/exploitation decision process may offer performance benefits. One approach experimented with here is to reduce the probability of exploration periodically throughout the evaluation runs.

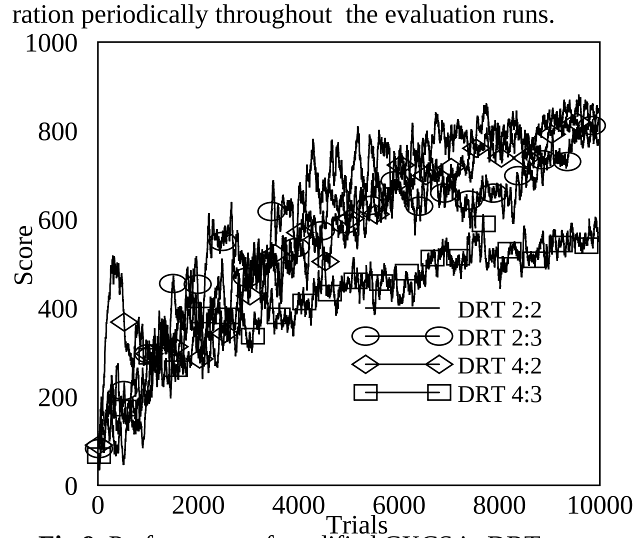

Fig 9. Performance of modified CXCS in DRTs.

The initial probability is 0.5 and every 100 trials this is reduced by one tenth of the current value. This simulated annealing of the system over time offers the performance component increasing opportunities to evaluate the corporate rule structures. Due to the persistent nature of corporations this has become a somewhat serial affair, and so perhaps the performance component now requires a reduction in the relative rate of genetic activity. Performance plots (Figure 9) show that this modification offers some benefit to CXCS in the harder of the four delayed reward tasks. Apart from the variable exploration rate, the system is identical to the system presented in section 6.1.

Within the population, accurate and thus fit corporations are present which collectively map the entire problem space of these tasks, i.e. for each map all possible routes tend to be represented along with reasonable prediction values for these routes, as represented by the predictions of component rules within the corporations.

One final consideration is the evaluation of the modified CXCS model in Woods2. It is important to ensure that the new mechanisms do not disrupt system stability in a less controlled environment than the time constrained delayed reward tasks. Figure 10 shows performance in Woods2 of the modified CXCS. All parameters are as in section 5.2. At the end of 4,000 trials the system reaches food in an average of two steps to food. This plot is comparable with the earlier plot (Figure 2) for CXCS with the original configuration; improvements for delayed reward tasks have no significant effect on stimulus response tasks here.

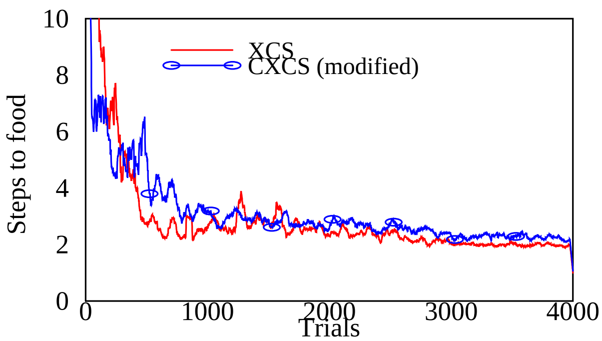

Fig 10. Performance in Woods2

7. Conclusions

This work has shown that it is possible to include the corporate concept in a system based on XCS and achieve similar benefits to those gained in ZCS. The resultant system, CXCS, demonstrates reasonably equivalent abilities to XCS in tackling Woods2, an environment which offers many opportunities for generalization. It is also capable of making some internal associations necessary to solve the delayed reward tasks presented here.

It has further been shown that performance improvements in the delayed reward tasks can be gained by re-evaluating corporate rule fitnesses according to their consistency in maintaining control of the system, and also by reducing the exploration rate during training. The modified CXCS was tested in Woods2 and it was shown that the modifications do not excessively disrupt performance in this environment.

From rule-base inspection, the modified fitness calculation appears to encourage increased generalization within corporations, certainly for "followers", and to an extent which accuracy permits. This is most apparent when examining the rule-base after the system has been trained in Woods2 (not shown). A full investigation of the generalization abilities of our CXCS represents future work.

Acknowledgments

The authors thank Pier Luca Lanzi for useful guidance during the development of XCS.

References

Booker, L. (1985) "Improving the performance of Genetic Algorithms in Classifier Systems." *Proceedings of the First International Conference on Genetic Algorithms and their Applications.* (pp. 80-92). Lawrence Erlbaum Assoc.

Frey, P.W., & Slate, D.J. (1991). "Letter recognition using Holland-style adaptive classifiers." *Machine Learning*, 6, 161-182.

Holland, J. H. (1975) (ed.) *Adaptation in Natural and Artificial Systems.* Univ. of Michigan Press, Ann Arbor.

Holland, J. H., Holyoak, K. J., Nisbett, R. E. & Thagard, P.R. (1986) (eds.) *Induction: Processes of Inference, Learning and Discovery.* MIT Press.

Smith, R. E. (1992) "Memory exploitation in learning classifier systems." *Evolutionary Computation*, 2(3): 199-220.

Tomlinson, A. & Bull, L. (1998) "A Corporate Classifier System." In Eiben, A.E., Back, T., Schoenauer, M. & Schwefel, H.P.(Eds.) *Parallel Problem Solving from Nature, PPSN V*, (pp. 550-559), Springer.

Tomlinson, A. & Bull, L. (1999) "On Corporate Classifier Systems: Increasing the Benefits From Rule Linkage." in Banzhaf, W., Daida, J., Eiben, A. E., Garzon, M. H., Honavar, V., Jakiela, M. & Smith, R. E.(Eds) *GECCO-99: Proceedings of the Genetic and Evolutionary Computation Conference*, Morgan Kaufmann, *to appear*.

Wilson, S. W. & Goldberg, D. E. (1989) "A critical review of classifier systems." In Schaffer, J. D. (Ed.) *Proceedings of the Third International Conference on Genetic Algorithms*, (pp. 244-255), Morgan Kaufmann.

Wilson, S. W. (1985) "Knowledge growth in an artificial animal." *Proceedings of the First International Conference on Genetic Algorithms and Their Applications*, (pp. 16-23). Hillsdale, NJ: Lawrence Erlbaum Assoc.

Wilson, S. W. (1994) "ZCS: A zeroth level classifier system." *Evolutionary Computation*, 2 (1): 1-18.

Wilson, S.W. (1995) "Classifier Fitness Based on Accuracy." *Evolutionary Computation*, 3(2): 149-176.